

UNIVERSIDAD AUTÓNOMA DE CHIHUAHUA
FACULTAD DE CONTADURÍA Y ADMINISTRACIÓN
SECRETARÍA DE INVESTIGACIÓN Y POSGRADO



MODELO PARA CLASIFICAR EMPRESAS DE ÉXITO
EMPLEANDO ESTADÍSTICAS MULTIVARIADAS
EN EL ANÁLISIS FINANCIERO

POR

CARLOS CRISTIAN DE LA ROSA FLORES

DISERTACIÓN COMO REQUISITO PARA OBTENER EL GRADO DE
DOCTOR EN ADMINISTRACIÓN

CHIHUAHUA, CHIH., MÉXICO

DICIEMBRE DE 2020



Universidad Autónoma de Chihuahua
Facultad de Contaduría y Administración
Secretaría de Investigación y Posgrado



La Tesis “**MODELO PARA CLASIFICAR EMPRESAS DE ÉXITO, EMPLEANDO ESTADÍSTICAS MULTIVARIADAS EN EL ANÁLISIS FINANCIERO**” que presenta Carlos Cristian De la Rosa Flores, como requisito parcial para obtener el grado de: **Doctor en Administración**, ha sido revisada y aprobada en la **Facultad de Contaduría y Administración** por los miembros del **Comité de Grado:**

Dra. Ana Isabel Ordoñez Parada
Directora de Tesis

Dr. Alfredo Roberto Velasco Calvo
Asesor de Área Mayor

Dra. Xóchitl Bustillos Varela
Asesor de Área Menor

Dra. Irma Leticia Chávez Márquez
Asesor de Estadística

Dra. Carmen Romelia Flores Morales
Asesor de Metodología

Por las autoridades de la Facultad:

Dr. Héctor Martínez Lara
Secretario de Investigación y Posgrado

M.F. Luis Raúl Sánchez Acosta
Director de la
Facultad de Contaduría y Administración.

© Derechos Reservados
Carlos Cristian De la Rosa Flores
Facultad de Contaduría y
Administración - UACH

DICIEMBRE DE 2020

Dedicatoria

A mis padres, Raquel y Carlos, quienes me inculcaron que todo sueño es posible de hacer realidad, con el esfuerzo y dedicación del día a día.

A mi esposa Teresa, quien siempre ha estado a mi lado en situaciones favorables y desfavorables de la vida que compartimos, enseñándome que el amor todo lo vence.

A mi hija Natalia, la bendición más grande de mi vida, el motor para seguir adelante ante cualquier adversidad y la motivación para siempre mejorar en cualquier aspecto.

A todos los estudiantes que sigan con esa inquietud de cada vez construir más conocimiento con la intención de solucionar problemas que afecten a nuestra sociedad.

Agradecimientos

Agradezco a Dios por darme salud y la oportunidad de siempre estar rodeado de gente con mucho talento, permitiéndome desarrollar mejores ideas y decisiones en cada aspecto de mi vida.

Gracias a la Universidad Autónoma de Chihuahua, Facultad de Contaduría y Administración, por darme la oportunidad de ser parte de su planta docente y poder entonces influenciar de manera positiva a las generaciones de estudiantes que estarán llevando a nuestro país a la cima del éxito.

A mis asesores, Dr. Roberto Alfredo Velazco Calvo por sus enseñanzas desde que fui su alumno al cursar mi Licenciatura; a la Dra. Irma Leticia Chávez Márquez por sus valiosas aportaciones que enriquecieron esta investigación; a la Dra. Xóchitl Bustillos Varela por cada consejo destinado a la conformación de esta tesis; a la Dra. Carmen Romelia Flores Morales quien desde que fui su alumno generó en mí, el interés por la investigación académica;

A mi directora de tesis la Dra. Ana Isabel Ordoñez Parada, por su invaluable apoyo y seguimiento en toda ocasión en la que hemos tenido la oportunidad de colaborar en equipo.

Resumen

El objetivo de la presente investigación fue diseñar un modelo que permita realizar una clasificación de las empresas de éxito, por medio de la aplicación de técnicas estadísticas multivariadas al análisis financiero. El enfoque fue cuantitativo, documental, de carácter no experimental y longitudinal, empleando el método analítico sintético. Se analizaron 93 empresas que cotizaron en la bolsa de valores del Nasdaq durante el periodo 2005 – 2020, obtenidas a partir de un muestreo probabilístico. Se aplicó un análisis de componentes principales, justificado por la prueba Kaiser-Meyer-Olkin y esfericidad de Barlett, obteniéndose 4 nuevas variables capaces de resumir la variabilidad total en un 81.88%, dando pie a construir una escala de medición para determinar la probabilidad de éxito de una empresa, justificado por un análisis de regresión logística binaria. Se utilizó el análisis clúster y el de K-medias para la formación de grupos de éxito, basado en el criterio cúbico de clasificación y el dendrograma obtenido. Con el análisis discriminante y regresión lineal múltiple, se identificaron las variables más significativas para el éxito de una empresa, siendo estas la razón corriente, razón de margen de utilidad y la razón de rendimiento sobre activos. Como resultado final se obtuvo un modelo con la metodología a implementar en el análisis financiero recomendable a aplicar en empresas que cotizan en mercados, tales como la Bolsa Mexicana de Valores y la Bolsa de Valores de New York.

Palabras clave: Estadística multivariada, razones financieras, bolsa de valores.

Abstract

The objective of the research was to design a model that allows a classification of successful companies, through the application of multivariate techniques to financial analysis. The approach was quantitative, documentary, of a non-experimental nature and longitudinal design, using the synthetic analytical method. 93 companies that were listed on the Nasdaq stock exchange during the period 2005-2020 were analyzed, obtained from a probability sampling. An analysis of principal components was applied, justified by the Kaiser-Meyer-Olkin test and Barlett's sphericity, obtaining 4 new variables capable of summarizing the total variability in 81.88%, leading to the construction of a measurement scale to determine the probability of success of a company, justified by a binary logistic regression analysis. Cluster analysis and K-means analysis were used for the formation of success groups, based on the cubic classification criterion and the dendrogram obtained. With the discriminant analysis and multiple linear regression, the most significant variables for the success of a company were identified, these being the current ratio, the profit margin ratio and the return on assets ratio. As a final result, a model was obtained with the methodology to be implemented in the financial analysis recommended to be applied in companies that are listed in other markets, such as the BMV or the NYSE.

Key words: Multivariate statistics, financial ratios, stock market.

Índice general

Resumen	V
Abstract.....	V
Índice general	1
Índice de figuras	3
Índice de tablas	4
I. Introducción	6
I.1. Antecedentes	9
I.1.1. El análisis financiero	10
I.1.2. Parámetros para medir el éxito en una empresa	11
I.1.3. La estadística multivariada	14
I.2. Problema de investigación.....	16
I.2.1. Pregunta general.....	17
I.2.2. Preguntas específicas	17
I.3. Justificación.....	17
I.4. Objetivos del estudio.....	18
I.4.1. Objetivo general	18
I.4.2. Objetivos específicos.....	18
I.5. Formulación de la hipótesis	19
I.5.1. Hipótesis específicas	19
II. Fundamentación	20
II.1. Marco teórico y conceptual	20
II.1.1. Clasificación de las razones financieras	21
II.1.2. Mercados Financieros y su clasificación.....	29
II.1.3. La Bolsa de Valores del Nasdaq y sus aspectos legales	31
II.1.4. Clasificación de la estadística Multivariada	35
II.1.5. Análisis de componentes principales.....	37
II.1.6. Análisis discriminante	38
II.1.7. Análisis de grupos jerárquico	40
II.1.8. Análisis de grupos no jerárquico.....	41
II.1.9. Análisis de regresión lineal.....	43
II.1.10. Análisis de regresión logística	45
II.2. Estado del arte.....	48



III. Criterios metodológicos	50
Naturaleza de la investigación	50
Tipo	50
Carácter	51
Diseño	51
Método	51
Modo	51
Lugar	51
Población, unidad de análisis, muestra y marco muestral	52
Variables	71
Análisis estadístico	73
IV. Resultados y discusión	74
IV.1. Estadísticos descriptivos	74
IV.2. Prueba de medida Kaiser-Meyer-Olkin y prueba de esfericidad de Barlett	75
IV.3. Análisis de componentes principales	76
IV.4. Análisis discriminante	87
IV.5. Análisis clúster	98
IV.6. Análisis K-medias	103
IV.7. Regresión lineal múltiple	108
IV.8. Regresión logística binaria	114
V. Conclusiones	117
V.1. Recomendaciones	132
Bibliografía	133



Índice de figuras

Figura 1: Gráfica de sedimentación	77
Figura 2. Dendrograma	101
Figura 3. Gráficas de residuos	110
Figura 4. Modelo de clasificación de empresas de éxito	125



Índice de tablas

Tabla 1. Clasificación de las razones financieras	20
Tabla 2. Cálculo de la muestra	52
Tabla 3. Codificación de razones financieras.....	71
Tabla 4: Empresas del grupo de éxito	73
Tabla 5: Estadísticos descriptivos	74
Tabla 6: KMO y Esfericidad de Barlett	75
Tabla 8: Correlaciones de los 3 componentes principales.....	78
Tabla 9: Agrupación de razones financieras por los componentes principales	79
Tabla 10: Empresas ordenadas en función del componente principal 1	80
Tabla 11: Empresas ordenadas en función del componente principal 2	81
Tabla 12: Empresas ordenadas en función del componente principal 3	82
Tabla 13: Empresas ordenadas en función del componente principal 4	83
Tabla 14: Cuartiles por grupo de componente principal	85
Tabla 15: Niveles de clasificación	85
Tabla 16: Empresas de éxito en clasificación de componentes principales	86
Tabla 17. Medias de valores por grupo	88
Tabla 18. Prueba de igualdad de medias de grupos	89
Tabla 19. Prueba M de Box.....	90
Tabla 20. Autovalores	91
Tabla 21. Coeficientes de la función discriminante	91
Tabla 22. Matriz de estructuras	92
Tabla 23. Centroides de grupos analizados	93
Tabla 24. Resultados de clasificación	94
Tabla 25. Prueba de igualdad de medias de grupos	94
Tabla 26. Prueba M de Box.....	95
Tabla 27. Autovalores	95
Tabla 28. Coeficientes de la función discriminante	96
Tabla 29. Matriz de estructuras	96
Tabla 30. Centroides de grupos analizados	97
Tabla 31. Resultados de clasificación	97
Tabla 32. Matriz de autovalores de covarianza clúster	98
Tabla 33. Matriz de proximidades	100
Tabla 34. Estadísticos de decisión.....	100



Tabla 35. Primer clúster formado	102
Tabla 36. Segundo clúster formado	103
Tabla 37. Centros de clústeres iniciales.....	104
Tabla 38. Historial de iteraciones	105
Tabla 39. Centros de clústeres asignados por variable	106
Tabla 40. Resumen de agrupación	106
Tabla 41. Primer clúster formado por K-medias	107
Tabla 42. Segundo clúster formado por K-medias	107
Tabla 43. Matriz de correlaciones de Pearson.....	108
Tabla 44. ANOVA de regresión lineal	110
Tabla 45. Significancia de coeficientes de regresión	111
Tabla 48. Significancia de coeficientes de regresión (2)	113
Tabla 49. Coeficiente de determinación (2)	114
Tabla 50. Variables para regresión logística	114
Tabla 51. ANOVA regresión logística	115
Tabla 52. Probabilidad de éxito	116
Tabla 53. Clasificación de empresas con base en Regresión logística.....	116

I. Introducción

El avance y los beneficios que han sido producto del exponencial crecimiento de las Tecnologías de la Información y Comunicación (TIC), han impactado en diversas áreas de estudio, siendo de gran utilidad en el área del análisis del comportamiento de los mercados financieros, en especial el mercado de capitales. Debido a la volatilidad que se presenta en este tipo de mercados, es importante desarrollar estudios orientados a encontrar modelos que permitan minimizar el riesgo en el que incurre un inversionista al colocar sus recursos monetarios en alguna de las empresas que cotizan en las diversas bolsas de valores (Mures, 2013).

Previamente al momento de invertir es posible revisar el perfil de la empresa que cotiza en este tipo de mercados, ya que además de la vasta información que puede encontrarse en las diversas instituciones financieras, es posible tener acceso a la información financiera, ya que uno de los requisitos para cotizar en el mercado de capitales, es convertirse en una empresa pública, por lo que se adquiere la obligación de publicar sus estados financieros al público inversionista de forma periódica (BMV, 2018).

El análisis financiero de una empresa es de vital importancia para lograr uno de los principales objetivos de la administración financiera, que es el de maximizar la riqueza de los accionistas. El estudio de las finanzas proporciona distintas herramientas para poder evaluar el desempeño de las operaciones de la empresa, siendo una de las principales, el uso de las razones financieras, con las cuales se obtienen distintos indicadores, que a su vez proporcionan información sobre los niveles liquidez, rentabilidad, apalancamiento y el uso eficiente de los activos (Servín y Silva, 2011).

Sin embargo, toda empresa a través del tiempo sufre altibajos, por ello no todas llegan a tener el mismo crecimiento y éxito, por lo que es importante analizar los

indicadores financieros para clasificar empresas entre sus semejanzas, facilitando con ello la toma de decisiones para el inversionista. Las crisis financieras recientes han demostrado, que son capaces de afectar a los diversos sectores empresariales, hasta poder hacerlos llegar a la quiebra, como sucedió durante la crisis de las empresas de internet en el año 2001, o bien la crisis hipotecaria del 2008, ambas surgidas en Estados Unidos (Salgado, 2017).

Para poder lidiar con los efectos adversos de estos acontecimientos, y poder permanecer en el mercado de forma exitosa, la estabilidad financiera de cada empresa, juega un rol sumamente valioso.

Finalmente, con la recopilación de información de empresas disponibles para la inversión y datos sobre condiciones de los mercados, los inversionistas requieren analizar cuidadosamente la información para tomar una decisión respecto a donde colocar su capital, por lo que la relación de riesgo y rendimiento está inmersa en la decisión final (Gutiérrez & Medina, 2018).

Al existir una amplia gama de empresas favorables para invertir en los mercados financieros, el que existan metodologías para reducir información de las compañías, al agruparlas con base a similitudes de desempeño financiero, puede agilizar la toma de decisiones de inversión, minimizando el riesgo (Contreras, Bronfman, & Arenas, 2015).

El análisis financiero, puede complementarse con el uso de herramientas estadísticas, en específico, las técnicas multivariadas, con la finalidad de reducir la información obtenida de indicadores que miden niveles de liquidez, rendimiento, actividad y apalancamiento, permitiendo agrupar empresas en grupos de éxito, lo cual mejora la toma de decisiones por parte de los inversionistas al poder identificar que grupo de empresas comparten similitudes en su desempeño financiero, logrando una mejor diversificación del capital invertido.

El objetivo de la presente investigación fue diseñar un modelo que permita realizar una clasificación de las empresas de éxito, por medio de la aplicación de técnicas estadísticas multivariadas al análisis financiero. El desarrollo de la investigación se abordó en varios apartados.

En un primer momento se plantearon los antecedentes del problema a investigar, describiendo la importancia del análisis financiero y la determinación de lo que es el éxito en una empresa. Posteriormente se detallaron los principales aspectos de la problemática a analizar, en donde se describió la manera en que otros investigadores han utilizado la estadística multivariada en el área financiera.

Una vez descrito el problema, se plantearon los objetivos bajo los que se rige la presente investigación, así como la justificación y delimitación de esta, pasando así a la generación de las hipótesis pertinentes.

En el apartado de revisión se hizo la exploración de la literatura en la que se definen los marcos teórico y conceptual para fundamentar el estudio realizado, a través del conocimiento que se tiene del tema, así como la definición de conceptos utilizados durante la investigación.

En el tercer apartado se definieron los criterios metodológicos bajo los cuales se realizó la investigación, siendo ésta de naturaleza cuantitativa, documental, de carácter no experimental y longitudinal, empleando el método analítico sintético. Por último, se describieron los resultados, conclusiones y recomendaciones obtenidas, donde tras analizar las principales razones financieras de una muestra de 93 empresas que cotizaron desde el segundo semestre del año 2015 hasta el primer semestre del año 2020 (15 años en total) en la bolsa del Nasdaq, se diseñó un modelo basado en las principales técnicas estadísticas multivariadas para llegar a clasificar empresas de éxito.



I.1. Antecedentes

Los mercados financieros internacionales, se han caracterizado por tener una serie de altibajos a lo largo del tiempo, los cuales repercuten directamente en los resultados financieros de las empresas. El mercado financiero más grande del mundo se ubica en Estados Unidos, siendo el NYSE y el Nasdaq, las principales bolsas de valores.

El Nasdaq es un mercado global que vende y compra acciones de compañías a través de Internet, es decir, una bolsa de valores que no utiliza intermediarios, sino que todas las transacciones se hacen en línea para mayor seguridad. Fue creado por la Asociación Nacional de Corredores de Valores (NASD) para permitir a los inversores intercambiar valores en un sistema computarizado, rápido y transparente, y comenzó a operar el 8 de febrero de 1971. El término "Nasdaq" también se usa para referirse al Nasdaq Composite, un índice de más de 3,000 acciones listadas en el mercado Nasdaq que incluye la tecnología más importante del mundo y gigantes de la biotecnología como Apple, Google, Microsoft, Oracle, Amazon, Intel y Amgen.

Las crisis financieras recientes han demostrado, que son capaces de afectar a los diversos sectores empresariales, hasta poder hacerlos llegar a la quiebra, como sucedió durante la crisis de las empresas de internet en el año 2001, o bien la crisis hipotecaria del 2008, ambas surgidas en Estados Unidos.

Para poder lidiar con los efectos adversos de estos acontecimientos, y poder permanecer en el mercado de forma exitosa, la estabilidad financiera de cada empresa, juega un rol sumamente valioso.

I.1.1. El análisis financiero

El análisis financiero empresarial comienza partir de la información financiera de las empresas, ésta se encuentra en los diversos estados financieros, los cuales reflejan las operaciones de las compañías, así como los derechos (activos) y obligaciones (pasivos) que tienen, es decir, los estados financieros contienen la información que se necesita para efectuar el estudio, esta información está compuesta por cifras y datos de diversa índole. “El cuerpo de información que describe hasta la más pequeña empresa es enorme y comprende las operaciones internas de la compañía y sus relaciones con el mundo exterior. Para que sea útil esta información se debe organizar en un conjunto de datos suficientemente limitado, fácil de entender y coherente” (Schall, 2000).

Por eso el análisis de estados financieros es útil, pues destaca los puntos fuertes y débiles de la compañía, sin embargo, estos datos por sí solos no proporcionan las respuestas a la situación en que se encuentra la liquidez, la solvencia, la actividad, el apalancamiento y la rentabilidad de las empresas, por consiguiente, el rol que tienen las razones financieras es vital para completar este análisis.

El objetivo de las razones financieras es mostrar cómo se puede reordenar la información proveniente de los estados financieros para convertirla en indicadores capaces de proporcionar información acerca de las cuatro áreas más importantes del desempeño financiero y, en general, de la empresa. El análisis financiero por este medio, es una técnica muy efectiva para reducir la información (Ibarra, 2006).

La aplicación de las razones financieras se justifica de acuerdo a las siguientes tres ventajas: permite reducir el número de variables necesarias, permite comparar empresas de diferente dimensión y se logran calcular valores de referencia sectoriales (Block, 2013).

En el origen de las razones financieras, comienzan a utilizarse como una herramienta de análisis y evaluación de la información contable de las empresas, a partir del estudio titulado “Accounting ratios and the prediction of failure: some behavioral evidence” (Libby, 1975). Estas razones fueron originalmente desarrolladas para el análisis de crédito de corto plazo, y su origen puede ser remontado a finales del siglo XIX (Walsh, 2001).

Los modelos utilizados para clasificar y describir la situación financiera de las empresas han utilizado diferentes métodos estadísticos, sin embargo, las técnicas estadísticas multivariadas han sido poco utilizadas en el análisis financiero. Los modelos que usan razones financieras e integran un conjunto de métodos estadísticos, se enfocan hacia el contenido de la información contable, la cual es expresada en los coeficientes de las razones financieras. Estos modelos que se basan en el uso de razones financieras, miden principalmente la relación del costo de inversión con la rentabilidad de la empresa; pero, también existen diferentes criterios para considerar cuál es el indicador más apropiado para juzgar a las empresas exitosas desde el punto de vista financiero (Mures, 2013).

Las empresas requieren varios elementos para una situación financiera exitosa, entre ellos se encuentran: liquidez necesaria para pagar sus deudas, que el costo de financiamiento sea menor a la rentabilidad de la inversión, nivel de solvencia adecuado para cubrir sus deudas, que la rotación de inventarios sea la necesaria para proporcionar rentabilidad a la empresa, la estructura de activos corresponda al nivel de ventas. Estas características deben incluirse en los modelos de análisis y medición del desempeño financiero de las empresas.

I.1.2. Parámetros para medir el éxito en una empresa

El vocablo éxito, puede interpretarse de muchas formas, por lo tanto, para calificar a una empresa como “exitosa” primero se debe establecer el parámetro o medida estándar de comparación que permita definir si se cumple con esta característica.



Existen diversos indicadores que han sido utilizados en otras investigaciones, para medir el desempeño financiero y entonces clasificar empresas como exitosas. Entre los principales indicadores se encuentran los siguientes; la bursatilidad, el valor de la acción, la generación de valor, la existencia de utilidades, la expansión de la empresa y la existencia de alianzas estratégicas.

Bursatilidad

Mide la facilidad que tienen las acciones, para transferir su propiedad fácilmente dentro de las bolsas de valores donde se comercian (Díaz, 2005). Entre los factores que integran a la bursatilidad destacan: su frecuencia de operación, que la empresa emisora pague buenos dividendos y que la empresa tenga buenas utilidades por acción. En este sentido, la bursatilidad se asocia a las empresas con utilidades suficientes para pagar dividendos a los tenedores de acciones, por eso, es posible que los tenedores de las acciones puedan venderlas fácilmente en los mercados bursátiles y por lo tanto considerarse este indicador como una forma para determinar el éxito de una compañía (Wilmer, 2015).

Valor de la acción

La teoría financiera considera que maximizar la riqueza de los accionistas se refleja principalmente en maximizar el valor de mercado de las acciones de la empresa. Lo cual está determinado por tres factores: la cantidad de flujos de efectivo que se espera se generen para beneficio de los accionistas, el tiempo y el riesgo de los flujos de efectivo (Gitman & Zutter, 2015).

El valor de las acciones de propiedad proporciona un buen índice para medir el grado de efectividad en el desempeño de una compañía. Es por esta razón, que la meta de la administración financiera está generalmente expresada en términos de la maximización del valor de las acciones de propiedad de la empresa, en resumen, maximizar precio de las acciones. Ospina (2015) coincide en que el objetivo primario de los accionistas es la maximización de la riqueza de los accionistas, que equivale a maximizar el precio de las acciones, por lo tanto, es indicador de éxito en

el cumplimiento de las metas principales de las empresas. En la obra más importante de Collins (2002) en la que analiza variables tales como, el liderazgo, trabajo en equipo y un concepto que él denomina como “el círculo del erizo”, define como empresa exitosa, a aquella en la cual el valor de mercado de sus acciones haya duplicado el rendimiento mercado donde cotizaron en un periodo de 15 años, ya que con este lapso de tiempo las empresas habrían pasado por diferentes externalidades que pudieran impactar en sus resultados.

Generación de valor

Según Ross (2009), el valor es el reflejo de la capacidad que tiene una empresa para aprovechar al máximo el capital con que cuenta, que es escaso y requiere del mejor manejo para que pueda generar nuevos recursos, además de costearse a sí mismo. Desde este punto de vista el objetivo de los accionistas no sólo es la maximización de las utilidades, sino la maximización del valor en la empresa. Como consecuencia, se tendrá mayor cantidad de utilidades para los accionistas y además la confianza de que los rendimientos obtenidos por la empresa permiten recuperar adecuadamente los costos de la empresa, es así que Costin (2017) y Bonilla (2010) afirman que la maximización del valor requiere de una perspectiva más amplia que la maximización de las utilidades, por lo que el indicador del EVA, resulta ser una variable importante en la determinación de éxito en una compañía.

Expansión de la empresa

Las empresas atraviesan por un ciclo de vida, y en el transcurso de éste llegan a dominar su mercado local, posteriormente desean expandirse a otros países. Esta expansión o presencia fuera del país de origen de la empresa es para mejorar los flujos de efectivo y, de esta forma, incrementar la riqueza de los accionistas. Algunas estrategias incluyen la penetración de mercados extranjeros, que son muy distintas a los locales, por lo que se desarrollan oportunidades de éxito para mejorar los flujos de efectivo de la empresa, ya que éstos son muy diferentes de los locales, pues crean oportunidades para mejorar los flujos de efectivo de la empresa (Ávila & Alarcón, 2014).



El que una empresa haya logrado ampliar sus horizontes y alcances, es indicador de que la administración de la compañía va recorriendo el sendero del éxito (Puerto, 2010).

Alianzas estratégicas

El establecimiento de alianzas estratégicas hace que las empresas tengan un factor de ayuda para obtener utilidades altas. Una alianza estratégica pone en relación facetas específicas de los negocios de dos o más empresas para entonces crear sinergia y llegar al éxito (Rojas, Rincón, & Mesa, 2015).

Según Díaz (2018) una sociedad mercantil que aumenta la eficacia de las estrategias competitivas de las empresas participantes al hacer posible el intercambio mutuamente beneficioso de las tecnologías, las cualificaciones, o los productos que poseen, por lo tanto, estas alianzas permiten reaccionar rápidamente a las necesidades del mercado y pueden ser factor de éxito a corto, mediano y largo.

I.1.3. La estadística multivariada

La estadística es un proceso que, a partir de los objetivos de una investigación, trata de explicar la variabilidad de un conjunto de datos. Cuando se trata de estudiar la variabilidad de los valores de cada variable de manera aislada, se tiene un estudio univariado y es así como se usan: las tablas de frecuencias, los gráficos de sectores circulares, los gráficos de barras, los histogramas, y los resúmenes numéricos como la moda, la mediana, la media, la varianza, la desviación estándar, los percentiles, etcétera. Los conceptos de probabilidad, de variable aleatoria y varianza de una variable aleatoria, así como la consideración de modelos de distribuciones como la binomial, la Poisson, la distribución normal, la t de Student, la F de Snedecor, la chi-cuadrada, etcétera, son muy importantes en los estudios univariados para inferir propiedades de las muestras aleatorias de toda la población (Véliz, Análisis multivariante: Métodos estadísticos para la investigación, 2017).



En la actualidad gracias al avance de las TIC, es posible generar información la cual a su vez se va almacenando en distintos sitios en donde puede llegar a tenerse acceso para su consulta e inclusive para su descarga y análisis. Esto mismo ha propiciado que la información que se desea analizar sobre algún tópico sea cada vez de un volumen mayor, por lo que es necesario recurrir a técnicas de análisis que permitan compactar la información para generar estudios con un grado mayor de objetividad. En la investigación. Es común recolectar de un grupo de individuos los valores de muchas variables a la vez y la información que se extrae de estos datos puede ser el resultado de analizar cada una de las variables de manera aislada; sin embargo, son los análisis conjuntos de todas las variables a la vez los que proporcionan una información más amplia y revelan patrones de la estructura conjunta de los datos (Rueda, Moya, & Aranada, 2011).

El análisis multivariante o multivariado, proporciona una serie de técnicas estadísticas con este propósito. Hair (2000) establece que algunas de estas técnicas son de carácter exploratoria, pero de mucha utilidad, ya que los patrones de los datos disponibles que a través de ellas se generan, suelen servir como bases para sustentar y explicar propiedades y teorías que más adelante pueden ser objeto de evaluaciones de carácter inferencial y que sirven para toda la población.

La estadística multivariada permite establecer niveles de asociación y agrupamiento, entre una gran serie de datos y variables, con lo que se logra organizar la información en nuevas estructuras, permitiendo su presentación y organización en una forma, más asequible, reduciendo su dimensión, sin perder demasiada información inicial, hasta componer una estructura de información lo más completo posible del conjunto de datos originales (Morelos, 2013).

Diferentes investigaciones han implementado este tipo de estadística en el área financiera, mismas que se detallarán en la sección de estado del arte de esta investigación.

I.2. Problema de investigación

En el área bursátil, en específico en el mercado de capitales, es de vital importancia el realizar un análisis histórico del desempeño de las empresas a través del tiempo, tanto a corto, mediano y largo plazo, con la finalidad de obtener información que permita tomar las mejores decisiones en cuanto a la colocación de sus recursos se refiere. Diversos estudios han explicado los factores que puedan explicar el éxito de las empresas que cotizan en bolsa, sin embargo, desde la perspectiva de los indicadores financieros que se obtienen de un análisis por medio de razones financieras, es poca la información que puede encontrarse.

Las herramientas que nos brinda el estudio de las finanzas, no son lo suficientemente robustas para llegar a establecer causalidad y efecto de un vasto número de variables, por lo que el uso de distintas técnicas estadísticas, en específico, las técnicas multivariadas, pueden lograr este objetivo de forma eficiente.

Debido a la falta de un modelo que establezca la causa y el efecto de un número bastante amplio de variables para explicar el éxito de las empresas que cotizan en bolsa y por lo tanto llegar a crear una clasificación de ellas, se plantean las siguientes preguntas de investigación:



1.2.1. Pregunta general

PG.- ¿De qué manera se pueden clasificar las empresas de éxito empleando las técnicas estadísticas multivariadas en el análisis financiero?

1.2.2. Preguntas específicas

PE1.- ¿Cuáles son los valores de las principales razones financieras de las empresas que cotizan en la bolsa del Nasdaq?

PE2.- ¿Cuáles son las empresas de éxito que cotizan en la bolsa del Nasdaq?

PE3.- ¿Cuáles son las razones financieras que determinan el éxito de una empresa?

PE4.- ¿Cuál es la mejor técnica estadística multivariada que puede aplicarse a las razones financieras para lograr una clasificación de empresas de éxito?

1.3. Justificación

El llegar a crear un modelo que permita determinar cuáles son las razones financieras que permiten conseguir el éxito para una empresa que cotiza en bolsa, y lograr entonces una clasificación de ellas, sería de gran utilidad tanto para los inversionistas que buscan generar un retorno de sus recursos, como a otras empresas que busquen seguir el camino al éxito que han trazado sus semejantes.

Para llegar a la creación del modelo, será necesario contrastar 6 análisis distintos de la estadística multivariada, analizando las ventajas y desventajas que puede presentar cada uno de ellos, con lo que al finalizar el estudio se generará un conocimiento mucho más amplio de este tipo de metodologías, que podrán ser aplicadas en el análisis financiero, con lo que se podrán generar modelos que logren apoyar la toma de decisiones



I.4. Objetivos del estudio

I.4.1. Objetivo general

OG. - Diseñar un modelo para clasificar empresas de éxito, por medio de la aplicación de técnicas estadísticas multivariadas al análisis financiero.

I.4.2. Objetivos específicos

OE1.- Calcular los valores de las principales razones financieras de las empresas que cotizan en la bolsa del Nasdaq.

OE2.- Identificar las empresas de éxito que cotizan en la bolsa del Nasdaq.

OE3.- Determinar las principales razones financieras que permitan establecer el éxito de una empresa.

OE4.- Identificar la mejor técnica estadística multivariada que puede aplicarse a las razones financieras para lograr una clasificación de empresas de éxito.



I.5. Formulación de la hipótesis

Hipótesis General

HG. - La aplicación de técnicas estadísticas multivariadas a las razones financieras, permite generar un modelo para clasificar empresas de éxito.

I.5.1. Hipótesis específicas

HE1.- Los valores de las principales razones financieras de las empresas que cotizan en la bolsa del Nasdaq son muy similares entre sí.

HE2.- Las empresas que tienen una tasa de rendimiento superior en al menos un 50% respecto a la tasa de rendimiento promedio de la bolsa del Nasdaq, son consideradas exitosas.

HE3.- Las principales razones financieras que determinan el éxito de las empresas, pertenecen al rubro liquidez, rendimiento, actividad y apalancamiento.

HE4.- La mejor técnica estadística multivariada que puede aplicarse a las razones financieras para lograr una clasificación de empresas de éxito es la de Análisis de componentes principales, Análisis discriminante, Análisis clúster), Análisis K-medias, Regresión lineal múltiple y Regresión logística binaria.

II.1. Marco teórico y conceptual

Ya que se han definido en forma específica los objetivos sobre los cuáles se desarrolla la presente investigación, se procede a mostrar las diversas teorías y los conceptos principales, que en conjunto robustecen la validación de esta tesis.

Un estudio de Horrigan (1965), comenzó por presentar una clasificación que obtuvo de una compilación de varios autores del tema que en su mayoría coinciden en la clasificación por consenso.

La Tabla 1 muestra las principales clasificaciones de las razones financieras, es necesario tener en cuenta que, consultando a diferentes autores del campo de las finanzas, pueden encontrarse distintas combinaciones entre estos indicadores, a fin de calcular diversos datos, aunque la mayoría de los autores coinciden en que estas razones son suficientes, ya que concentran la información más relevante de la empresa.

Tabla 1. Clasificación de las razones financieras

Rubro	Razón financiera
Liquidez	Razón circulante
	Razón rápida
Rentabilidad	Razón de rotación de inventario
	Razón de días de venta en inventario
	Razón de rotación de cuentas por cobrar
	Razón de días en cuentas por cobrar
	Razón de rotación de activos totales
Actividad	Razón de margen de utilidad
	Razón de margen de rendimiento sobre activos
	Razón de margen de rendimiento sobre capital
Apalancamiento	Razón de margen de deuda total
	Razón de cobertura de intereses

Fuente: Elaboración propia

II.1.1. Clasificación de las razones financieras

El uso de las razones financieras permite realizar comparaciones e investigar acerca de las relaciones que existen entre los diversos elementos de la información financiera de las empresas (Ross, Westerfield, & Jaffe, 2012).

La información utilizada para el cálculo de estos indicadores proviene de los estados financieros que emiten las empresas, tales como el estado de resultados, balance general y estado de flujo de efectivo.

Dependiendo del autor consultado, es posible encontrar opiniones distintas en cuanto a la importancia y en ocasiones incluso en cuanto al cálculo para la obtención de las razones financieras, sin embargo, importantes autores como (Ross, Westerfield, & Jaffe, 2012), (Block, Hirt, & Danielsen, 2013) y (Gitman & Zutter, 2015) han coincidido en que la agrupación y cálculo de las razones financieras se define de la siguiente forma:

Razones de liquidez

Este conjunto de razones financieras tiene como finalidad, el brindar información sobre la solvencia a corto plazo o nivel de liquidez de una empresa, es decir medir la capacidad de una empresa para pagar sus cuentas a corto plazo. Este grupo de indicadores utilizan las cuentas de activo circulante y pasivo circulante, que tienen origen en el balance general de la empresa. La información arrojada por estos indicadores no es útil únicamente para los administradores de la empresa, sino también para agentes externos como lo pueden ser los proveedores, ya que ellos venden sus productos o servicios, siendo por lo general a crédito en periodos de tiempo cortos (semanas, quincenas o meses), así como resulta también de interés para los bancos, ya que pueden otorgar créditos a corto plazo.



Una de las razones más útiles y ampliamente utilizada, es la denominada “razón circulante”, donde su cálculo se realiza de la siguiente forma:

$$\text{Razón circulante} = \frac{\text{Activo circulante}}{\text{Pasivo circualnte}}$$

La principal característica que presentan los activos y pasivos financieros, es su facilidad de conversión a efectivo (por ejemplo, ante una contingencia económica, la empresa puede vender su inventario para hacerse llegar de recursos y solventar sus obligaciones financieras), motivo por el cual esta razón financiera es de gran utilidad para la administración de la empresa. La interpretación de este indicador puede ser en unidades monetarias o bien en veces. Suponiendo que el resultado de ese indicador fuera 1.72, puede interpretarse de la siguiente forma:

- La empresa tiene \$1.72 pesos en activos circulantes por cada peso que tiene el pasivo circulante.
- La empresa es capaz de cubrir sus obligaciones a corto plazo 1.72 veces, haciendo uso de su activo circulante.

Es importante mencionar que, si el resultado obtenido fuera menor a 1, la empresa estaría presentando bajos niveles de liquidez, por lo que, ante una contingencia financiera, la empresa presentaría problemas para hacer frente a sus obligaciones a corto plazo. Mientras mayor sea el resultado de calcular la razón circulante, la empresa estará presentando mayores niveles de solvencia a corto plazo.

Por otro lado, es posible conocer la capacidad de la empresa para cumplir con sus obligaciones en el corto plazo, pero sin la necesidad de recurrir a sus inventarios. Esto es importante conocer, ya que en caso de alguna problemática que presente la empresa y tener que recorrer a la venta del inventario, podría implicar problemas adicionales, tales como el abastecimiento de productos para venta con los clientes. Por lo tanto, el indicador que proporciona información de utilidad para estén tipo de



decisiones en la empresa es la denominada “razón rápida” donde su cálculo se realiza de la siguiente forma:

$$\text{Razón rápida} = \frac{\text{Activo circulante} - \text{Inventario}}{\text{Pasivo circulante}}$$

Esta razón financiera también es conocida como la “prueba del ácido”, en referencia a una analogía, en el que, al verter ácido en el inventario, éste quedaría en condiciones no óptimas para su uso, por lo que la empresa no podría recurrir a él. Al igual que la razón circulante, la interpretación puede ser en unidades monetarias o en veces. Suponiendo que el resultado de ese indicador fuera 1.18, puede interpretarse de la siguiente forma:

- La empresa tiene \$1.18 pesos en activos circulantes (sin tener en cuenta el inventario) por cada peso que tiene el pasivo circulante.
- La empresa es capaz de cubrir sus obligaciones a corto plazo 1.18 veces, haciendo uso de su activo circulante, sin tener que recurrir al inventario.

Pueden existir ocasiones en que, al hacer el cálculo de la razón circulante y la prueba del ácido, los resultados sean muy similares, eso se debe a que existen empresas que cuentan con niveles muy bajos de inventarios (ejemplo, una empresa que se dedica a la reparación de equipos de cómputo), mientras que existen otras que tienen altos niveles de inventario (una empresa que vende equipos de cómputo), es decir, el giro de la empresa influenciará de manera importante en los indicadores calculados.

Razones de actividad

Para una correcta administración de los activos que posee la empresa para realizar sus operaciones, es necesario medir el grado de eficiencia de ellos. Las razones financieras de esta categoría, también evalúan la rapidez con que varias cuentas llegan a convertirse en efectivo, un dato sumamente relevante que permite a los



administradores planificar mejor el tiempo de entrada y salida de los recursos monetarios de la empresa. Los indicadores de actividad utilizan principalmente las cuentas del balance general, como lo es la cuenta de inventarios, cuentas por cobrar y cuentas por pagar.

La denominada “razón de rotación de inventario”, se calcula con cuentas tanto del estado de resultados como con datos del balance general, por lo su cálculo queda de la siguiente forma:

$$\text{Razón de rotación de inventario} = \frac{\text{Costo de los bienes vendidos}}{\text{Inventario}}$$

A diferencia de las razones de liquidez, esta razón no se interpreta en unidades monetarias, sino en veces. Ejemplificando un resultado, donde este fuera 3.8, tomando como referencia que esta razón se calcula de forma anual (aunque puede calcularse en cualquier otro lapso de tiempo), la interpretación es que 3.8 veces en el año, la empresa rotó o renovó la totalidad de su inventario, esto, a consecuencia de las ventas. Por lo tanto, mientras mayor sea el valor de esta razón financiera, mayor es la eficacia en la administración de los inventarios.

Por otro lado, el cálculo de esta razón financiera, da pie al cálculo de otra, que brinda información más puntual acerca del tiempo en que rota el inventario en la empresa. La “razón de días de venta en el inventario” brinda a detalles cuantos días demora la empresa en renovar sus inventarios durante un año, por lo que la fórmula utilizada es la siguiente:

$$\text{Razón de días de venta en el inventario} = \frac{365}{\text{Razón de rotación de inventario}}$$

Siguiendo el ejemplo anterior, al dividir la cantidad de 365 (cantidad que representa los 365 días que existen en un año) entre el valor de la razón de inventario que fue



3.8, se obtiene la cantidad de 96.05, interpretándose que cada 95 días la empresa vende la totalidad de sus inventarios, por lo que debe estar preparado en esos lapsos de tiempo para la adquisición de uno nuevo, evitando un desabastecimiento que pueda impactar de forma negativa en sus ventas.

Sin embargo, si se utilizaran únicamente estas primeras dos razones de actividad para medir la eficiencia de los activos, se estaría aún lejos de tener información suficiente para llegar a una correcta administración. Por lo tanto, la “razón de rotación de cuentas por cobrar” abona de manera sustancial a la información sobre la eficiencia que están presentando los activos de la empresa. El cálculo de este indicador, es de la siguiente forma:

$$\text{Razón de rotación de cuentas por cobrar} = \frac{\text{Ventas a crédito}}{\text{Cuentas por cobrar}}$$

Esta razón mide la rapidez con la que se cobran las ventas que fueron realizadas a crédito durante un año de operaciones de la empresa. Suponiendo que el resultado de este indicador es 12.7, la interpretación es que 12.7 veces en el año se cobraron las cuentas por concepto de ventas a crédito. Sin embargo, la “razón de días en cuentas por cobrar” brinda información más detallada respecto al tiempo que demora la empresa en hacer efectivas sus cuentas por cobrar. El cálculo está dado de la siguiente forma:

$$\text{Razón de días en cuentas por cobrar} = \frac{365}{\text{Razón de rotación de cuentas por cobrar}}$$

Siguiendo con el ejemplo anterior, al dividir la cantidad de 365 (cantidad que representa los 365 días que existen en un año) entre el valor de la razón de rotación de cuentas por cobrar que fue 12.7, se obtiene la cantidad de 28.74, interpretándose que cada 28 días, las ventas a crédito de la empresa se cobran en este lapso de tiempo.



Finalmente, en este rubro de razones de actividad, es posible tener un indicador que muestre la eficiencia general que están brindando todos los activos de la empresa. La “razón de rotación de activos totales”, involucra todos los activos, por lo que es muy útil para tener un panorama general. El cálculo de este indicador se muestra a continuación:

$$\text{Razón de rotación de activos totales} = \frac{\text{Ventas}}{\text{Activos totales}}$$

Suponiendo que el resultado obtenido es de 0.88, en este caso la interpretación no sería muy útil medirlo en veces, sino en unidades monetarias, ya que este resultado indica que por cada peso que se tiene de activos, se generan 0.88 pesos en ventas, por lo que mientras mayor sea este indicador, mayor es la eficacia en la administración de los activos totales.

Razones de rentabilidad

Sin lugar a duda la medición de la rentabilidad en las finanzas, es algo primordial establecido en sus principios más básicos. Esta clasificación de indicadores, es de vital monitoreo por parte de la administración de la empresa, ya que en todo momento se busca maximizar el valor del ente económico, al presentarse disminuciones en los indicadores de rentabilidad, puede ser una situación que no únicamente se debe a una contracción del mercado, sino a una mala administración por parte de la gerencia.

La “razón de margen de utilidad” es un indicador que, de manera general, brinda una primera perspectiva del rendimiento que se está generando, siendo su cálculo el siguiente:

$$\text{Razón de margen de utilidad} = \frac{\text{Utilidad neta}}{\text{Ventas}}$$

Para una correcta interpretación de este indicador, es necesario multiplicarlo por 100, para obtener el resultado de forma porcentual, por lo tanto, teniendo como ejemplo que se obtiene la cantidad de 0.22, significa que únicamente el 22% de las



ventas totales, se convierten en utilidad neta, es decir el 78% restante de los ingresos por ventas, se consumen en cuentas tales como gastos operativos e impuestos. Otra forma de interpretarse es que, por cada peso de las ventas, únicamente 22 centavos llegan a ser parte de la utilidad neta generada.

En esta clasificación de indicadores, se vuelve a utilizar la cuenta de activos totales, para de esta forma conocer la rentabilidad que están generando en su conjunto, por lo tanto, la “razón de rendimiento sobre activos” que también es conocida como ROA (return on assets), está formulada de la siguiente manera:

$$\text{Razón de margen de rendimiento sobre activos} = \frac{\text{Utilidad neta}}{\text{Activos totales}}$$

De igual forma que la razón previamente analizada, la interpretación puede ser de forma porcentual o bien, en unidades monetarias. Suponiendo que se obtiene un resultado de 0.14, la interpretación es que la totalidad de activos presentan un rendimiento del 14%, tomando como referencia la generación de utilidades netas, en otras palabras, quiere decir que por cada peso que se tiene en los activos, estos contribuyen generar 14 centavos de utilidad.

Uno de los objetivos principales del administrador financiero, es el de maximizar la riqueza de los accionistas, por lo que la “razón de rendimiento sobre el capital”, también conocido como ROE (return on equity), brinda información sobre el capital invertido en las compañías.

El cálculo de este indicador, se muestra a continuación:

$$\text{Razón de margen de rendimiento sobre capital} = \frac{\text{Utilidad neta}}{\text{Capital total contable}}$$

Nuevamente, el resultado de este indicador puede ser interpretado en unidades monetarias o bien en forma porcentual, por lo que, en un ejemplo, al obtenerse un resultado de 0.11, significa que el rendimiento que presentan los accionistas de la empresa en el capital invertido, es de 11%, o bien en otras palabras significa que,



por cada peso invertido en la empresa, han generado 11 centavos. Este indicador es tan útil que incluso, nuevos inversionistas pueden tomar la decisión sobre invertir o no en alguna empresa, teniendo como referencia el análisis de esta razón financiera.

En este punto es importante mencionar que tanto el ROA como el ROE, reflejan valores respecto a cantidades tomadas en el valor en libros de los activos y del valor de las acciones y no como tal en el valor de mercado que puede diferir de forma considerable, según el caso de cada compañía.

Razones de apalancamiento

Los indicadores de esta agrupación, tienen el objetivo de determinar la capacidad de la empresa para hacer frente a sus obligaciones de largo plazo, es decir, hacer frente al apalancamiento financiero, por lo que también estas razones son llamadas. Razones de solvencia a largo plazo. El que una empresa se apalanque financieramente, significa que estará presentando cierto riesgo al presentar alguna disminución en ventas, por lo que la información oportuna de estos indicadores ayuda a mejorar la toma de decisiones respecto a la obtención de créditos y el uso de costos fijos en las operaciones de la compañía.

Para tener una visión general de la situación de apalancamiento, existe la “razón de deuda total”, donde el cálculo es el siguiente:

$$\text{Razón de margen de deuda total} = \frac{\text{Pasivo total}}{\text{Activo total}}$$

Como se explicó previamente en el cálculo de otros indicadores, el resultado obtenido puede interpretarse de forma porcentual o bien en forma de unidades monetarias, por lo que en un ejemplo de obtener un resultado de 0.34, esto quiere decir que los pasivos totales representan un 34% de los activos totales, es decir que por cada peso que se tiene en el activo, existen 34 centavos de pasivo. Cuando el indicador supera la unidad, quiere decir que el total de pasivos supera la cantidad de activos que posee la empresa.



Al estar hablando de endeudamiento, es inherente que los intereses están involucrados en los niveles que apalancamiento financiero que puede tener un ente económico, por lo que medir la capacidad que tiene la empresa para solventarlos es muy importante. La “razón de cobertura de intereses” tiene como objetivo el determinar el número de veces que la empresa puede cubrir el pago de intereses, tomando como referencia la UAI, por lo que el cálculo del indicador se realiza de la siguiente manera:

$$\text{Razón de cobertura de intereses} = \frac{UAI}{\text{Intereses}}$$

Generalmente la interpretación de este indicador se hace en número de veces. Teniendo un ejemplo de que el resultado obtenido es 8.2, eso quiere decir que es posible pagar los intereses de la deuda con la cantidad de utilidad operativa generada, por lo que mientras mayor sea el valor de este indicador, menos riesgo se presenta en cuanto a que la empresa tenga alguna complicación para cumplir con este compromiso con sus acreedores. Esta razón puede ser de particular interés para la banca, previo al otorgamiento de un crédito.

II.1.2. Mercados Financieros y su clasificación

La mayoría de las empresas que van logrando un crecimiento y expansión a través del tiempo, van incrementando su necesidad de obtener fondos, por lo que según Gaytán Cortés (2018), las principales fuentes externas a las que se acuden son: Instituciones financieras: las cuales sirven como intermediarios para canalizar los ahorros de los individuos, empresas y gobiernos, hacia préstamos e inversiones. Dentro de la diversa gama de instituciones financieras que pueden encontrarse en el mercado, destacan los bancos comerciales, las asociaciones de ahorro y préstamos, las cooperativas de crédito, los bancos de ahorro, las empresas de seguros, los fondos de inversión y los fondos de pensiones.



Mercados financieros; son espacios (físicos o virtuales) en los que proveedores y solicitantes de fondos negocian directamente. Los mercados financieros son aquéllos en los que se intercambian activos con el propósito principal de movilizar dinero a través del tiempo. Están integrados fundamentalmente por los mercados de deuda, los mercados de acciones y el mercado cambiario (Rueda A. , 2005).

El concepto de mercado financiero se refiere en sentido general a cualquier mercado organizado en el que se negocien instrumentos financieros del mercado de deuda, del accionario, del de cambios, del de metales y del de derivados. Para efectos del presente estudio, resulta relevante únicamente el mencionar los principales dos mercados que lo componen:

Mercados de deuda: El Gobierno Federal, los gobiernos estatales o locales y las empresas paraestatales o privadas pueden necesitar financiamiento, ya sea para realizar un proyecto de inversión o para mantener sus propias actividades. Estas entidades pueden conseguir los recursos a través de un préstamo; solicitando un crédito a un banco o a través de la emisión de un instrumento de deuda. El mercado de deuda es la infraestructura donde se emiten y negocian los instrumentos de deuda. El mercado de deuda también se conoce con otros nombres dependiendo del tipo de instrumentos de deuda negociado. Por ejemplo, si en el mercado se negocian principalmente instrumentos de deuda que pagan una tasa fija entonces se denomina mercado de renta fija (Ross S. A., 2009).

Mercado de capitales, es aquel en donde se realizan la compra-venta de títulos valor, representativos de activos financieros de empresas y otras unidades económicas como son las acciones, obligaciones y títulos de deuda de largo plazo. Este mercado otorga a los inversionistas la posibilidad de participar como socios (en parte proporcional a lo invertido), en el capital de las mejores empresas de México. En contraparte a las empresas les otorga la posibilidad de colocar parte de su capital entre un gran número de inversionistas, con el objeto de financiar capital de trabajo y/o la expansión de la misma empresa (Gitman & Zutter, 2015).

II.1.3. La Bolsa de Valores del Nasdaq y sus principales aspectos legales

Nasdaq es la bolsa de valores electrónica automatizada más grande de Estados Unidos. Las siglas Nasdaq significan "National Association of Securities Dealers Automated Quotation". En el Nasdaq cotizan más de 3,300 empresas del sector de tecnología, como electrónica, informática, telecomunicaciones y biotecnología, teniendo en total un valor de capitalización de mercado de casi 22,000 millones de dólares. Este corporativo es dirigido por la directora ejecutiva Adena Friedman, desde el año 2017 (Nasdaq, 2020).

Respecto al marco legal que regula el mercado bursátil en Estados Unidos, es importante hacer un recorrido en los diferentes acontecimientos y regulaciones que se dieron en el tiempo, para dar pie a una de las bolsas de valores con mayor relevancia en el sector financiero.

En 1933 fue emitida la denominada "securities act 1933". Esta regulación fue expedida con posterioridad y con ocasión de la gran depresión y el colapso de los mercados de octubre de 1929, los cuales favorecieron una serie de actos del Congreso de los Estados Unidos que durante una década produjeron el conjunto de actos conocido como las leyes de valores federales (Lopez, 2008). La primera ley federal de valores expedida fue el "Federal Securities Act de 1933" que reguló la oferta pública y venta de valores en el comercio interestatal. Con anterioridad a la gran depresión, se dieron muchos abusos en el manejo del mercado de valores de los Estados Unidos y para efectos de conjurar estos abusos y lograr un mercado más confiable y transparente, se expidió el Securities Act de 1933, cuya base de mayor impacto fue la que estableció la normatividad respecto al proceso de revelación de información que toda empresa debía emitir. Esta emisión de 1933 reguló las denominadas ofertas primarias, las cuales consistían en el proceso en que los valores eran ofrecidos por primera vez al público, pero este acto también cubrió las denominadas ofertas secundarias. Puede decirse entonces que el aspecto más importante de esta



primera regulación, fue que gracias a ella, se regularon las emisiones de valores hechas a los inversionistas y requiriendo que el emisor realizara un registro de todos los valores que se emitían al público por primera vez. Por lo anterior, se requirió que el emisor debía hacer el registro de lo que dentro de la legislación de valores se le denominó “informe de registro” (registration statement) ante la SEC. Con la actualización de nuevas funciones de la Securities and Exchange Commission en 1934, se derivaron 4 divisiones encargadas de supervisar los aspectos más relevantes en cuanto a fiscalización de operaciones primordiales en la operación y cuidado de la información de empresas que estuvieran en situación de registro y venta de valores (Lopez, 2008).

- División de cumplimiento: esta división se responsabiliza de investigar las violaciones a las leyes en materia de valores. La SEC cuenta con acciones de cumplimiento y acciones administrativas a ser impuestas cuando esta cree que una violación a las normas ha sido cometida por parte de emisores registrados, sus oficiales y empleados, broker-dealers registrados o miembros de bolsas de valores o de asociaciones autoregulatorias.
- División de finanzas corporativas: esta división se encarga de revisar que los documentos de registro cumplan con los requisitos de revelación de información contenidos en las leyes de valores. También prepara las guías de revelación a ser publicadas por la SEC y apoya a los usuarios en la elaboración de sus revelaciones de información.
- División de regulación de mercados: esta división expide las políticas relacionadas con las bolsas de valores, los mercados relacionados a operaciones bursátiles y los broker-dealers.
- División de administración de inversiones: esta división se encarga de supervisar que todas las transacciones emitidas entre oferentes y

demandantes de valores, se realicen de forma correcta, cumpliendo con todas las normas de seguridad bancarias requeridas para de esta forma garantizar la seguridad de los recursos monetarios de todas las entidades.

Con estas primeras regulaciones generales en el mercado de valores de Estados Unidos, dieron pie a los primeros acontecimientos de la formación de valores del Nasdaq, mismos que se explican a continuación:

El primer antecedente de la bolsa del Nasdaq, comenzó antes de la segunda guerra mundial, ya que, en 1938, el Maloney Act creó la National Association of Securities Dealers (NASD), asociación que comenzó en operaciones de compra y venta de títulos accionarios. La NASD, siguió con crecimientos moderados a través de los años., sin embargo, el origen del Nasdaq, se debe a una petición del Congreso de los Estados Unidos a la comisión que regula la bolsa (Securities and Exchange Commission, SEC) de que realizara un estudio sobre la seguridad de los mercados. La elaboración de este informe detectó que los mercados no regulados eran poco transparentes, por lo que la SEC propuso su automatización y de ahí surgió el NASDAQ Stock Market (Nikolsko-Rzhevskaya, Nikolsko-Rzhevskyy, & Black, 2020). Esta bolsa de valores, empezó a operar el 9 de febrero de 1971, siendo primer mercado de valores electrónico del mundo, donde en su primer día de operaciones, se negociaron más de 2,500 títulos accionarios. En 1985, se creó el índice bursátil Nasdaq-100, que incluyó a 100 de las compañías no financieras internacionales más grandes. Muchas compañías, incluidas las de atención médica, telecomunicaciones y venta minorista, etc., figuraron en el índice. Los pesos de las compañías listadas se basaron en su capitalización de mercado (Nasdaq, 2020). En 1991, Nasdaq inició un servicio internacional para que los inversores y las empresas de otras áreas también pudieran participar en este mercado, además, la bolsa presentó un nuevo sistema de red que permitió a empresas e inversores negociar acciones durante el horario comercial europeo.



En 1994, Nasdaq superó a la Bolsa de Nueva York en volumen de acciones anuales muchas veces, esto debido principalmente a su popularidad entre los inversores de empresas emergentes de alta tecnología, como lo fue Microsoft, al aparecer en cotizaciones de esta bolsa. El sitio web oficial de Nasdaq (<https://www.nasdaq.com>) comenzó a funcionar en 1996, proporcionando toda la información que los inversores podrían necesitar sobre todas las acciones cotizadas en ese mercado, por lo que rápidamente se convirtió en el sitio web más popular para los inversores, dando una mayor seguridad y certeza en sus operaciones (Nasdaq, 2020)

En 1999, se convirtió en el mercado bursátil más grande del mundo basado en el volumen en dólares, lo que le permitió cambiar de ubicación a Time Square en la ciudad de Nueva York. El año 2004 fue un periodo clave para esta bolsa de valores ya que el gigante Google, llegaba a cotizar en la bolsa de valores por primera ocasión. Google ganó poco más de \$100 dólares por acción en el primer día de su oferta pública (CNN-money, 2004).

Durante 2007, Nasdaq estableció una oficina en Beijín, que es la capital de China, lo que permitió incluir a 52 empresas de ese país, ampliando las fronteras de inversión de una manera sorprendente. En la actualidad Nasdaq tiene inscritas aproximadamente 3,200 valores, de los cuales 335 son empresas fuera de los EEUU, provenientes de 35 países que representan todos sectores de la industria (Nasdaq, 2020).

Según el sitio web oficial del Nasdaq, existen 3 indicadores bursátiles, que concentran todas las empresas enlistadas, a continuación, se detalla cada uno:

Índice Nasdaq-100: éste incluye cien de las empresas más grandes que se encuentran listadas en esta bolsa basada en el volumen de venta que presentan las acciones de dichas compañías. Este índice refleja a aquellas empresas dedicadas a las telecomunicaciones, hardware y software, pero no contiene compañías financieras o que se dediquen a inversiones puramente (Nasdaq, 2020).

Índice Nasdaq Composite: es el grupo de acciones que se encuentran listadas dentro de esta bolsa electrónica e incluye más de 3 mil compañías sin importar el giro. En éste se pueden incluir compañías financieras y de inversiones, así como de tecnología en general (Nasdaq, 2020).

Índice Nasdaq de Biotecnología: lista a las empresas farmacéuticas y de biotecnología que estén listadas dentro del Nasdaq Composite. Es requisito que las compañías que se quieran sumar a este índice coticen únicamente dentro del Nasdaq y tener una operatividad superior a las 100 mil acciones (Nasdaq, 2020).

II.1.4. Clasificación de la estadística Multivariada

Existen diferentes técnicas multivariadas posibles de aplicar según las características específicas del estudio, por lo tanto, Álvarez (2006), clasifica a la estadística multivariante, de las siguientes formas:

- Según la estructura de la matriz de datos:
Los métodos pueden clasificarse según sean las variables o los individuos de la matriz de datos base la estructura refiere a si las variables o los individuos son diferentes, o si pertenecen a un grupo o más grupos de variables o de individuos, por lo tanto, pueden existir los siguientes casos:
 1. Sin ninguna estructura en particular
 2. Una estructura entre variables
 3. Una estructura entre individuos
 4. Ambas estructuras
- Según el objetivo deseado:
Los métodos pueden clasificarse en dos grandes grupos; los descriptivos y los inferenciales, pero se pueden ser clasificados de una forma más específica, como se muestra a continuación:

1. Reducción de datos o simplificación estructural: el problema de estudio se debe presentar tan simplemente como se pueda, sin sacrificar información valiosa, y esto hará la interpretación más sencilla.
 2. Selección y agrupamiento: se crean grupos de individuos o variables similares basándose en las características que se midieron. En este caso se requieren reglas para clasificar los individuos en grupos bien definidos.
 3. Investigación de la dependencia entre variables: se está interesado en relaciones entre variables. Las relaciones entre variables deben ser halladas con el propósito de predecir los valores de una o más variables sobre la base de las observaciones de otras variables.
 4. Construcción de hipótesis y prueba de ellas: se desean probar algunas hipótesis estadísticas específicas, formuladas en función de los parámetros de poblaciones multivariadas. Esto debe realizarse para validar las suposiciones o para reforzar algunos supuestos a priori.
- Según la naturaleza de los datos:
 1. Las variables del estudio son cuantitativas
 2. Las variables del estudio son cualitativas
 3. El estudio tiene variables cuantitativas y cualitativas

Varios autores como Pérez (2004) , Castro & Galindo (2000), y De la garza, Morales & Gonzáles (2017), coinciden en que las técnicas estadísticas multivariadas de mayor aplicación para la solución de problemas de diferentes ámbitos, tales como en el área médica, biología, industrial, financiera y económica, son las siguientes:

- Análisis de componentes principales
- Análisis discriminante
- Análisis de grupos jerárquico (análisis clúster)
- análisis de grupos no jerárquicos (análisis K-medias)
- Regresión lineal múltiple

- Regresión logística binaria

Por lo tanto, en la presente investigación se procedió a consultar a diferentes autores, para describir a profundidad cada una de estas técnicas estadísticas multivariadas.

II.1.5. Análisis de componentes principales

El análisis de componentes principales es una herramienta estadística ampliamente utilizada en distintas áreas del conocimiento, en especial en casos donde se tiene un amplio número de datos y variables de estudio. Según Balakrishnama (1999), el ACP es un algoritmo para clasificar elementos en grupos, mediante el cálculo de la varianza y la covarianza, es decir, se calcula la varianza entre cada conjunto y la covarianza entre conjuntos, teniendo como resultado una mejor separación entre cada grupo obtenido.

El ACP se ha utilizado de manera creciente desde hace varias décadas, en una gran variedad de las distintas áreas del conocimiento (Hair J. , 1999).

El ACP realiza dos acciones fundamentales: cuantifica las variables originales y reduce la dimensionalidad de los datos. Cuando se tiene éxito aplicando esta herramienta, cada variable estará muy bien representada (presentará correlación elevada) en una dimensión y deficientemente representada (presentará correlaciones bajas) en las demás, (Jhonson, 2002).

Según Sanguansat (2012), pueden citarse los objetivos específicos que el ACP presenta; explicar la variación del conjunto multivariado (p) con un número reducido de nuevas variables (r), transformar variables originales, en general correlacionadas, en nuevas variables no correlacionadas y finalmente, evaluar, o en su caso, eliminar variables base con poca contribución.

Cada componente principal es una nueva variable, producto de la combinación lineal de las variables originales (Abdi & Williams, 2010).

La transformación lineal para obtener un componente principal se logra ponderando los valores de las variables originales por los coeficientes de un eigenvector de la matriz de covarianzas o correlaciones, por lo que cada eigenvector tiene tantos coeficientes como variables tenga el conjunto de datos del cual se obtiene la matriz. El número de componentes principales que se pueden definir para una matriz de datos “n × p” es igual a “p”, además de que conservan la variabilidad original. La suma de varianzas de los eigenvalores, es igual a la suma de las varianzas de las variables originales, es decir la suma de los eigenvalores es igual a la traza de la matriz de covarianzas.

Un componente principal se define como una combinación lineal de los valores particulares de la entidad “i” (Jhonson, 2002):

$$Y_i = b_1 x_1 + b_2 x_2 + \dots b_p x_p$$

Generalizando:

$$Y_{ij} = b_{1j} x_{i1} + b_{2j} x_{i2} + \dots b_{pj} x_{ip}$$

En términos matriciales:

$$Y_j = Xb_j$$

Donde:

- Y_j es un vector $n \times 1$ que contiene los valores en las n entidades para el componente principal “j”.
- X es la matriz “n x p” de datos originales
- b_j es el vector “p × 1” de coeficientes del componente principal “j” que define la combinación lineal.

II.1.6. Análisis discriminante

Uno de los diversos objetivos que tienen los investigadores de diversos campos del conocimiento, es la clasificación de elementos en determinadas categorías. Por lo



general, el problema de clasificación se presenta en el siguiente panorama: se dispone de dos o más grupos de elementos para los cuales se conocen los valores de diversas variables numéricas y la tarea es reconocer a la o las variables que mejor discriminan los elementos, adicionalmente se establecen reglas que permiten clasificar elementos en uno de los grupos disponibles a partir de los valores que toman en las variables. Las principales técnicas estadísticas que permiten discriminar elementos de una población, están: el análisis discriminante lineal y la regresión logística binaria.

El análisis discriminante lineal es una técnica de clasificación multivariada análoga al análisis de regresión lineal, en donde la variable dependiente no es numérica, sino categórica.

Los principales objetivos de este tipo de análisis, son los siguientes:

1. Construir, a partir de las variables independientes, nuevas variables que mejor discriminen o separen a los elementos en los grupos determinados por las categorías de la variable dependiente. Si la variable dependiente tiene K categorías, el número de estas variables que se puede construir es igual al mínimo entre el número de variables independientes y $k-1$. Si la variable dependiente tiene dos categorías, se podrá construir una nueva variable discriminante.
2. Identificar las variables independientes que permitan una mayor diferenciación de los grupos predeterminados.
3. Generar reglas de clasificación que permitan ubicar nuevos elementos en una de las categorías a partir de sus características numéricas descritas por las variables independientes originales.

El concepto esencial del análisis discriminante fue desarrollado por R.A. Fisher, quien introdujo la técnica, transformando las observaciones multivariadas y resumiéndolas en las llamadas *funciones discriminantes de Fisher*.

Fisher encontró que la combinación lineal que mejor satisface la identificación de variables discriminantes en una ecuación, está dada por:

$$Z = (\mu_1 - \mu_2)^t \sum^{-1} X$$

Donde μ_1 y μ_2 son los vectores de medias de las variables $X_1, X_2, \dots, X_p$ en los grupos respectivos, X^t es el vector $(X_1, X_2, \dots, X_p)$ y $\sum$ es la matriz de covarianzas de tales variables.

Con la función discriminante, se construye una nueva variable construida a partir de las variables independientes originales, con lo que se logra la mejor separación de grupos,

II.1.7. Análisis de grupos jerárquico

El análisis de grupos o análisis clúster, es un método matemático que forma parte de la estadística multivariada; este método es principalmente utilizado para la formación de grupos con características similares a partir de las similitudes o disimilitudes que se presentan entre pares (Everitt, 2002).

Este análisis trata de situar los casos (individuos) en grupos homogéneos, conglomerados o clústeres, no conocidos de antemano, pero sugeridos por la propia esencia de los datos, de manera que individuos que puedan ser considerados similares sean asignados a un mismo grupo, mientras que individuos diferentes se localicen en agrupaciones distintas (De la fuente, 2011).

El análisis clúster define grupos tan distintos como sea posible dependiendo de los datos obtenidos del estudio, la creación de grupos basados en similitud de casos exige una definición de este concepto, o de su complementaria distancia entre individuos (Hennig, 2003).

La creación de grupos basados en similitud de casos exige una definición de este concepto, o de su complementaria distancia entre individuos. La variedad de formas de medir diferencias multivariantes o distancias entre casos proporciona diversas posibilidades de análisis. El empleo de ellas, y el de las que continuamente siguen apareciendo, así como de los algoritmos de clasificación, o diferentes reglas matemáticas para asignar los individuos a distintos grupos, depende del fenómeno estudiado y del conocimiento previo de posible agrupamiento que de él se tenga.

El algoritmo que corresponde a esta técnica se resume en los siguientes pasos:

1. Se parte de tantos conglomerados (clústeres o también llamados grupos) como elementos existan.
2. Se calculan las distancias entre los conglomerados iniciales
3. Con los conglomerados más próximos se forma un nuevo grupo
4. Con los nuevos elementos se procede como en los pasos 2 y 3 hasta obtener un solo grupo formado con todos los elementos.

Los pasos que se siguen para llevar a cabo la partición de los elementos se representan mediante un diagrama llamado dendrograma.

Las técnicas para formar clústeres son diversas, según Véliz (2017), pueden dividirse principalmente en dos tipos; las técnicas jerárquicas aglomerativas con las que es posible formar grupos sucesivos partiendo de tantos grupos como elementos se tengan disponibles, hasta formar un único grupo con todos los elementos. Por otro lado, existen las técnicas no jerárquicas, con las que se forma un número preestablecido de clústeres, destacando el método de K-medias, que permite la generación de un número de “K” clústeres previamente determinados. Por lo tanto, es importante revisar las características de esta variante del análisis de grupos.

II.1.8. Análisis de grupos no jerárquico

El algoritmo de las K-medias, el más importante desde los puntos de vista conceptual y práctico, parte también de unas medias arbitrarias y, mediante pruebas sucesivas, contrasta el efecto que sobre la varianza residual tiene la asignación de cada uno de los casos a cada uno de los grupos. El valor mínimo de varianza determina una configuración de nuevos grupos con sus respectivas medias. Se asignan otra vez todos los casos a estos nuevos centroides en un proceso que se repite hasta que ninguna transferencia puede ya disminuir la varianza residual; o se alcance otro criterio de parada: un número limitado de pasos de iteración o, simplemente, que la diferencia obtenida entre los centroides de dos pasos consecutivos sea menor que un valor prefijado (Moradi Fard, Thonet, & Gaussier, 2020).

El procedimiento configura los grupos maximizando, a su vez, la distancia entre sus centros de gravedad. Como la varianza total es fija, minimizar la residual hace máxima la factorial o intragrupo. Y puesto que minimizar la varianza residual es equivalente a conseguir que sea mínima la suma de distancias al cuadrado desde los casos a la media del clúster al que van a ser asignados, es esta distancia euclídea al cuadrado la utilizada por el método (Gengyuan , 2018).

En resumen, la técnica no jerárquica de K-medias, permite la formación de conglomerados cuyo número K es previamente fijado por el investigador. Los pasos de este análisis pueden resumirse de la siguiente forma:

1. Se forma al azar K grupos y para cada grupo se calculan los centroides o puntos cuyas coordenadas son las medias aritméticas de las variables del estudio.
2. Usando la distancia euclidiana, para cada elemento se calcula su distancia a cada uno de los centroides, reasignándolos al grupo cuyo centroide es el más cercano. Los nuevos centroides de los nuevos grupos formados se recalculan.

3. Si la distancia de los centroides iniciales y los nuevos centroides es pequeña, o si se ha complementado un número fijo de iteraciones, el proceso culmina, de otro modo, se repite el paso dos.

II.1.9. Análisis de regresión lineal

El modelo de regresión lineal es uno de los modelos estadísticos más utilizados en el área de la investigación, en distintos campos del conocimiento. El modelo estadístico de regresión lineal se usa para estudiar la relación lineal que pueda existir entre una variable llamada dependiente y un grupo de variables llamadas independientes o predictoras. El modelo de regresión fue introducido por Laplace (1749-1827) y Gauss (1777-1855), pero posteriormente fue usado por Galton (1822-1911) para explicar la relación entre las estaturas de los padres y las estaturas de los hijos, encontrando lo que denominó como regresión a la media (Reding, López, & Zamora, 2011).

Con esta expresión quería indicar que, en promedio, los hijos de padres altos no eran tan altos como sus padres y que los hijos de padres bajos eran más altos que estos.

Si el modelo tiene una variable independiente, entonces se tiene un modelo de regresión lineal simple, en caso de que se tenga 2 o más variables independientes, se trata de un modelo de regresión lineal múltiple.

Los modelos de regresión lineal se utilizan en diferentes campos y su uso puede resumirse en:

- El análisis de la influencia de ciertas variables en los valores de otra variable.
- La predicción de los valores de una variable cuando se conocen los valores de otras variables.

El modelo de regresión lineal simple relaciona los valores X y Y de dos variables numéricas, mediante la relación:

$$Y = \beta_0 + \beta_1 X + \epsilon$$

En esta relación, Y se llama variable dependiente, X se denomina variable independiente predictora y ϵ es una variable aleatoria que representa la diferencia entre la parte estructural $\beta_0 + \beta_1 x$ y el valor correspondiente Y . La variable ϵ representa errores de medida; sin embargo, a menudo también representa valores de variables no consideradas en el modelo, la inclusión de esta variable aleatoria es primordial pues permite modelar relaciones que no son perfectamente lineales (Milla, 2011).

Para completar el modelo, se supone que los errores representados por esta variable no son correlacionados, que tienen una distribución normal con media 0 y varianza constante σ^2 , y que no son correlacionados con los valores de X .

A partir de estos supuestos, el modelo de regresión puede expresarse como:

$$E(Y | x) = \beta_0 + \beta_1 x$$

Indicando esta relación que para cada valor x de X , el valor esperado de los valores de Y es igual a $\beta_0 + \beta_1 x$.

La pendiente β_1 se interpreta como el cambio del promedio de los valores de Y es para un cambio unitario de x .

Una vez que se han establecido las variables y suponiendo que un modelo de regresión lineal simple puede usarse para corroborar la relación que pudiera existir entre las variables, los pasos a realizarse para la aplicación del modelo son las siguientes:

- Estimación del modelo
- Adecuación del modelo
- Verificación de los supuestos del modelo
- Utilización del modelo

Sin embargo, en diferentes casos de estudio, es posible que no sea suficiente la variable independiente para explicar linealmente una variable dependiente, por lo

tanto, será agregar más variables al análisis. Los modelos de este tipo se llaman modelos de regresión lineal múltiple, mismos que pueden expresarse de la siguiente forma:

$$Y = \beta_0 + \beta_1x + \beta_2x + \beta_3x + \epsilon$$

Donde ϵ es una variable aleatoria con distribución normal de media 0, varianza σ^2 y con valores independientes.

Al igual que el modelo de regresión lineal simple, este modelo permite evaluar la relación que pudiera existir entre un conjunto de variables independientes y una variable dependiente para estudiar la influencia en esta última variable. Para ello será necesario estimar los parámetros del modelo a partir de valores recogidos obtenidos de cada variable independiente.

Como en el caso de la regresión lineal simple, será necesario verificar si se cumplen los supuestos del modelo, por lo que el histograma o la gráfica de ajuste de probabilidad normal juegan un papel importante en esta función, así como de igual manera deben considerarse los diagramas de dispersión de los valores de Y versus los residuales (Baeza & Vázquez, 2014).

Por último, un aspecto a verificar en este análisis, es el referente a la multicolinealidad, problema que existe cuando entre las variables independientes de un modelo, existe una asociación entre ellas. Existen algunas señales que ayudan a detectar multicolinealidad, entre ellas se tiene: la correlación significativa entre dos variables independientes, o la falta de coherencia de los signos de los coeficientes de regresión.

II.1.10. Análisis de regresión logística

El modelo de regresión logística es una extensión del modelo de regresión lineal. En el modelo de regresión lineal se tratan de explicar los valores de una variable

dependiente continua en términos de una combinación lineal de valores independientes continuas, mientras que el modelo de regresión logística binaria se tiene una variable dependiente Y binaria dicotómica, como la compra o no compra de un producto, y se intenta explicar o expresar la probabilidad de que ocurra una de las categorías en términos de una función de la combinación lineal de un grupo de variables, que pueden ser numéricas o categóricas (Solano & Ramírez, 2007).

Las aplicaciones de este tipo de análisis se llevan en diferentes campos del conocimiento, por ejemplo, en la salud se usa para estudiar el riesgo que tiene una persona de tener o no una determinada dolencia, a partir de su perfil médico; en finanzas se utiliza para el riesgo que tiene una entidad bancaria de que sus clientes no paguen un préstamo, a partir de su perfil crediticio.

Por sus características, los modelos de regresión logística permiten dos finalidades:

- Cuantificar la importancia de la relación existente entre cada una de las variables y la variable dependiente, lo que lleva implícito también clarificar la existencia de interacción y confusión entre variables respecto a la variable dependiente, es decir, conocer la odds ratio para cada variable.
- Clasificar individuos dentro de las categorías (presente/ausente) de la variable dependiente, según la probabilidad que tenga de pertenecer a una de ellas dada la presencia de determinadas variables.

En el modelo de regresión logística binaria la variable dependiente debe tomar exactamente dos valores (Sí-No, 0-1, Verdadero-Falso, etc.). Las variables independientes pueden estar a nivel de intervalo o ser categóricas; si son categóricas, deben ser variables dummy o estar codificadas como. En el caso de que exista alguna variable independiente cuantitativa no es necesario que sigan la Ley Normal, pese a que la solución puede ser más estable si los predictores tienen una distribución normal multivariante (Ramírez, Islas, & Valverde, 2011). A pesar de estas condiciones de aplicación poco restrictivas, es muy recomendable que el

modelo esté bien especificado teóricamente. Al igual que otras formas de regresión, la multicolinealidad entre los predictores puede sesgar las interpretaciones. Antes de aplicar el modelo, es conveniente analizar la asociación de la variable respuesta con las variables explicativas, así como entre las propias variables explicativas para poder interpretar los coeficientes obtenidos de forma adecuada. Este análisis previo fundamentará la elección de las variables independientes que se añaden en el modelo (Jhonson, 2002).

En resumen, el modelo debe ser el más reducido que explique los datos (principio de parsimonia), y que además sea congruente e interpretable.

La regresión logística consiste en obtener una función lineal de las variables independientes que permita clasificar a los individuos en una de las dos subpoblaciones o grupos por los dos valores de la variable dependiente. Por lo tanto, la función lineal es muy parecida a la de un modelo de regresión lineal, pero con algunas diferencias que se describirán más adelante (Véliz, 2017).

A partir de los coeficientes de regresión β de las variables independientes introducidas en el modelo se puede obtener directamente la odds ratio de cada una de ellas, que corresponde al riesgo de tener el resultado o efecto evaluado para un determinado valor x respecto al valor disminuido en una unidad ($x - 1$). En otras palabras, si la variable independiente es una variable cuantitativa, la odds ratio que se obtiene representa la probabilidad del evento predicho que tiene un individuo con un valor x frente a la probabilidad que tiene un individuo con valor ($x - 1$).

El modelo logístico se puede escribir de otras formas equivalentes que permiten calcular directamente la probabilidad del proceso binomial para los distintos valores de las variables incluidas en el modelo. Se le denomina función logística la cual calcula la probabilidad de que un individuo pertenezca a una subpoblación, misma que se representa de la siguiente forma:

$$p = \frac{1}{1 + e^{-y}} = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x + \dots + \beta_n x)}}$$

Donde y es la función lineal del modelo de regresión logística, e es la base de logaritmos neperianos y $\beta_1 \dots \beta_n$ son los coeficientes de la ecuación de regresión, que en su conjunto se calcula el resultado para p es la probabilidad de que un caso de estudio pertenezca a una de las dos categorías binarias de la variable dependiente.

II.2. Estado del arte

Existen investigaciones encaminadas a la aplicación de algunas técnicas multivariadas en el ámbito financiero, como lo es la investigación de Quiroga (2016), titulada “Aplicación de dos técnicas del análisis multivariado en el mercado de valores” en la que se utilizaron dos variantes del análisis de componentes principales además del uso del análisis discriminante para lograr una clasificación de 88 empresas emisoras de la Bolsa Mexicana de valores. Una investigación parecida al presente estudio, es el que lleva por título “Análisis de componentes principales funcionales en series de tiempo económicas” (Chávez, 2015), los resultados obtenidos ilustraron la capacidad de los componentes principales para destacar características importantes de un conjunto de series de tiempo económicas, donde fue posible identificar cuatro componentes que reflejan una tendencia general al crecimiento de las ventas a lo largo del período analizado y el efecto de las crisis de 1994- 1995, 2001 y 2008-2009, además de obtener 3 grupos de empresas distintos de acuerdo con lo explicado por los componentes.

Una investigación que utilizó técnicas de análisis multivariado aplicado a variables como lo son las razones financieras, fue el que lleva por título “Aplicación de técnicas estadísticas multivariantes en el tratamiento de la información económico-financiera”, (Mures, 2013) en la que se logró identificar las razones financieras que



tuvieron mayor influencia en dos grupos de empresas en los que el autor previamente clasificó en empresas fracasadas y empresas no fracasadas.

Otra investigación titulada “Selección de índices financieros mediante técnicas estadísticas de los análisis multivariantes” de Pérez (2013), demostró la metodología a aplicar por medio del análisis de componente principales, es útil para la reducción de indicadores de rentabilidad y riesgo financiero.

Respecto a trabajos de investigación en donde se haya aplicado el análisis clúster en el ámbito del análisis financiero, existen documentos que sirvieron de base para otros estudios relacionados como es la publicación que lleva por título “Alternative accounting measures as predictors of failure”, (Beaver, 1968) y “Accounting ratios and the prediction of failure: some behavioral evidence” (Libby, 1975), en los que en ambas investigaciones se hacen agrupaciones de los indicadores financieros de las empresas que se analizan. Trabajos recientes como el denominado “Financial inclusion index: proposal of a multidimensional measure for Mexico” (2013) y “Análisis de indicadores financieros del sector manufacturero del cuero y marroquinería” (2015) evidencian que el análisis clúster es útil para resumir la información referente al desempeño financiero de las organizaciones y que es posible incluso crear indicadores que sean útiles para la toma de decisiones.

Otras investigaciones recientes han involucrado el análisis multivariado, con la finalidad de reducir información obtenida de indicadores financieros, tales como el trabajo titulado “Análisis del comportamiento bursátil de las principales bolsas financieras en el mundo usando el análisis multivariado (análisis de componentes principales PCA) para el periodo de 2011 a 2014”, (2015), en la que los autores justifican la utilización del ACP al presentarse complejidad al tratar de analizar y pronosticar varios indicadores bursátiles pertenecientes a diferentes mercados. Otro ejemplo en el uso de análisis multivariado fue requerido para compactar información financiera, se muestra en la publicación que lleva por título “Análisis de los determinantes de revelación de instrumentos financieros mediante el uso de componentes principales: el caso de empresas mexicanas” en donde los autores

justifican la utilización del ACP para poder robustecer la identificación de variables asociadas con el nivel de cumplimiento de la información financiera a revelar por parte de empresas mexicanas (Vásquez & Mora, 2016).

Un trabajo enfocado a analizar los mercados accionarios, en los que el ACP fue requerido para reducir la información obtenida por los autores, es la que lleva por nombre “The application of principal component analysis and factor analysis to stock markets returns”, (2019) en donde se crea un índice de rentabilidad basado en los resultados anuales de las compañías analizadas. Por último, una reciente publicación titulada “Financial statement analysis: Principal component analysis (PCA) approach case study on China telecoms industry”, (2019), en el cual, se analizaron 18 indicadores que midieron la rentabilidad, la liquidez, el crecimiento de la cartera de clientes y la eficiencia de la administración, de una muestra de empresas obtenidas del sector de comunicaciones de China, concluyendo que solo 12 de dichos indicadores, impactan de manera significativa en el área financiera de este tipo de industria.

III. Criterios metodológicos

Para el desarrollo de la presente investigación, se utilizaron los siguientes criterios metodológicos:

Naturaleza de la investigación

La naturaleza o enfoque de la investigación fue cuantitativa, puesto que las variables independientes (razones financieras) son de tipo continuas, mientras que la variable dependiente (éxito) también fue de tipo continua ya que se obtuvo mediante el cálculo de los diferenciales de precios de las acciones de cada empresa analizada.

Tipo

El tipo de investigación fue de tipo aplicada, debido a que aborda un problema en su etapa inicial y genera alternativas de solución, que permita determinar cuáles son las razones financieras que tienen como consecuencia, el éxito de una empresa.

Carácter

La investigación fue de carácter no experimental ya que no se manipularon las variables de estudio, se trabajó con los estados financieros de las empresas analizadas.

Diseño

El diseño fue longitudinal, correlacional-causal. Longitudinal, ya que se recolectaron datos de los últimos 15 años de las empresas del estudio, que han estado cotizando en la bolsa de valores del Nasdaq. Correlacional-causal, ya que se analizó la asociación de cause efecto entre las 12 variables independientes y la variable dependiente.

Método

El método utilizado será el analítico-sintético, ya que éste consiste en la desmembración de un todo. Descomponiéndolo en sus partes o elementos para observar las causas y efectos, para después relacionar cada relación, mediante la elaboración de una síntesis general del fenómeno estudiado.

Modo

El modo de investigación fue documental ya que la información financiera de las empresas de estudio, se encuentra publicada tanto en los sitios web de cada una de ellas, como en la página web de la bolsa de valores donde cotizan, así como en otros sitios especializados de finanzas y economía.

Lugar

La investigación se llevó a cabo en la ciudad de Chihuahua, Chihuahua, durante los meses de junio de 2017 a octubre de 2020.

Población, unidad de análisis, muestra y marco muestral

El marco muestral de la investigación fue la bolsa de valores del Nasdaq, cuya población se compone un total de 3,300 empresas que cotizan en ella.

La unidad de análisis fueron las empresas que cotizaron en la bolsa de valores del Nasdaq, con las siguientes características:

1. Se encuentren en los sectores (tecnología, farmacéutica, biotecnología y servicios financieros).

El tamaño de la muestra fue el resultado de aplicar un muestreo no probabilístico con un nivel de confianza del 95% y una precisión o error máximo del 10%. El cálculo completo se muestra en tabla 2.

Tabla 2. Cálculo de la muestra

Población	N =	3,300
Alfa (Máximo error tipo I)	$\alpha =$	0.05
Nivel de Confianza	$1-\alpha =$	0.95
Valor Z	Z =	1.96
Probabilidad de éxito	p =	0.5
Complemento de p	q =	0.5
Precisión	d =	0.10
Tamaño de la muestra	n =	93

Al realizar los cálculos correspondientes, se obtuvo que el tamaño de la muestra fue 93 empresas en total. A continuación, se muestran las empresas que fueron elegidas de forma aleatoria, con base al número obtenido en el cálculo de la muestra, junto con una breve descripción de cada una de ellas, además de mostrar las claves con las que fueron codificadas para introducir los datos al programa estadístico SAS.



Empresa: Advanced Micro Devices, Inc.

Codificación: AMD

Compañía estadounidense de semiconductores con sede en Santa Clara, California, que desarrolla procesadores de computación y productos tecnológicos similares de consumo. Sus productos principales incluyen microprocesadores, chipsets para placas base, circuitos integrados auxiliares, procesadores embebidos y procesadores gráficos para servidores, estaciones de trabajo, computadores personales y aplicaciones para sistemas embebidos.

Empresa: Asea Brown Boveri

Codificación: ABB

Corporación multinacional, cuya sede central queda en Zürich, Suiza y cuyos mayores negocios son los de tecnologías en generación de energía eléctrica y en automatización industrial. ABB opera en más de cien países y emplea a más de 135,000 personas.

Empresa: Abbott Laboratories

Codificación: ABT

Compañía farmacéutica estadounidense, fundada en Chicago en 1888 por Wallace Calvin Abbott, con sede en Abbott Park, al norte de Chicago (Illinois). Hoy Abbott cuenta con aproximadamente 90,000 empleados.

Empresa: Automatic Data Processing, Inc.

Codificación: ADP

Empresa proveedora de soluciones de gestión de capital humano para empleadores, que ofrece soluciones a empresas de varios tamaños. La Compañía también ofrece soluciones de subcontratación de procesos comerciales. Sus segmentos incluyen Servicios para empleadores y Servicios para organizaciones profesionales de empleadores.

Empresa: Amgen Inc.



Codificación: AMGN

Compañía de biotecnología estadounidense, con sede central en Thousand Oaks (California), Localizada en Conejo Valley, es una de las empresas más importantes de la zona, siendo una sociedad anónima.

Empresa: ASML Holding N.V.

Codificación: ASML

Empresa neerlandesa dedicada a la fabricación de máquinas para la producción de circuitos integrados. Se trata del mayor proveedor del mundo de sistemas de fotolitografía para la industria de los semiconductores.

Empresa: Activision Blizzard, Inc.

Codificación: ATVI

Activision Blizzard, Inc., anteriormente conocida como Activision, Inc. es una empresa estadounidense de videojuegos que posee Activision creadora de la saga Call of Duty, Guitar Hero, Tony Hawk Skylanders entre otras y Blizzard Entertainment.

Empresa: Becton, Dickinson and Company

Codificación: BDX

Empresa multinacional estadounidense de tecnología médica que fabrica y vende dispositivos médicos, sistemas de instrumentos y reactivos. BD también ofrece servicios de consultoría y análisis en determinadas geografías.

Empresa: Bristol-Myers Squibb Company

Codificación: BMY

Compañía farmacéutica, con sede en la ciudad de Nueva York. La empresa se formó en 1989, tras la fusión de sus predecesores de Bristol-Myers y la Corporación de Squibb.

Empresa: Boston Scientific Corporation

Codificación: BSX

Fabricante de dispositivos médicos utilizados en especialidades médicas intervencionistas, que incluyen radiología intervencionista, cardiología intervencionista, intervenciones periféricas, neuromodulación, intervención neurovascular, electrofisiología, cirugía cardíaca, cirugía vascular, endoscopia, oncología, urología y ginecología.

Empresa: Biogen Inc.**Codificación: BIIB**

Compañía multinacional estadounidense de biotecnología con sede en Cambridge, Massachusetts, que se especializa en el descubrimiento, desarrollo y suministro de terapias para el tratamiento de enfermedades neurológicas a pacientes de todo el mundo.

Empresa: Baxter International Inc.**Codificación: BAX**

Empresa multinacional estadounidense de atención médica con sede en Deerfield, Illinois. La empresa se centra principalmente en productos para tratar la hemofilia, las enfermedades renales, los trastornos inmunitarios y otras afecciones médicas crónicas y agudas.

Empresa: Caterpillar Inc.**Codificación: CAT**

Corporación de Estados Unidos con sede central en Peoria, Illinois. Caterpillar es el fabricante más grande del mundo de maquinaria para la construcción y equipos de minería, motores diésel y turbinas industriales de gas.

Empresa: CVS Health Corporation**Codificación: CVS**

Compañía estadounidense de atención médica que posee CVS Pharmacy, una cadena de farmacias minoristas, CVS Caremark, un administrador de beneficios de



farmacia, Aetna, un proveedor de seguros de salud, entre muchas otras marcas. La sede de la compañía se encuentra en Woonsocket, Rhode Island.

Empresa: Canadian National Railway Company

Codificación: CNI

Compañía ferroviaria canadiense de transporte de mercancías, siendo la más importante de Canadá tanto por la extensión de su red como por la magnitud de negocios.

Empresa: CSX Corporation

Codificación: CSX

Empresa enfocada al transporte ferroviario y bienes raíces en Norteamérica, entre otras industrias. La compañía se estableció en 1980 como parte de la fusión de Chessie System y Seaboard Coast Line Industries. Los diversos ferrocarriles del antiguo Chessie System y Seaboard Coast Line Industries que ahora son propiedad de CSX Corporation finalmente se fusionaron en una sola línea en 1986 y se conoció como CSX Transportation.

Empresa: Danaher Corporation

Codificación: DHR

Danaher Corporation es un conglomerado estadounidense diversificado a nivel mundial con sede en Washington, D.C. La empresa diseña, fabrica y comercializa productos y servicios profesionales, médicos, industriales y comerciales.

Empresa: DHI Group, Inc.

Codificación: DHX

DHI Group, Inc. proporciona datos, conocimientos y conexiones laborales a través de servicios especializados para comunidades profesionales, que incluyen tecnología y autorización de seguridad, servicios financieros y energía. La Compañía opera a través de su segmento centrado en la tecnología.

Empresa: Ecolab Inc.**Codificación: ECL**

Proveedor global estadounidense de soluciones y servicios de agua, higiene y prevención de infecciones para los mercados de alimentos, atención médica, hotelería e industrial.

Empresa: Emerson Electric Co.**Codificación: EMR**

Corporación multinacional estadounidense con sede en Ferguson, Missouri, Estados Unidos, fabrica productos y brinda servicios de ingeniería para una amplia gama de mercados industriales, comerciales y de consumo.

Empresa: The Estée Lauder Companies Inc.**Codificación: EL**

Fabricante y comercializador multinacional estadounidense de prestigiosos productos para el cuidado de la piel, maquillaje, fragancias y cuidado del cabello, con sede en Midtown Manhattan, Nueva York.

Empresa: Eaton Corporation plc**Codificación: ETN**

Compañía de gestión de energía. Eaton adquirió Cooper Industries plc en noviembre de 2012. Los ingresos en 2012 de las compañías combinadas fueron de \$ 21,8 mil millones en una base pro forma. Eaton cuenta con aproximadamente 102.000 empleados y vende sus productos a clientes en más de 175 países.

Empresa: Exelon Corporation**Codificación: EXC**

Compañía gigante en la generación y distribución de electricidad con sede en Chicago. Fue creada en octubre de 2000 por la fusión de PECO Energy Company, de Filadelfia y Unicom, de Chicago. Exelon tiene 5,2 millones de consumidores eléctricos y, en Filadelfia, 460.000 clientes de gas natural.

Empresa: Eversource Energy**Codificación: ES**

Compañía de energía con sede en Hartford, Connecticut y Boston, Massachusetts, con varias subsidiarias reguladas que ofrecen servicios de electricidad, gas natural y agua al por menor a aproximadamente 4 millones de clientes en Connecticut, Massachusetts y New Hampshire.

Empresa: Freeport Inc.**Codificación: EW**

Compañía estadounidense de tecnología médica con sede en Irvine, CA, que se especializa en válvulas cardíacas artificiales y monitoreo hemodinámico.

Empresa: Edwards Lifesciences Corporation**Codificación: FCX**

Productor de más bajo costo de cobre del mundo y uno de los mayores productores mundiales de oro.

Empresa: FirstEnergy Corp.**Codificación: FE**

Empresa de servicios eléctricos con sede en Akron, Ohio. Sus subsidiarias y afiliadas están involucradas en la distribución, transmisión y generación de electricidad, así como en la administración de energía y otros servicios relacionados con la energía.

Empresa: Fresenius Medical Care AG & Co. KGaA**Codificación: FMS**

Empresa alemana especializada en la producción de suministros médicos, en especial para diálisis renales.

Empresa: Gilead Sciences, Inc.**Codificación: GILD**

Compañía biofarmacéutica estadounidense con sede en Foster City, California, que investiga, desarrolla y comercializa medicamentos.

Empresa: General Dynamics Corporation

Codificación: GD

Compañía conglomerado de empresas estadounidense del sector aeroespacial y militar, fruto de la fusión de numerosas compañías.

Empresa: General Mills, Inc.

Codificación: GIS

Compañía estadounidense, principalmente relacionada con productos alimenticios, que tiene su sede en Golden Valley, Minnesota, un suburbio de Mineápolis.

Empresa: Corning Incorporated

Codificación: GLW

Compañía estadounidense de tecnología multinacional que se especializa en vidrio especializado, cerámica y materiales y tecnologías relacionados, incluida la óptica avanzada, principalmente para aplicaciones industriales y científicas.

Empresa: CGI Inc.

Codificación: GIB

Empresa canadiense de integración de sistemas y consultoría de tecnología de la información global con sede en Montreal, Quebec, Canadá.

Empresa: Genuine Parts Company

Codificación: GPC

Empresa dedicada a la distribución de piezas de repuesto para automóviles, piezas de repuesto industriales, productos de oficina y materiales eléctricos / electrónicos.

Empresa: Garmin Ltd.

Codificación: GRMN



Empresa que desarrolla y fabrica dispositivos de GPS para el ámbito civil, principalmente para tránsito terrestre, aunque también naval y aéreo. Garmin Ltd., con sede en George Town.

Empresa: Hologic, Inc.

Codificación: HOLX

Compañía de tecnología médica enfocada principalmente en la salud de las mujeres; vende dispositivos médicos para diagnóstico, cirugía e imagen médica.

Empresa: HEICO Corporation

Codificación: HEI

Empresa aeroespacial y electrónica que se centra en nichos de mercado. Los productos HEICO se encuentran en aviones, naves espaciales, equipos de defensa, equipos médicos y sistemas de telecomunicaciones.

Empresa: Halliburton Company

Codificación: HAL

Corporación estadounidense que, dedicada a la prestación de servicios en los yacimientos petroleros, actualmente desempeña su labor en más de setenta países.

Empresa: HP Inc.

Codificación: HPQ

Empresa estadounidense con sede en Palo Alto, California, surgida de la separación de Hewlett-Packard, líder en venta de impresoras en el mundo y la segunda a nivel mundial en venta de computadores personales y portátiles.

Empresa: IDEXX Laboratories, Inc.

Codificación: IDXX

Empresa estadounidense que se dedica al desarrollo, fabricación y distribución de productos y servicios para los mercados de animales de compañía, veterinaria, ganadería y avicultura, análisis de agua y productos lácteos.

Empresa: Illumina, Inc.**Codificación: ILMN**

Empresa estadounidense. Incorporada en abril de 1998, Illumina desarrolla, fabrica y comercializa sistemas integrados para el análisis de la variación genética y la función biológica.

Empresa: Incyte Corporation**Codificación: INCY**

Compañía farmacéutica estadounidense con sede en Alapocas, Delaware. La compañía fue fundada en Palo Alto, California en 1991 y se hizo pública en 1993. Incyte tiene un medicamento, Jakafi, que ha sido aprobado por la Administración de Alimentos y Medicamentos de EE. UU.

Empresa: Johnson Controls International plc**Codificación: JCI**

Empresa en Irlanda con sede en Cork, Irlanda, que produce equipos contra incendios, HVAC y de seguridad para edificios.

Empresa: KLA Corporation**Codificación: KLAC**

Empresa de bienes de capital con sede en Milpitas, California. Suministra sistemas de gestión de rendimiento y control de procesos para la industria de los semiconductores y otras industrias relacionadas con la nanoelectrónica.

Empresa: Korea Electric Power Corporation**Codificación: KEP**

Empresa de bienes de capital con sede en Milpitas, California. Suministra sistemas de gestión de rendimiento y control de procesos para la industria de los semiconductores y otras industrias relacionadas con la nanoelectrónica.

Empresa: L3-Harris Technologies, Inc.**Codificación: LHX**

Compañía de tecnología estadounidense, contratista de defensa y proveedor de servicios de tecnología de la información que produce sistemas y productos C6ISR, equipos inalámbricos, radios tácticas, aviónica y sistemas electrónicos, equipos de visión nocturna y antenas terrestres y espaciales para uso en el gobierno, defensa y comercio, se especializan en soluciones de vigilancia.

Empresa: Alliant Energy Corporation

Codificación: LNT

Compañía tenedora de servicios públicos que se incorporó en Madison, Wisconsin en 1981. Consta de cuatro subsidiarias: Interstate Power and Light Company es una empresa de servicios públicos que genera y distribuye electricidad y distribuye y transporta gas natural en Iowa.

Empresa: Medtronic plc

Codificación: MDT

Empresa estadounidense de dispositivos médicos con domicilio en Irlanda que genera la mayoría de sus ventas y beneficios del sistema de atención médica de los EE. UU., Pero tiene su sede en la República de Irlanda a efectos fiscales.

Empresa: Micron Technology, Inc.

Codificación: MU

Empresa multinacional con sede en Boise, Idaho, Estados Unidos, que produce muchas formas de semiconductores. Esto incluye DRAM, SDRAM, memoria flash, y SSD. Sus productos de consumo se comercializan bajo la marca Crucial Technology y Lexar.

Empresa: NVIDIA Corporation

Codificación: NVDA

Empresa multinacional especializada en el desarrollo de unidades de procesamiento gráfico y tecnologías de circuitos integrados para estaciones de trabajo, ordenadores personales y dispositivos móviles.

Empresa: Novo Nordisk A/S

Codificación: NVO

Empresa farmacéutica danesa. Surge de la fusión de las compañías danesas Novo Industri A/S y Nordisk Gentofte A/S en 1989. Es una de las empresas líderes en el segmento de las insulinas y cuidados de la diabetes en general.

Empresa: NextEra Energy, Inc.**Codificación: NEE**

Empresa energética la cual se encuentra dentro de la lista Fortune 500 con aproximadamente 45,900 megavatios de capacidad de generación, sus ingresos son de más de \$ 17 mil millones y aproximadamente tiene 14,000 empleados en todo Estados Unidos y Canadá.

Empresa: Nutrien Ltd.**Codificación: NTR**

Empresa canadiense de fertilizantes con sede en Saskatoon, Saskatchewan. Es el mayor productor de potasa y el tercer mayor productor de fertilizantes nitrogenados del mundo. Cuenta con 1,500 tiendas minoristas y más de 20,000 empleados.

Empresa: ONEOK, Inc.**Codificación: OKE**

Empresa diversificada Fortune 500 con sede en Tulsa, Oklahoma. ONEOK se fundó en 1906 como Oklahoma Natural Gas Company, pero cambió su nombre corporativo a ONEOK en diciembre de 1980.

Empresa: Omnicom Group Inc.**Codificación: OMC**

Empresa diversificada Fortune 500 con sede en Tulsa, Oklahoma. ONEOK se fundó en 1906 como Oklahoma Natural Gas Company, pero cambió su nombre corporativo a ONEOK en diciembre de 1980.

Empresa: ON Semiconductor Corporation**Codificación: ON**



Empresa proveedora de semiconductores y otros productos derivados para dispositivos lógicos, discretos y personalizados para aplicaciones de automoción, comunicaciones, informática, consumo, industrial, iluminación LED, médicas, militares / aeroespaciales y de energía.

Empresa: OGE Energy Corp.

Codificación: OGE

Compañía de servicios eléctricos regulada que atiende a más de 843,000 clientes en Oklahoma y Arkansas, incluidos 1,5 millones de personas en el área metropolitana de la ciudad de Oklahoma.

Empresa: Pfizer Inc.

Codificación: PFE

Empresa farmacéutica estadounidense que después de diversas fusiones llevadas a cabo con Pharmacia and Upjohn y Parke Davis, es el laboratorio líder a nivel mundial en el sector farmacéutico. La sociedad tiene su sede central en Nueva York.

Empresa: Plug Power Inc.

Codificación: PLUG

Plug Power Inc. es una empresa estadounidense dedicada al diseño y fabricación de sistemas de pilas de combustible de hidrógeno que reemplazan las baterías convencionales en equipos y vehículos que funcionan con electricidad.

Empresa: QUALCOMM Incorporated.

Codificación: QCOM

Compañía estadounidense fundada en 1985 que produce chipsets para la tecnología móvil CDMA y W-CDMA. Y son los creadores y diseñadores de los procesadores Snapdragon También es responsable del cliente de correo electrónico Eudora.

Empresa: QIAGEN N.V.

Codificación: QGEN

Compañía proveedora de tecnologías de muestra y ensayo para diagnóstico molecular, pruebas aplicadas, investigación académica y farmacéutica. Consolidada bajo el holding holandés QIAGEN N.V., la compañía opera más de 35 oficinas en más de 25 países.

Empresa: Quidel

Codificación: QDEL

Fabricante estadounidense de productos de diagnóstico para la salud que se venden en todo el mundo. El 8 de mayo de 2020, la Administración de Alimentos y Medicamentos de EE. UU.

Empresa: Qurate Retail, Inc.

Codificación: QRTEB

Empresa que posee una red estadounidense de televisión abierta y un canal insignia de compras que se especializa en compras televisadas en el hogar.

Empresa: Raytheon Technologies Corporation

Codificación: RTX

Conglomerado multinacional estadounidense con sede en Waltham, Massachusetts. La compañía es uno de los fabricantes de defensa y aeroespacial más grandes del mundo por ingresos y capitalización de mercado.

Empresa: Regeneron Pharmaceuticals, Inc.

Codificación: REGN

Compañía estadounidense de biotecnología con sede en Tarrytown, Nueva York, según el sitio web de la compañía.

Empresa: Roper Technologies, Inc.

Codificación: ROP



Compañía estadounidense de biotecnología con sede en Tarrytown, Nueva York, según el sitio web de la compañía.

Empresa: Rockwell Automation, Inc.

Codificación: ROK

Compañía proveedora estadounidense de automatización industrial y tecnología de la información. Las marcas incluyen el software Allen-Bradley y FactoryTalk.

Empresa: ResMed Inc.

Codificación: RMD

Empresa de equipos médicos con sede en San Diego, California. Proporciona principalmente dispositivos médicos que se pueden conectar a la nube para el tratamiento de la apnea del sueño, la enfermedad pulmonar obstructiva crónica y otras afecciones respiratorias.

Empresa: Rogers Communications Inc.

Codificación: RCI

Compañía canadiense de comunicaciones y medios. Opera principalmente en los campos de comunicaciones inalámbricas, televisión por cable, telefonía y conectividad a Internet, con importantes activos adicionales de telecomunicaciones y medios de comunicación.

Empresa: Sanofi

Codificación: SNY

Compañía farmacéutica europea nacido en 2004 de la fusión de la francesa Sanofi-Synthelabo y la francoalemana Aventis (que fue, a su vez, el resultado de la fusión de la empresa alemana Hoechst y la francesa Rhône Poulenc). Sanofi-Aventis es una de las compañías farmacéuticas líderes en el mundo, la primera en Europa y la tercera a nivel mundial.

Empresa: Stryker Corporation

Codificación: SYK



Compañía multinacional estadounidense de tecnologías médicas con sede en Kalamazoo, Michigan.

Empresa: Taiwan Semiconductor Manufacturing Company Limited

Codificación: TSM

Empresa de fundición de semiconductores más grande del mundo, con la sede y sus principales operaciones localizadas en el Parque científico e industrial de Hsinchu en Hsinchu, Taiwán.

Empresa: Texas Instruments Incorporated

Codificación: TXN

Empresa estadounidense con sede en Dallas que desarrolla y comercializa semiconductores y tecnología para ordenadores.

Empresa: TOTAL SE

Codificación: TOT

Grupo empresarial del sector petroquímico y energético con sede mundial en La Défense. Su actividad se encuentra presente en más de 130 países, empleando a unas 111 000 personas.

Empresa: TC Energy Corporation

Codificación: TRP

Compañía de energía de América del Norte, con sede en Calgary, Alberta, Canadá, que desarrolla y opera infraestructura energética en Canadá, Estados Unidos y México. La compañía opera tres negocios principales: gasoductos de gas natural, tuberías de líquidos y energía.

Empresa: TE Connectivity Ltd.

Codificación: TEL

Líder global de componentes electrónicos de ingeniería, soluciones de redes, sistemas de telecomunicación marinos y bienes de consumo en más de 150 países.

Empresa: Universal Health Services, Inc.

Codificación: UHS

Compañía estadounidense Fortune 500 que brinda servicios hospitalarios y de atención médica, con sede en King of Prussia, Pennsylvania. En 2019, sus ingresos anuales fueron de \$ 11.37 mil millones.

Empresa: United Microelectronics Corporation**Codificación: UMC**

Compañía taiwanesa que se dedica al desarrollo de circuitos integrados. Fue fundada en 1980, siendo la primera empresa productora de semiconductores de Taiwán, como una empresa derivada del Industrial Technology Research Institute patrocinado por el gobierno.

Empresa: UGI Corporation**Codificación: UGI**

Empresa de distribución de energía eléctrica y gas natural con sede en King of Prussia, Pensilvania, con amplias operaciones en Estados Unidos y Europa. UGI es propietaria de AmeriGas, el mayor comercializador de propano de Estados Unidos.

Empresa: United Therapeutics Corporation**Codificación: UTHR**

Empresa de biotecnología centrada en el desarrollo y la comercialización de productos únicos para abordar las necesidades médicas insatisfechas de los pacientes con enfermedades crónicas y potencialmente mortales.

Empresa: Verizon Communications Inc.**Codificación: VZ**

Compañía global de banda ancha y telecomunicaciones y parte del Índice Dow Jones. La compañía fue fundada en 1983 como Bell Atlantic con un tamaño que abarcaba de Nueva Jersey a Virginia, y surgió como parte de la desintegración, en 1984 de AT&T en siete "Baby Bells".

Empresa: Vertex Pharmaceuticals Incorporated**Codificación: VRTX**

Compañía biofarmacéutica estadounidense con sede en Boston, Massachusetts. Fue una de las primeras empresas de biotecnología en utilizar una estrategia explícita de diseño racional de medicamentos en lugar de la química combinatoria.

Empresa: Vodafone Group Plc**Codificación: VOD**

Operador de telefonía móvil, telefonía fija, banda ancha y televisión digital, multinacional con sede central en Newbury, Berkshire, Reino Unido.

Empresa: Valero Energy Corporation**Codificación: VLO**

Fabricante y comercializador internacional Fortune 500 de combustibles para el transporte, otros productos petroquímicos y energía. Tiene su sede en San Antonio, Texas, Estados Unidos.

Empresa: Varian Medical Systems, Inc.**Codificación: VAR**

Fabricante estadounidense de software y tratamientos de oncología radioterápica con sede en Palo Alto, California. Sus dispositivos médicos incluyen aceleradores lineales y software para tratar el cáncer y otras condiciones médicas con radioterapia, radiocirugía, terapia de protones y braquiterapia.

Empresa: Waste Management, Inc.**Codificación: WM**

Empresa estadounidense de gestión de residuos, residuos integrales y servicios ambientales en América del Norte. Fundada en 1968, la compañía tiene su sede en First City Tower en Houston, Texas.

Empresa: Wipro Limited



Codificación: WIT

Corporación multinacional india que brinda servicios de tecnología de la información, consultoría y procesos comerciales. Tiene su sede en Bangalore, Karnataka, India. En 2013, Wipro separó sus negocios que no son de TI y formó Wipro Enterprises, de propiedad privada.

Empresa: Watsco, Inc.

Codificación: WSO

Distribuidor de equipos de aire acondicionado, calefacción y refrigeración y piezas y suministros relacionados en los Estados Unidos. Watsco se fundó hace más de 60 años como fabricante de piezas, componentes y herramientas utilizados en la industria HVAC / R.

Empresa: Wabtec Corporation

Codificación: WAB

Compañía estadounidense formada por la fusión de Westinghouse Air Brake Company y MotivePower Industries Corporation en 1999. Tiene su sede en Pittsburgh, Pensilvania, EE. UU.

Empresa: Xilinx, Inc.

Codificación: XLNX

Compañía de tecnología estadounidense, primeramente, es un distribuidor de dispositivos lógicos programables. Es reconocida por inventar los FPGA y también por ser la primera compañía con modelos de manufactura fabless.

Empresa: Zimmer Biomet Holdings, Inc.

Codificación: ZBH

Empresa de dispositivos médicos que cotiza en bolsa. Fue fundada en 1927 para producir tablas de aluminio. La empresa tiene su sede en Varsovia, Indiana, donde forma parte del grupo empresarial de dispositivos médicos.

Empresa: Zebra Technologies Corporation

Codificación: ZBRA

Fabricante de impresoras de etiquetas de código de barras, impresoras/codificadoras de etiquetas RFID e impresoras de tarjetas. La empresa tiene la base situada en Vernon Hills, Illinois, Estados Unidos de América

Variables

En esta sección se especifican las variables independientes y variable dependiente de la investigación.

Variables independientes

Para el estudio se utilizaron 12 razones financieras de las 93 empresas seleccionadas, que fueron calculados en base a los estados financieros de periodos semestrales.

Se analizó un periodo total de 15 años, obteniendo primeramente los indicadores de las razones financieras de forma semestral, para luego obtener un promedio anual.

La codificación utilizada para el ingreso de los datos al programa estadístico SAS, referente a las razones financieras, se muestra en la tabla 3:

Tabla 3. Codificación de razones financieras

Razones de liquidez		Razones de actividad	
RC	Razón circulante	RRI	Razón de rotación de inventario
RA	Razón rápida	RRDI	Razón de días de venta en inventario
		RCXC	Razón de rotación de cuentas por cobrar
		RDCXC	Razón de días en cuentas por cobrar
		RRA	Razón de rotación de activos totales
Razones de rentabilidad		Razones de apalancamiento	
RMU	Razón de margen de utilidad	RAT	Razón de margen de deuda total
ROA	Razón de margen de rendimiento sobre activos	RCI	Razón de cobertura de intereses
ROE	Razón de margen de rendimiento sobre capital		

Fuente: Elaboración propia

Variable dependiente

Para efecto de determinar que una empresa ha sido exitosa, el criterio fue que la tasa de rendimiento anual de la empresa haya superado al menos el 50% la tasa de rendimiento anual de la bolsa del NASDAQ, en los últimos 15 años.

Por lo tanto, se procedió a formar dos grupos de empresas:

- Grupo de éxito
- Grupo de no éxito

Para determinar si una empresa fue exitosa o no, se procedió a calcular la media geométrica de rendimiento del Índice Nasdaq-100 y el de las 93 empresas del estudio, por lo que el criterio para determinar el éxito fue el siguiente:

Si es la media geométrica de rendimiento de la empresa fue mayor al menos en un 50%, respecto a la media geométrica del índice Nasdaq-100, la empresa es del grupo de empresas exitosas.

La tasa de crecimiento anual para las empresas que cotizan en el Nasdaq, fue de un 9.91%, es con base en el crecimiento que tuvo el índice Nasdaq-100, en los 15 años que contempla la presente investigación. Por lo tanto, para que una empresa fuera agrupada en el grupo de empresas de éxito, debió tener una tasa de crecimiento de al menos un 14.85%, por lo que después de realizar el cálculo de la media geométrica de crecimiento para las 93 empresas, se obtuvo un total de 25 empresas que superaron en al menos un 50% el crecimiento general de la bolsa. La tabla 4 muestra las 25 empresas clasificadas en el grupo de éxito.

Tabla 4: Empresas del grupo de éxito

Tasa de rendimiento anual	Empresa
28.67%	ILMN
26.80%	REGN
24.71%	NVO
23.76%	NVDA
22.43%	HEI
21.73%	ASML
21.18%	EW
20.75%	VRTX
19.54%	IDXX
18.92%	QDEL
17.96%	CSX
17.73%	RMD
17.49%	ROP
17.29%	GIB
17.21%	TSM
17.18%	DHX
16.66%	OKE
16.42%	DHR
16.21%	EL
15.78%	ATVI
15.8%	NEE
15.36%	INCY
14.96%	GILD
14.89%	WAB
14.83%	LHX

Fuente: Elaboración propia con el paquete estadístico SAS

Análisis estadístico

Una vez obtenidos los valores promedios de los últimos 15 años de las razones financieras de las 93 empresas se procedió a correr las siguientes pruebas y análisis estadísticos:

- Prueba de medida Kaiser-Meyer-Olkin (KMO)
- Prueba de esfericidad de Barlett
- Análisis de componentes principales
- Análisis discriminante
- Análisis de grupos jerárquico (análisis clúster)
- análisis de grupos no jerárquicos (análisis K-medias)
- Regresión logística binaria
- Regresión lineal múltiple

IV. Resultados y discusión

IV.1. Estadísticos descriptivos

En la primera fase de la investigación, se obtuvieron un total de 1,116 observaciones (12 razones financieras promediadas de forma anual para cada una de las 93 empresas seleccionadas), la tabla 1 muestra los estadísticos descriptivos de para las 12 variables independientes del estudio.

La tabla 5 muestra los estadísticos descriptivos para las 12 razones financieras analizadas en las 93 empresas seleccionadas, la columna 1 muestra la codificación de la razón analizada, la columna 2 muestra el conteo total, la columna 3 muestra la media de los valores obtenidos, la columna 4 muestra la desviación estándar, la columna 5 muestra la varianza, la columna 6 muestra el coeficiente de variación, la columna 7 muestra el valor mínimo obtenido calculados y finalmente la columna 8 muestra el valor máximo obtenido para cada una de las variables.

Tabla 5: Estadísticos descriptivos

Variable	Conteo total	Media	Desv.Est.	Varianza	Coef. Var	Mínimo	Máximo
RC	93	2.161	1.215	1.475	56.19	0.544	5.325
RA	93	1.4336	0.8691	0.7553	60.62	0.3046	4.7397
RRI	93	8.41	12.34	152.37	146.75	1.1	77.18
RRDI	93	89.14	66.96	4483.96	75.12	4.73	332.08

RCXC	93	14.01	64.23	4125.61	458.43	2.03	626.13
RDCXC	93	68.47	65.1	4238.31	95.09	18.22	475.8
RRA	93	10.6901	0.3659	0.1339	53.02	0.191	2.3786
RMU	93	39.673	7.754	60.126	61.19	0.88	34.086
ROA	93	69.198	5.971	35.654	72.83	1.159	45.3
ROE	93	19.02	14.43	208.2	75.87	2.49	82.72
RAT	93	9.3578	0.2023	0.0409	56.54	0.0066	1.2478
RCI	93	4.742	1.002	1.004	21.12	1.092	9.04

Fuente: Elaboración propia con el paquete estadístico SAS

IV.2. Prueba de medida Kaiser-Meyer-Olkin y prueba de esfericidad de Barlett

Primeramente, se le aplicó al conjunto de datos, la medida de adecuación de muestreo de Kaiser-Meyer-Olkin (KMO) y la prueba de esfericidad de Barlett, utilizando el paquete estadístico SAS, para cerciorarse de que los datos del estudio fueran aptos para analizarse por medio de una técnica estadística multivariada de reducción de variables.

Diversos autores sugieren que este tipo de pruebas, componen el primer filtro previo a la aplicación de otras técnicas estadísticas multivariadas (Méndez & Rondón, 2012) y (Urrutia & Palomino, 2010).

La medida de adecuación de muestreo de Kaiser-Meyer-Olkin es un estadístico que indica la proporción de varianza en las variables que podría ser causada por factores subyacentes. Los valores altos (cerca de 1.0) generalmente indican que un análisis de reducción de variables, puede ser útil con los datos. Si el valor es inferior a 0.50, los resultados de este tipo de análisis no serían confiables, por otro lado, la prueba de esfericidad de Bartlett prueba la hipótesis de que la matriz de correlación es una matriz de identidad, lo que indicaría que las variables no están relacionadas y por lo tanto, no son adecuadas para la detección de estructuras. Los valores pequeños (menos de 0.05) del nivel de significancia indican que un análisis de reducción de variables, es útil con los datos (SPSS statics, s.f.). Los resultados de la aplicación de estas pruebas se muestran en la tabla 6:

Tabla 6: KMO y Esfericidad de Barlett

Medida Kaiser-Meyer-Olkin de adecuación de muestreo		0.8572
Prueba de esfericidad de Barlett	Chi cuadrada	282.93
	gl	91
	Sig.	0.000

Fuente: Elaboración propia con el paquete estadístico SAS

Con los resultados obtenidos se encuentra que es viable aplicar un análisis de reducción de variables, debido a que la medida KMO presenta un valor muy cercano a 1.0 (0.8572), mientras que la significancia de la prueba de esfericidad de Barlett, presenta valor de $p < 0.05$, con lo que se rechaza la hipótesis nula de que la matriz de correlación es una matriz identidad.

IV.3. Análisis de componentes principales

En la segunda fase del análisis se procedió a aplicar distintas técnicas estadísticas multivariadas, comenzando con el análisis de componentes principales, para lo cual el primer paso para detectar que esta técnica estadística es apta para resumir la variabilidad del conjunto de datos y por lo tanto realizar agrupaciones o clasificaciones, es analizar los eigenvalores obtenidos.

La tabla 7 muestra las cantidades de los eigenvalores obtenidos para cada uno de los 12 componentes principales obtenidos, después de correr el procedimiento PROC PRINT del paquete estadístico SAS. A mayor eigenvalor, indica que el componente principal al cual pertenece, es capaz de resumir una mayor proporción de la variabilidad total explicada por el conjunto de todas las variables que se analizan. La segunda columna de la tabla muestra un alto valor para el primer componente principal teniendo un eigenvalor 3.3260, el cual con base al valor mostrado en la columna 5, es capaz de resumir la variabilidad total del estudio en un 25.02%. Los Eigenvalores 2, 3 y 4 también obtienen altos valores, siendo de 1.729, 1.4096 y 1.2113 respectivamente, por lo que al observar el porcentaje acumulado de resumen variabilidad total, estos 4 Eigenvalores son capaces de

resumir el 81.88% de la variabilidad total conformada por las 12 variables que se analizan en total.

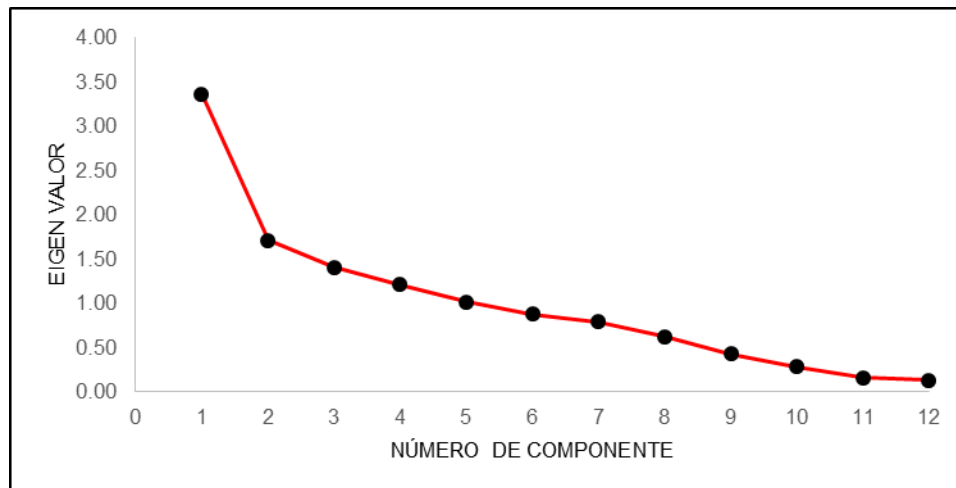
Tabla 7: Eigenvalores

NUMERACIÓN	AUTOVALOR	DIFERENCIA	PROPORCIÓN	ACUMULADO
1	3.3620	1.6492	0.2502	0.2502
2	1.7129	0.3033	0.2127	0.4629
3	1.4096	0.1983	0.1950	0.6579
4	1.2113	0.1977	0.1609	0.8188
5	1.0136	0.1374	0.0559	0.8747
6	0.8762	0.0857	0.0345	0.9092
7	0.7905	0.1662	0.0322	0.9414
8	0.6242	0.1986	0.0320	0.9734
9	0.4257	0.1442	0.0211	0.9945
10	0.2815	0.1190	0.0025	0.9970
11	0.1624	0.0323	0.0022	0.9992
12	0.1302	0.0000	0.0008	1.0000

Fuente: Elaboración propia con el paquete estadístico SAS

Una mejor forma de apreciar la magnitud de los eigenvalores, es por medio de una gráfica de sedimentación. En la figura 1, se aprecia el alto valor del primer eigenvalor obtenido para el componente principal 1, el cual fue de 3.3620, representando el 25.02% de la variabilidad total. Sin embargo, teniendo en cuenta los primeros 4 eigenvectores, estos llegan a explicar el 81.88% de la variabilidad total, mientras que los demás 8 eigenvalores presentan valores muy inferiores, representando en su conjunto el 18.12% restante de la variabilidad total. Este tipo de gráficas permiten apreciar de forma visual, la magnitud y diferencia que existen entre los diversos Eigenvalores que pertenecen a los distintos componentes principales, permitiendo corroborar si es pertinente utilizar alguno para la agrupación de las variables.

Figura 1: Gráfica de sedimentación



Fuente: Elaboración propia con el paquete estadístico SAS

En la tabla 8 se muestran las correlaciones entre las 12 variables analizadas y los primeros 5 componentes principales obtenidos. Puede observarse que las 2 razones financieras del rubro de liquidez presentan altos valores de correlación con el componente principal 1, respecto a las razones del rubro de rentabilidad, estas razones presentan altos niveles de correlación con el componente principal 2, mientras que para las razones de actividad, 4 de las razones de este rubro presentan altos niveles de correlación con el componente principal 3, mientras que las 2 razones de apalancamiento presentan altos niveles de correlación con el componente principal 4. La significancia estadística de estos niveles de correlación se confirma al presentar valores de $p < 0.05$, con lo que se confirma que la técnica de ACP, es capaz de reducir el número de variables para analizar la información, ya que es posible utilizar los 4 componentes principales obtenidos para resumir la variabilidad que contienen las 12 variables originales del estudio.

Tabla 8: Correlaciones de los 3 componentes principales

INDICADOR	CP-1	CP-2	CP-3	CP-4	CP-5	CP-6
RC	0.9546	0.1264	-0.0151	-0.1009	0.0480	-0.4089
P value	<.0001	0.1235	0.1244	0.0944	0.1244	0.1247
RA	0.8422	0.0581	0.0381	-0.0793	0.0309	-0.4741
P value	0.0001	0.1420	0.1431	0.1086	0.1431	0.1434

RRI	-0.2216	0.0161	0.8676	-0.4602	-0.3325	0.1287
P value	0.1392	0.1402	<.0001	0.1053	0.1388	0.1391
RRDI	0.3414	-0.0829	0.8235	0.2292	0.1470	0.2416
P value	0.1359	0.1368	<.0001	0.1038	0.1368	0.1372
RCXC	0.1099	-0.2758	0.6699	0.1612	-0.1872	-0.0906
P value	0.1426	0.1437	0.0001	0.1121	0.1478	0.1481
RDCXC	0.0756	0.3981	0.9834	-0.3277	-0.1664	0.3411
P value	0.1498	0.1509	<.0001	0.1234	0.1626	0.1630
RRA	-0.1081	-0.0042	0.7502	-0.1232	0.8056	0.1138
P value	0.1363	0.1373	0.0001	0.1012	0.1333	0.1336
RMU	0.3749	0.7975	0.0898	0.1483	-0.3271	0.3306
P value	0.1390	0.0001	0.0001	0.1001	0.1320	0.1323
ROA	0.3797	0.7435	0.3581	-0.0247	0.1285	0.3291
P value	0.1223	0.0001	0.0001	0.1102	0.1452	0.1455
ROE	0.1705	0.9232	0.3031	0.2959	0.0400	0.3350
P value	0.0856	<.0001	0.0001	0.1069	0.1408	0.1412
RAT	-0.1452	0.5923	0.1400	0.9371	-0.1661	-0.2555
P value	0.0848	0.0933	0.1026	<.0001	0.1549	0.1553
RCI	-0.2603	-0.1547	-0.0750	0.8889	-0.0148	0.0196
P value	0.0839	0.0923	0.1016	<.0001	0.1348	0.1351

Fuente: Elaboración propia con el paquete estadístico SAS

Teniendo la justificación de la utilización del ACP, por medio de la aplicación de la medida de KMO, junto con la prueba de esfericidad de Barlett y los altos valores de correlación entre las razones financieras y los 4 componentes principales obtenidos, se concluye que el componente principal 1 resume la variabilidad de las razones financieras del rubro de liquidez, el componente principal 2 resume la variabilidad de las razones financieras del rubro de rentabilidad, el componente principal 3 resume la variabilidad de las razones financieras del rubro de administración y componente principal 4 resume la variabilidad de las razones financieras del rubro de apalancamiento. Esta información se muestra en la tabla 9:

Tabla 9: Agrupación de razones financieras por los componentes principales

Componente principal	Grupo de razones financieras que agrupa
Componente principal 1	Liquidez
Componente principal 2	Rentabilidad
Componente principal 3	Actividad

Componente principal 4	Apalancamiento
------------------------	----------------

Fuente: Elaboración propia con el paquete estadístico SAS

La tabla 10 muestra valores obtenidos por cada una de las empresas en función componente principal 1. Debido a las altas correlaciones presentadas en las matrices de correlación previamente presentadas, es posible resumir la variabilidad de las razones del rubro de liquidez a través del componente principal 1, lo que permite interpretar que la empresa que presenta menores niveles de rendimiento es la codificada como WM, al presentar un valor de -3.22678, mientras que la empresa que presenta mayores niveles de rendimiento es la empresa codificada como DHX, al presentar un valor de 4.5958, es decir el mayor de las 93 empresas analizadas.

Tabla 10: Empresas ordenadas en función del componente principal 1

EMPRESA	CP-1	EMPRESA	CP-1	EMPRESA	CP-1
PLUG	-3.22678	GD	-0.99334	CEO	0.55636
FE	-3.02082	ECL	-0.97359	SNY	0.5677
OKE	-2.85123	RTX	-0.90361	NTR	0.85925
AMD	-2.85076	BSX	-0.87846	QGEN	0.89104
VOD	-2.65981	LHX	-0.79047	HEI	0.99927
VZ	-2.4905	IDXX	-0.7029	MDT	1.15445
KEP	-2.31766	HOLX	-0.65367	RMD	1.24405
ADP	-2.27059	EMR	-0.61713	QDEL	1.33224
UGI	-2.24579	ABB	-0.60612	PFE	1.4366
ES	-2.23809	CNI	-0.59834	SYK	1.56657
VLO	-2.0734	FCX	-0.58608	ASML	1.70727
RCI	-1.98561	TEL	-0.58313	KLAC	1.7356
OGE	-1.85255	ROP	-0.54463	QCOM	1.81021
TRP	-1.84252	WIT	-0.47179	TXN	2.04628
CVS	-1.8241	WAB	-0.44145	GLW	2.11723
CAT	-1.66561	ATVI	-0.43918	VRTX	2.18338
HPQ	-1.66474	WSO	-0.35654	EW	2.26914
GIS	-1.61348	HAL	-0.2806	BIIB	2.2973
EXC	-1.60541	GPC	-0.26452	REGN	2.31822
LNT	-1.56754	ABT	-0.24596	ZBH	2.38734
NEE	-1.53978	DHR	-0.23137	NVDA	2.38857
UHS	-1.43485	UMC	-0.21153	GILD	2.59897
OMC	-1.3962	MU	-0.18652	GRMN	2.66434

FMS	-1.34778	BDX	0.07104	JCI	2.95125
ON	-1.30662	EL	0.12269	AMGN	3.07746
QRTEB	-1.26328	ROK	0.20508	INCY	3.33127
CSX	-1.17225	ZBRA	0.28584	UTHR	3.44856
TOT	-1.14621	ILMN	0.28762	TSM	3.68434
GIB	-1.10905	BMV	0.44112	NVO	3.72764
ETN	-1.09235	BAX	0.46986	DHX	3.94108
NOXL	-1.01499	VAR	0.47921	XLNX	4.5958

Fuente: Elaboración propia con el paquete estadístico SAS

La tabla 11 muestra los valores obtenidos por cada una de las empresas en función del componente principal 2. Debido a las altas correlaciones presentadas en las matrices de correlación previamente presentadas, es posible resumir la variabilidad de las razones del rubro de rentabilidad a través del componente principal 2, lo que permite interpretar que la empresa que presenta menores niveles de este rentabilidad es la codificada como TSM, al presentar un valor de -4.46169, mientras que la empresa que presenta mayores niveles de rentabilidad es la empresa codificada como INCY, al presentar un valor de 6.99111, es decir el mayor de las 93 empresas analizadas.

Tabla 11: Empresas ordenadas en función del componente principal 2

EMPRESA	CP-2	EMPRESA	CP-2	EMPRESA	CP-2
TSM	-4.46169	BSX	-0.41468	QGEN	0.05417
CEO	-1.38172	CSX	-0.396	MDT	0.07837
BIIB	-1.13441	SYK	-0.37722	FE	0.09545
GLW	-1.1119	ES	-0.37019	EXC	0.1
CNI	-1.05005	GPC	-0.36743	VAR	0.10309
UMC	-0.93446	RMD	-0.35109	WAB	0.11192
PFE	-0.8615	TRP	-0.30366	BMV	0.12007
SNY	-0.85763	WSO	-0.30291	RTX	0.16197
UTHR	-0.85681	ADP	-0.29831	UHS	0.17458
CVS	-0.79897	WIT	-0.29699	EMR	0.20605
QCOM	-0.72769	NEE	-0.2947	ECL	0.23826
ZBH	-0.6732	ROP	-0.28825	FMS	0.25751
VLO	-0.66413	ATVI	-0.27528	GILD	0.28994
NVDA	-0.6555	QDEL	-0.27269	OKE	0.32603
GRMN	-0.6479	TEL	-0.2636	LHX	0.33007

OGE	-0.56722	VOD	-0.22247	KLAC	0.50518
QRTEB	-0.5449	XLNX	-0.22058	DHX	0.52969
DHR	-0.54016	ABT	-0.20548	ABB	0.5474
EW	-0.52255	NVO	-0.19728	HAL	0.57682
NTR	-0.50989	AMGN	-0.19647	GD	0.58507
TOT	-0.50919	ETN	-0.14	WM	0.61403
ASML	-0.50865	REGN	-0.13507	ZBRA	0.7224
KEP	-0.50458	BDX	-0.10779	VZ	0.73817
GIB	-0.50146	NOXL	-0.10412	RCI	0.74569
HEI	-0.47757	ILMN	-0.06738	ROK	0.97516
FCX	-0.47107	HOLX	-0.05317	HPQ	0.98148
LNT	-0.46666	GIS	-0.03085	CAT	2.0272
TXN	-0.43824	UGI	-0.02345	ON	2.18617
VRTX	-0.43645	AMD	-0.01849	OMC	2.90911
MU	-0.4361	EL	0.01934	JCI	6.93685
IDXX	-0.41572	BAX	0.02524	INCY	6.99111

Fuente: Elaboración propia con el paquete estadístico SAS

La tabla 12 muestra los valores obtenidos por cada una de las empresas en función componente principal 3. Debido a las altas correlaciones presentadas en las matrices de correlación previamente presentadas, es posible resumir la variabilidad de las razones del rubro de actividad a través del componente principal 3, lo que permite interpretar que la empresa que presenta menores niveles de actividad es la codificada como REGN, al presentar un valor de -2.48014, mientras que la empresa que presenta mayores niveles de actividad es la empresa codificada como TSM, al presentar un valor de 6.06692, es decir el mayor de las 93 empresas analizadas.

Tabla 12: Empresas ordenadas en función del componente principal 3

EMPRESA	CP-3	EMPRESA	CP-3	EMPRESA	CP-3
REGN	-2.48014	HEI	-0.47223	OGE	0.3927
ZBH	-2.33372	WAB	-0.44101	TOT	0.41153
SNY	-1.76091	BDX	-0.40976	RCI	0.46995
PFE	-1.69868	EW	-0.40391	XLNX	0.48124
QGEN	-1.58145	ZBRA	-0.38947	NVDA	0.51085
BSX	-1.5509	HAL	-0.38918	VRTX	0.51257
OMC	-1.49428	FE	-0.37901	QCOM	0.55218
CAT	-1.42263	NEE	-0.37625	GPC	0.60522

ILMN	-1.38597	FMS	-0.36572	UGI	0.62013
HOLX	-1.19898	ROP	-0.35554	ATVI	0.62196
MDT	-1.18606	NOXL	-0.34905	VOD	0.7383
AMGN	-1.14072	BAX	-0.3459	GIS	0.75297
ASML	-1.06355	EXC	-0.33426	CSX	0.75552
MU	-0.98504	TRP	-0.32291	TSM	0.76651
ABT	-0.96895	LNT	-0.30945	CEO	0.7867
QDEL	-0.91947	GD	-0.30308	ROK	0.8507
RMD	-0.9003	RTX	-0.26033	ON	0.88274
ABB	-0.89113	GLW	-0.20549	IDXX	0.9754
SYK	-0.78563	ETN	-0.20464	OKE	1.06338
TEL	-0.73307	GRMN	-0.18256	CVS	1.07088
KLAC	-0.7241	ES	-0.10995	CNI	1.10561
UMC	-0.70087	ECL	-0.10069	TXN	1.14618
FCX	-0.68163	WSO	0.02347	ADP	1.42959
BIIB	-0.67829	EMR	0.17616	VZ	1.44764
DHR	-0.67222	EL	0.20423	JCI	1.6817
UTHR	-0.64144	NTR	0.22635	INCY	1.84362
BMJ	-0.5992	LHX	0.24751	VLO	2.07446
VAR	-0.58346	GILD	0.30546	WIT	2.39051
KEP	-0.53927	NVO	0.30985	WM	2.4815
QRTEB	-0.50005	GIB	0.36888	DHX	2.56652
AMD	-0.485	UHS	0.37989	HPQ	6.06692

Fuente: Elaboración propia con el paquete estadístico SAS

La tabla 13 muestra valores obtenidos por cada una de las empresas en función componente principal 1. Debido a las altas correlaciones presentadas en las matrices de correlación previamente presentadas, es posible resumir la variabilidad de las razones del rubro de apalancamiento a través del componente principal 4, lo que permite interpretar que la empresa que presenta menores niveles de apalancamiento es la codificada como WIT, al presentar un valor de -3.67974, mientras que la empresa que presenta mayores niveles de apalancamiento es la empresa codificada como ON, al presentar un valor de 4.42791, es decir el mayor de las 93 empresas analizadas.

Tabla 13: Empresas ordenadas en función del componente principal 4

EMPRESA	CP-4	EMPRESA	CP-4	EMPRESA	CP-4
---------	------	---------	------	---------	------

VAR	-3.67974	JCI	-0.32069	QRTEB	0.37581
VOD	-2.72011	EMR	-0.31672	ZBH	0.38517
WM	-1.7885	ILMN	-0.31339	ES	0.39487
NOXL	-1.44711	ETN	-0.28972	VZ	0.40079
GD	-1.17143	TOT	-0.27577	EXC	0.42554
VLO	-1.1559	GLW	-0.22976	OGE	0.47265
OMC	-1.12523	EW	-0.18016	CNI	0.4798
UMC	-1.09305	ROP	-0.16925	UGI	0.53561
WIT	-1.08575	CEO	-0.15667	OKE	0.54129
GRMN	-1.01202	XLNX	-0.14097	ON	0.55576
QDEL	-0.9829	KEP	-0.12667	CSX	0.55616
REGN	-0.90524	ECL	-0.11759	BIIB	0.56232
ABB	-0.89367	SYK	-0.10344	LNT	0.61256
HAL	-0.89304	SNY	-0.09141	EL	0.61929
NVDA	-0.89081	ABT	-0.08845	QCOM	0.66091
HPQ	-0.84434	RTX	-0.04367	NVO	0.68501
ATVI	-0.82426	MDT	-0.0397	ZBRA	0.72673
RMD	-0.64688	UHS	-0.02699	IDXX	0.73189
MU	-0.63759	BSX	0.01784	VRTX	0.74625
TEL	-0.63313	BMY	0.02668	TRP	0.88389
GPC	-0.62991	ASML	0.07094	KLAC	0.923
QGEN	-0.55882	NTR	0.09454	GILD	1.0038
GIB	-0.49194	FCX	0.12753	NEE	1.03222
WAB	-0.49038	HOLX	0.16323	FE	1.12373
FMS	-0.48317	LHX	0.20963	GIS	1.23333
DHR	-0.46446	ADP	0.24812	CAT	1.33906
CVS	-0.44081	UTHR	0.25239	TSM	1.47279
HEI	-0.42034	BAX	0.3042	AMGN	1.63411
WSO	-0.41104	TXN	0.31764	RCI	1.91337
INCY	-0.37419	ROK	0.37285	AMD	2.56554
DHX	-0.37374	BDX	0.37377	PFE	4.42791

Fuente: Elaboración propia con el paquete estadístico SAS

Una de las aplicaciones que tiene el ACP, es la de formar rankings para de esta forma organizar y delimitar las variables de los análisis que se elaboren. Utilizando los 4 grupos formados con base a los diferentes componentes principales, se procedió a segmentar las empresas, utilizando la medida de posición de cuartiles. La tabla 14 muestra los valores de los cuartiles obtenidos para cada uno de los grupos de empresas agrupadas en función de cada componente principal.

Tabla 14: Cuartiles por grupo de componente principal

Cuartil	CP-1	CP-2	CP-3	CP-4
Q1	-1.372	-0.503	-0.68	-0.491
Q2	-0.439	-0.222	-0.303	-0.044
Q3	1.384	0.168	0.613	0.538

Fuente: Elaboración propia con el paquete estadístico SAS

Una vez que se calcularon los cuartiles para cada grupo de empresas, se utilizaron dichos valores para poder clasificar los niveles de liquidez, rentabilidad, actividad y apalancamiento de las 93 empresas del estudio. Se asignaron 4 niveles diferentes en función de cada cuartil, mismo que se muestran en la tabla 15.

Tabla 15: Niveles de clasificación

Nivel	Cuartil
Bajo	Q1
Moderado	Q2
Alto	Q3
Muy alto	>Q3

Fuente: Elaboración propia con el paquete estadístico SAS

Una vez teniendo calculada la variable dependiente del presente estudio, que es el éxito en las empresas, se procedió a fusionar la información obtenida de la fase 1 del análisis con la fase 2, para de esta forma contrastar la relación entre las clasificaciones de empresas mediante la técnica de análisis de componentes principales y la clasificación de empresas de éxito generada a partir de su rendimiento superior a la media del mercado Nasdaq.

Se identificó que el 40% de las empresas de éxito presentaron un nivel muy alto de liquidez, el 28% de ellas presentó un nivel alto, mientras que el 24% presentó niveles moderados y finalmente el 8% presentó niveles bajos.

Con respecto a los niveles de rentabilidad, el 20% de las empresas del grupo de éxito presentó niveles muy altos de este rubro, mientras que otro 20% presentó niveles altos. Por otro lado, el 40% de este grupo de empresas de éxito presentó niveles moderados de rentabilidad y el 20% restante presentó niveles bajos.

En cuanto a los niveles de actividad el 28% de las empresas de éxito presentaron un nivel muy alto, correspondiendo un 20% en un nivel alto, mientras que el 28% fue moderado y 24% restante presentó niveles bajos.

Finalmente, para el rubro de apalancamiento, el 36% de las empresas de éxito presentó niveles muy altos en este rubro, mientras que solo el 8% presentó niveles altos. Por otro lado, el 56% restante presentó niveles moderados y bajos de apalancamiento, distribuidos de forma equitativa en este porcentaje.

La tabla 16 resume los niveles de las 4 de liquidez, rentabilidad, actividad y apalancamiento de las empresas de éxito distribuidas en cada agrupación.

Tabla 16: Empresas de éxito en clasificación de componentes principales

Liquidez			
Nivel	Cuartil	Número de empresas	Porcentaje
Bajo	Q1	2	8%
Moderado	Q2	6	24%
Alto	Q3	7	28%
Muy alto	>Q3	10	40%
Total		25	100%
Rentabilidad			
Nivel	Cuartil	Número de empresas	Porcentaje
Bajo	Q1	5	20%
Moderado	Q2	10	40%
Alto	Q3	5	20%
Muy alto	>Q3	5	20%
Total		25	100%
Actividad			
Nivel	Cuartil	Número de empresas	Porcentaje
Bajo	Q1	6	24%

Moderado	Q2	7	28%
Alto	Q3	5	20%
Muy alto	>Q3	7	28%
Total		25	100%
Apalancamiento			
Nivel	Cuartil	Número de empresas	Porcentaje
Bajo	Q1	7	28%
Moderado	Q2	7	28%
Alto	Q3	2	8%
Muy alto	>Q3	9	36%
Total		25	100%

Fuente: Elaboración propia con el paquete estadístico SAS

IV.4. Análisis discriminante

En esta parte del análisis se aplicó otra técnica del análisis multivariado, que es el análisis discriminante. En este punto fue necesario utilizar las 12 razones financieras (variables independientes) y las empresas del grupo de éxito (la variable dependiente éxito, fue calculada en la fase 2 del análisis.

Al igual que en el análisis de componentes principales, primero es necesario aplicar una serie de pruebas a los datos para poder determinar la viabilidad de la aplicación de esta técnica estadística.

La tabla 17 muestra los valores de las medias obtenidas para cada una de las 12 variables independientes del estudio, clasificadas en dos grupos, siendo el grupo 1, el que concentra las 25 empresas de éxito, mientras que el grupo 2 concentra las empresas que no éxito. Este conjunto de información, es con lo que se da inicio al análisis discriminante, ya que es necesario comenzar a identificar, cuáles son variables que tienen capacidad para clasificar a la variable dependiente en diferentes grupos. A simple vista pueden observarse ciertas diferencias en los promedios entre cada grupo, sin embargo es necesario que se presenta una significancia estadística para poder afirmarlo.

Tabla 17. Medias de valores por grupo

Grupo	Variable	Media
Grupo 1	RC	2.7312
	RA	1.6110
	RRI	6.0378
	RRDI	103.2591
	RCXC	32.0244
	RDCXC	73.8144
	RRA	0.6836
	RMU	15.4658
	ROA	11.5101
	ROE	21.1626
	RAT	0.3317
	RCI	4.4376
Grupo 2	RC	1.9518
	RA	1.3684
	RRI	9.2843
	RRDI	83.9472
	RCXC	7.3885
	RDCXC	66.4991
	RRA	0.6926
	RMU	11.5314
	ROA	6.9810
	ROE	18.2315
	RAT	0.3674
	RCI	4.8541

Fuente: Elaboración propia con el paquete estadístico SAS

La tabla 18 muestra la prueba de igualdad de medias entre los grupos. Esta prueba se utiliza para conocer si el promedio de cada una de las variables independientes del modelo, difiere de forma estadísticamente significativa entre los grupos del análisis, es decir entre el grupo de empresas exitosas y empresas no exitosas.

En la primera columna de la tabla se indica la razón financiera que a la que se le está aplicando la prueba de igualdad de medias, la segunda columna contiene el estadístico de Lambda de Wilks.

El estadístico Lambda de Wilks expresa la proporción de variabilidad total que no es ocasionada por las diferencias entre los grupos, con esto se permite contrastar la hipótesis nula de que las medias de los grupos son iguales, es decir, es un estadístico que mide el poder discriminante de un conjunto de variables (Mendoza Méndez & Dorantes Coronado, 2016). El valor Lambda de Wilks es una de las medidas de mayor importancia de las funciones discriminantes, ya que también es útil para demostrar la significancia estadística de la función discriminante (Escobar Jaramillo, 2008).

La tercera columna es el estadístico de contraste con la que se realiza la siguiente prueba de hipótesis:

H0: Las medias de los grupos son iguales

H1: Las medias de los grupos son diferente

Finalmente, la cuarta columna, contiene los P-Value para cada una de las razones financieras analizadas. Por lo tanto, en las celdas que contienen un valor inferior a 0.05 ($P\text{-Value} < 0.05$), la hipótesis nula se rechaza, obteniendo como resultado que únicamente en las medias de las razones financieras RC, RA, RMU y ROA, existe una diferencia estadísticamente significativa entre los grupos de las empresas exitosas y las no exitosas, mientras que, para el resto de las razones financieras, la hipótesis de igualdad de medias, no es rechazada.

Este análisis es el primer paso para detectar la viabilidad de las variables discriminantes en el modelo, por lo que hasta este punto.

Tabla 18. Prueba de igualdad de medias de grupos

Variable	Lambda de Wilks	F	P-Value
RC	0.818	8.108	0.005
RA	0.885	1.43	0.011
RRI	0.986	1.268	0.263
RRDI	0.983	1.529	0.219
RCXC	0.971	2.74	0.101
RDCXC	0.997	0.229	0.634

RRA	0.986	0.011	0.917
RMU	0.899	4.906	0.029
ROA	0.886	11.745	0.001
ROE	0.992	0.752	0.388
RAT	0.994	0.567	0.454
RCI	0.966	3.237	0.075

Fuente: Elaboración propia con el paquete estadístico SAS

La tabla 19 muestra los resultados de M de Box, la cual se utiliza para contrastar la hipótesis nula de igualdad de las matrices de covarianzas son iguales. Uno de los supuestos del análisis discriminante es que todos los grupos proceden de la misma población y que, por lo tanto, las matrices de covarianzas correspondientes a cada grupo son iguales entre sí.

Por lo tanto, la prueba de hipótesis queda de la siguiente forma:

H0: Las matrices de covarianzas entre cada grupo son iguales.

H1: Las matrices de covarianzas entre cada grupo son diferentes.

El resultado de la prueba indica que el supuesto de igualdad de covarianzas entre los grupos se cumple, por lo que el valor de significancia, es superior a 0.05 (P-Value >0.05), por lo que con esta prueba se indica que es factible aplicar un análisis discriminante a las variables de estudio.

Tabla 19. Prueba M de Box

M de Box	685.508
F	7.03
gl1	78
P-Value	0.047

Fuente: Elaboración propia con el paquete estadístico SAS

En la tabla 20 se muestra el auto valor para la función discriminante utilizada en el modelo. El autovalor obtenido para la función discriminante obtiene un valor

relativamente bajo. Un autovalor de una función se interpreta como la parte de variabilidad total de la nube de puntos proyectada sobre el conjunto de todas las funciones atribuible a la función. Si su valor es grande, la función tiene una amplia capacidad de discriminación. En la siguiente columna se indica el valor de la correlación canónica obtenida.

La correlación canónica de esta tabla, se refiere a la medición de la asociación entre la variable dependiente y el grupo de variables independientes del modelo. Si su valor es grande (próximo a 1), es otro indicador de que el modelo es viable aplicar y que por lo tanto la función tiene una amplia capacidad de discriminación. En este caso la correlación canónica obtenida es de un valor de 0.487, porque la capacidad de las variables no es muy alta.

Tabla 20. Autovalores

Función	Autovalor	% de varianza	% acumulado	Correlación canónica
1	0.31	100	100	0.487

Fuente: Elaboración propia con el paquete estadístico SAS

La tabla 21 muestra los coeficientes de la función discriminante obtenida, utilizando todas las variables del estudio, puede observarse que existen varios coeficientes con valores muy cercanos a 0, siendo un indicador de que dichas variables tienen una muy baja capacidad de discriminación. Esta parte tiene relación con la prueba de igualdad de medias, en la que solo 4 variables independientes presentaron diferencias significativas, por lo que hasta este punto es posible identificar que la función discriminante obtenida no es la más óptima, por lo que será necesario realizar algunos ajustes para poder construir un mejor modelo. En este punto la función discriminante quedaría expresada de la siguiente forma:

Tabla 21. Coeficientes de la función discriminante

Variable	Coeficiente
RC	1.185
RA	-1.467
RRI	-0.009

RRDI	0
RCXC	0.005
RDCXC	-0.007
RRA	-0.417
RMU	-0.013
ROA	0.109
ROE	0.004
RAT	0.753
RCI	-0.348
(Constante)	0.874

Fuente: Elaboración propia con el paquete estadístico SAS

La tabla 22 muestra la matriz de estructuras, la cual muestra las correlaciones entre las variables discriminantes y la función discriminante calculada, en orden descendente. Al tratarse de correlaciones, significa que mientras más cercanos sean los valores cercanos a uno (en valor absoluto), más alta es la capacidad discriminante de la variable. Aquí es posible confirmar lo que otras pruebas han indicado, que no todas las variables son viables para incluir en el modelo, ya que es posible identificar que hay variables como RRA y RDCXC, que presentan niveles de correlación prácticamente nulos, mientras que variables como RRDI, ROE y RCXC, presentan correlaciones muy bajas. Por otro lado, las variables RC, RA, RMU y ROA, presentan los valores más altos, aunque indicando correlaciones moderadas con respecto a la función discriminante.

Tabla 22. Matriz de estructuras

RC	0.645
RA	0.536
RMU	0.417
ROA	0.339
RCXC	0.312
RRDI	0.233
RCI	0.225
RRI	-0.212
ROE	0.163
RAT	-0.142

RDCXC	0.09
RRA	-0.02

Fuente: Elaboración propia con el paquete estadístico SAS

La tabla 23 muestra el valor de los centroides para ambos grupos de estudio, puede observarse que, para el grupo de empresas de éxito, el centroide se encuentra en valores positivos, mientras que para las empresas de no éxito se encuentran en valores negativos, siendo esto un indicador de que existe una formación de 2 grupos distintos. Mientras más amplia sea esta diferencia, mayor será la confiabilidad de del uso de un análisis discriminante.

Tabla 23. Centroides de grupos analizados

Grupo	Valor del centroide
1	0.909
2	-0.334

Fuente: Elaboración propia con el paquete estadístico SAS

Finalmente, después de haber aplicado todas las pruebas de validación para la función discriminante, en la tabla 24 se calculan los resultados de clasificación del modelo. Utilizando a la función discriminante descrita anteriormente en la tabla 99, se obtiene de las 25 empresas clasificadas como exitosas, 14 de ellas fueron nuevamente clasificadas en dicho grupo, mientras que 11 de ellas fueron clasificadas en el grupo de empresas no exitosas, por lo que el modelo clasificó de forma correcta al 56% de las empresas exitosas.

Por otro lado, aplicado la función discriminante para el grupo de empresas no exitosas, se clasificaron nuevamente a 57 de las 68 empresas pertenecientes en este grupo, mientras que 11 fueron clasificadas en el grupo de empresas exitosas, por lo que el modelo clasificó de forma correcta al 83.8% de las empresas no exitosas.

De forma general la función discriminante clasificó de forma correcta al 76.3% de las empresas, tomando en cuenta su clasificación original utilizada para crear el modelo, porcentaje que podría considerarse aceptable según el nivel que se

considere tomar como óptimo para utilizar el modelo, sin embargo con base en algunas pruebas ya aplicadas durante este desarrollo (como la prueba de igualdad de medias) y otros resultados obtenidos (como la matriz de estructuras), es posible hacer una nueva corrida en el modelo de análisis discriminante, pero utilizando menos variables, para de esta forma tratar de construir una nueva función discriminante que sea más efectiva y por lo tanto su utilización permita generar resultados más confiables.

Tabla 24. Resultados de clasificación

Grupos	Grupo 1	Grupo 2	Total
Grupo 1	14	11	25
Grupo 2	11	57	68
Grupo 1	56%	44%	100%
Grupo 2	16%	84%	100%

Fuente: Elaboración propia con el paquete estadístico SAS

Se procedió a correr nuevamente el análisis discriminante, pero utilizando únicamente 4 variables independientes, las cuales fueron elegidas ya que presentaron significancia estadística en la prueba de igualdad de medias presentado en el análisis que contempló las 12 variables. La tabla 25 muestra que las 4 variables utilizadas, presentan diferencia significativa ($P\text{-Value} < 0.05$), es decir la hipótesis nula de igualdad de medias se rechaza, con lo que se concluye que las medias de las variables RC, RA, ROA y RMU son estadísticamente diferentes entre el grupo de las empresas exitosas y el grupo de empresas no exitosas, por lo que es el indicio de que es viable calcular una función discriminante.

Tabla 25. Prueba de igualdad de medias de grupos

Variable	Lambda de Wilks	F	P-Value
RC	0.721	8.108	0.000
RA	0.795	1.43	0.00
RRI	0.726	11.745	0.001

RRDI	0.733	4.906	0.029
------	-------	-------	-------

Fuente: Elaboración propia con el paquete estadístico SAS

La tabla 26 presenta los nuevos resultados para la prueba de hipótesis de igualdad de las matrices de covarianzas, que al igual que en el primer análisis aplicado, se corroboró la viabilidad de utilizar el análisis discriminante, ya que la hipótesis nula no se rechaza y tomando en cuenta únicamente las 4 variables independientes, la significancia estadística mejora, ya que el P-Value de la prueba es muy cercano a 0.

Tabla 26. Prueba M de Box

M de Box	77.832
F	7.266
gl1	10
P-Value	0.001

Fuente: Elaboración propia con el paquete estadístico SAS

En la tabla 27 se muestran los valores del autovalor de la función discriminante y la correlación canónica, en donde ambos valores mejoran de forma sustancial, ya que al existir un nivel de correlación de 0.921 indica que se tiene una relación muy fuerte entre la variable dependiente y el conjunto de variables independientes, por lo que en esto aporta que la función discriminante contará con una gran capacidad de clasificación.

Tabla 27. Autovalores

Función	Autovalor	% de varianza	% acumulado	Correlación canónica
1	0.899	100	100	0.921

Fuente: Elaboración propia con el paquete estadístico SAS

Al utilizar únicamente 4 variables para calcular la función discriminante, ésta se reduce en extensión, por lo que únicamente estará conformada por los 4

coeficientes que muestra la tabla 28. La función queda expresada de la siguiente forma:

Tabla 28. Coeficientes de la función discriminante

Variable	Coeficiente
RC	1.163
RA	1.467
ROA	0.841
RMU	0.948
(Constante)	0.948

Fuente: Elaboración propia con el paquete estadístico SAS

La tabla 29 muestra la matriz de estructuras, la cual indica los niveles de correlación que tiene cada variable discriminante con la función discriminante, se ordenan de forma descendente, por lo que esto también implica el orden de importancia que tiene cada variable en el modelo, en este caso la variable RC es la que tiene una mayor capacidad para discriminar, ya que presenta el valor más alto de correlación, seguido de la variable RA que presenta un valor similar. Seguido, las variables ROA y RMU presentan valores de correlación más pequeños, pero siguen siendo altos, ya que, si se compara con anterior matriz de estructuras, existían valores de correlación muy bajos (entre -0.3 y 0.3), incluso algunos cercanos a 0.

Con los resultados de esta matriz, se tiene evidencia de que la función discriminante será eficiente para la clasificación de la variable dependiente.

Tabla 29. Matriz de estructuras

RC	0.932
RA	0.914
ROA	0.867
RMU	0.848

Fuente: Elaboración propia con el paquete estadístico SAS

Con el nuevo análisis, puede apreciarse que la distancia entre los centroides de los grupos mejora, ya que el centroide para las empresas de éxito se ubica en un valor positivo de 0.987 mientras que, para el grupo de las empresas no exitosas, el centroide es de un valor de -0.454, con lo que se tiene evidencia de que existe diferencia entre ambos grupos.

Tabla 30. Centroides de grupos analizados

Grupo	Valor del centroide
1	0.987
2	-0.454

Fuente: Elaboración propia con el paquete estadístico SAS

En la tabla 31 se reflejan los beneficios de utilizar únicamente 4 variables en para la aplicación del análisis discriminante, ya que en los resultados de clasificación para el grupo de empresas exitosas, 23 de ellas fueron clasificadas de forma correcta utilizando la nueva función discriminante, es decir el 92% de estas empresas se clasifican en el grupo 1, mientras que para el grupo de las empresas no exitosas, 60 de ellas se clasifican de forma correcta, es decir el 88% de estas empresas se clasifican en el grupo 2.

De forma general, la nueva función discriminante pasó de clasificar de forma correcta al 76.3% de las empresas totales, a un 89.2%, por lo que con esto se evidencia de que es necesario someter a los datos a diferentes pruebas estadísticas para lograr la mejor función discriminante posible.

Tabla 31. Resultados de clasificación

Grupos	Grupo 1	Grupo 2	Total
Grupo 1	23	2	25
Grupo 2	8	60	68
Grupo 1	92%	8%	100%
Grupo 2	12%	88%	100%



IV.5. Análisis clúster

En esta parte del análisis, se corrió un análisis clúster jerárquico, para de esta forma revisar las diferencias de agrupamiento que se tuvieron en la fase 2 y 3 de la investigación.

Al igual que los demás análisis estadísticos, es necesario revisar varias pruebas para poder validar que el análisis sea viable de implementar con los datos del estudio que se trate. En el paquete estadístico SAS se realiza el análisis por medio del procedimiento PROC CLUSTER.

La primera información a analizar, es la matriz de auto valores de covarianza clúster que se muestra en la tabla 32. Este es el primer indicio de que es factible realizar agrupaciones con este análisis, ya que mientras más variabilidad concentren los autovalores, en menos cantidad de ellos, se interpreta que la agrupación es más factible de realizarse que si estuviera dispersa entre varios de ellos. En este caso el primer autovalor representa una variabilidad acumulada del 49.31% mientras que el segundo autovalor representa un 40.88% (esto se refleja en la columna número cuatro de la tabla), por lo que en sumatoria llegan a concentrar el 90.19% de la variabilidad total, por lo tanto, esto indica que la agrupación por el análisis clúster es capaz de generar distintas agrupaciones de los casos de estudio, es decir, para las 93 empresas de la presente investigación.

Tabla 32. Matriz de autovalores de covarianza clúster

Número	Autovalor	Diferencia	Proporción	Acumulada
1	4763.73074	814.75265	0.4931	0.4931
2	3948.97809	3391.36709	0.4088	0.9019
3	557.611	357.55851	0.0577	0.9597
4	200.05249	81.9002	0.0207	0.9804
5	118.1523	66.94551	0.0122	0.9926
6	51.20678	33.53506	0.0053	0.9979
7	17.67173	16.21573	0.0018	0.9997
8	1.456	0.69905	0.0002	0.9999
9	0.75694	0.61933	0.0001	1

10	0.13761	0.04386	0	1
11	0.09376	0.07814	0	1
12	0.01562		0	1

Fuente: Elaboración propia con el paquete estadístico SAS

La siguiente información que se obtiene es la matriz de proximidades, misma que muestra en la tabla 33. Esta matriz muestra las distancias cuadradas que se encuentran entre cada empresa, derivada de la varianza en cada una de las 12 variables independientes. Debido a que la tabla es demasiado extensa ya que compara las distancias entre cada una de las 93 empresas, se muestra únicamente una parte de la misma para que la interpretación de los datos pueda simplificarse. La tabla contiene una diagonal de ceros, debido a que, en cada una de esas celdas, se realiza la comparación de una compañía con ella misma, por lo que la distancia cuadrada que existe entre ellas es cero. Sin embargo, las demás celdas contienen una diversidad de valores, debido a que estas cantidades difieren según la distancia que existe entre cada empresa, por ejemplo, la distancia que existe entre la ABB y ABT, es por un valor de 13.405, a diferencia de ABB y ATVI, donde la distancia es de 65.581, interpretándose con esto, que existe una mayor similitud en cuanto a la variabilidad de las razones financieras de la empresa ABB y ABT, que con ABB y AVTI.

Esto indicaría que existe una mayor probabilidad de que ABB y ABT, sean agrupadas en un mismo clúster. Un ejemplo contrastante de este punto serían los valores de distancia cuadrática que se tienen entre la empresa ATVI y ASML, donde el valor obtenido es 399.156, indicando que existe una menor posibilidad de que estas empresas estén situadas en un mismo clúster. Esta misma comparación se va creando para cada una de las compañías de la investigación.

Tabla 33. Matriz de proximidades

Empresa	AMD	ABB	ABT	ADP	AMGN	ASML	ATVI
AMD	0.000	27.732	19.128	21.585	465.267	224.562	30.406
ABB	27.732	0.000	13.405	80.347	447.823	204.659	65.581
ABT	19.128	13.405	0.000	70.099	333.155	131.278	80.681
ADP	21.585	80.347	70.099	0.000	586.360	325.032	19.362
:AMGN	465.267	447.823	333.155	586.360	0.000	47.470	703.400
ASML	224.562	204.659	131.278	325.032	47.470	0.000	399.156
ATVI	30.406	65.581	80.681	19.362	703.400	399.156	0.000

Fuente: Elaboración propia con el paquete estadístico SAS

Siguiendo con el análisis de la información previa a la generación de grupos, la tabla 34 muestra una información fundamental, ya que contiene los estadísticos de agrupamiento aglomerativo para realizar la formación de clústeres (SAS, 2020). Debido a la extensión de la tabla que contiene la información de las 93 empresas, se muestra la última sección de ésta ya que es la información de mayor utilidad. La tabla muestra en la primera columna el número de clústeres a formar, que en este caso serían 92 posibles hasta llegar a únicamente 1. Basado en el criterio cúbico de agrupación (CCC), considerado como uno de los estadísticos más útiles para determinar el número óptimo de clústeres generados (Núñez & Escobedo, 2011) se elige la formación de 4 clústeres que integran las 21 empresas del análisis.

Tabla 34. Estadísticos de decisión

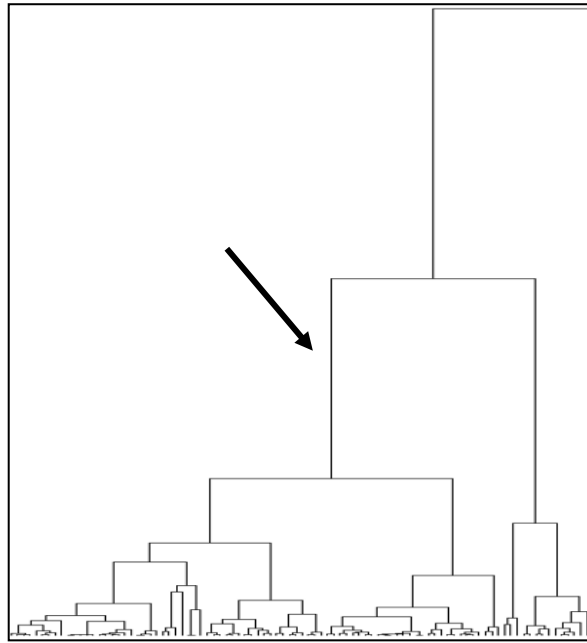
NCL	Conglomerados unidos		SPRSQ	RSQ	ERSQ	CCC	PSF	PST2
23	CL34	GILD	0.0012	0.981	0.971	16.8	154	4.9
22	CL70	CL37	0.0012	0.98	0.971	16.7	154	4.2
21	CL27	DHX	0.0013	0.978	0.954	16.7	153	10.4
20	CL25	CL33	0.0016	0.977	0.913	16.4	151	6.1
19	CL26	ZBH	0.0018	0.975	0.911	16.1	150	5.8
18	CL29	CL43	0.002	0.973	0.909	15.5	148	2.2
17	CL21	CL22	0.002	0.971	0.908	15.3	147	9.6
16	CL30	CL49	0.0022	0.969	0.903	15.2	147	6.6
15	CL23	REGN	0.0023	0.966	0.897	15.1	148	8.7
14	CL92	ON	0.0033	0.963	0.891	14.7	146	6.3
13	CL18	CL24	0.0033	0.96	0.884	14.5	146	930
12	CL17	CL41	0.0042	0.956	0.877	14.1	144	12.5

11	CL20	CL16	0.0057	0.95	0.868	13.5	140	14.4
10	CL28	CL15	0.0058	0.944	0.859	13.1	139	11.8
9	CL32	OMC	0.0061	0.938	0.848	12.9	140	7.9
8	CL9	CL14	0.0072	0.931	0.835	12.7	141	14.8
7	CL13	CL31	0.0101	0.921	0.82	12.2	141	4.4
6	CL12	CL8	0.0156	0.905	0.802	9.23	137	31
5	CL6	CL11	0.0212	0.884	0.778	8.43	133	17.1
4	CL19	CL10	0.0309	0.853	0.745	7.53	128	20.8
3	CL5	CL7	0.0317	0.821	0.697	7.84	136	23.6
2	CL3	CL4	0.0707	0.751	0.618	7.33	135	41.1
1	CL2	TSM	0.3196	0.431	0.382	1.79	68.9	115
	1	CL2	0.4309	0	0	0	0	68.9

Fuente: Elaboración propia con el paquete estadístico SAS

Al correr el análisis clúster en el paquete estadístico SAS, es posible generar un dendrograma, el cual muestra de manera visual la formación de grupos que se van generando. Estas formaciones es el resultado de los valores de la matriz de proximidad que van enlazando cada empresa con su nivel de proximidad. Visualmente en este dendrograma se muestran los 92 clúster posibles de formarse. La decisión de cuantos clústeres debe formarse, depende tanto de los estadísticos de decisión, como del interés que tenga el investigador o analista de la información. En la figura 2 se indica también el punto de unión entre dos grandes grupos de conglomerados, el de la izquierda se visualiza que contiene un número mayor de empresas agrupadas con respecto al grupo de la derecha. El dendrograma es una herramienta muy útil para de manera visual apreciar que tan dispersos están los casos de estudio, sin embargo, por eso es necesario correr otro procedimiento para conocer la forma en que quedará formado cada clúster.

Figura 2. Dendrograma



Fuente: Elaboración propia con el paquete estadístico SAS

La tabla 35 y muestran la conformación de los dos clústeres formados con base al CCC, indicando que las empresas que integran cada uno de dichos clústeres, presentan más semejanzas entre las 12 razones financieras analizadas, quedando conformado el clúster 1 por un total de 18 empresas y el clúster 2 por 18 empresas

Tabla 35. Primer clúster formado

Empresa	Empresa	Empresa
ABB	GD	QGEN
ABT	GIB	QRTEB
ADP	GILD	RCI
AMD	GIS	ROK
AMGN	GLW	RTX
ASML	GPC	SNY
ATVI	GRMN	SYK
BAX	HAL	TEL
BDX	HOLX	TRP
BIIB	HPQ	TSM
BMJ	INCY	TXN
BSX	JCI	UGI
CAT	KLAC	UHS
CEO	LHX	UMC



CNI	LNT	UTHR
CSX	MDT	VAR
CVS	MU	VLO
DHR	NOXL	VOD
ECL	NTR	VZ
EMR	NVDA	WIT
ES	OGE	WM
ETN	OMC	WSO
FCX	ON	XLNX
FE	PFE	ZBH
FMS	QCOM	ZBRA

Fuente: Elaboración propia con el paquete estadístico SAS

Tabla 36. Segundo clúster formado

Empresa	Empresa
DHX	NVO
EL	OKE
EW	QDEL
EXC	REGN
HEI	RMD
IDXX	ROP
ILMN	TOT
KEP	VRTX
NEE	WAB

Fuente: Elaboración propia con el paquete estadístico SAS

IV.6. Análisis K-medias

Para comprobar la formación de clústeres y ver las posibles diferencias entre los resultados, se optó por utilizar al análisis de K-medias, disponible en el paquete estadístico SAS. En este caso las observaciones se asignan a unos clústeres que el propio análisis configura sin que haya una relación entre unos grupos y otros. Se necesita que el experto, fije de antemano el número de clústeres en que desea agrupar los datos. Como esto no se sabe a priori con exactitud, lo que se suele hacer es repetir el análisis con diferentes números de clústeres, hasta conseguir la mejor interpretación o la clasificación que más se ajuste al objetivo del problema.

Una clasificación se considera correcta si la dispersión dentro de cada grupo formado, sea la mejor posible. Esta condición se denomina, criterio de varianza, la cual lleva a seleccionar una configuración cuando la suma de las varianzas dentro de cada grupo sea mínima.

Para la generación de K-medias, se inicia con una agrupación de observaciones en un número predefinido de conglomerados que el paquete SAS asigna, para luego evaluar cada observación, trasladándola hacia el conglomerado más cercano. Cuando un conglomerado cambia, debido a que se pierde o gana una observación, el programa calcula de nuevo el centroide del conglomerado (Minitab, 2020).

El valor mínimo de varianza determina una configuración de nuevos grupos con sus respectivas medias y se deben asignar otra vez todos los casos a estos nuevos centroides. Este proceso se repite hasta que ningún cambio pueda disminuir más la varianza intra-grupos, o se alcance otro criterio límite, que puede ser un número limitado de iteraciones.

Al correr el análisis no jerárquico de K-medias, la primera información que muestra, se muestra en la tabla 37, que corresponde a los valores de los centroides iniciales con los que comienza el algoritmo de K-medias, asignando de manera previa la generación de dos grupos, que fueron los que se formaron en el análisis clúster, previamente desarrollado en la presente investigación, para de esta forma ver las posibles diferencias en agrupación que genera cada método.

Tabla 37. Centros de clústeres iniciales

Variable	Clúster 1	Clúster 2
RC	4.9461	3.1297
RA	3.3139	2.4099
RRI	5.2557	7.8453
RRDI	69.4482	46.5249
RCXC	9.6874	626.1271
RDCXC	475.7950	18.2241
RRA	0.4283	0.5230

RMU	4.3400	34.0856
ROA	16.3400	17.8327
ROE	82.7234	22.8891
RAT	1.2478	0.0654
RCI	1.0923	4.1156

Fuente: Elaboración propia con el paquete estadístico SAS

Continuando con el análisis, se muestra la tabla 38, que muestra las iteraciones que realiza el programa estadístico, para calcular las distancias de los centroides entre cada grupo. Puede apreciarse que conforme se van llevando diferentes iteraciones, el calor de los centroides se va modificando, en donde al llegar a la quinta iteración se obtienen valores de cero. Esto último significa que los valores de los centroides obtenidos en la cuarta iteración, fueron los alcanzados como límite y que por lo tanto existieron cambios en cada la agrupación original del análisis jerárquico. Cabe mencionar que el programa estadístico SAS, tiene programado inicialmente realizar los cálculos con 10 iteraciones, por lo que el investigador puede modificar esta configuración a más iteraciones, con la finalidad tratar de encontrar una diferencia en la agrupación de variables, sin embargo, conforme más iteraciones se den, es probable que baje la significancia de las variables respecto a su aportación en la discriminación entre grupos.

Tabla 38. Historial de iteraciones

Iteración	Clúster 1	Clúster 2
1	412.4510	371.2059
2	329.9608	296.9647
3	300.2643	270.2379
4	240.2115	216.1903
5	0.0000	0.0000

Fuente: Elaboración propia con el paquete estadístico SAS

Después de haber corrido las iteraciones correspondientes para obtener los centroides por grupo, se muestran las distancias de centroides por variable de estudio con respecto a la de cada grupo en que son concentradas. Mientras más

pequeñas sean las distancias, puede inferirse que son las variables que más aportan para poder realizar la diferencia entre clústeres y que por lo tanto puedan generarse.

Tabla 39. Centros de clústeres asignados por variable

Variable	Clúster 1	Clúster 2
RC	11.1508	0.8452
RA	10.4230	0.7921
RRI	8.4177	7.8453
RRDI	49.6017	46.5249
RCXC	7.3576	6.1271
RDCXC	29.0117	18.2241
RRA	3.6920	8.5230
RMU	12.3554	0.9855
ROA	8.0938	0.8759
ROE	18.9773	7.8891
RAT	21.9852	21.8744
RCI	9.7490	9.1156

Fuente: Elaboración propia con el paquete estadístico SAS

Finalmente, previo a obtener los clústeres formados por el análisis de K-medias, la tabla 40 muestra el resumen de agrupación, indicando que en total el clúster 1 está conformado por 68 empresas, mientras que el clúster 2 está conformado por 25 empresas, que en sumatoria representan las 93 empresas de la investigación. La celda de perdidos, contiene un valor de 0, significando que no existieron empresas que hayan quedado fuera de alguna agrupación, cuestión que se pudo haber dado si la empresa presentara valores muy distintos a las del resto de las compañías analizadas.

Tabla 40. Resumen de agrupación

Clúster	1	68
	2	25
Válidos		93
Perdidos		0



En las tablas 41 y 42 se muestran los dos clústeres formados por el análisis no jerárquico de K-medias, teniendo como resultado final 68 empresas agrupadas en el clúster 1 mientras que en el clúster 2 se compone de 25 empresas en total.

Tabla 41. Primer clúster formado por K-medias

Empresa	Empresa	Empresa
ABB	GIS	RCI
ABT	GLW	ROK
ADP	GPC	RTX
AMD	GRMN	SNY
AMGN	HAL	SYK
ASML	HOLX	TEL
BAX	HPQ	TRP
BDX	JCI	TXN
BIIB	KLAC	UGI
BMJ	LHX	UHS
BSX	LNT	UMC
CAT	MDT	UTHR
CNI	MU	VAR
CSX	NOXL	VLO
CVS	NTR	VOD
DHR	NVDA	VZ
ECL	OGE	WIT
ES	OMC	WM
ETN	ON	WSO
FCX	PFE	XLNX
FE	QCOM	ZBH
FMS	QGEN	ZBRA
GD	QRTEB	

Fuente: Elaboración propia con el paquete estadístico SAS

Tabla 42. Segundo clúster formado por K-medias

Empresa	Empresa	Empresa
ATVI	HEI	REGN
CEO	IDXX	RMD

DHX	ILMN	ROP
EL	INCY	TOT
EMR	KEP	TSM
EW	NEE	VRTX
EXC	NVO	WAB
GIB	OKE	
GILD	QDEL	

Fuente: Elaboración propia con el paquete estadístico SAS

IV.7. Regresión lineal múltiple

Para esta parte del análisis se procedió a correr un análisis de regresión lineal múltiple, utilizando como variables independientes las 12 razones financieras, mientras que, para la variable dependiente, fue utilizada la tasa de rendimiento anual de cada empresa.

Previo a correr el análisis se procedió a verificar la matriz de correlaciones que se muestra en la tabla 43. La tabla muestra las asociaciones existentes entre cada una de las variables, junto con su valor de significancia. Puede observarse que existe una fuerte asociación positiva entre las variables independientes RCA, RA, RMU, ROA y ROE, con la variable dependiente RENDIMIENTO, todas estas presentando una significancia estadística ($P\text{-value} < 0.05$). También puede observarse una fuerte asociación entre la variable RC y RA, siendo esto un indicador de que existan problemas de multicolinealidad en el modelo a construir.

Tabla 43. Matriz de correlaciones de Pearson

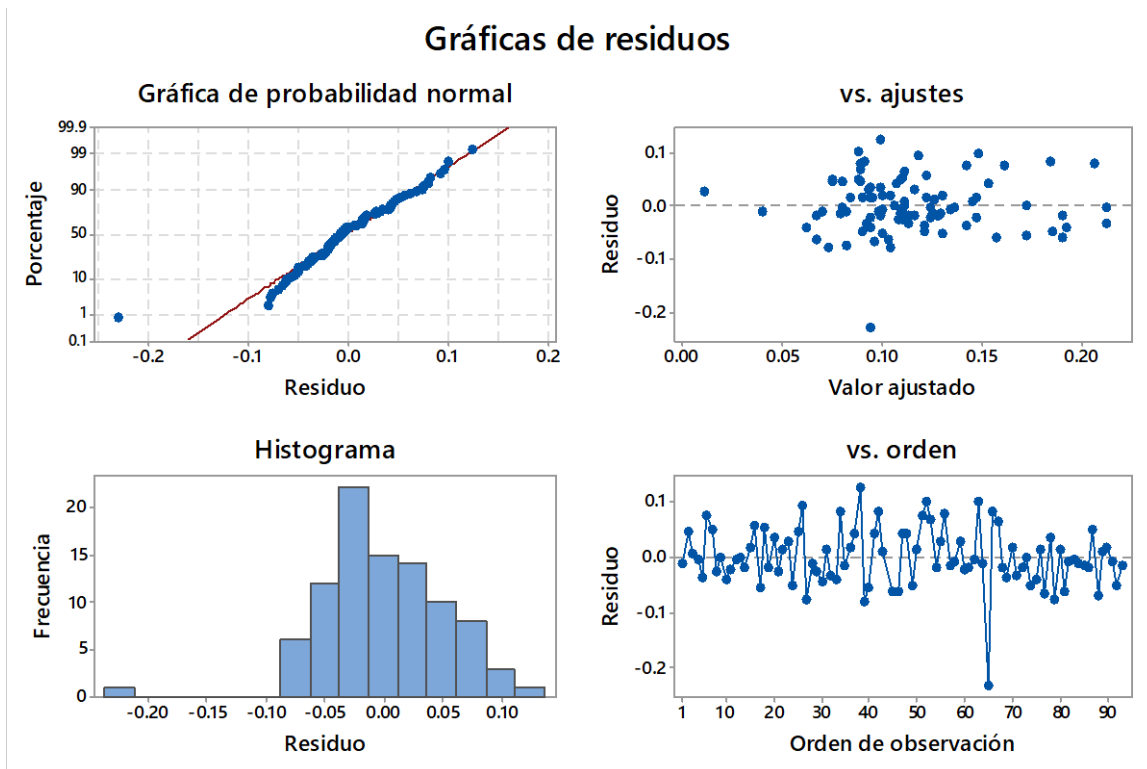
		REND	RC	RA	RRI	RRDI	RCXC	RDCXC	RRA	RMU	ROA	ROE	RAT	RCI
REND	CC	1	0.884	0.078	0.159	0.191	0.097	-0.087	0.084	0.789	0.801	0.712	0.135	0.026
	P-value		0.004	0.002	0.131	0.068	0.357	0.408	0.424	0.002	0.001	0.004	0.2	0.803
RC	CC	0.884	1	.850**	.278**	.416**	0.082	.347**	0.132	.383**	.488**	.238*	0.102	.408**
	P-value	0.004		0	0.007	0	0.438	0.001	0.209	0	0	0.022	0.335	0
RA	CC	0.078	.850**	1	-.233*	.373**	0.118	.249*	0.139	.398**	.471**	0.2	0.141	.359**
	P-value	0.002	0		0.026	0	0.261	0.017	0.187	0	0	0.056	0.179	0
RRI	CC	-0.159	-.278**	-.233*	1	.471**	-0.002	-0.05	0.006	-0.14	0.102	0.051	0.076	0.088

	P- value	0.131	0.007	0.026	-	0	0.984	0.638	0.954	0.183	0.335	0.627	0.469	0.407
RRDI	CC	0.191	.416**	.373**	.471**	1	-0.081	0.024	-0.12	.405**	.244*	0.101	.300**	0.129
	P- value	0.068	0	0	0		0.442	0.819	0.253	0	0.019	0.339	0.004	0.222
RCXC	CC	0.097	0.082	0.118	0.002	0.081	1	-0.089	-0.03	.285**	0.169	0.03	0.142	0.063
	P- value	0.357	0.438	0.261	0.984	0.442		0.399	0.776	0.006	0.108	0.774	0.178	0.55
RDCXC	CC	-0.087	.347**	.249*	-0.05	0.024	-0.089	1	0.157	0.162	0.145	.626**	.592**	.533**
	P- value	0.408	0.001	0.017	0.638	0.819	0.399		0.135	0.123	0.167	0	0	0
RRA	CC	0.084	-0.132	0.139	0.006	-0.12	-0.03	-0.157	1	.376**	0.102	0.074	0.133	0.037
	P- value	0.424	0.209	0.187	0.954	0.253	0.776	0.135		0	0.334	0.482	0.207	0.723
RMU	CC	0.789	.383**	.398**	-0.14	.405**	.285**	-0.162	.376**	1	.525**	0.103	.396**	0.149
	P- value	0.002	0	0	0.183	0	0.006	0.123	0		0	0.329	0	0.158
ROA	CC	0.801	.488**	.471**	0.102	.244*	0.169	0.145	0.102	.525**	1	.388**	0.132	.311**
	P- value	0.001	0	0	0.335	0.019	0.108	0.167	0.334	0		0	0.211	0.003
ROE	CC	0.712	.238*	0.2	0.051	0.101	0.03	.626**	0.074	0.103	.388**	1	.463**	-0.15
	P- value	0.004	0.022	0.056	0.627	0.339	0.774	0	0.482	0.329	0		0	0.154
RAT	CC	-0.135	-0.102	0.141	0.076	.300**	-0.142	.592**	0.133	.396**	0.132	.463**	1	0.067
	P- value	0.2	0.335	0.179	0.469	0.004	0.178	0	0.207	0	0.211	0		0.523
RCI	CC	-0.026	-.408**	.359**	0.088	0.129	-0.063	-.533**	0.037	0.149	.311**	-0.15	0.067	1
	P- value	0.803	0	0	0.407	0.222	0.55	0	0.723	0.158	0.003	0.154	0.523	

Fuente: Elaboración propia con el paquete estadístico SAS

Seguido a correr el modelo de regresión lineal se pocedió a verificar los supuestos del modelo de regresion lineal, mismos que se visualizan en la tabla 3. Puede observarse que de manera general los datos tienden a presentar una distribución normal ya que la mayoría d ela sobbservaciones se alinean de forma correcta en la gráfica de probabilidad normal. Por otro lado, el supeusto de homogenidad de varinzas se cumple al ver el gráfico de residuos, en donde no se refleja una tendencia de dichos residuos producto del moodelo de regresión generado. El hsitograma generado evidencia una tendencia a una distribución normal, ademas de que el gráfico de residuo cotra orden no refleja alguna estacionalidad en los datos. Todos estos son evidencia de la factibilidad de correr el analisis de regresion multiple.

Figura 3. Gráficas de residuos



Fuente: Elaboración propia con el paquete estadístico SAS

Para continuar con la verificación de variables a utilizar en el modelo del análisis de regresión múltiple, en la tabla 44 se muestran los resultados del ANOVA. Con estos resultados se corrobora que 5 de las 12 variables independientes presentan significancia estadística ($P\text{-values} < 0.05$), esto significa que las demás 8 variables que no presentaron significancia estadística, no estarían presentando una influencia en la variable dependiente, sin embargo, para conocer la eficiencia del modelo, es necesario conocer otros indicadores estadísticos, como lo es el coeficiente de determinación del modelo.

Tabla 44. ANOVA de regresión lineal

Fuente	GL	SC Ajust.	MC Ajust.	Valor F	Valor p
Regresión	12	0.123789	0.010316	3.3	0.001
RC	1	0.064063	0.064063	20.47	0.001
RA	1	0.051703	0.051703	16.52	0.002
RRI	1	0.000976	0.000976	0.31	0.578
RRDI	1	0.000025	0.000025	0.01	0.929
RCXC	1	0.000427	0.000427	0.14	0.713
RDCXC	1	0.003971	0.003971	1.27	0.263
RRA	1	0.002871	0.002871	0.92	0.341
RMU	1	0.004417	0.004417	15.41	0.002
ROA	1	0.000607	0.000607	14.33	0.002
ROE	1	0.003271	0.003271	10.98	0.004
RAT	1	0.0034	0.0034	1.09	0.3
RCI	1	0.001569	0.001569	0.5	0.481

Casi por finalizar la validación del análisis de regresión lineal múltiple, en la tabla 45 se muestran los coeficientes de regresión del modelo obtenido, donde únicamente los coeficientes de las variables RC, RA, RMU, ROA y ROE, presentan significancia estadística para entonces formar la ecuación final del modelo. Sin embargo, las variables RC y RA, presentan valore VIF superiores a 1, lo que es un indicio de que existen problemas de multicolinealidad, es decir que estas variables están correlacionadas entre sí, por lo que es necesario realizar ajustes al modelo.

Tabla 45. Significancia de coeficientes de regresión

Término	Coef	Valor p	VIF
Constante	0.0678	0.263	
RC	0.464	0.001	1.29
RA	0.535	0.002	1.75
RRI	-0.000306	0.578	0.34
RRDI	0.000011	0.929	0.86
RCXC	0.000036	0.713	0.18
RDCXC	-0.000251	0.263	0.52
RRA	0.0219	0.341	0.05
RMU	0.167	0.002	0.45
ROA	0.65	0.002	0.23

ROE	0.721	0.004	0.39
RAT	0.0528	0.3	0.43
RCI	-0.00664	0.481	0.22

Fuente: Elaboración propia con el paquete estadístico SAS

Finalmente, se obtiene el coeficiente de determinación y el coeficiente de determinación ajustado, éstos se reflejan en la tabla 46. El coeficiente de determinación, refleja el porcentaje de variabilidad de la variable dependiente que es explicado por el total de las variables independientes, mientras que el coeficiente de determinación ajustado, se interpreta de la misma forma, solo que toma en cuenta el ajuste realizado por el hecho de incluir más variables en el modelo. Por lo tanto, tomando el coeficiente de determinación ajustado, tomando en cuenta las 12 variables del estudio, éstas explican la variabilidad total en 48.55%. Debido a que en el análisis se detectaron variables no significativas, además de detectar problemas de multicolinealidad, se procede a correr nuevamente el análisis de regresión, aplicando ajustes en el número de variables a utilizar.

Tabla 46. Coeficiente de determinación

R-Cuadrada	R-Cuadrada (ajustado)
63.37%	60.25%

Fuente: Elaboración propia con el paquete estadístico SAS

Después de correr nuevamente el análisis de regresión múltiple, se obtiene la significancia estadística de las 4 variables utilizadas, donde todas presentan P-valores inferiores a 0.05, con lo que se determina que las 4 variables deben ser incluidas en el modelo. Cabe mencionar que las variables RC y RA, originaron problemas de multicolinealidad en el modelo, por lo que se incluyó únicamente la variable RC, ya que presentó mayor nivel de correlación con la variable dependiente.

Tabla 47. ANOVA de regresión lineal (2)

Fuente	GL	SC Ajust.	MC Ajust.	Valor F	Valor p
Regresión	4	0.058561	0.01464	4.08	0.004
RC	1	0.014365	0.014365	4	0.000
RMU	1	0.003009	0.003009	3.84	0.000
ROA	1	0.002154	0.002154	3.62	0.000
ROE	1	0.003476	0.003476	2.97	0.040

Fuente: Elaboración propia con el paquete estadístico SAS

Con el nuevo análisis elaborado, se obtienen nuevos coeficientes de regresión, mismos que quedan validados al presentar significancia estadística ($P\text{-Value} < 0.05$). Por lo tanto, la ecuación final del modelo de regresión lineal múltiple, queda conformada de la siguiente manera:

$$REND = 1.20 + .1226 RC + 0.902 RMU + 0.409 ROA + 0.522 ROE + e$$

Tabla 48. Significancia de coeficientes de regresión (2)

Término	Coef	Valor p	VIF
Constante	1.2	0.000	
RC	0.1226	0.000	1.34
RMU	0.902	0.001	1.48
ROA	0.409	0.020	1.8
ROE	0.522	0.041	1.18

Fuente: Elaboración propia con el paquete estadístico SAS

Como resultado final del análisis de regresión múltiple utilizando las 4 variables anteriormente mencionadas, el coeficiente de determinación y el coeficiente de determinación ajustado, mejoran sustancialmente, ya que cualquiera de los dos (al ser muy similares) casi alcanza a explicar el 80% de la variabilidad total de la variable dependiente, ya que el R-Cuadrada que se muestra en la tabla 49 tiene un valor de 79.87%, mientras que el R-Cuadrada ajustado, tiene un valor del 78.82%.

Tabla 49. Coeficiente de determinación (2)

R-Cuadrada	R-Cuadrada (ajustado)
79.87%	78.82%

Fuente: Elaboración propia con el paquete estadístico SAS

IV.8. Regresión logística binaria

Para la etapa final de todas las series de pruebas estadísticas aplicadas en la presente investigación, se optó por aplicar una regresión logística binaria, con la intención de determinar la probabilidad de éxito de las empresas de estudio. Para este análisis se utilizaron las clasificaciones obtenidas por medio de la aplicación de componentes principales realizadas en la presente investigación, donde se obtuvieron 4 niveles para cada uno de los rubros de las razones financieras analizadas.

La tabla 50 muestra el resumen de las variables para introducirlas al paquete estadístico SAS. La primera columna es la codificación utilizada para introducir los datos al paquete estadístico, la segunda columna corresponde al rubro a analizar (corresponde a los 4 rubros a los que pertenecen las razones financieras de la investigación y el rubro de éxito obtenido según los cálculos previamente elaborados en esta investigación), la tercera columna indica el tipo de variables que es, mientras que la última columna corresponde a los 4 niveles posibles que puede tomar la variable.

Tabla 50. Variables para regresión logística

Codificación	Rubro analizado	Tipo de variable	Niveles
LQ	Liquidez	Independiente	Muy alto
			Alto
			Moderado
			Bajo
AC	Actividad	Independiente	Muy alto
			Alto
			Moderado

			Bajo
RE	Rendimiento	Independiente	Muy alto
			Alto
			Moderado
			Bajo
AP	Apalancamiento	Independiente	Muy alto
			Alto
			Moderado
			Bajo
EX	Éxito	Dependiente	Sí
			No

Fuente: Elaboración propia con el paquete estadístico SAS

Al igual que otros análisis elaborados en esta investigación es necesario corroborar las variables que se utilizan para cada uno, por lo tanto, la tabla 51 muestra el ANOVA para los 4 rubros analizados. Según la significancia del estadístico Chi-cuadrado, es la forma de determinar la significancia de cada variable, en donde se muestra que las 4 variables utilizadas, son viables de utilizar en el análisis ($P\text{-Value} < 0.05$), por lo tanto, es posible obtener la ecuación logística y por lo tanto determinar las probabilidades asociadas con cada variable dependiente y la variable dependiente.

Tabla 51. ANOVA regresión logística

Fuente de variación	Chi - cuadrada	P -Value
LQ	16.84	0.001
AC	8.47	0.043
RE	14.77	0.001
AP	7.45	0.039

Fuente: Elaboración propia con el paquete estadístico SAS

Con la validación de las variables el resultado de la ecuación de regresión logística queda conformada de la siguiente forma:

$$Y' = 4.854 + 1.658 LG + 1.546 AC + 1.224 RE + .985 AP + e$$

Por lo tanto, al sustituir los coeficientes en la ecuación creada, el programa estadístico SAS, calcula la combinación con la probabilidad más alta para la variable dependiente, que en este caso fue determinada como la probabilidad de éxito debido al rendimiento superior al de la bolsa del Nasdaq. La tabla 52 muestra que, teniendo un nivel de liquidez muy alto, un nivel de actividad alto, un nivel de rendimiento alto y un nivel bajo de apalancamiento, existe una probabilidad del 86.13% de que una empresa sea del grupo de empresas exitosas.

Tabla 52. Probabilidad de éxito

LQ	AC	RE	AP	Probabilidad
Muy alto	Alto	Alto	Bajo	86.13%

Fuente: Elaboración propia con el paquete estadístico SAS

Tabla 53. Clasificación de empresas con base en Regresión logística

Empresa	Empresa
WAB	REGN
CEO	IDXX
TOT	ILMN
EL	INCY
EMR	KEP
EW	VRTX
EXC	NVO
GIB	OKE
ROP	QDEL

Fuente: Elaboración propia con el paquete estadístico SAS

V. Conclusiones

La primera hipótesis específica de la investigación fue que los valores de las razones financieras de las empresas que cotizan en la bolsa de valores del Nasdaq, son muy similares entre sí, esta hipótesis se rechaza, ya que después de haber procesado los estados financieros de las 93 empresas de la investigación, se obtuvieron los estadísticos descriptivos que evidencian una variabilidad importante entre el conjunto de empresas. Para la razón corriente se obtuvo un valor de 2.16 con un coeficiente de variación de 56.19%, para la razón ácida se obtuvo un valor de 1.43 con un coeficiente de variación de 60.62%, para la razón de rotación de inventario el valor fue de 8.41 con un coeficiente de variación de 146.75%, para la razón de rotación de días en el inventario el valor obtenido fue de 89.14 días con un coeficiente de variación de 75.12%, para el valor de la razón de cuentas por cobrar se obtuvo un valor de 14.01 con un coeficiente de variación de 458.83%, para la razón de rotación en días de las cuentas por cobrar se obtuvo un valor de 68.47 días con un coeficiente de variación de 95.09%, para la razón de rotación de activo se obtuvo un valor de 10.69 con un coeficiente de variación de 53.02%, para la razón de margen de utilidad el valor obtenido fue de un 39.67% con un coeficiente de variación de 61.19%, para la razón de retorno sobre activos el valor fue de 69.19% con un coeficiente de variación de 72.83%, para la razón de rotación sobre patrimonio el valor fue de 19.02% con un coeficiente de variación de 75.87%, para la razón de apalancamiento total se obtuvo un valor de 32.35% con un coeficiente de variación de 56.54% y finalmente para la razón de cobertura de intereses el valor fue de 4.74 veces con un coeficiente de variación de 21.12%.

La segunda hipótesis específica fue que las empresas que tienen una tasa de rendimiento superior en al menos un 50% respecto a la tasa de rendimiento promedio de la bolsa del Nasdaq, son consideradas exitosas. Esta hipótesis se acepta ya que con base en ese criterio que ha sido utilizado en otras investigaciones, fue factible conocer la tasa de rendimiento anualizada para cada empresa, misma que se calculó utilizando la media geométrica mensual de crecimiento en el valor de las acciones. Con este dato fue posible formar dos grupos

empresas que fueron clasificados como empresas exitosas y empresas no exitosas, donde el primer grupo estuvo conformado por un total de 25 empresas, mientras que el segundo grupo estuvo conformado por un total de 68 empresas. El criterio de agrupación fue que, para ser considerada una empresa exitosa, el rendimiento anual de la empresa debió superar en al menos el 50% el rendimiento de la bolsa del Nasdaq. El rendimiento de la bolsa del Nasdaq fue medido a través del índice Nasdaq 100, que tuvo una tasa de crecimiento de 9.91%, con lo que se interpreta que el rendimiento promedio de las empresas que cotizan en ella, fue este valor.

Ya que se obtuvieron los valores de las razones financieras y las tasas de rendimiento de las 93 empresas, se procedió a correr diferentes técnicas estadísticas multivariadas, para de esta forma determinar cuáles fueron las principales razones financieras que tuvieron influencia en la tasa de rendimiento de las empresas exitosas, además de poder explorar las ventajas que tiene la incorporación de la estadística multivariada al análisis financiero.

El primer paso fue corroborar que el grupo de datos correspondiera a ser factible de ser analizado por una técnica de reducción de variables y, por ende, pudiera efectuarse una separación de grupos, por lo que se aplicó la prueba de Kaiser-Meyer-Olkin y la prueba de esfericidad de Barlett, mismos que confirmaron la viabilidad de la aplicación de los análisis estadísticos a aplicarse. Seguido a esto, se procedió a correr el análisis de componentes principales, teniendo como resultado la formación de 4 nuevas variables denominadas eigenvalores, mismas que en su conjunto fueron capaces de explicar 81.88% de la variabilidad total de los datos. Estos 4 eigenvalores presentaron altos niveles de correlación con un rubro diferente de las 12 razones financieras que fueron calculadas en la investigación. Ya que a cada eigenvalor está asociado con el número de componente principal que se utiliza para el cálculo, se concluyó que el componente principal 1 resume la información del rubro de razones de liquidez, el componente principal 2 resume la información del rubro de razones de rentabilidad, el componente principal 3 resume la información del rubro de razones de actividad y finalmente el componente

principal 4 resume la información del rubro de razones de apalancamiento. Seguido a esto se ordenó cada empresa en función del valor de cada componente principal y utilizando las medidas de posición de cuartiles se creó una escala de Likert de 4 niveles (bajo, moderado, alto y muy alto) para identificar el nivel de liquidez, rentabilidad, actividad y apalancamiento de cada empresa.

En función de estas escalas citadas, se obtuvo que la empresa Xilinx, Inc, fue la que obtuvo el nivel más alto de liquidez. Esta compañía de tecnología estadounidense, es una de las principales distribuidoras de dispositivos lógicos programables. Es reconocida por inventar los FPGA y también por ser la primera compañía con modelos de manufactura fabless. En este mismo rubro, se encontró que la empresa que obtuvo el menor nivel de liquidez fue la empresa Plug Power Inc., compañía estadounidense que se dedica al diseño y fabricación de sistemas de pilas de combustible de hidrógeno que reemplazan las baterías convencionales en equipos y vehículos que funcionan con electricidad.

En el rubro de rentabilidad la empresa que obtuvo el mayor nivel fue la empresa farmacéutica Incyte Corporation, que se dedica a la investigación, fabricación y venta de medicamentos para el área oncológica principalmente y la autoinmunidad. Por otro lado, la empresa que presentó el menor nivel de rentabilidad fue la compañía Taiwan Semiconductor Manufacturing Company, que es una de las empresas de semiconductores para diversos equipos electrodomésticos.

En la clasificación que realizó en el rubro de actividad, la empresa que ocupó el primer lugar fue HP Inc, Empresa estadounidense con sede en Palo Alto, California, surgida de la separación de Hewlett-Packard, líder en venta de impresoras en el mundo y la segunda a nivel mundial en venta de computadores personales y portátiles. En última instancia de este rubro, el menor nivel de actividad lo presentó la empresa de biotecnología Regeneron Pharmaceuticals, compañía estadounidense de biotecnología con sede en Tarrytown, Nueva York.

Finalmente, en la el rubro de apalancamiento, la empresa que presentó mayores niveles fue la farmacéutica Pfizer, compañía diversificada en investigación y desarrollo farmacéutico en muchas áreas de la salud, desde enfermedades básicas, hasta la incursión en enfermedades de nuevo origen como el Covid-19. La empresa que obtuvo el menor nivel de esta categoría fue Varian Medical Systems, fabricante estadounidense de software y tratamientos de oncología radioterápica con sede en Palo Alto, California. Sus dispositivos médicos incluyen aceleradores lineales y software para tratar el cáncer y otras condiciones médicas con radioterapia, radiocirugía, terapia de protones y braquiterapia.

En otra parte de los análisis de la investigación, se utilizó la regresión logística binaria para determinar las probabilidades de éxito de una empresa, en función de los niveles de liquidez, rentabilidad, actividad y apalancamiento que presentaron. En este análisis se utilizó la clasificación que se obtuvo con el análisis de componentes principales, es decir se utilizó como variable independiente el nivel de Bajo, Moderado, Alto y Muy alto que obtuvo cada empresa para entonces determinar la probabilidad de que esto se diera. El análisis mostró que los 4 rubros presentaron significancia estadística, es decir los 4 rubros en su conjunto determinan la probabilidad de que la empresa sea exitosa. De todas las combinaciones posibles, el valor más alto fue de un 86.13%. Con base a este análisis, se concluyó que, si la empresa presenta altos niveles de liquidez, altos niveles de actividad, altos niveles de rendimiento y bajo nivel de apalancamiento, la empresa tiene una probabilidad del 86.13% de ser exitosa, es decir, superar el rendimiento promedio de las empresas que cotizan en la misma bolsa de valores que es el Nasdaq.

Otro de las técnicas estadísticas implementados fue el análisis clúster, donde primeramente fue necesario corroborar la viabilidad de efectuar este tipo de análisis para lo cual se procedió a calcular la matriz de autovarianza clúster donde al presentar valores altos en los primeros autovalores, es la primera forma de validar esta prueba. Seguido a esto se calculó la matriz de proximidades, misma que

muestra las diferentes distancias de los centroides de cada empresa analizada, al poder apreciarse distancias alejadas de forma significativa, se validó nuevamente la aplicación de este análisis. Para la formación de clústeres, el criterio cúbico cuadrado, con lo que se formaron dos grupos distintos. En el dendrograma se apreció una buena separación entre los dos grupos formados, por lo que el primer clúster quedó conformado por un total de 75 empresas, mismo que concentró al 40% de las empresas exitosas. En este sentido, el segundo clúster formado quedó integrado por un total de 18 empresas, mismo que concentró al 60% de las empresas exitosas.

Dando continuidad a este análisis se corrió la técnica estadística de K-medias, la cual una de las diferencias que presenta con respecto al análisis clúster para la formación de grupos, es que en esta es necesario tener a priori un número de grupos a crear, por lo tanto, debido a que previamente se realizó el análisis clúster y bajo un criterio estadístico se optó por la formación de dos grupos, en este análisis se toma como referencia dicho valor, por lo que al generar también dos grupos. Ya que el algoritmo de K-medias se basa en número de iteraciones para realizar los cambios en el cálculo de los centroides, en el proceso se llegó a 4 iteraciones, dando como resultado la formación de 2 grupos muy parecidos a los generados con el análisis clúster. El primer grupo quedó conformado por un total de 68 empresas, mientras que el segundo grupo quedó conformado por un total de 25 empresas. Pareciera ser que esto último no fue un cambio significativo, sin embargo, al revisar las empresas que conformó cada uno de los grupos, se identificó que el primero concentró el 20% de las empresas de éxito, mientras que el segundo concentró el 80% de las empresas de éxito.

Teniendo en cuenta estos cambios, se concluye que es necesario que una vez que se aplica un análisis de grupos jerárquico, es necesario también aplicar un análisis de grupos no jerárquico, para de esta forma evaluar los posibles cambios que pueden generarse en las agrupaciones y el investigador puede decidir cuál agrupación se adapta a sus objetivos de investigación.

En la fase donde se implementó el análisis discriminante, se comenzó con hacer probar la igualdad de las medias en los dos grupos generados por esta técnica estadística, en la que en esta primera instancia, por medio del estadístico de Lambda de Wilks, que prueba la hipótesis nula de igualdad de medias entre grupos, se obtuvo que únicamente la razón circulante, la razón ácida, la razón de margen de utilidad y la razón de rendimiento sobre activos, presentó una diferencia estadística, por lo que se interpreta que únicamente el promedio de estas variables son diferentes entre el grupo 1 y el grupo 2. Seguido a esto se corrió la prueba M de box con la finalidad de contrastar la hipótesis nula de igualdad de matrices de covarianza, misma que no fue rechazada, por lo que con esta prueba se concluye que el análisis discriminante es factible de realizarse. El siguiente paso fue obtener las correlaciones canónicas entre la variable dependiente y el conjunto de variables independientes del modelo generado por el análisis discriminante, donde se obtuvo un valor de 0.487, siendo éste un valor bajo de asociación. En este punto fue posible entonces calcular la función discriminante, donde al obtener la matriz de estructuras se obtienen las correlaciones entre las variables discriminantes y la función discriminante generada, con lo que se obtiene que las 5 variables más importantes para realizar la discriminación de grupos fueron; la razón corriente, la razón ácida, la razón de margen de utilidad, la razón de retorno sobre activos y la razón de rendimiento sobre patrimonio. Por último, para conocer los resultados de clasificación utilizando la función discriminante generada, se obtuvo que las 25 empresas clasificadas como exitosas, 14 de ellas fueron nuevamente clasificadas en dicho grupo, mientras que 11 de ellas fueron clasificadas en el grupo de empresas no exitosas, por lo que el modelo clasificó de forma correcta al 56% de las empresas exitosas. Por otro lado, aplicado la función discriminante para el grupo de empresas no exitosas, se clasificaron nuevamente a 57 de las 68 empresas pertenecientes en este grupo, mientras que 11 fueron clasificadas en el grupo de empresas exitosas.

Debido a que, en esta primera aplicación del análisis discriminante, no todas fueron significativas con el estadístico de Lambda de Wilks, la correlación canónica no fue alta y el resultado de clasificación generado por función discriminante únicamente

tuvo efectividad de un 56% respecto al grupo de empresas exitosas, se hizo una nueva corrida del análisis utilizando únicamente las variables que presentaron significancia con el estadístico de Lambda de Wilks.

Con la nueva corrida del análisis, en la prueba M de box no se rechazó la hipótesis nula, el valor de la correlación canónica incrementó de forma sustancial a un 92.1%, la distancia entre los centroides de los grupos aumentó y los valores de correlación presentados en la matriz de estructuras también aumentaron. Estos valores mejorados finalmente reflejan su efectividad en los resultados de clasificación obtenidos con la función discriminante que quedó conformada de la siguiente manera:

Con dicha ecuación se obtuvo que, para el grupo de empresas exitosas, 23 de ellas fueron clasificadas de forma correcta utilizando la nueva función discriminante, es decir el 92% de estas empresas se clasifican en el grupo 1, mientras que, para el grupo de las empresas no exitosas, 60 de ellas se clasifican de forma correcta, es decir el 88% de estas empresas se clasifican en el grupo 2.

De forma general, la nueva función discriminante pasó de clasificar de forma correcta al 76.3% de las empresas totales, a un 89.2%, por lo que se concluye con base en este análisis, que a razón corriente, la razón ácida, la razón de margen de utilidad y la razón de rendimiento sobre activos, son las que tienen influencia y por lo tanto determinan el éxito de una empresa que cotiza en la bolsa de valores del Nasdaq.

El último análisis implementado fue el de regresión lineal múltiple, donde en primera instancia se procedió a correr una matriz de correlaciones con las 12 variables independientes que fueron las razones financieras y la variable dependiente fue la tasa de rendimiento que obtuvo cada empresa. En primera instancia se calculó la matriz de correlaciones para todas las variables del análisis y posteriormente a calcular a gráfica de probabilidad normal, el histograma y el gráfico de valor residuos vs ajustes, y residuos vs orden, con lo que se corroboró que los datos eran factibles de analizarse por esta técnica estadística. El ANOVA de regresión lineal mostró que la razón corriente, la razón ácida, la razón de margen de utilidad, la razón de

rendimiento sobre activos y la razón de rendimiento sobre activos, presentaron significancia estadística para entonces formar la ecuación final del modelo. Sin embargo, las variables razón corriente y la razón ácida, presentan valores VIF superiores a 1, con lo que se detectó problemas de multicolinealidad. Con este primer análisis, se obtuvo que, basado en el coeficiente de determinación, las variables del modelo explicaron en un 60.25% la variabilidad de la dependiente, que fue la tasa de rendimiento, por lo tanto, se decidió correr nuevamente el modelo utilizando únicamente las variables significativas.

Para la nueva corrida del análisis se utilizaron se decidió no utilizar únicamente la razón corriente, la razón de margen de utilidad, la razón de rendimiento sobre activos y la razón de rendimiento sobre patrimonio. La razón ácida fue eliminada del análisis ya que presentó altos niveles de correlación con la razón corriente por lo que se originó un problema de multicolinealidad, pero la razón corriente la que quedó incorporada en el análisis ya que presentó mayores niveles de correlación con la variable dependiente. Las demás variables no fueron incorporadas en este segundo análisis debido a la falta de significancia estadística mostrada en el ANOVA. Con estos ajustes, el modelo de regresión lineal múltiple, explicó en un 78.82% el comportamiento de la tasa de rendimiento a través del valor de la razón corriente, la razón de margen de utilidad, la razón de rendimiento sobre activos y la razón de rendimiento sobre patrimonio.

Con base a los 6 análisis estadísticos multivariados, se acepta la hipótesis específica tres, que consistió en que las principales razones financieras que determinan el éxito de las empresas, pertenecen al rubro liquidez, rendimiento, actividad y apalancamiento, ya que la combinación de los niveles de cada rubro, influye en el éxito de la empresa, siendo las de mayor peso, las razones financieras del rubro de liquidez y las de rendimiento.

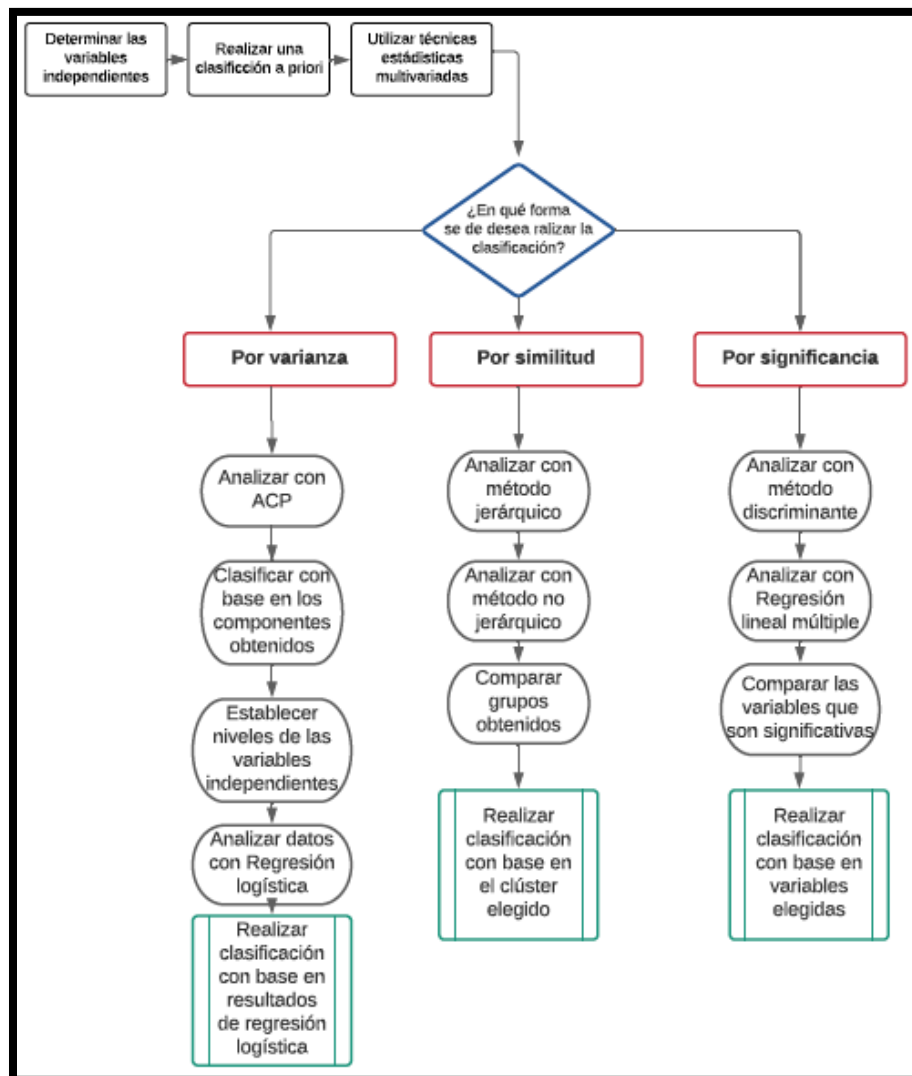
Respecto a la hipótesis específica cuatro, que consistió en que la mejor técnica estadística multivariada que puede aplicarse a las razones financieras para lograr una clasificación de empresas de éxito es la de Análisis de componentes principales, Análisis discriminante, Análisis clúster, Análisis K-medias, Regresión

lineal múltiple y Regresión logística binaria, se acepta, ya que cada una de ellas es capaz de agrupar e identificar empresas de éxito.

Finalmente la hipótesis general de la investigación fue que la aplicación de técnicas estadísticas multivariadas a las razones financieras, permite generar un modelo para clasificar empresas de éxito, se acepta, ya que se logró diseñar un modelo con una metodología específica seccionada en tres formas distintas.

En la figura 4, se muestra el resumen del modelo propuesto, que se obtuvo con base en los resultados obtenidos después de correr 6 análisis distintos.

Figura 4. Modelo de clasificación de empresas de éxito



El modelo presentado en la figura 4 muestra una metodología a seguir para lograr una clasificación de empresas de éxito, el cual comienza con la detección de las variables independientes que en el caso de la investigación fueron las razones financieras. Seguido a esto, será necesario realizar una clasificación a priori de éxito, utilizando la variable dependiente, que para el caso de la investigación fue la tasa de rendimiento de las empresas, pero también se recomienda aplicar esta metodología utilizando otra variable que sea utilizada para medir el éxito de una compañía.

Una vez que tengan bien definidos estos puntos, se procede a aplicar 3 combinaciones distintas de técnicas multivariadas, esto en función de la forma que se desee realizar la clasificación. El modelo propone tres formas distintas, las cuales de detallan en 10 puntos principales para llevarse a cabo de manera efectiva.

Clasificación por varianza:

Se propone realizar una clasificación por medio de una probabilidad de éxito, para lo cual es necesario implementar dos técnicas estadísticas multivariadas: el análisis de componentes principales y la regresión logística binaria.

Es importante mencionar que previo a aplicar esta combinación de análisis, es necesario realizar las pruebas de KMO y esfericidad de Bartlett, esto debido a que en este tipo de análisis es necesario verificar que la varianza pueda ser causada por factores subyacentes.

A continuación, se muestran los pasos esenciales a desarrollar para lograr esta clasificación:

1. Obtención de los eigenvalores: debe analizarse la proporción de variabilidad que es explicada por cada eigenvalor, para de esta forma determinar cuántos componentes principales son capaces de resumir la variabilidad del conjunto

de datos que se esté analizando. Se recomienda que el conjunto de eigenvalores represente al menos un 75% de la variabilidad total.

2. Elaborar la gráfica de sedimentación: este gráfico puede ser un apoyo visual para determinar apreciar la magnitud y la importancia que tiene cierto componente principal respecto a los demás componentes que se calculan en el análisis.
3. Obtener la matriz de correlaciones de componentes principales con las variables independientes: con esta matriz se determina que componente principal representa a cada conjunto de variables independientes del estudio. Esto se obtiene identificando cual es el mayor nivel de correlación (ya sea positivo o negativo) que tiene cada grupo de variables con cada componente.
4. Agrupación de casos en función de componentes principales elegidos: en esta parte es necesario agrupar cada caso de estudio de acuerdo al valor que obtiene cada uno en el componente principal elegido, esto permitirá analizar cada grupo de variables de una forma más compacta, además de ser el paso previo para generar escalas o niveles de agrupación.
5. Formación de niveles de agrupación: se sugiere utilizar una escala de Likert, basada en una medida de posición, tal como los cuartiles, deciles o percentiles, la medida a utilizar, dependerá de los niveles que se deseen implementar. Ya que se ha creado la escala de agrupamiento, se debe asignar a cada empresa, el nivel que ocupa en dicha escala, este es el paso previo a aplicar el análisis de regresión logística.
6. Análisis de regresión logística: las variables a introducir deben ser la empresa, el nivel que obtuvo en la escala creada y la variable dependiente categórica del estudio (ejemplo: éxito-no éxito).
7. ANOVA de regresión logística: verificar en este apartado la significancia estadística de cada variable independiente, para proseguir con el análisis o bien remover variables que no aporten al análisis.
8. Obtener la ecuación de regresión logística: esta ecuación puede ser utilizada para compararla con alguna otra que se haya obtenido en otros análisis.



9. Tabla de probabilidades de éxito: obtener la tabla de probabilidades de éxito que se obtiene a partir de la ecuación de regresión logística, en esta tabla deben verificarse las combinaciones que existen para el conjunto de variables independientes y las probabilidades asociadas de éxito en la variable dependiente y por lo tanto elegir la combinación la combinación que mayor probabilidad se calcule.
10. Realizar la clasificación de empresas de éxito: con base a la combinación de variables independientes elegidas se procede a realizar la clasificación de cada empresa en el grupo de éxito o no éxito.

Clasificación por similitud:

Se propone realizar una clasificación basado en la similitud de las empresas, respecto a las variables independientes utilizadas para el estudio. Para lograr este objetivo es necesario implementar dos técnicas estadísticas multivariadas: el análisis clúster y el análisis de K-medias.

Es importante notar que en este tipo de análisis no es necesario realizar las pruebas de KMO y esfericidad de Barlett, esto debido a que en este tipo de análisis no es necesario verificar que la varianza pueda ser causada por factores subyacentes, sin embargo, será necesario realizar otro tipo de pruebas.

1. Matriz de autovalores de covarianza clúster: esta matriz proporciona los primeros indicios de que es factible realizar un análisis clúster con el conjunto de datos. A mayor proporción de la variabilidad acumulada en menos variables independientes, mayor es la factibilidad.
2. Matriz de proximidades: esta matriz debe obtenerse para para explorar las distancias cuadradas que se encuentran entre cada empresa, derivada de la varianza, si las distancias son muy cortas al compararlas entre cada empresa, esto es un indicio de que puede no existir una buena separación



de grupos, es decir tenderían a ser homogéneos y por lo tanto no podrían realizarse clústeres.

3. Estadísticos de decisión: pueden utilizarse distintos estadísticos para determinar el número de grupos a formar, sin embargo, se recomienda emplear el criterio cubico cuadrado.
4. Dendrograma: se obtiene este tipo de gráfico con la intención de apreciar el número de clústeres posibles a formarse.
5. Realizar clasificación: con base en los dos últimos puntos redactados, se generan los clústeres determinados y se analiza la forma en que quedaron conformados (analizar el número de empresas exitosas que conforma cada clúster).
6. Análisis de K-medias: para corroborar o ver diferencias en los grupos generados por el análisis clúster, se debe correr el análisis no jerárquico de K-medias. En este análisis debe tenerse a priori un número de clústeres a formar, ya que los cálculos de agrupación parten de este dato, por lo que es importante que primero se lleve a cabo el análisis clúster para tener esta información (de no ser así tendría que realizarse muchas corridas del análisis y hacer comparaciones para la determinación del número óptimo de clústeres a formar).
7. Centroides iniciales: esta información la muestra el programa después de correr análisis, la cual muestra el valor de los centroides para cada variable, separadas por el número de clústeres formados inicialmente. Esta información sirve para contrastar los nuevos centroides que se formarán conforme avancen los cálculos del análisis.
8. Historial de iteraciones: el análisis de K-medias, realizará un cálculo de iteraciones para calcular los nuevos centroides, generalmente los programas estadísticos realizan hasta 10 iteraciones, sin embargo, este número puede aumentarse en las opciones del programa.
9. Realizar clasificación: con base al último cálculo elaborado se pueden generar los clústeres y verificar las empresas que se encuentran en cada



uno, para de esta forma verificar las empresas de éxito que se agrupan en cada uno.

10. Comparación y elección de clústeres: finalmente después de correr ambas técnicas estadísticas, se debe realizar una comparación entre los clústeres formados para elegir cual información es más útil para el investigador. Es importante notar que no siempre existirán diferencias significativas entre los clústeres formados, esto dependerá del conjunto de datos que se utilicen.

Clasificación por significancia:

Se propone realizar una clasificación basado en la significancia estadística que presenten las variables independientes del estudio. Para lograr este objetivo es necesario implementar dos técnicas estadísticas multivariadas: el análisis discriminante y el análisis de regresión lineal múltiple.

Es importante notar que en este tipo de análisis no es necesario realizar las pruebas de KMO y esfericidad de Barlett, esto debido a que en este tipo de análisis no es necesario verificar que la varianza pueda ser causada por factores subyacentes, sin embargo, será necesario realizar otro tipo de pruebas.

1. Estadístico de Lambda de Wilks: se calcula este estadístico para contrastar si las medias de los grupos son iguales, este es el primer paso para comenzar la validación de variables significativas para la clasificación de empresas de éxito o no éxito.
2. Prueba M de box: con esta prueba se contrasta si igualdad de covarianzas entre los grupos se cumple, de no ser así, el análisis discriminante será descartado y por lo tanto debe procederse a realizar un análisis con otra técnica estadística.
3. Correlación canónica: se calcula este valor que va desde -1.00 hasta 1.00, interpretándose como el valor de asociación entre el grupo de variables independientes y la variable dependiente. Este valor será útil al momento en

que se tengan que realizar ajustes en el análisis, si este nivel de asociación mejora, por lo tanto, la función discriminante mejora.

4. Matriz de estructuras: se calculan los coeficientes de la función discriminante para de esta forma conocer cuáles son las variables significativas para que una empresa pueda ser exitosa, el orden de importancia se conoce debido al valor de correlación que presenta cada variable con la función discriminante.
5. Resultados de clasificación: en esta parte final se evalúa la efectividad de la función discriminante, por lo que debe evaluarse el porcentaje de empresas exitosas que logran clasificarse de manera correcta con base en esta ecuación obtenida. Se recomienda correr de nuevo el análisis con base en las variables más significativas con el estadístico de Lambda de Wilks, y verificar los nuevos valores de correlación canónica, matriz de estructuras y los resultados de clasificación, para ver si existen mejoras sustanciales y entonces realizar la mejor agrupación de empresas exitosas.
6. Análisis de regresión múltiple: después de correr el análisis discriminante será necesario correr este análisis de regresión para contrastar las variables significativas y poder elegir mejor clasificación. En este análisis a diferencia del discriminante, al introducir los datos de la variable dependiente no será la variable categórica de éxito o no éxito, será la variable numérica de la tasa de rendimiento con la que se llegó a determinar dicha variable categórica.
7. Matriz de correlación y comprobación de supuestos: es necesario primero verificar las correlaciones entre todas las variables del análisis, para de esta forma tener un indicio de que tan factible puede ser el modelo (si no existe correlación entre variables dependientes y la dependiente, los resultados pueden llegar a ser de poca confianza y utilidad). Para poder dar certeza a los resultados del modelo de regresión lineal múltiple, debe revisarse que el conjunto de datos se ajuste a una probabilidad normal, que exista homogeneidad de varianzas y que no existe un patrón de estacionalidad en los datos.
8. ANOVA de regresión lineal: en este análisis se verifica la significancia estadística de las variables independientes, este punto es importante ya que

es posible detectar variables que no están influyendo en la variable dependiente, por lo que es un aspecto a considerar para correr de nuevo el análisis tomando únicamente en cuenta las variables significativas.

9. Significancia de coeficientes de regresión y coeficiente de determinación: conocer los coeficientes de la ecuación de regresión permite determinar la influencia que tiene cada variable independiente en la dependiente, también aquí deben detectarse si existen problemas de multicolinealidad, revisando los valores VIF. En caso de detectar esta problemática, debe correrse de nuevo el análisis utilizando solo las variables significativas y las variables que no estén correlacionadas entre sí. Por otro lado, el coeficiente de determinación mide la capacidad que tiene la ecuación de regresión lineal múltiple para explicar el comportamiento del rendimiento de la empresa, por lo que con este coeficiente se pueden verificar las mejoras en la ecuación al cada vez que se corra el análisis.
10. Realizar la clasificación de empresas de éxito: después de haber corrido ambos análisis es necesario contrastar las variables que fueron significativas y hacer comparaciones entre las ecuaciones obtenidas (función discriminante y ecuación de regresión lineal) para determinar cuáles variables se deben elegir para realizar una clasificación de empresas de éxito.

V.1. Recomendaciones

El modelo presentado fue construido con base en las razones financieras como variables que determinan el éxito de una empresa, sin embargo, en este punto también se recomienda que este modelo pueda aplicarse a otro tipo de variables independientes que puedan ser causantes del éxito, tomando como medida el rendimiento de la misma o bien tomando en cuenta un indicador distinto como lo han hecho otros investigadores.

También se sugiere replicar este modelo en empresas que cotizan en otros mercados como en la Bolsa Mexicana de Valores, o el New York Stock Exchange,



esto con la finalidad de extender el uso de la estadística multivariada en el campo del análisis financiero.

Bibliografía

- Abdi, H., & Williams, L. J. (2010). *Principal component analysis* (Vol. 2). doi: <https://doi.org/10.1002/wics.101>
- Álvarez Suárez, M. M. (2006). Análisis Multivariante: clasificación, organización y validación de resultados. *Breaking frontiers and barriers in engineering; educational, research and practice*. Mayaguezm Puerto Rico: LACCEI.
- Ávila, H. L., & Alarcón, M. G. (2014). Factores de éxito en empresas innovadoras. *innovaciones de negocios*, 21, 143-167.
- Baeza, S. R., & Vázquez, L. J. (2014). Transición de un modelo de regresión lineal múltiple predictivo, a un modelo de regresión no lineal simple explicativo con mejor nivel de predicción: Un enfoque de dinámica de sistemas. *Revista Facultad De Ingeniería*, 71, 59-71.
- Balakrishnama, S. (1999). Linear discriminant analysis for signal processing problems. *Proceedings IEEE Southeastcon'99. Technology on the Brink of 2000*, 78 - 81.
- Beaver, W. (1968). Alternative Accounting Measures as Predictors of Failure. *The Accounting Review*, 43, 113-122. Obtenido de www.jstor.org/stable/244122
- Block, S., Hirt, G., & Danielsen, B. (2013). Fundamentos de administración financiera. México D.F.: McGraw-Hill.
- Block. (2013). *Fundamentos de administración financiera* (Vol. 14). Ciudad de México: MC Graw-Hill.
- BMV. (1 de 12 de 2018). *Bolsa Mexicana de Valores*. Obtenido de <https://www.bmv.com.mx/es/mercados/capitales>
- Bonilla, F. (2010). El valor económico agregado en el valor del negocio. *Revista nacional de administración*, 55-70.

- Cardona, O. J. (2015). Análisis de indicadores financieros del sector manufacturero del cuero y marroquinería. *Informador técnico*, 79(2), 156-168. doi: <https://doi.org/10.23850/22565035.160>
- Castro posada, J. A., & Galindo Vilardón, M. P. (2000). *Estadística multivariante, análisis y correlaciones*. Madrid España: Amarú ediciones.
- Chávez, C. (2015). Análisis de componentes principales funcionales en series de tiempo económicas. *Revista Internacional de Gestión del Conocimiento y la Tecnología*, 3, 5-9. Obtenido de <https://www.upo.es/revistas/index.php/gecontec/article/view/1694>
- CNN-money. (2004). Obtenido de <https://money.cnn.com/2004/08/19/technology/goog/>
- Collins, J. (2002). *Empresas que sobresalen; por qué unas sí pueden mejorar la rentabilidad y otras no*. Colombia: Norma.
- Contreras, O. E., Bronfman, R. S., & Arenas, C. E. (2015). Estrategia de inversión optimizando la relación rentabilidad-riesgo: evidencia en el mercado accionario colombiano. *Estudios Gerenciales*, 31(137), 383-392. doi:<https://doi.org/10.1016/j.estger.2015.07.005>
- Costin, D. (2017). Economic Value Added – A general review of the concept. *Ovidius: Economic Sciences Series*, 17(1), 168-173.
- De la fuente, S. (2011). Análisis de conglomerados. Madrid, España.
- De la Garza García, J., Morales Serrano, B., & González Cavazos, B. A. (2017). *Análisis estadístico multivariante: un enfoque teórico y práctico* (Vol. 7). México D.F.: Mc Graw Hill.
- Díaz, G. M. (2018). La transferencia de conocimiento en las alianzas estratégicas para la innovación. *Revista de ciencias de la gestión*. doi:<https://doi.org/10.18800/360gestion.201803.006>
- Díaz, M. (2005). Una propuesta para un sistema de análisis técnico para compra y venta de acciones en la Bolsa Mexicana de Valores. *Contaduría y administración*, 25-48.
- Escobar Jaramillo, L. A. (2008). Indicadores ambientales sintéticos: Una aproximación conceptual desde la estadística multivariante. *Gestión y*



- ambiente*, 11(1). Obtenido de
<https://www.redalyc.org/articulo.oa?id=169414452008>
- Everitt, B. (2002). *Cluster analysis*. E.U.A.: Wiley.
- Fierro, O. (15 de noviembre de 2010). *Derecho* . Recuperado el 15 de abril de 2013, de
<http://www.google.com/cse?cx=002866702082184079149%3Ahvjb6-5ekau&q=requisitos+para+publicar+en+Synthesis&sa=Buscar#gsc.tab=0&gsc.q=requisitos%20para%20publicar%20en%20Synthesis&gsc.page=1>
- Flores, C. (2011). *Tecnología de Información*. Madrid: Pearson.
- Flores, C. (2013). *Metodología de la Investigación*. México: Pearson.
- Gaytán Cortés, J. (2018). Clasificación De Los Riesgos Financieros. *Mercados Y Negocios*, 38, 126-136.
- Gengyuan , L. (2018). Big data-informed energy efficiency assessment of China industry sectors based on K-means clustering. *Journal of Cleaner Production*, 183, 304-313.
- Gitman, L., & Zutter, C. (2015). *Principios de administración financiera* (12 ed.). México: Pearson.
- González, R. (2013). *Trata de Personas*. Madrid: Mc Graw Hill.
- Gutiérrez, S. D., & Medina, H. S. (2018). Estimación robusta de la matriz de covarianza para la selección óptima de portafolios de inversión. *DYNA*, 85, 328-336. doi:<http://dx.doi.org/10.15446/dyna.v85n207.74883>
- Hair, J. (1999). *Análisis multivariante*. Madrid.
- Hair, J. (2000). *Análisis multivariante* (Vol. 6). Madrid España: Prentice Hall.
- Hair, J. (2000). *Análisis multivariante* (Vol. 6). Madrid España: Prentice Hall.
- Hennig, C. (2003). *Handbook of cluster analysis*. (C. Press, Ed.) E.U.A.
- Horrigan, J. (Julio de 1965). Some Empirical Bases of Financial Ratio Analysis. *The Accounting Review*(3), 558 - 568. Obtenido de
www.jstor.org/stable/243060
- Ibarra, M. A. (2006). Una perspectiva sobre la evolución en la utilización de las razones financieras o ratios. *Pensamiento y gestión: Revista de la división de ciencias administrativas de la universidad del norte*(21), 234-71.

Obtenido de <https://biblat.unam.mx/es/revista/pensamiento-gestion/articulo/una-perspectiva-sobre-la-evolucion-en-la-utilizacion-de-las-razones-financieras-o-ratios>

- Jhonson, R. (2002). *Applied Multivariate Statical Analysis* (Vol. 5).
- Libby, R. (1975). Accounting Ratios and the Prediction of Failure: Some Behavioral Evidence. *Journal of Accounting Research*.
doi:<https://doi.org/10.2307/2490653>
- López Pérez, C. (2004). *Técnicas de análisis multivariante de datos*. Madrid, España: Pearson.
- Lopez, Á. A. (2008). Introducción a la regulación de valores en los Estados Unidos. *Mercatoria*, 7(2).
- Mantilla, R. (2008). *Derecho Mercantil*. Mexico: Mc Graw Hill.
- Mbona, R., & Yusheng, K. (2019). Financial statement analysis: Principal component analysis (PCA) approach case study on China telecoms industry. *Asian Journal of Accounting Research*., 4(2), 233-245. doi: <https://doi.org/10.1108/AJAR-05-2019-0037>
- Méndez, M. C., & Rondón, S. M. (2012). Introducción al análisis factorial exploratorio. *Revista Colombiana de Psiquiatría*, 41(1), 197-207.
doi:[https://doi.org/10.1016/S0034-7450\(14\)60077-9](https://doi.org/10.1016/S0034-7450(14)60077-9)
- Mendoza Méndez, R., & Dorantes Coronado, E. J. (2016). Discriminant Function Analysis as interpretive tool for the Study on Mobile Phone Overuse in students of the Bachelor's degree in Administrative Data Processing of the Autonomous University of Mexico State, University Center UAEM Temascaltepec. *cenid*. doi:10.23913/ride.v7i14.282
- Milla, M. (2011). Estudio de métodos de prueba en el análisis de regresión lineal multivariado usando correcciones tipo Bartlett. *Aporte Santiaguino*, 4, 57-63.
- Minitab. (2020). *Soporte Minitab*. Recuperado el 2020, de <https://support.minitab.com/es-mx/minitab/18/help-and-how-to/modeling-statistics/multivariate/how-to/cluster-k-means/methods-and-formulas/cluster-k-means/>

- Moradi Fard, M., Thonet, T., & Gaussier, E. (2020). Jointly clustering with k-Means and learning representations. *Pattern Recognition Letters*, 138, 185-192.
- Morelos, G. J. (2013). Análisis multivariado de la productividad y rendimientos financieros de empresas industriales en Cartagena, Colombia. *Apuntes del Cenes*(32). doi:<https://doi.org/10.19053/22565779.2074>
- Mures, Q. J. (2013). Aplicación de técnicas estadísticas multivariantes en el tratamiento de información económico-financiera. *Estadística española*, 55, 337-364. Obtenido de <https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=2ahUKEwiOsvqst4HoAhUEPK0KHSYPCwQQFjAAegQIBRAB&url=https%3A%2F%2F Dialnet.unirioja.es%2Fdescarga%2Farticulo%2F6014736.pdf&usg=AOvVaw2OygpmoPwlajExKJBgQtn0>
- Nasdaq. (2020). Obtenido de <https://www.nasdaq.com/news-and-insights/company-intel>
- Nasdaq. (2020). *NASDAQ BACKGROUND*. Obtenido de <http://ir.nasdaq.com/news-releases/news-release-details/nasdaq-january-2020-volumes>
- Nasdaq. (2020). *NASDAQ PRESS RELEASES*. Obtenido de <https://www.nasdaq.com/market-activity/quotes/press-releases>
- Navia, C.E.,M. Ossa. (Enero de 2001). *Revista de Estudios Sociales*. Recuperado el 24 de mayo de 2013, de Universidad de los Andes. Facultad de Ciencias Sociales: <http://res.uniandes.edu.co/view.php/203/pdf/view.php>
- Nikolsko-Rzhevskaya, O., Nikolsko-Rzhevskyy, A., & Black, J. (2020). The life of U's: Order revisions on NASDAQ. *Journal of Banking & Finance*. doi:<https://doi.org/10.1016/j.jbankfin.2019.105724>
- Núñez, C. C., & Escobedo, L. D. (2011). Uso correcto del análisis clúster en la caracterización de germoplasma vegetal. *Agronomía Mesoamericana*, 2, 415-427. doi:<https://doi.org/10.15517/am.v22i2.8746>
- Ospina, H. J. (2015). Medidas dinámicas de predictibilidad en el índice s&p 500 y sus determinantes. *Innovar : Revista de Ciencias Administrativas y Sociales*, 56(1), 69-82.

- Pérez Grau, S. L. (2013). Selección de índices financieros mediante técnicas estadísticas del análisis multivariante. *Entramado*, 9. Obtenido de <http://www.scielo.org.co/pdf/entra/v9n1/v9n1a08.pdf>
- Pérez, R. (2011). *Metodología*. México: Pearson.
- Puerto, B. D. (2010). La globalización y el crecimiento empresarial a través de estrategias de internacionalización. *Pensamiento y gestión*., 28, 171-195.
- Quiroga, C., & Villalobos, A. (2015). Análisis del comportamiento bursátil de las principales bolsas financieras en el mundo usando el análisis multivariado (análisis de componentes principales PCA) para el periodo de 2011 a 2014. *CEA*, 25-36. doi:<https://doi.org/10.22430/24223182.122>
- Quiroga, J. C. (2016). Aplicación de dos técnicas del análisis multivariado en el mercado de valores mexicano. *Revista de métodos cuantitativos para la economía y la empresa*, 104-119. Obtenido de <https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=2ahUKEwj6r-OtUlHoAhUI2qwKHSuCDUoQFjAAegQIBBAB&url=https%3A%2F%2Fwww.upo.es%2Frevistas%2Findex.php%2FRevMetCuant%2Farticle%2Fview%2F2341%2F1917&usg=AOwVaw0PRH339XLfie-HwKYkrN-C>
- Ramírez, V. G., Islas, M. J., & Valverde, R. B. (2011). Asociación espuria en el modelo de regresión logística con series de tiempo. *Agrociencia*, 45(5), 583-591.
- Reding, B. A., López, A. J., & Zamora, M. M. (2011). ¿Cómo y cuándo realizar un análisis de regresión lineal simple? Aplicación e interpretación. *Producción y ciencia: Revista Mexicana*, 395-40.
- Rodriguez, C. (2012). *Derecho Penal*. México: Pearson.
- Rojas, L. M., Rincón, L. C., & Mesa, L. S. (2015). Alianzas estratégicas: Alternativas generadoras de valor. *Universidad y empresa*, 27, 281-302.
- Ross, S. A. (2009). *Finanzas corporativas* (Vol. 9). México: Mc Graw-Hill.
- Ross, S., Westerfield, R., & Jaffe, J. (2012). *Finanzas corporativas* (9 ed.). México: Mc Graw Hill.



- Rueda, A. (2005). *Para entender la bolsa financiamiento e inversión en el mercado de valores*. México: Cengage Learning.
- Rueda, V. M., Moya, M. L., & Aranada, S. M. (2011). Aplicación de técnicas estadísticas multivariadas en perfilación y segmentación. *Universitarias Scientiarum*, 254-262.
- Salgado, M. (2017). La Crisis Del Crédito Hipotecario Subprime: Una Revisión. *Realidad*, 128, 243-60. doi:<https://doi.org/10.5377/realidad.v0i128.3249>
- Samreen, F., Tayyab , R., & Rafia , S. (2019). The application of principal component analysis and factor analysis to stock markets returns. *International Journal of Advanced Research*, 7, 97-105. doi:<http://dx.doi.org/10.21474/IJAR01/9008>
- Sanguansat, P. (2012). Two-Dimensional Principal Component Analysis and Its Extensions. En *Principal component analysis*. doi:<https://doi.org/10.5772/36892>
- SAS. (2020). *SAS Documentation*. Obtenido de https://documentation.sas.com/?docsetId=statug&docsetTarget=statug_cluster_syntax01.htm&docsetVersion=15.1&locale=en
- Schall, L. y. (2000). *Introduction to financial management* (Vol. 6). Nueva York: Mc Graw Hill.
- Servín y Silva, F. H. (2011). Estimación de la volatilidad de los precios de las acciones de la BMV mediante el modelo CARR: El caso de AMX-L. *Contaduría Y Administración*, 234, 173-96. Obtenido de http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S0186-10422011000200009
- Solano, D. O., & Ramírez, T. A. (2007). Análisis de diagnóstico en el modelo de regresión logística. *Pesquimat*, 10. doi:<https://doi.org/10.15381/pes.v10i1.9431>
- SPSS statics, 2. (s.f.). *IBM Knowledge center*. Recuperado el 2020, de https://www.ibm.com/support/knowledgecenter/SSLVMB_23.0.0/spss/tutorials/fac_telco_kmo_01.html

- Urrutia, M. J., & Palomino, L. P. (2010). Componentes principales en la determinación de estaciones con patrones homogéneos de temperatura en el chocó. *Scientia et Technica*, 2(45), 257-262. Obtenido de <https://dialnet.unirioja.es/servlet/articulo?codigo=4544856>
- Vásquez, Q. N., & Mora, R. J. (2016). Análisis de los determinantes de revelación de instrumentos financieros mediante el uso de componentes principales: el caso de empresas mexicanas. *Cuadernos De Contabilidad*, 44, 467-492. doi:<https://doi.org/10.11144/Javeriana.cc17-44.adri>
- Véliz, C. C. (2017). *Análisis multivariante: Métodos estadísticos para la investigación* (Vol. 1). México D.F.: CENGAGE Learning.
- Véliz, C. C. (2017). *Análisis multivariante; métodos estadísticos para la investigación* (Primera ed.). Buenos Aires, Argentina: Cengage.
- Walsh, C. (2001). *Ratios fundamentales de gestión empresarial*. Madrid, España: Pearson.
- Wilmer, E. J. (2015). Metodología para la toma de decisiones de Inversión en portafolio de acciones utilizando la técnica multicriterio AHP. *Contaduría y administración*, 60(2), 324-366.
- Zulaica, P. C. (2013). Financial inclusion index: proposal of a multidimensional measure for Mexico. *Revista Mexicana de Economía y Finanzas*, 8, 157-180. doi:<https://doi.org/10.21919/remef.v8i2.46>