

Manejo de la información

Breve descripción:

En este documento se comprenderá el manejo de la información y sus tipos de análisis, los cuales son necesarios para la toma de decisiones, dirigidos a minimizar la incertidumbre y el riesgo en el mercado. Igualmente se abordarán las variedades de presentación de la información de mercados, su procesamiento, desde su recolección, depuración, análisis y construcción de conocimiento de mercados.

Tabla de contenido

Introducción.....	4
1. La información	6
1.1 Datos, información y conocimientos	6
1.2 Importancia de la información	7
1.3 Fuentes de información.....	8
1.4 Técnicas de recolección de la información	9
1.5 Análisis de la información.....	10
2. Variables de información	13
2.1 Tipos de variables	13
2.2 Clasificación de las variables.....	15
3. Procesamiento de la información.....	17
4. Análisis de datos	19
4.1 Antecedentes	19
4.2 Tipos de análisis.....	19
4.3 Técnicas de análisis de datos	22
5. Técnicas de evaluación	28
6. Informes y documentos.....	38
6.1 Informes.....	38

6.2 Tipos de informes	38
Síntesis	40
Material complementario.....	41
Glosario.....	42
Referencias bibliográficas	44
Créditos.....	48

Introducción

En el siguiente video se presenta la introducción al manejo de la información de los clientes y sus características y necesidades, con esto la empresa puede tomar buenas decisiones de mercado.

Video 1. manejo de la información



Manejo de la información introducción

Enlace de reproducción del video

Síntesis del video: manejo de la información

Manejo de la información

Para que la empresa tome decisiones de mercado, debe tener claro a quiénes va a servir, cuáles son sus características y necesidades. Para ello, es fundamental el

manejo expedito de la información. En el desarrollo de este componente se tratan aspectos del manejo de la información, la interpretación basada en los diferentes tipos de análisis cuantitativos y cualitativos. Ello define la propuesta de la empresa a cada uno de los segmentos de mercado seleccionados. Asimismo, se realiza un recorrido por la recolección de la información y sus diferentes técnicas de recogida, operacionalización de variables, además de los pasos para definir el procesamiento de la información: preparación, interpretación y almacenamiento. Luego se mira a qué estrategias de análisis de datos tradicionales y modernos recurren las empresas de hoy para alcanzar una mayor cobertura de mercado. Finalmente, los tipos de informes que se construyen con el proceso anteriormente expuesto y sus elementos básicos. Bienvenidos.

1. La información

En el actual siglo XXI, una de sus principales características es la valía de la información para la toma de decisiones tanto empresariales, como ambientales y sociales; la sociedad vive en un mundo de permanente bombardeo de información en noticias de prensa, radio y televisión, pero con la masificación de la web y de las redes sociales, hoy en día el ser humano vive rodeado de información útil, pero también de información bastante inservible o inútil.

Ahora bien, en los desempeños en que cotidianamente se desenvuelve la gran mayoría de las profesiones, oficios y artes, la información es trascendente en la toma de decisiones para el desempeño laboral. Para que este proceso de decisión sea lo más acertado posible es necesario determinar primero lo que es dato, información y generación de conocimientos.

1.1 Datos, información y conocimientos

En el presente milenio la sociedad científica ha caracterizado y denominado a la población como la sociedad de la información, según las expresiones de Ruiz de Querol (2007), y es que ha sido tanta la exposición de la sociedad a todo tipo de información y en todos los momentos, que se ha convertido en una desventaja competitiva no hacer parte de esta sociedad del conocimiento, para esto es importante entender como los datos se convierten en conocimiento en el contexto contemporáneo.

El análisis de datos e información se da en el ser humano

- **Datos (el inicio):** es la diferenciación simbólica de una variable que puede ser clasificada como cualitativa o cuantitativa, indicando un valor representativo a dicha variable para su análisis, estudio y clasificación. Por

ejemplo: Juan pesa 80 kilogramos, la expresión 80 es el valor representativo del peso de Juan que se mide en kilogramos.

- **Información:** es la denominación que se le da a la agrupación organizada de datos significativos que describen sucesos, eventos y hechos, con cuyo procesamiento se puede construir un mensaje descriptivo y / o analítico.
- **Conocimiento (el fin):** es la explicación e interpretación de la información producto de los procesos cognitivos de comparar, analizar, combinar o experimentar, permitiendo al sujeto la argumentación y la aplicación de la información.

Datos - información y conocimientos.

1.2 Importancia de la información

Es importante estar debidamente informado sobre un propósito, porque no solamente permite desenrollar un punto de vista o postura respecto del mismo, sino también porque permite adelantarse a un hecho o estar prevenidos sobre determinada situación.

Según Florencia (2013) en sus blogs expresa “Gracias a la información se puede conceptuar, significar o encontrarle sentido a las cosas que rodean o acontecen en la sociedad, para luego de concebir el conocimiento debido, tomar decisiones”.

Entonces, la información es necesaria y vital en muchos aspectos de la vida personal y profesional, y en los tiempos presentes se le atribuye a la misma una especial valoración en diversos contextos ya que sin duda alguna en la actualidad la humanidad está atravesando la era de la información.

1.3 Fuentes de información

Por fuente de información se entiende, cualquier instrumento o en un sentido más amplio, recurso, que permita satisfacer una necesidad informativa. La expresión fuente se entiende entonces, como el documento, sitio, lugar, página, domicilio, persona natural o jurídica, en donde se encuentra la información que necesita un demandante de la misma.

Fuentes de información primaria y secundaria

Obtener información confiable y válida requiere cuidado y dedicación, pues de ello depende la confiabilidad y validez del estudio, estos datos o información que va a recolectarse en el medio, a través del cual se responden las preguntas de investigación y se logran los objetivos del estudio. Los datos deben ser pertinentes y suficientes para lo cual es necesario definir las fuentes y técnicas adecuadas para su recolección.

- **Fuentes de información primarias**

Son todas aquellas personas, empresas, organizaciones, los acontecimientos que aquejan o benefician a la sociedad, el propio medio ambiente natural de las donde se obtienen información directa, es decir, de donde se origina la información, también son conocidas como fuentes de primera mano o desde el lugar de los hechos.

- **Fuentes de información secundarias**

Son todas aquellas revistas, libros, los documentales, los noticieros, medios de información y los documentos escritos (todo medio impreso físico o digital) que ofrecen información sobre el tema por investigar, pero que no son la fuente original de los hechos o situaciones, sino que los referencian.

1.4 Técnicas de recolección de la información

Debe entenderse que los procesos de recolección de la información constan una gran multiplicidad de técnicas o herramientas para la recopilación de la misma, las técnicas más usadas son:

Técnicas de recolección de la información

- **Encuesta**

La encuesta, según Bernal (2000), “una de las técnicas más reconocidas y utilizadas, a pesar de que ha perdido credibilidad por el sesgo que hay en las personas encuestadas y por el uso cada vez más frecuente de TIC’S” (p.172).

Por tanto, la encuesta está fundada en un agregado de interrogantes que son planeadas con la finalidad de obtener información de las personas en referencia a una temática o hecho específico.

- **Entrevista**

La entrevista es una técnica orientada a obtener una conversación estructurada con las personas con apropiación de la información, la entrevista a pesar de manejar un cuestionario base con el cual se establecerá una comunicación dialógica con el entrevistado, permite la flexibilidad de ampliar la respuesta del entrevistado con nuevas preguntas.

- **Observación directa**

La observación directa cada día cobra mayor fuerza y credibilidad, y cada vez tiende a generalizarse más, dado que permite la obtención de información directa y confiable, siempre que medie un procedimiento sistematizado y supervisado, para lo cual hoy en día se utilizan medios

audiovisuales muy completos y con alto componente tecnológico en las industrias 4.0, permite estudios de comportamiento del consumidor, del personal en el puesto de trabajo, estímulos basados en percepciones sensoriales en pruebas de productos, entre otros.

- **Análisis de documentos**

El análisis de documentos se utiliza bajo los criterios de fichas bibliográficas que analizan material impreso o digital, su principal uso está en la elaboración del marco teórico del estudio.

- **Búsqueda web**

La búsqueda web o consulta de bases de datos, páginas web, bloog's, en la red virtual se ha convertido en técnica de recolección de información por excelencia de toda la sociedad.

1.5 Análisis de la información

La información tiene diversas formas de ser examinada y analizada, de ello, dependerá las conclusiones que sean pertinentes al proceso de investigación que se está desarrollando, por tanto, a continuación se presenta la conceptualización de las formas de análisis de la información.

Análisis de la información

Investigación cualitativa

Según Sampieri (2006) “se fundamentan más en un proceso inductivo (explorar y describir, y luego generar perspectivas teóricas), van de lo particular a lo general en el análisis de fenómenos sociales” (p.8).

Según Báez (2009) en su definición agrega el componente comercial muy de moda en los presentes contextos ofreciendo el concepto de investigación cualitativa como “el conjunto de todas las cosas que se hacen para seguir la pista a los mercados y encontrar los rasgos que distinguen a las personas (cliente, consumidor, organización...) y a las cosas (productos, bienes, servicios, sectores de actividad) sus propiedades y atributos, sean estos naturales y / o adquiridos” (p.35).

Ventajas y desventajas de la investigación cualitativa

- **Ventajas**

- Permite profundización.
- Evaluación correspondencia verbal - corporal.
- Mayor rapidez.

- **Desventajas**

- No cuantifican.
- Poco personal especializado.

Investigación cuantitativa

Es un método de investigación en el que, según Sampieri (2006), el investigador recolecta datos numéricos de los objetos, fenómenos o participantes, que estudia y analiza mediante procedimientos estadísticos.

Ventajas y desventajas de la investigación cuantitativa

- **Ventajas**

- Permite hacer proyecciones.
- Mayor variedad de opiniones.

- Hace más tangible la relación causa / efecto.

- **Desventajas**

- Altos costos.
- Mayor tiempo.
- No profundiza.
- No registra reacciones.
- Requiere más personal.

Para concretar la investigación cualitativa pretende profundizar en los conceptos y valores, detectar motivadores y valorar las reacciones y para su ejecución requiere un soporte de psicología. Por otro lado en la investigación cuantitativa pretende cuantificar, establecer proporciones y diferencias e importancia relativa y para su ejecución requiere un soporte estadístico.

2. Variables de información

Según Espinoza (2018), las variables son elementos que se ponen en dependencia a un referente teórico y según lo que se puede pronosticar en la población de estudio, por tanto, la variable es una propiedad que puede cambiar de valor en cada situación o de un individuo a otro y cuyo cambio o alteración es capaz de medirse en torno a una escala.

2.1 Tipos de variables en el análisis de datos

Existen varios tipos de variables y cada una de ellas representa una característica, número o cantidad que puede ser medida o contada.

Tipos de variable

Obtener información confiable y válida requiere cuidado y dedicación, pues de ello depende la confiabilidad y validez del estudio, estos datos o información que va a recolectarse en el medio, a través del cual se responden las preguntas de investigación y se logran los objetivos del estudio. Los datos deben ser pertinentes y suficientes para lo cual es necesario definir las fuentes y técnicas adecuadas para su recolección.

Variables numéricas

Son aquellas que personifican números y con ellas se logran realizar procedimientos aritméticamente. Por ejemplo, ventas mensuales puede representar a las ventas de productos que hay en un mes cualquiera. También son conocidas como variables cuantitativas, que a su vez se clasifican en discretas y continuas.

- **Variables discretas**

Son aquellas cuyos valores corresponden a clases independientes, no admiten valores intermedios y por lo tanto en ellas no tiene sentido hablar

de promedios estadísticos, un ejemplo clásico, es la respuesta a la pregunta cuántos hermanos tienen los 5 amigos de Javier Martínez, la respuesta a esta pregunta es 3, 2, 1, 0, 2; es preciso aclarar que no cabe en las respuesta un amigo de Javier Martínez que tenga 1.5 hermanos o alguna respuesta semejante.

- **Variables continuas**

Son variables cuyos valores cambian en carácter progresivo y no a irrupciones o de otra condición, tienen una disposición que no puede alterarse y sobre ellas se pueden hacer análisis estadísticos descriptivos mediante medidas de tendencia central (media, mediana y moda) y medidas de dispersión. Por ejemplo, la estatura de los jugadores de la selección Colombia de fútbol, la respuesta sería que el jugador A mide 1.80 m, el jugador B 1,82 m, el jugador C 1.79 m, el jugador D 1.85 m y el jugador E 1.78 m; sobre esa estatura es típico poder calcular medidas de tendencia central estadísticas, como el promedio de estatura que sería 1.80 m.

Variables categóricas

Son un número definido de valores diferentes o categorías (por ejemplo, raza, género, sexo o religión). Las variables categóricas pueden ser nominales u ordinales.

- **Variable ordinal**

Es un tipo de variable estadística de tipo cualitativo que expresa con palabras una cualidad de naturaleza ordenable como por ejemplo primer puesto, segundo puesto, tercer puesto.

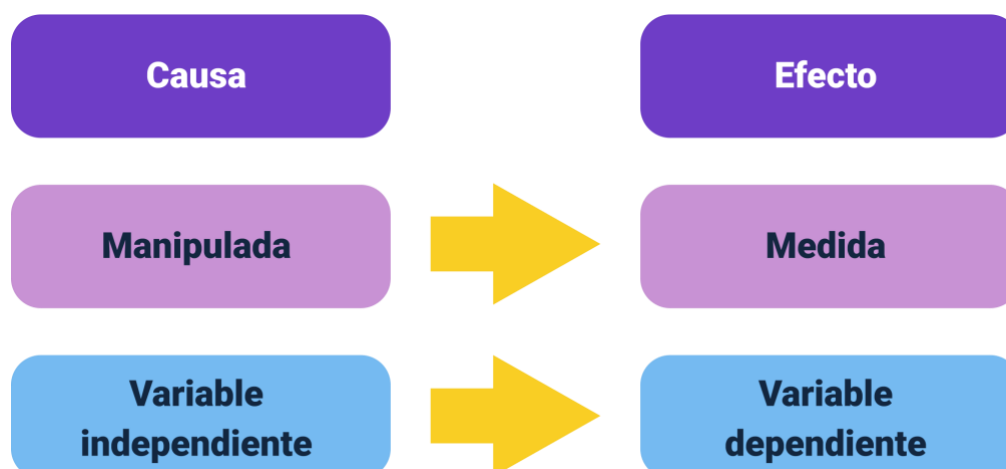
- **Variables nominal**

Es un tipo de variable estadística de tipo cualitativo que expresa con nombre una cualidad no necesariamente ordenable, como puede ser por ejemplo los colores de ojos de un salón de clases.

2.2 Clasificación de las variables

Se pueden identificar 2 tipos de variables que participan en el manejo de la información de un proceso investigativo y pueden señalar relaciones comportamentales al medir determinado número de sucesos:

Figura 1. Causa y efecto



Nota. <https://bit.ly/3uLW1v1>

La imagen que se presenta organiza la información en dos columnas claramente definidas: causa y efecto. En la columna de causa, se encuentran los términos "manipulada" y "variable independiente", que indican cómo ciertos factores pueden ser alterados o controlados en un experimento. Por otro lado, en la columna de efecto, se mencionan las palabras "medida" y "variable dependiente", que se refieren a los

resultados que se obtienen como consecuencia de las variaciones en la variable independiente.

- **Variable dependiente:** es definida por Bernal (2003) como el “resultado” o “efecto” producido por la variable independiente.
- **Variable independiente:** es definida como todo aquel aspecto, hecho, situación, rasgo, que se considera como “la causa de” en una relación entre variables.

3. Procesamiento de la información

Es un proceso que implica:

- Determinación clara de los objetivos que ameritan la necesidad de información.
- Tener una población plenamente identificada para la recolección de la información.
- Identificar las técnicas de recolección de la información debidamente validadas.

A continuación se presenta una cadena de pasos que pueden emplearse para la recaudación de los datos necesarios para responder a los objetivos:

- **Recogida de los datos**

Esta parte del procesamiento de la información requiere previamente haber perfilado los instrumentos de recolección que hayan sido validados y que generen confiabilidad. La recolección de datos se realiza siguiendo tres (3) pasos:

- Elegir y confeccionar un instrumento de recolección de datos validado.
- Utilizar el instrumento de recolección de datos validado en las fuentes de información previamente escogidas.
- Una vez recopilada la información obtenida se debe codificar.

- **Procesamiento de los datos**

Debe efectuarse mediante el uso de herramientas estadísticas apoyadas en

software reconocidos por la comunidad científica, que hoy fácilmente se encuentran en el mercado. Para realizar la actividad de procesar los datos se siguen unos pasos de los cuales Bernal (2003) postula lo siguiente:

- Obtener la información de la población o muestra objeto de la investigación.
- Definir las variables o criterios para ordenar los datos obtenidos.
- Definir las herramientas estadísticas y software reconocidos por la comunidad científica.
- Introducir los datos en el computador y activar el programa para que procese la información.
- Imprimir los resultados. (p.179).

- **Interpretación de los datos**

Se refiere a la ejecución de procesos por medio de los cuales se examinan los datos con el fin de llegar a una conclusión informada y una etapa esencial del procesamiento de datos.

- **Almacenamiento de los datos**

Es el proceso de custodia o protección de la información a través del empleo de un conjunto de técnicas desarrollada enfáticamente para guardar esos datos y garantizar su acceso en el momento en que se requiera.

4. Análisis de datos

El análisis de datos es el proceso de examinar conjuntos de datos para encontrar tendencias y sacar conclusiones sobre la información que contienen; en la investigación cuantitativa apela a técnicas de cálculo de variables y a la estadística como técnica específica en la ordenación y la exégesis de datos cuantitativos o cualitativos.

4.1 Antecedentes

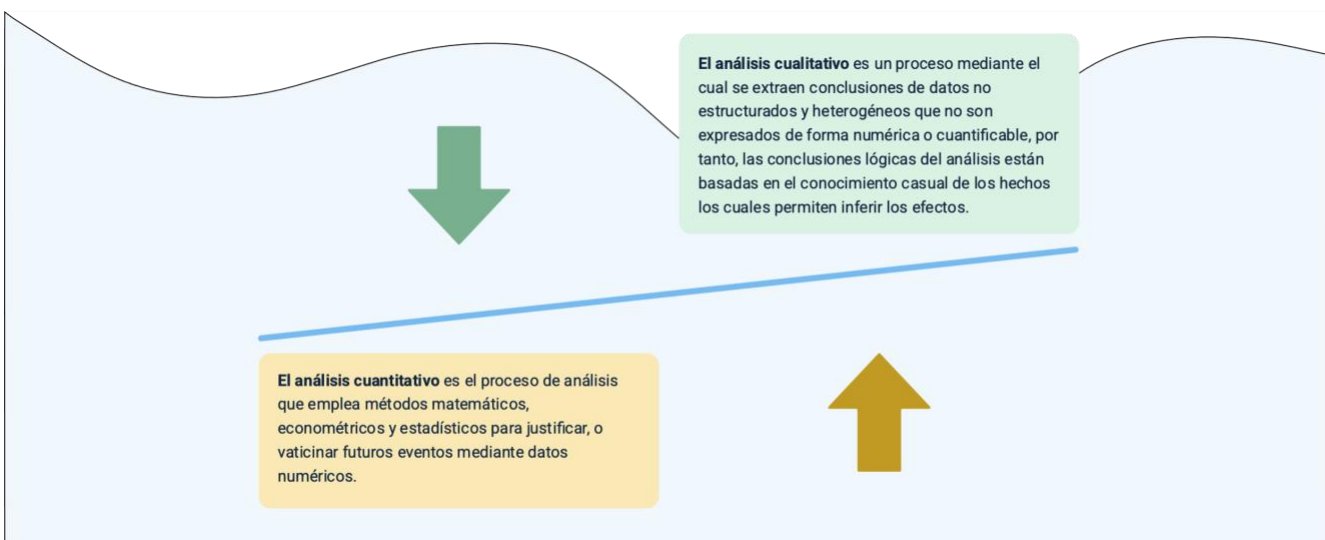
Son todos los trabajos realizados previamente sobre el tema que se va a investigar y que aportan información relevante para nuestro trabajo. La página web de Quality Leadership University (QLE) indica que tradicionalmente en el análisis de antecedentes se verifica:

- Los resultados de otras investigaciones o estudios.
- La importancia de estos con relación al tema planteado.
- Título y autores del trabajo de referencia.
- El planteamiento del problema, los objetivos e hipótesis planteadas por estos estudios o investigadores.
- Metodología utilizada.

4.2 Tipos de análisis de datos

Son varios los tipos de análisis de datos que se pueden realizar dependiendo de la información que se dispone, a continuación se destacan los más importantes.

Figura 2. Tipos de análisis de datos



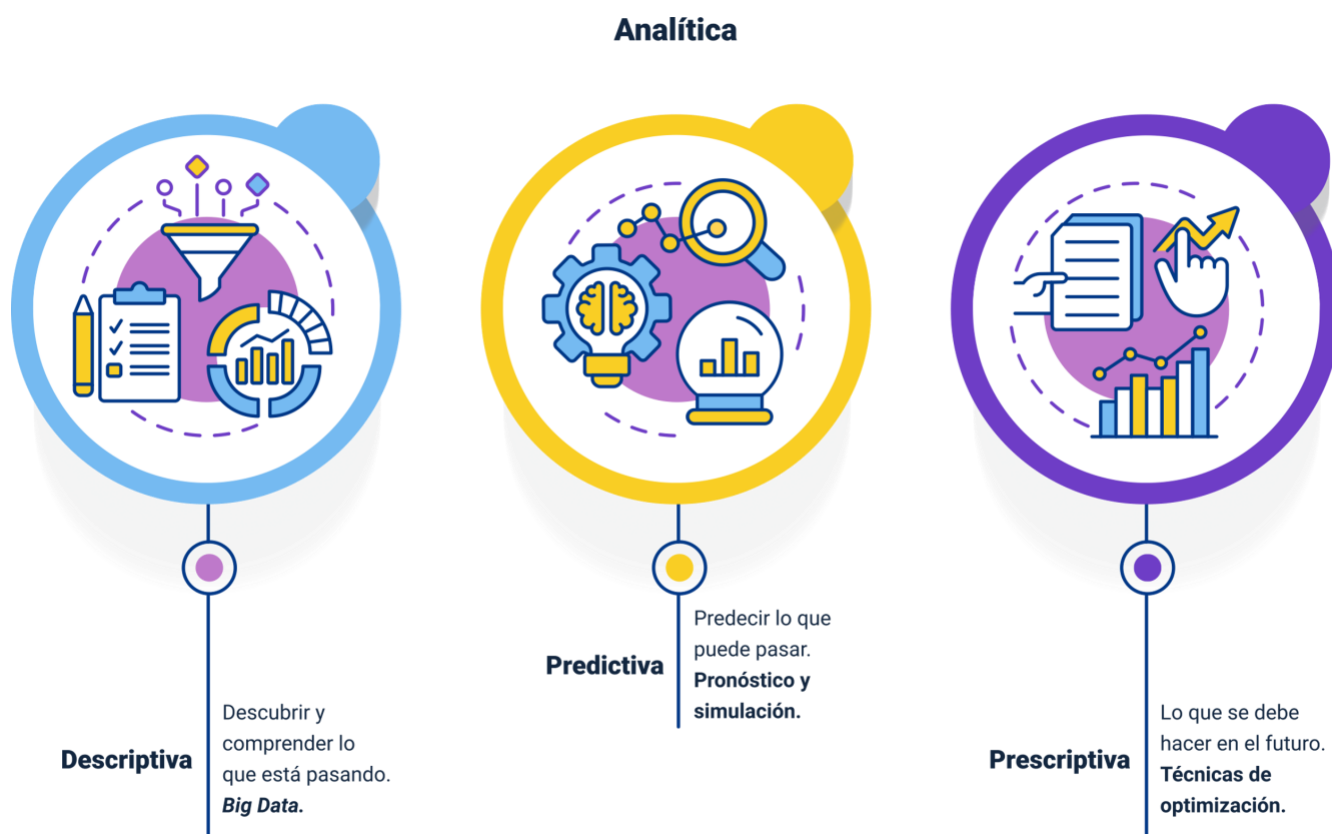
La imagen presenta una línea recta con dos flechas que indican direcciones opuestas: una flecha hacia abajo, acompañada de la inscripción "análisis cualitativo", y otra hacia arriba, que dice "análisis cuantitativo". Esta representación simboliza la complementariedad entre ambos enfoques en la investigación; el análisis cualitativo se centra en la comprensión profunda de fenómenos a través de métodos descriptivos, mientras que el análisis cuantitativo se enfoca en la medición y el análisis de datos numéricos. Juntos, estos enfoques ofrecen una visión más completa y enriquecedora de la realidad estudiada.

De acuerdo con el Blog Questionpro (2021): también se recurre a otros tipos de análisis entre los cuales está el análisis descriptivo, que como su nombre lo indica, consiste en describir las tendencias claves en los datos existentes y observar las situaciones que conduzcan a nuevos hechos, está basada en la estadística descriptiva que conlleva a las medidas de tendencia central (media, mediana y moda) esta información es cotejada con las medidas de dispersión (desviación estándar y error estándar) para su validación.

Por su parte, el análisis predictivo se conceptúa por parte del mismo blog Questionpro (2021) como una forma de análisis estadístico que se encomienda de conseguir información nueva o histórica y utilizarla para pronosticar patrones de conducta.

Por último, está el análisis prescriptivo se encarga de encontrar una solución entre una gama de variantes con el objetivo de optimizar los recursos y aumentar la eficiencia operativa, es decir, ayuda a entender qué tenemos que hacer para obtener los resultados que queramos en el futuro.

Figura 3. Análisis descriptivo, predictivo y prescriptivo



Nota: Adaptada de David Higueta (2.016)

La imagen presenta tres iconos que representan diferentes enfoques de análisis de datos, cada uno acompañado de una breve descripción. El primer icono, un bombillo, se asocia con el análisis descriptivo, que se enfoca en describir y comprender lo que está sucediendo en un contexto determinado, utilizando herramientas como "Big Data". El segundo icono, una lupa, simboliza el análisis predictivo, que tiene como objetivo anticipar lo que podría ocurrir en el futuro a través de pronósticos y simulaciones. Por último, el tercer icono, una gráfica de barras, representa el análisis prescriptivo, que se centra en determinar las acciones que se deben tomar en el futuro, empleando técnicas de optimización. Juntos, estos enfoques ofrecen un marco integral para la toma de decisiones informadas en diversos ámbitos.

4.3 Técnicas de análisis de datos

Las técnicas de análisis de datos se definen como el proceso de explorar los datos metódicamente con el objetivo de destacar aquella información que sea más útil, con el auge y la aplicación de las tecnologías basadas en las industrias 4.0 los análisis de datos se hacen cada vez más precisos, disminuyendo la incertidumbre y el riesgo de errores.

Minería de datos

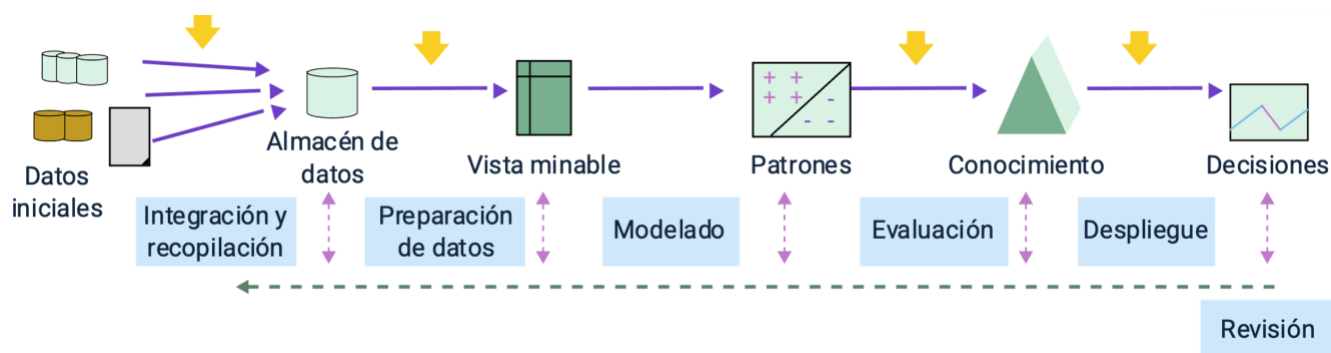
Minería de datos es definida como el proceso de encontrar singularidades, elementos comunes y correspondencias en grandes agregados de datos para augurar conclusiones, empleando una amplia variedad de técnicas. En conformidad con la página web de SAS la minería de datos le permite:

- Filtrar todo el ruido caótico y repetitivo en sus datos. (datos que se repiten)
- Entender qué es relevante y luego hacer un buen uso de esa información para evaluar resultados probables. (datos significativos).

- Acelerar el ritmo de la toma de decisiones informadas.

La minería de datos tiene seis fases: comprensión del negocio, comprensión de los datos, preparación de datos, modelado, evaluación y despliegue.

Figura 4. Ámbito, Objetivos de negocio



Nota. Adaptada de conceptos (2018)

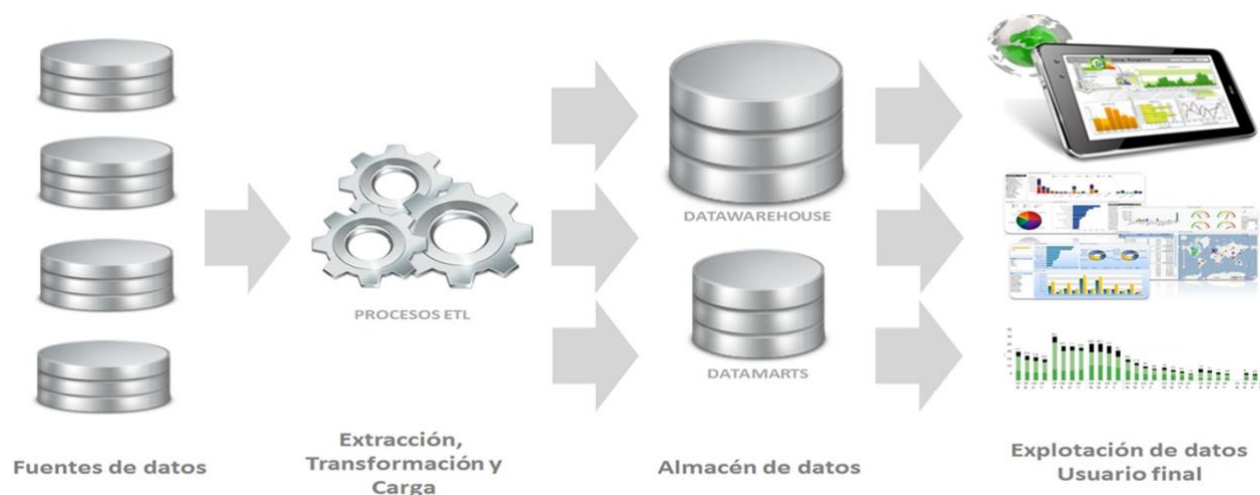
La imagen ilustra un diagrama de flujo que utiliza figuras geométricas para representar el proceso del análisis de datos. Este proceso comienza con los "datos iniciales" y avanza a través de etapas como "almacén de datos", "vista minable", "patrones", "conocimiento" y, finalmente, "decisiones". Además, en un flujo opuesto, se presentan conceptos complementarios que incluyen "revisión", "despliegue", "evaluación", "modelado", "preparación de datos", "integración" y "recopilación". Esta disposición resalta la naturaleza cíclica y dinámica del análisis de datos, donde cada etapa es fundamental para transformar la información bruta en conocimientos útiles que guían la toma de decisiones.

Business intelligence

Se conceptúa al conjunto de metodologías, aplicaciones, prácticas y capacidades enfocadas a la creación y administración de información que permite tomar mejores

decisiones a los usuarios de una organización (Conesa y Curto, 2010, p18). Estas acciones pueden ser procesos, técnicas, metodologías, herramientas y tecnologías que generan información para que sirva de base para la toma de decisiones. En otras palabras la inteligencia de negocios o business intelligence (BI), mezcla lo más profundo del análisis de negocios con el trabajo del big data y la minería de datos, para obtener la información más expedita que permita la mejor decisión posible en los escenarios posibles, minimizando siempre de esta forma los riesgos posibles.

Figura 5. Proceso business intelligence



El diagrama ilustra un flujo que abarca varios conceptos clave en el proceso de manejo de datos. Comienza con las "fuentes de datos", que son el punto de partida para la recopilación de información. A continuación, se lleva a cabo la "extracción", seguida de la "transformación" y "carga" de los datos, etapas esenciales para preparar la información para su análisis. Luego, los datos se almacenan en un "almacén de datos", donde se organizan y gestionan de manera eficiente. Finalmente, se llega a la "explotación de datos", que permite a los usuarios finales acceder y utilizar la

información para tomar decisiones informadas. Este flujo destaca la importancia de cada etapa en la conversión de datos en conocimientos valiosos.

Aprendizaje automático

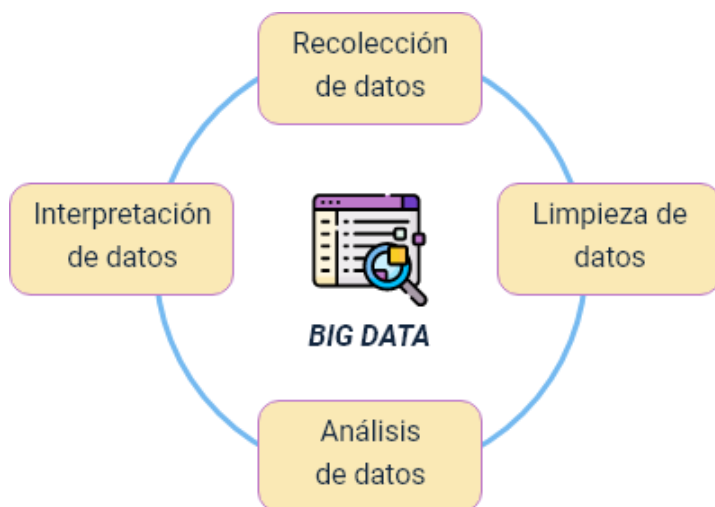
Es la parte de la inteligencia artificial (IA) que se concentra en desarrollar sistemas que aprenden, u optimizan el rendimiento, en dependencia de los datos que ejecuten. Inteligencia artificial es un término amplio que se refiere a sistemas o máquinas que imitan la inteligencia humana, se puede decir que, el Machine Learning (como también es conocido el aprendizaje automático) es una disciplina del ámbito de la Inteligencia Artificial que, a través de algoritmos o guarismos matemáticos, otorga a los computadores de la capacidad de asemejar estándares en datos masivos y elaborar predicciones (para hacer análisis predictivos).

Análisis del Big Data

El análisis de ‘grandes datos’ es el proceso de inspeccionar grandes cantidades de datos de una diversidad de tipos (big data) para descubrir patrones escondidos u ocultos, correspondencias inéditas u otra información útil, para generar información que posteriormente se convierta en conocimiento y aprendizaje en la toma asertiva de decisiones. Según Rodal Montero (2020) define la “analítica de big data como la recolección, gestión y análisis a alta velocidad de grandes, dinámicos y heterogéneos volúmenes de datos generados por usuarios y máquinas, que debido a su tamaño y complejidad superan las capacidades de procesamiento de las herramientas de software tradicionales, por lo que requieren técnicas innovadoras para su procesamiento y tratamiento”. (p.65).

Las fases del proceso de la analítica de grandes datos es la recolección de grandes magnitudes de datos, posteriormente existe todo un proceso de limpieza de esos datos, evitando los valores que constituyan ruido.

Figura 6. Big Data



Nota. Adaptada de Raul Pulido (2.018)

La imagen presenta las palabras "BIG DATA" en el centro de un círculo, rodeadas de conceptos clave que destacan las etapas fundamentales del manejo de grandes volúmenes de información. Alrededor de "BIG DATA", se mencionan la "recolección de datos", que se refiere al proceso de reunir información de diversas fuentes; la "limpieza de datos", que implica la depuración y corrección de errores en los datos recopilados; el "análisis de datos", donde se examinan y procesan los datos para extraer patrones y tendencias; y la "interpretación de datos", que consiste en dar sentido a los resultados obtenidos y traducirlos en información útil. Esta disposición resalta la interconexión de estas etapas en el ciclo de vida del Big Data, enfatizando su importancia para la toma de decisiones informadas.

Para Big Data Internacional Campus (2020) Data mining es el proceso de identificación de toda la información que es relevante y es extraída de grandes cantidades de datos. El objetivo de esta extracción es descubrir patrones y tendencias estructurando la información que se ha obtenido de una manera que sea comprensible para su utilización. El data mining reúne las ventajas de varias áreas como la inteligencia artificial, la estadística, las bases de datos, la computación gráfica y el procesamiento masivo.

La principal diferencia entre big data y data mining (grandes datos y minería de datos) se halla en que data mining se refiere a operaciones que implican búsquedas sofisticadas para la obtención de datos específicos y concretos dentro del gran mundo de los datos (big data) Ejemplificando la data mining es observar, analizar y estudiar los datos con microscopios, mientras que el big data utilizará telescopios para observar, analizar y estudiar los datos.

5. Técnicas de evaluación

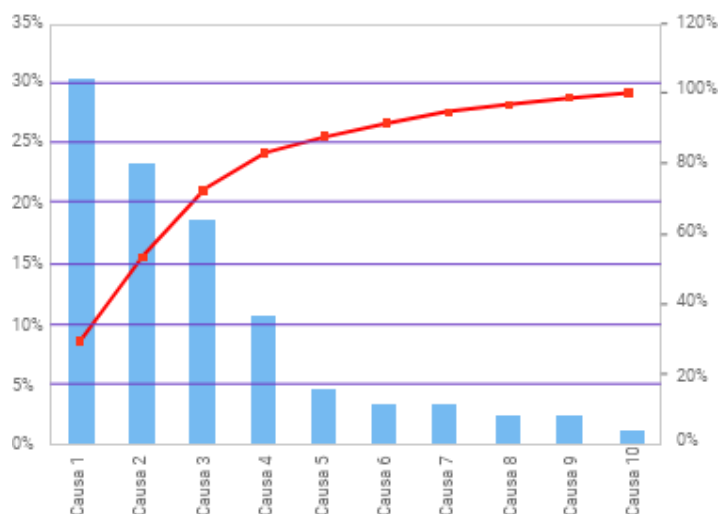
Bernal (2000) describe las siguientes técnicas de evaluación de la información que pueden efectuarse:

Análisis de Pareto

Es una técnica para analizar orígenes de problemas y las procedencias de sus causas, de igual forma, es una técnica que permite catalogar gráficamente la información de mayor a menor preeminencia, con el fundamento de identificar los problemas más trascendentales en los que se debe enfocar y solucionar, tiene como procedimiento generalizado:

- Identifica el problema que deseas analizar, es prioritario identificar el problema que está afectando el hecho, evento, situación a investigar.
- Agrupar los datos que ayudan a evaluar el problema, identificando en las categorías o problemas (eje de las y) y la frecuencia de ocurrencia de esos problemas (eje de las x).
- Ordenar los datos en una tabla de forma decreciente de la frecuencia (de mayor a menor), luego calcular sus porcentajes y acumulados.
- Construir una gráfica de barras con estos datos.

Figura 7. Análisis de Pareto



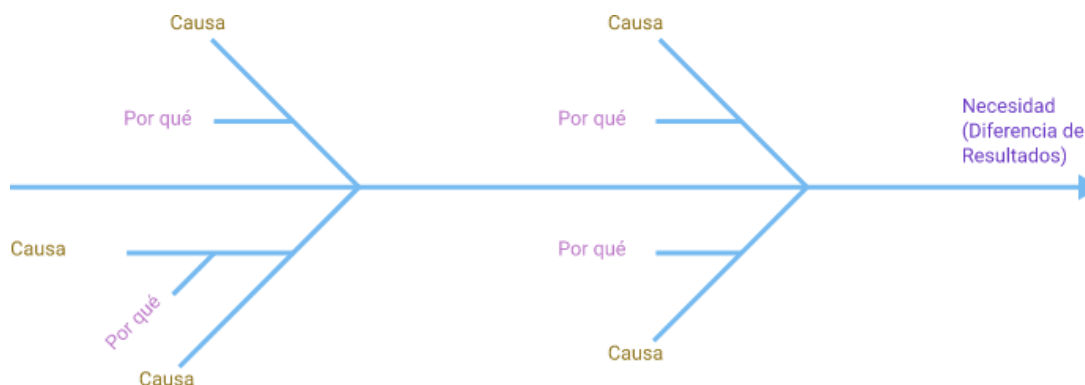
Nota. Adaptada de Jose Marcondes (2020)

La figura ilustra un gráfico que consta de varias barras azules, acompañadas por una línea roja que conecta puntos rojos. Las barras azules representan diferentes valores o categorías, proporcionando una comparación visual clara entre ellos. La línea roja, al unir los puntos rojos, destaca una tendencia o evolución a lo largo del tiempo, permitiendo observar patrones o cambios significativos en los datos. Esta combinación de elementos visuales facilita la interpretación de la información y ayuda a identificar relaciones entre las variables presentadas.

Diagrama de causa efecto

Bernal, (2003): (espina de pescado) es una gráfica mediante la cual los miembros de un equipo representan, categorizan y evalúan todos los posibles motivos de un resultado o reacción; por lo general, se expresa como un problema para resolver. Se le conoce como diagrama de Ishikaw. (p.180).

Figura 8. Diagrama de causa y efecto



Nota. Adaptada de IBM (2007)

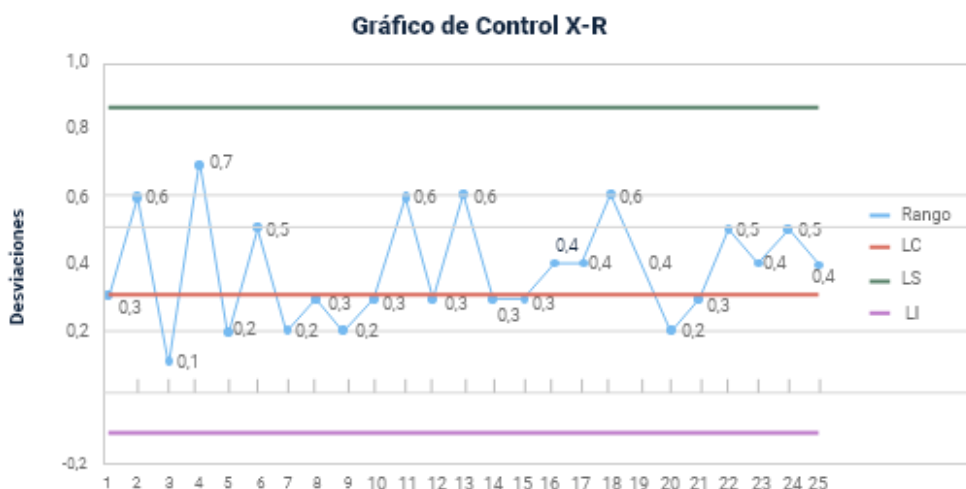
La figura ilustra un gráfico en forma de espina de pescado, también conocido como diagrama de Ishikawa. La estructura del gráfico permite visualizar de manera clara y sistemática las relaciones entre las causas y el efecto, facilitando así la identificación de áreas de mejora.

Gráficas de control

Se recurren a ellas cuando existe la necesidad de control de calidad de procesos. Según Levin y Rubin (1996) “estas graficas reciben el nombre de diagramas de control y son de varios tipos :

- Diagramas o diagramas de control para medias de procesos.
- Diagramas R o diagramas de control para variabilidad en los procesos.
- Diagramas o diagramas de control para atributos, (p.528).

Figura 9. Gráfico de control X - R.



Nota. Adaptada de Grupo 10 (2015)

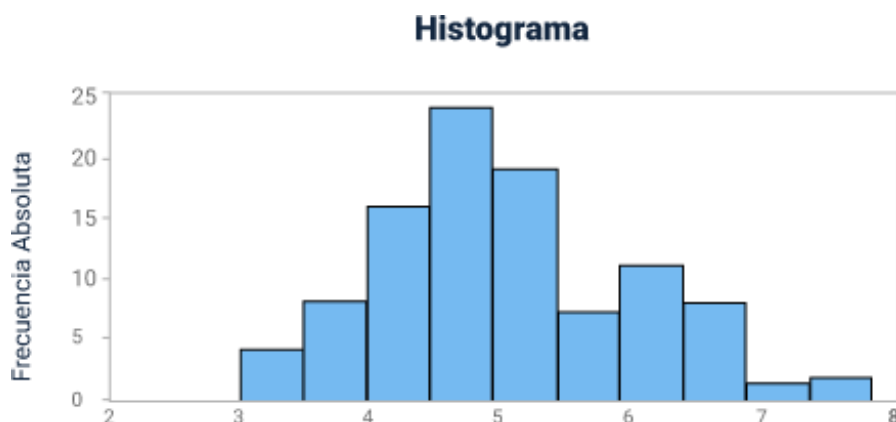
La gráfica muestra una serie de puntos dispuestos en un patrón de zigzag irregular, conectados por una línea azul. Esta disposición resalta las fluctuaciones y variaciones en los datos a lo largo de un intervalo determinado. La línea azul que une los puntos permite visualizar claramente las tendencias y cambios en los valores, facilitando la identificación de patrones o comportamientos específicos en la información presentada. Este tipo de representación gráfica es útil para analizar datos que presentan movimientos no lineales, proporcionando una comprensión más profunda de la dinámica subyacente.

Distribución de frecuencias y representaciones gráficas

De acuerdo con Mason y Lind (1994) es el conjunto de datos en categorías que muestran el número de observaciones de cada categoría, en otras palabras, una distribución de frecuencias indica el número de veces que ocurre cada valor o dato en una tabla de resultados. (p.24).

- Histogramas son medios gráficos para la representación de la distribución de frecuencias.

Figura 10. Histograma



Nota. Adaptada de Marta (2019)

La gráfica exhibe una serie de barras en color azul, que representan diferentes valores o categorías. Este diseño permite realizar comparaciones visuales entre las distintas barras, facilitando la identificación de tendencias y patrones en los datos. El uso del color azul aporta claridad y uniformidad a la presentación, haciendo que la información sea fácilmente interpretable. Este tipo de representación es especialmente útil para resaltar diferencias y similitudes entre los elementos analizados, proporcionando una visión general efectiva de la información.

- Polígonos de frecuencia son representaciones gráficas que reflejan una imagen veloz de los más importantes rasgos de los datos de una distribución de frecuencias.

Figura 11. Polígonos de frecuencia



Nota. Adaptada de Khan Academy (2018)

La gráfica muestra una serie de puntos conectados por una línea roja que crea una forma similar a una montaña. Este diseño resalta las variaciones en los datos, permitiendo observar claramente los picos y valles a lo largo del recorrido. La línea roja, al seguir el contorno de los puntos, facilita la identificación de tendencias y cambios significativos en la información presentada. Este tipo de representación es útil para analizar datos que exhiben fluctuaciones, proporcionando una visualización clara de la dinámica y el comportamiento de las variables a lo largo del tiempo.

Medidas de tendencia central

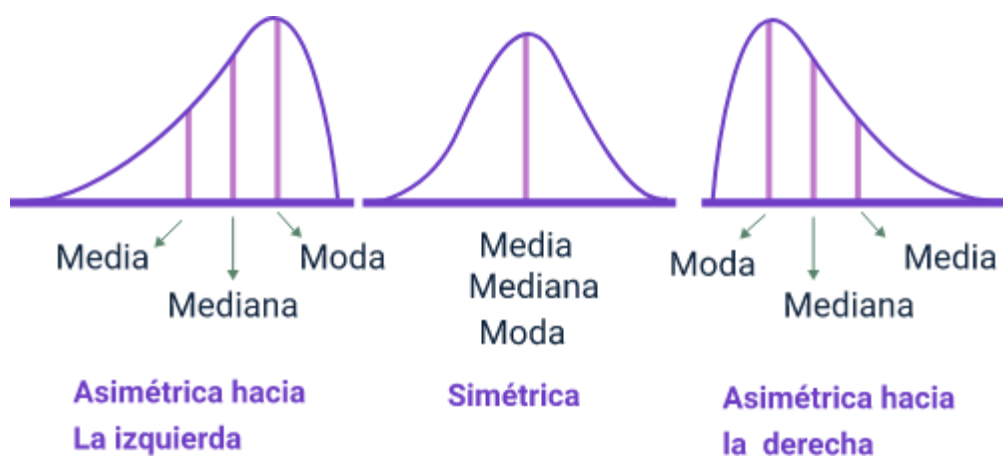
Son medidas estadísticas que pretenden resumir en un solo valor a un conjunto de valores. Se reconocen de manera universal las siguientes:

- La media o promedio aritmético es la sumatoria de un conjunto de frecuencias divididos por la cantidad total de los mismos.

- La moda se identifica con el valor con la frecuencia que más se repite en una tabla de distribución de datos.
- La mediana es el valor que resulta una vez organizado la distribución de frecuencias de manera ascendente o descendente y se encuentre justo en la mitad de los datos.

El concepto de simetría que se aprecia en la figura 12, no es más que la distancia en la cual se encuentran la moda y la mediana con respecto a la media aritmética, mientras más cerca se encuentren mayor confiabilidad tiene el análisis de la población, mientras más lejos se encuentre menor confiabilidad.

Figura 12. Tendencia central



Nota. Adaptada de estudiar física.wordpress (2009)

La figura muestra tres gráficas con forma de montaña: una asimétrica hacia la izquierda, otra simétrica y una más asimétrica hacia la derecha. En cada una de estas gráficas se destacan las medidas de tendencia central, incluyendo la media, la mediana y la moda. Esta representación permite comparar visualmente cómo se distribuyen los datos en cada caso, evidenciando las diferencias en la forma y la dispersión de las

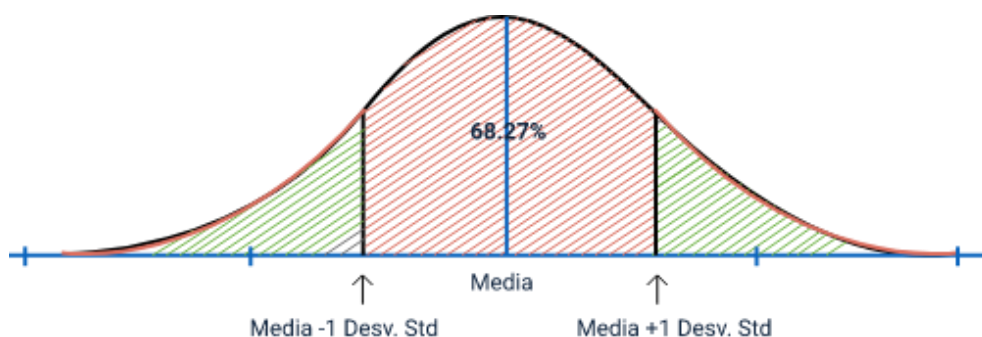
distribuciones. Al señalar estas medidas, se facilita la comprensión de la posición relativa de los valores en cada gráfica, lo que resulta útil para el análisis estadístico y la interpretación de los datos.

Medidas de dispersión

Hablan cómo se dividen los datos en torno a una de las medidas de tendencia central conceptuadas anteriormente, y son un valor añadido a estas últimas para representar más fidedignamente un conjunto de datos. Dentro de las medidas de dispersión más reconocidas están:

- La varianza es la suma de las desviaciones de la media elevada al cuadrado, dividida entre el número de observaciones menos uno, en otras maneras de explicar, es una medida de dispersión que significa la movilidad de una serie de datos en relación a su promedio o media aritmética.
- La desviación estándar es la cuantía media en que cada una de las frecuencias individuales varía respecto a la media aritmética del conjunto de datos; dicho de otras maneras, cuanto mayor sea la dispersión mayor es la desviación estándar.

Figura 13. Medidas de dispersión



Nota. Adaptada de matemáticas con tecnología

La figura muestra una gráfica con forma de montaña en la que se indican tres puntos clave: la media - 1, la media, y la media + 1. Esta representación permite visualizar cómo se distribuyen los datos en relación con la media, destacando las variaciones que se encuentran por debajo y por encima de este valor central. Al marcar estos puntos, se facilita la comprensión de la dispersión de los datos y se proporciona un contexto adicional para el análisis estadístico, ayudando a identificar patrones y tendencias en la información presentada.

Pruebas estadísticas

Son diversas las pruebas estadísticas que pueden aplicarse al conjunto de datos, a continuación se relacionan algunas de las más utilizadas.

- Prueba t de student es un medidor estadístico para poblaciones menores o iguales a 30.
- Prueba z es una prueba para distribuciones normales, que muestra la posibilidad que un data aparezca reflejado en la distribución de frecuencias.
- Análisis de varianza es una test estadístico para examinar si dos o más grupos de datos se diferencian entre sí, analizando sus medidas de tendencia central y sus varianzas.
- Análisis de covarianza según Bernal (2006) es una prueba que se utiliza para analizar y determinar la relación entre una variable dependiente y dos o más variables independientes. (p.181).
- Chi cuadrado es una prueba estadística que permite demostrar si dos fracciones de población pueden considerarse iguales, es decir, si no presentan diferencias significativas.

En la actualidad todos los ordenadores facilitan la verificación de todas estas pruebas estadísticas de la información, a través de Excel, o de programas especiales como stapgraphic o SSPS, entre otros validados por la comunidad científica.

6. Informes y documentos

Los informes y documentos son parte fundamental del proceso que se llevará a cabo.

6.1 Informes

Es un escrito que tiene como resultado anunciar y dar balance de una situación desde diferentes perspectivas, es decir, de algo sobre lo que se está efectuando una investigación exhaustiva; entre las principales características que debe tener el informe, la primera es la claridad en el lenguaje para que tenga comprensión por parte del lector, la segunda característica del informe escrito es la objetividad en reflejar los hechos basados en información comprobable, la tercera característica es la estructuración respondiendo a criterios narrativos y presentar una estructura mínima en cuanto introducción, desarrollo y conclusiones.

6.2 Tipos de informes

En la contemporaneidad en la cual recurre el mundo moderno, para cualquier administrativo o comercial de una empresa debe ser importante saber dirigir, saber presidir una reunión, dictar conferencias, también lo es escribir informes, luego entonces, es imprescindible desarrollar estas habilidades de comunicación escrita en este tipo de personal.

Por tanto, la calidad de la información que se transmite en los informes debe ser tal que facilite la comprensión del texto, que transmita el mensaje apropiado y que facilite el flujo de trabajo en los procesos de la empresa.

Según el contenido

Se pueden encontrar 3 grandes tipos de informes de acuerdo con la especificidad del contenido del informe

- **Informes técnicos:** son informes basados en exploraciones sobre fenómenos sociales, psicológicos o económicos adaptados para las organizaciones que se nutren de esos datos.
- **Informes científicos:** se define como un informe escrito y publicado que describe resultados originales de una investigación. (Lam, 2016).
- **Informes de divulgación:** es un escrito generalmente breve, que puede explicar hechos, ideas, conceptos, descubrimientos vinculados al quehacer científico y tecnológico.

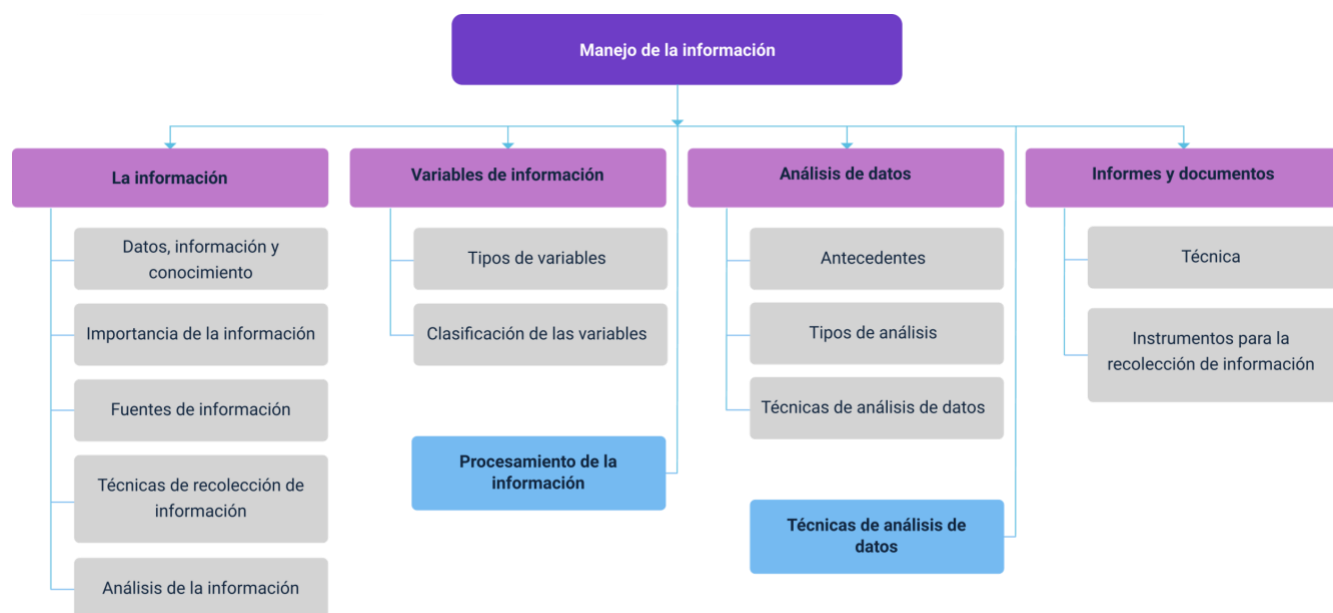
Según la estructura

Se destacan con preeminencia 3 tipos de informes

- **Informes persuasivos:** tienen como finalidad convencer a los lectores de las conclusiones e ideas desarrolladas en el proceso de trabajo.
- **Informes expositivos:** es una cadena de hechos relatada sin análisis e interpretaciones del autor, debe contener referencias que ayuden a comprender lo que se narra.
- **Informe analítico:** es un tipo de informe o evaluación de un conjunto particular de circunstancias relacionadas con el efecto que producen, por lo tanto, determinan y desarrollan las causas de un hecho o fenómeno.

Síntesis

El manejo de la información es un tema fundamental para un asesor comercial ya que dentro de sus competencias debe dominar temas como el procesamiento de la información, el análisis de datos y la realización de informes, en el siguiente diagrama se relacionan los conceptos y temas más significativos estudiados en este componente:



Material complementario

Tema	Referencia	Tipo de material	Enlace del recurso
1.1 La información	Pérez, J. (2008) concepto de información. https://definicion.de/informacion/	Página web	https://definicion.de/informacion/
3. Variables de la información	Parra, A. (2023, 7 agosto). Cuáles son los tipos de variables en una investigación. QuestionPro. https://www.questionpro.com/blog/es/tipos-de-variables-en-una-investigacion/	Página web	https://www.questionpro.com/blog/es/tipos-de-variables-en-una-investigacion/
4. Procesamiento de la información	Espinoza Freire, E. E. (2018). Las variables y su operacionalización en la investigación educativa. Parte I.	Página web	http://scielo.sld.cu/scielo.php?pid=S1990-86442018000500039&script=sci_arttext&tlng=en
5. Análisis de datos	Universidad del sur.(2.018), Análisis de datos	Página web	https://www.mundoposgrado.com/que-es-el-analisis-de-datos/
6. Informes	Equipo editorial etece. (2021) Concepto, informes	Página web	https://concepto.de/informe/

Glosario

Base de datos: las bases de datos son una herramienta fundamental en el entorno informático hoy en día y tienen aplicación en la práctica totalidad de campos como los negocios, organizaciones y en la vida personal. Su aplicabilidad responde a la necesidad de gestionar datos que provean información pertinente.

Conocimiento: es la explicación e interpretación de la información producto de los procesos cognitivos de comparar, analizar, combinar o experimentar, permitiendo al sujeto la argumentación y la aplicación de la información.

Datos: son la diferenciación simbólica de una variable que puede ser clasificada como cualitativa o cuantitativa, indicando un valor representativo a dicha variable para su análisis, estudio y clasificación.

Encuesta: las encuestas son un método de investigación y recopilación de datos utilizadas para obtener información a un grupo de individuos sobre diversos temas.

Entrevista: la entrevista es una charla en la que una persona hace a otra una serie de preguntas sobre un tema determinado.

Información: es la denominación que se le da a la agrupación organizada de datos significativos que describen sucesos, eventos y hechos, con cuyo procesamiento se puede construir un mensaje descriptivo y / o analítico.

Información cualitativa: es un método de estudio y análisis que se propone evaluar e interpretar información obtenida a través de recursos como entrevistas, conversaciones, memorias y registros documentales.

Información cuantitativa: método estructurado de recopilación y análisis de información que se obtiene a través de diversas fuentes. Este proceso se lleva a cabo con el uso de herramientas estadísticas y matemáticas con el propósito de cuantificar el problema de investigación.

Observación: es una técnica que consiste en observar atentamente el fenómeno, hecho o caso, tomar información y registrarla para su posterior análisis.

Referencias bibliográficas

Análisis de big data—ArcGIS Velocity | Documentación. (s. f.).

<https://doc.arcgis.com/es/velocity/analyze/perform-big-data-analysis.htm>

Bernal T cesar Augusto. (2.003) Metodología de la investigación aplicada para administración y economía. Editorial Prentice Hill.

Biblioguías: Gestión del Conocimiento (GDC): 1.3 Datos, Información y Conocimiento. (2017). <https://biblioguías.cepal.org/c.php?g=738015&p=5789030>

Big Data Internacional Campus (2.020). Data mining vs Big data.

<https://www.campusbigdata.com/blog/data-mining-vs-big-data/>

Conceptos. (2018). Minería de datos.

<https://mineriadedatos1.blogspot.com/2018/10/conceptos.html>

Conesa Caralt Jordi, Curto Diaz Josep. (2010). Introducción al Bussines Intelligence. Editorial UOC.

Dertiano Victor. (2014). ¿Qué es Business Intelligence? Beneficios del BI.

<https://bigeek.medium.com/qué-es-business-intelligence-509d6a12fdf3>

Espinoza Freire, Eudaldo. (2018). Las variables y su operacionalización en la investigación educativa. Scielo.

http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S1990-86442018000500039

Fernández Jorge. (2016). Investigación Cualitativa Vs Investigación Cuantitativa. Investigación y docencia.

<https://www.investigacion360.com/2017/02/investigacion-cualitativa-vs-investigacion-cuantitativa.html>

Florencia. (2013). Importancia de la información. <https://significado.com/im-informacion/>

Grupo 10. (2015). Tipos de gráficas de control.

<http://grupo10direccion.blogspot.com/2015/06/tipos-de-graficas-de-control-1.html>

Hernández Macario. (2008). El control estadístico del proceso.

<https://optyestadistica.wordpress.com/2008/06/11/el-control-estadistico-de-proceso-3-de-3/>

Hernández Noelia. (2019). Esto es lo que puede hacer la Inteligencia Artificial por la educación. Educación 3.0.

<https://www.educaciontrespuntocero.com/tecnologia/poencial-ia-en-educacion/>

Hidalgo Nuchera Antonio, León Serrano Gonzalo, Pavon Morote Julián. (2013). la gestión de la innovación y la tecnología en las organizaciones. Editorial Pirámide.

Higueta David. (2016). Analítica y productividad aplicada al ruteo de vehículos.

<https://www.bancolombia.com/empresas/capital-inteligente/tendencias/innovacion/analitica-y-productividad>

IBM. (2007). Análisis de causa raíz: el diagrama de espina de pescado.

<https://managing-ils-reporting.itcilo.org/es/herramientas/root-cause-analysis-the-fishbone-diagramme/>

Khan academy. (2018) Polígonos de frecuencia. <https://es.khanacademy.org/math/2-secundaria-pe/xf4e5558599a475b6:probabilidad-y-estadistica-2sec/xf4e5558599a475b6:preguntas-estadisticas-representacion-de-datos-a-traves-de-histogramas-y-poligonos-de-frecuencia/a/81817-articulo-polgonos-de-frecuencia>

Lam Diaz Rosa. (2016). La redacción de un artículo científico.
http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S0864-02892016000100006

Lind Douglas A, Marchal William E, Wathen Samuel A. (2.008). Estadística aplicada a los negocios y la economía. MC Graw Hill Interamericana.

Marcondes José. (2020). Diagrama ou grafico de Pareto.
<https://gestaodesegurancaprivada.com.br/diagrama-ou-grafico-de-pareto-conceito/>

Marta. (2019). ¿Qué es un Histograma? Super prof.
<https://www.superprof.es/apuntes/escolar/matematicas/estadistica/descriptiva/histograma.html>

Quality Leadership University. (2020). Ejemplos de antecedentes de un anteproyecto.
<https://qlu.ac.pa/ejemplos-antecedentes-anteproyecto/>

Questionpro. (2021). ¿Qué es el análisis descriptivo?
<https://www.questionpro.com/blog/es/analisis-descriptivo/>

Rodal Montero Enrique, 2.020, Industrias 4.0. Editorial Pirámide.

Samperi Hernández Roberto, 2.006, Metodología de la investigación. Editorial Mc Graw Hill.

SAS. (2021). minería de datos, ¿Qué es y porqué es importante? https://ecoredsena-santander.github.io/631101_CF3_ASESORIA_COMERCIAL/#/referencias

Suarez R. Pedro Alejandro, 2.001, Metodología de la investigación diseños y técnicas. Orión editores.

Temas de enfermería. (2012). La media aritmética en medidas de tendencia central. <https://saludyenfermeria.org/1952-2/>

Créditos

Nombre	Cargo	Centro de Formación y Regional
Milady Tatiana Villamil Castellanos	Responsable del ecosistema	Dirección General
Claudia Johanna Gómez Pérez	Responsable de línea de producción	Centro Agroturístico - Regional Santander
Laureano E. Ruidiaz Polo	Instructor grado 16	Centro de Comercio y Servicios - Regional Bolívar
Zenith Chinchilla Ruedas	Diseñadora instruccional	Centro de Diseño y Metrología - Regional Distrito Capital
Carolina Coca Salazar	Revisora metodológica y pedagógica	Centro de Diseño y Metrología - Regional Distrito Capital
Sandra Patricia Hoyos Sepúlveda	Revisión y corrección de estilo	Centro para la Industria de la Comunicación Gráfica - Regional Distrito Capital
Carolina Jiménez Suescún	Evaluadora instruccional	Centro Agroturístico - Regional Santander
Yazmin Rocio Figueroa Pacheco	Diseñadora de contenidos digitales	Centro Agroturístico - Regional Santander
Leonardo Castellanos Rodríguez	Desarrollador full stack	Centro Agroturístico - Regional Santander
María Alejandra Vera Briceño	Animadora y productora multimedia	Centro Agroturístico - Regional Santander
Yineth Ibette Gonzalez Quintero	Validadora de recursos educativos digitales	Centro Agroturístico - Regional Santander
Andrea Ardila Chaparro	Evaluadora para contenidos inclusivos y accesibles	Centro Agroturístico - Regional Santander

