

INTRODUCCIÓN A DATA ANALYTICS



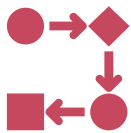
Data Analytics

A magnifying glass is positioned over a bar chart. The chart has a light blue background with white grid lines. The x-axis is labeled with 'Q1', 'Q2', 'Q3', and 'Q4'. The y-axis has a label '1,00'. There are two data series: one represented by blue bars and another by green bars. The blue bars are consistently taller than the green bars. The magnifying glass is focused on the 'Q2' and 'Q3' bars. Overlaid on the left side of the image is a black box containing the text 'DEFINICIÓN DE DATA ANALYTICS' in white, bold, sans-serif font.

DEFINICIÓN DE DATA ANALYTICS

- Es el proceso sistemático de examinar datos crudos para identificar patrones, tendencias, correlaciones y obtener conclusiones útiles. A través de este proceso, convertimos grandes volúmenes de datos en información accionable, lo que permite una mejor toma de decisiones basada en hechos en lugar de suposiciones

CICLO DEL ANÁLISIS DE DATOS



1. **Recolección de Datos:** Se recopilan datos de diversas fuentes. Por ejemplo, datos de ventas de una tienda, comentarios de usuarios en redes sociales, información de sensores en una fábrica.



2. **Limpieza de Datos:** Esto implica eliminar duplicados, corregir errores, manejar valores faltantes y asegurar la consistencia. Por ejemplo, eliminación de datos incompletos o erróneos en una base de clientes.



3. **Transformación y Análisis:** Consiste en transformar los datos para hacerlos más utilizables y aplicar técnicas estadísticas para descubrir patrones e *insights*. Por ejemplo, Agrupar ventas por región y producto.



4. **Visualización:** Los resultados deben presentarse de manera clara y comprensible mediante gráficos o *dashboards*. Por ejemplo, un *dashboard* que muestre las métricas clave del negocio en tiempo real.



5. **Toma de Decisiones:** Los tomadores de decisiones pueden basar sus estrategias en datos sólidos para mejorar resultados y ajustar acciones

IMPORTANCIA DEL ANÁLISIS DE DATOS

El análisis de datos es más que solo números y estadísticas; es una herramienta crítica para mejorar el rendimiento y la competitividad de las empresas. Algunas de las formas en las que impactamos son:

- **Optimización Operacional:** A través del análisis de datos, las empresas pueden identificar ineficiencias en sus procesos y optimizar el uso de recursos.
- **Toma de Decisiones Informada:** Data Analytics permite a las empresas tomar decisiones basadas en hechos y cifras, minimizando el riesgo de decisiones incorrectas.
- **Personalización y Mejora de la Experiencia del Cliente:** Las empresas pueden personalizar productos y servicios para ajustarse a las necesidades individuales.
- **Ventaja Competitiva:** Las empresas que dominan el análisis de datos pueden anticipar tendencias, lo que les permite reaccionar más rápido y captar oportunidades antes que sus competidores.

TIPOS DE ANÁLISIS DE DATOS



DESCRIPTIVO: QUÉ OCURRIÓ
(ANÁLISIS DE TENDENCIAS Y
RESÚMENES HISTÓRICOS).



DIAGNÓSTICO: POR QUÉ
OCURRIÓ (IDENTIFICACIÓN DE
CAUSAS Y PATRONES).



PREDICTIVO: QUÉ PODRÍA
OCURRIR (MODELOS DE
PREDICCIÓN Y MACHINE
LEARNING).



PRESCRIPTIVO: QUÉ DEBERÍA
OCURRIR (RECOMENDACIONES
BASADAS EN DATOS).

ANÁLISIS DESCRIPTIVO



Este análisis se utiliza para resumir datos históricos, mostrando tendencias y patrones. Ayuda a organizar grandes cantidades de datos en resúmenes que son fáciles de entender y comunicar. La pregunta clave que nos debemos hacer es, ¿Qué ocurrió?



Proporciona una imagen clara de lo que ha pasado hasta ahora, sin explicar el porqué de los eventos.



Las herramientas que usaremos para esto son: gráficas de barras, tablas de frecuencia, medidas de tendencia central (media, mediana, moda), histogramas.

ANÁLISIS DIAGNÓSTICO



Este tipo de análisis se enfoca en entender las causas detrás de los eventos pasados. Va más allá del análisis descriptivo y profundiza en la identificación de patrones y correlaciones. La pregunta que nos deberemos hacer es, ¿Por qué ocurrió?



Descubre las razones y factores subyacentes que llevaron a ciertos resultados. Responde preguntas como: ¿Qué variables están relacionadas? ¿Qué eventos causaron el resultado observado?



Las herramientas que usaremos son el análisis de correlación, análisis de regresión, etc.

ANÁLISIS PREDICTIVO



El análisis predictivo utiliza datos históricos y algoritmos avanzados (como machine learning) para hacer predicciones sobre futuros eventos o comportamientos. Se basa en patrones previamente identificados para proyectar resultados futuros. La pregunta que nos tenemos que hacer es, ¿Qué podría ocurrir?



Ayuda a prever lo que podría suceder en función de las tendencias y comportamientos pasados. Este análisis permite a las empresas prepararse para eventos futuros.



Las herramientas que usaremos en este tipo de análisis son, modelos de regresión, árboles de decisión, redes neuronales, algoritmos de machine learning como *random forest* o modelos de series temporales.

ANÁLISIS PRESCRIPTIVO



Este análisis va un paso más allá del predictivo, ya que no solo predice lo que podría suceder, sino que también recomienda acciones a tomar. Proporciona soluciones basadas en datos y modelos de optimización, ayudando a tomar decisiones más precisas. La pregunta que nos haremos será, ¿Qué debería ocurrir?

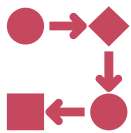


Su objetivo es ofrecer recomendaciones sobre la mejor acción a tomar para maximizar el éxito o minimizar riesgos. Su enfoque se basa en sugerir el curso de acción más adecuado en base a los datos disponibles.



Las herramientas que usaremos son, algoritmos de optimización, simulaciones, modelos avanzados de AI, análisis de escenarios.

CICLO DEL ANÁLISIS DE DATOS



1. **Recolección de Datos:** Se recopilan datos de diversas fuentes. Por ejemplo, datos de ventas de una tienda, comentarios de usuarios en redes sociales, información de sensores en una fábrica.



2. **Limpieza de Datos:** Esto implica eliminar duplicados, corregir errores, manejar valores faltantes y asegurar la consistencia. Por ejemplo, eliminación de datos incompletos o erróneos en una base de clientes.



3. **Transformación y Análisis:** Consiste en transformar los datos para hacerlos más utilizables y aplicar técnicas estadísticas para descubrir patrones e *insights*. Por ejemplo, Agrupar ventas por región y producto.



4. **Visualización:** Los resultados deben presentarse de manera clara y comprensible mediante gráficos o *dashboards*. Por ejemplo, un *dashboard* que muestre las métricas clave del negocio en tiempo real.



5. **Toma de Decisiones:** Los tomadores de decisiones pueden basar sus estrategias en datos sólidos para mejorar resultados y ajustar acciones

LIMPIEZA DE DATOS



Excel

PROCESO LIMPIEZA DE DATOS

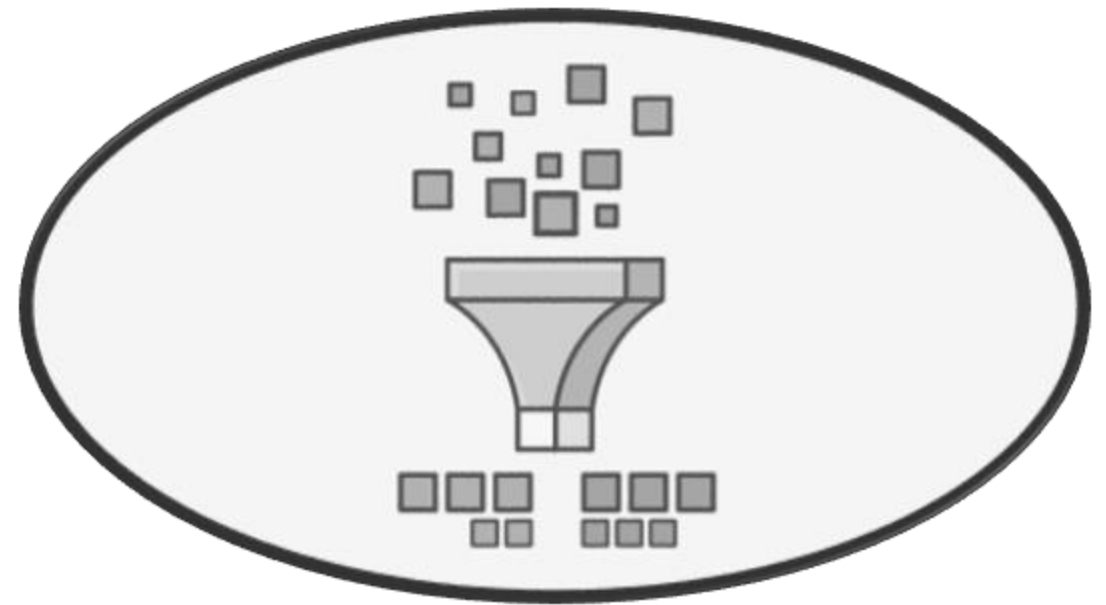
Eliminar datos irrelevantes

Establecimiento de tipos de datos

Normalizar datos

Manejo de duplicados

Tratamiento de valores nulos



No podemos limpiar lo que no conocemos.

¿QUÉ ES EL ANÁLISIS EXPLORATORIO DE DATOS?

El **Análisis Exploratorio de Datos (EDA)** es un proceso crucial en el análisis de datos cuyo objetivo principal es explorar, resumir y visualizar los datos para identificar patrones, detectar anomalías, probar hipótesis y entender las relaciones entre las variables.

El propósito principal es obtener una comprensión inicial de los datos, lo que permite **tomar decisiones informadas sobre la limpieza y preprocesamiento**. Es el momento para detectar problemas como datos faltantes, outliers, o errores de formato que podrían influir en la calidad del análisis.

FLUJO DE TRABAJO

EDA

Identificación de problemas (valores faltantes, outliers, inconsistencias)

Limpieza de Datos

Análisis Predictivo/Prescriptivo.

OBJETIVOS DEL EDA



1. Identificación de Patrones y Relaciones en los Datos.

Esto incluye la identificación de distribuciones comunes, tendencias, y correlaciones entre las variables. Por ejemplo, determinar si las ventas aumentan durante los meses de verano o si hay una correlación entre la edad y el ingreso en un análisis demográfico.



2. Detección de Outliers (Valores Atípicos):

Nos ayuda a identificar estos valores, que pueden ser errores o datos excepcionales que requieran un tratamiento especial. Por ejemplo, un precio de producto más alto que los demás o un cliente con un gasto mucho mayor al promedio.



3. Evaluar la Calidad de los Datos:

Evaluar la calidad de los datos, identificando problemas como valores faltantes, duplicados, o datos erróneos. Este paso es importante para planificar las acciones de limpieza y transformación de los datos.



4. Preparación de los Datos para Modelado Posterior:

Preparar los datos para el modelado y análisis avanzado, como machine learning o análisis predictivo. Esto incluye normalizar, transformar y dividir los datos en subconjuntos utilizables.



5. Formular Hipótesis:

Permite generar hipótesis basadas en los patrones y relaciones observados en los datos. Estas hipótesis pueden guiar futuros análisis o investigaciones más profundas. Por ejemplo, "los clientes con mayores ingresos tienen una mayor propensión a comprar productos premium".