

TECNICATURA SUPERIOR EN

Ciencia de Datos e Inteligencia Artificial

Módulo Análisis de datos

Proyecto Final

1. NOMBRE DEL PROYECTO: Análisis descriptivo y visualización de ventas de PepsiCo en Argentina.

Estudio basado en datos recolectados de distribuidores a nivel nacional durante un período determinado.

Integrantes:

- Eglimar Ramirez
- Francisco Oviedo
- Jonathan Guillen
- Lucas Ledesma
- Maia Majzum
- Jonathan Manuel Palomeque
- Valentina Pich

2. TIPO DE PROYECTO: Tecnológico.
3. ESPACIO CURRICULAR: Análisis de datos (Procesamiento y Estadísticas).
4. EJES TEMÁTICOS/ RED DE CONCEPTOS:

1. Procesamiento de Datos

● Descripción:

Análisis y transformación de datos crudos utilizando bibliotecas especializadas como Pandas. El trabajo incluye la importación, limpieza, normalización y estructuración de conjuntos de datos reales para facilitar el análisis estadístico.

- **Competencias a fortalecer:**

- Capacidad para manipular grandes volúmenes de datos con eficiencia.
 - Manejo de entornos como Jupyter Notebook y Google Colab.
 - Aplicación de buenas prácticas en la documentación y organización de los datos.
-

2. Estadística Descriptiva y Exploración de Datos

- **Descripción:**

Aplicación de técnicas de estadística descriptiva para obtener conclusiones significativas sobre los datos. Se trabaja con medidas de tendencia central, dispersión, tablas de frecuencia y gráficos estadísticos para interpretar y comunicar resultados.

- **Competencias a fortalecer:**

- Interpretación de variables cuantitativas y cualitativas.
 - Aplicación de análisis descriptivos e inferenciales.
 - Evaluación de la calidad de los datos y análisis de valores faltantes o atípicos.
-

3. Programación y Automatización del Análisis

- **Descripción:**

Implementación de scripts en Python para automatizar el análisis, filtrado y visualización de datos. Se utilizan librerías como Pandas, NumPy y Matplotlib para desarrollar soluciones replicables y escalables.

- **Competencias a fortalecer:**

- Estructuración de código limpio y modular.
 - Integración de análisis estadístico con programación.
 - Capacidad para reutilizar y documentar código de análisis.
-

4. Visualización y Comunicación de Resultados

- **Descripción:**

Elaboración de gráficos y reportes visuales que permitan comunicar resultados de forma clara, precisa y comprensible. Uso de gráficos de barras, histogramas, diagramas de caja, entre otros.

- **Competencias a fortalecer:**

- Selección adecuada de visualizaciones según tipo de variable.
 - Uso de Matplotlib, Seaborn o herramientas similares.
 - Redacción de conclusiones coherentes con los datos representados.
-

5. PROBLEMÁTICAS/ NECESIDADES:

En el contexto actual, las empresas multinacionales como PepsiCo enfrentan el desafío de comprender con precisión el rendimiento de sus productos en distintas regiones del país, especialmente a través de su red de distribuidores. La falta de una visión integrada y depurada de los datos de ventas dificulta la toma de decisiones estratégicas respecto a stock, distribución, campañas comerciales o desarrollo de nuevos productos.

Desde el rol de analistas de datos, surge la necesidad de limpiar, estructurar y analizar estos datos para detectar patrones de consumo, identificar productos destacados y comparar resultados entre zonas geográficas.

6. FUNDAMENTACIÓN:

Este proyecto representa una problemática real del sector comercial de consumo masivo: la necesidad de transformar grandes volúmenes de datos en información útil.

El potencial del análisis de este conjunto de datos reside en su capacidad para relevar el comportamiento de ventas a nivel nacional, reconociendo diferencias regionales y productivas.

Es altamente relevante para nuestro perfil profesional porque nos permite aplicar competencias de análisis descriptivo, programación, procesamiento y visualización de datos, tal como exige el campo de la Ciencia de Datos en entornos reales.

Además, tiene impacto en la comunidad al facilitar mejores decisiones logísticas y comerciales que pueden influir en el acceso a productos, la eficiencia en el consumo de recursos y la reducción de costos.

7. VISIÓN DEL PROYECTO:

Construir un análisis integral y visualmente comprensible de las ventas de productos PepsiCo a lo largo de las diferentes regiones de Argentina, a partir de un dataset sin depurar, utilizando herramientas de procesamiento de datos y estadística descriptiva.

El proyecto aspira a consolidar un enfoque metodológico profesional para abordar problemas reales con datos reales, reforzando tanto el criterio técnico como el sentido crítico para interpretar información.

-OBJETIVOS:

- **Objetivo general:**

Analizar y visualizar los datos de ventas de productos PepsiCo distribuidos en Argentina, a partir de un dataset sin depurar, aplicando herramientas de procesamiento de datos y estadística descriptiva para detectar patrones relevantes y comunicar resultados útiles para la toma de decisiones.

- **Objetivos específicos:**

- 1- Limpiar y estructurar el conjunto de datos para su posterior análisis.
- 2- Analizar las variables numéricas y categóricas y detectar valores nulos o mal cargados.
- 3- Calcular y representar medidas estadísticas descriptivas de las ventas por producto, región y rubro.
- 4- Detectar productos más vendidos y su distribución por área geográfica.
- 5- Desarrollar visualizaciones efectivas que permitan comunicar los hallazgos de forma clara.

- **Metas:**

- Asegurar la integridad y consistencia de los datos.
- Completar el proyecto en el plazo estipulado.
- Analizar datos de ventas del último mes de 20 distribuidores de diferentes zonas geográficas del país.