

Investigación sobre Análisis Exploratorio de Datos en Python

Objetivo

El objetivo de esta tarea es que cada alumno investigue y comprenda todos los aspectos del Análisis Exploratorio de Datos (EDA) utilizando Python. Cada alumno deberá explorar las herramientas y técnicas disponibles en Python, y documentar sus hallazgos en un documento en Notion.

Temas a Investigar

Cada alumno debe investigar y documentar los siguientes temas:

1. Importación y limpieza de datos:

- Técnicas para manejar datos faltantes, duplicados y valores atípicos.
- Transformación y normalización de datos.

2. Visualización de datos:

- Uso de gráficos de dispersión, histogramas, diagramas de caja y bigotes (box plots), y gráficos de barras para explorar la distribución y relaciones de los datos.

3. Análisis estadístico descriptivo:

- Cálculo de medidas de tendencia central (media, mediana, moda) y dispersión (rango, varianza, desviación estándar).

4. Correlación y análisis de variables:

- Métodos para calcular la correlación entre variables y su interpretación.
- Análisis de multicolinealidad y su impacto en el modelo.

5. Análisis de distribuciones y normalidad:

- Evaluación de la distribución de las variables.
- Pruebas de normalidad y su relevancia en el análisis.

6. Técnicas de reducción de dimensionalidad:

- Uso de técnicas como el Análisis de Componentes Principales (PCA) para simplificar el conjunto de datos.

7. Análisis de grupos y segmentación:

- Técnicas de clustering como k-means y análisis jerárquico para identificar grupos naturales dentro de los datos.

8. Uso de herramientas y librerías de EDA:

- Aplicación de herramientas como pandas, matplotlib, seaborn y otros paquetes de Python para realizar EDA de manera eficiente.

Instrucciones

1. Investigación:

- Investiga cada uno de los temas utilizando recursos en línea, libros y documentación oficial de las librerías de Python.
- Prueba ejemplos prácticos de código en Python para comprender mejor el uso de las herramientas y técnicas.

2. Documentación en Notion:

- Crea una página en Notion para documentar tu investigación.
- La documentación debe incluir:
 - **Introducción** a cada tema.
 - **Herramientas y librerías** utilizadas para cada tema.
 - **Ejemplos de código** comentados para ilustrar las técnicas y herramientas.
 - **Conclusiones y recomendaciones** basadas en tu investigación y análisis.

3. Entrega:

- Sube el enlace a tu documento en Notion a classroom, la fecha de entrega es el viernes 2 de Agosto.
- La presentación se realizará en clase en la fecha asignada.

¡Buena suerte y feliz investigación!