

Objetivo: Este taller tiene como objetivo enseñar y practicar conceptos básicos y avanzados de estadística utilizando la biblioteca Pandas en Python. Los participantes trabajarán en equipo para resolver problemas y analizar datos reales.

Actividades:

- **Instrucciones:**

1. Descarguen un dataset de [Kaggle](#) o utilicen el dataset proporcionado.
2. Carguen el dataset en un DataFrame de Pandas.
3. Realicen una exploración inicial de los datos (head, describe, info).

- **Preguntas Guía:**

- ¿Cuántas filas y columnas tiene el dataset?
- ¿Qué tipo de datos contiene cada columna?
- ¿Hay valores nulos o datos faltantes?

1. Actividad: Estadísticas Descriptivas

- **Instrucciones:**

1. Calcular la media, mediana y moda de las columnas numéricas.
2. Calcular la desviación estándar y la varianza.
3. Calcular los percentiles (25%, 50%, 75%).

- **Preguntas Guía:**

- ¿Cuál es la media de cada columna numérica?
- ¿Cuál es la desviación estándar más alta y más baja entre las columnas numéricas?
- ¿Qué nos dicen los percentiles sobre la distribución de los datos?

2. Actividad 2: Visualización de Datos

- **Instrucciones:**

1. Utilizar Pandas y Matplotlib/Seaborn para crear gráficos de barras, histogramas y boxplots.
2. Visualizar la distribución de una columna numérica.
3. Comparar la media y la mediana con un boxplot.

- **Preguntas Guía:**
 - ¿Qué patrones o tendencias pueden observar en los gráficos?
 - ¿Hay algún outlier en los datos? ¿Cómo pueden identificarlo?

3. Actividad 3: Análisis de Correlación

- **Instrucciones:**
 1. Calcular la matriz de correlación entre las columnas numéricas.
 2. Crear un heatmap para visualizar las correlaciones.
- **Preguntas Guía:**
 - ¿Qué columnas están altamente correlacionadas?
 - ¿Cómo pueden interpretar una correlación negativa?

4. Discusión y Conclusiones

- **Instrucciones:**
 1. Cada uno presenta brevemente sus hallazgos.
 2. Discusión sobre los desafíos encontrados y cómo se resolvieron.
- **Preguntas Guía:**
 - ¿Qué aprendieron sobre el análisis de datos utilizando Pandas?
 - ¿Qué otras herramientas o métodos les gustaría explorar en el futuro?