# **Proyecto Integrador: Análisis Estadístico y Visualización de Datos del Naufragio del Titanic**

## **Descripción:**

Los estudiantes realizarán un análisis estadístico y visualización de datos utilizando Python, basándose en el famoso conjunto de datos del naufragio del Titanic disponible en la plataforma Kaggle. Este conjunto de datos incluye variables como el sexo, la edad, la clase de pasaje, el puerto de embarque y si sobrevivió o no la persona.

## **Objetivos:**

* Aplicar los conceptos básicos de programación en Python adquiridos durante las primeras sesiones del curso.
* Utilizar técnicas de estadística descriptiva y probabilidad para analizar los datos.
* Crear visualizaciones de los datos utilizando Matplotlib.

## **Instrucciones:**

1. [Descarga el conjunto de datos del Titanic de Kaggle y familiarízate con las variables disponibles.](https://www.kaggle.com/competitions/titanic/data)
2. Utiliza Python para cargar los datos en un dataframe y realiza una limpieza inicial si es necesario (por ejemplo, manejo de datos faltantes).
3. Realiza un análisis estadístico descriptivo de las variables de interés. Esto puede incluir cálculos de medidas de tendencia central, medidas de dispersión, y cualquier otra estadística relevante.
4. Utiliza técnicas de probabilidad para hacer inferencias sobre los pasajeros. Por ejemplo, puedes calcular la probabilidad de supervivencia en función de distintas variables (clase de pasaje, edad, sexo, etc.).
5. Crea visualizaciones de los datos utilizando Matplotlib. Esto puede incluir gráficos de barras, gráficos de cajas, histogramas, gráficos de dispersión, etc., dependiendo de lo que sea más apropiado para tus datos.
6. Redacta un informe que presente tus hallazgos. Asegúrate de interpretar los resultados de tus análisis en el contexto del naufragio del Titanic.

## **Rúbrica para el Proyecto Integrador: Análisis Estadístico y Visualización de Datos del Naufragio del Titanic**

**1. Código en Python (25 puntos)**

* **Claridad y organización (10 puntos):** El código está bien organizado y es fácil de seguir. Las variables y funciones tienen nombres descriptivos. Se proporcionan comentarios claros y útiles.
* **Corrección (15 puntos):** El código se ejecuta sin errores y realiza las tareas requeridas de manera efectiva.

**2. Análisis Estadístico (25 puntos)**

* **Descriptivo (10 puntos):** Se realizan cálculos correctos de estadísticas descriptivas relevantes.
* **Probabilidad e inferencia (15 puntos):** Se utilizan técnicas de probabilidad e inferencia de manera apropiada para hacer conclusiones a partir de los datos.

**3. Visualizaciones de Datos (25 puntos)**

* **Calidad técnica (10 puntos):** Las visualizaciones son técnicamente correctas y se presentan de manera efectiva.
* **Relevancia (15 puntos):** Las visualizaciones son relevantes para el análisis y ayudan a aclarar los hallazgos.

**4. Interpretación y Comunicación de los Resultados (25 puntos)**

* **Claridad de la escritura (10 puntos):** El informe está bien escrito y es fácil de seguir. Los resultados del análisis se presentan de manera clara y concisa.
* **Interpretación de los resultados (15 puntos):** Los resultados del análisis se interpretan de manera correcta y se discuten en el contexto del naufragio del Titanic.