*NOMBRE DE LA INSTITUCIÓN*

**PROYECTO DE ANÁLISIS DE DATOS**

***Análisis de datos de* UFO Sightings**

**Presentado por:**

**Laureano D. Rojas C.**

**Supervisor:**

**Carlos Cáceres**

**Fecha Inicio:**

**06-03-2025**

**Fecha Entrega:**

**15-03-2025**

**Descripción breve del proyecto:**

**Tabla que contiene más de 80.000 registros de avistamientos de ovnis entre 1949 y 2014, incluidos ciudad, estado, país, latitud/longitud, forma, duración, fecha/hora y comentarios**

**2025**

**Modelo de Análisis de Datos**

**1. Definición del Problema**

**✅ Problema a resolver:**

**Ninguno ya que solo estamos haciendo un análisis descriptivo**

**❓ Preguntas clave:**

* **¿En qué cuidad se registraron más avistamientos OVNI?**
* **¿En qué país se registraron más avistamientos OVNI?**
* **¿Cuál fue la forma más común del OVNI avistado?**
* **¿Cuál fue la hora con más frecuencia de avistamientos?**

**⚡ Contexto y objetivo:**

**El contexto es que tenemos datos de avistamientos OVNI desde 1949 hasta 2014 y queremos saber en qué ciudades y países hubo más avistamientos ovni, así como la forma más común del objeto y por último la hora con más frecuencia de avistamientos.**

**2. Recolección de Datos**

**📂 Fuentes de datos:**

Los datos provienen de la web de Kaggle:

https://www.kaggle.com/datasets/sahityasetu/ufo-sightings

**🌐 Descripción de los datos:**

Este dataset está compuesto por una sola tabla la cual a su vez tiene las siguientes columnas:

Datetime: Es la fecha y hora en el que ocurrió el avistamiento.

City: Es la cuidad donde ocurrió el avistamiento.

State: Es el estado donde ocurrió el avistamiento.

Country: Es el país donde ocurrió el avistamiento.

Shape: Es la forma del objeto avistado.

Duration (Seconds): Es la duración en segundos del avistamiento.

Duration (Hours/min): Es la duración en horas y minutos del avistamiento.

Comments: Son los comentarios acerca del avistamiento.

Date posted: Es la fecha de publicación.

Latitude: Forma parte de las coordenadas de la ubicación.

Longitude: Forma parte de las coordenadas de la ubicación.

**3. Limpieza y Preparación de Datos**

**💡 Procesos aplicados:**

* Eliminación de duplicados ✅/❌
* Manejo de valores nulos (indicar estrategia: eliminar, imputar, etc.) ✅/❌
* Conversión de tipos de datos ✅/❌
* Normalización y transformaciones ✅/❌

**📊 Variables seleccionadas:**

Las columnas que vamos a usar son:

City

Country

Shape

Duration (Seconds)

**4. Exploración y Análisis Descriptivo**

**📅 Estadísticas básicas:**

* Promedio, mediana, moda
* Máximos y mínimos
* Distribuciones de los datos

**🌈 Visualizaciones clave:**

**5. Modelado y Análisis Inferencial**

**🔢 Métodos utilizados:**

* Pruebas de hipótesis ✅/❌
* Regresión ✅/❌
* Machine Learning (indicar técnicas si aplica) ✅/❌
* Segmentación o clustering ✅/❌

**6. Interpretación y Comunicación de Resultados**

**💬 Principales hallazgos:**

Las conclusiones que sacamos con el análisis son las siguientes:

* El tipo de filme mejor calificado fue el Documental, con una puntuación promedio de 6.73.
* El país con más filmes estrenados es Estados Unidos, con un total de 42387 filmes estrenados.
* Los cinco años en los que se estrenaron más filmes fueron, 2018, 2017, 2016, 2015 y 2014.
* El género con más filmes fue

**📈 Visualizaciones finales:**

**7. Toma de Decisiones y Acción**

**🔍 Recomendaciones basadas en datos:**

Según el análisis hecho basado en los datos, podemos concluir que, un Documental sería una buena elección ya que este recibió mejor promedio de puntuación en cuanto a tipos de filme que debemos hacer

**🌐 Impacto esperado:**

El impacto que esperamos es que al momento de estrenar un filme del tipo documental, tengamos la misma aceptación que con los filmes del tipo Documental de los datos analizados.

**8. Monitoreo y Mejora Continua**

**⏳ Métricas de seguimiento:**

Debemos estrenar el filme y obtener los resultados o el impacto esperado

**📅 Actualizaciones futuras:**

Debemos de recolectar los datos obtenidos después de estrenar el filme para así ir actualizando nuestros datos e ideas.

**💪 Este documento te ayudará a estructurar tu proyecto paso a paso. Completa cada sección con la información de tu análisis y avanza de manera organizada.**