

EDWIN GIOVANNY CUENCA MORENO
GESTION DE DATOS - MAESTRIA EN IOT
PUJ

- **QUERY 1 - Sample_mflix /movies**

Descripción general de los datos:

Para este desarrollo se realizará un análisis basado en el país, género y año de cada película. Se realizó la consulta de los datos de prueba de MongoDB Atlas correspondientes a "Sample_mflix /movies", se realizó la extracción de las columnas País, Género y Año.

En el análisis de los registros se encontraron inconsistencia tales como campos nulos y algunos años que al final incluían el carácter "é" lo cual generaría inconvenientes al momento de realizar el análisis.

En total se tienen 23530 registros, para el análisis se realizó ajuste de formato a los campos de País y Género en String y el Año en Int.

De esta manera es posible responder a preguntas como:

- Cuál es el Top 3 de los géneros más producidos:
En el lugar número uno se encuentra el genero Drama con 3621 producciones, en el segundo lugar esta el genero de Comedia con 1147 producciones y en el tercer lugar esta una fusión entre los géneros de Comedia y Drama con 1124 producciones.
- En qué año se tuvo la mayor producción de películas:
El año en el que se realizó el mayor numero de producciones fue en el año 2013 con un total de 788 películas.
- Cual es Top 3 de los países con mayor número de producción de películas:
En el lugar número uno se encuentra Estados Unidos con 8959 producciones, en el segundo lugar se encuentra Reino Unido con 1123 producciones y en el tercer lugar esta Francia con 839 producciones.

- **QUERY 2 - Sample_geospatial/shipwrecks**

Descripción general de los datos:

Para este desarrollo se realizará un análisis basado en datos de eventos de Naufragio, Profundidad y Condiciones de los elementos encontrados. Se realizó la consulta de los datos de prueba de MongoDB Atlas correspondientes a "Sample_geospatial /shipwrecks", se realizó la extracción de las columnas Naufragio, Profundidad y Condiciones. Adicionalmente la data contiene elementos como las coordenadas ya que con estos sería posible visualizar mediante un mapa los puntos exactos de ocurrencia.

En el análisis de los registros se encontraron inconsistencia tales como campos nulos.

En total se tienen 11095 registros, para el análisis se realizó ajuste de formato a los campos de Naufragio, Profundidad y Condiciones en String.

De esta manera es posible responder a preguntas como:

- Cuál es el tipo de Naufragio que presenta una mayor cantidad de eventos:
En el primer lugar con una cantidad de 6578 registros, se encuentran los Naufragios de tipo sumergidos con la condición de peligrosos.
- Cuál es la condición que más se presenta en este tipo de eventos:
La condición que mas se presenta es siempre bajo el agua completamente sumergido para un total de 8399 registros.
- Cuál es el porcentaje que se tienen de datos con resultado de "Profundidad Desconocida":
Se tiene un total de 5572 registros correspondientes al 69.61% de eventos ocurridos donde se tiene Profundidad Desconocida.

- **QUERY 3 - Sample_restaurants/restaurants**

Descripción general de los datos:

Para este desarrollo se realizará un análisis basado en datos de eventos de Cocina, Restaurante y Ciudad de los elementos encontrados. Se realizó la consulta de los datos de prueba de MongoDB Atlas correspondientes a "Sample_restaurants/restaurants", se realizó la extracción de las columnas Cocina, Restaurante y Ciudad. Adicionalmente la data contiene elementos como una categoría de calificación y el año de referencia que serían importantes para analizar cuáles de estos presentan una mayor calidad.

En el análisis de los registros se encontraron inconsistencia tales como campos nulos.

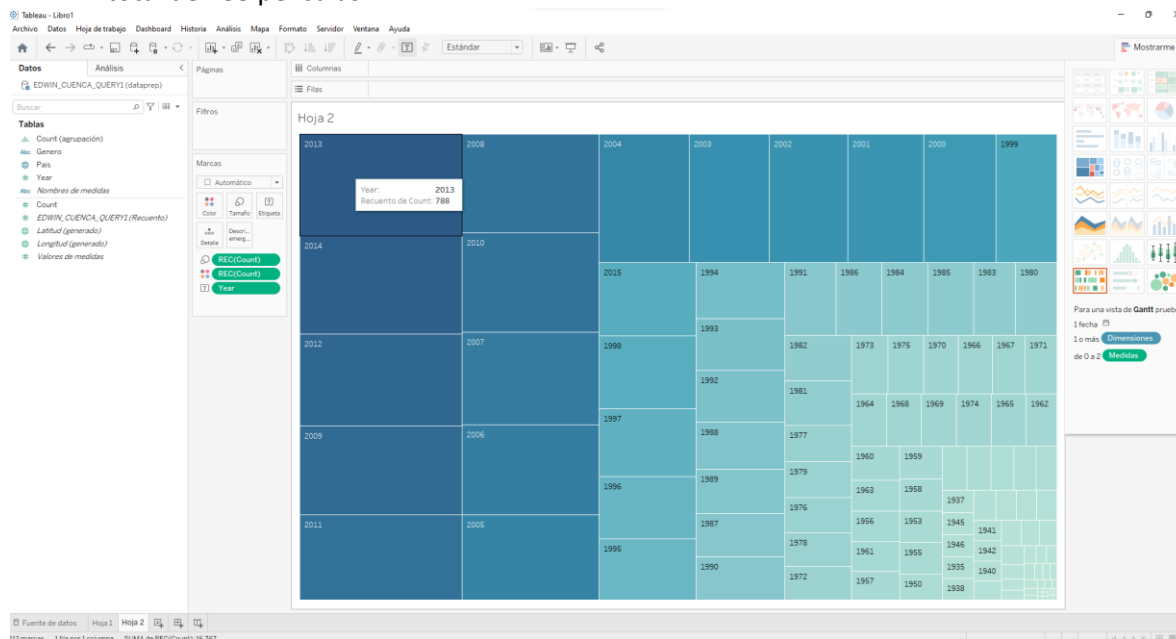
En total se tienen 25359 registros, para el análisis se realizó ajuste de formato a los campos de Cocina, Restaurante y Ciudad en String.

De esta manera es posible responder a preguntas como:

- Cuál es el Top 3 de las ciudades con mayor cantidad de restaurantes:
En el primer lugar se encuentra la ciudad de Manhattan con 10259 restaurantes, en segundo lugar, se encuentra la Ciudad de Brooklyn con 6086 restaurantes y en tercer lugar esta la ciudad de Queens con 5656 restaurantes.
- Cuál es la cocina más representativa:
La Cocina mas representativa es la American representada en 6183 restaurantes
- Cuál es el Top 3 de restaurantes:
En primer lugar, se encuentra Subway con 421 restaurantes, en segundo lugar se encuentra Starbucks con 223 restaurantes y en tercer lugar McDonald's con 208 restaurantes.

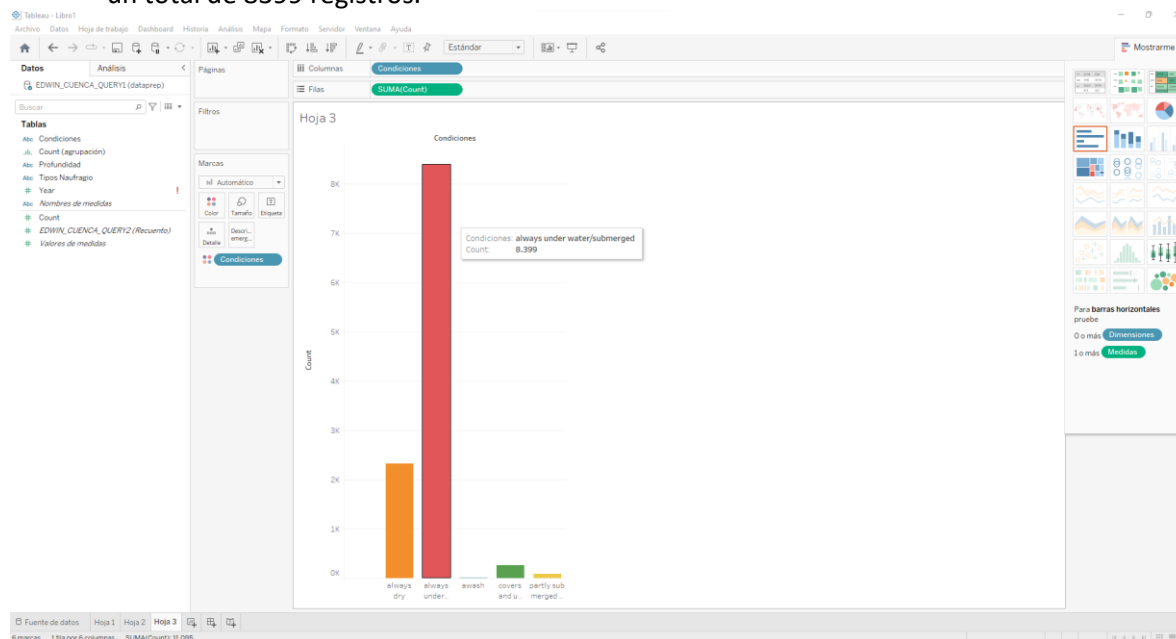
Visualización Tableau – QUERY1

- En qué año se tuvo la mayor producción de películas:
El año en el que se realizó el mayor número de producciones fue en el año 2013 con un total de 788 películas.



Visualización Tableau – QUERY2

- Cuál es la condición que más se presenta en este tipo de eventos:
La condición que más se presenta es siempre bajo el agua completamente sumergido para un total de 8399 registros.



Visualización Tableau – QUERY3

- Cuál es el Top 3 de las ciudades con mayor cantidad de restaurantes:
En el primer lugar se encuentra la ciudad de Manhattan con 10259 restaurantes, en segundo lugar, se encuentra la Ciudad de Brooklyn con 6086 restaurantes y en tercer lugar esta la ciudad de Queens con 5656 restaurantes.

