**Presentación Prueba Técnica Analista de Datos.**

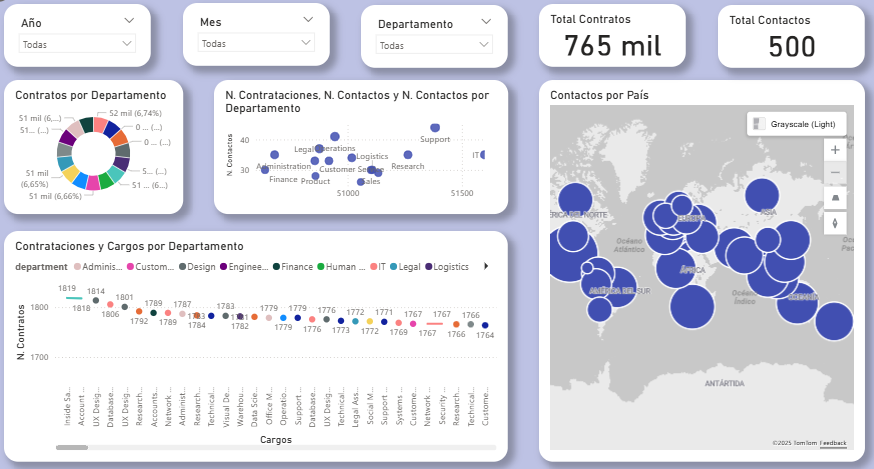
Candidato: Cristhian Alexander Torres Polanco

Repositorio: <https://github.com/cristhianalextorres/dataAnalytics>

En el siguiente documento se responde la rúbrica de la prueba en mención.

El proyecto abarcó desde la carga y depuración de datos en Python hasta su integración en una base de datos SQLite conectada a Power BI.  
Tras limpiar y estructurar los datasets, se construyó un dashboard interactivo que permite analizar las contrataciones por departamento, cargo y red de contactos, ofreciendo una visión clara y accionable del proceso de contratación.

1. **Definición de Pregunta de Negocios.**
   1. ¿Qué departamento tiene mayor impacto en el volumen de contrataciones?
   2. ¿Qué cargos generan mayor impacto en las contrataciones totales?
   3. ¿Existe relación entre el número de contactos y el volumen de contrataciones por departamento?
   4. ¿En qué países o regiones se concentra la red de contactos?
2. **Se desarrolló un Tablero de Contrataciones de acuerdo a la data suministrada:**



* 1. ¿Qué departamento tiene mayor impacto en el volumen de contrataciones?

R/ No hay un Departamento que tenga un volumen significativo con respecto a los otros, levemente, el departamento de IT es el de mayor contratación con 51599 con una participación del 6,74%

* 1. ¿Qué cargos generan mayor impacto en las contrataciones totales?

R/ Como en Departamento. En Cargos no tenemos diferencias muy grandes, el cargo Inside Sales Associate Jr.

* 1. ¿Existe relación entre el número de contactos y el volumen de contrataciones por departamento?

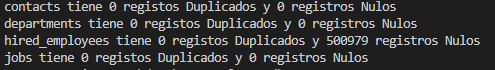
R/ Se detalla una correlación Positiva leve en el volumen de Contratación y el número de contactos, se puede permitir priorizar perfiles críticos o con alta rotación.

* 1. ¿En qué países o regiones se concentra la red de contactos?

R/ La mayoría de contactos están en Sur América y Europa. Esto también ayuda observar perfiles críticos.

1. **Perfilamiento y calidad de los datos.**

Durante el proceso de preparación se aplicaron análisis de duplicados, valores nulos y consistencia en los campos claves.



* 1. **Duplicados**: No se encontraron registros duplicados.
  2. **Vacíos**: Se encontraron 500.979 en hired\_employees, se eliminaron las filas incompletas.
  3. **Consistencia:** Se evidenció desalineación entre los campos *country* y *region* del dataset contacts, posiblemente por errores de origen o mezcla de codificaciones.
  4. **Formato:** El campo datetime de hired\_employees fue convertido correctamente a tipo fecha para habilitar análisis temporales.

**Recomendaciones:**

1. Estandarizar los valores de *country* y *region* para mejorar precisión geográfica.
2. Definir llaves primarias y foráneas en origen para evitar duplicidad. En el ejercicio se definieron llaves foráneas al cargar a la BD SQLite.
3. Realizar validación de recolección de datos para contrarrestar los datos incompletos o vacíos.