

## Prueba Técnica Ingeniería de datos

El reto consiste en realizar un proceso ETL que permita conocer datos de los países con los que Colombia ha realizado algún tipo de tratado. A continuación se explican las bases de datos que se trabajarán y las actividades a realizar.

Las instrucciones y bases de datos se encontraran en el repositorio git (<https://github.com/sebas317/testDataEnginner>).

### Bases de datos:

1. **Tratados\_internacionales\_de\_Colombia.csv** esta base de datos contiene la lista de tratados internacionales realizados por colombia con un estado o una organización. El campo que contiene los estados con los que se realiza el tratado es (**Estados-Organismos**). <https://github.com/sebas317/testDataEnginner>
2. <https://restcountries.eu/rest/v2/all> la anterior API contiene la información básica de los países del mundo. Esto nos permitirá conocer la información básica de los países que han realizado algún tratado internacional con Colombia.

### Actividades a realizar:

1. Construir una tabla uniendo las dos fuentes de datos que permita conocer la información de cada país incluido en el tratado. La tabla debe tener los siguientes campos:

nombre_campo	tipo_dato
Nombre del Tratado	Texto
Bilateral	Boolean
Lugar de Adopción	Texto
Fecha de Adopción	date formato (YYYY-MM-DD)
Estados-Organismos	Texto
Temas	Texto
Naturaleza del Tratado	Texto
Depositario	Texto
Suscribió Por Colombia	Texto
Vigente	Boolean
Fecha Ley Aprobatoria	date formato (YYYY-MM-DD)
Número Ley Aprobatoria	Texto
Sentencia Fecha Ley	date formato (YYYY-MM-DD)
Sentencia Número	Texto

Decreto Fecha Diario Oficial	date formato (YYYY-MM-DD)
Decreto Número Diario Oficial	Texto
País del tratado	Texto
Código de llamadas	Texto
Capital	Texto
Región	Texto
Subregión	Texto
Población	Texto
Área	integer
Zona horaria	Texto
Monedas	Texto separado por comas
Idiomas	Texto separado por comas
Cantidad fronteras	integer
Diferencia horas zona horaria	decimal

- Calcular la diferencia en horas entre Colombia y el país firmante del tratado teniendo en cuenta su zona horaria.
- Para los campos que registra valores **(NO REGISTRA)** se deben almacenar como NULL.

#### Entregables:

1. Repositorio GIT con el código de la ETL que realice la unión de las bases de datos con los campos mencionados en la tabla anterior.
  2. Archivo parquet con los datos de la tabla estructurada.
2. Crear una tabla en base de datos postgres que almacene la información automáticamente desde el script ETL de las transformaciones. Los datos almacenados en la tabla deben tener las siguientes condiciones:

#### Entregables:

1. Script de creación de la tabla para motor de base de datos postgres.
  2. Evidencia del cargue de la tabla en el motor de base de datos postgres a través de la ETL (Pantallazo).
3. Realizar un script SQL que permita tomar de la tabla creada en el **punto 2**:
- La fecha de adopción del primer tratado de cada país
  - Cantidad de tratados que se tienen con cada país
  - Diferencia en meses desde la fecha de adopción de este tratado hasta la fecha actual
  - Se deben tener en cuenta únicamente los países con tratados vigentes.
  - El resultado de este query SQL debe almacenarse en una nueva tabla de nombre **agrupacion\_por\_pais** con los siguientes campos:

nombre_campo	tipo_dato
país	texto
cantidad de tratados vigentes	integer
cantidad de tratados no vigentes	integer
fecha adopción primer tratado	date
diferencia en meses desde el primer acuerdo a la fecha actual	integer

**Entregables:**

1. Script que realice: creación de la tabla **agrupacion\_por\_pais** y el volcado de los datos agrupados en la tabla mencionada anteriormente.