

Instrucciones de ejecución

Crear contenedores

1. Crear imagen principal

```
sh build_image.sh
```

ⓘ Nota:

Asegúrate que Docker esté corriendo

2. Correr imagen con Postgresql de fondo

```
sh run_db_image.sh
```

3. Corre el contenedor principal:

```
sh run_image.sh
```

ETL (Extract, Transform, Load)

1. Cargar y limpiar archivo excel de base de datos SIPRI sobre gasto militar. Única sección que usa pandas. (Ya fue ejecutado, se puede saltar.)

```
python src/etl/load_excels.py
```

2. ETL principal: Corre limpieza, transformación, eliminación de nulos, estandarización de nombres, etc. y escribe a base de datos.

- a. Ejecutar programa y escribir a base de datos

```
sh run_etl.sh
```

b. Verificar base de datos

```
sh run_verify_db.sh
```

3. Programa de Aprendizaje Automático:

a. Ejecutar programa

```
sh run_ml.sh
```

b. Ver análisis de resultados

```
cat analisis_de_resultados.txt
```

Estructura de Base de Datos ms_di:

- country VARCHAR(100)
- constant_2023_usd_military FLOAT
- current_usd_military FLOAT
- share_gdp_military FLOAT
- per_capita_military FLOAT
- share_govt_spend_military FLOAT
- exports_percent_gdp FLOAT
- gdp_per_capita_2015_usd FLOAT
- gdp_per_capita_current_usd FLOAT
- gcf_percent_gdp FLOAT
- life_expectancy_female FLOAT
- life_expectancy_male FLOAT
- life_expectancy_total FLOAT
- net_migration FLOAT

- trade_percent_gdp `FLOAT`

1. Se puede conectar a la base de datos para correr queries:

```
sh connect_to_db.sh
```

❗ **Nota:**

Contraseña es testPassword

Nombre de la base de datos es ms_di

Ejemplo: `SELECT * FROM ms_di WHERE country = 'Albania';`

Pruebas unitarias

1. Se puede correr pruebas unitarias para la sección de ETL.

```
pytest src/etl/test.py
```