

ETL Pipelines for Multi-Platform DaaS

1. Power BI:

- a. Extract: Use the **Power BI REST API** to retrieve data from datasets.
- b. Transform: Clean and format the data as required.
- c. Load: Push the transformed data to your platform using appropriate endpoints.

2. SQL Databases:

- a. Extract: Use **pyodbc** or **SQLAlchemy** to connect to the database and run queries.
- b. Transform: Process the data to fit your needs.
- c. Load: Insert the processed data into your platform's database.

3. Excel:

- a. Extract: Use **pandas** or **openpyxl** **OR** **Apache NiFi** or **Talend** to read data from Excel files.
- b. Transform: Clean and shape the data.
- c. Load: Load the data into your platform's storage.

4. Tableau:

- a. Extract: Utilize **Tableau's REST API** **OR** **Alteryx** or **Tableau Prep** to pull data.
- b. Transform: Adjust the data as required.
- c. Load: Send the processed data to your platform.

5. Metabase:

- a. Extract: Use **Metabase's API** to fetch data **OR** Use a tool like **Stitch** or **Fivetran** to extract data.
- b. Transform: Format and clean the data.
- c. Load: Import the data into your platform.

6. IBM Watson:

- a. Extract: Access data using **Watson's APIs** **OR** **IBM Cloud Pak**
- b. Transform: Process and prepare the data.
- c. Load: Load it into your platform.

7. MongoDB:

- a. Extract: Use a MongoDB driver like `pymongo` to connect and fetch data.
- b. Transform: Process the JSON-like documents.
- c. Load: Insert the transformed data into your platform's database.