ETL
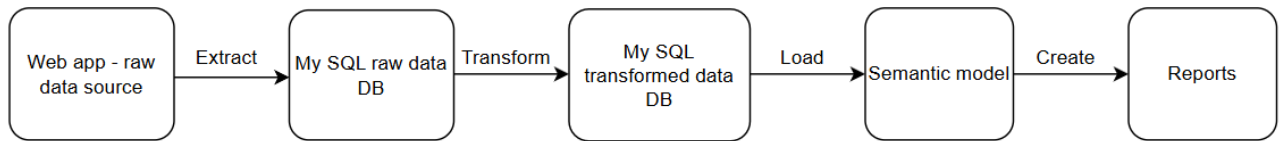

This process can be executed in two ways: manually and automatically.

- **Manual Process:**
  This is run the first time the process is installed, and any time a data transformation or other part of the process is updated.
  The code for the manual process is stored in a Jupyter Notebook, allowing easy execution of specific code sections for particular use cases.

- **Automatic Process:**
  This runs nightly and supplements the data with information from the previous day.
  The code for the automatic process is compiled into an .exe file.

Automatic Process Flow:

1. At a scheduled time each night, a Scheduler runs the .exe file, which:

   - Retrieves the previous day's data from the web application via an API and adds it to the raw_data MySQL database.

   - Transforms the previous day's data and stores it in the transformed_data MySQL database.

2. After the above steps are completed, Power BI's automatic data refresh (also scheduled) pulls data from the transformed_data MySQL database and updates the semantic model.

3. Power BI reports are automatically updated with the latest data.