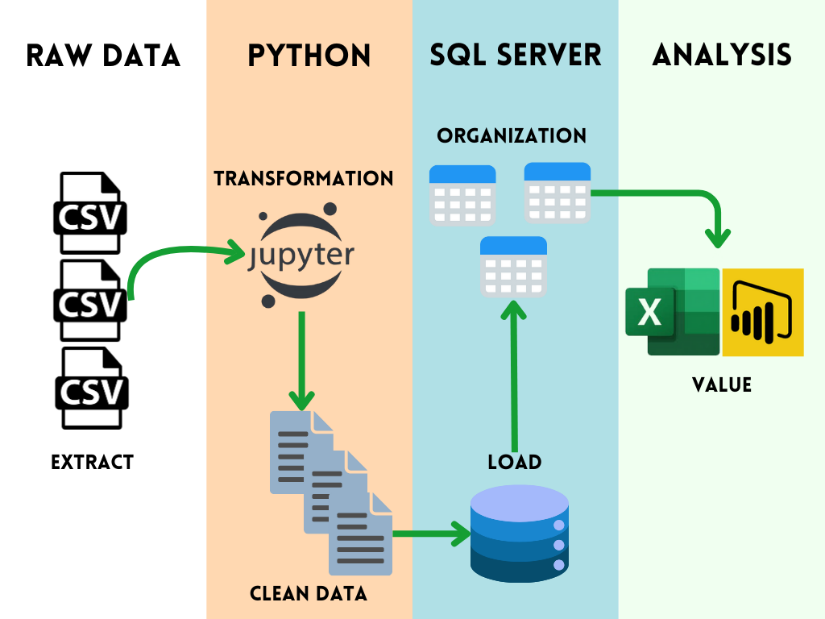
ETL ON SQL AND PYTHON



The project aimed to reduce data analysis and processing times through an automated ETL system. This will allow for the improvement of data transformation and cleaning, creating a database with relevant and accessible information for the client.

**Here are the questions I was interested in answering**

How can I carry out the ETL process on two different platforms?

What are the most relevant variables for analysis?

How can I manage the information to reduce processing times?

What information is most relevant to the end customer, and how can I make it easy for them to access?

**I took the following steps to create my analysis**

Load the raw data into Jupyter Notebook.

Perform data cleansing, selecting relevant variables, and generating new files.

Load the new files into SQL Server.

Modify variables, create new tables, and establish relationships between them.

Query relevant databases in Power Query - Excel.

**Here are my takeaways**

Due to the need for context, the process had to be created in Jupyter Notebook and SQL Server.

The weight of the initial information was considerably reduced.

A process was established that allows weekly data updates.

An effective connection was established so that the client can perform its respective analysis in Excel.