**Objective of the Project:**

* Extract data from the Wikipedia and Kaggle Data files, Transforming the datasets by cleaning them up and joining them together and Load the clean dataset into a SQL database which can be used for the Hackathon organized By Britta from Amazing Prime Video

**Process:**

* Extracting data from Wikipedia Movies JSON and the Kaggle data.
* Clean the Wikipedia, Kaggle & Ratings Data by filtering, parsing, translating, sorting, merging etc.
* Once we merged the clean data from Wikipedia, Kaggle and Ratings files create a database and load on SQL

**Results:**

* **Write an ETL function to read three data files**

<https://github.com/sachinnabar/MoviesETL2/blob/main/Wikipedia%20movies%20JSON.PNG>

Text, email

Description automatically generated

<https://github.com/sachinnabar/Movies-ETL2/blob/main/Kaggle%20Data.PNG>

A screenshot of a computer

Description automatically generated with medium confidence

<https://github.com/sachinnabar/Movies-ETL2/blob/main/Ratings%20Data.PNG>

Graphical user interface, application, Word

Description automatically generated

* **Extract and Transform the Wikepedia Data**

<https://github.com/sachinnabar/Movies-ETL2/blob/main/Wiki_movies_df.PNG>

Graphical user interface, text, application, email

Description automatically generated

<https://github.com/sachinnabar/Movies-ETL2/blob/main/Wiki_movies_df_column.PNG>

Graphical user interface, text, application, email

Description automatically generated

* **Extract and Transform the Kaggle Data**

<https://github.com/sachinnabar/Movies-ETL2/blob/main/movies_with_ratings_df.PNG>

Graphical user interface, text, application, email

Description automatically generated

<https://github.com/sachinnabar/Movies-ETL2/blob/main/movies_df.PNG>

Graphical user interface, text, application, email

Description automatically generated

* **Create Movie Database**

<https://github.com/sachinnabar/Movies-ETL2/blob/main/movie_database.PNG>

Text

Description automatically generated