Source: https://api.balldontlie.io/v1/players

API_ETL

Extract

Transform

Load

- Data Retrieval: This step involves pulling data from , such as APIs BallDontLie API for NBA player stats. It includes handling API requests, managing responses, and ensuring that data collected is complete and accurate.
- Initial Processing: In this step, data is initially checked and pre-processed using scripts in a Jupyter Notebook. This includes tasks like parsing JSON data from the API, filtering relevant fields, and performing preliminary data quality checks to identify and discard corrupted or irrelevant data entries.

• Dataframe Creation: Initializes dataframes from raw data extracted via API, setting up structured, tabular formats that facilitate easier data manipulation and analysis.

- Engine Setup: Establish a connection to your target database using SQLAlchemy's create_engine.
 Configure the connection string with appropriate credentials to ensure secure and efficient database access.
- Data Loading: Utilize the to_sql method from Pandas to transfer data from your dataframe directly into the database. Configure options such as
- if_exists='replace' to handle pre-existing data by either replacing it or appending to it, depending on your requirements.