Code File Snippets:





Sample downloaded JSON File from API:



Sample processed CSV File:



Analysed output:

The output gives the Maximum Currency Rate, Maximum Currency Trading Date, Minimum Currency Rate, Minimum Currency Trading Date, Average Currency Rate, Percentage changed based on the MINIMUM Currency rate.

[('BaseCurrency', 'TargetCurrency', 'MaxCurrRate', 'MaxTradeDate', 'MinCurrRate', 'MinTradeDate', 'AvgCurrRate', 'PercentageCurrRate'),

('AUD', 'NZD', '1.09443405', '2024-04-05 00:00:00', '1.07123743', '2024-03-10 00:00:00', '1.0844122506451614', '3.2556183474084137')]

High level Approach:

1. To download the JSON file using a API
   1. ADF
   2. Python
2. Preprocess/Validate
   1. Python (ADB)
3. Analysis:
   1. SQL
   2. Python

Output: ?????

Unifinished\*\*\*\*\*\*\*\*

1. Azure Data Factory (ADF):

o Define a JSON dataset in ADF to read data from the Exchange Rates API.

o Create a pipeline that uses the ADF copy activity to ingest data from the API into Azure Data Lake Storage Gen2 (ADLS Gen2).

2. Azure Databricks:

o Use Databricks notebooks to transform the JSON data.

o Create a Parquet file in ADLS Gen2 to store the curated data.

3. Azure Data Lake Storage Gen2 (ADLS Gen2):

o Store the raw JSON data ingested from the API.

o Store the transformed data after processing in Databricks.

