Case Study – Answers

Rose Alcolea – Product Manager

To ensure that all aspects of the assignment are fully addressed, I have provided a direct answer to each assignment. However, for a more comprehensive and thorough approach, I would recommend referring to the Jupyter Notebook in the repository.

1. An ETL process diagram or action list which would load the Sales Transactions table from IT database to BI database.

Diagram:

Diagram

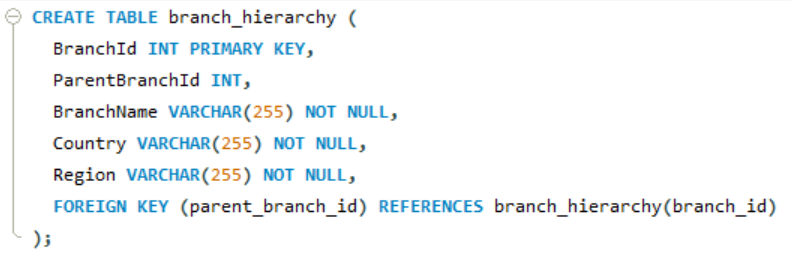
Description automatically generated

1. A Branch hierarchy from the provided file. Imagine that this would be used by other processes in future and become the company’s master dataset which the BI department would need to maintain, potentially using it for data applications developed in house.

Diagram

Description automatically generated

* + Please provide the DDL statement of Branch tables you would create.



* + Please keep in mind that new branches can be created, and old branches can be closed.

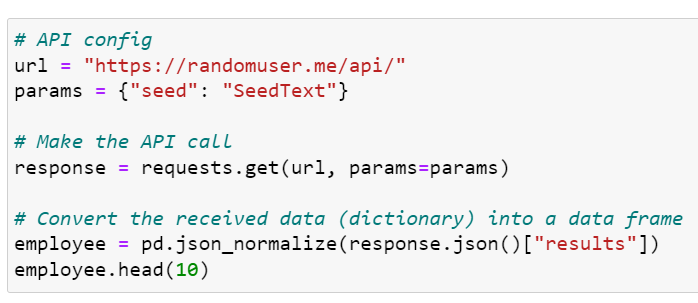
To handle new branches and closed branches, we can update the branch\_hierarchy table as needed. For example, if a new branch is created, we would add a new row to the branch\_hierarchy table with the appropriate values. If a branch is closed, we would update the parent\_branch\_id field of its child branches to reflect the new parent branch, or set it to null if the branch has no parent.

1. A script which would calculate a dataset/table where we could see all the customers and their underlying child customers.

Graphical user interface, text, application, email

Description automatically generated

1. A script to call the Employee API to get details about users



1. A data model or a view combing all provided data sources which would give the Sales amount per product, per location, per customer, per user
   * It does not matter how you return this result, an excel file would work as well

Graphical user interface, text, application, table

Description automatically generated

Text, table

Description automatically generatedTable

Description automatically generatedTable

Description automatically generatedText, table

Description automatically generated

* + You may also visualize this dataset in BI tool such as Power BI

Chart, treemap chart

Description automatically generated

**Conclusion:**

Throughout this project, a variety of data sources were utilized and processed using Python to clean and transform the data. The resulting data was then merged into a single table and loaded into Power BI to create a dashboard featuring clear and concise data visualizations. This process allowed for more efficient analysis and decision-making based on the provided data. While there is always room for improvement, this project highlights the benefits of using Python and Power BI to transform and analyse complex datasets.