

# README

## Data Science Assignment Submission

This repository contains the deliverables for the Data Science assignment focused on analyzing an eCommerce Transactions dataset. Below is a description of each file and its purpose.

### Files Included

1. **FirstName\_LastName\_EDA.pdf**
  - A detailed report summarizing the exploratory data analysis (EDA) performed on the dataset.
  - Includes key insights, visualizations, and actionable recommendations for improving business strategies.
2. **FirstName\_LastName\_EDA.ipynb**
  - Jupyter Notebook containing the code for performing EDA.
  - Covers data loading, preprocessing, visualizations, and insights generation.
3. **FirstName\_LastName\_Lookalike.csv**
  - A CSV file containing the top 3 most similar customers for the first 20 customers based on their profiles and transaction history.
  - Format: CustomerID, Lookalikes (list of similar CustomerIDs with similarity scores).
4. **FirstName\_LastName\_Lookalike.ipynb**
  - Jupyter Notebook implementing the Lookalike Model.
  - Includes data preprocessing, similarity calculation, and recommendations generation.
5. **FirstName\_LastName\_Clustering.pdf**
  - A detailed report summarizing the results of customer segmentation using clustering techniques.
  - Includes the number of clusters, DB Index value, cluster characteristics, and visualizations.
6. **FirstName\_LastName\_Clustering.ipynb**
  - Jupyter Notebook implementing the clustering analysis.
  - Covers data preprocessing, K-Means clustering, evaluation (DB Index), and visualization.

### Instructions to Run

1. **Environment Setup:**
  - Install the required Python libraries:
  - `pip install pandas numpy scikit-learn matplotlib seaborn`
2. **Run Jupyter Notebooks:**
  - Open the .ipynb files in Jupyter Notebook or JupyterLab.
  - Execute each cell sequentially to reproduce the results.
3. **Outputs:**

- **Running the Lookalike Model and Clustering notebooks will generate the following outputs:**
  - **FirstName\_LastName\_Lookalike.csv:** Contains lookalike customer recommendations.
  - **Cluster visualizations will be displayed within the notebook.**

## **Contact Information**

**For any questions or issues, feel free to reach out to:**

- **Name:** Cheedalla Rahul
- **Email:** cheedallarahul9876@gmail.com

**Thank you for reviewing this submission!**