Data Science Project

**Group name: Data Explorers**

**Specialization:** Data Science

**Team members Details:**

|  |  |  |  |
| --- | --- | --- | --- |
| **Name** | **Email** | **Country** | **College\Company** |
| Jalpa Deepak Patel | pjalpa015@gmail.com | Canada | University of Windsor |
| Dhvanilkumar Kiritkumar Prajapati | dkprajapati46@gmail.com | Canada | University of Ottawa |
| Sri Ram Prasad Commuri | [csriramprasad@gmail.com](mailto:csriramprasad@gmail.com) | India | Data Glacier |

**Project Title:** Bank Customer Segmentation

**Problem Description**: Create customer segments/categories to send personalized Christmas offers. Develop a feature to group customers in one category and others in different categories.

**Business Understanding:** Send customized offers belonging to relevant customer categories. Customer data has been provided to understand trends of customer categories.

**Project lifecycle**: The tasks needed to implement include –

1. Business understanding
2. Data understanding
3. EDA
4. Feature engineering
5. Model building
6. Model evaluation
7. Presentation and documentation

Data Intake Report

Name: Bank Customer Segmentation

Report date: 17th September 2022

Internship Batch: LISUM12

Version: 1.0

Data intake by: Dhvanilkumar Kiritkumar Prajapati

Data intake reviewer: Jalpa Patel

Data intake reviewer: Sri Ram Commuri

Data storage location: <https://drive.google.com/drive/folders/1bfCpJIKmp6IHxiLPWvOS2nU1dc24pViB>

**Tabular data details: cust\_seg**

|  |  |
| --- | --- |
| **Total number of observations** | 1000000 |
| **Total number of files** | 1 |
| **Total number of features** | 47 |
| **Base format of the file** | .csv |
| **Size of the data** | 154 MB |

**Proposed Approach:**

* Column headers are in Spanish; convert them to English names for ease of programming and understanding.
* Perform data cleaning and pre-processing.