

# Data Intake Report

## Data Glacier

### Team Members:

- Ray Ng
  - [rayng2018@gmail.com](mailto:rayng2018@gmail.com)
  - Canada
  - University of British Columbia
  - Specialization: Data Science
- Rita Uzoka
  - [rita.uzoka@yahoo.com](mailto:rita.uzoka@yahoo.com)
  - United Kingdom
  - Sheffield Hallam University
  - Specialization: Data Science
- Fatemeh Bagheri
  - [f.bagheri13@gmail.com](mailto:f.bagheri13@gmail.com)
  - France
  - Université Jean Monnet St Etienne - Université de Lyon
  - Your Specialization: Data Science

Name: Customer Segmentation

Report date: 18 July 2022

Internship Batch: LISUM10

Version: 1.0

Data intake by: Ray Ng

Data intake reviewer:

Data storage location:

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

### Tabular data details:

Total number of observations	1000000
Total number of files	1
Total number of features	48
Base format of the file	.csv

Size of the data	154 MB
------------------	--------

**Note: Replicate same table with file name if you have more than one file.**

### **Proposed Approach:**

**Problem: Customer Segmentation** - XYZ bank wants to roll out Christmas offers to their customers. But the bank does not want to roll out the same offer to all customers, instead they want to roll out personalized offers to particular sets of customers. If they manually start understanding the category of customer then this will not be efficient and also they will not be able to uncover the hidden pattern in the data ( pattern which groups certain kinds of customer in one category). Bank approached ABC analytics company to solve their problem. Bank also shared information with ABC analytics that they don't want more than 5 groups as this will be inefficient for their campaign.

**Business Understanding:** ABC analytics should try to understand certain patterns in customer behaviour, such as: more loyal customers tend to...; depending on customer type and customer relation type do customers have more funds, pensions, loans, etc.; and how has each customer's financial behaviour changed from 2015-01-28 to 2015-02-28. ABC must find at most five groups in which customers share common behaviors.

### **Project life cycle along with deadlines:**

S/N	Project Task	Deadlines
1	Business Understanding	19th July 2022
2	Data Understanding(Type of Data and Problems in the data)	26th July 2022
3	Data Cleansing and Transformation	2nd August 2022
4	Exploratory Data Analysis (This should include the EDA recommendations as well)	9th August 2022

5	EDA Recommendation (ppt and proposed modelling technique)	16th August 2022
6	Model Selection and Model Building	23rd August 2022
7	Final Project Report and Code	30th August 2022