**BI MANAGER TECHNICAL TEST - SYSTEM DOCUMENTATION**

The system is designed to read a CSV file sent in on a monthly basis. The file is intended to be loaded in a pre-existing database with multiple tables. Python being the key programming language for this solution, the file is expected to be broken down and loaded in multiple tables under one schema IBRD\_UG.

After the file loading process is done and the file is complete, a number of results set in Excel workbooks are created in the directory “OUTPUT” in the home path of the source code.

The following results were selected to be output after the file is processed:

* Average for original Principal amount
* Average for cancelled amount
* Average for undisbursed amount
* Average for disbursed amount
* Average for Repaid To IBRD
* Average for Due To IBRD
* Average for Exchange Adjustment
* Average for Sold 3rd Party
* Average for Repaid 3rd Party
* Average for due 3rd Party
* Average for Loans Held
* All Loan types submitted in current raw file
* Loans taken by each Country
* Count for all Loan statuses in file
* Missing Values: Count for loan numbers without guarantor
* Missing Values: Count for loan numbers without borrower name
* Maximum amount taken by a country
* Minimum amount taken by a country

**TOOLS AND RESOURCES USED:**

* Python 3.8
* PG Admin 4
* PostgresSQL (For its high-level large data processing for data modelling and data science)
* SQL WorkBench

Experian being in the financial sector, the selected data is key in multiple ways.

We can use such data results and much more when the service scales up to determine the borrowing behavior of a client by analyzing how much thy borrow, how frequently they borrow, how much they borrow, how often do they default and by how much and so much more. With this data in our hands we can develop tools and systems that assist our users i.e. Banks and other financial institutions to better understand who they lend to.

Additionally, we can use said data to educate and help potential borrowers to better manage credit facilities there by creating a sustainable borrowing habit that suits their needs including understanding potential risks for both borrowers and the lender

**Current and Future Technical specs for the system.**

The system is expected to encrypt emails when notifying the data vendor that the file is processed. I stuck to and would still recommend an SFTP server to move such files for they are very confidential and email servers are hard to monitor or control especially from the vendor’s side

The system will have more notification functionality to improve error reporting and strict data validation. For example, the raw file completely has no CURRENCY supplied which makes the amounts unclear to any 3rd party who would be interested in understanding the results from the file.

GitHub repository link:

<https://github.com/yosightpro/PAI_RUN>