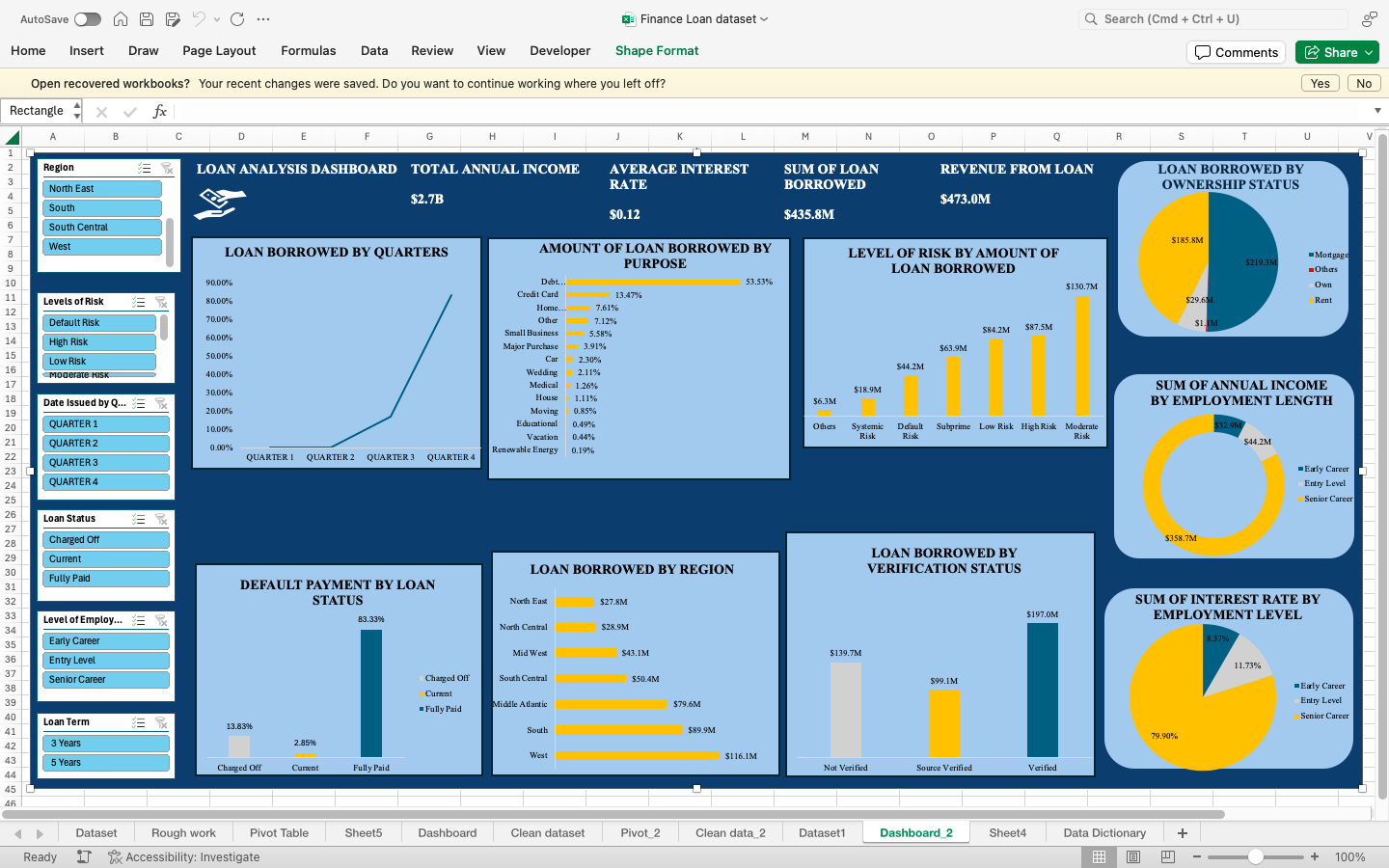
**The Classic Data Analysis by Oguanyia Cynthia**



The Classic Data provides a comprehensive report on a loan dataset that contains the loan's purpose, interest rate, duration of employment, loan status, and degree of risk. Through an evaluation of borrower profiles, loan purposes, risk levels, and geographical performance, this analysis seeks to optimise the loan portfolio. In order to inform strategic decision-making, the project focusses on comprehending the distribution of loan amounts, interest rates, risk categories, and payback performance. Examples of this include identifying high-risk locations and refining loan assessment processes to lower default rates. maximising income by focussing on lucrative markets, such as senior career professionals and areas with strong loan demand.   
The project's ultimate goal is to increase the loan portfolio's general health while maximising profitability and lowering risk.

**Introduction**

The classic Data Analysis is an Internship Café with some experience in Microsoft offices. It was created under the supervision of the great “The Data Immersed (TDI)”, a financial services company specializing in providing loans to individuals and small businesses to aid in meeting up with some necessary basic needs at a particular point in time. The datasets provides information on the interest rate for the loan borrowed, the length of employment, ownership status and the loan approval status, this dataset was provided by The Data Immersed (TDI).

In this report, useful insights will be drawn by improving loan verification processes, identifying high-risk areas, and maximizing revenue using Microsoft excel.

**Problem Statement**

A sizeable amount of the loan portfolio is devoted to credit card payments and debt consolidation, with a high concentration of loans in riskier categories. The existing distribution significantly favours moderate- and high-risk borrowers, which may raise the possibility of defaults notwithstanding possibilities in other loan purposes and locations. Furthermore, there is a discrepancy in the loan verification procedures; borrowers who are not confirmed are granted substantially smaller loan amounts. Moreover, under-representation exists in some areas and at some employment levels, which may restrict opportunities for growth and diversification.   
To increase profitability and lower default rates, the challenge is to optimise the loan portfolio by lowering risk exposure, enhancing credit verification, broadening the range of loans offered, and focussing on underserved market segments.

**Data Overview**

The dataset offers a thorough picture of the loan portfolio by emphasising borrower profiles, regional distribution, risk tolerance, and lending reasons. Verified and senior career borrowers make up the most profitable sectors of the portfolio, which is strongly concentrated in debt-related loans and exposed to high-risk borrowers. The Classic Data consists of multiple columns, such as:

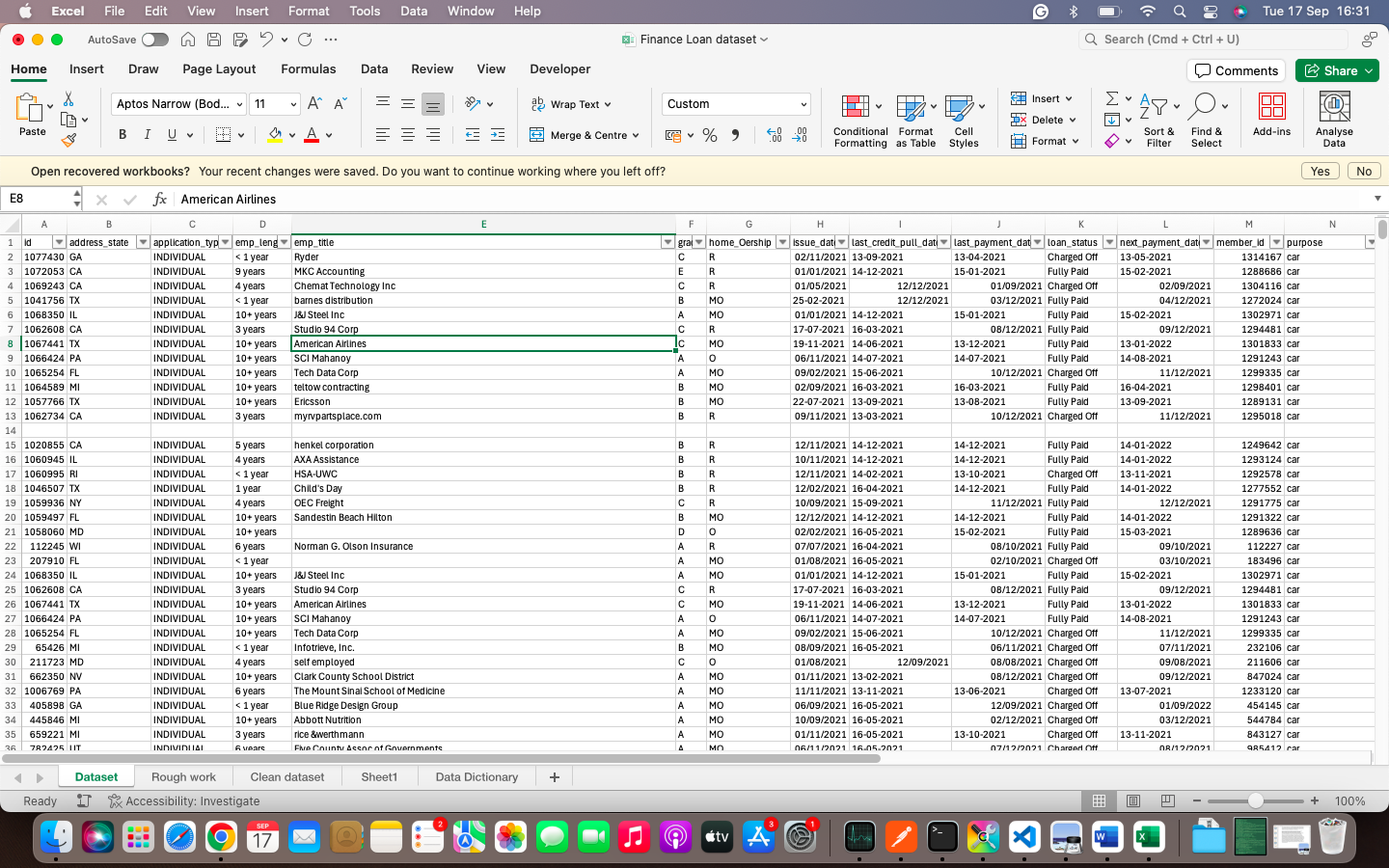
1. **Total Annual Income**: The dashboard shows **$2.7B** in total annual income, indicating the revenue generated from loan repayments across different borrower segments.
2. **Average Interest Rate**: The average interest rate is **12%**, representing the cost of borrowing for customers. This rate likely varies across different risk levels and loan types.
3. **Sum of Loan Borrowed**: A total of **$435.8M** has been borrowed across various purposes, regions, and risk levels.
4. **Revenue from Loans**: The revenue generated from loans is **$473M**,which could include interest payments, fees, and other associated loan income.
5. **Loan Distribution by Purpose**: **53.53%** of loans are for debt consolidation. Other significant categories include credit card repayment **(13.47%)**, home improvement **(7.16%)**, and small business loans **(5.8%)**.
6. **Risk Levels by Loan Amount**: Loans categorized under moderate risk and high risk account for the largest shares, with **$130.7M** and **$87.5M** respectively. Lower risk and subprime borrowers contribute smaller amounts, highlighting risk concentration.
7. **Loan Verification Status**: Fully verified borrowers receive the highest loan amount of **$197.0M**, followed by source verified (**$99.1M**), and non-verified borrowers (**$139.7M**).
8. **Employment and Loan Characteristics**: Borrowers in senior career positions contribute **$358.7M** to the sum of annual income and pay **79.9%** of the total interest rate, making them a highly valuable segment.
9. **Loan Status and Default Rates**: Fully paid loans have a high repayment success rate (83.33%), while charged-off loans show a significant default rate (13.83%).

**Data Transformation**

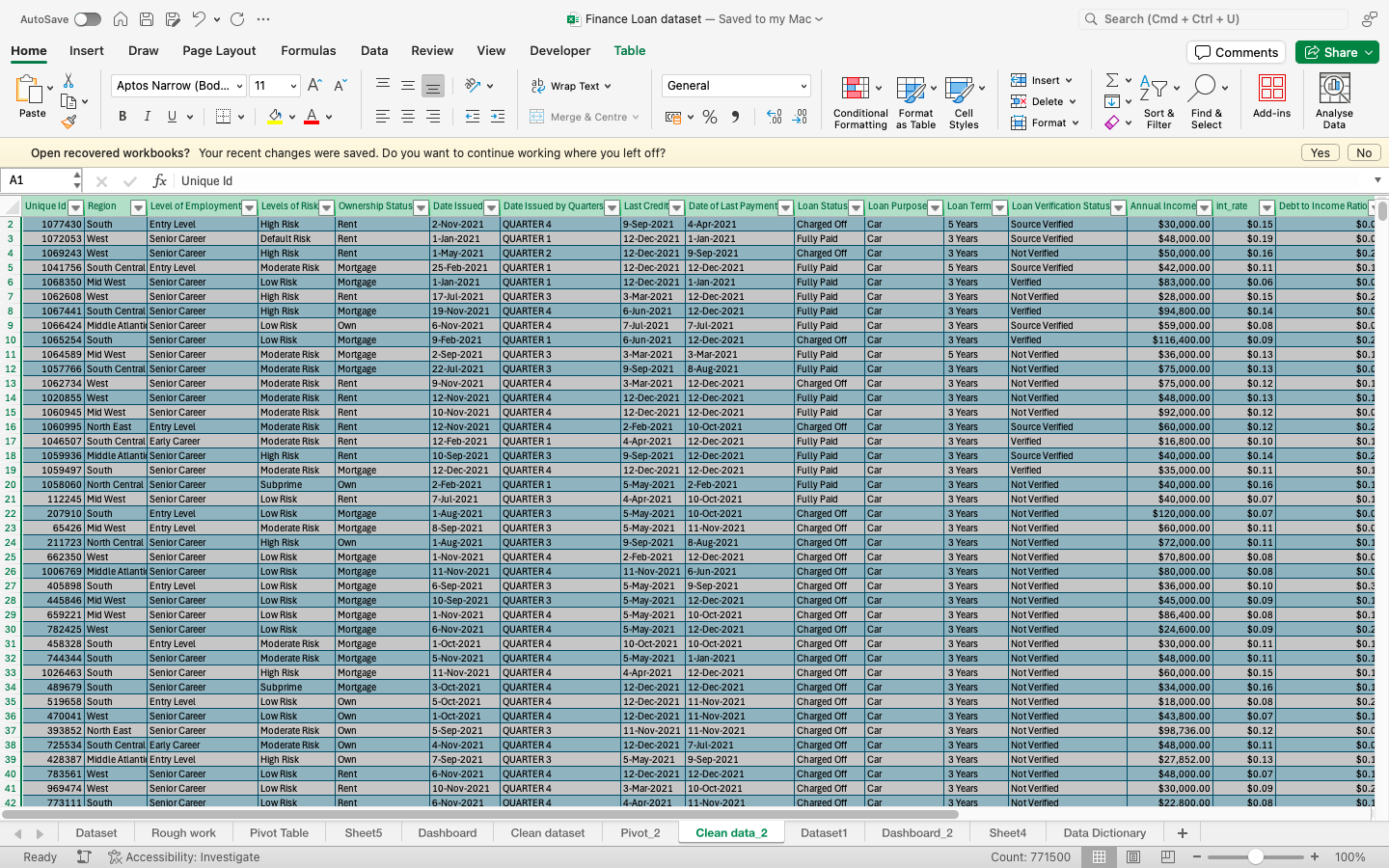
The dataset was downloaded from the zip file provided by TDI and was load in excel where the cleaning was done following these steps:

* **Managing Missing Values**: Relevant placeholders were used in place of missing values.
* **Verifying for Duplicate Values**: To guarantee the highest output, duplicate values were eliminated from the dataset.
* **Deleting Empty Rows and Columns**: Rows and columns that contained no data were eliminated.
* **Fixing Inconsistent Data:** To guarantee consistency, standardise date formats and other field formats.
* **Data Type Conversion:** Verify that fields, like dates, are in the appropriate data type (e.g., convert dates to datetime format).
* **Filtering Irrelevant Data**: Discarded any columns or rows that weren't required for the analysis.
* **Recording Data Cleaning Steps:** For transparency and repeatability, keep a thorough record of all data cleaning procedures.
* **Renaming Columns to Meaningful Names**: Used descriptive names for tables and columns to enhance clarity.

**The dataset before cleaning**



**The dataset after cleaning**



**Key Performance Indicators (KPIs)**

KPIs were added to tell us how well loan business is doing. In our analysis performance was based on key metrics:

* **Total Annual Income: $2.7B** Represents the total income generated from all loans, reflecting the revenue from borrower repayments, including interest and fees.
* **Average Interest Rate: 12%** The overall average interest rate charged across the loan portfolio. This rate influences profitability and borrower affordability.
* **Sum of Loan Borrowed: $435.8M** The total amount of money borrowed across all loans, indicating the size of the loan portfolio.
* **Revenue from Loans: $473M** The total revenue earned from the loans, highlighting the financial return from the portfolio, which may include interest, fees, and other charges.

**Recommendation**

Based on the visual data in this loan dashboard, here are some key recommendations:

1. **Diversify Loan Portfolio to Reduce Risk:** The "Level of Risk by Amount of Loan Borrowed" chart shows a significant amount of loans classified under high and moderate risk categories. Diversifying the loan portfolio by encouraging more low-risk loans may reduce potential defaults and increase financial stability.
2. **Focus on Growing Specific Loan Purposes:** A large portion of loans is being borrowed for debt (53.53%) and credit card repayment (13.47%). Diversifying into smaller categories like renewable energy, education, and medical loans could promote long-term value creation, and possibly lower risk.
3. **Target High-Income Borrowers in Specific Regions:** The "Loan Borrowed by Region" chart indicates that the South and West regions have the largest loan amounts. Since these regions show high borrowing power, consider offering tailored loan products or incentives to these areas to expand market share.
4. **Verification Impact on Loan Amounts:** The "Loan Borrowed by Verification Status" chart shows that source-verified and fully verified loans significantly exceed those that are not verified. Encouraging borrowers to complete verification can lead to larger loan amounts while also managing default risks.

**Conclusion**

In summary, the dashboard reveals a portfolio with a significant share of high-risk loans but also shows strong opportunities for growth, particularly by focusing on verified borrowers, senior career professionals, and geographical regions with high loan demand.