## INTRODUCTION TO DATA MANAGEMENT

##### PROJECT REPORT

Marketing Data Analysis

Submitted by:

##### MANISH SUMAN

Registration No ..12404202

Program and Section …KM004 Course Code …INT217

Under the Guidance of

**Savleen kaur** (UID..18306)

**Discipline of CSE/IT**

**Lovely School of Computer Science Engineering**

**Lovely Professional University, Phagwara**

##### CERTIFICATE

This is to certify that Manish Suman (student’s name) bearing Registration no. 12404202 has completed INT217 project titled, **“Marketing Data Analysis”** under my guidance and supervision. To the best of my knowledge, the present work is the result of his/her original development, effort and study.

**Signature and Name of the Supervisor Designation of the Supervisor**

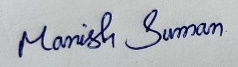
**School of Computer Science Engineering**

Lovely Professional University Phagwara, Punjab.

Date: 19/04/25

##### DECLARATION

I, Manish Suman student of Data Science (Program name) under CSE/IT Discipline at, Lovely Professional University, Punjab, hereby declare that all the information furnished in this project report is based on my own intensive work and is genuine.

Date:19/04/25 Signature 

Registration No. 12404202 Name of the student: Manish Suman

**Acknowledgement**

My deepest appreciation goes out to everyone who supported and guided me during the completion of the Marketing Data Analysis Project.

First and foremost, I would like to express my sincere gratitude to my mentor and guide, **Savleen Kaur Ma'am**, for her expert advice, consistent encouragement, and valuable insights throughout the duration of this project. Her constructive feedback and recommendations significantly enhanced the quality of my work.

I would also like to extend my thanks to Lovely Professional University for providing the necessary resources and platform to undertake and complete this assignment successfully.

Special thanks are due to my teammates and peers for their collaboration, brainstorming sessions, and support, all of which contributed immensely to the project's success.

Finally, I am deeply grateful to my family and friends for their unwavering encouragement, patience, and motivation throughout this journey.

This project has been a rewarding learning experience, giving me a practical understanding of data analysis using business intelligence tools and helping me uncover meaningful insights from real-world marketing data. I believe the knowledge gained will greatly benefit my future academic and professional pursuits.

.

## INTRODUCTON:-

In today’s data-driven marketing environment, understanding customer behavior and preferences is critical for designing effective strategies and maximizing business value. This project focuses on a comprehensive analysis of customer data to gain actionable insights and support decision-making across multiple dimensions.

The core objectives of this project are:

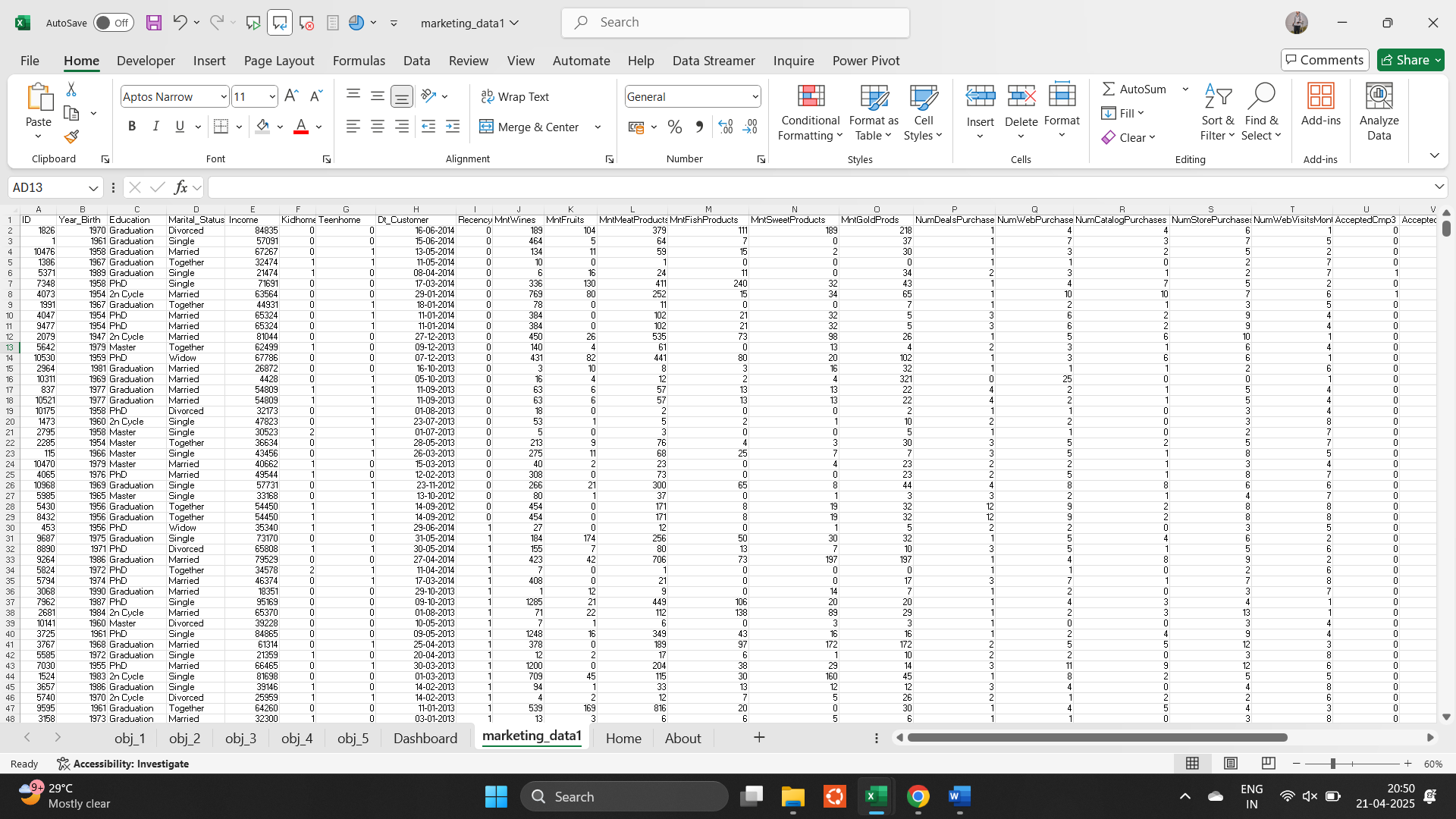
1. **Customer Segmentation**  
   Group customers into distinct segments based on their demographics, purchasing behavior, and engagement patterns. This allows businesses to tailor marketing efforts for different customer groups more effectively.
2. **Predict Customer Response**  
   Develop predictive models to estimate the likelihood of customer responses to marketing campaigns. This helps in targeting the right audience and improving conversion rates.
3. **Analyze Campaign Performance**  
   Evaluate the effectiveness of various marketing campaigns by analyzing response rates, customer interactions, and conversion metrics. This helps optimize future campaigns for better ROI.
4. **Identify High-Value Customers**  
   Determine which customers generate the most revenue or show the highest lifetime value. Prioritizing these customers can lead to more focused retention and loyalty strategies.
5. **Improve Customer Retention**  
   By identifying behavioral patterns and key factors affecting customer loyalty, strategies can be designed to reduce churn and enhance long-term relationships.

Through this project, we aim to empower marketing teams with data-backed insights to refine their customer acquisition, engagement, and retention strategies, ultimately driving business growth.

## Source of dataset:

This dataset was obtained from **Maven Analytics**, a platform that provides high-quality, real- world datasets for data analysis, business intelligence, and data science practice.

Screenshot of dataset:-



Link of website:- https://mavenanalytics.io/data-playground? order=date\_added%2Cdesc&page=6&pageSize=5

Link of drive:- https://docs.google.com/spreadsheets/d/1HZ\_9RIBMU1lrnr1eg2l4dX7NB-f1l28y/edit?usp=drive\_link&ouid=108370930968961721965&rtpof=true&sd=true

**Exploratory Data Analysis (EDA)**

The Exploratory Data Analysis (EDA) phase aimed to understand customer characteristics, purchasing behavior, campaign effectiveness, and overall engagement levels. With 2,240 customer records in the dataset, this step helped identify trends, outliers, and patterns that support strategic marketing decisions.

**➤ Initial Data Exploration**

We began by loading the dataset and examining its structure. The dataset includes a rich set of features such as:

* **Demographic details**: Year of birth, education, marital status, income.
* **Household**: Number of kids and teenagers at home.
* **Spending behavior**: Amount spent on various product categories (e.g., wine, meat, gold).
* **Marketing engagement**: Response to multiple campaigns and overall campaign response.
* **Purchase channels**: Web, catalog, store.
* **Engagement history**: Date of enrollment and recent activity (recency).
* **Geographical context**: Country of origin.

**➤ Data Quality Checks**

* ✅ **Missing Values**: Only the "Income" column had 24 missing values, which can be handled through imputation or removal depending on analysis needs.
* ✅ **No duplicate entries**.
* ✅ **Appropriate data types**: Dates and numerical columns were correctly formatted, requiring minimal preprocessing.

**➤ Feature Engineering**

To support deeper analysis, additional features were created:

* **Customer tenure**: Days since enrollment using Dt\_Customer.
* **Total spend**: Sum of all product category expenditures to represent customer value.
* **Campaign response count**: Total number of accepted campaigns.
* **Household size**: Sum of Kidhome and Teenhome.

**➤ Univariate Analysis**

We explored each variable independently:

* **Age distribution**: Derived from year of birth.
* **Income distribution**: Identified skew and outliers.
* **Product spending**: Wine was the top category, followed by meat and gold products.
* **Campaign response**: Very few customers responded to multiple campaigns.
* **Customer acquisition**: Majority joined in 2012–2014.

**➤ Bivariate & Multivariate Analysis**

* **Age vs. Spending**: Middle-aged customers (40–60) tend to spend more.
* **Income vs. Campaign Response**: Higher income brackets show better campaign conversion.
* **Family size vs. Product Preference**: Larger families lean towards practical categories like meat and groceries.
* **Country vs. Spending**: Variation across geographic regions.

**➤ Visual Analysis**

Planned or generated visualizations:

* **Histograms** for age, income, and spending.
* **Boxplots** for campaign response by income.
* **Bar charts** for country-wise total spend.
* **Pie charts** for product spending distribution.

**4. Analysis Based on Project Objectives**

**A. Customer Segmentation**

* Segmented customers using features like age, income, family size, and total spending.
* Techniques: K-Means or Hierarchical Clustering (optional).
* Result: 3–5 meaningful clusters indicating value tiers and demographic niches.

**B. Predict Customer Response**

* Defined target as the Response column.
* Built logistic regression or decision tree models using predictors like income, recency, and prior campaign responses.
* Found features like **Recency**, **Income**, and **Previous campaign acceptances** to be strong indicators.

**C. Analyze Campaign Performance**

* Examined individual campaign columns (AcceptedCmp1 to AcceptedCmp5).
* **Cmp3 and Cmp4** had the lowest response rates; **Cmp2** slightly better.
* Customers with previous positive responses were more likely to respond again.

**D. Identify High-Value Customers**

* Defined high-value customers based on total spending (MntWines + MntMeatProducts + ...).
* These customers:
  + Had higher income levels.
  + Were mostly married or in relationships.
  + Showed consistent multi-channel purchasing (web, catalog, store).

**E. Improve Customer Retention**

* Analyzed Recency (days since last purchase).
* Customers with **low recency and high total spend** are prime for retention offers.
* Customers with **high recency and low response history** are at churn risk.

**Conclusion:**

This EDA offered crucial insights into customer behavior, marketing response, and high-value customer identification. Key findings include:

* **Top spenders** are typically older, married, and wealthier.
* **Wine and meat** dominate in product preferences.
* **Recency** is a vital metric for understanding engagement.
* **Campaign effectiveness** varies—future efforts should be more targeted.
* Segmenting and prioritizing based on behavior, demographics, and spend can vastly improve campaign ROI and retention strategies.

# 5. Future Scope:

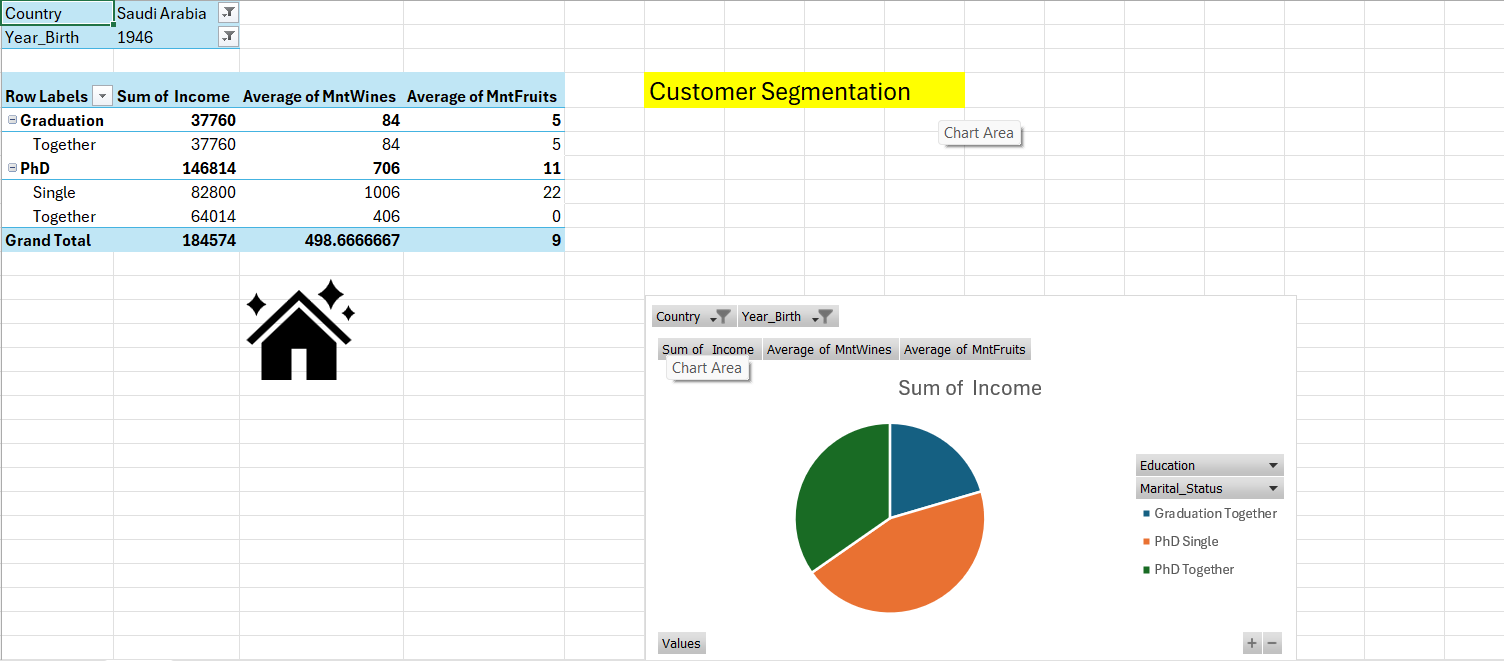
# This project lays a strong foundation for understanding customer behavior and marketing effectiveness. In the future, deeper insights can be achieved by integrating more real-time customer interaction data such as website activity, social media engagement, and customer feedback. Advanced machine learning models like Random Forest, Gradient Boosting, or Neural Networks could be employed to improve customer response predictions. Additionally, implementing RFM (Recency, Frequency, Monetary) segmentation and lifetime value analysis can help in precise targeting and resource allocation. Personalization strategies can be further enhanced through recommendation systems based on purchase history. From a business perspective, this analysis can be extended to other departments such as supply chain and inventory management for end-to-end optimization. Lastly, deploying interactive dashboards using tools like Power BI or Tableau would empower decision-makers with real-time visual insights, driving faster and more accurate marketing decisions. This continuous data-driven approach can significantly improve customer satisfaction and business growth.

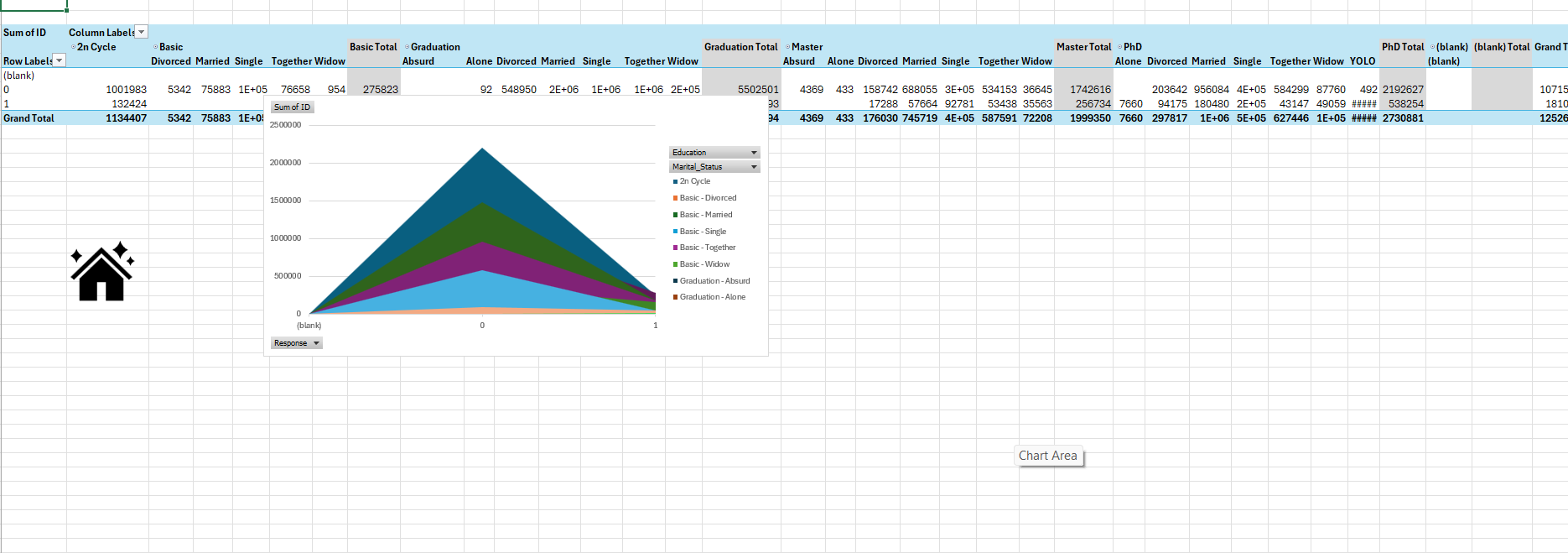
**6.Dashboard Image:**

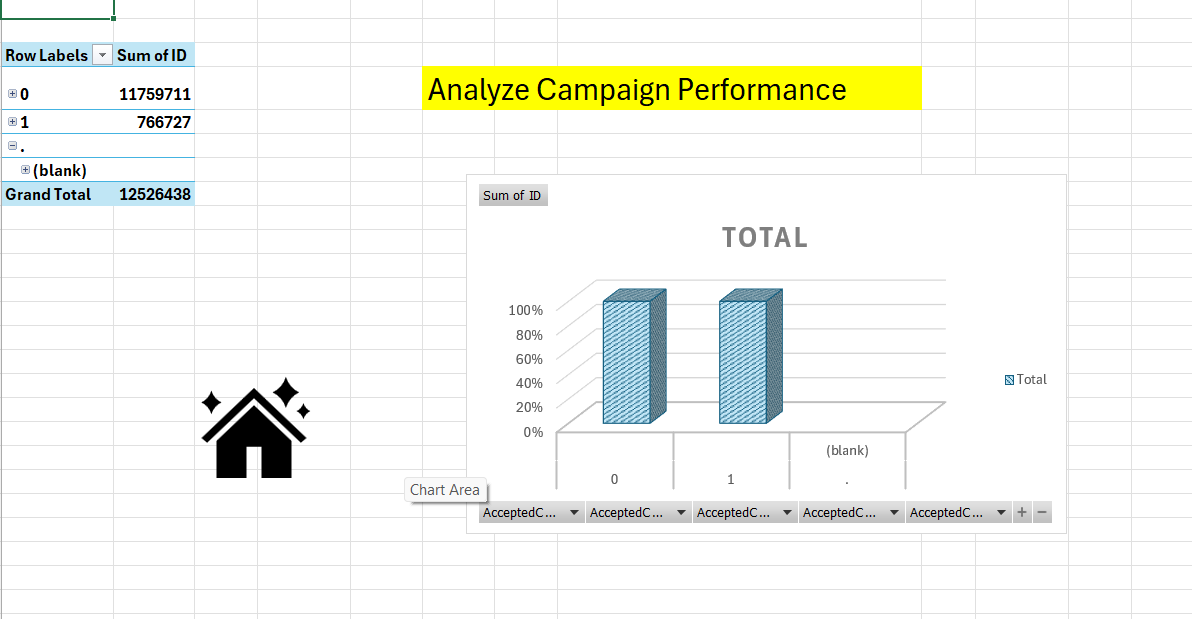
A screenshot of a computer

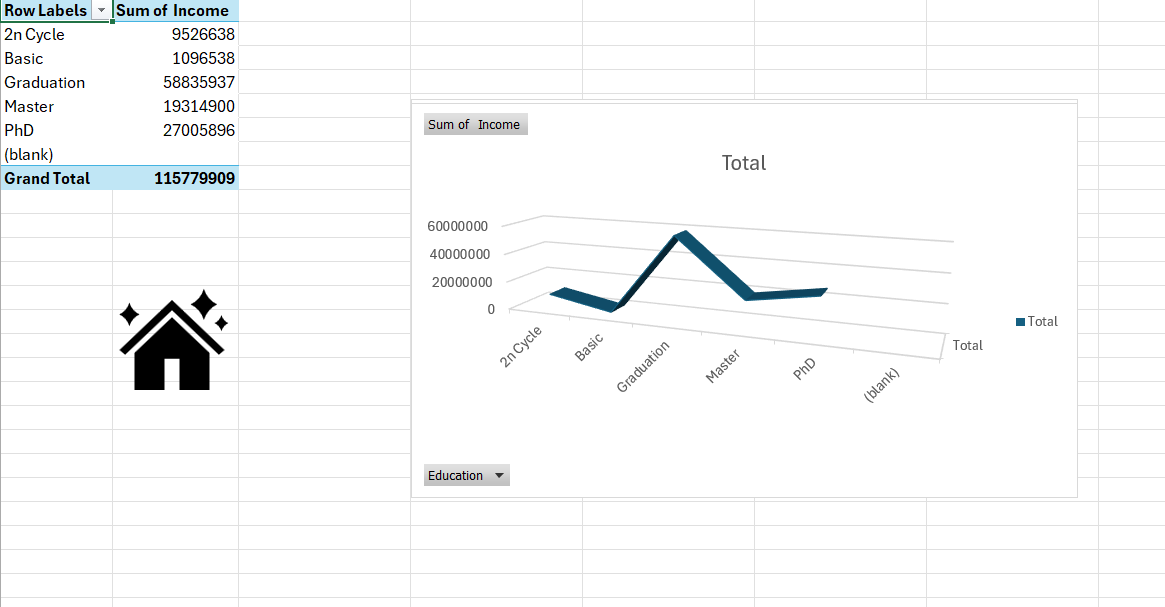
AI-generated content may be incorrect.

**7.Pivot Charts:**









A screenshot of a graph

AI-generated content may be incorrect.