Personal Finance Analysis Using Linear Regression in Power-BI

# Saleti Sumalatha

*Assistant Professor*

*Computer Science and Engineering SRM University AP* [sumalatha.s@srmap.edu.in](mailto:sumalatha.s@srmap.edu.in)

# *Lakshmi Sai Prakash Paladugu*

*Computer Science and Engineering SRM University AP saiprakash\_p@srmap.edu.in*

# *Sricharan Gudi*

*Computer Science and Engineering SRM University AP* *sricharan\_gudi@srmap.edu.in*

# *Sai Gopal Lanka*

*Computer Science and Engineering SRM University AP gopal\_lanka@srmap.edu.in*

# *Mounish Sai Mididodla*

*Computer Science and Engineering SRM University AP mounishsai\_m@srmap.edu.in*

***Abstract*—The project aims to conduct a comprehensive analysis of personal finances using Power BI as the visualization tool. Through this endeavor, we seek to provide valuable insights into individual financial habits, trends, and potential areas for improvement. By leveraging the power of data analytics, we aim to empower individuals to make informed decisions about their finances, ultimately leading to greater financial stability and well-being.**

**Keywords:- Power BI, Data analytics, Linear regression.**

1. Introduction

The problem of personal finance management is pervasive and affects individuals across all demographics. In today's fast-paced world, managing personal finances effectively has become increasingly challenging due to the multitude of financial instruments, expenses, and investment options available. Moreover, the lack of financial literacy among many individuals further exacerbates this issue.

Data science offers a powerful solution to this problem by enabling the analysis of large datasets comprising financial transactions, expenses, income sources, and investment portfolios. By applying various statistical and machine learning techniques, we can uncover hidden patterns, correlations, and insights within the data. These insights can then be translated into actionable recommendations for individuals to optimize their financial strategies.

Power BI serves as an excellent tool for visualizing the results of our data analysis. Its intuitive interface and interactive dashboards allow users to explore complex financial data effortlessly. Through visually compelling charts, graphs, and reports, we can effectively communicate key findings and empower individuals to take control of their financial futures.

1. Literature Review

The literature review for this project on personal finance analysis and expense tracking using Power BI underscores the critical importance of comprehending and effectively managing individual finances. It emphasizes that achieving financial literacy and stability is not only essential for personal well-being but also has a profound impact on overall economic health. By leveraging advanced tools like Power BI, individuals can delve deep into their spending habits, understand their financial behavior, and subsequently, develop more effective budgeting strategies. This enables them to make well-informed financial decisions tailored to their unique circumstances. Moreover, the review highlights how technology, particularly platforms like Power BI, plays a pivotal role in advancing personal finance management. These tools offer sophisticated analytics and visualization capabilities, empowering users to gain actionable insights from their financial data. However, it also acknowledges the importance of addressing challenges such as data security and privacy concerns, ensuring that individuals feel confident and secure in leveraging technology for managing their finances. Ultimately, the overarching goal is to promote financial well-being and foster economic resilience at both individual and societal levels

**Proposed Methodology :**

1. Pre-processing:

In this step, we import the dataset, examine its characteristics, look for null values, and determine the kinds of values that were included.The dataset has a 12 by 39 shape.None of the dataset's values are null.It includes monthly income for a maximum of two years as columns which should be changed to rows.The dataset consists of six expenses rows (food, housing rent, groceries, gas, shopping, etc.), four savings rows (funds, fixed deposits, etc.), and two income rows (salary, source 2).The dataset needs to be changed right now so that it fits into the linear regression.

Next data is to merge various sources of income, savings, and expenses into a useful dataset.The salary and sources are first chosen, then they are combined into a single row and given the name Total\_income. Subsequently, the funds, fixed deposits, and additional rows pertaining to savings are aggregated and transformed into total\_savings. This process is repeated for the expenses.We now remove the last few rows.The rows will then be converted into columns.At last, we possess three distinct columns that are suitable for performing regression analysis.

III. Dataset Description

The dataset employed in the personal finance dashboard project is a detailed record of an individual’s financial transactions spanning three years, from January 2018 to January 2021. The data are structured monthly and encompass various components critical for personal financial planning and analysis. These components are categorized into Income, Savings, and Expenses, with each category further divided into specific types as described below:

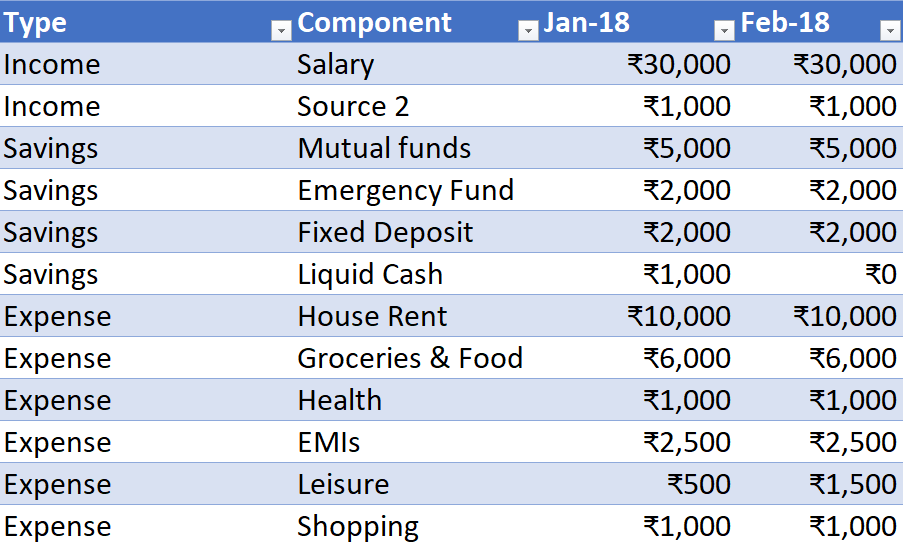


Fig. 1. Dataset.

**Income:**

1**. Salary:** This is the primary source of income, recorded monthly. There are observed increases in salary at certain intervals, reflecting possible promotions or pay raises.

2. **Source 2:** Represents additional income from a secondary source, which also shows variations over the study period, indicating possible changes in additional earnings or side income.

**Savings:**

1. Mutual Funds: Monthly contributions to mutual funds, with noticeable adjustments in the contribution amount during the period.

2. **Emergency Fund**: Represents allocations towards an emergency fund, showing fluctuations that might correspond to varying monthly savings capabilities or emergency withdrawals.

3. **Fixed Deposit**: Monthly amounts allocated to fixed deposits, with some months showing no deposits.

4. Liquid Cash: Tracks cash liquidity, showing both deposits and withdrawals, reflecting the ebb and flow of cash on hand.

**Expenses:**

1. **House Rent**: Regular payments with periodic increases, possibly reflecting rent hikes.

2. **Groceries & Food**: Monthly expenses that show variability, potentially indicating changes in market prices or consumption patterns.

3. **Health**: Expenses related to health and wellness, with notable variations possibly due to unforeseen medical expenses or changes in health insurance costs.

4. **EMIs:** Payments towards liabilities or loans, which increase over time, potentially due to additional loans or changes in interest rates.

5. **Leisure:** Discretionary spending on leisure activities, showing variation likely reflecting lifestyle changes.

6. **Shopping**: Expenditure on shopping, with some variability, possibly indicative of non-regular purchasing patterns or discretionary spending controls.

This dataset provides a comprehensive overview of the financial behavior of an individual over three years, capturing the dynamics of income streams, savings habits, and expenditure patterns. The granularity of the data supports detailed analysis and visualization in a personal finance dashboard, enabling insights into financial health, spending trends, and potential areas for financial optimization. This rich dataset is pivotal for the project, allowing the development of tailored financial strategies based on historical financial behavior patterns.

IV. ALGORITHM

In general, we have used linear regression to predict income,expenses or savings . Linear regression can only give few insights about the persons personal finance as it includes other relations between different features .So we have created a dashboard based on the clients adjustments .Here we load the dataset and implement some formlaes to create the visually appalling insights which can be understood easily.

V. EXPLORATORY DATA ANALYSIS  
  
In the course of our research, we leveraged Microsoft Power BI, a leading business intelligence tool, to conduct exploratory data analysis (EDA) on a comprehensive personal finance dataset. This dataset detailed the financial transactions of an individual over a three-year period, encompassing diverse aspects such as income, expenses, and savings across multiple categories.

Power BI is an analytics platform that offers extensive capabilities for data integration, processing, and visualization. It enables users to import data from various sources, transform and clean the data using its powerful query editor, and then create a suite of dynamic reports and dashboards. These features allow for deep dives into the data, supporting both granular analyses at the transaction level and high-level overviews. Power BI's real-time update capability and interactive visualizations significantly enhance the analytic experience, promoting an intuitive understanding of complex datasets.

Exploratory Data Analysis represents a critical initial phase in data analysis where data is summarized and examined to uncover patterns, spot anomalies, test hypotheses, and check assumptions through statistical graphics and other data visualization methods. EDA is pivotal for gaining preliminary insights before formal modeling commences, guiding the direction of further analysis.

In our project, we applied EDA using Power BI to:

Visualize Trends: Through Power BI, we visualized income trends, expense distributions, and savings patterns over time, employing a variety of charts and graphs that elucidate changes and trends within the financial data.

Identify Variabilities: By creating interactive visualizations, we analyzed the variability in expenditures and income over different periods, helping us understand volatility and the typical financial cycle of the individual.

Generate Summary Statistics: Power BI facilitated the computation of descriptive statistics that provided insights into the central tendencies and dispersions of the financial figures, crucial for financial health assessment.

VI RESULTANT GRAPHS

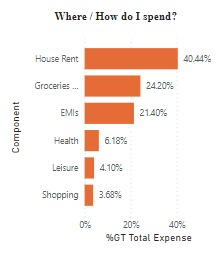


Fig. 2. Bar Graph of Expenditures

Figure 2 presents a bar graph illustrating the distribution of expenditure across various categories within the personal finance dataset. This visualization effectively highlights the proportional spending in each category as a percentage of the total expenditures over the three-year period covered by the dataset.

Total Expense = CALCULATE(SUM(FinData[Value]),FinData[Type]="Expense")

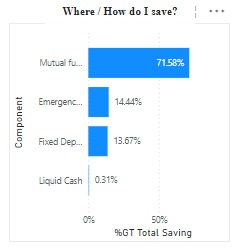


Fig. 3. Bar Graph of Savings

The individual's entire savings are shown in graph 2, with mutual funds ranking highest at 71.58 percent, emergency funds coming in second at 14.44 percent, fixed deposit, and liquid cash in order of preference.The formula "Total Saving =CALCULATE(SUM(FinData[Value]),FinData[Type]="Savings")"is used to obtain the graph.

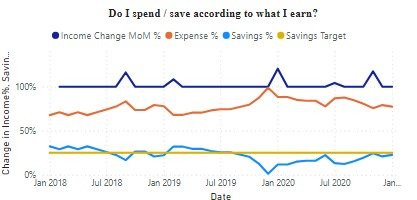


Fig. 4. Comparison of Savings and Expenditure with respective Income

Figure 4 represents the Income change of month on month and change in expenditure behavior with respect to the change in income. It also shows the change in savings behavior with respect to the change in income and also shows a line which represents the savings target that needs to be meet each month based on user's preference.

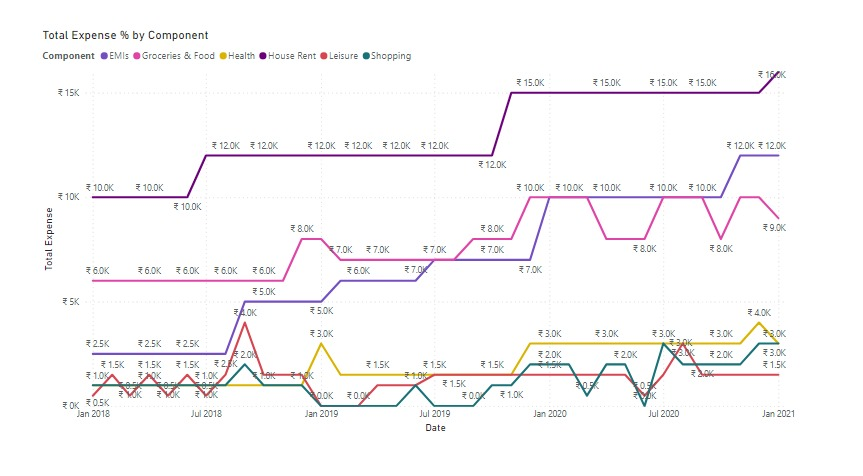


Fig. 5. Total Expenditure of Each Component

The graph shows the total expense by component over the past year. The components are EMIs, Groceries & Food, Health, House Rent, Leisure, Shopping. The total expense seems to be around ₹15,000 per month. Groceries & Food to be the biggest expense, costing around ₹3,000 per month. House Rent is the second biggest expense, costing around ₹2,500 per month. It is difficult to say which month was the most expensive but December and June seem to be the most expensive months as the total expense goes upto ₹16,000.



Fig. 6. Total Saving of Each Component

The total savings seems to fluctuate throughout the timeframe. It starts at around ₹7,000 in Jan 2018, increases to around ₹20,000 in Jul 2018, then decreases again to around ₹7,000 in Jan 2019. It increases again to around ₹20,000 in Jul 2020 and stays around that level until Jan 2021.

Emergency Fund seems to have the most money allocated to it, with a balance of around ₹70,000 throughout the timeframe.

Mutual Funds seems to have the least amount of money allocated to it, with a balance that fluctuates between ₹0 and ₹2,000.

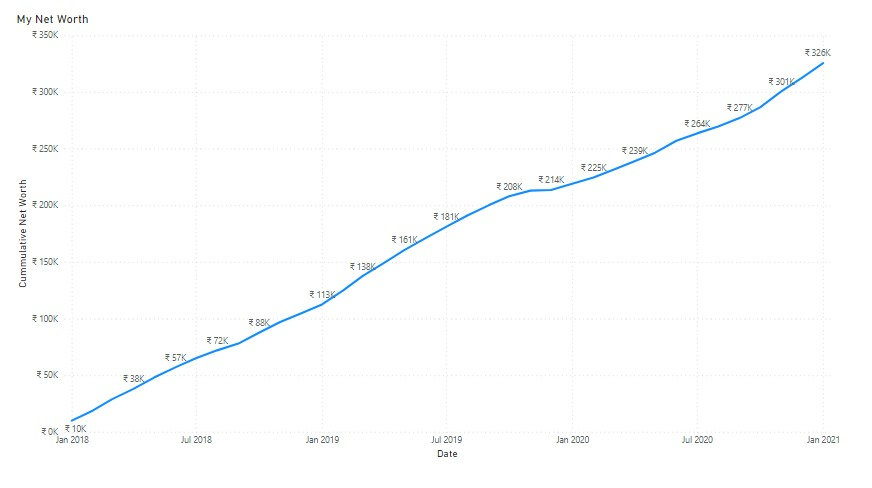
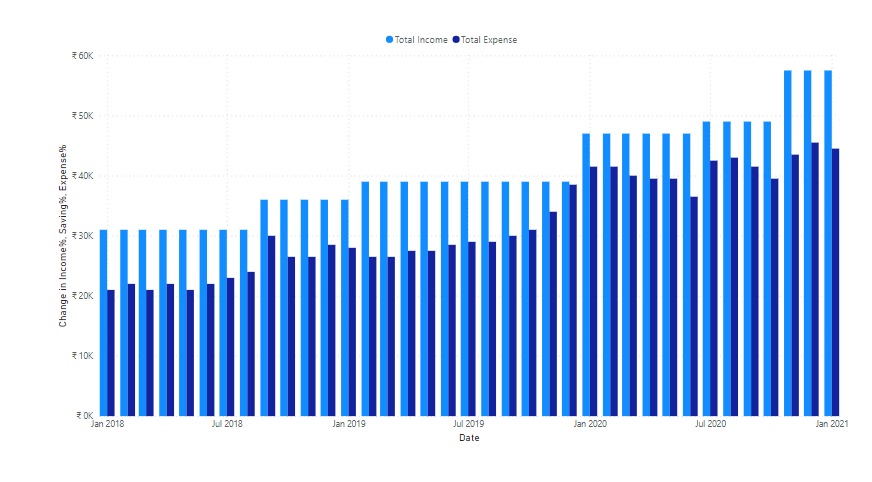


Fig. 7. Total Net Worth

Fig 7 takes the historic data of the person's income and their net worth is calculated this is similar to the linear regression algorithm. A significant increase in net worth over the time period. Overall, the net worth has more than tripled over three and a half years. The net worth appears to have increased steadily over the time period shown in the graph.It is important to note that this graph only shows net worth, and does not show income or expenses. It is also possible that there were months where the net worth decreased but they are not shown in this graph.    
 Fig. 8. total income and total expenses against the change in income

VII.CONCLUSION

Finally, we would like to point out that various persons may have different financial planning styles and terminology, therefore we should employ a variety of visualization tools when conducting personal finance analyses rather than relying just on machine learning techniques.In order to get insights from the data, use the dash boards.

VIII. References

1. Hickok, S., and J. Hung 1992 Explaining the persistence of the U.S. trade

deficit in the late 1980s. Federal Reserve Bank of New York Quarterly

Review 16(4):29-46.

1. W. F. Sharpe (1978). Major investment styles. Journal of Portfolio Management, Winter: 68–75.
2. Bandura, A. (1977). Social learning theory. Englewood Cliffs, NJ: Prentice-Hall
3. Kindleberger, C.P., and D.B. Audretsch 1983 The Multinational Corporation

in the 1980s. Cambridge, Mass.: MIT Press.

1. Caves, R.E., J.A. Frankel, and R.W. Jones 1990 World Trade and Payments: An Introduction. Glenview, Ill.: Scott, Foresman and Company.
2. Financial Accounting Standards Board 1990a Disclosure of Information about Financial Instruments with Off-Balance-Sheet Risk and Financial Instruments with Concentrations of Credit Risk. Statement of Financial Accounting Standards No. 105. (March). Stamford, Conn.
3. Grubert, H., T. Goodspeed, and D. Wenson 1991 Explaining the Low Taxable Income of Foreign-Controlled Companies in the United States. (November) Unpublished paper, Tax Policy Group in the Office of International Taxation, U.S. Department of the Treasury.