

Digital Finance Manager

Meghana Avadhanam

INTRODUCTION

Managing personal finances can be a daunting task, particularly when it comes to tracking expenses. An expense tracker helps individuals monitor and control their spending habits by keeping a record of their expenses. The ability to understand expenses is critical for individuals and businesses to effectively manage their finances. However, financial data can be overwhelming and difficult to understand, especially for those who are not trained in finance or accounting. Visualizations can play a significant role in simplifying financial data and making it more accessible to a broader audience. This project aims to utilize visualizations to help users better understand their expenses. Visual analytics can be a powerful tool in helping individuals and families better understand their spending habits and make informed decisions about their finances. By analysing various factors such as stores, areas, and websites that are more expensive, users can identify areas where they can cut down on expenses and make changes to their spending habits. Moreover, visual analytics can provide an easy-to-understand view of financial data, making it more engaging and accessible to users. This can help users to stay motivated and engaged with their financial management and can provide a clearer picture of their overall financial health.

1 INSIGHT NEEDS

Insight needs that are supported in this problem would be categories/ clusters, orders, comparisons (or more). Answering questions like what time of the year is the highest expenditure? Or, over the past 10 months, what is the expenditure trend observed? Stakeholders may want to understand how much money is being spent, where it is being spent (geospatial), and on what it is being spent (categories of expenses). This information can help them identify areas where expenses can be reduced, optimized, or eliminated. Stakeholders may require insights on how actual spending compares to the budgeted amounts (finding gaps and outliers). This can help them identify overspending or under-spending in specific categories, departments, or teams. Insights into historical spending patterns and trends can help stakeholders make more informed decisions about future budgets and expenses.

1.1 Stakeholder Analysis

Some of the stakeholders and their interests are: 1. Project team: The project team members are stakeholders who have a direct interest in

the success of the project. They are responsible for ensuring the project is delivered on time and within budget. Their interest lies in developing a high-quality expense tracking visualization that meets the user's needs and improves the user experience. 2. Users: The users of the expenses tracking visualization are stakeholders who have a direct interest in the project's success. They will use the visualizations tool to track their expenses and manage their finances. Their interest lies in having a user-friendly and intuitive tool that provides them with clear and concise information on their expenses. These users can include:

- Companies
- Individuals
- Family households
- College students

2 DATA ACQUISITION

Datasets can be custom created or obtained from public data repositories. Data acquisition of expenses data typically involves gathering and collecting information about various expenses incurred by individuals or organizations. There are several ways to acquire expense data, including:

1. Manual entry: This involves manually entering expense data into a spreadsheet or accounting software. It can be time-consuming but allows for a high level of control and accuracy.
2. Kaggle datasets : for example <https://www.kaggle.com/datasets/grosvenpaul/family-income-and-expenditure> Links to an external site.
3. Automated expense tracking: There are a variety of software tools available that can automatically track and categorize expenses
4. Public surveys
5. Government/ public websites: for example https://data.worldbank.org/indicator/NE.CON.PRVT.CD?name_desc=true&locations=US

In this project, student expenses dataset from public resource (Kaggle) was taken for analysis.

2.1 Description of Data

The dataset is a csv file with 100 records of Bachelors students in various years of college, and their split-up expenses – month-wise and category-wise. The columns include Gender, Age, Study_year, Living, Scholarship, Part_time_job, Transporting, Smoking, Drinks, Games_&_Hobbies, Cosmetics_&_Self-care, Monthly_Subscription, June_expenses_\$, July_expenses_\$, August_expenses_\$, September_expenses_\$, October_expenses_\$, November_expenses_\$, Total_Expenses, EXPENSE_high. This data represents the expenses of college students and will help visualizing their expenditure track. The month wise expenses fields display variations but the total expenditure column sums up the totals which range from \$893 to \$2078, which is a big variation.

Table 1. Student Expenses Data Analysis

Gender	Age	Study_year	Living	Scholarship
Female	18	2	Home	No
Female	18	2	Home	No
Female	18	2	Home	No
Female	23	4	Home	No
Male	21	4	Home	No
Male	21	4	Home	No
Female	18	1	Home	Yes
Male	22	3	Hostel	No
Male	22	3	Home	No
Male	18	1	Home	Yes
Male	22	3	Hostel	No
Male	19	2	Home	No
Female	18	1	Home	Yes
Male	22	3	Hostel	No
Male	19	1	Home	No

3 ANALYSIS METHODS

The regular analysis methods for visualizing data and understanding the insight needs include temporal, geospatial, topical and network analysis. Although the stakeholders can be assisted utilizing all the existing analysis methods, dataset and resource constraints will lead us to analyse a few methods such as temporal and topical.

3.1 Temporal Analysis

Temporal analysis of student expenditure involves examining how students' spending patterns change over time. This analysis can provide insights into how students manage their finances, what expenses are most important to them, and how their spending habits change in response to various factors such as changes in income or lifestyle.

One way to conduct a temporal analysis of student expenditure is to track expenses over a specific period, such as a semester or academic year.

Since we already have data that has attributes representing the split-up monthly expenses for a semester, we can simply plot a histogram, bar plot, line graph or scatter plot to understand the temporal trend/pattern in the student expenditure.

4 VISUALIZATIONS

The following visualizations are created with the help of Microsoft Excel Visualization tool and Python Pandas and Matplotlib module.

4.1.1 Prototype 1 – Bar graph

A bar plot is a common way of visualizing data that shows the frequency or count of a categorical variable. Temporal patterns can be understood from a bar plot by adding a time dimension to the data. This bar plot demonstrates the averages of monthly expenses over a semester for 100 students belonging to different college years – with expenses such as living at home/ hostel, transportation, monthly subscriptions, part-time job, smoking, drinking habits, etc.

Graphic symbols – numerals, line, text, point

Graphic variables – color, numerals

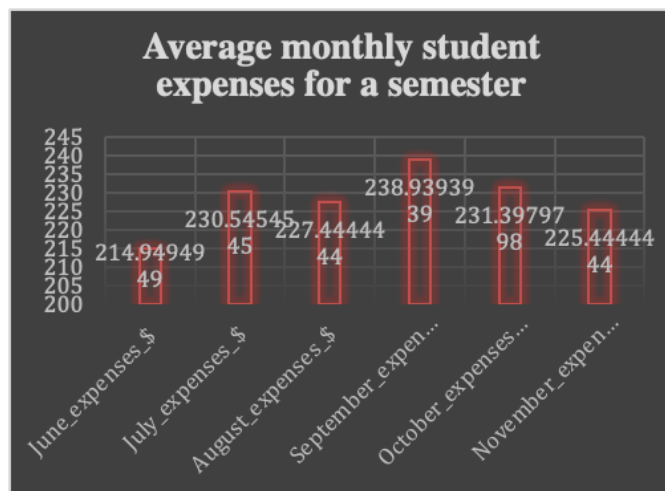


Fig. 1. Bar graph

4.1.2 Prototype 2 – Box plot

Box plots are a useful visualization tool for understanding the distribution of a dataset. They provide a graphical representation of the median, quartiles, and outliers of the data. In order to understand

temporal patterns from a box plot, you need to consider the time dimension of the data.

One way to represent temporal patterns using box plots is to create a series of box plots for each time interval or category of interest.

This box plot is a combination of 7 different graphs over 7 months of the Fall semester. The range of expenses seems to be highest in the month of September and least in the months of October, June and July. It is safe to assume that students tend to spend more when they are in college than they are at home or if they've recently come from home.

Graphic symbols – numerals, line, text, point

Graphic variables – numeral

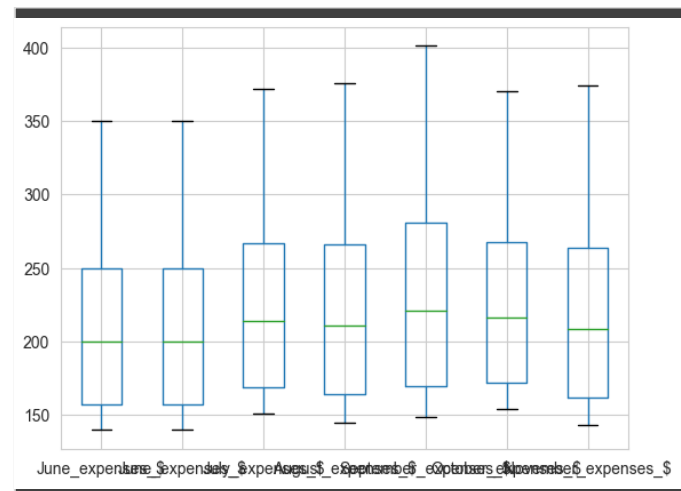


Fig. 2. Box plot.

5 INTERPRETATION OF RESULTS

From the first prototype graph, we can see that the average expenses monthly for students range from \$214 in June to \$239 in September. It can be inferred that the start of semester has the least expense, but the month before fall break has the highest. This can indicate pre-break costs like shopping for fall break or airplane ticket booking etc. It is wise to know that these expenses are the subtraction of part-time income and savings. The median seem to range between \$200 and \$225 (~\$25) and look similar in patterns; contradicting the range of total expenses ~\$400 which makes us think there is a huge variation/anomaly to find in the data.

The second graph shows us the median, range, and quartiles for monthly student expenditure. If we observe a similar kind of box-plot combination for the next semester or next Fall, we can conclude the cyclic nature of expenses in college. Since the granularity of the data is monthly, we can expect to find cycles on an annual or semi-annual basis.

A good conclusion from both the graphs would be that September is the time when student should be mindful of their expenses.

ACKNOWLEDGMENTS

I wish to thank Professor Andreas Bueckle to give me this opportunity of understanding visual analytics and applying it to real world issues.

REFERENCES

- [1] K. Summers and K. T. Block. Design and Evaluation of Visualization Tools for Personal Finance Management.
- [2] M. J. Salmeron, C. T. Calero, and J. L. Garrido. Design and Implementation of a Mobile Expense Tracker