Introduction and Motivation

As a student living in Turkey, for a couple of years official numbers have shown a very high hyper-inflation environment. It is discussed that inflation erodes purchasing power and alters consumer behavior, often pushing individuals to prioritize essential over discretionary spending. However, I have observed that my personal experience may not reflect this assumption.

Null Hypothesis (H_0):

> **As a student at Sabancı University, in high-inflation conditions, there is no significant shift in my spending habits from discretionary to essential categories.**

Alternative Hypothesis (H_1):

> **As a student at Sabancı University, in high-inflation conditions, there is a significant shift in my spending habits from discretionary to essential categories.**

Exploratory Data Analysis

1 - Build a data frame

Motivated by this, I embarked on a journey to analyze my personal spending history during active school terms from fall 2022 to the current fall 2024. Those files had the transactions at my bank account and attributed with the following fields:
- Tarih
- Tutar
- Bakiye
- Açıklama

I needed to do categorizing for each transaction into essential and discretionary subcategories, and see the relations between the different sub-categories.
- Cafeteria food
- Market – grocery
- Education
- Savings
- Entertainment
- Dining out - food
- Dining out - coffee
- Fashion
- Transportation
- Others

I also had my monthly irregular income put into classification by
- Allowance
- Grant

✅ The raw data is made focused on the school terms specifically. Based on the words inside the Açıklama column, there detected the type of the sub category and it is assigned to the regarding transaction.

✅ Since the Cafeteria food is on the basis of a large charge-up and being served long-lasting, it needed a distributed scale of those school-id cafeteria charge-ups. Starting from the day the charge-up occurs, up until the day the same type transaction holds, every day is equally allocated a share. Only the days that Dining out - food is happened, are passed.

- For some days, shares fell short between the charge-up days, it caused very small shares a day far from the reality for the time being. To scale this up, the average cafeteria spending a day (slightly less than the average spending) is assigned to the years and if transaction money falls short on those time periods, those assigned realistic values replaced the initial shares.
- P.S. This supposedly realistic values are determined by first-hand experience, the prices for those times are considered.
- P.S. Some part of categorization is handled by python code whereas 10% of all data are categorized manually and corrected.

2 - Clean the data frame

AÃ§Ä±klama	Sub-category	General Category
Hmr Restoran GÄ±da	Dining out - food	Discretionary
Starbucks Card MobÄ±l	Dining out - coffee	Discretionary
Distributed Cafeteria food	Cafeteria food	Essential
YDS GOOGLE *ChatGPT	Education	Essential
YDS GOOGLE *ChatGPT	Education	Essential
Distributed Cafeteria food	Cafeteria food	Essential
7036 Ä°stanbul Sabanc	Market â€" grocery	Essential
Self GÄ±da SanayÄ± Ve	Dining out - food	Discretionary
SabancÄ± UnÄ±versÄ±tesÄ	Cafeteria food	Essential
Starbucks Card MobÄ±l	Dining out - coffee	Discretionary
Distributed Cafeteria food	Cafeteria food	Essential
Hmr Restoran GÄ±da	Dining out - food	Discretionary
TaksÄ±de Pos	Transportation	Essential
Belbim	Transportation	Essential

Above – final_categorized_transactions.csv

- All data points are merged into the day they belong to.

Sub-category	General Category	dayOfWeek
Dining out - food	Discretionary	Saturday
Dining out - coffee	Discretionary	Saturday
Cafeteria food	Essential	Saturday
Education	Essential	Thursday
Education	Essential	Thursday
Cafeteria food	Essential	Thursday
Market â€" grocery	Essential	Tuesday
Dining out - food	Discretionary	Tuesday
Cafeteria food	Essential	Monday
Dining out - coffee	Discretionary	Monday
Cafeteria food	Essential	Monday
Dining out - food	Discretionary	Sunday
Transportation	Essential	Sunday

Above - final_categorized_dayOfWeek.csv

- The corresponding day is mapped to the day of the week as a new field.
- Each day has a total tutar (total spending) value for all transactions that day.
- Each day has a discretionary share value, representing the proportion of discretionary spending over the total spending.

| Tarih	dayOfWeek	Total Spending	Total Income	Discretionary Amount	Essential Amount	Share of Discretionary
1.11.2022	Tuesday	-143	2500	-103	-40	0.72027972
2.11.2022	Wednesday	-110	0	-110	0	1
3.11.2022	Thursday	-79.25	0	0	-79.25	0
4.11.2022	Friday	-188.4	750	-75	-113.4	0.398089172
5.11.2022	Saturday	-89.5	0	-25	-64.5	0.279329609
6.11.2022	Sunday	-160	2	-75	-85	0.46875
7.11.2022	Monday	-110	0	-50	-60	0.454545455
8.11.2022	Tuesday	-80.9	0	-20.9	-60	0.258343634
9.11.2022	Wednesday	-365.47	0	-305.47	-60	0.835827838
10.11.2022	Thursday	-135	0	-75	-60	0.555555556
11.11.2022	Friday	-1920.41	0	-1707.91	-212.5	0.889346546
13.11.2022	Sunday	-225	0	-75	-150	0.333333333
15.11.2022	Tuesday	-122.75	2600	-13	-109.75	0.105906314
16.11.2022	Wednesday	-159.9	0	-159.9	0	1
17.11.2022	Thursday	-179.45	0	-75	-104.45	0.417943717

Above - final_categorized_merged.csv

Those datasets have shown the raw relation for my spending history and raised some questions. Can I find any effect of the inflation on my volume of spending and the change occurred per month? I made inspection against official inflation statistics from TÜİK. For the inflation rates of annual and monthly scale (TÜFE), I tried to reach out data from the TÜİK's website by data scraping method. Cumulative inflation along with the monthly inflation will be laid down with my own spending history and be analyzed for the claim I proposed at the beginning.

This project serves as an opportunity to explore the interplay between economic factors and personal financial behavior, challenging conventional assumptions and gaining insights into how young individuals adapt—or resist adapting—to economic turbulence.

At the end, I will try to put my findings to shed light on a broader perspective, such that if grants we can collect help us do well during this hyper-inflation condition.

3 - Explore Global Properties
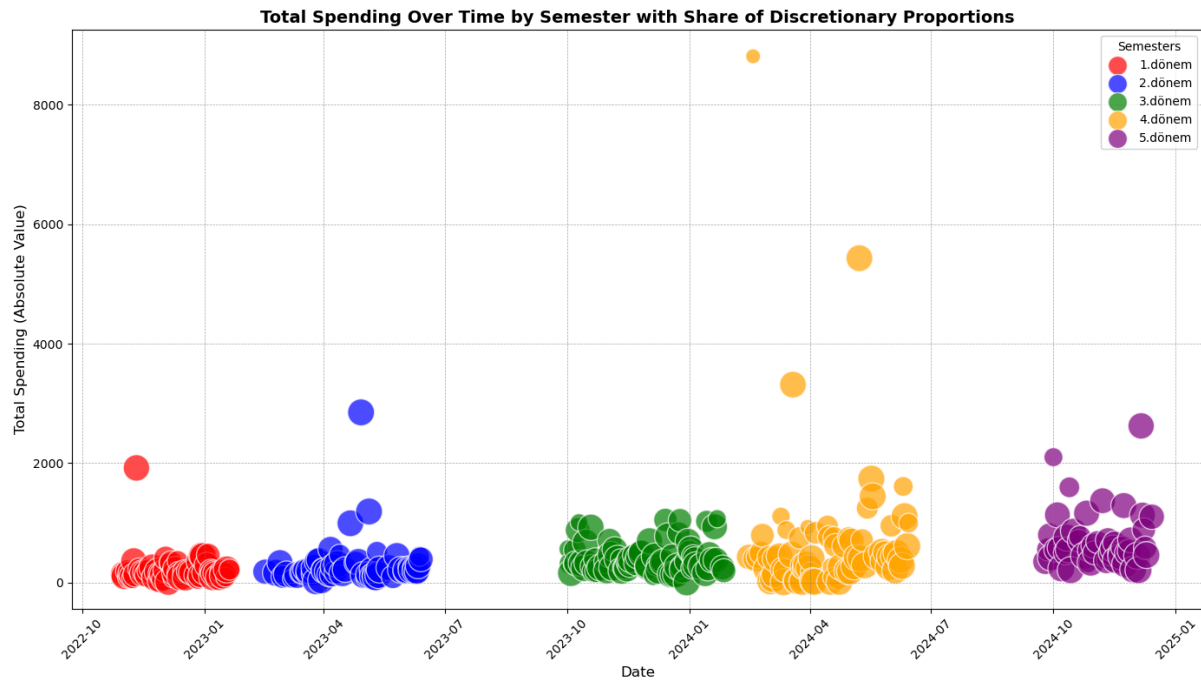
Using the final_categorized_merged.csv file, I put a chart that lays over the semester-based relation over the volume of spending and the degree to which the share of discretionary has.

This chart is freed of overlies and focused on the mostly populated range for all 5 of semester.
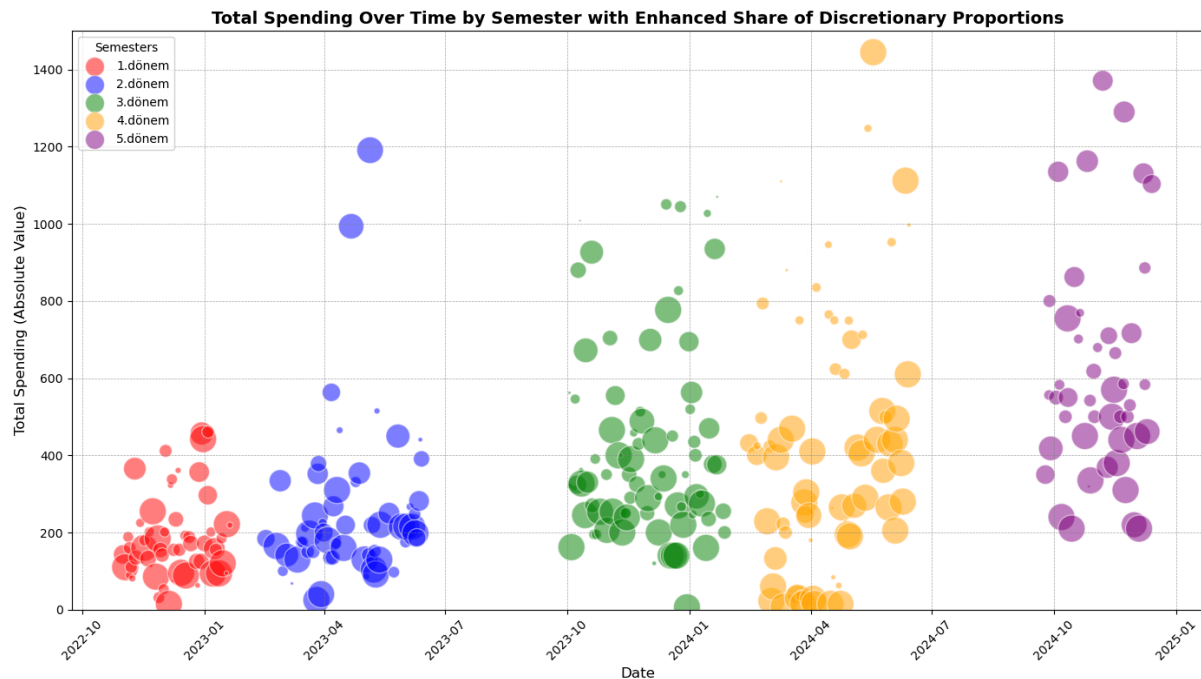
All of the days in my spending history are notated with a coloured, differently sized circles. The size of a circle is dependent on the ratio of discretionary spending over the total for that day.

Findings:
- The volume of the spending increases by time, the center of masses move upward
- The spring semesters (blue and yellow) have clusters emerged at the very bottom around the April time. This may be indicating to the spring break Sabancı University has at that time of the year, when I go to my hometown and have less personal spending.

Histogram with Statistical Highlights:

- This graph shows the distribution of total spending during the term
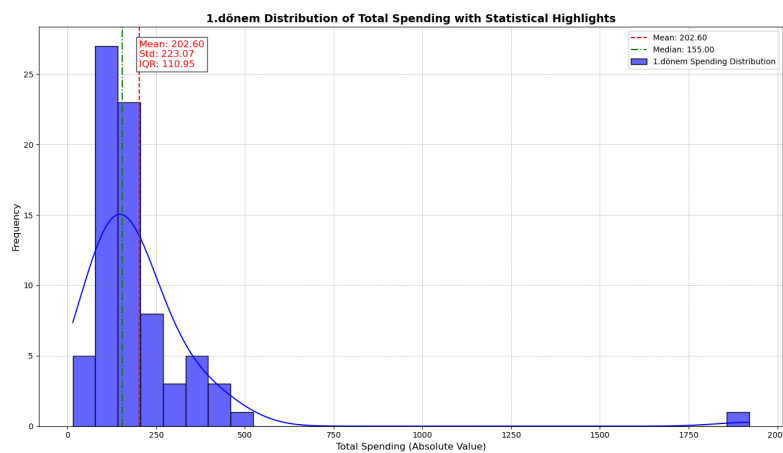
- indicates a right-skewed distribution. This suggests that there are some days with significantly higher spending than the typical day.

- A long tail on the right side of the histogram suggests occasional high-expenditure days (outliers), possibly due to large discretionary purchases or essential payments (e.g., tuition fees, travel)
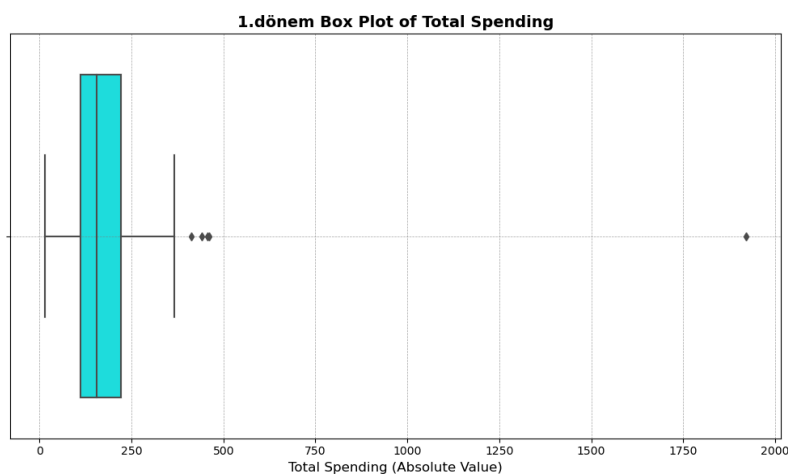
- The presence of **outliers** confirms occasional days with extremely high spending compared to the typical pattern.
- The range between Q1 and Q3 suggests moderate spending variability mostly, but the extended whiskers shows some days significantly deviate from this pattern.
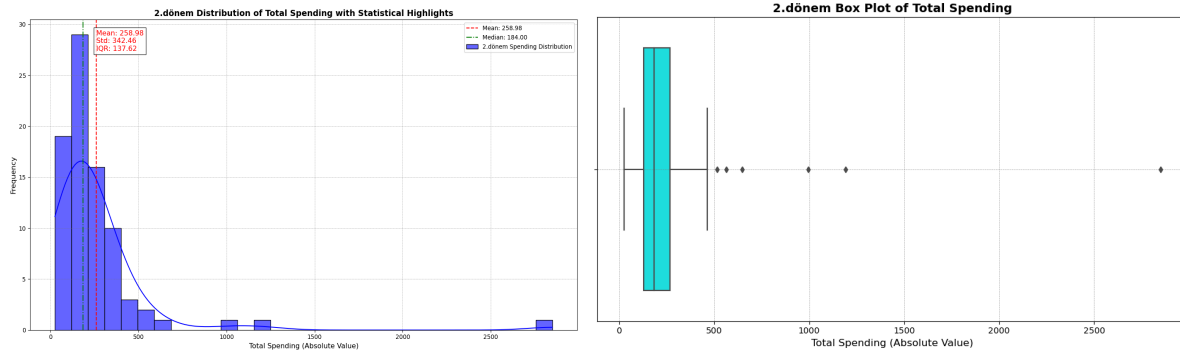

● Occasionally the similar plot with the first semester
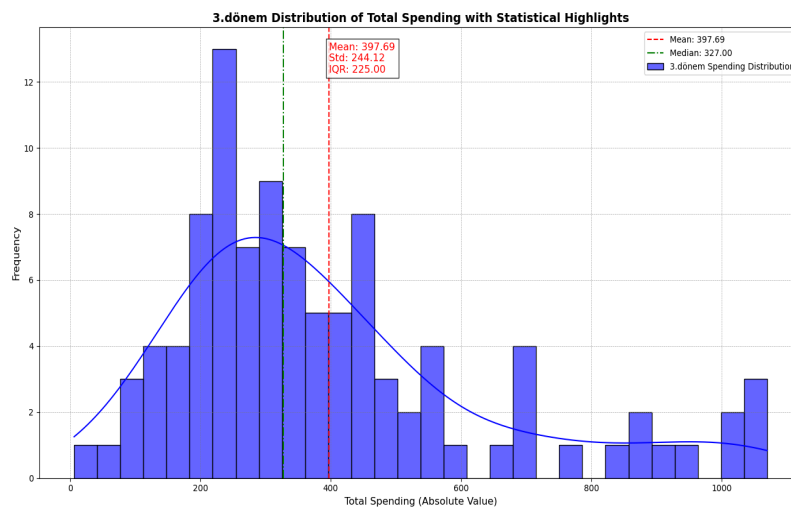
● The histogram demonstrates a notable upward shift in spending levels compared to earlier terms. This could be due to inflation, which increases essential spending, or a shift in spending habits toward higher discretionary spending.
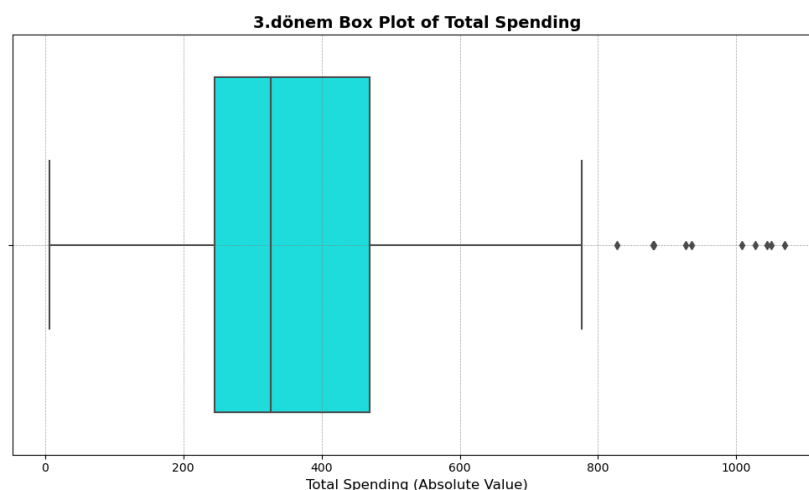
- **Outliers:** There are several high-spending outliers beyond 800, with one approaching 1000+. These outliers may correspond to significant discretionary or essential expenses, such as large purchases or fees.
- **Spread:** The box plot shows a much larger interquartile range than earlier terms, indicating higher spending variability during this period.
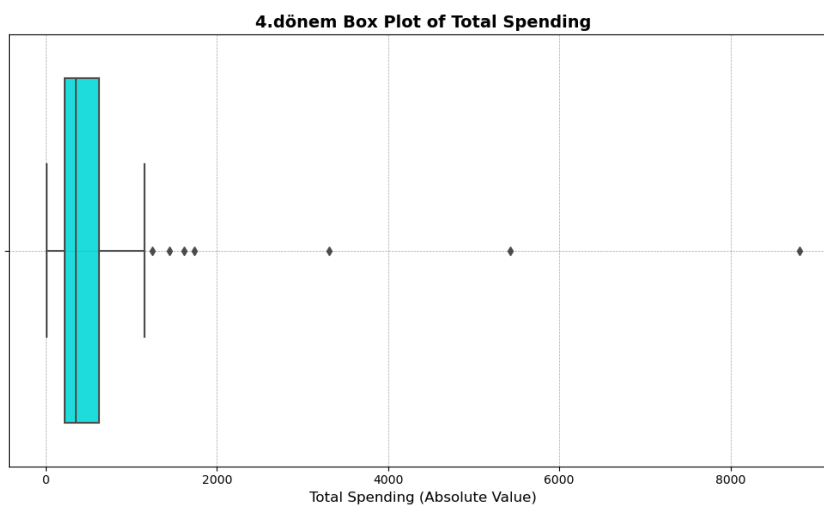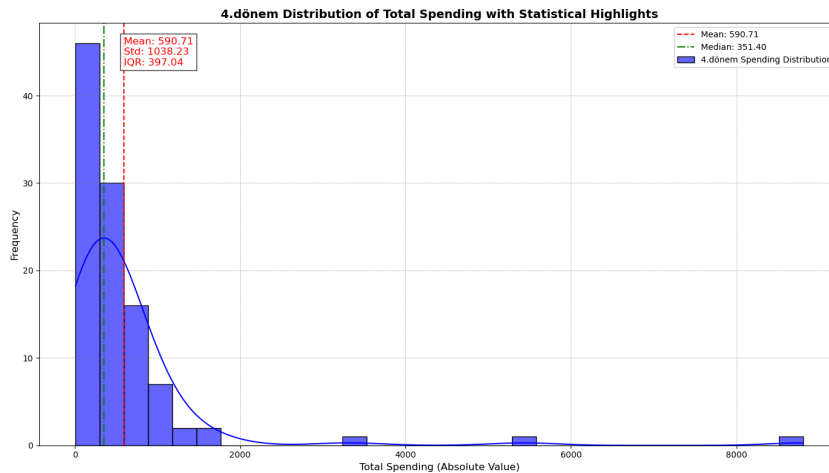
- There is drastically larger overlies laying over that makes the data seem shrinked. However, the volume of spending also takes much more space by the x-axis values.

5.dönem Distribution of Total Spending with Statistical Highlights
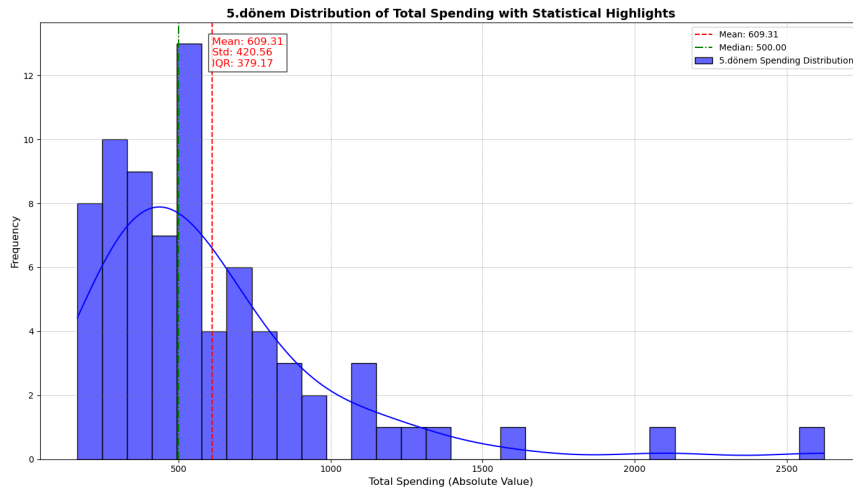
- The median value of **500.00** is closer to the mean, suggesting a more balanced distribution but still slightly skewed to the right.
- The standard deviation (420.56) is the highest so far, indicating substantial variability.
- The interquartile range has grown considerably compared to previous terms, implying a larger spread of typical spending days.

5.dönem Box Plot of Total Spending
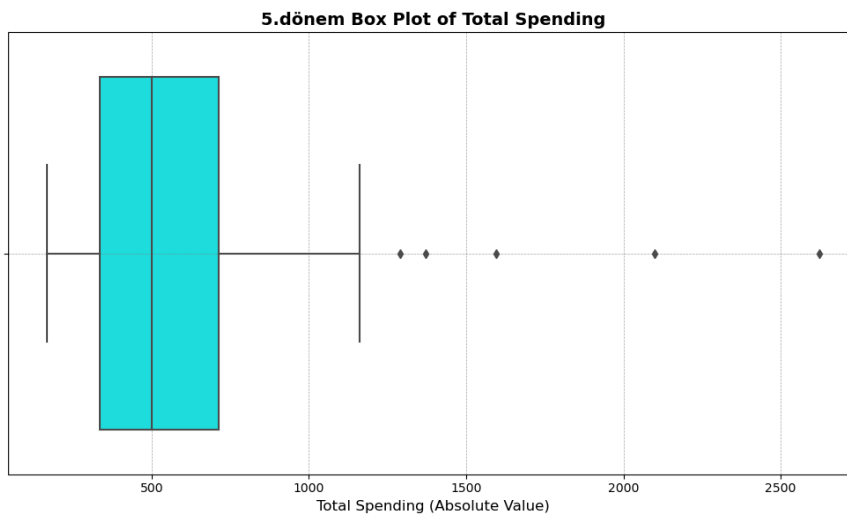
- The interquartile range is the widest among the terms analyzed, with the whiskers extending to a broader range of values before hitting outliers.
- While the typical daily spending is centered around 500.00, the growing variability implies that spending behavior is becoming less predictable.

Heatmaps : volume of spending vs. discretionary share
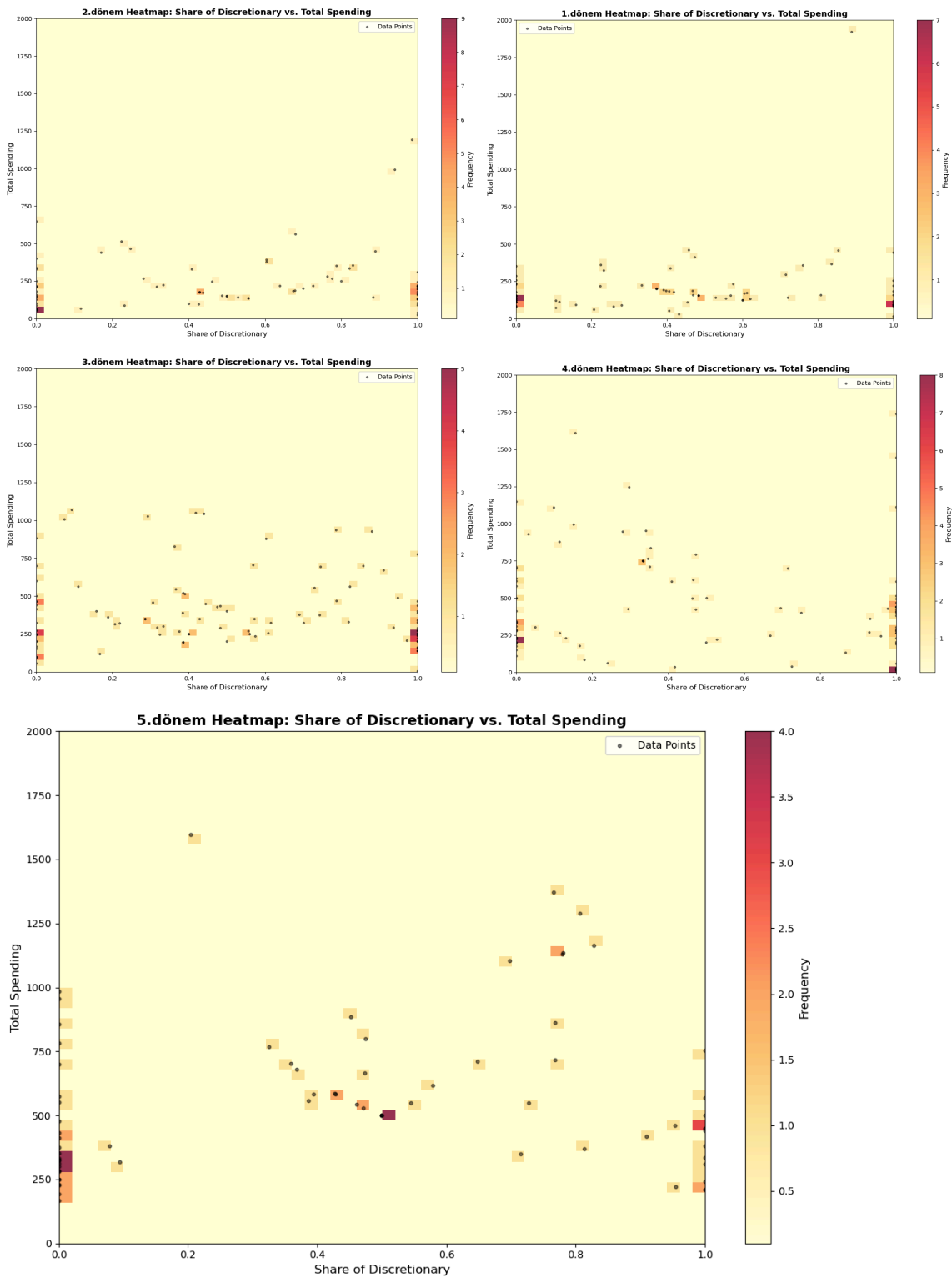
1. **Spending Trends Over Time**
   - Total spending has progressively increased across terms, consistent with rising inflation and/or changing financial behavior.
   - Higher spending days (outliers) have become more frequent, reflecting the impact of **inflation on essential purchases** or **occasional discretionary splurges**.
2. **Discretionary Spending Patterns**
   - Discretionary spending remains relevant, but its share is generally lower (<0.5) for high total spending days. It also may indicate to the impact of the inflation on essential purchases as mentioned above.
   - In terms of the share of discretionary, the outside regions take more frequency over time. Essential extremums have higher total spending volumes(5th: around 700) whereas discretionary extremum remains at more moderate scale(5th: around 500).
3. **Relationship Between Discretionary Share and Total Spending**
   - High total spending days are primarily driven by essentials (low discretionary shares).
   - Discretionary-heavy spending days (>0.8 share) are associated with smaller total spending amounts, indicating limited indulgence in purely non-essential expenditures. This might be caused by budget-management against discretionary.
- Despite rising total spending, students seem to maintain a balance between discretionary and essential expenditures, with high total spending driven mostly by essentials.
- The increasing variability in spending reflects the evolving spending priorities.

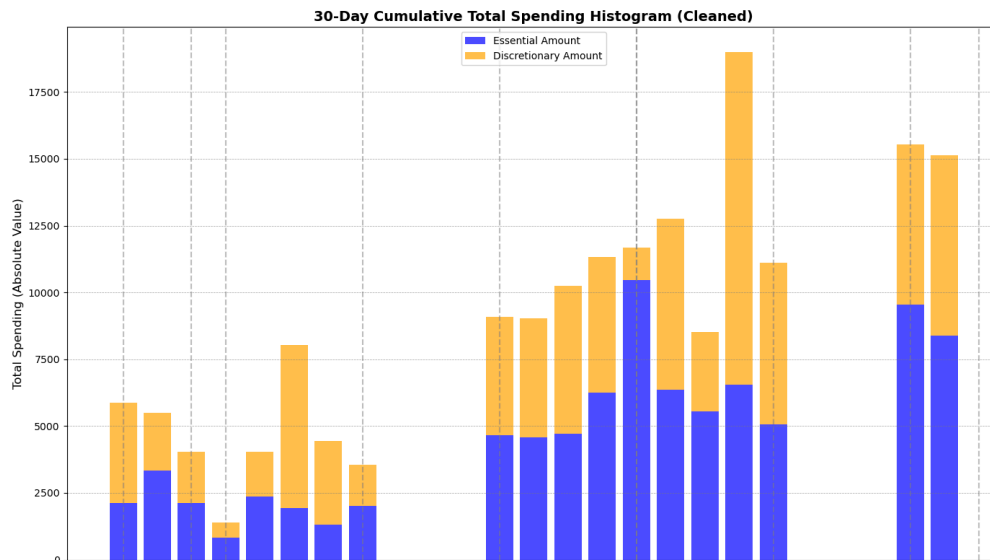## Analysis of Cumulative Spending Histograms

note that: the vertical lines along the chart indicate to the start and end dates of the school semester since this project particularly focuses on the school terms.
- This puts 5 separate school semester distinct from each other.
- The small gap is for the semester (winter) break as the large is for summer.

### 15-Day Cumulative Spending Histogram
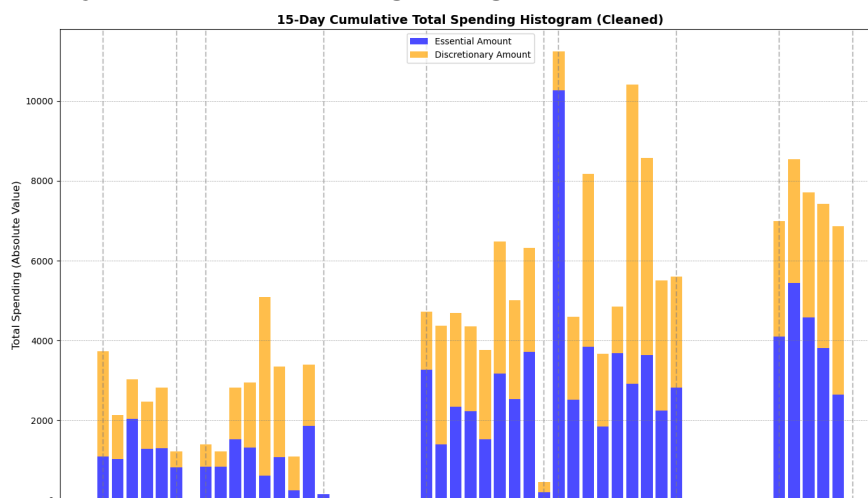
- Spending varies significantly across periods, with some showing spikes where discretionary spending dominates. Spending spikes occur around the **first 15 days** of any month, indicating to the **dates the income (grant) transactions** occur.
- Blue (essential) spending remains more consistent, while yellow (discretionary) spending shows larger fluctuations, indicating more flexibility.
- Peaks in spending may align with the **start or end of school periods**.

30-Day Cumulative Total Spending Histogram (Cleaned)

30-Day Cumulative Spending Histogram

- This chart aggregates spending over a longer 30-day cumulative period
- Total spending increases over time, particularly **in the later school terms**, which likely reflects the compounding impact of inflation or changing spending behavior.
- Discretionary spending continues to show substantial variability, while **essential spending grows more consistently.**
- Longer periods show clearer dominance of discretionary spending in the total budget during peaks, suggesting that **non-essential** purchases are **more affected by the timing of income** or other financial factors.

Inflation Rate Trends and Analysis
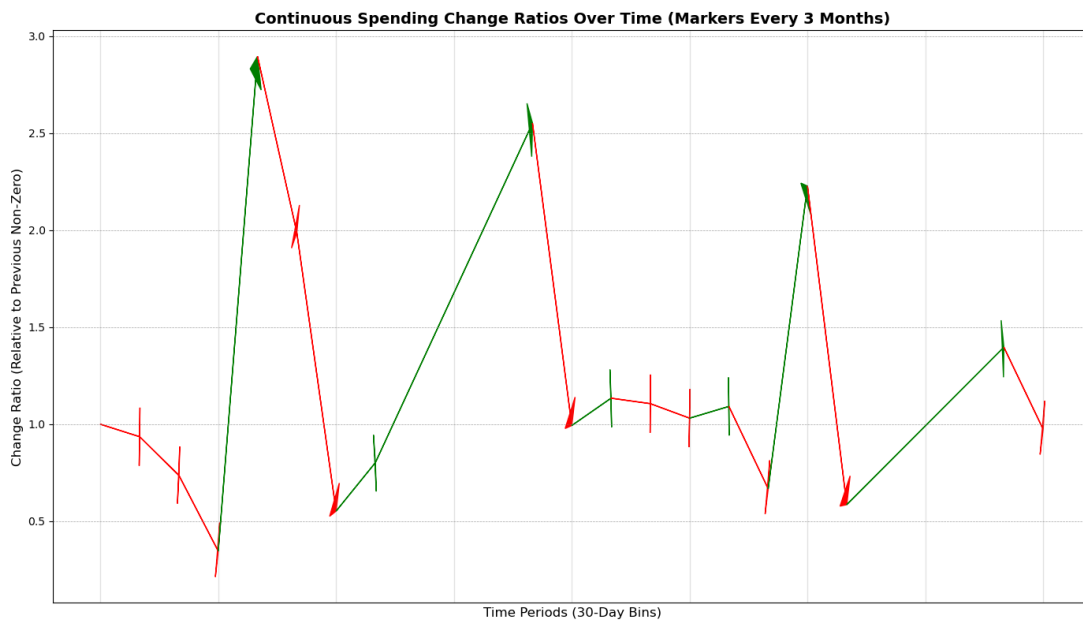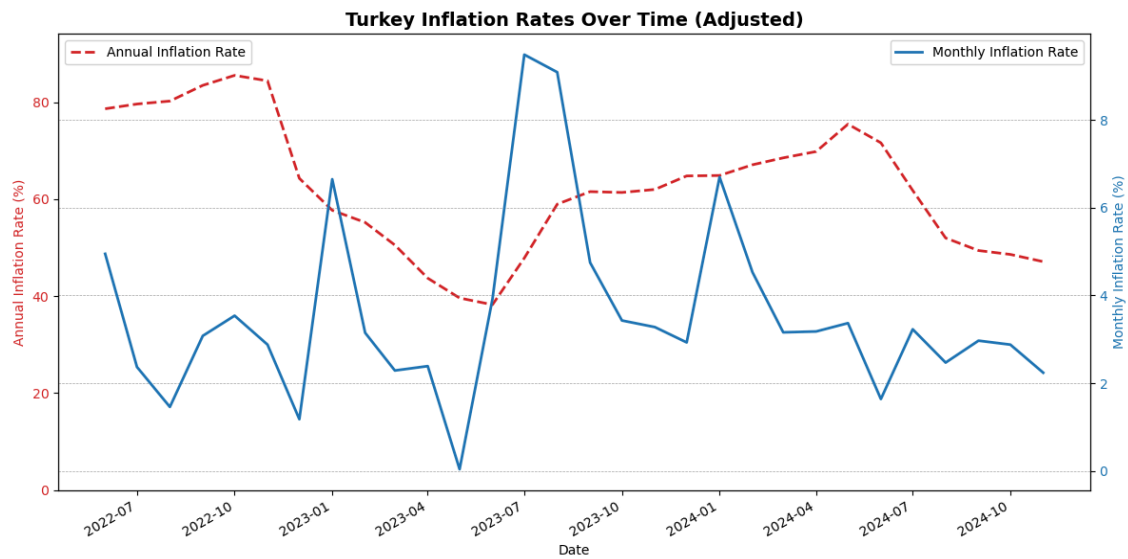
The first chart is from the analysis of the scraped data from TÜİK's official website. There two distinct variable are laid over:

- The red is the annual inflation rate which is calculated based on the prices of the previous year's same date.
- The blue chart is for the monthly inflation which is found by the increase in the prices from the last month's same products.

The second chart is from the cumulative histogram of bins of 30-days. I computed the change of the volume of the spending from the previous month (parallel with the logic of the computation of inflation), The lines are set positive or negative by the type of change.

- it seems that the increase in the volume caught up with the inflation rates, especially at the peaks of 2023-01, 2023-07, 2024-01

```
1    Annual Inflation Rate,Monthly Inflation Rate,Date
2    47.09.,2.24,Kasım-2024
3    48.58.,2.88,Ekim-2024
4    49.38.,2.97,Eylül-2024
5    51.97.,2.47,Ağustos-2024
6    61.78.,3.23,Temmuz-2024
7    71.60.,1.64,Haziran-2024
8    75.45.,3.37,Mayıs-2024
9    69.80.,3.18,Nisan-2024
10   68.50.,3.16,Mart-2024
11   67.07.,4.53,Şubat-2024
12   64.86.,6.70,Ocak-2024
13   64.77.,2.93,Aralık-2023
14   61.98.,3.28,Kasım-2023
15   61.36.,3.43,Ekim-2023
16   61.53.,4.75,Eylül-2023
17   58.94.,9.09,Ağustos-2023
18   47.83.,9.49,Temmuz-2023
19   38.21.,3.92,Haziran-2023
20   39.59.,0.04,Mayıs-2023
21   43.68.,2.39,Nisan-2023
22   50.51.,2.29,Mart-2023
23   55.18.,3.15,Şubat-2023
24   57.68.,6.65,Ocak-2023
25   64.27.,1.18,Aralık-2022
26   84.39.,2.88,Kasım-2022
27   85.51.,3.54,Ekim-2022
28   83.45.,3.08,Eylül-2022
29   80.21.,1.46,Ağustos-2022
30   79.60.,2.37,Temmuz-2022
31   78.62.,4.95,Haziran-2022
```

monthly_inflation_rates_list_2022_to_2024.txt

- The inflationary environment in Turkey from 2022 to 2024 has been characterized by extreme annual rates and significant monthly spikes.

- These trends likely caused noticeable shifts in consumer spending, with essentials dominating during peak inflation and discretionary spending gradually returning as inflation stabilized.

note that:
- So far, it seems that inflation to a great degree affected the volume of the spendings. The change at the monthly spendings happened accordingly. It might have caused due to the parallel increase at my income.
- However, there is still no concrete finding showing a drastic change at the share of discretionary spendings.
- Independent of inflation, the timeline of the special days (e.g. start-end date of the school, holidays) has some peaks of the discretionary. This does not help my hypothesis.

4 - Explore Group Properties

Network Graph

Spending Patterns by Semester and Day of Week
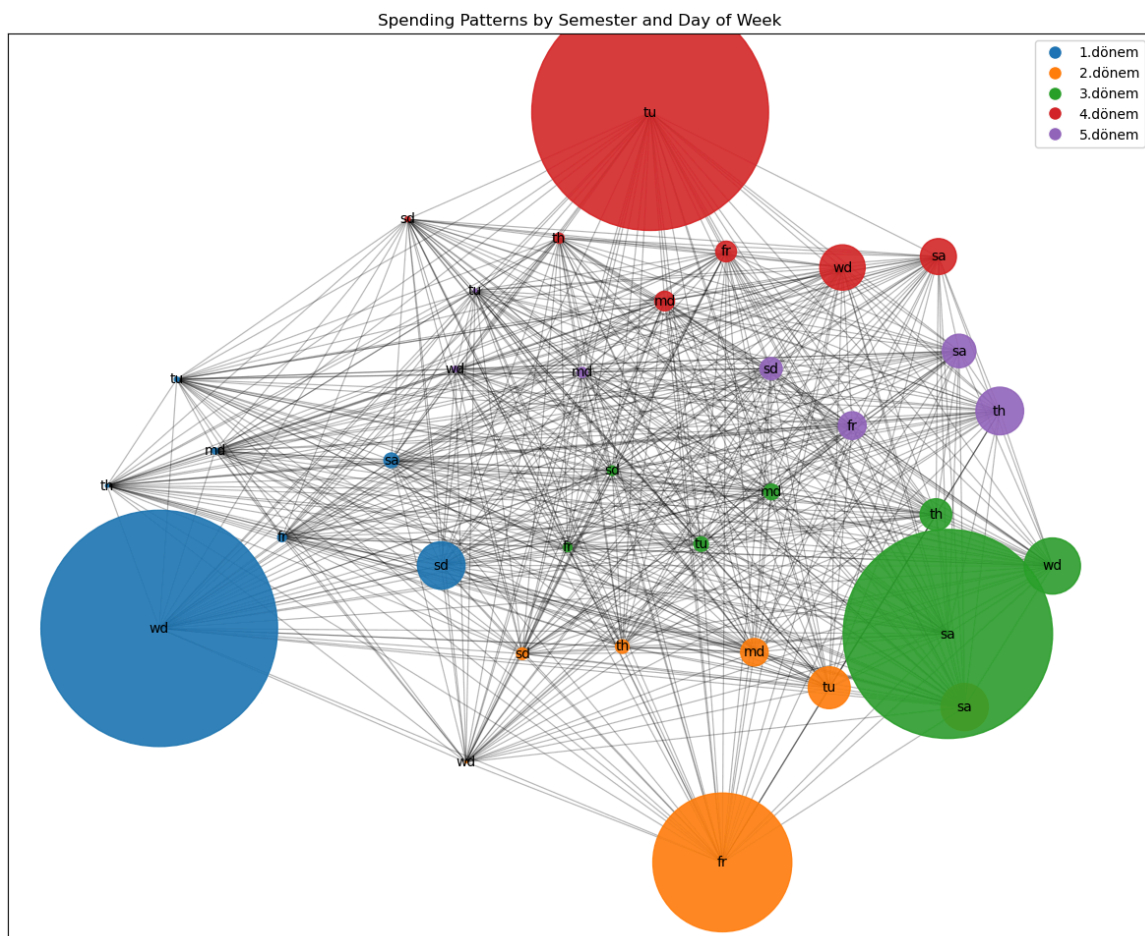
Node Sizes (Discretionary Ratio):

- Larger nodes represent higher **average discretionary spending ratios** for specific days and semesters.

Node Distribution Across Semesters:

- Nodes are colored by semester, showing that **discretionary spending varies across terms** and specific days of the week.

Edge Weights (Relationships Between Nodes):

Edges represent relationships based on:

- **Similarity in discretionary spending** (discretionary ratio).
- **Day of the week grouping** (e.g., weekday vs. weekend spending behavior).
- **Semester effects** (same semester).

Heavier edge connections are present between nodes of the same semester or similar days, suggesting **consistency in spending patterns within terms.**

- Weekends : Friday to Sunday

show consistently higher discretionary spending, indicated by the larger nodes and closer grouping

- Weekdays : Monday to Thursday

tend to have smaller nodes, reflecting a lower proportion of discretionary spending

Findings:

- 2nd (orange) and 3rd (green) semesters are likely to have larger discretionary share on weekends such that they clustered close, however 1st (blue) and 4th (red) puts forward something different and diverged from the mass 2 and 3 formed.
- 5th semester (purple) does not have a distinguishing discretionary share for those days of weeks. The largest share is on Thursday, on which I have no classes in my schedule. (is it a coincidence?)

Analysis of the Subcategory Spending Histograms

2. Proportional Percentage Chart of Subcategories

Proportional Percentage Chart of Sub-categories Over 15-Day Bins

Proportional Percentage Chart of Sub-categories Over 30-Day Bins
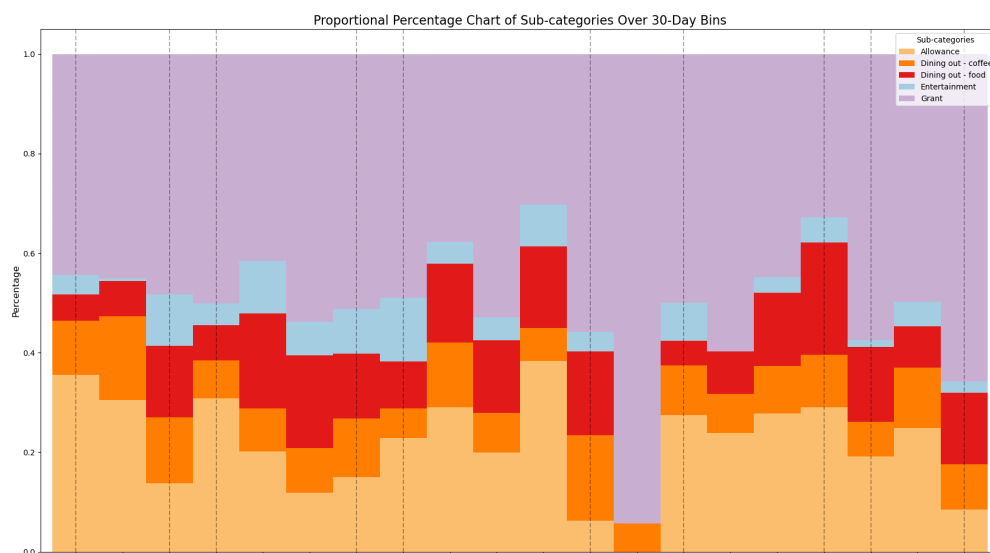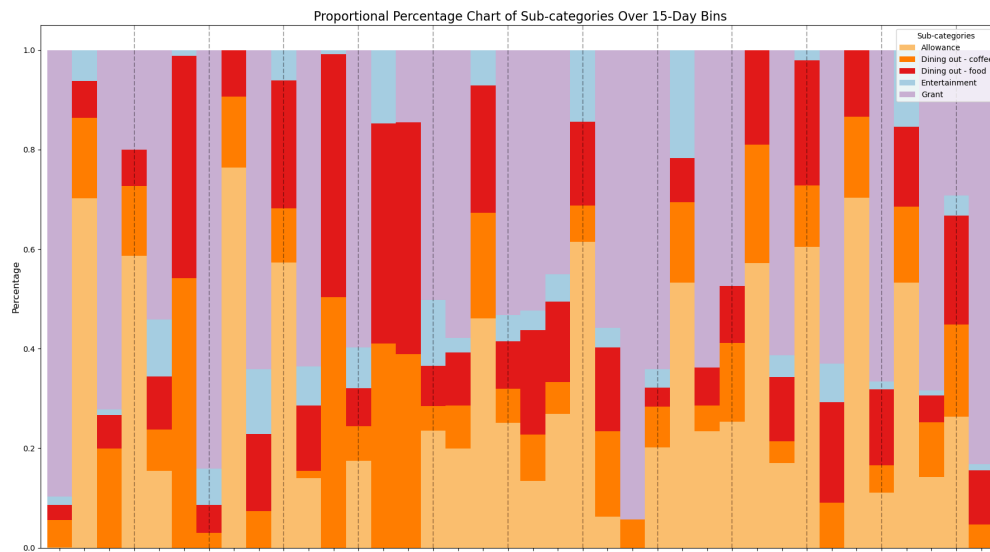
- It highlights shifts in relative importance of subcategories over time.

- Grants dominate the proportion in certain bins, especially at the **start of terms**, indicating that essential income forms the **base of financial stability.**
- Allowances also contribute significantly, particularly in the middle of terms, showing reliance on regular inflows for both essential and discretionary needs.

- Vertical lines shows the start and the end times of the school.
- Dining out (red and orange) takes up larger proportions during specific intervals, particularly near the **end of terms**.
- Entertainment spending fluctuates but retains a steady proportion.

note that:
Essential inflows (grants and allowances) form the base of spending, with discretionary expenditures layered on top. This reflects **financial stability derived from predictable income sources**.

3. Proportional Percentage Chart of Subcategories - Only Spendings

Proportional Percentage Chart of Sub-categories Over 30-Day Bins
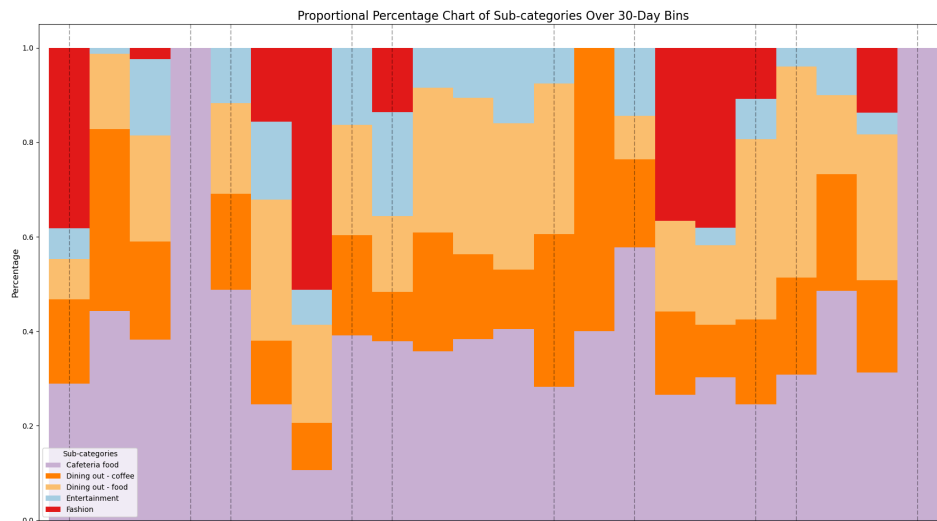
- Cafeteria food (essential) comes as the dominant factor.
- Spendings for fashion does not get rare and kind of hold a steady place during the semesters.
- Dining out - food and coffee seem the main parts of discretionary spendings. Since entertainment (theaters, parties, travelling) takes always the smallest share across the terms.

note that:
- It was discussed that the discretionary spending rate over the term goes steady, however, in the current economic condition the major share for the discretionary spending is dining - out (anywhere other than the cafeteria) and coffee shops, where student can find as only places to socialize properly.
- The main source of stable discretionary rate seems like the steady income discussed here.

Hypothesis Testing

After a careful consideration, in order to put a statistical foundation for the hypothesis testing, I shrinked the final_categorized_merged formerly month based, latter semester-wise, then formed two lists of data.

MonthYear	Total Monthly Spending	Average Discretionary Share
11.2022	-5867.773333	0.476950834
12.2022	-5931.623333	0.437465199
1.2023	-3598.293333	0.41180721
2.2023	-1396.89	0.498864144
3.2023	-4271.04	0.447707899
4.2023	-7820.96	0.515935937
5.2023	-4621.27	0.560446643
6.2023	-3384.98	0.45611878
10.2023	-10530.17	0.468673095
11.2023	-8436.59	0.525116212
12.2023	-11196.36667	0.541308718
1.2024	-10003.31333	0.493537757
2.2024	-13150.54	0.308515034
3.2024	-12901.71	0.5948906
4.2024	-9327.6	0.431311277
5.2024	-18926.25	0.505052408
6.2024	-8308.65	0.592038095
9.2024	-2290.78	0.496976321
10.2024	-18610.51	0.389489171
11.2024	-15349.67333	0.466929863
12.2024	-9447.426667	0.515308528

Semester	Total Monthly Spending	Average Discretionary Share
1st Semester	-15397.69	0.442074
2nd Semester	-20098.25	0.495052
3rd Semester	-40166.44	0.507159
4th Semester	-49464.21	0.530823
5th Semester	-43407.61	0.457243

Average Discretionary Share vs. Total Monthly Spending

Average Discretionary Share vs. Total Spending by Semester
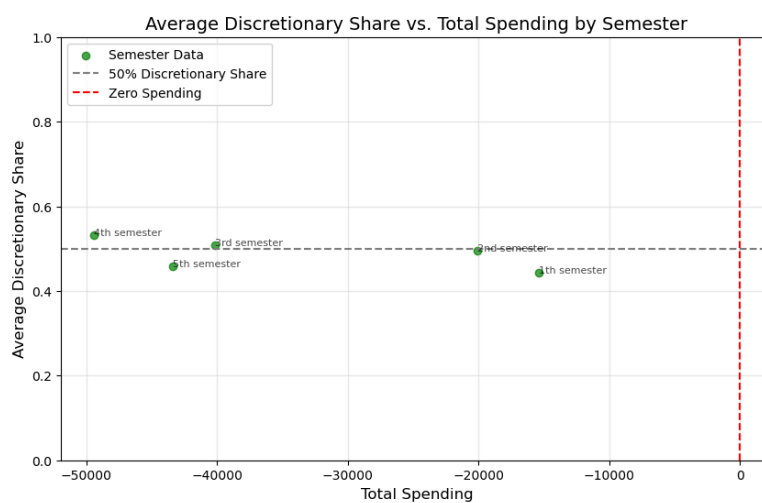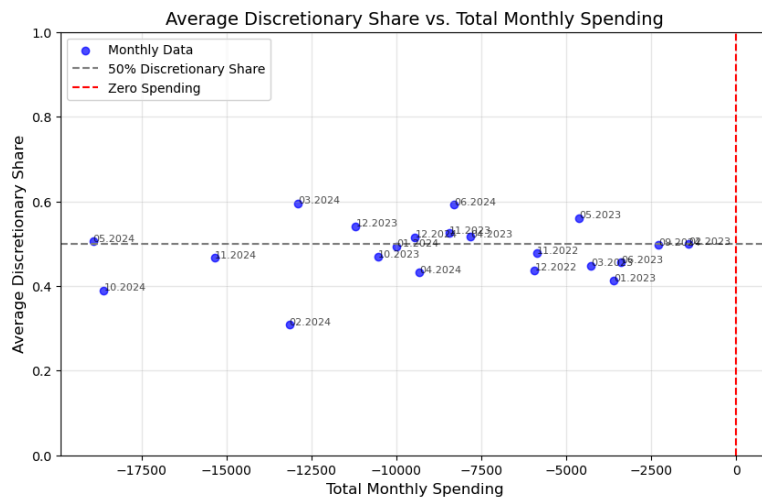
For those charts, the relation between data points across the x and y axis is approximately linear. This means it paves the way for Pearson correlation coefficient which is closely focused on the linear relations of the systems.

- First one is for the monthly discretionary rates
- Second is semester-based so there is scarce of data points, present ones put a linear relation though
- The linear relation for these two has got the total spending per particular period of time ineffective for the discretionary share attribute.

- Pearson test also puts the following statistics:
  - Monthly Spending :
    - correlation coefficient   = 0.1174261645459552
    - p-value   = 0.6122084602151447
  - Semester Spending :
    - correlation coefficient   = -0.5689070256803843
    - p-value   = 0.31687742685132503
- Both have p-values larger than significance level 0.05. This prevents from having a proof that supports the claim of the alternative hypothesis.
- **I failed to reject Null Hypothesis.**

As a student at Sabancı University, in high-inflation conditions, there is no significant shift in my spending habits from discretionary to essential categories.

Machine Learning

Step 1 : Baseline Decision Tree Model

A Decision Tree Classifier was trained to predict economic periods using features like total spending, income, and discretionary share. The initial model achieved an accuracy of 30% on the test set.

Decision Tree for Economic Period Prediction

Step 2: Optimized Decision Tree

Using GridSearchCV for hyperparameter tuning, the decision tree model was optimized with the following best parameters:

- max_depth: None
- min_samples_leaf: 2
- min_samples_split: 10

The optimized model achieved a cross-validation score of 42% and a test accuracy of 33%.
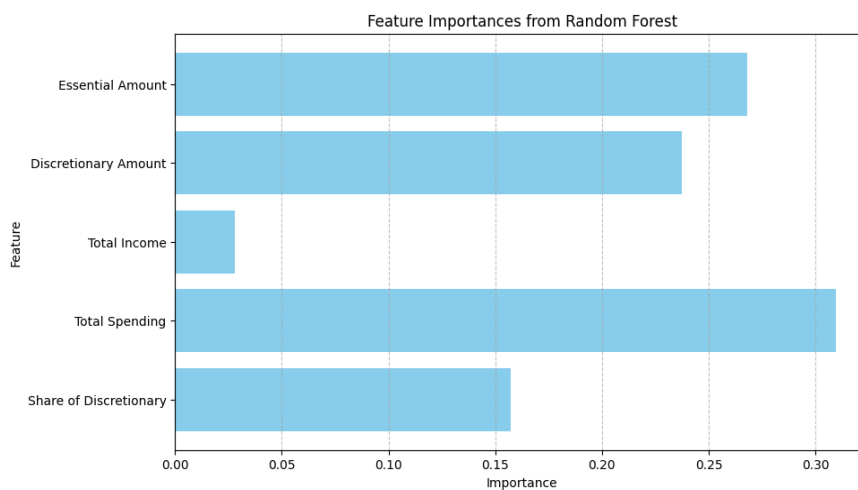
Tuned Decision Tree
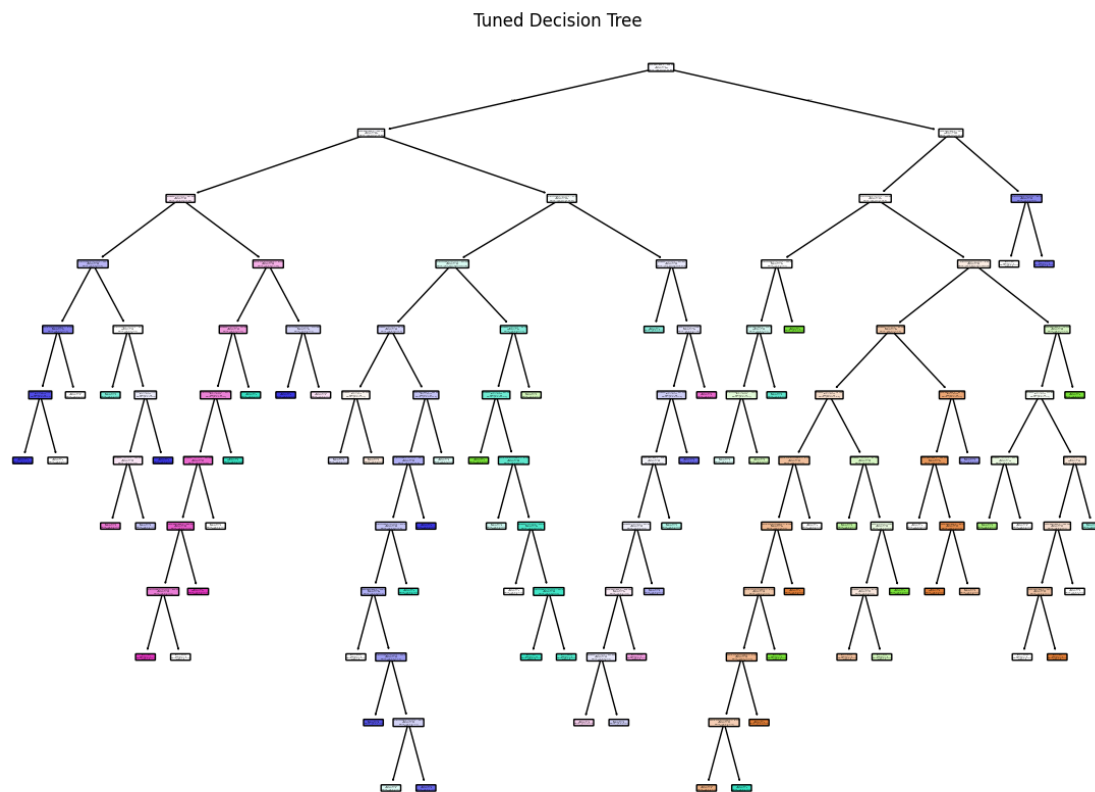
Step 3: Random Forest Model

A Random Forest Classifier was trained and tuned using GridSearchCV. The best parameters:

- n_estimators: 50
- max_depth: 5
- min_samples_split: 10
- min_samples_leaf: 1
- bootstrap: False

The Random Forest achieved a cross-validation score of 45% and a test accuracy of 39%.

Feature Importances from Random Forest

Discussion

The findings of this project indicate that the discretionary spending rate over the terms remained steady despite the high-inflation conditions. The primary contributors to discretionary spending, such as dining out (excluding the cafeteria) and coffee shops, highlight the social significance of these expenditures. These locations serve as essential venues for students to socialize, making them less susceptible to significant reductions even in challenging economic conditions.

One of the key reasons for the stability in the discretionary spending rate is the steady and predictable income stream, primarily from essential inflows like grants and allowances. These income sources form the financial foundation, ensuring that essential expenditures are covered first, while discretionary spending is layered on top. This stability reflects how predictable income sources help maintain financial equilibrium.

Inflation has undeniably impacted the volume of total spending. Monthly spending trends show that as inflation rises, the volume of expenditures increases correspondingly, potentially driven by a parallel increase in income. However, this increase in spending volume has not resulted in a substantial shift in the proportion of discretionary spending. The findings suggest that the discretionary spending share remains relatively consistent, with no drastic changes observed over time.

Independent of inflation, the timeline of special events—such as the start and end of semesters or holidays—shows peaks in discretionary spending. For instance, the spring semesters exhibit clusters of lower spending during April, which aligns with the spring break at Sabancı University. This period is typically characterized by reduced personal spending as I spend more time in my hometown. These event-driven fluctuations do not support the hypothesis that inflation has caused a shift in spending priorities.

Despite the rising total spending driven primarily by essential expenditures, the balance between discretionary and essential categories has been maintained. The increasing variability in spending reflects the influence of inflation and evolving spending priorities. However, the data does not provide concrete evidence of a significant shift from discretionary to essential spending.

Conclusion

The hypothesis posited that in high-inflation conditions, as a student at Sabancı University, there would be no significant shift in spending habits from discretionary to essential categories. The findings support this hypothesis, as the discretionary spending rate has remained steady. Inflation has influenced the total spending volume but has not led to a notable reallocation of spending priorities. Therefore, the null hypothesis cannot be rejected.