
SWE-TRANS
www.swe-trans.com
02-617-4481-2
081-836-6439
CASIO
NATURAL DISPLAY
Taxes 60
loan 800

SPENDING PATTERN OF MUMBAIKARS

TABLE OF CONTENTS

Title: To study the Spending pattern of Mumbaikars.

Purpose of the survey:

As consumers, both their level of spending and the way they allocate their spending changes. So called "life events" such as getting a first job, marriage, having children, and retirement can all have profound effects on spending patterns. By analysing the spending patterns, we can infer what consumers are currently purchasing and how much they are willing to pay in the future in specific product or service categories. Income and expenditure insights, such as who the consumers are, where they live, how much money they earn and what they spend their money on, influence the spending patterns. Last year, the outbreak of Covid–19 has severely affected the spending patterns as well.

This survey provides the information we need to analyse the spending pattern of people based on various demographic factors such as age, income, and region. This type of information can help segment the market, assess consumer demand, help prioritise needs from wants and pinpoint opportunities.

Objectives:

- To study the expenditure of the target group on needs and wants.
- To ascertain the different avenues of savings and investments.
- To identify the demographic factors that affect expenditure.
- To analyse the impact of pandemic on the economic aspect of the target group.

Target Audience:

The target audience for our survey is the people residing in Mumbai belonging to the age group of 16-60 years who are earning.

Confidentiality Note:

The information collected through this survey is purely for academic purposes and will not be used for any other purposes.

Questionnaire:

Section I

1. Age (Completed years) _____
2. Gender
    - Male
    - Female
    - Others

3. Are you an earning individual?
    - Yes
    - No (goes to end of questionnaire)

4. Are you a resident of Mumbai?
    - Yes
    - No (goes to end of questionnaire)

5. Where do you live?
    - Mumbai city
    - Mumbai Suburban

6. What is your employment status?
    - Full time employed
    - Part time employed
    - Entrepreneur
    - Self-employed/Freelancing
    - Student
    - Pensioner
    - Other (specify) _____

7. Which of these describes your personal income **per annum** (in lakh) ?
    - below 2.5

- 2.5 - 5
- 5 - 10
- above 10

Section II

8. How do you divide your income in the following terms? (in percentage)
    - Essentials _____
    - Non-essentials _____
    - Savings _____

9. How do you like to dissect your discretionary income (income that is left after spending on your basic necessities)?
    - Save the money
    - Invest the money somewhere for yields
    - Partially save and partially invest

10. What is your preferred way of investing?
    - None
    - Building portfolios(stock market)
    - Real estate investing
    - Mutual Funds
    - Government securities
    - Others (specify) _____

11. What is your preferred way of saving?
    - Insurance and annuities
    - Fixed deposits
    - Precious collectibles(Eg.: gold,silver)
    - Others (specify) _____

12. How has the pandemic affected you? (Select all that are applicable)
    - ☐ Pay cut
    - ☐ Pay hike
    - ☐ Job loss

- ☐ Job change
- ☐ No change
- ☐ Others (specify) _____

13. In which category have your expenses **increased** the most after the pandemic?

_____

- Groceries
- Rent/Maintenance
- Insurance/Healthcare
- EMI
- Education
- Transportation
- Services (househelp)
- Recreation
- Others

14. In which category have your expenses **decreased** the most after the pandemic?

_____

- Groceries
- Rent/Maintenance
- Insurance/Healthcare
- EMI
- Education
- Transportation
- Services (househelp)
- Recreation
- Others

15. Exhaustion of your savings since the pandemic is _____

- More than before
- Same as before
- Less than before

16. How did your last month's expense vary from your average monthly expense?

- Significant increase

- Increase
- No change
- Decrease
- Significant decrease

17. Please select the option which best represents your spending habits.
(You may use each rating more than once)

	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree
Your savings are exhausted easily					
Your monthly expenses exceed your monthly income					
You are satisfied with your spending pattern					
Your monthly expenses match your budgeting plan					

18. If your income increased by 10%, how would you utilize it?

- Saving more
- Spend more

19. If your income decreased by 10%, how would you compensate for it?
    - Cutting expenses
    - Use savings

20. If you have to make a major purchase, how would you do it?
    - Use savings
    - Take loan
    - Borrow from others

Thank you for spending your precious time.

Methodology:

Survey was conducted using the services of forms.app and the total collected responses were 109. Responses were collected on the basis of snowball sampling.The irrelevant data, i.e., non- resident of Mumbai and non-earning responses, and erroneous data was removed from the data. The valid data left was 52 responses for which the data was cleaned.

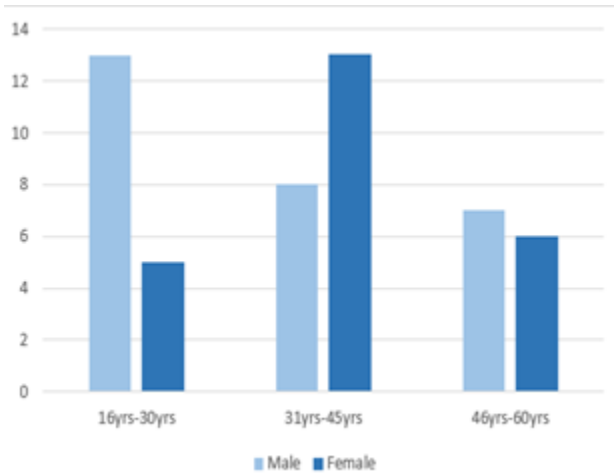
Python Code for Data Cleaning:

```
import pandas as pd
df = pd.read_excel("SpendingData.xlsx")
del df['Index']   #unnecessary column deleted
#giving short column names
df.columns=['age','gender','earning','resident','region','employment'
,'salary','essential','non-essential','savings','discretionary_income
','investments_type','savings_type','pandemic_job','incr_expense','de
cr_expense','exhaustion_savings','last_month_exp_change','Your
savings are exhausted easily','Your monthly expenses exceed your
monthly income','You are satisfied with your spending pattern','Your
monthly expenses match your budgeting
plan','income_incr','income_decr','major_purchase']
#replacing user-typed responses with other
df['investments_type'] = df['investments_type'].replace({'Stock
Derivatives':'other', 'Nonr': 'None', 'gold & investment policies':
'other'})
#replacing user-typed responses with other
df['savings_type'] =
df['savings_type'].replace({'MF':'Other','Crypto, Stocks
':'Other','Mf': 'Other','PPF, EPF': 'Other','Deposit in
bank':'Other','Provident fund': 'Other','Banking':'Other' })
#filling null values with zero
df['essential'].fillna(0,inplace= True)
df['non-essential'].fillna(0,inplace= True)
df['savings'].fillna(0,inplace= True)
```

Analysis:

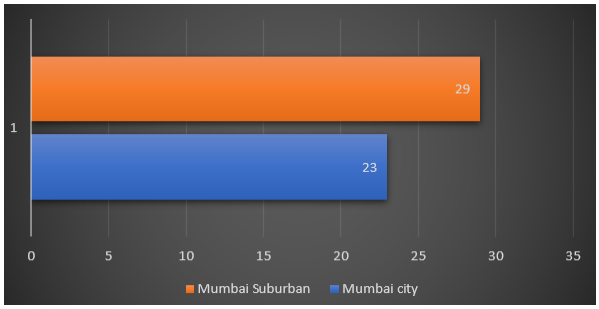
1. **Demographics of respondents:**

Demographic category	**Total**	**%**
Overall	52	100
Age (years)		
16-30	18	34.62
31-45	21	40.38
46-60	13	25
Gender		
Male	24	46.15
Female	28	53.85



The sample consisted of about equal representation of males and females with a fairly distributed representation in the three age categories of 16-30, 31-45 and 46-60 years.

Location	Total	%
Mumbai City	23	44.23
Mumbai Suburban	29	55.77

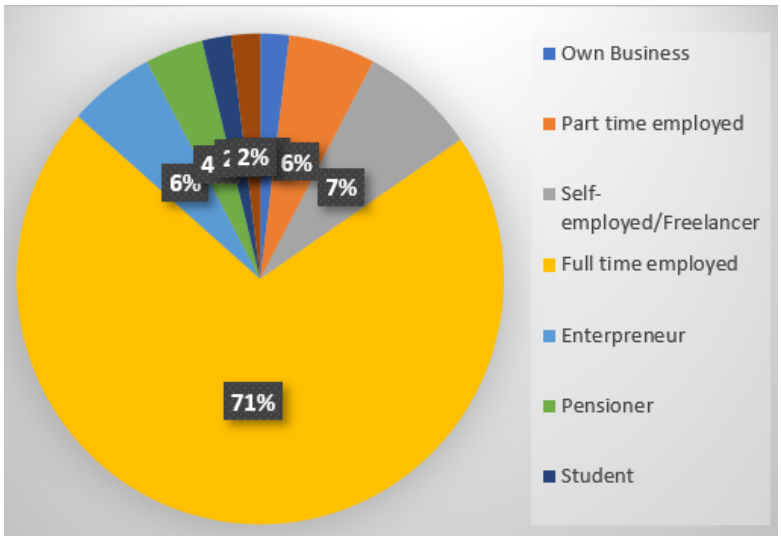


Almost half (56%) of the sample were from Mumbai Suburban and 44% from Mumbai City

Demographic Category	Total	%
Overall	52	100
Annual Income (in lakhs)		
Below 2.5	15	28.86
2.5 - 5	13	25
5 - 7.5	7	13.46
7.5 - 10	9	17.30
Above 10	8	15.38

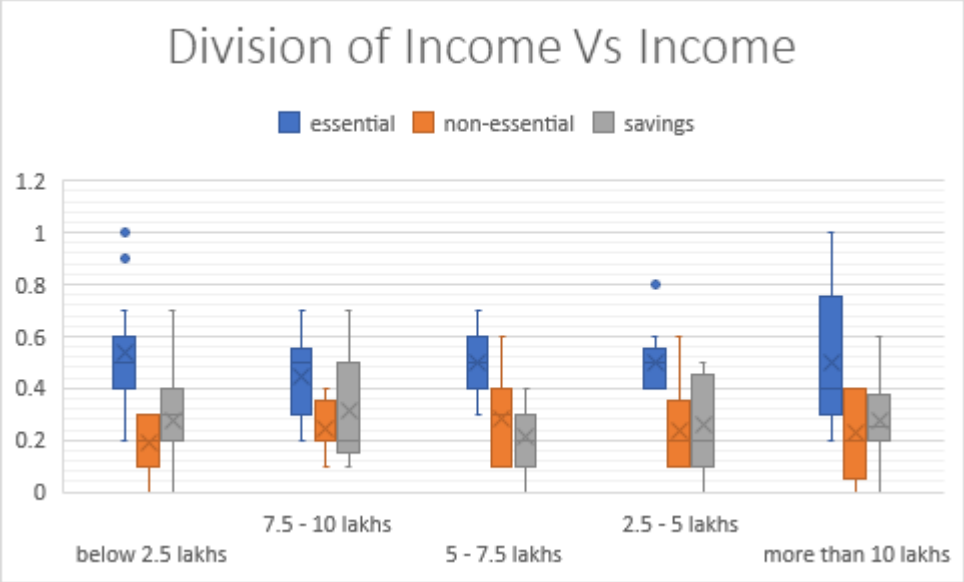
- Most of the respondents were from below 2.5 lakhs followed by 25% between 2.5 to 5 lakhs and 13% from 5-7.5 lakhs.17% were from 7.5 to 10 lakhs and the rest with above 10 lakhs.

Employment Status		
Own Business	1	1.92
Part time employed	3	5.77
Self-employed/Freelancer	4	7.69
Full time employed	37	71.15
Entrepreneur	3	5.77
Pensioner	2	1.92
Student	1	1.92
Other	1	

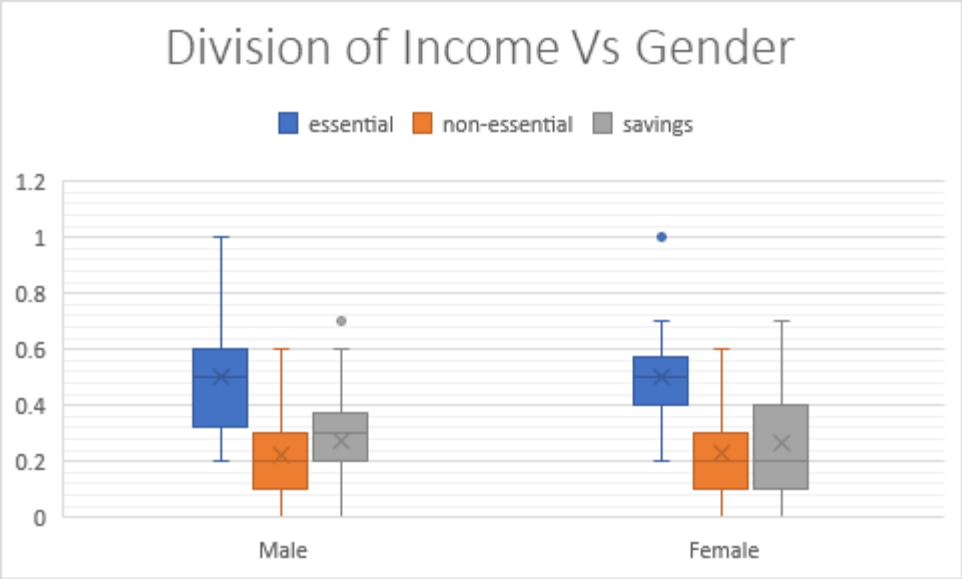


71% of the population was full time employed and the least were from student,own Business and Other.Equal no of respondents(5%) were from part time employed and Entrepreneur.

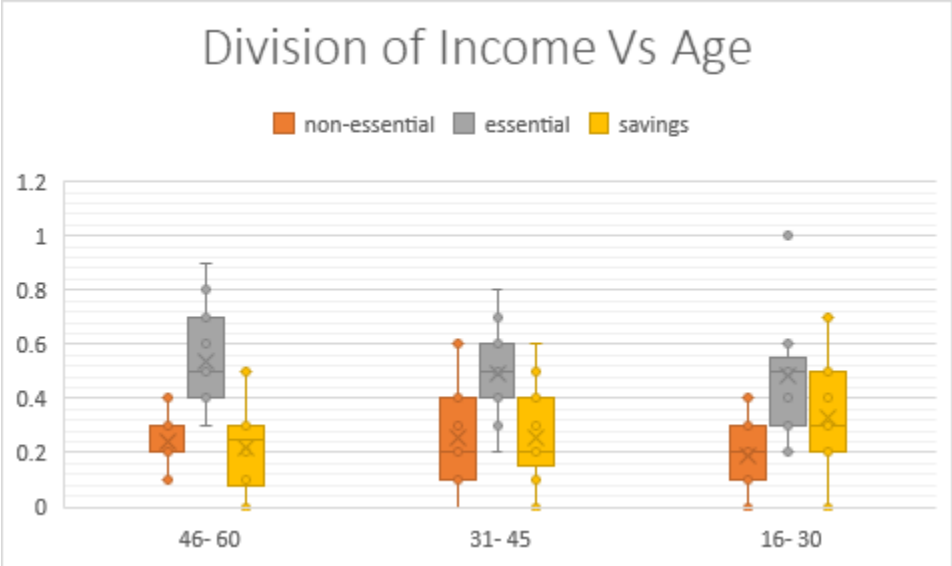
2. **Box and whisker plot analysis for-**
   Allocation of income against various demographic variables



From the above graph, we can see that whatever the income is, most of the income is used for essential things. Income used for non-essentials and savings are seen to be varying. People with income less than 2.5 lakhs are seen to generally save more than they spend on non-essential things. People earning more than 10 lakhs have the widest spending range in essentials (30%-75%), while they have the shortest spending range on savings (20%-37%). Candidates who save the least are from the income category- 5-7.5 lakhs.

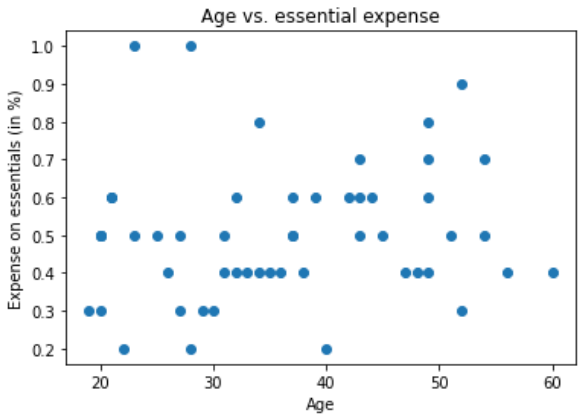


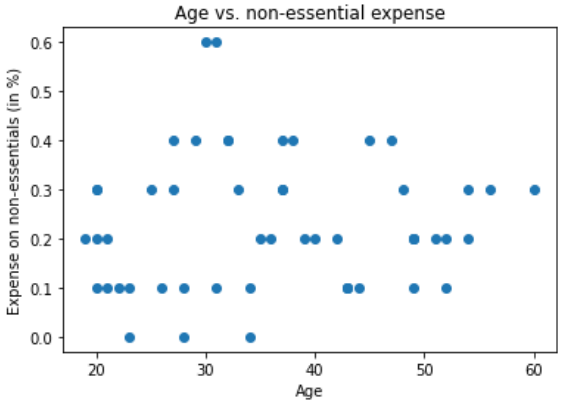
Both male and female candidates have spent most of their income on essential things. Both genders spend almost the same amount on non-essential things. More number of male candidates are shown to save more than female candidates. Most females save 10%- 40% of their income, while most males save 20% - 37% of their income.

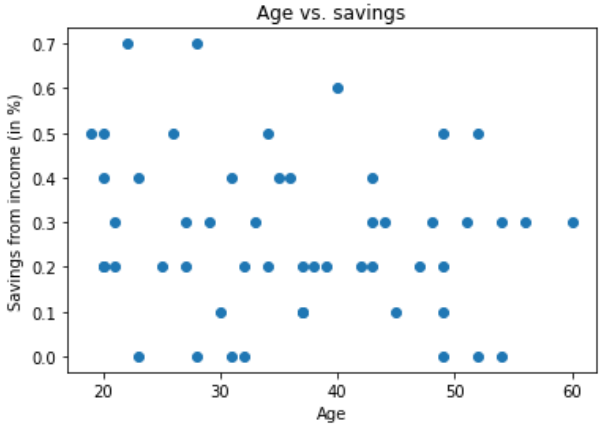


There is a significant difference between the amount used for essentials compared to non-essentials and savings in ages 31-60. The candidates who are aged between 16 and 30 are seen to be saving more of their income (20%-50%) than on non-essential things compared to the older candidates (10%-40%).

3. Scatter Plot of age vs. Allocation of income







- In the above scatter plots, the points are spread throughout without any significant clustering, which implies the spending habit of a person doesn't depend on their age.
- This goes against the general notion that younger people spend more on trendy items and older people save more. Rather, here the spending is based on the individual's nature.

Python Code:

```
x = df.iloc[:,[0]]      #selects age column
y1 = df.iloc[:,[7]]      #selects income percentage on essential
column
plt.scatter(x,y1)        #scatter plot
plt.xlabel("Age")
plt.ylabel("Expense on essentials (in %)")
plt.title("Age vs. essential expense")
```

```
plt.show()

y2 = df.iloc[:,[8]]  #selects income percentage on non-essential
column
plt.scatter(x,y2)      #scatter plot
plt.xlabel("Age")
plt.ylabel("Expense on non-essentials (in %)")
plt.title("Age vs. non-essential expense")
plt.show()

y3 = df.iloc[:,[9]]    #selects income percentage on savings column
plt.scatter(x,y3)    #scatter plot
plt.xlabel("Age")
plt.ylabel("Savings from income (in %)")
plt.title("Age vs. savings")
plt.show()
```

4. Contingency table analysis of discretionary income with income

Attributes:

A : Income groups

B : Preference of saving, investing or both

Hypothesis:

H_0: A & B are independent.

H_1: B is dependent on A.

Observed frequencies:

How do you like to dissect your discretionary income (income that is left after spending on your basic necessities)?	Invest the money somewhere for yields	Partially save and partially invest	Save the money
Which of these describes your personal income per annum (in lakh) ?			
2.5 - 5 lakhs	4	5	4
5 - 7.5 lakhs	2	1	4
7.5 - 10 lakhs	2	6	1
below 2.5 lakhs	2	3	10
more than 10 lakhs	0	7	1

Expected frequencies:

	How do you like to dissect your discretionary income (income that is left after spending on your basic necessities)?		
How do you like to dissect your discretionary income (income that is left after spending on your basic necessities)?	Invest the money somewhere for yields	Partially save and partially invest	Save the money
Which of these describes your personal income per annum (in lakh) ?			
2.5 - 5 lakhs	2.500000	5.500000	5.000000
5 - 7.5 lakhs	1.346154	2.961538	2.692308
7.5 - 10 lakhs	1.730769	3.807692	3.461538
below 2.5 lakhs	2.884615	6.346154	5.769231
more than 10 lakhs	1.538462	3.384615	3.076923

Results:

	Chi-square test	results
0	Pearson Chi-square (8.0) =	18.3924
1	p-value =	0.0185
2	Cramer's V =	0.4205

α= 0.05

Decision criteria: Reject the null hypothesis if the **p-value** is less than **α**

Result: **p-value < α**

0.0185 < 0.05

Reject null hypothesis

Cramer's V is a way of calculating correlation in tables which have more than 2x2 rows and columns. It is used as a post-test to determine strengths of association after chi-square has determined significance.

Decision criteria:

Phi and Cramer's V	Interpretation
>0.25	Very strong
>0.15	Strong
>0.10	Moderate
>0.05	Weak
>0	No or very weak

Result: **Cramer's V > 0.25**

0.4205 > 0.05

There is a strong association between A & B.

Conclusion: As the **p-value** is less than **α**, therefore we reject the null hypothesis and conclude that the preference of saving, investing or both of Mumbaikars strongly depends on their income.

Python code: A

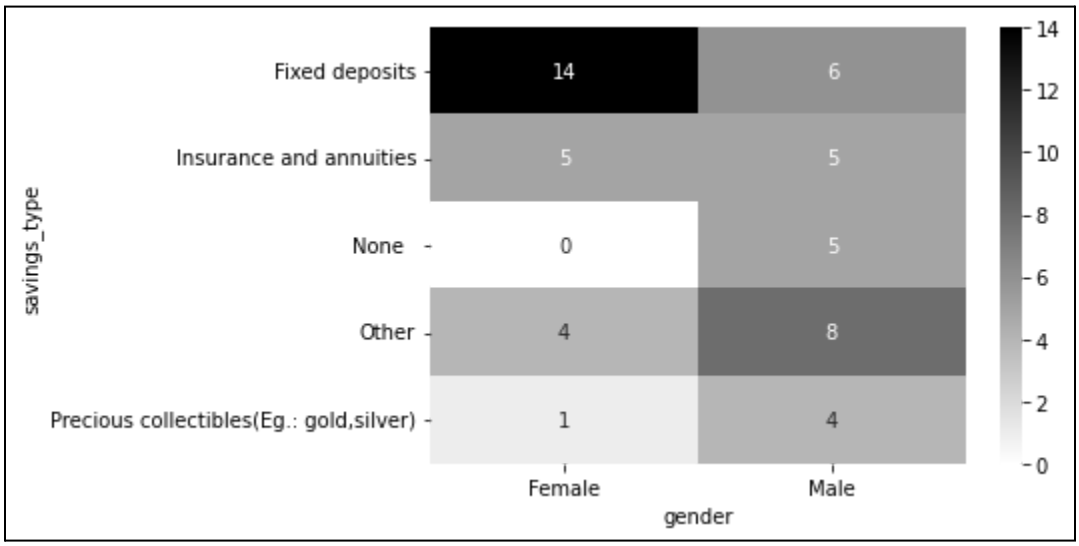
```
import pandas as pd
import researchpy as rp
import scipy.stats as ss

df = pd.read_excel(r'C:\SpendingData.xlsx')

rp.summary_cat(df[["Which of these describes your personal    income
per annum (in lakh) ?", "How do you like to dissect your
discretionary income (income that is left after spending on your
basic necessities)?"]])

crosstab = pd.crosstab(df["Which of these describes your personal
income per annum (in lakh) ?"], df["How do you like to dissect your
discretionary income (income that is left after spending on your
basic necessities)?"])
#observed frequencies
crosstab
```

```
cross_tab, test_results, expected = rp.crosstab(df["Which of these
describes your personal income per annum (in lakh) ?"],df["How do you
like to dissect your discretionary income (income that is left after
spending on your basic necessities)?"],test =
"chi-square",expected_freqs = True, prop = "cell")

#expected frequencies
expected
#test results
test_results
```

5. Comparison for savings with demographic factors (gender)

H_0: There is no significant relationship between gender and saving type
H_1: There is a significant relationship between gender and saving type

Table:



Test Statistic: $\chi^2 = 11.091$

p-value = 0.026

α= 0.05

Decision criteria: Reject the null hypothesis if the **p-value** is less than **α**

Result: **p-value < α**
0.026 < 0.05
Reject null hypothesis

Conclusion: As the **p-value** is less than **α**, therefore we reject the null hypothesis and conclude that there is a significant relationship between the gender and saving type of Mumbaikars

.

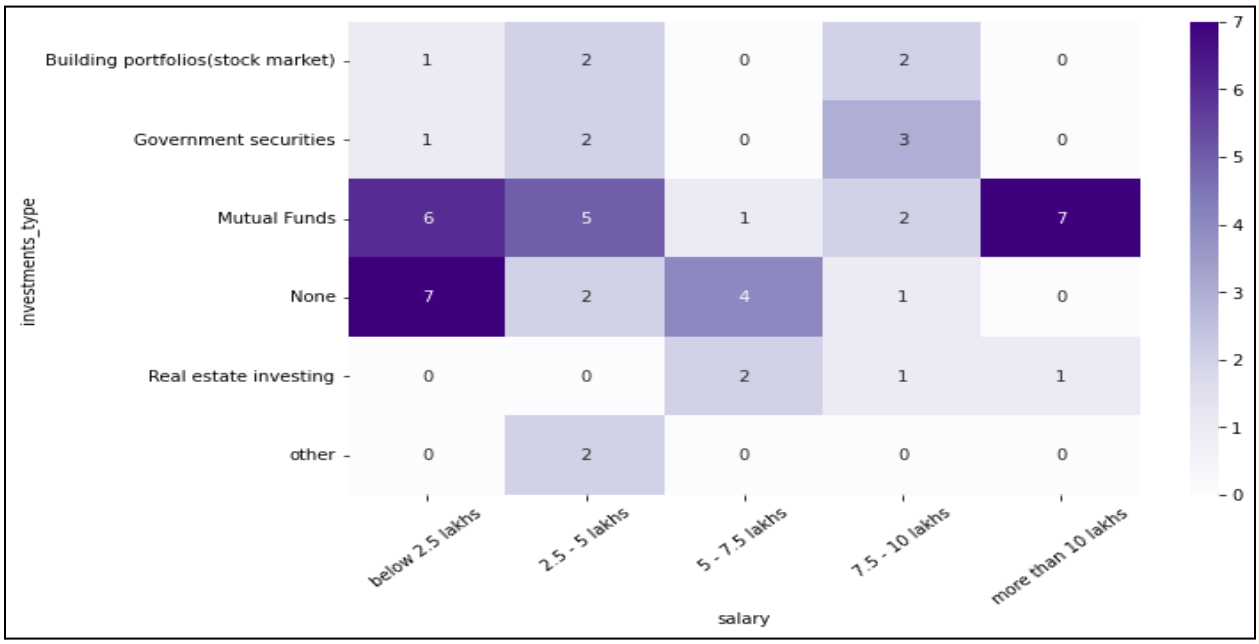
Python Code:

```
#Comparison for savings with demographic factors (gender)

# contingency table
saving_gender= pd.crosstab(df['savings_type'], df['gender'])
sns.heatmap(sav_gen, annot = True, cmap = 'binary')    #heatmaps
from scipy.stats import chi2_contingency
stats, p, dof, expected = chi2_contingency(saving_gender)

alpha = 0.05
if p > 0.05:
    print('variable are independent')
else:
    print('variable are dependent')
```

6. Comparison for investing with demographic factors(Salary)

H_0: There is no significant relationship between salary and investment type
H_1: There is a significant relationship between salary and investment type

Table:



Test Statistic: $\chi^2 = 36.429$

p-value = 0.014

α = 0.05

Decision criteria: Reject the null hypothesis if the **p-value** is less than **α**

Result: **p-value < α**
0.014 < 0.05
Reject null hypothesis

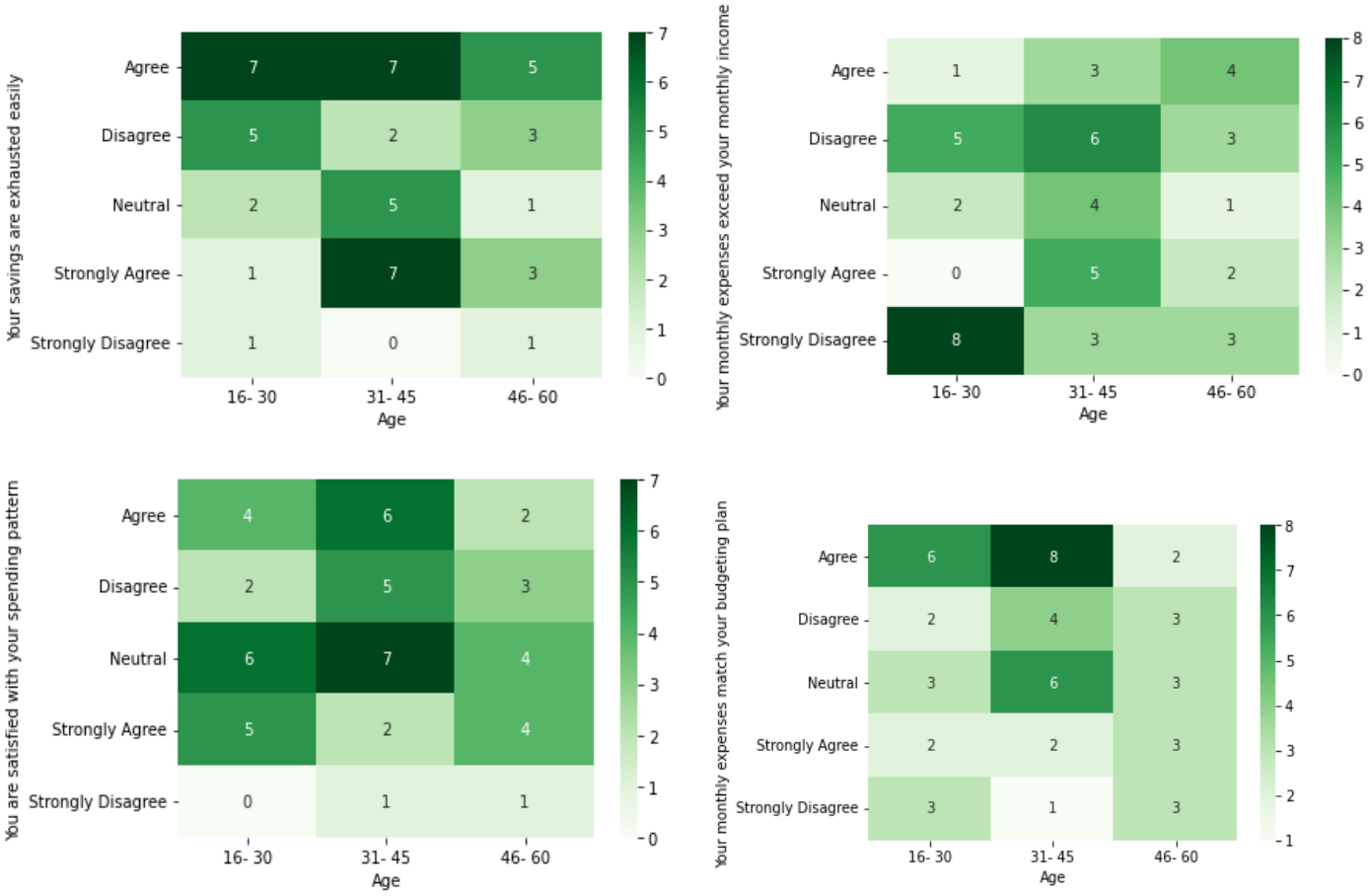
Conclusion: As the **p-value** is less than **α**, therefore we reject the null hypothesis and conclude that there is a significant relationship between the salary of Mumbaikers and how they do investment

Python Code:

```
#Comparison for investing with demographic factors(Salary)

#contingency table
investment_salary = pd.crosstab(df['investments_type'], df['salary'])
hp = sns.heatmap(inv_sal, annot = True, cmap = 'Purples')  #heatmaps
hp.set_xticklabels(a.get_xticklabels(), rotation = 40)
from scipy.stats import chi2_contingency
stats1, p1, dof1, expected1 = chi2_contingency(investment_salary)

alpha = 0.05
if p > 0.05:
    print('variable are independent')
else:
    print('variable are dependent')
```

7. Likert Scale Analysis

Heatmaps were made to compare the opinions of respondents against their age. For this age was divided into 3 categories: 15-30 which is the transition phase from studying to the working and focus is on the individual expenses, 31-45 where major decisions (in terms of spending) are made- eg. having a family or buying a property, and 46-60 where people are preparing for retirement.



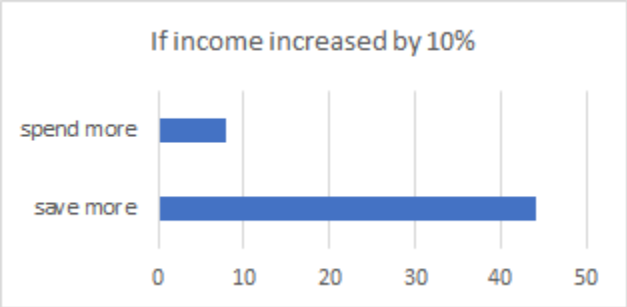
- The last age group has the least people in agreement that their savings are easily exhausted, as their main focus is on saving.
- Maximum respondents in the first age group strongly disagree that their monthly expenses exceed their income, as their expenses at their age is minimum.
- Majority of people in the first and second age groups are neutral with their satisfaction about their spending, while there is a mixed response for the last age group.
- Respondents below 45 years, affirm that their expenditure is according to their budgeting plan, which shows their proficiency in managing their expenses.

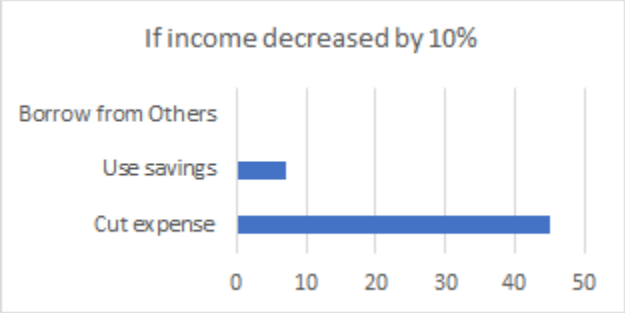
Python Code:

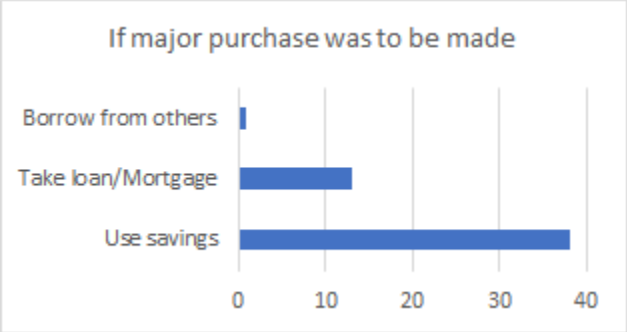
```
df1 = df   #copy of original dataframe
l1 = []
for i in range(len(df1.index)):
  temp = df.at[i,'age']
  if (temp< 30 and temp >=15):
     l1.append('16- 30')
  elif (temp< 45 and temp >=30):
     l1.append('31- 45')
  else:
     l1.append('46- 60')
df1['age_cat'] = l1     #add a new column with age categories

#Age category vs. statement opinion statement
#selects values
contingency_tab = pd.crosstab(df1['Your savings are exhausted
easily'],df1['Age'])   sns.heatmap(contingency_tab, cmap ='Greens',
annot = True) #heatmaps

contingency_tab = pd.crosstab(df1['Your monthly expenses exceed your
monthly income'],df1['Age'])
sns.heatmap(contingency_tab, cmap ='Greens', annot = True)

contingency_tab = pd.crosstab(df1['You are satisfied with your
spending pattern'],df1['Age'])
sns.heatmap(contingency_tab, cmap ='Greens', annot = True)

contingency_tab = pd.crosstab(df1['Your monthly expenses match your
budgeting plan'],df1['Age'])
sns.heatmap(contingency_tab, cmap ='Greens', annot = True)
```

8. Perceptions of respondents:

The respondents were given three "what-if" situations and were asked for how they would respond to it.



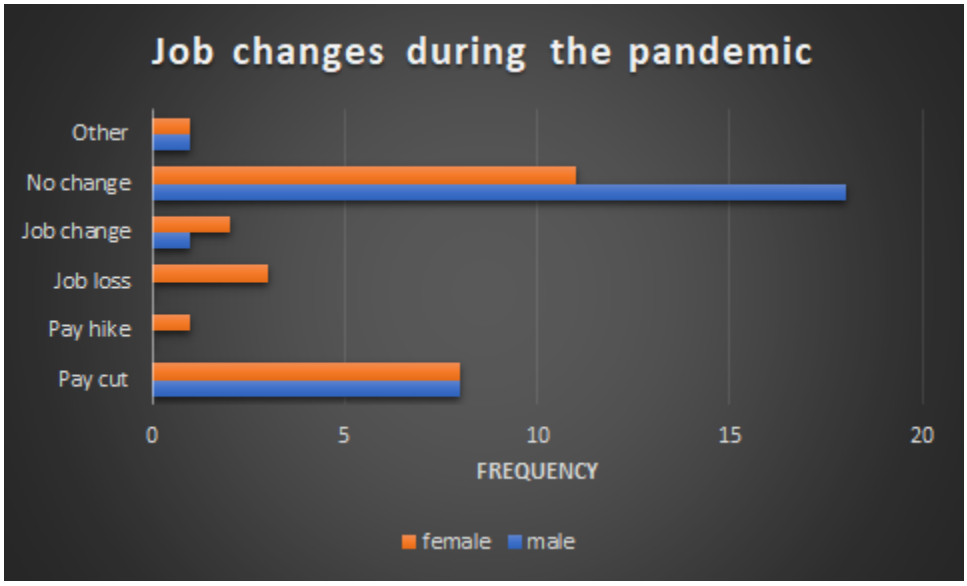




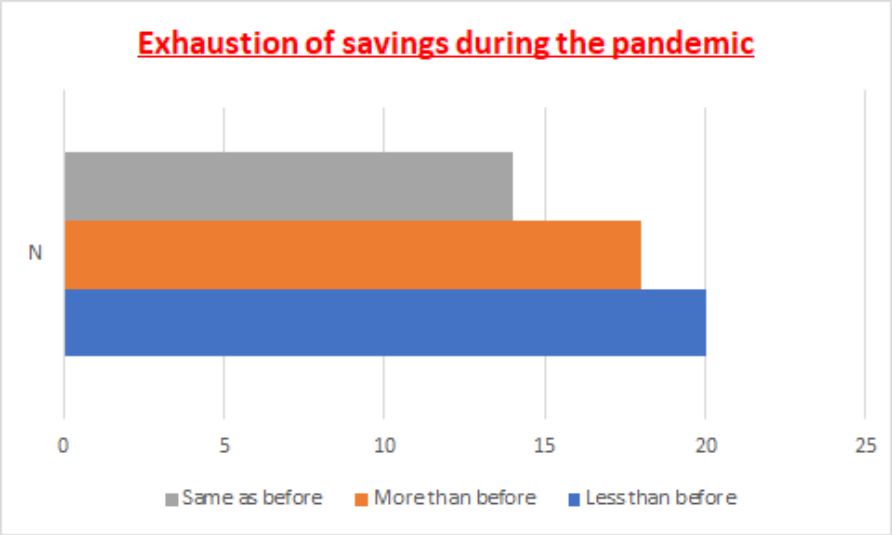
- Saving money is the topmost priority for most Mumbaikars and is used only for major purchases. This can be inferred as cutting savings is less favourable if income decreases and if income increases saving it is preferred by most.
- People prefer to manage their financial difficulties either by themselves or by the services of the bank, rather than borrowing from others.

9. Impact of Pandemic on Spending Pattern:

Due to the pandemic there have been drastic changes in lifestyle of people, which has affected their spending habits. To measure this aspect, we have performed an analysis on income of Mumbaikars on the basis of gender.

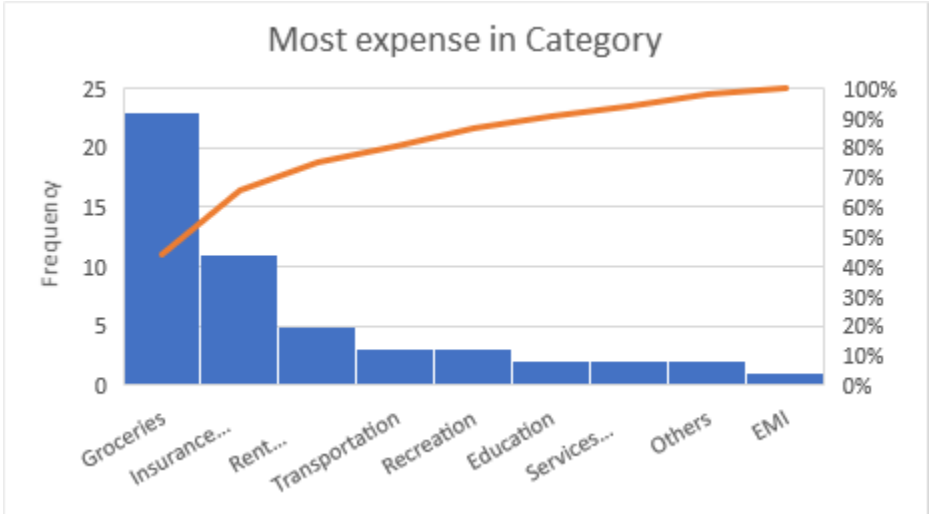


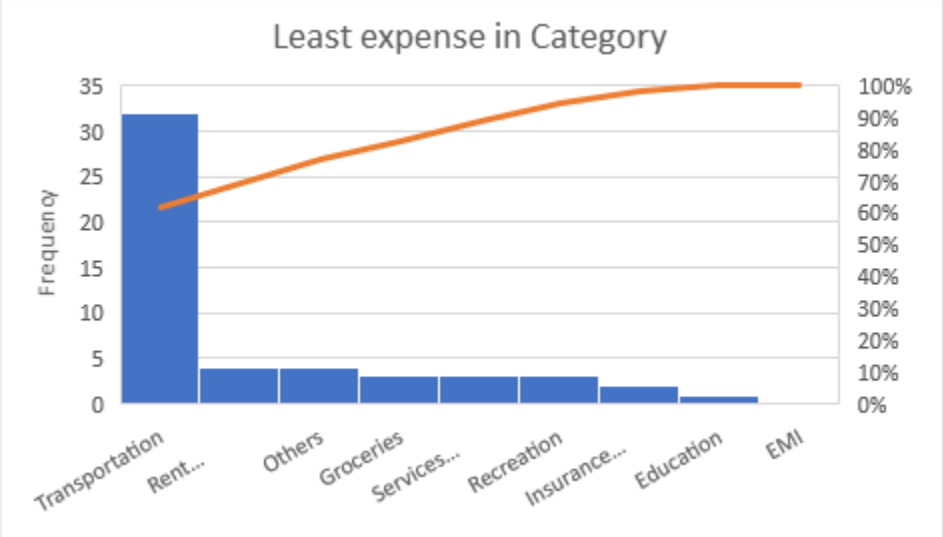
- The maximum number of our respondents both male and female, had “No change” in their income.
- The problem faced by the majority of respondents is pay cut, while job loss and job change have equivalent respondents.
- In all cases excluding “No change”, it is observed that female respondents have been affected more as compared to males.



- It is no surprise that the result of category “Less than before” has the maximum frequency as it is obvious that transportation was a standstill, not just for the bread-winner of the family but also for his members. Also outings or vacations of any form were nullified due to the pandemic which eventually led to more savings at the end of each month
- The category “More than before” clearly implies that the average number of respondents faced financial crisis or debts to be paid which was impossible as they probably would be facing job loss or pay cut which caused their savings to be exhausted at a faster pace. Also, during the pandemic Mumbai was in the RED ZONE of coronavirus so, these respondents may probably have additional medical expenses.

Further, the categories which had the most and least expenditure during the lockdown are displayed below using Pareto charts.





- During the pandemic the categories of Groceries, Insurance/Healthcare and rent/Maintenance contributed to 80% of the most expenses while, Transportation, rent/Maintenance and other miscellaneous expenses contributed to 80% of the least expenses among all the categories.
- From the above, the supermarkets, insurance firms and healthcare institutions would have seen a boost in profits during the pandemic.
- Household cost generally consumes the maximum income of a Mumbaikar but, here it is noted to have mildly affected people's expenses. The reason for the low expense could be that people have relocated to places with lower cost or people already own a house.

Conclusion:

The aim of this study was to observe various patterns of spending and saving of individuals who earn and live in Mumbai. The average budget of a Mumbaikar follows the 50/20/30 budget rule as the average amount spent on essential items is 50%, on non-essential items is 23%, while the remaining 27% is either saved or invested. People in the age category of 16-30 years have minimal expenses, while people of 31-45 years have numerous expenses as their expenses often are higher than their earnings, and the people of 46-60 years concentrate on saving. It was observed that most of the people earn below 5 lakhs in this sample. In general, everyone tends to spend more on essentials over spending on non-essentials and saving.

Earnings of any individual had a clear and pressing impact on whether they choose to save, invest or both. In addition, gender and income have a significant relation with the mode of saving and investing, respectively. It was also found that males are more likely to save than females. Given a raise, people tend to save rather than spend, while in case of a pay cut, they choose to cut down on expenses over utilising savings. Most respondents, including youngsters, are mature in their financial decisions as saving is their top priority and their expenses are managed as per their budget. This is crucial as it will ensure funds for emergencies and financial security for retirement.

In an extension of the above findings, we further examined the effects or its lack thereof of pandemic on the spending and saving patterns. Many of the respondents had to tackle the issues of a deducted income, irrespective of their gender while some of them even reported to have gone through no change in the aspects of job and salary. During the lockdown, there was a significant increase of expenditure on groceries while they enjoyed the benefits of a depleted transportation cost.

References:

1) https://courses.lumenlearning.com/odessa-introstats1-1/chapter/contingency-tables/
2) Think Stats: Probability and Statistics for Programmers- By Allen B. Downey
3) Practical Statistics for Data Scientists- By Peter Bruce and Andrew Bruce
4) https://www.newyorkfed.org/medialibrary/media/research/microeconomics/interactive/downloads/sce_household-spending-questionnaire.pdf