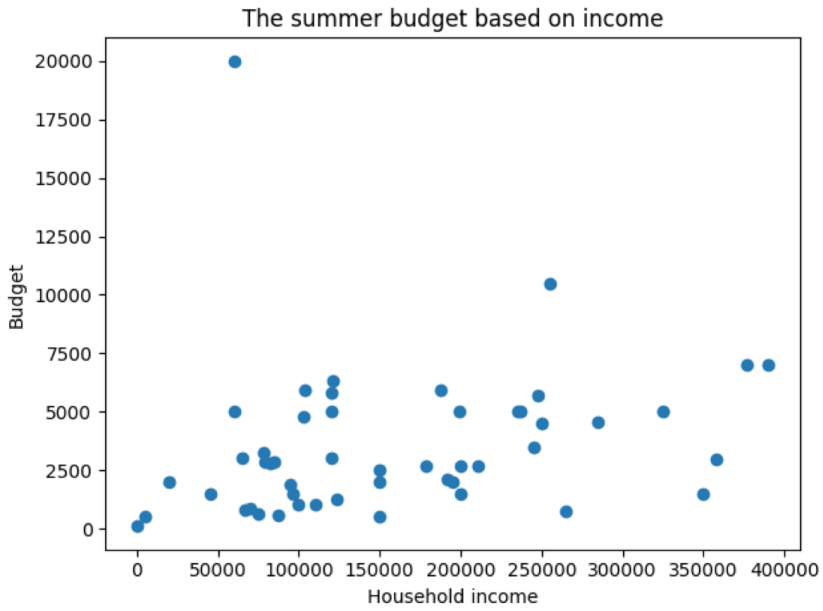
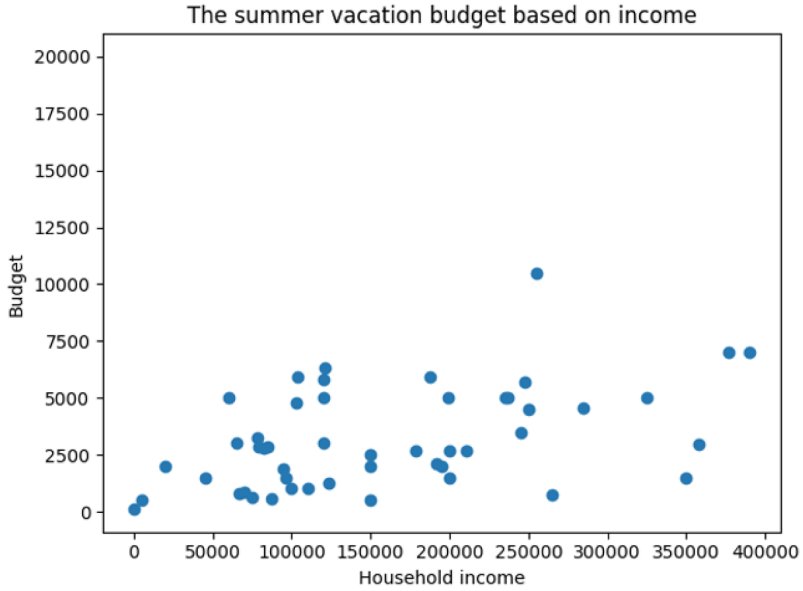
Scatter Plot (Household income & Sumer Budget)

```
# Scatter Plot for Income and Summer Budget
x = df.income
y = df.budget
plt.title("The summer budget based on income")
plt.scatter(x,y)
plt.show()
```

Before



After



Analysis

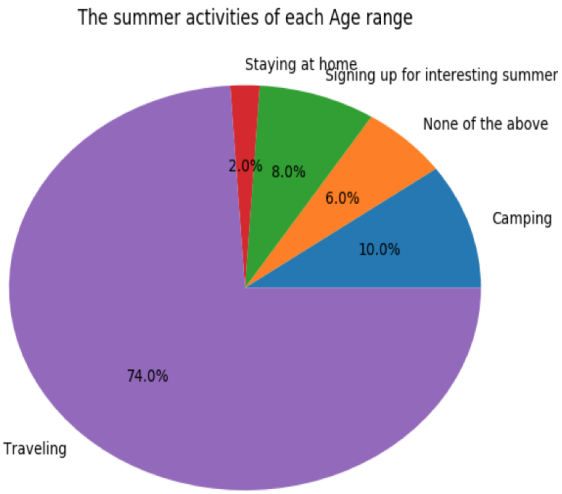
x-axis: Household income

y-axis: vacation Budget

From scatter plot we conclude that **strong positive linear correlation between x and y.**
It Means that if household income is Increase then budget also increased

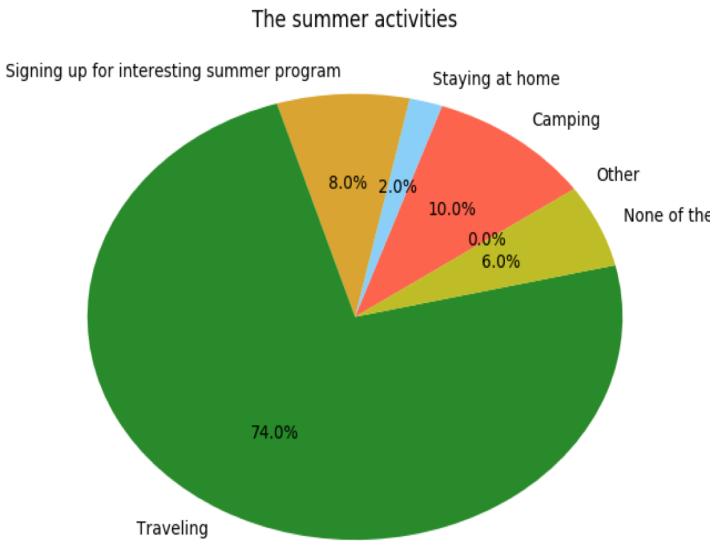
Pie chart (Age & Summer Vacation)

```
vac_col = df.groupby(["vacation"]).count().reset_index()
x = vac_col['Age']
y = vac_col['vacation']
age_col = df.index
plt.title("The summer activities of each Age range")
plt.pie(x, labels = y, autopct='%1.1f%%')
plt.show()
```



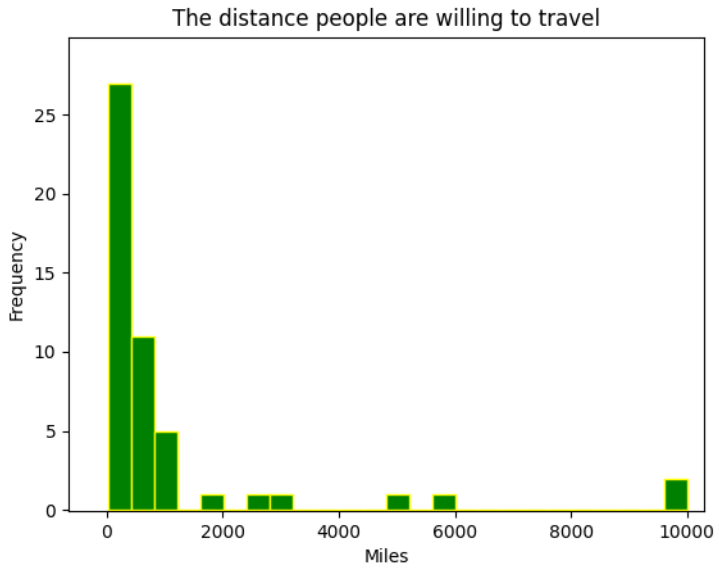
Pie chart (Summer Vacation)

```
vac_type = pd.read_csv('/Users/maithai/Desktop/NPU/IT501/Project/Part
B/Vacation_types.csv')
vac_type_col = vac_type['vacations']
vac_results = vac_type['vacation_results']
colors = ['tomato', 'lightskyblue', 'goldenrod', 'forestgreen', 'y','pink']
plt.title("The summer activities")
plt.pie(vac_results, labels = vac_type_col, colors = colors, autopct='%1.1f%%', startangle=35)
plt.show()
```



Histogram (Summer Miles)

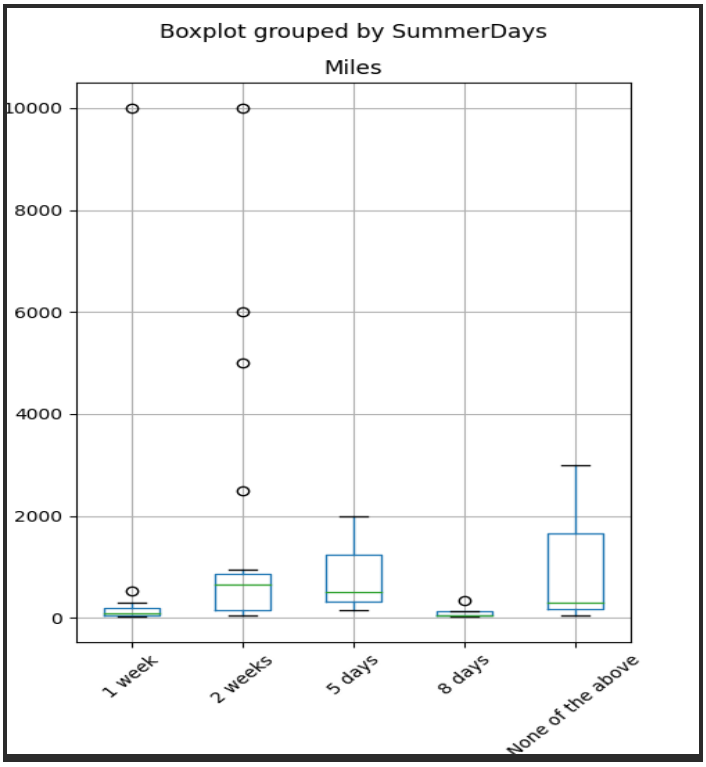
```
plt.hist(df['summer_miles'], bins=25, color='green',
edgecolor="yellow")
plt.xlim([50,11000])
plt.ylim([0,30])
plt.title("The distance people are willing to travel")
plt.xlabel("Miles")
plt.ylabel("Frequency")
plt.show()
```



Boxplot (Miles & Summer Days)

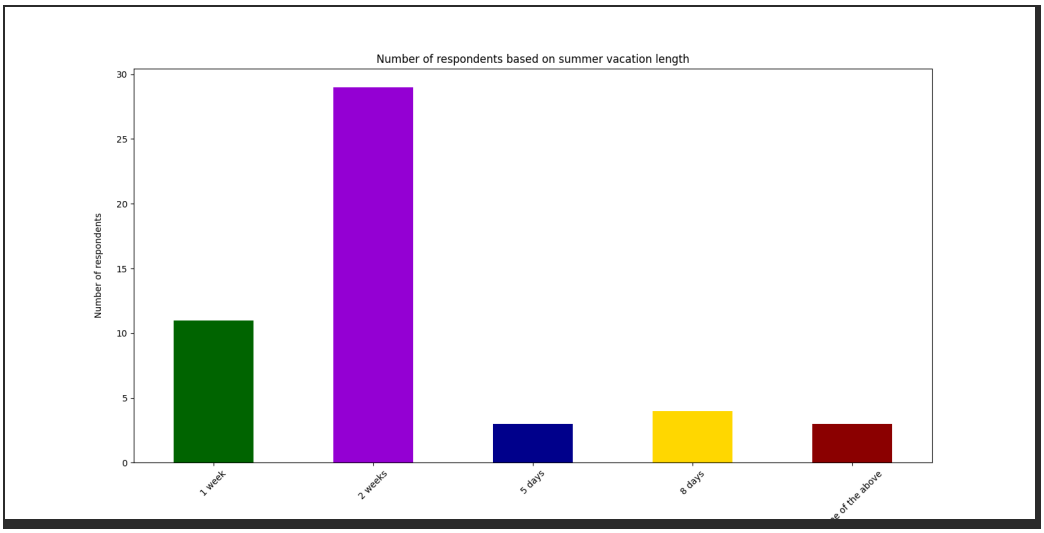
```
import pandas as pd
import matplotlib.pyplot as plt

df = pd.read_csv("Summer Mania.csv", encoding="ISO-8859-1")
miles_days = df.boxplot(column=['Miles'], by='SummerDays',
figsize=(5,7),grid=True)
plt.ylabel('Miles')
plt.xticks(rotation=45)
plt.show()
```



Bar Plot (Length of Summer Vacation)

```
df1=(df.groupby('SummerDays').size())
colors = ['darkgreen','darkviolet','darkblue','gold','darkred']
df1.plot(kind='bar', figsize=(5,7), rot=45, color = colors)
plt.xlabel("The length of summer vacation",labelpad=14)
plt.ylabel("Number of respondents",labelpad=14)
plt.title("Number of respondents based on summer vacation length")
plt.show()
```


Number of respondents based on summer vacation length
Number of respondents
0
5
10
15
20
25
30
1 week
2 weeks
5 days
8 days