Question 2:

This question is an extension of Question 1. Complete the steps in the **Jupyter notebook used in the PREVIOUS question**:

Step 1: Perform chi-square test for correlation analysis between categorical attributes "Pclass" and "Survived" in the titanic dataset. To complete this step, just run the code block under Problem 2 in the Jupyter notebook.

Step 2: Using the chi-square test result, answer the following question.
For a significance level of **0.05**, can you reject the null hypothesis that the two attributes are independent? Explain briefly. (Hint: Calculate the degrees of freedom and use the chi-square distribution table to get the critical value).

Note: Edit your answer in the Jupyter notebook and upload in the previous question.

Your Answer:

Yes you can reject the null hypothesis because there are two degrees of freedom, as calculated by the number of rows in the contingency table minus one times the number of columns minus one, and the critical value for two degrees of freedom and a significance level of 0.05 is 5.99 which is less than the calculated test value of 102.89.

Question 3:

Using the VOO 5-Year Historical Quotes csv file provided in the data folder in hw2_Fall23.zip

Download hw2_Fall23.zip

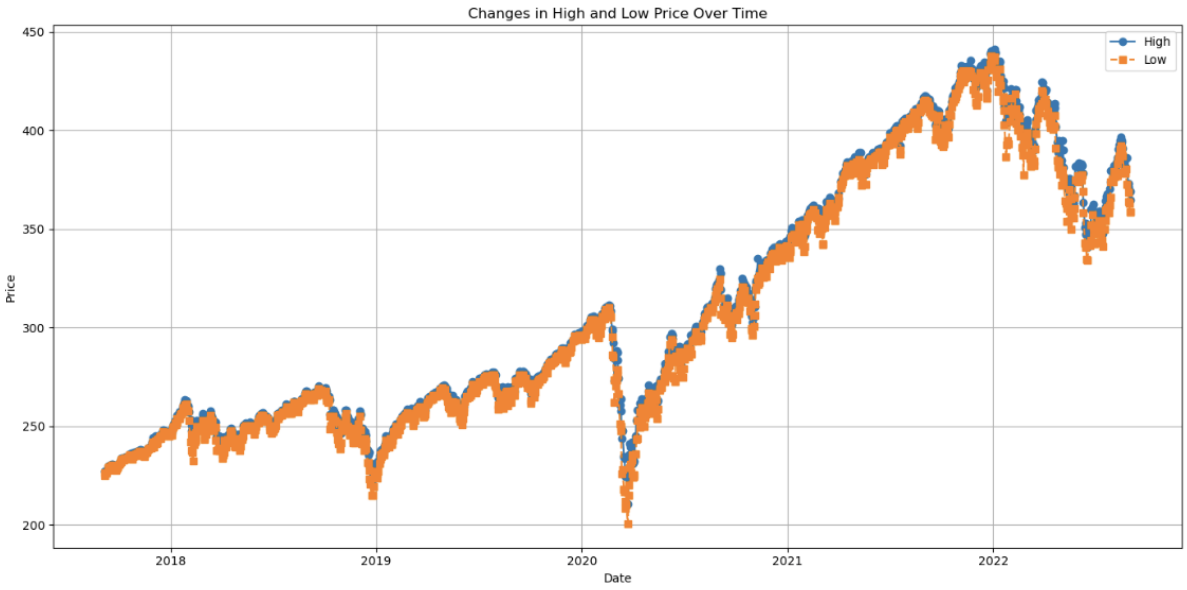
and the same Jupyter Notebook provided in the previous question, plot the following:

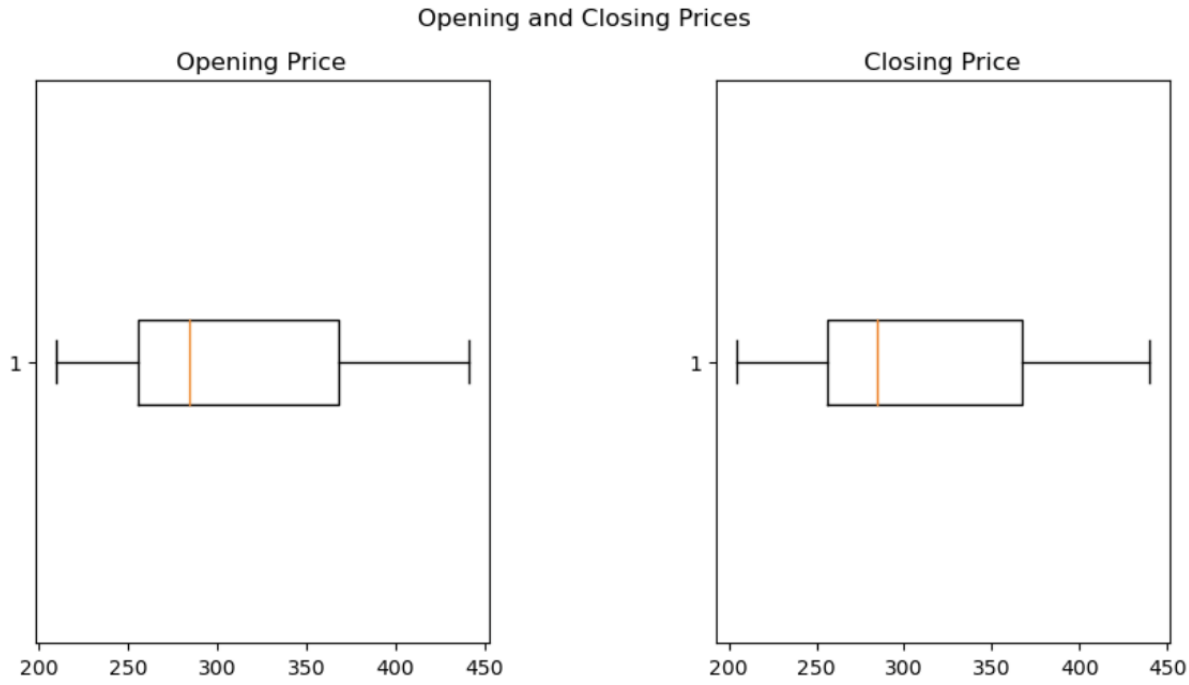
(a) A single plot showing the temporal change of the “High” and “Low” attributes.

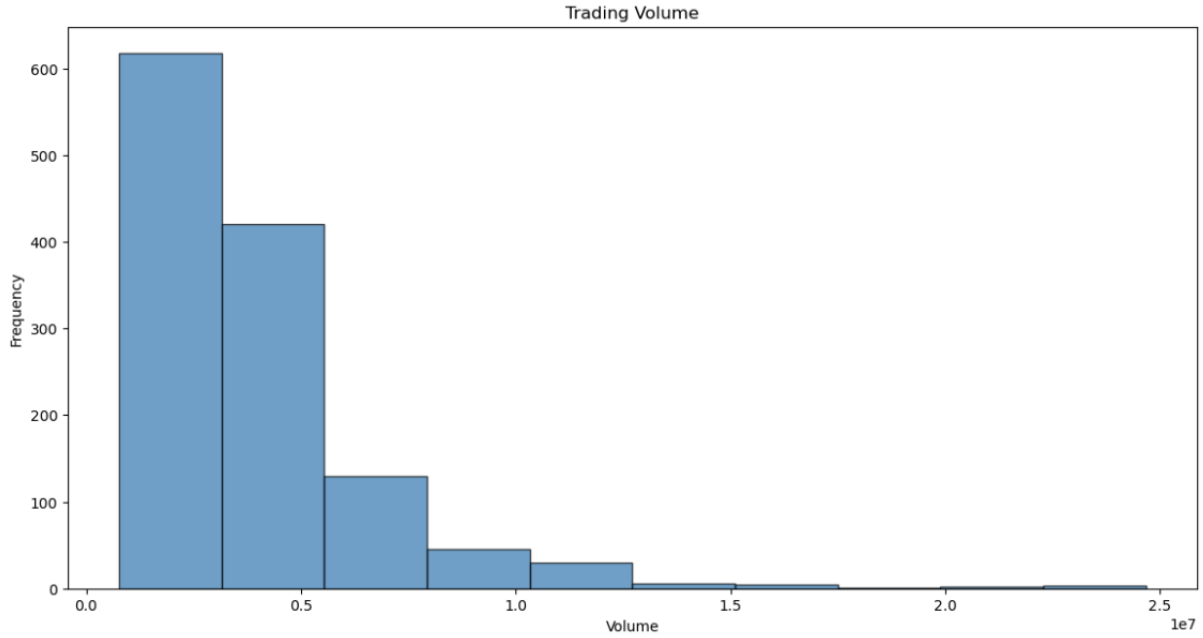
(b) A boxplot for the “Open” and “Close/Last” attributes.

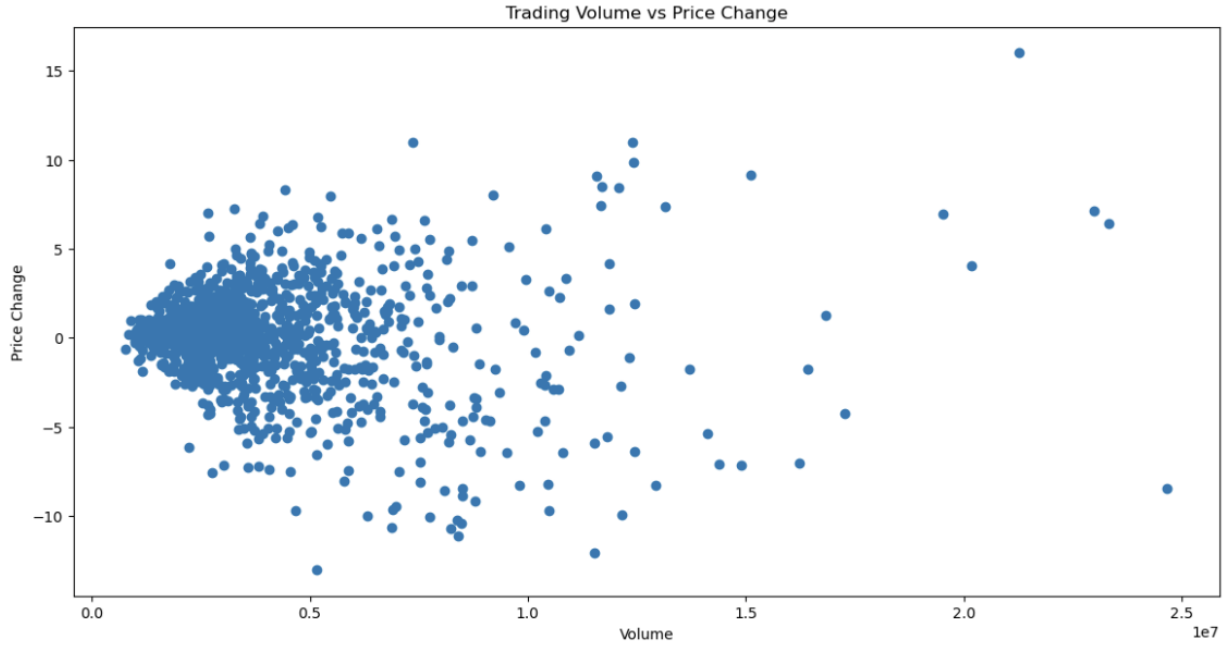
(c) The 10-bin equal-width histogram for the “Volume” attribute.

(d) Any other plot that interests you (using this VOO dataset).


Changes in High and Low Price Over Time
High
Low
Price
Date
450
400
350
300
250
200
2018
2019
2020
2021
2022


Opening and Closing Prices
Opening Price
Closing Price
1
1
200
250
300
350
400
450
200
250
300
350
400
450


Trading Volume
Frequency
Volume
1e7
0
100
200
300
400
500
600
0.0
0.5
1.0
1.5
2.0
2.5


Trading Volume vs Price Change
Price Change
Volume
1e7
15
10
5
0
−5
−10
0.0
0.5
1.0
1.5
2.0
2.5