Scenario

We want to see whether there is any difference in the average frequency of purchases by the customer from various cities.

Hypothesis

H0: There is no significant difference in average frequency of purchases among the customers from different cities.

Ha: There is a significant difference in average frequency of purchases among the customers from different cities.

Sig. Level

0.05 or 5%

Appropriate Test

One way ANOVA

(If normality assumed)

Performing Test

```
from scipy import stats

Chicago = preprocessed_data.query(‘City == "Chicago"’)['Frequency_of_Purchases’]
New_york = preprocessed_data.query(‘City == ”New York"’)['Frequency_of_Purchases’]
Houston = preprocessed_data.query(‘City == ”Houston"’)['Frequency_of_Purchases’]
Los_angeles = preprocessed_data.query(‘City == ”Los Angeles"’)['Frequency_of_Purchases']

t_statistic, p_value = stats.f_oneway(Chicago, New_York, Houston, Los_angeles )
print("P-value:", p_value)
```

Decision & Conclusion

```
alpha = 0.05

if p_value < alpha:

    print("Reject the null hypothesis. There is a significant difference in average frequency of purchases among the customers from different cities.")

else:

    print("Fail to reject the null hypothesis. There is no significant difference in average frequency of purchases among the customers from different cities.")
```