Project Title: Data Analysis for Detection of Credit Card Fraud

List at least 5 (five) data points that are required for the analysis and detection of a credit card fraud. (3 marks)

1 - The transaction values

2 - The units purchased above the average

3 - Frequency of orders

4 - The shipping address

5 - The IP address

Refer to the data table below and identify 3 (three) errors/issues that could impact the accuracy of your findings. (3 marks)

Table

Description automatically generated

1 - The missing data on the IP column, row 5

2 - The missing data on the IP column, row 9

3 - The missing data on the Transaction value column, row 3

4 - The transaction date column is not standardized and uniform

Refer to the data table below and identify 2 (two) anomalies or unexpected behaviors, that would lead you to believe the transaction may be suspect. (2 marks)

Table

Description automatically generated

- Johnp is making larger purchases than usual after 3-6-20, and its heavy on electronics

- Johnp has shifted to in-store purchases along with a PO box, and his IP address changed

- Ellend changed the PO Box in July and made a significant purchase of a high-end laptop.

Briefly explain your key take-away from the provided data visualization chart. (1 mark)

Chart, line chart

Description automatically generated

The transaction value on 2 is trending higher than normal for both ellend and johnp, and johnp’s consumption varies considerably. Making them both suspect of fraud.

Identify the type of analysis that you are performing when you are analyzing historical credit card data to understand what a fraudulent transaction looks like.

Descriptive and summary statistics - By analyzing the past data we can find average and variances of customer purchases with the data points, thus determining anomalies in pattern behavior.