

|  |
| --- |
| Stopping Credit Card Fraud using data analysis |
| data Anaysis in action |
|  |
| **Andrea Plunkett Jackson** |**Coursera** | **September 23, 2023** |

# Introduction to Data Analytics: FInal Assignment

1. The primary data points required for the analysis and detection of credit card fraud can be placed within three categories – transactional history, purchase behavior pattern and source data. The breakdown of each category as shown below.

Transactional history

1. Amount of Purchase
2. Items Purchased

Purchase Behavioral Patterns

1. Frequency of Purchases
2. Address of Delivery

Source of Data

1. Location of Purchase
2. Device Used
3. IP address of Order Placement online
4. According to the data table three errors/issues that could impact accuracy of findings would be
   * + 1. Varied address formatting
       2. Inconsistent transaction date entry
       3. Missing information ( ex. IP address and Transaction Value)

A table with numbers and numbers

Description automatically generated

1. Two anomalies or unexpected behaviors that lead me to believe the transaction may be suspect are missing and change of IP address for johnp in addition to change of delivery method/address. These outliers in combination to the change in transactional value increase illuminating suspected fraud.

A table with numbers and letters

Description automatically generated with medium confidence

1. The key takeaways from the visualization chart (shown below) are the exponential increase in transaction values purchased for johnp between TRN#3 and TRN #5 as well as ellend for TRN#2. Sudden changes in purchase behavior pattern and transactional data can be an indication of fraud.

A graph of a transaction value

Description automatically generated

1. The type of analysis I used when performing an analysis of historical credit card data to understand the appearance of fraudulent transactions is diagnostic analytics. The analytic type helps in the identifying patterns and trends in order to examine outliers based on the primary data points discussed on question one,