

Name: Ismi Ana S

Date: October 25th,2025

Course 1: Introduction to Data Analytics

USING DATA ANALYSIS FOR DETECTING CREDIT CARD FRAUD

Companies today are employing analytical techniques for the early detection of credit card frauds, a key factor in mitigating fraud damage. The most common type of credit card fraud does not involve the physical stealing of the card, but that of credit card credentials, which are then used for online purchases.

In this case, I will take a position as Data Analyst to work in the Credit Card Division of a bank. And my first assignment is to join the team in using data analysis for the early detection and mitigation of credit card fraud. In order to prescribe a way forward, that is, suggest what should be done in order for fraud to get detected early on, I need to understand what a fraudulent transaction looks like. And for that I will start by looking at historical data.

Here is a sample data set that captures the credit card transaction details for a few users.

IP Address	User ID	Account Number	Age	Shipping Address	Transaction Date	Transaction Time	Transaction Value	Product Category	Units Purchased
3.56.123.0	johnp	25671147	32	1542, Orchid Lane, WA 98706, US	15-5-20	15:00:05	\$121.58	Clothing	1
3.56.123.0	johnp	25671147	32	1542, Orchid Lane, WA 98706, US	10-6-20	10:23:10	\$79.23	Electronics	2
3.56.123.0	johnp	25671147	32	1542, Orchid Lane, WA 98706, US	1-6-20	07:12:45		Home Décor	1
1.186.52.7	johnp	25671147	32	In-store	3-6-20	01:11:10	\$2,009.99	Electronics	10
	johnp	25671147	32	In-store	2020-06-03	01:15:12	\$4,131.00	Electronics	15
1.186.52.7	johnp	25671147	32	P.O. Box 1049	03-06-2020	01:22:24	\$3,010.50	Tools	20
1.58.167.2	davidg	51422789	47	90 Robinson Blvd, Alberta, 97602, Canada	15 May 2020	17:02:08	\$234.20	Furniture	1
1.58.167.2	davidg	51422789	47	90 Robinson Blvd, Alberta, 97602, Canada	18 May 2020	19:12:45	\$141.00	Kithcen Supplies	3
	davidg	51422789	47	90 Robinson Blvd, Alberta, 97602, Canada	01 June 2020	17:34:15	\$157.25	Car Spares	2
1.58.167.2	davidg	51422789	47	90 Robinson Blvd, Alberta, 97602, Canada	13 June 2020	18:02:10	\$59.99	Kithcen Supplies	1
172.165.10.1	ellend	11568528		P.O. Box 1322	07 June 2020	15:53:12	\$99.99	Clothing	1
172.165.10.1	ellend	11568528		P.O. Box 1322	08 June 2020	17:15:30	\$53.15	Beauty	1
1.167.255.10	ellend	11568528		P.O. Box 5401	02 July 2020	00:05:10	\$4,895.00	Laptop	1

Descriptive techniques of analysis, that is, techniques that help me gain an understanding of what happened, include the identification of patterns and anomalies in data. Anomalies signify a variation in a pattern that seems uncharacteristic, or, out of the ordinary. Anomalies may occur for perfectly valid and genuine reasons, but they do warrant an evaluation because they can be a sign of fraudulent activity.

1. List the data points that are required for the analysis and detection of a credit card fraud.

- Card holder /
- Transaction time
- IP address
- Customer Id
- Transaction value
- Device model
- Transaction date
- Shipping address
- Location

2. Identify any errors/issues that could impact the accuracy of findings, based on a data table.

Data cleaning (or cleansing) is the process of identifying and correcting or removing errors, inconsistencies, and inaccuracies in raw data to improve its quality and make it ready for analysis. I need to identify and fix issues in the data that can lead to false or incomplete findings. Some of the issues that

could impact the accuracy of findings include: Missing transaction value & Missing IP Address, then a Date format inconsistency, that I will need to standardize data formats.

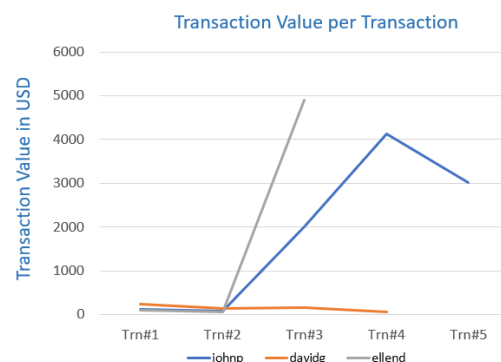
3. Identify any anomalies, or unexpected behaviors, that would lead to believe the transaction may be suspect, based on a data table provided.

Some of anomalies in the provided data set include:

- A sudden change in delivery preference. Significantly higher *Transaction Value* where *Shipping Address* has been changed from home/office address to P.O. Box.
- Higher Transaction Value and increased frequency of transactions. A change in frequency of orders placed, for example: a customer who typically places a couple of orders a month, suddenly makes numerous transactions within a short span of time, sometimes within minutes of the previous order.
- A mismatched IP Address, or an IP Address that is not from the general location or area of the billing address: *IP Address* change and significantly higher *Transaction Value* and *IP Address* change and *Shipping Address* change
- Orders that are significantly higher than a user's average transaction.
- Bulk orders of the same item with slight variations such as color or size—especially if this is atypical of the user's transaction history.

4. Briefly explain your key take-away from the provided data visualization chart.

The graph below samples one such visualization that would be use to capture a trend hidden in the sample data set shared earlier on in the case study.



The visualization depicts the transaction values per transaction for all three users. The key take-away from this visualization is the sharp rise in the transaction values for users johnp and ellend, which may be indicative of an anomaly.

Name: Ismi Ana S

Date: October 25th, 2025

Course 1: Introduction to Data Analytics
