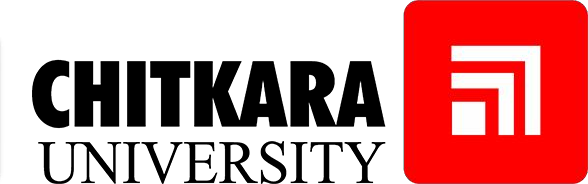
Artificial Intelligence and Machine Learning

Project Report Semester-IV (Batch-2022)

Case Study - Amazon Purchases Dataset [URL:-](file:///C:\Users\admin\Downloads\-)

<https://drive.google.com/file/d/1NwFI1H9H7duuIHzYykNYHLsvTen2a-Jq/view?usp=sharing>



**Supervised By: Submitted By:**

Rajeev Bhardwaj Tanisha

2210990891

**G-13**

**Department of Computer Science and Engineering**

Chitkara University Institute of Engineering & Technology,

**Chitkara University, Punjab**

**Description about Case Study:-**

* Display Top 10 Rows:

How can we view the top 10 rows of our dataset?

* Display Last 10 Rows:

How can we view the last 10 rows of our dataset?

* Check Datatype of Each Column:

How do we check the datatype of each column in our dataset?

* Check Null Values in the Dataset:

How can we identify and handle null values in our dataset?

* Number of Rows and Columns in the Dataset:

How do we determine the total number of rows and columns in our dataset?

* Highest and Lowest Purchase Price in Dataset:

Where can we find the highest and lowest purchase prices in our dataset?

* Average Purchase Price:

How can we calculate the average purchase price from our dataset?

* People with French as Their Language:

Where can we find information on the number of people with 'fr' (French) as their language?

* Job Title Contains 'Engineer':

How do we filter and display individuals whose job title contains the term 'Engineer'?

* Find Email by IP Address (132.207.160.22):

Where can we find the email of the person with the IP address 132.207.160.22?

* Mastercard Holders with Purchases Above 50:

How many people have Mastercard as their Credit Card Provider and made a purchase above 50?

* Find Email by Credit Card Number (4105595335494659):

Where can we find the email of the person with the Credit Card number 4105595335494659?

* People Purchasing During AM and PM:

How many people make purchases during the AM and how many during the PM?

* Credit Cards Expiring in 2020:

How many people have a credit card that expires in the year 2020?

* Top 5 Most Popular Email Providers:

Where can we find information on the top 5 most popular email providers in the dataset (e.g., Gmail.com, Yahoo.com, etc.)?.

# Library:-.

* Pandas

# Methods:-

1. **Display Top 10 Rows:-**

Method: Use the head() function in Python (Pandas) or a similar function in other data analysis tools.

# Display Last 10 Rows:

Method: Use the tail() function in Python (Pandas) or a similar function in other data analysis tools.

# Check Datatype of Each Column:

Method: Utilize the dtypes attribute in Python (Pandas) or similar functions in other tools

# Check Null Values in the Dataset:-

Method: Use isnull() or info() function in Python (Pandas) or similar functions in other tools.

# Number of Rows and Columns in the Dataset:

Method: Use the shape attribute in Python (Pandas) or similar functions in other tools.

# Highest and Lowest Purchase Price in Dataset:

Method: Use max() and min() functions on the Purchase Price column in Python (Pandas) or similar functions in other tools**.**

# Average Purchase Price:

Method: Use mean() function on the Purchase Price column in Python (Pandas) or similar functions in other tools.

# People with French as Their Language:

Method: Filter the dataset based on the 'Language' column where the value is 'fr'.

# Job Title Contains 'Engineer':

Method: Filter the dataset based on the 'Job Title' column for entries containing the term 'Engineer'.

# Find Email by IP Address (132.207.160.22):

Method: Search the dataset for the row where the IP address matches and retrieve the corresponding email.

# Mastercard Holders with Purchases Above 50:

Method: Filter the dataset for entries with 'Mastercard' as the Credit Card Provider and 'Purchase Price' above 50.

# Find Email by Credit Card Number (4105595335494659):

Method: Search the dataset for the row where the Credit Card number matches and retrieve the corresponding email.

# People Purchasing During AM and PM:

Method: Use the time-related information in the dataset to categorize purchases as AM or PM and count accordingly.

# Credit Cards Expiring in 2020:

Method: Extract the year from the 'Credit Card Expiry' column and filter for entries with the year 2020.

# Top 5 Most Popular Email Providers:

Method: Extract the email domain from the 'Email' column and use value\_counts() or similar functions to find the top 5.