## Artificial Intelligence and Machine Learning

Project Report

Semester-IV (Batch-2022)

**Case Study**: - Amazon Purchases Dataset

[Url:-](about:blank) <https://drive.google.com/file/d/1ryUQ-dZVByn58ATs7oDJub-kEcqSw93w/view?usp=drive_link>

A red and white sign

Description automatically generated with low confidence

| **Supervised By:**  **Rajeev Bhardwaj** | **Submitted By:**  **Udaibir Singh Bhathal**  **Roll Number: 2210990916**  **Group: 14** |
| --- | --- |

**Department of Computer Science and Engineering**

## Chitkara University Institute of Engineering & Technology,

## Chitkara University, Punjab

**Description about Case Study: -**

* Read dataset Amazon Purchases
* Display Top 10 rows
* Display the Last 10 rows
* Check the datatype of Each column
* Check null values in the Dataset
* How many rows and columns are in our dataset
* Highest and lowest purchase price in the dataset
* Average purchase price
* How many people have French 'fr' as their Language
* The job title contains engineer

**Library: -**

* Pandas

**Methods: -**

1. read\_csv():

Description: Reads a CSV file and converts it into a data frame.

1. tail():

Description: Displays the last few rows of the data frame.

1. head():

Description: Displays the first few rows of the data frame.

1. shape():

Description: Returns the shape (number of rows, number of columns) of the data frame.

1. info():

Description: Provides basic information about the data frame, such as column types and missing values.

1. isnull():

Description: Returns True/False for each value in the data frame, indicating whether the value is missing (NaN) or not.

1. sum():

Description: Calculates the sum of values in each column of the data frame.

1. contains():

Description: Checks if a specified substring or value is present in a column of the data frame.

1. max():

Description: Returns the maximum value in a column of the data frame.

1. min():

Description: Returns the minimum value in a column of the data frame.

1. mean():

Description: Calculates the mean (average) value of a column in the data frame.

1. len():

Description: Returns the number of rows in the data frame