## Artificial Intelligence and Machine Learning

Project Report

Semester-IV (Batch-2022)

**Case Study**: - Employee Salaries Assignment

[Url:-](about:blank) <https://drive.google.com/file/d/1kPp2MePpJKrft1PPID-r8erK58K7Vofi/view?usp=sharing>

A red and white sign

Description automatically generated with low confidence

**Supervised By: Submitted By:**

Rajeev Thakur Yashika Taneja

Roll Number: -2210990979

Group - 14

**Department of Computer Science and Engineering,** Chitkara University Institute of Engineering & Technology,

**Description about Case Study: -**

* Display first 10 rows of the Dataset
* Check Last 10 Rows of the Dataset
* Find Shape of our Dataset (Number of rows and number of columns)
* Getting info about our dataset like total number of rows, total number of columns,datatypes of each column and memory requirement.
* Check null values in the dataset
* Drop ID, Notes, Agency, Status columns.
* Find occurrence of the employee name(top 5)
* Find the number of unique job titles.
* Total number of jobs titles contains captains.
* Display all the employee names from fire department
* Find minimum,maximum, average basepay.
* Replace not provided in employeename column to Nan
* Find top 5 Most common jobs
* Find Average BasePay of Employee having Job title Accountant
* Find Average BasePay of All Employee Per job title
* Find average basepay of all employee per year
* Display name of the person having the highest basepay
* How much Albert Pardini make (include benefits)?
* Find the job title of Albert Pardini
* Drop the rows having more than 5 missing values

**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. **drop():**

Description: Removes specific rows or columns from the data frame.

1. **value\_counts():**

Description: Counts the unique values in a specific column of the data frame.

1. **nunique():**

Description: Returns the count of unique values in a specific 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

1. **value\_counts():**

Description: Counts the unique values in a specific column of the data frame.

1. **apply():** Description: Applies a function to transform the values in the data frame.
2. **replace()**:Description : This method is used to replace a specified value or values with another value in a dataframes or series.