## Front End Engineering-II /Artificial

## Intelligence and Machine Learning

Project Report

Semester-IV (Batch-2022)

Database Salaries

A red and white sign

Description automatically generated with low confidence

**Supervised By: Submitted By:**

MR. RAJIV BHARDWAJ TAMANNA

2210990885 (G-13)

**Department of Computer Science and Engineering**

## Chitkara University Institute of Engineering & Technology,

## Chitkara University, Punjab

**Description About the case study**

# **Display Top 10 rows**

# **Display Last 10 rows**

# **Check shape of our dataset(number of rows and number of columns)**

# **Total number rows,total number of columns,datatypes of each column and memory requirement**

# **Check null values in the dataset**

# **Drop id,notes,agency and status columns**

# **Find occurrence of the employee name(top5)**

# **Find the number of unique job titles**

# **Total number of jobs titles contains captain**

# **Display all the employee names from fire department**

# **Find minimum,maximum and averge basepay**

# **Replace 'not provided' in employeename column to nan**

# **Drop the rows having more than 5 missing values**

# **Find the job title of albert pardini**

# **How much albert pardini make (inculde denefits)?**

# **Display name of the person havinng the highest basepay**

# **Find averge basepay of all employee per year**

# **Find averge basepay of all employee per job title**

# **Find averge basepay of employee havig job title accountant**

# **Find top5 most comman jobs**

**Library**

**library used in this case study is pandas**

**Method**

**read\_csv():**

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

**tail():**

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

**head():**

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

**shape():**

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

**info():**

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

**isnull():**

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

**sum():**

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

**drop():**

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

**value\_counts():**

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

**nunique():**

**Description: Returns the count of unique values in a specific column of the data frame.**

**contains():**

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

**max():**

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

**min():**

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

**mean():**

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

**replace():**

**Description: Replace values in dataframe or series.**

**Loc[]:**

**Description: This index in pandas is used to access a group of rows and columns in a dataframe by lables.**

**groupby():**

**Description: Enables us to perform operations on grouped data based on specific conditions.**