1. PROJECT NO:

- 02

2. TITLE:

- San Francisco City Employee Salary Dataset

3. OBJECTIVES:

- Learn how to show top and bottom data
- Find the shape, info and describe the dataset
- Check null values and drop and fill
- Drop rows for a specific null values
- Total unique job and job contain specific titles
- People from specific department
- Convert string to float and perform mean, max. min
- Replace any value
- Calculation of each year, genre
- Using group by

4. DESIGN/ALGORITHM/FLOWCHART:

5. QUESTION:

1. Display Top 10 Rows of The Dataset
2. Check Last 10 Rows of The Dataset 3.
3. Find Shape of Our Dataset (Number of Rows and Number of Columns)
4. Getting Information About Our Dataset Like Total Number Rows, Total Number of Columns, Datatypes of Each Column and Memory Requirement
5. Check Null Values in The Dataset
6. Drop ID, Notes, Agency, and Status Columns
7. Get Overall Statistics About the Dataframe
8. Find Occurrence of the Employee Names (Top 5)
9. Find The Number of Unique Job Titles
10. Total Number of Job Titles Contain Captain
11. Display All the Employee Names from Fire Department
12. Find Minimum, Maximum, and Average BasePay
13. Replace 'Not Provided' in EmployeeName' Column to NaN
14. Drop The Rows Having 5 Missing Values
15. Find Job Title of ALBERT PARDINI
16. How Much ALBERT PARDINI Make (Include Benefits)?
17. Display Name of the Person Having the Highest BasePay
18. Find Average BasePay of All Employee Per Year
19. Find Average BasePay of All Employee Per JobTitle
20. Find Average BasePay of Employee Having Job Title ACCOUNTANT
21. Find Top 5 Most Common Jobs

6. IMPLEMENTATION:

```
# d.value_counts('Year')
d = x.copy()

d = d.groupby('Year')
for i, j in d:
  # print(i)
  # print(j['BasePay'].mean())
  print(j['BasePay'].mean())
```

```
63595.95651677314
65436.40685742263
68509.83215550712
66557.43774991475
```

```
n = x.copy()
n = n.groupby('Year')
```

7. OUTPUT:

```
JobTitle
Transit Operator                7036
Special Nurse                   4389
Registered Nurse                3736
Public Svc Aide-Public Works    2518
Police Officer 3                2421
Name: count, dtype: int64
```

8. ANALYSIS:

9. DISCUSSION:

- My Semester final exam is running. Tomorrow is my EEE-201 final exam. And I'm practicing Pandas library which is very weird. Bay the way, this exercise made me more confident about data cleaning with drop, fill and group and replace.

10. CONCLUSION:

11. REFERENCE:

- **Dataset:**
- **Tutorial:** YouTube Link
- **Online Resources:** Kaggle Dataset
- **Others:**