```
#MINI_PROJECT_01
#EXPLORATORY DATA ANALYSIS ON EMPLOYEES TABLE
#dataset - 'https://gist.githubusercontent.com/kevin336/acbb2271e66c10a5b73aacf82ca82784/raw/e38afe62e088394d61ed30884dd50a6826eee0a8/employe

import pandas as pd
df = pd.read_csv('https://gist.githubusercontent.com/kevin336/acbb2271e66c10a5b73aacf82ca82784/raw/e38afe62e088394d61ed30884dd50a6826eee0a8/e
df
```

```
#1. How many employees have ST.CLERK and ST.MAN as their job id:-
df.JOB_ID.value_counts()

#RESULT- 16 EMPLOYEES HAVE ST_CLERK AS THEIR EMPLOYEE ID
#        5 EMPLOYEES HAVE ST_MAN AS THEIR EMPLOYEE ID
```

```
#2. i WANT TO PRINT ONLY THOSE ROWS WHOSE SALARY_COL IS GREATER THAN RS. 8000:-
df[df.SALARY > 8000]
```

```
#3. Gives the info. about the dataframe
df.info()
```

```
#4. Number of not-null values in each row
df.count()
```

```
#5. Check the datatype of each column
df.dtypes
```

```
EMPLOYEE_ID        int64
FIRST_NAME        object
LAST_NAME         object
EMAIL             object
PHONE_NUMBER      object
HIRE_DATE         object
JOB_ID            object
SALARY             int64
COMMISSION_PCT    object
MANAGER_ID        object
DEPARTMENT_ID      int64
dtype: object
```