EMPLOYEES SALARY ANALYSIS

INTRODUCTION

HELLO, I'M **RITIK RAHI**, AN ASPIRING DATA ANALYST SHOWCASING MY SQL SKILLS THROUGH THIS PROJECT.

OBJECTIVE

The main objective of this project is to analyze employee salary data and extract valuable insights. The dataset was first uploaded into **Jupyter Notebook**, and using **SQLAlchemy**, it was imported into **MySQL** for storage and analysis. This project aims to understand salary distribution, identify pay gaps, and analyze factors influencing employee compensation
to assist in data-driven decision-making.

ABOUT THIS PROJECT

Employee salary analysis is essential for organizations to ensure fair compensation, identify salary trends, and optimize payroll management. This project utilizes **Jupyter Notebook** for data preprocessing, **SQLAlchemy** for database connectivity, and MySQL for performing advanced queries and salary-based analytics.

TOOLS USED

Jupyter Notebook: For data preprocessing and cleaning
SQLAlchemy: For connecting and importing data into MySQL
MySQL: For data extraction, transformation, and analysis

KEY INSIGHTS AND FEATURES

- Salary distribution analysis across different job roles and experience levels
- Comparison of employee salaries based on department and location
- Gender pay gap analysis to assess fairness in compensation
- Trends in salary growth over time for career progression insights
- Impact of bonuses and incentives on total compensation

CONCLUSION

This project demonstrates how **Jupyter Notebook, SQLAlchemy**, and **MySQL** can be effectively used for employee salary analysis. By leveraging data analytics, organizations can gain better insights into salary structures, ensure fair compensation policies, and make data-driven HR decisions.

QUESTIONS

1 Find all employees whose TotalPay is greater than 5,00,000.
2 Get Employee Name, Job Title, and Total Pay
3 Get Employees Who Earn More Than ₹1,80,000 as TotalPay
4 Find Employees Who Earn More Than the Average Total Pay for Their Department
5 Retrieve Employee Names with Their Job Titles
6 Find the Employee Who Earns the Highest Salary in Each Department
7 Find the Second Highest Salary
8 Find the Employees Earning More Than the Average Salary
9 Find Employees Who Have Never Received a Bonus (OtherPay = 0)
10 Total Salary Paid for Each Job Title
11 Which employees have OvertimePay greater than the average OvertimePay?
12 Count number of Employees in Each department(JobTitle)
13 Find the Highest Salary, Lowest Salary, and Average Salary in Each Department
14 Find Job Roles with More Than 100 Employees
15 Find Total Salary(TotalPayBenefits) Paid Per Job Title
16 Find Average, Minimum, and Maximum Base Pay
17 Find the Department with the Highest Average Salary

```sql
-- Find all employees whose TotalPay is greater than 5,00,000.

select * from salary where TotalPay > "500000"
```

Employee_ID	Job_ID	BasePay	OvertimePay	OtherPay	Benefits	TotalPay	TotalPayBenefits	Year
1	1	167411.18	0.00	400184.25	Not Provided	567595.43	567595.43	2011
2	2	155966.02	245131.88	137811.38	Not Provided	538909.28	538909.28	2011

```sql
-- Get Employee Name, Job Title, and Total Pay
select t2.EmployeeName,t3.JobTitle,t1.TotalPay
from salary t1
inner join employees t2
on t1.Employee_ID=t2.Employee_ID
inner join jobs t3
on t3.Job_ID=t1.Job_ID
limit 10
```

EmployeeName	JobTitle	TotalPay
WALLACE WONG	ADMINISTRATIVE ENGINEER	149322.23
SEVERINO CARANTO-JR	ADMINISTRATIVE ENGINEER	136679.61
TIMOTHY LEUNG	ADMINISTRATIVE ENGINEER	134976.96
AMY BROWNELL	ADMINISTRATIVE ENGINEER	130171.24
SUSAN LEONG	ADMINISTRATIVE ENGINEER	130171.08
CHARLES YU	ADMINISTRATIVE ENGINEER	130171.05
RAUL MOSUELA	ADMINISTRATIVE ENGINEER	130171.02
LORA BOWMAN	ADMINISTRATIVE ENGINEER	128923.30
JOHN ROSS	ADMINISTRATIVE ENGINEER	117194.48

```sql
-- Get Employees Who Earn More Than ₹1,80,000 as TotalPay
select t1.EmployeeName,t2.TotalPay
from salary t2
inner join employees t1
on t1.Employee_ID=t2.Employee_ID
where t2.TotalPay > 180000
```

EmployeeName	TotalPay
NATHANIEL FORD	567595.43
GARY JIMENEZ	538909.28
ALBERT PARDINI	335279.91
CHRISTOPHER CHONG	332343.61
PATRICK GARDNER	326373.19
DAVID SULLIVAN	316285.74
ALSON LEE	315981.05
DAVID KUSHNER	307899.46
MICHAEL MORRIS	303427.55
JOANNE HAYES-WHITE	302377.73
ARTHUR KENNEY	299494.17
PATRICIA JACKSON	297608.92
EDWARD HARRINGTON	294580.02
JOHN MARTIN	292671.62
DAVID FRANKLIN	286347.05
RICHARD CORRIEA	286213.86

```sql
-- Find Employees Who Earn More Than the Average Total Pay for Their Department
select t1.EmployeeName,t2.TotalPay,t3.JobTitle
from salary t2
inner join employees t1
on t1.Employee_Id= t2.Employee_ID
inner join jobs t3
on t3.Job_Id=t2.Job_ID
where TotalPay >(select avg(TotalPay) from salary)
```

EmployeeName	TotalPay	JobTitle
NATHANIEL FORD	567595.43	GENERAL MANAGER-METROPOLITAN TRANSIT ...
GARY JIMENEZ	538909.28	CAPTAIN III (POLICE DEPARTMENT)
ALBERT PARDINI	335279.91	CAPTAIN III (POLICE DEPARTMENT)
CHRISTOPHER CHONG	332343.61	WIRE ROPE CABLE MAINTENANCE MECHANIC
PATRICK GARDNER	326373.19	DEPUTY CHIEF OF DEPARTMENT,(FIRE DEPART...
DAVID SULLIVAN	316285.74	ASSISTANT DEPUTY CHIEF II
ALSON LEE	315981.05	BATTALION CHIEF, (FIRE DEPARTMENT)
DAVID KUSHNER	307899.46	DEPUTY DIRECTOR OF INVESTMENTS
MICHAEL MORRIS	303427.55	BATTALION CHIEF, (FIRE DEPARTMENT)
JOANNE HAYES-WHITE	302377.73	CHIEF OF DEPARTMENT, (FIRE DEPARTMENT)

```sql
-- Retrieve Employee Names with Their Job Titles
select t1.EmployeeName,t2.JobTitle
from salary t3
inner join jobs t2
on t2.Job_ID=t3.Job_ID
inner join employees t1
on t1.Employee_ID= t3.Employee_ID
```

EmployeeName	JobTitle
WALLACE WONG	ADMINISTRATIVE ENGINEER
SEVERINO CARANTO-JR	ADMINISTRATIVE ENGINEER
TIMOTHY LEUNG	ADMINISTRATIVE ENGINEER
AMY BROWNELL	ADMINISTRATIVE ENGINEER
SUSAN LEONG	ADMINISTRATIVE ENGINEER
CHARLES YU	ADMINISTRATIVE ENGINEER
RAUL MOSUELA	ADMINISTRATIVE ENGINEER
LORA BOWMAN	ADMINISTRATIVE ENGINEER
JOHN ROSS	ADMINISTRATIVE ENGINEER
SUSAN CURRIN	ADMINISTRATOR, SFGH MEDICAL CENTER

```sql
-- Find the Employee Who Earns the Highest Salary in Each Department
select EmployeeName,JobTitle,TotalPay
from( select t1.EmployeeName,t2.JobTitle,t3.TotalPay,
rank() over ( partition by t2.Job_ID order by t3.TotalPay desc) As Ranks
from employees t1
inner join salary t3
on t1.Employee_ID=t3.Employee_ID
inner join jobs t2
on t2.Job_ID=t3.Job_ID) ranked
where Ranks= 1
```

EmployeeName	JobTitle	TotalPay
NATHANIEL FORD	GENERAL MANAGER-METROPOLITAN TRANSIT ...	567595.43
GARY JIMENEZ	CAPTAIN III (POLICE DEPARTMENT)	538909.28
CHRISTOPHER CHONG	WIRE ROPE CABLE MAINTENANCE MECHANIC	332343.61
PATRICK GARDNER	DEPUTY CHIEF OF DEPARTMENT,(FIRE DEPART...	326373.19
DAVID SULLIVAN	ASSISTANT DEPUTY CHIEF II	316285.74
ALSON LEE	BATTALION CHIEF, (FIRE DEPARTMENT)	315981.05
DAVID KUSHNER	DEPUTY DIRECTOR OF INVESTMENTS	307899.46
JOANNE HAYES-WHITE	CHIEF OF DEPARTMENT, (FIRE DEPARTMENT)	302377.73
ARTHUR KENNEY	ASSISTANT CHIEF OF DEPARTMENT, (FIRE DEP...	299494.17
EDWARD HARRINGTON	EXECUTIVE CONTRACT EMPLOYEE	294580.02

```sql
-- Find the Second Highest Salary
select max(TotalPay) As Second_Highest_Salary
from salary
where TotalPay<(select max(TotalPay) from salary)
```

Second_Highest_Salary
538909.28

```sql
-- Find the Employees Earning More Than the Average Salary
select EmployeeName,TotalPay
from ( select t1.EmployeeName, t2.TotalPay, avg(t2.TotalPay) over () As Average_Salary
from Employees t1
inner join Salary t2
on t1.Employee_ID=t2.Employee_ID) sub
where TotalPay > Average_Salary
```

EmployeeName	TotalPay
NATHANIEL FORD	567595.43
GARY JIMENEZ	538909.28
ALBERT PARDINI	335279.91
CHRISTOPHER CHONG	332343.61
PATRICK GARDNER	326373.19
DAVID SULLIVAN	316285.74
ALSON LEE	315981.05
DAVID KUSHNER	307899.46
MICHAEL MORRIS	303427.55
JOANNE HAYES-WHITE	302377.73

```sql
-- Find Employees Who Have Never Received a Bonus (OtherPay = 0)
select t1.EmployeeName,t2.OtherPay
from salary t2
inner join employees t1
on t1.Employee_ID=t2.Employee_ID
where t2.OtherPay=0
```

EmployeeName	OtherPay
EDWARD HARRINGTON	0.00
EDWIN LEE	0.00
TRENT RHORER	0.00
BARBARA GARCIA	0.00
MONIQUE MOYER	0.00
EDWARD REISKIN	0.00
SONALI BOSE	0.00
JACKSON WONG	0.00
JOHN RAHAIM	0.00
MICHAEL CARLIN	0.00
DEBRA JOHNSON	0.00
DENNIS HERRERA	0.00
SUSAN BUCHBINDER	0.00
JOHN BROWN	0.00

```sql
-- Total Salary Paid for Each Job Title
select t1.JobTitle,sum(t2.TotalPay)
from jobs t1
inner join salary t2
on t1.Job_ID=t2.Job_ID
group by t1.JobTitle
```

JobTitle	sum(t2.TotalPay)
ADMINISTRATIVE ENGINEER	1187780.97
ADMINISTRATOR, SFGH MEDICAL CENTER	257124.44
AIRPORT COMMUNICATIONS OPERATOR	310647.20
AIRPORT ECONOMIC PLANNER	1407442.68
AIRPORT ELECTRICIAN	1848033.08
AIRPORT ELECTRICIAN SUPERVISOR	234177.52
AIRPORT MECHANICAL MAINTENANCE SUPERV...	128534.17
AIRPORT OPERATIONS SUPERVISOR	422794.20
AIRPORT SAFETY OFFICER	97733.68
ANESTHETIST	2404645.02
APPRENTICE STATIONARY ENGINEER	122543.20
ARCHITECT	1454219.97
ARCHITECTURAL ADMINISTRATOR	411161.64

```sql
-- Which employees have OvertimePay greater than the average OvertimePay?
select t1.EmployeeName,t2.JobTitle,t3.OvertimePay
from salary t3
inner join jobs t2
on t2.Job_Id=t3.Job_ID
inner join employees t1
on t1.Employee_ID=t3.Employee_ID
where t3.OvertimePay>(select avg(OvertimePay)from salary)
```

EmployeeName	JobTitle	OvertimePay
GARY JIMENEZ	CAPTAIN III (POLICE DEPARTMENT)	245131.88
ALBERT PARDINI	CAPTAIN III (POLICE DEPARTMENT)	106088.18
CHRISTOPHER CHONG	WIRE ROPE CABLE MAINTENANCE MECHANIC	56120.71
ALSON LEE	BATTALION CHIEF, (FIRE DEPARTMENT)	89062.90
MICHAEL MORRIS	BATTALION CHIEF, (FIRE DEPARTMENT)	86362.68
ARTHUR KENNEY	ASSISTANT CHIEF OF DEPARTMENT, (FIRE DEP...	71344.88
PATRICIA JACKSON	CAPTAIN III (POLICE DEPARTMENT)	87082.62
DAVID FRANKLIN	BATTALION CHIEF, (FIRE DEPARTMENT)	74050.30
RICHARD CORRIEA	COMMANDER III, (POLICE DEPARTMENT)	73478.20
SEBASTIAN WONG	CAPTAIN, EMERGENCYCY MEDICAL SERVICES	119397.26
MARTY ROSS	BATTALION CHIEF, (FIRE DEPARTMENT)	69626.12
GEORGE GARCIA	CAPTAIN, FIRE SUPPRESSION	93200.58
VICTOR WYRSCH	BATTALION CHIEF, (FIRE DEPARTMENT)	77896.14

```sql
-- Count number of Employees in Each department( JobTitle)
select t1.JobTitle,count(t2.Employee_ID) As Employee_Count
from employees t2
inner join salary t3
on t2.Employee_ID=t3.Employee_ID
inner join jobs t1
on t1.Job_ID=t3.Job_ID
group by t1.JobTitle
order by Employee_Count desc
```

JobTitle	Employee_Count
REGISTERED NURSE	905
POLICE OFFICER III	727
FIREFIGHTER	718
POLICE OFFICER I	430
POLICE OFFICER II	343
ATTORNEY (CIVIL/CRIMINAL)	298
SERGEANT III (POLICE DEPARTMENT)	293
DEPUTY SHERIFF	287
FIRE FIGHTER PARAMEDIC	235
TRANSIT OPERATOR	214
LIEUTENANT, FIRE DEPARTMENT	170
INSPECTOR III, (POLICE DEPARTMENT)	161
TRANSIT SUPERVISOR	139

```sql
-- Find the Highest Salary, Lowest Salary, and Average Salary in Each Department
select distinct t1.JobTitle As Department,
first_value(t2.TotalPay) over ( partition by t1.JobTitle order by t2.TotalPay desc) As Highest_Salary,
first_value(t2.TotalPay) over ( partition by t1.JobTitle order by t2.TotalPay asc) As Lowest_Salary,
avg(t2.TotalPay) over ( partition by t1.JobTitle) As Average_Salary
from Salary t2
inner join jobs t1
on t1.Job_ID=t2.Job_ID
```

Department	Highest_Salary	Lowest_Salary	Average_Salary
ADMINISTRATIVE ENGINEER	149322.23	117194.48	131975.663333
ADMINISTRATOR, SFGH MEDICAL CENTER	257124.44	257124.44	257124.440000
AIRPORT COMMUNICATIONS OPERATOR	106776.25	101361.47	103549.066667
AIRPORT ECONOMIC PLANNER	111811.72	101590.74	108264.821538
AIRPORT ELECTRICIAN	118617.32	98061.21	108707.828235
AIRPORT ELECTRICIAN SUPERVISOR	118585.40	115592.12	117088.760000
AIRPORT MECHANICAL MAINTENANCE SUPERV...	128534.17	128534.17	128534.170000

```sql
-- Find Job Roles with More Than 100 Employees
select t1.JobTitle,count(t2.Employee_ID) As Employee_Count
from employees t2
inner join salary t3
on t2.Employee_ID=t3.Employee_ID
inner join jobs t1
on t1.Job_ID=t3.Job_ID
group by t1.JobTitle
Having count(t2.Employee_ID) > 100
order by Employee_Count desc
```

JobTitle	Employee_Count
REGISTERED NURSE	905
POLICE OFFICER III	727
FIREFIGHTER	718
POLICE OFFICER I	430
POLICE OFFICER II	343
ATTORNEY (CIVIL/CRIMINAL)	298
SERGEANT III (POLICE DEPARTMENT)	293

```sql
-- Find Total Salary(TotalPayBenefits) Paid Per Job Title
select t1.JobTitle,sum(TotalPayBenefits) As Total_Salary
from salary t2
inner join jobs t1
on t1.Job_ID=t2.Job_ID
group by t1.JobTitle
order by Total_Salary desc
```

JobTitle	Total_Salary
REGISTERED NURSE	110570137.13
FIREFIGHTER	106198022.61
POLICE OFFICER III	97108031.65
POLICE OFFICER I	50661863.47
SERGEANT III (POLICE DEPARTMENT)	44557827.75
ATTORNEY (CIVIL/CRIMINAL)	43484263.64
POLICE OFFICER II	43403241.16
FIRE FIGHTER PARAMEDIC	33990711.54
DEPUTY SHERIFF	31994300.10
LIEUTENANT, FIRE DEPARTMENT	27905707.74
INSPECTOR III, (POLICE DEPARTMENT)	24896913.98
TRANSIT OPERATOR	23062023.82
TRANSIT SUPERVISOR	18834418.84

```sql
-- Find the Department with the Highest Average Salary
select JobTitle,Average_Salary
from ( select t1.JobTitle,avg(t2.TotalPay) As Average_Salary,
rank() over( order by avg(t2.TotalPay) desc) As Ranks
from salary t2
inner join jobs t1
on t1.Job_ID=t2.Job_ID
group by t1.JobTitle) ranked
where Ranks =1
```

JobTitle	Average_Salary
GENERAL MANAGER-METROPOLITAN TRANSIT ...	399211.275000

```sql
-- Find Average, Minimum, and Maximum Base Pay

select  round(max(BasePay),0) As Maximum_BasePay, round(min(BasePay),0) As Minimum_BasePay, round(avg(BasePay),0) As Average_BasePay
from salary
```

Maximum_BasePay	Minimum_BasePay	Average_BasePay
294580	0	112177