Personal Project- SQL

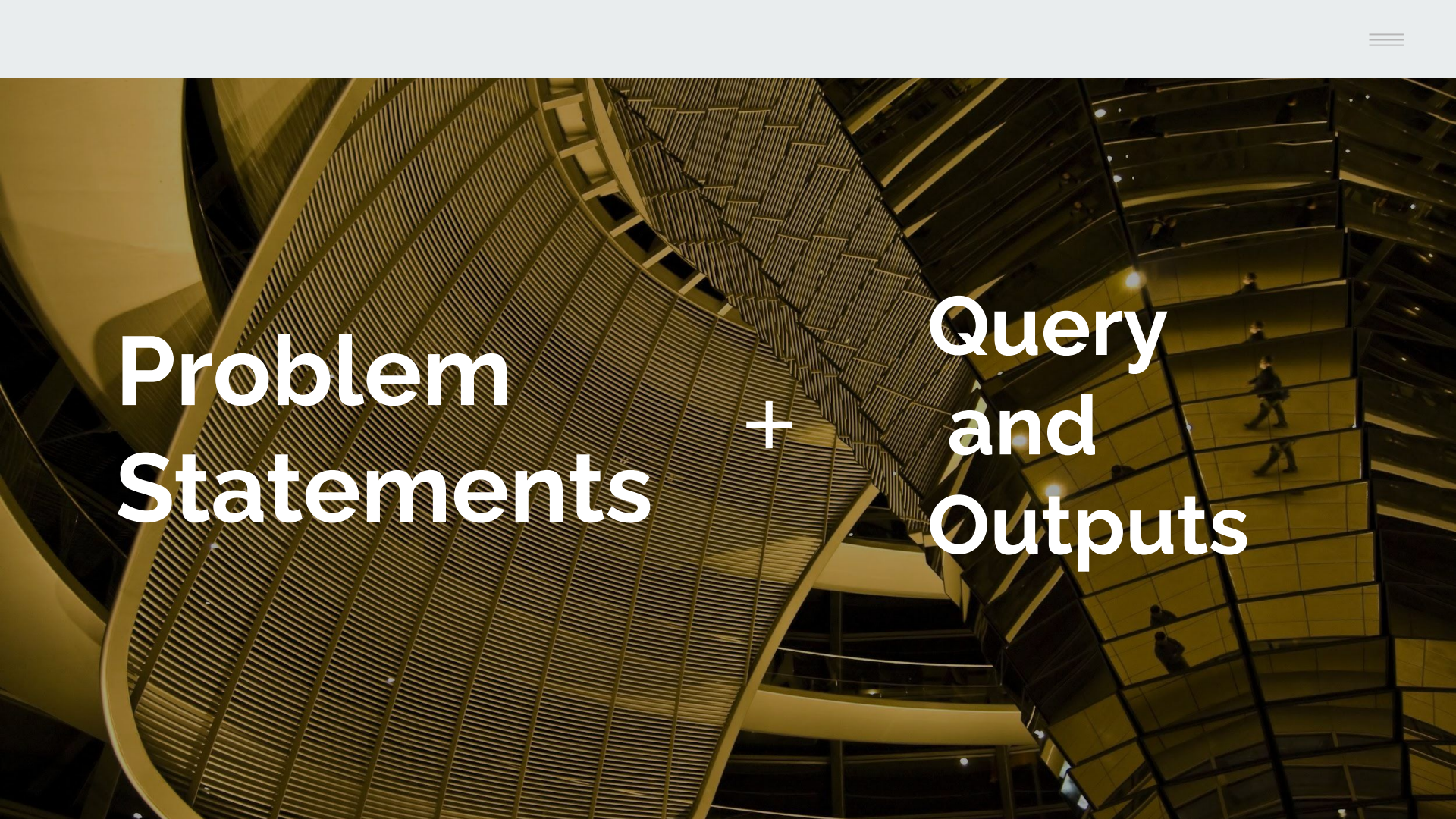Created by- Anshul Ajay Jethani

Overview

In this project, I analyzed engineers dataset (https://github.com/datacharmer/test_db) and developed 10 SQL queries to extract meaningful insights. The queries addressed various aspects of the data, such as filtering engineers based on specific attributes, aggregating salary statistics, and identifying patterns in work experience by years and many more. Through this process, I gained a deeper understanding of SQL query structure, data manipulation, and optimization techniques.

Problem Statements + Query and Outputs

1. Show the list of employees by department

Question: Write a query to list all employees along with their respective department names. Include employee number, first name, last name, department number, and department name.

Query :

```
select e.emp_no, e.first_name , e.last_name, d.dept_no, d.dept_name from employees as e
join dept_emp as de
on e.emp_no = de.emp_no
join departments as d
on de.dept_no = d.dept_no
```

Output :

emp_no	first_name	last_name	dept_no	dept_name
10011	Mary	Sluis	d009	Customer Service
10038	Huan	Lortz	d009	Customer Service
10049	Basil	Tramer	d009	Customer Service
10060	Breannda	Billingsley	d009	Customer Service
10088	Jungsoon	Syrzycki	d009	Customer Service
10098	Sreekrishna	Servieres	d009	Customer Service
10112	Yuichiro	Swick	d009	Customer Service
10115	Chikara	Rissland	d009	Customer Service
10126	Kayoko	Valtorta	d009	Customer Service
10128	Babette	Lamba	d009	Customer Service
10137	Maren	Hutton	d009	Customer Service
10154	Abdulah	Thibadeau	d009	Customer Service
10164	Jagoda	Braunmuhl	d009	Customer Service

2. Current and Past Salaries of an Employee

Question: Write a query to retrieve all the salary records of a given employee (by employee number). Include employee number, salary, from_date, and to_date.

Query :

```
select max(salary) as salary , emp_no , from_date, to_date from salaries
where emp_no = '10044'
GROUP BY emp_no, from_date, to_date
order by salary desc
LIMIT 1
select * from salaries
```

Output :

salary	emp_no	from_date	to_date
58345	10044	2002-05-19	9999-01-01

3. Employees with Specific Titles

Question: Write a query to find all employees who have held a specific title (e.g., 'Engineer'). Include employee number, first name, last name, and title.

Query :

```sql
SELECT e.emp_no,first_name,last_name,title FROM employees as e
JOIN titles  as t
ON e.emp_no = t.emp_no
where t.title = 'Engineer';
```

Output :

emp_no	first_name	last_name	title
10004	Chirstian	Koblick	Engineer
10009	Sumant	Peac	Engineer
10010	Duangkaew	Piveteau	Engineer
10012	Patricio	Bridgland	Engineer
10014	Berni	Genin	Engineer
10018	Kazuhide	Peha	Engineer
10020	Mayuko	Warwick	Engineer
10022	Shahaf	Famili	Engineer
10023	Bojan	Montemayor	Engineer
10026	Yongqiao	Berztiss	Engineer
10027	Divier	Reistad	Engineer
10028	Domenick	Tempesti	Engineer

4. Departments with Their Managers

Question: Write a query to list all departments along with their current managers. Include department number, department name, manager's employee number, first name, and last name.

Query :

```sql
SELECT d.dept_no,d.dept_name,e.emp_no,e.first_name,e.last_name FROM departments as d
JOIN dept_manager as dm
ON d.dept_no = dm.dept_no
JOIN employees as e
ON e.emp_no = dm.emp_no;
```

Output :

dept_no	dept_name	emp_no	first_name	last_name	
d009	Customer Service	111692	Tonny	Butterworth	
d009	Customer Service	111784	Marjo	Giarratana	
d009	Customer Service	111877	Xiaobin	Spinelli	
d009	Customer Service	111939	Yuchang	Weedman	
d005	Development	110511	DeForest	Hagimont	
d005	Development	110567	Leon	DasSarma	
d002	Finance	110085	Ebru	Alpin	
d002	Finance	110114	Isamu	Legleitner	
d003	Human Resources	110183	Shirish	Ossenbruggen	
d003	Human Resources	110228	Karsten	Sigstam	
d001	Marketing	110022	Margareta	Markovitch	
d001	Marketing	110039	Vishwani	Minakawa	

5. Employee Count by Department

Question: Write a queryto count the number of each departments, include department number, department name and employee count.

Query :

```
SELECT d.dept_no,dept_name,count(de.emp_no) as emp_count FROM departments as d
JOIN dept_emp as de
ON d.dept_no = de.dept_no
GROUP BY dept_no,dept_name;
```

Output :

dept_no	dept_name	emp_count
d009	Customer Service	23580
d005	Development	85707
d002	Finance	17346
d003	Human Resources	17786
d001	Marketing	20211
d004	Production	73485
d006	Quality Management	20117
d008	Research	21126
d007	Sales	52245

6. Employees' Birthdates in a Given Year

Question: Write a query to find all employees born in a specific year (e.g., 1953). Include employee number, first name, last name, and birth date.

Query :

```
SELECT emp_no,first_name,last_name,birth_date FROM employees
WHERE year(birth_date) = 1953;
```

Output :

emp_no	first_name	last_name	birth_date
10001	Georgi	Facello	1953-09-02
10006	Anneke	Preusig	1953-04-20
10011	Mary	Sluis	1953-11-07
10019	Lillian	Haddadi	1953-01-23
10023	Bojan	Montemayor	1953-09-29
10026	Yongqiao	Berztiss	1953-04-03
10035	Alain	Chappelet	1953-02-08
10051	Hidefumi	Caine	1953-07-28
10059	Alejandro	McAlpine	1953-09-19
10067	Claudi	Stavenow	1953-01-07
10100	Hironobu	Haraldson	1953-04-21
10103	Akemi	Birch	1953-11-26

7. Employees Hired in the Last 5 Years

Question: Write a query to find all employees hired in the last 50 years. Include employee number, first name, last name, and hire date.

Query :

```sql
SELECT emp_no,first_name,last_name,hire_date FROM employees
WHERE hire_date >= date_sub(curdate(),INTERVAL 50 YEAR);
```

Output :

emp_no	first_name	last_name	hire_date
10001	Georgi	Facello	1986-06-26
10002	Bezalel	Simmel	1985-11-21
10003	Parto	Bamford	1986-08-28
10004	Chirstian	Koblick	1986-12-01
10005	Kyoichi	Maliniak	1989-09-12
10006	Anneke	Preusig	1989-06-02
10007	Tzvetan	Zielinski	1989-02-10
10008	Saniya	Kalloufi	1994-09-15
10009	Sumant	Peac	1985-02-18
10010	Duangkaew	Piveteau	1989-08-24
10011	Mary	Sluis	1990-01-22
10012	Patricio	Bridgland	1992-12-18

8. Average Salary by Department

Question: Write a query to calculate the average salary for each department. Include department number, department name, and average salary.

Query :

```sql
SELECT d.dept_no,d.dept_name,avg(s.salary)  as avg_salary FROM departments as d
JOIN dept_emp as de
ON de.dept_no = d.dept_no
JOIN salaries as s ON de.emp_no = s.emp_no
GROUP BY d.dept_no,d.dept_name;
```

Output :

dept_no	dept_name	avg_salary
d001	Marketing	71913.2000
d002	Finance	70489.3649
d003	Human Resources	55574.8794
d004	Production	59605.4825
d005	Development	59478.9012
d006	Quality Management	57251.2719
d007	Sales	80667.6058
d008	Research	59665.1817
d009	Customer Service	58770.3665

9.Gender Distribution in Each Department

Question: Write a query to find the gender distribution (number of males and females) in each department. Include department number, department name, count of males, and count of females.

Query :

```sql
SELECT d.dept_no,d.dept_name,
SUM(CASE when e.gender = 'M' THEN 1 ELSE 0 END ) as male_count,
 SUM(CASE when e.gender = 'F' THEN 1 ELSE 0 END ) as female_count
 FROM  departments as d
JOIN dept_emp as de ON de.dept_no = d.dept_no
JOIN employees as e on de.emp_no = e.emp_no
GROUP by d.dept_no,d.dept_name;
```

Output :

dept_no	dept_name	male_count	female_count
d009	Customer Service	14132	9448
d005	Development	51449	34258
d002	Finance	10331	7015
d003	Human Resources	10711	7075
d001	Marketing	12174	8037
d004	Production	43936	29549
d006	Quality Management	12039	8078
d008	Research	12687	8439
d007	Sales	31391	20854

10. Longest Serving Employees

Question: Write a query to find the employees who have served the longest in the company. Include employee number, first name, last name, and number of years served.

Query :

```sql
SELECT emp_no,first_name,last_name,
timestampdiff(YEAR,hire_date,CUrdate()) as year_served
 FROM employees
 ORDER By year_served desc
 LIMIT 10
```

Output :

emp_no	first_name	last_name	year_served
10126	Kayoko	Valtorta	39
10198	Pranav	Furedi	39
10137	Maren	Hutton	39
10098	Sreekrishna	Servieres	39
10195	Annemarie	Redmiles	39
10116	Dayanand	Czap	39
10076	Erez	Ritzmann	39
10156	Sumali	Fargier	39
10009	Sumant	Peac	39
10048	Florian	Syrotiuk	39

Thank you.