



PSYLIQ

INTERNSHIP PROJECT

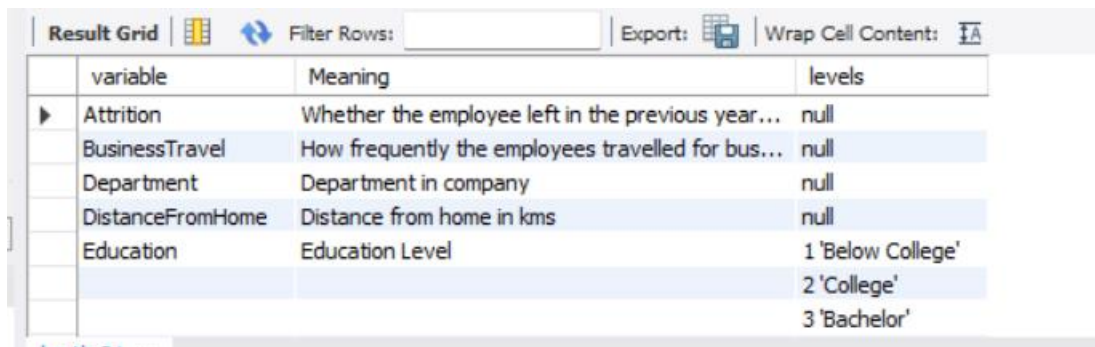
HR DATA ANALYSIS USING SQL

BY

SHARUMATHI C

LOADING THE DATA

```
LOAD DATA INFILE 'C:\\ProgramData\\MySQL\\MySQL Server 8.0\\Uploads\\General Data.csv'  
INTO TABLE hr_tb1  
FIELDS TERMINATED BY ','  
ENCLOSED BY ''  
LINES TERMINATED BY '\\n'  
IGNORE 1 ROWS;
```

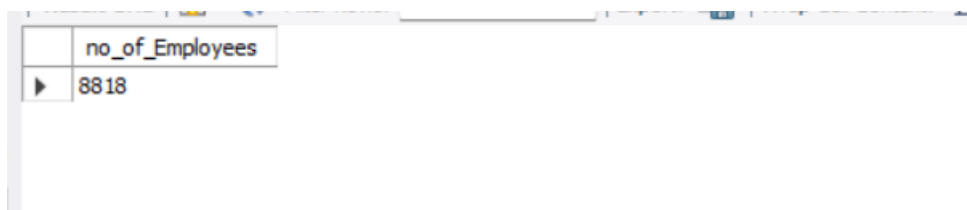


The screenshot shows a 'Result Grid' window with a table structure. The table has four columns: 'variable', 'Meaning', and 'levels'. The data is as follows:

variable	Meaning	levels
Attrition	Whether the employee left in the previous year...	null
BusinessTravel	How frequently the employees travelled for bus...	null
Department	Department in company	null
DistanceFromHome	Distance from home in kms	null
Education	Education Level	1 'Below College' 2 'College' 3 'Bachelor'

1. Retrieve the total number of employees in the dataset

```
SELECT COUNT(Employee_id) As no_of_Employees FROM hr_tb1;
```

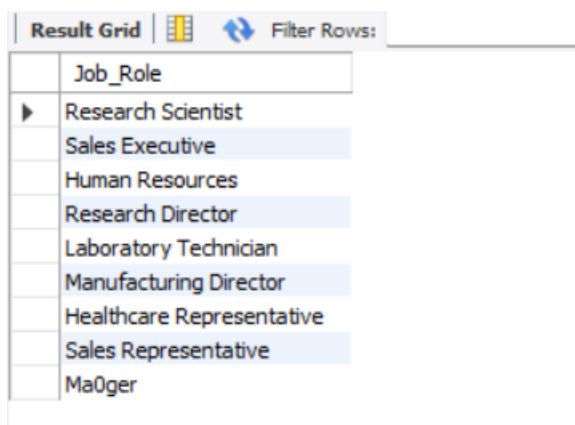


The screenshot shows a 'Result Grid' window with a single column 'no_of_Employees' and one row with the value '8818'.

no_of_Employees
8818

2. List all unique job roles in the dataset.

```
SELECT DISTINCT(Job_Role) FROM hr_tb1;
```

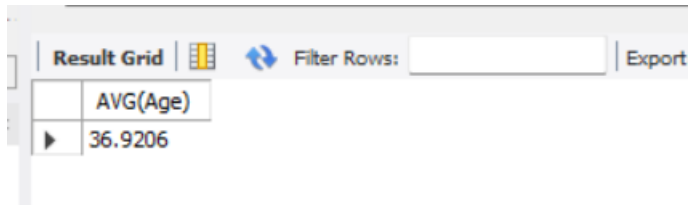


The screenshot shows a 'Result Grid' window with a single column 'Job_Role' and ten rows of unique job roles.

Job_Role
Research Scientist
Sales Executive
Human Resources
Research Director
Laboratory Technician
Manufacturing Director
Healthcare Representative
Sales Representative
Ma0ger

3. Find the average age of employees.

```
SELECT AVG(Age) FROM hr_tb1;
```

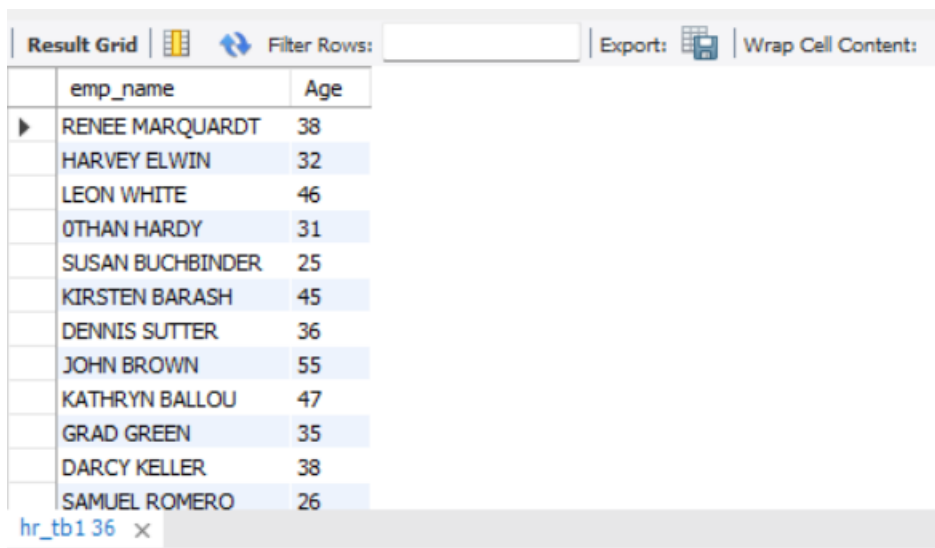


The screenshot shows a database interface with a 'Result Grid' tab. The grid contains two columns: 'AVG(Age)' and a value '36.9206'. There are also buttons for 'Filter Rows' and 'Export'.

AVG(Age)
36.9206

4. Retrieve the names and ages of employees who have worked at the company for more than 5 years.

```
SELECT emp_name, Age  
FROM hr_tb1  
WHERE Years_At_Company > 5;
```

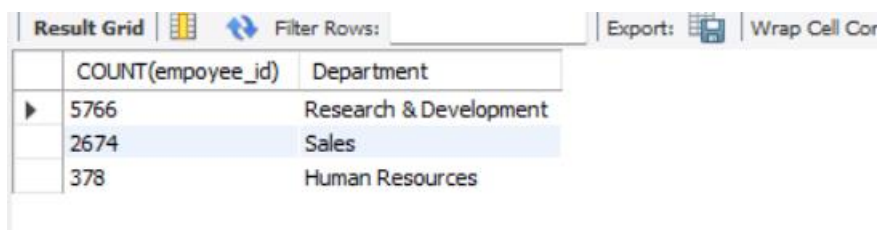


The screenshot shows a database interface with a 'Result Grid' tab. The grid contains two columns: 'emp_name' and 'Age'. The data is as follows:

emp_name	Age
RENEE MARQUARDT	38
HARVEY ELWIN	32
LEON WHITE	46
OTHAN HARDY	31
SUSAN BUCHBINDER	25
KIRSTEN BARASH	45
DENNIS SUTTER	36
JOHN BROWN	55
KATHRYN BALLOU	47
GRAD GREEN	35
DARCY KELLER	38
SAMUEL ROMERO	26

5. Get a count of employees grouped by their department

```
SELECT COUNT(employee_id), Department FROM hr_tb1 GROUP BY  
Department;
```

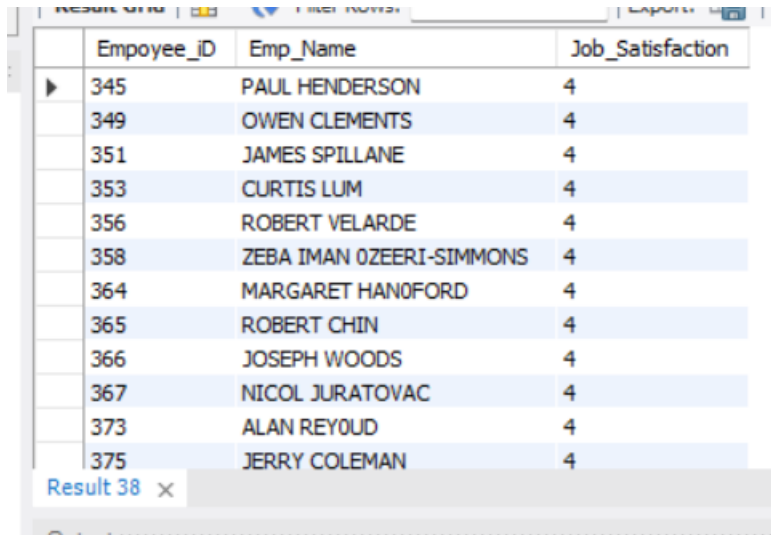


The screenshot shows a database interface with a 'Result Grid' tab. The grid contains two columns: 'COUNT(employee_id)' and 'Department'. The data is as follows:

COUNT(employee_id)	Department
5766	Research & Development
2674	Sales
378	Human Resources

6. List employees who have 'High' Job Satisfaction.

```
SELECT hr_tb1.Employee_ID,Emp_Name,hr_tb2.Job_Satisfaction  
FROM hr_tb1 JOIN hr_tb2 ON hr_tb1.Employee_ID =  
hr_tb2.Employee_ID ORDER BY Job_Satisfaction DESC;
```



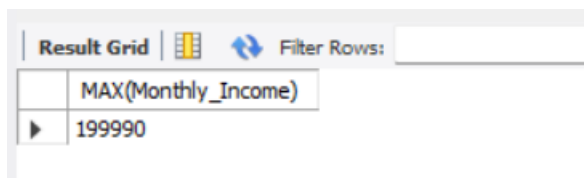
The screenshot shows a database query result grid with the following data:

Employee_ID	Emp_Name	Job_Satisfaction
345	PAUL HENDERSON	4
349	OWEN CLEMENTS	4
351	JAMES SPILLANE	4
353	CURTIS LUM	4
356	ROBERT VELARDE	4
358	ZEBIA IMAN OZEERI-SIMMONS	4
364	MARGARET HANFORD	4
365	ROBERT CHIN	4
366	JOSEPH WOODS	4
367	NICOL JURATOVAC	4
373	ALAN REYDUD	4
375	JERRY COLEMAN	4

Result 38 x

7. Find the highest Monthly Income in the dataset.

```
SELECT MAX(Monthly_Income) FROM hr_tb1;
```



The screenshot shows a database query result grid with the following data:

MAX(Monthly_Income)
199990

8. List employees who have 'Travel_Rarely' as their BusinessTravel type.

```
SELECT Emp_Name FROM hr_tb1 WHERE Business_Travel IN  
('Travel_Rarely');
```

Result Grid		Filter Rows:	Export:
Emp_Name			
▶ HARVEY ELWIN			
LEON WHITE			
DENNIS HERRERA			
DOOLD BRYANT			
OTHAN HARDY			
KIRSTEN BARASH			
DENNIS SUTTER			
JOHN BROWN			
DOOLD FIELDS			
LUIS HERRERA			
GEORGE FOURAS			
GRAD GREEN			

hr_tb1 40 x

9. Retrieve the distinct MaritalStatus categories in the dataset.

```
SELECT DISTINCT(Marital_Status) FROM hr_tb1;
```

Result Grid		Filter Rows:	Export:
Marital_Status			
▶ Single			
Married			
Divorced			

10. Get a list of employees with more than 2 years of work experience but less than 4 years in their current role.

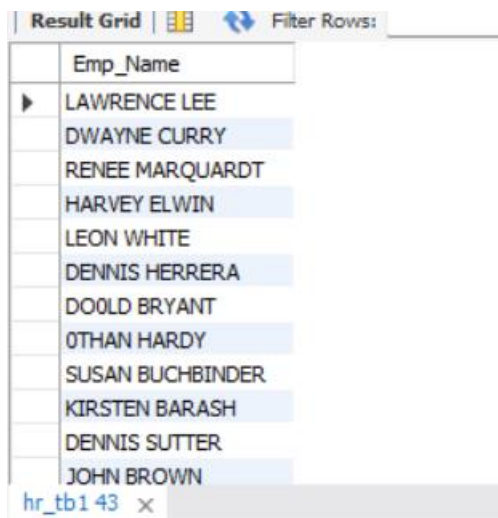
```
SELECT Emp_Name FROM hr_tb1 WHERE Total_Working_Years IN(2,4);
```

Result Grid		Filter Rows:
Emp_Name		
▶ MICHAEL AHERN		
ROBERT SERRANO		
TIMOTHY SULLIVAN		
RYAN KENNEDY		
JOHN GOLDBERG		
ROBERT STYLES		
JAY HUIISH		
PATRICIA O'CONNER		
BONNIE TAYLOR		
JAMES HARRIGAN		
CLYDE CHRISTOBAL		
AMEN CHOW		

hr_tb1 42 x

11. List employees who have changed their job roles within the company (JobLevel and JobRole differ from their previous job).

```
SELECT Emp_Name  
FROM hr_tb1  
WHERE Job_Level != Job_Role  
LIMIT 0, 50000;
```



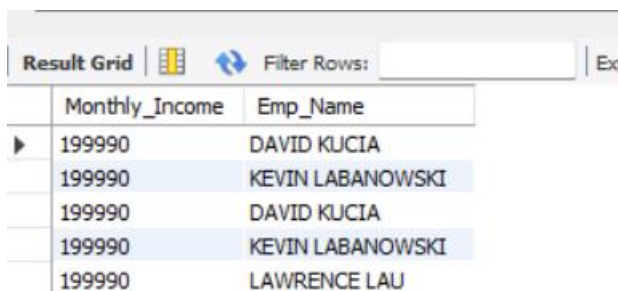
The screenshot shows a 'Result Grid' window with a 'Filter Rows' input field. The grid contains a single column titled 'Emp_Name'. The data is as follows:

Emp_Name
LAWRENCE LEE
DWAYNE CURRY
RENEE MARQUARDT
HARVEY ELWIN
LEON WHITE
DENNIS HERRERA
DOOLD BRYANT
OTHAN HARDY
SUSAN BUCHBINDER
KIRSTEN BARASH
DENNIS SUTTER
JOHN BROWN

At the bottom of the grid, there is a status bar that reads 'hr_tb1 43 x'.

13. Retrieve the top 5 employees with the highest MonthlyIncome

```
SELECT Monthly_Income, Emp_Name  
FROM hr_tb1  
ORDER BY Monthly_Income DESC  
LIMIT 5;
```



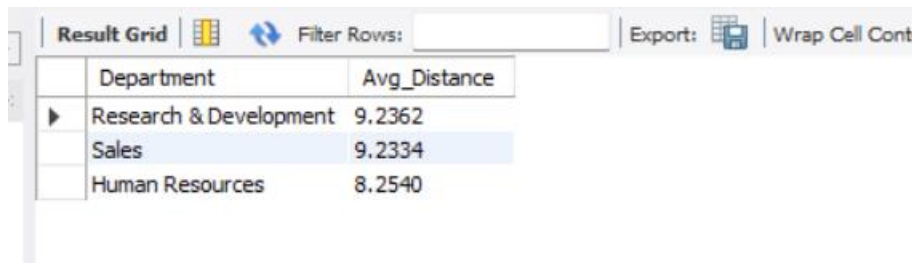
The screenshot shows a 'Result Grid' window with a 'Filter Rows' input field. The grid contains two columns: 'Monthly_Income' and 'Emp_Name'. The data is as follows:

Monthly_Income	Emp_Name
199990	DAVID KUCIA
199990	KEVIN LABANOWSKI
199990	DAVID KUCIA
199990	KEVIN LABANOWSKI
199990	LAWRENCE LAU

At the bottom of the grid, there is a status bar that reads 'Ex'.

12. Find the average distance from home for employees in each department.

```
SELECT Department,AVG(Distance_From_HOME)AS Avg_Distance FROM  
hr_tb1 GROUP BY Department;
```

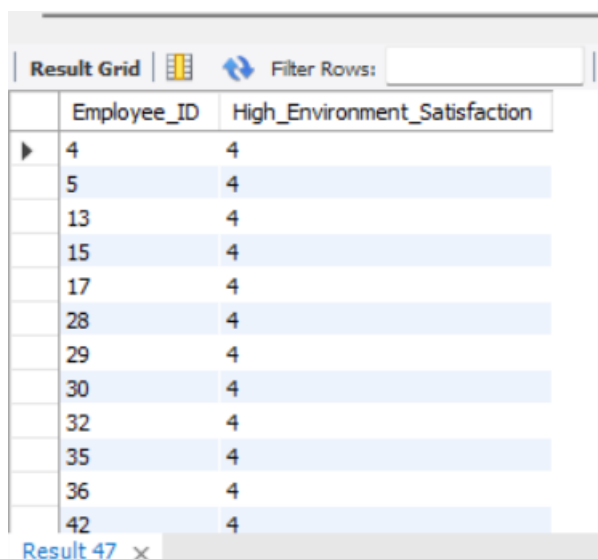


The screenshot shows a database query result grid with two columns: 'Department' and 'Avg_Distance'. The data is as follows:

Department	Avg_Distance
Research & Development	9.2362
Sales	9.2334
Human Resources	8.2540

13. List the employees with the highest and lowest Environment Satisfaction.

```
SELECT Employee_ID,Environment_Satisfaction AS High_Environment_Satisfaction  
FROM hr_tb2 Where Environment_Satisfaction = 4  
  
UNION ALL  
  
SELECT Employee_ID,Environment_Satisfaction AS Low_Environment_Satisfaction  
FROM hr_tb2 Where Environment_Satisfaction = 1 ;
```



The screenshot shows a database query result grid with two columns: 'Employee_ID' and 'High_Environment_Satisfaction'. The data is as follows:

Employee_ID	High_Environment_Satisfaction
4	4
5	4
13	4
15	4
17	4
28	4
29	4
30	4
32	4
35	4
36	4
42	4

14. Calculate the percentage of employees who have had a promotion in the last year

```
SELECT COUNT(*)*100/(SELECT COUNT(*) FROM hr_tb1) AS  
Promotion_percentage FROM hr_tb1 WHERE Years_Since_Last_Promotion=1;
```

Result Grid	Filter Rows:
Promotion_percentage	
24.2912	

15. Find the employees who have the same Job Role and Marital Status

```
SELECT a.employee_id, a.job_role, b.job_role, a.marital_status, b.marital_status from
```

```
hr_tb1 a join hr_tb1 b on a.Employee_ID=b.Employee_ID
```

```
where a.Job_Role=b.job_role and a. Marital_Status=b.Marital_Status;
```

Result Grid

Filter Rows:

Export:

Wrap Cell Content:

	employee_id	job_role	job_role	marital_status	marital_status
▶	2	Research Scientist	Research Scientist	Single	Single
	3	Sales Executive	Sales Executive	Married	Married
	4	Human Resources	Human Resources	Married	Married
	5	Sales Executive	Sales Executive	Single	Single
	6	Research Director	Research Director	Married	Married
	7	Sales Executive	Sales Executive	Single	Single
	8	Sales Executive	Sales Executive	Married	Married
	9	Laboratory Technician	Laboratory Technician	Married	Married
	10	Laboratory Technician	Laboratory Technician	Divorced	Divorced
	11	Laboratory Technician	Laboratory Technician	Married	Married
	12	Laboratory Technician	Laboratory Technician	Married	Married
	13	Sales Executive	Sales Executive	Single	Single

Result 48

×

17. List the employees with the highest Total Working Years who also have a Performance Rating of 4.

```
SELECT hr_tb1.Employee_ID,hr_tb1.Total_Working_Years,
```

```
Performance_Rating FROM hr_tb1 JOIN hr_tb2 ON
```

```
hr_tb1.Employee_ID =hr_tb2.Employee_ID WHERE
```

```
hr_tb2.Performance_Rating=4
```

```
ORDER BY hr_tb1.Total_Working_Years DESC ;
```


Result Grid			
	Employee_ID	Total_Working_Years	Performance_Rating
▶	786	35	4
	786	35	4
	786	35	4
	2256	35	4
	2256	35	4
	2256	35	4
	3726	35	4
	3726	35	4
	3726	35	4
	786	35	4
	786	35	4
	786	35	4

18. Calculate the average Age and JobSatisfaction for each BusinessTravel type.

```
select distinct Business_travel,avg(age)over(partition by Business_travel)as avgage from hr_tb1;
```

Result Grid		
	Business_travel	avgage
▶	Non-Travel	36.6200
	Travel_Frequently	36.4549
	Travel_Rarely	37.0876

19. Retrieve the most common Education Field among employees.

```
SELECT Education_Field FROM hr_tb1 GROUP BY Education_Field ORDER BY Education_Field DESC;
```

Education_Field	
▶	Technical Degree
	Other
	Medical
	Marketing
	Life Sciences
	Human Resources

20. List the employees who have worked for the company the longest but haven't had a promotion.

```
SELECT Employee_ID,MAX(Years_At_Company)AS Max_no_of_years FROM hr_tb1 WHERE Years_Since_Last_Promotion =0 GROUP BY Employee_ID
```

ORDER BY Max_no_of_years DESC ;

Result Grid			Filter Rows:
	Employee_ID	Max_no_of_years	
▶	705	33	
	2175	33	
	3645	33	
	216	31	
	1686	31	
	3156	31	
	878	26	
	2348	26	
	2010	26	
Result 52			