

HR Analytics Project

Objective

Analyze HR data to generate actionable insights related to employees, engagement, recruitment, and training programs.

Datasets Used

- employee_data.csv
- employee_engagement_survey_data.csv
- recruitment_data.csv
- training_and_development_data.csv

Tools & Methodology

- **SQL Queries:** All data analysis was performed using **MySQL Workbench**, where queries were executed to extract, filter, and aggregate HR data.
- **Data Visualization:** **Power BI Desktop** was used to create meaningful visualizations for selected results. Visuals include bar charts, column charts, and cards to highlight key insights. Large tables were **trimmed** in the document to focus on actionable insights.
- **Workflow:**
 1. Extract data from CSVs into MySQL tables.
 2. Run SQL queries for each analysis question.
 3. Load MySQL tables into Power BI.
 4. Create charts and visuals for questions where visualization adds clarity.
 5. Export charts as screenshots and paste into the document under the corresponding question.

⌚ Phase 1 – Basic Queries (SELECT, WHERE, Filters)

Q1: List all employees' first and last names

Query:

```
SELECT FirstName, LastName  
FROM employee_data;
```

Result:

FirstName	LastName
Uriah	Bridges
Paula	Small
Edward	Buck
Michael	Riordan
Jasmine	Onque

Q2: Show all employees in the "IT/IS" department

Query:

```
SELECT *  
FROM employee_data  
WHERE DepartmentType = 'IT/IS';
```

Result:

EmpId	FirstName	LastName	StartDate	ExitDate	Title
3442	Kaylah	Moon	09-Jul-19	16-Jun-22	Area Sales Manager
3443	Kristen	Tate	05-Apr-21	12-May-23	Area Sales Manager
3454	Leon	Beard	11-Sep-22		Area Sales Manager
3466	Clayton	Walker	13-Apr-22	10-Apr-23	Area Sales Manager
3477	Sonny	Horne	07-Apr-19		Area Sales Manager

Insight: Identifies IT/IS department members. Useful for staffing analysis.

Q3: Count total employees**Query:**

```
SELECT COUNT(*) AS TotalEmployees  
FROM employee_data;
```

Result:

TotalEmployees
3000

Insight: Helps HR monitor total headcount for reporting and planning.

Q4: List terminated employees**Query:**

```
SELECT *  
FROM employee_data  
WHERE EmployeeStatus LIKE '%terminated%';
```

Result:

EmpID	FirstName	LastName	StartDate	ExitDate	EmployeeStatus
3564	Susan	Lundy	26-Jan-19	02-Dec-20	Voluntarily Terminated
3568	Quinn	Rerrick	11-Jul-23	29-Jul-23	Voluntarily Terminated
3570	Adrienne J	Homberger	19-Apr-19	26-Oct-21	Voluntarily Terminated
3571	Marilyn	Linares	07-Jul-22	17-Jul-23	Voluntarily Terminated
3575	George	Johnson	28-Apr-21	18-Jan-22	Voluntarily Terminated

Insight: Supports turnover and retention analysis.

Q5: Employees born after 1990-01-01**Query:**

```
SELECT *  
FROM employee_data  
WHERE STR_TO_DATE(DOB, '%d-%m-%Y') > '1990-01-01';
```

Result:

EmpID	FirstName	LastName	DOB
3429	Edward	Buck	06-10-1991
3430	Michael	Riordan	04-04-1998
3443	Kristen	Tate	08-04-1994
3446	Hector	Dalton	01-05-1996
3449	Gerald	Preston	18-09-1992

Insight: Identifies younger employees for workforce age analysis.

Q6: List female employees

Query:

```
SELECT *
FROM employee_data
WHERE GenderCode = 'Female';
```

Result:

EmpID	FirstName	LastName	GenderCode
3427	Uriah	Bridges	Female
3431	Jasmine	Onque	Female
3433	Latia	Costa	Female
3434	Sharlene	Terry	Female
3437	Myriam	Givens	Female

Insight: Helps analyze gender diversity in the workforce.

Q7: List applicants for "Software Engineer"

Query:

```
SELECT *
FROM recruitment_data
WHERE `Job Title` = 'Software Engineer';
```

Result:

Applicant ID	First Name	Last Name	Job Title
1530	Justin	Randolph	Software engineer

Insight: Helps track candidates in specific roles for recruitment planning.

Q8: Count applicants by gender

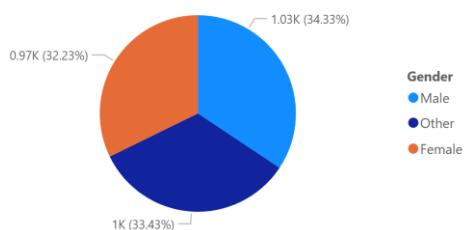
Query:

```
SELECT Gender, COUNT(*) AS TotalApplicants
FROM recruitment_data
GROUP BY Gender;
```

Result:

Gender	TotalApplications
Male	1030
Female	967
Other	1003

Count of Applicant ID by Gender



Insight: Provides gender distribution insights for recruitment diversity.

Q9: Trainings attended by Employee ID 1001

Query:

```
SELECT *
FROM training_and_development_data
WHERE `Employee ID` = 1001;
```

Result:

Employee ID	Training Date	Training Program Name
1001	21-Sep-22	Customer Service

Insight: Tracks individual employee's training history for skill development.

Q10: Employees and their supervisors

Query:

```
SELECT FirstName, LastName, Supervisor
FROM employee_data;
```

Result:

FirstName	LastName	Supervisor
Uriah	Bridges	Peter O'Neill
Paula	Small	Renee McCormick
Edward	Buck	Crystal Walker
Michael	Riordan	Rebekah Wright
Jasmine	Onque	Jason Kim

Insight: Shows reporting relationships within the organization.

Phase 2 – Aggregations & Joins (GROUP BY, AVG, SUM, JOINS)

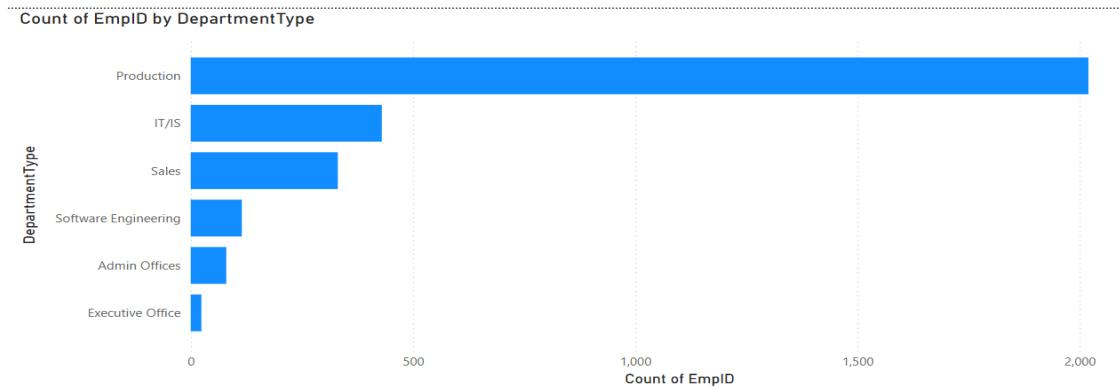
Q11: Employee count per department

Query:

```
SELECT DepartmentType, COUNT(*) AS TotalEmployees
FROM employee_data
GROUP BY DepartmentType;
```

Result:

DepartmentType	TotalEmployees
Production	2020
Sales	331
IT/IS	430
Executive Office	24
Software Engineering	115
Admin Offices	80



Insight: Helps HR assess staffing distribution across departments.

Q12: Average employee engagement score

Query:

```
SELECT AVG(`Engagement Score`) AS AvgEngagement
FROM employee_engagement_survey_data;
```

Result:

AvgEngagement
2.9397

Insight: Measures overall engagement for HR strategy.

Q13: Employees needing performance improvement

Query:

```
SELECT *
FROM employee_data
WHERE `Performance Score` = 'Needs Improvement';
```

Result:

EmpID	FirstName	LastName	Performance Score
3454	Leon	Beard	Needs Improvement
3457	Milton	Wall	Needs Improvement
3458	Cory	Robinson	Needs Improvement
3461	Lincoln	Compton	Needs Improvement
3462	Aliana	Nolan	Needs Improvement

Insight: Supports targeted coaching and development plans.

Q14: Employees with no training history

Query:

```
SELECT e.*
FROM employee_data e
LEFT JOIN training_and_development_data t
    ON e.EmpID = t.`Employee ID`
WHERE t.`Employee ID` IS NULL;
```

Result: no matching rows found

Insight: Highlights employees requiring training for skill development.

Q15: Total training days per employee

Query:

```
SELECT `Employee ID`, SUM(`Training Duration(Days)`)
AS
TotalTrainingDays
FROM training_and_development_data
GROUP BY `Employee ID`;
```

Result:

Employee ID	TotalTrainingDays
1001	4
1002	2
1003	4
1004	2
1005	4

Insight: Tracks training investment per employee.

Q16: Top 5 with the highest Engagement Score

Query:

```
SELECT `Employee ID`, `Engagement Score`
FROM employee_engagement_survey_data
ORDER BY `Engagement Score` DESC
LIMIT 5;
```

Result:

Employee ID	Engagement Score
1013	5
1024	5
1021	5
1008	5
1006	5

Insight: Identifies highly engaged employees for recognition programs.

Q17: Applicants with >5 years of experience

Query:

```
SELECT *
FROM recruitment_data
WHERE `Years of Experience` > 5;
```

Result:

Applicant ID	First Name	Last Name	Years of Experience
1001	Scott	Sheppard	8
1002	Stanley	Lewis	17
1003	Javier	Li	20
1004	Christopher	Johnston	8
1006	Christian	Maddox	18

Insight: Helps prioritize experienced candidates in recruitment.

Q18: Average satisfaction score per department

Query:

```
SELECT e.DepartmentType, AVG(s.`Satisfaction Score`) AS AvgSatisfaction
FROM employee_data e
JOIN employee_engagement_survey_data s
ON e.EmpID = s.`Employee ID`
GROUP BY e.DepartmentType;
```

Result:

DepartmentType	AvgSatisfaction
Software Engineering	3.1130
Admin Offices	2.5125
Production	3.0208
Executive Office	3.0833
IT/IS	3.0093
Sales	3.1329

Insight: Identifies departments needing engagement improvement.

Q19: Total training cost per program

Query:

```
SELECT `Training Program Name`, ROUND(SUM(`Training Cost`),2)
AS TotalCost
FROM training_and_development_data
GROUP BY `Training Program Name`;
```

Result:

Training Program Name	TotalCost
Customer Service	320575.04
Leadership Development	323902.03
Technical Skills	323072.61
Communication Skills	365023.24
Project Management	343313.17



Insight: Supports budgeting and ROI assessment for training programs.

Q20: Employee count per state

Query:

```
SELECT State, COUNT(*) AS TotalEmployees
FROM employee_data
GROUP BY State;
```

Result:

State	TotalEmployees
MA	2651
ND	4
FL	1
CT	40
CA	10

Insight: Helps plan regional HR initiatives.

 **Phase 3 – Advanced Analytics & Complex Queries**

Q21: Employees with above-average engagement**Query:**

```
SELECT e.* , s.`Engagement Score`  
FROM employee_data e  
JOIN employee_engagement_survey_data s  
ON e.EmpID = s.`Employee ID`  
WHERE s.`Engagement Score` > (SELECT AVG(`Engagement Score`)  
FROM employee_engagement_survey_data);
```

Result:

EmpID	FirstName	LastName	Engagement Score
3430	Michael	Riordan	5
3435	Jac	McKinzie	4
3436	Joseph	Martins	5
3437	Myriam	Givens	3
3438	Dheepa	Nguyen	4

Insight: Identifies highly engaged employees for benchmarking.

Q22: Active employees joined after 2018-01-01**Query:**

```
SELECT *  
FROM employee_data  
WHERE STR_TO_DATE(StartDate, '%d-%b-%y') > '2018-01-01'  
AND EmployeeStatus = 'Active';
```

Result:

EmpID	FirstName	LastName	StartDate
3427	Uriah	Bridges	20-Sep-19
3428	Paula	Small	11-Feb-23
3429	Edward	Buck	10-Dec-18
3430	Michael	Riordan	21-Jun-21
3431	Jasmine	Onque	29-Jun-19

Insight: Useful for workforce planning and onboarding analysis.

Q23: Employee count per JobFunctionDescription and gender

Query:

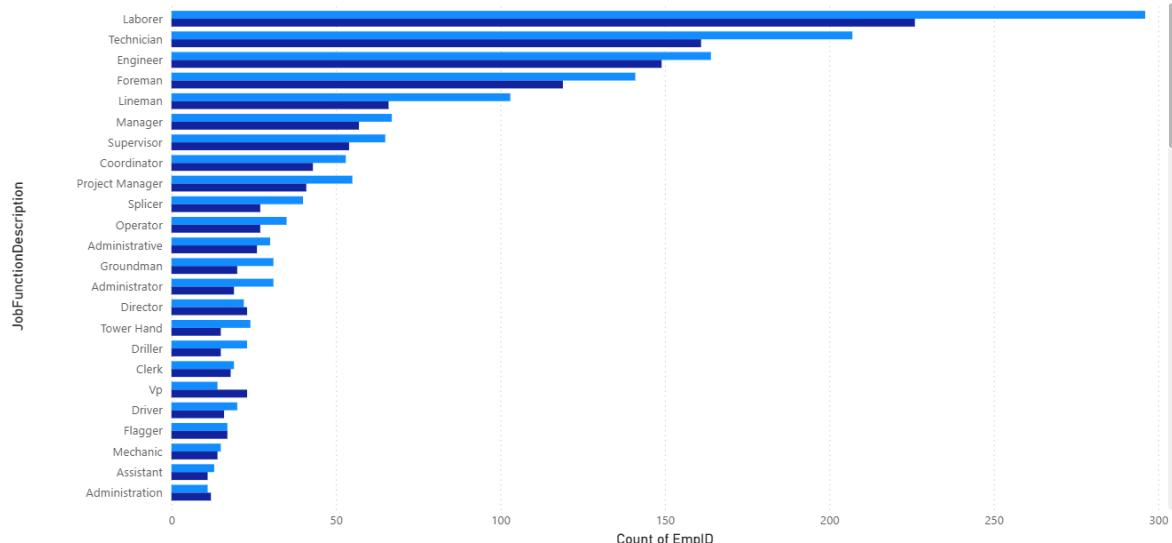
```
SELECT `JobFunctionDescription`, GenderCode, COUNT(*) AS TotalEmployees
FROM employee_data
GROUP BY `JobFunctionDescription`, GenderCode;
```

Result:

JobFunctionDescription	GenderCode	TotalEmployees
Accountant	Female	1
Accountant	Male	1
Accounting	Female	3
Accounting	Male	6
Administration	Female	11
Administration	Male	12
Administrative	Female	30

Count of EmpID by JobFunctionDescription and GenderCode

GenderCode ● Female ● Male



Insight: Analyzes role-wise gender diversity.

Q24: Employees with number of trainings attended

Query:

```
SELECT e.EmpID, e.FirstName, e.LastName, COUNT(t.`Employee ID`) AS TrainingsAttended
FROM employee_data e
LEFT JOIN training_and_development_data t
ON e.EmpID = t.`Employee ID`
GROUP BY e.EmpID, e.FirstName, e.LastName;
```

Result:

EmpID	FirstName	LastName	TrainingsAttended
3427	Uriah	Bridges	1
3428	Paula	Small	1
3429	Edward	Buck	1
3430	Michael	Riordan	1

Insight: Assesses training participation per employee.

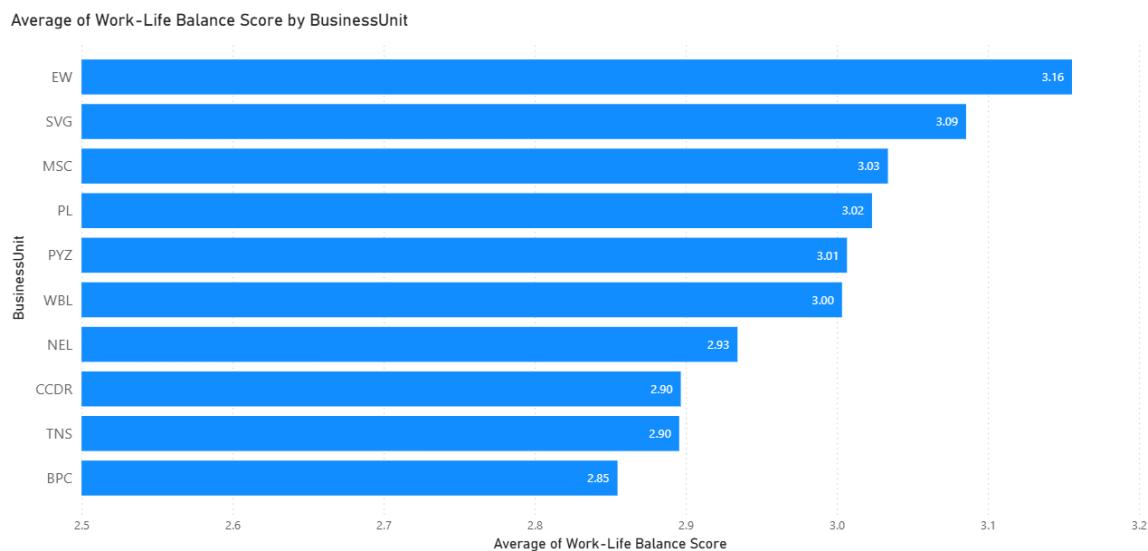
Q25: Average Work-Life Balance Score per BusinessUnit

Query:

```
SELECT e.BusinessUnit, AVG(s.`Work-Life Balance Score`) AS AvgWLB
FROM employee_data e
JOIN employee_engagement_survey_data s
ON e.EmpID = s.`Employee ID`
GROUP BY e.BusinessUnit;
```

Result:

BusinessUnit	AvgWorkLifeBalance
BPC	2.85
NEL	2.93
WBL	3.00
PYZ	3.01
SVG	3.09



Insight: Identifies business units needing work-life balance initiatives.

Q26: Employees whose last training outcome was "Failed"

Query:

```
SELECT t.`Employee ID`, e.FirstName, e.LastName, t.`Training Program Name`, t.`Training Outcome`
FROM training_and_development_data t
JOIN employee_data e
ON t.`Employee ID` = e.EmpID
WHERE t.`Training Outcome` = 'Failed';
```

Result:

Employee ID	FirstName	LastName	Training Outcome
3427	Uriah	Bridges	Failed
3429	Edward	Buck	Failed
3431	Jasmine	Onque	Failed
3436	Joseph	Martins	Failed
3437	Myriam	Givens	Failed

Insight: Highlights employees requiring additional support.

Q27: Employees who left within 1 year of joining**Query:**

```
SELECT EmpID, FirstName, LastName, StartDate, TerminationDate
FROM employee_data
WHERE DATEDIFF(STR_TO_DATE(TerminationDate, '%d-%b-%y'),
STR_TO_DATE(StartDate, '%d-%b-%y')) <= 365;
```

Result:

EmpID	FirstName	LastName	StartDate	ExitDate
3439	Bartholemew	Khemmich	25-May-22	27-Nov-22
3444	Bobby	Rodgers	28-Nov-21	04-Feb-22
3449	Gerald	Preston	10-May-23	27-May-23
3466	Clayton	Walker	13-Apr-22	10-Apr-23
3470	Esteban	Gilbert	14-Nov-18	28-Oct-19

Insight: Identifies short-tenure employees for retention analysis.

Q28: Show top 3 departments with the highest average engagement score**Query:**

```
SELECT e.DepartmentType, AVG(s.`Engagement Score`) AS
AvgEngagement
FROM employee_data e
JOIN employee_engagement_survey_data s
ON e.EmpID = s.`Employee ID`
GROUP BY e.DepartmentType
ORDER BY AvgEngagement DESC
LIMIT 3;
```

Result:

DepartmentType	AvgEngagement
Executive Office	3.3750
IT/IS	3.0256
Sales	2.9909

Q29: Employees with highest training cost**Query:**

```
SELECT t.`Employee ID`, e.FirstName, e.LastName,
ROUND(SUM(t.`Training Cost`),2) AS TotalCost
FROM training_and_development_data t
JOIN employee_data e
```

```

ON t.`Employee ID` = e.EmpID
GROUP BY t.`Employee ID`, e.FirstName, e.LastName
ORDER BY TotalCost DESC
LIMIT 5;

```

Result:

Employee ID	FirstName	LastName	TotalCost
1604	Ansley	Duke	999.97
2011	Naomi	Hardin	999.96
1665	Tania	Jennings	999.83
1768	Lucia	Hudson	999.79
3419	Stanley	Reyes	999.72

Insight: Reveals top training investments per employee.

Q30: Employees with engagement score below department average

Query:

```

SELECT e.EmpID, e.FirstName, e.LastName, e.DepartmentType,
s.`Engagement Score`
FROM employee_data e
JOIN employee_engagement_survey_data s
ON e.EmpID = s.`Employee ID`
WHERE s.`Engagement Score` < (
    SELECT AVG(s2.`Engagement Score`)
    FROM employee_data e2
    JOIN employee_engagement_survey_data s2
    ON e2.EmpID = s2.`Employee ID`
    WHERE e2.DepartmentType = e.DepartmentType
);

```

Result:

EmpID	FirstName	LastName	DepartmentType	Engagement Score
3541	Aaron	Weber	Sales	1
1232	Abbigail	Cameron	Production	2
3591	Abdullah	Ellison	Production	2
3637	Abdullah	Pruitt	Production	2
1296	Adalyn	Vaughan	Production	1

Insight: Identifies employees underperforming in engagement relative to peers.