

HR ANALYTICS



A DEEP DIVE INTO HR DATA WITH POWER BI

In this comprehensive report I am embarking on a journey to explore a rich dataset concerning employee demographics and performance. My goal is to extract valuable insights that can inform strategic decisions and empower our workforce.

I am having a wealth of information at my disposal, including details such as employee age, education level, job satisfaction, and work experience. This data holds immense potential to shed light on employee engagement, performance drivers, and areas for improvement.



HR ANALYTICS

FISCAL YEAR
2015

DEPARTMENT

Human
Resources

Research &
Development

Sales

ATTRITION

711



ATTRITION RATE

16.1%



TOTAL EMPLOYEES

4410



AVERAGE AGE

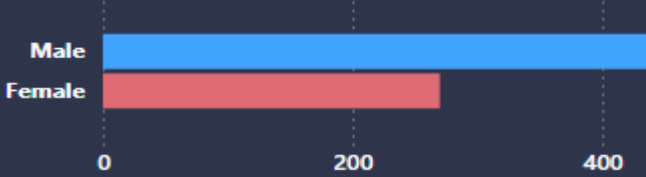
37



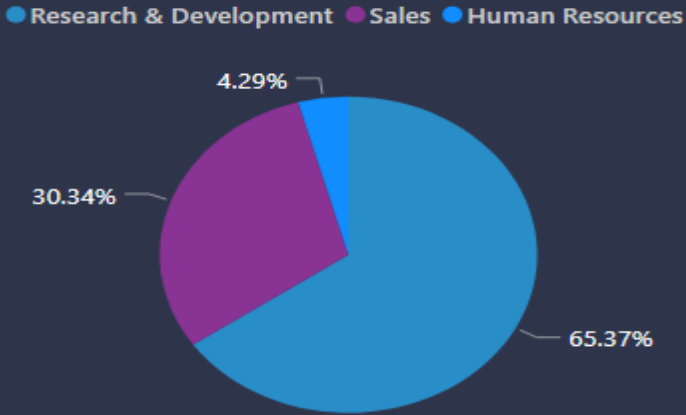
AVERAGE SALARY

65K

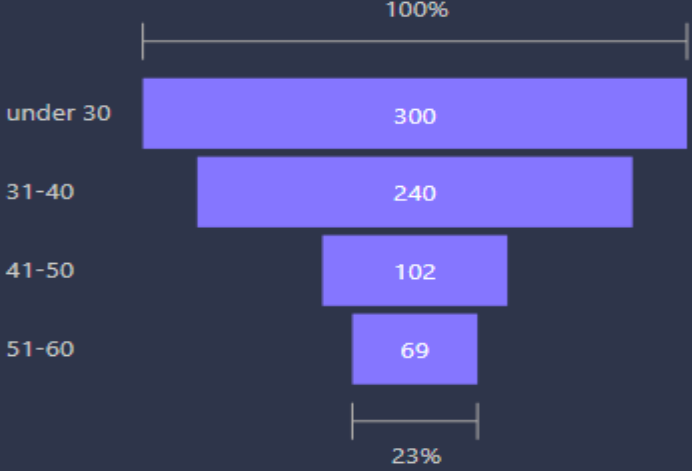
ATTRITION BY GENDER



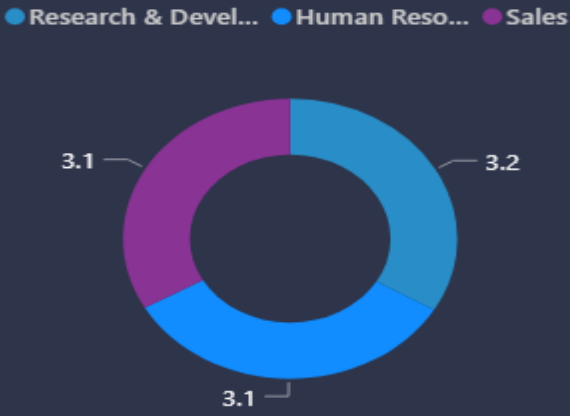
EMPLOYEES BY DEPARTMENT



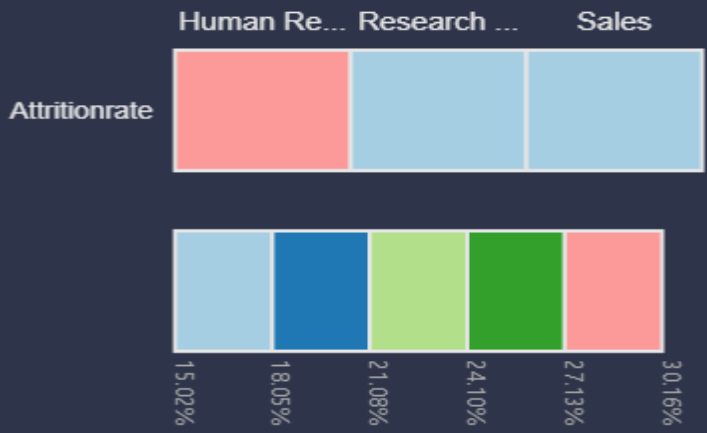
ATTRITION BY AGE GROUP



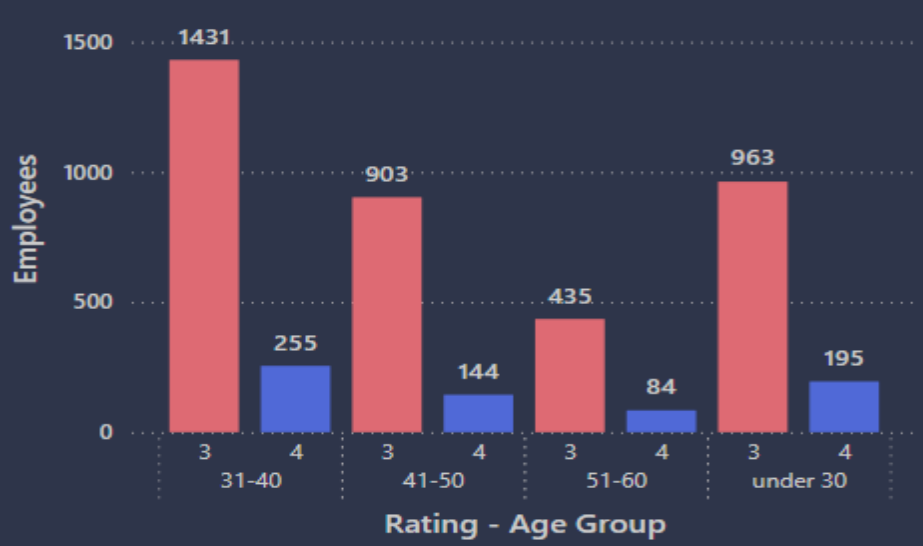
PERFORMANCE RATING BY
DEPARTMENT



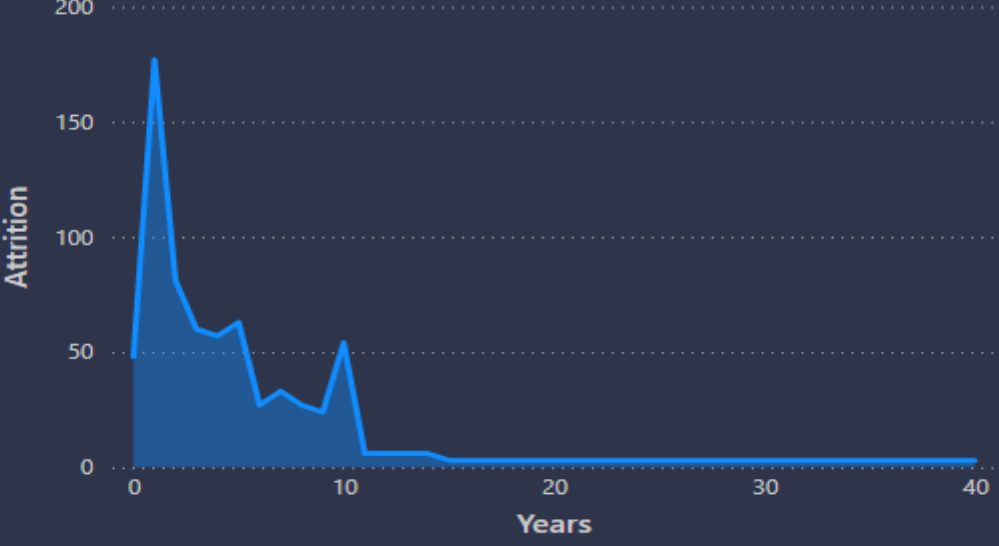
ATTRITION RATE BY DEPARTMENT



EMPLOYEES BY AGE GROUP & PERFORMANCE



ATTRITION BY YEARS AT COMPANY



JOB SATISFACTION

JobRole	1	2	3	4	NA	Total
Sales Executive	186	198	285	305	4	978
Research Scientist	176	189	252	256	3	876
Laboratory Technician	158	147	246	224	2	777
Manufacturing Director	72	87	133	139	4	435
Healthcare Representative	84	66	114	126	3	393
Manager	59	54	99	93	1	306
Sales Representative	69	30	72	77	1	249
Research Director	32	51	75	81	1	240
Human Resources	24	18	17	66	1	126
Total	860	840	1323	1367	20	4410

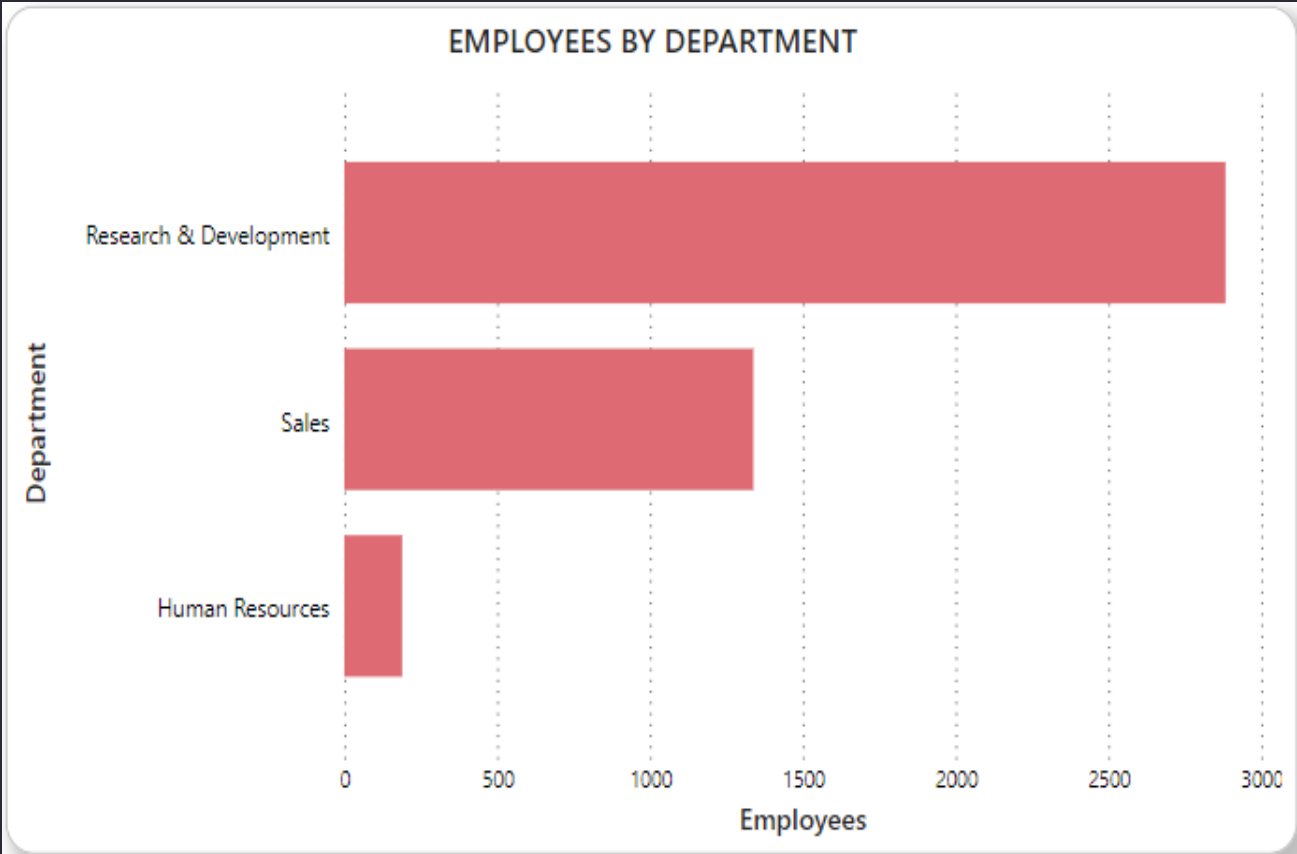
DATA OVERVIEW

The dataset under analysis comprises 4,410 rows and 30 columns, each representing a distinct aspect of the workforce. Below is a brief overview of the key columns present in the dataset:

- **Employee Basics:** Age, Gender, MaritalStatus, Education: Understand the age range, genderdistribution, marital status, and education levels of our workforce.
- **Job Details:** Department, JobRole, JobLevel: Explore the various departments. job roles, and hierarchical levels within the company.
- **Satisfaction and Engagement:** JobSatisfaction, EnvironmentSatisfaction, RelationshipSatisfaction.
- **WorkLifeBalance:** Gauge employee satisfaction and engagement levels in their roles and workplace.
- **Performance and Growth:** Performance Rating, Years SinceLastPromotion. Training TimesLastYear: Assess performance ratings, growth opportunities, and training engagement.
- **Compensation:** MonthlyIncome, StockOptionLevel, Examine salary and benefits, including monthlyincome and stock options.
- **Travel and Commute:** Business Travel, DistanceFromHome, Understand business travel frequency and commute distances.
- **Employment History:** NumCompanies Worked. TotalWorkingYears, PercentSalaryHike: Explor employees' overall professional journeys, including work history and salary hikes.

DATA QUESTIONS

Create a simple bar chart to visualize the distribution of employees by department.



Research & Development department has the largest workforce with nearly 3000 employees, indicating a strong focus on innovation and technical expertise. The Sales department follows with around 1500 employees, reflecting its importance in revenue generation. The Human Resources (HR) department has the fewest employees, approximately 300, highlighting its more administrative role. The significant disparity in workforce distribution suggests strategic prioritization towards R&D and Sales, which can inform HR analytics in optimizing resource allocation and improving organizational efficiency.

Create a slicer to allow users to filter employees based on their job role in Power BI?

EMPLOYEES BY JOB ROLE

JobRole

Manager

☐ Healthcare Representative

☐ Human Resources

☐ Laboratory Technician

☒ Manager

☐ Manufacturing Director

☐ Research Director

☐ Research Scientist

Explain what kind of join you would use to combine the employee data with the in-time and out-time data, and why.

A left outer join from Employee to In-Time would be suitable similarly a left outer join from Employee to Out-Time. To ensure you have all employees in the result, regardless of whether they have in-time or out-time records, a left outer join would be appropriate in both cases. This approach ensures that all employees are included in the resulting dataset, even if they don't have corresponding in-time or out-time records.

Merge

Select a table and matching columns to create a merged table.

general_data

int	EmployeeID	Gender	JobLevel	JobRole	MaritalStatus	MonthlyIncome	NumCom
1	1	Female	1	Healthcare Representative	Married	131160	1
1	2	Female	1	Research Scientist	Single	41890	0
1	3	Male	4	Sales Executive	Married	193280	1
1	4	Male	3	Human Resources	Married	83210	3

in_time

Employee_ID	Date	DateTime
1	02-01-2015	02-01-2015 09:43:45
1	05-01-2015	05-01-2015 10:08:48
1	06-01-2015	06-01-2015 09:54:26
1	07-01-2015	07-01-2015 09:34:31
1	08-01-2015	08-01-2015 09:51:09

Join Kind

Left Outer (all from first, matching from second)

☐ Use fuzzy matching to perform the merge

Fuzzy matching options

Estimating matches based on data previews

OKCancel

Merge

Select a table and matching columns to create a merged table.

general_data

int	EmployeeID	Gender	JobLevel	JobRole	MaritalStatus	MonthlyIncome	NumCom
1	1	Female	1	Healthcare Representative	Married	131160	1
1	2	Female	1	Research Scientist	Single	41890	0
1	3	Male	4	Sales Executive	Married	193280	1
1	4	Male	3	Human Resources	Married	83210	3

out_time

Employee_ID	Date	DateTime
1	02-01-2015	02-01-2015 16:56:15
1	05-01-2015	05-01-2015 17:20:11
1	06-01-2015	06-01-2015 17:19:05
1	07-01-2015	07-01-2015 16:34:55
1	08-01-2015	08-01-2015 17:08:32

Join Kind

Left Outer (all from first, matching from second)

☐ Use fuzzy matching to perform the merge

Fuzzy matching options

Estimating matches based on data previews

OKCancel

Create a calculated column to determine the age group of employees (e.g., under 30, 30-40, 40-50, over 50).

DAX QUERY:

```
1 Agegroup =  
2 VAR Age = [Age]  
3 RETURN  
4     SWITCH(  
5         TRUE(),  
6         Age <= 30, "under 30",  
7         Age >= 30 && Age < 40, "31-40",  
8         Age >= 40 && Age < 50, "41-50",  
9         Age >= 50, "51-60",  
10        BLANK()  
11    )
```

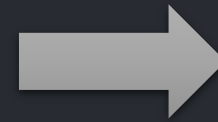


EmployeeID	Age	Agegroup
1	51	51-60
106	33	31-40
113	21	under 30
124	30	under 30
165	19	under 30
206	26	under 30
219	24	under 30
230	20	under 30
289	29	under 30
332	29	under 30
346	45	41-50
406	21	under 30

Calculate the average monthly income for employees and display it in a card visualization.

DAX QUERY :

```
AvgIncome = AVERAGE(general_data[MonthlyIncome])
```



AVERAGE INCOME

65.03K

Create a hierarchy for the date and time columns to allow for easy drill-down analysis.

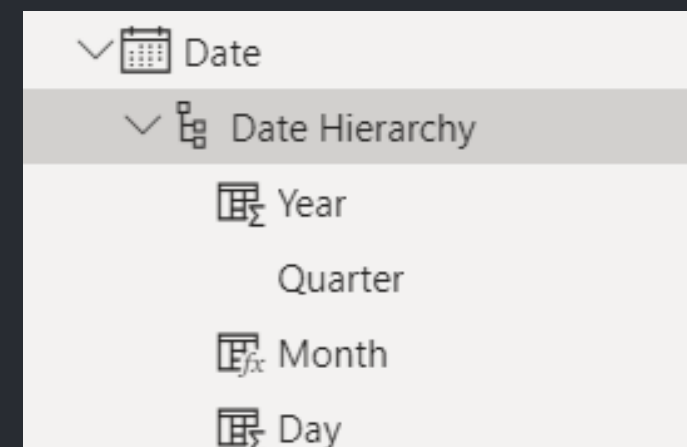
```
Calendarr = CALENDAR(MIN(in_time[DateTime]),MAX(in_time[DateTime]))
```

```
Year = year([Date])
```

```
Month = Format([Date],"MMMM")
```

```
Day = DAY([Date])
```

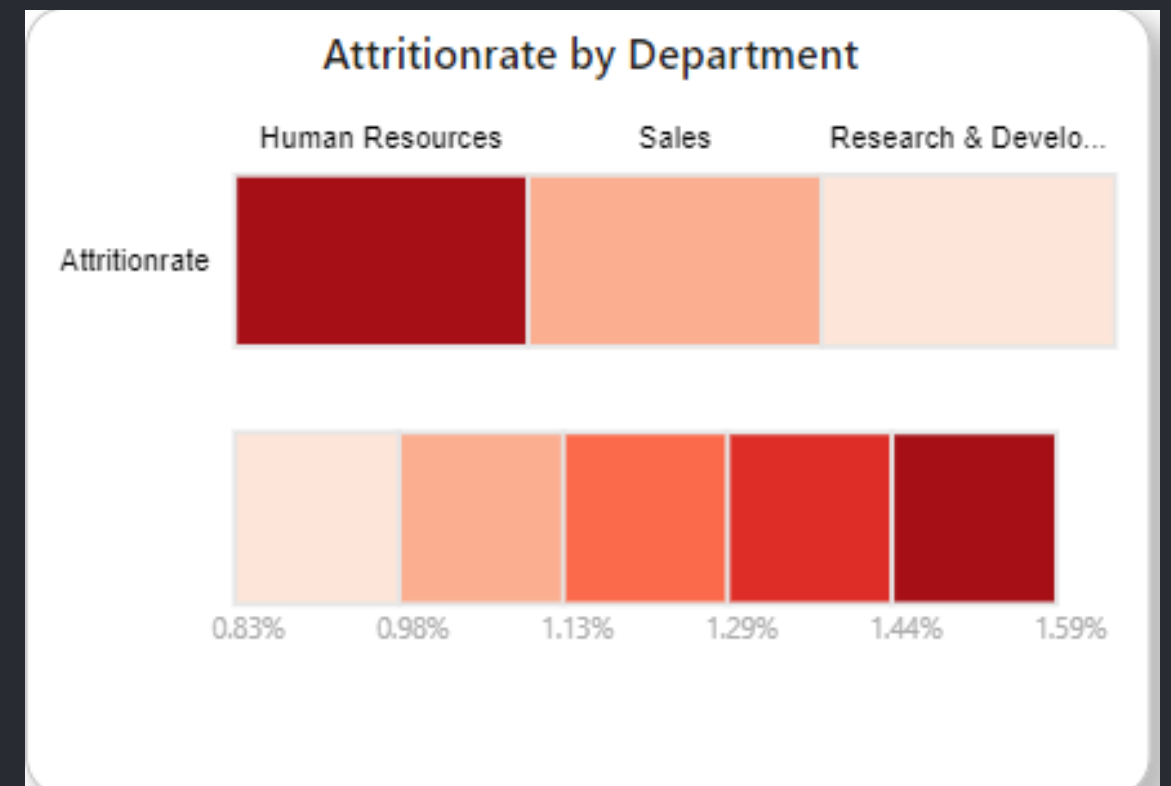
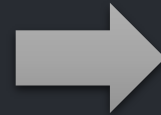
Date	Year	Month	Day
02-01-2015 00:00:00	2015	January	2
03-01-2015 00:00:00	2015	January	3
04-01-2015 00:00:00	2015	January	4
05-01-2015 00:00:00	2015	January	5
06-01-2015 00:00:00	2015	January	6
07-01-2015 00:00:00	2015	January	7
08-01-2015 00:00:00	2015	January	8
09-01-2015 00:00:00	2015	January	9
10-01-2015 00:00:00	2015	January	10
11-01-2015 00:00:00	2015	January	11
12-01-2015 00:00:00	2015	January	12
13-01-2015 00:00:00	2015	January	13
14-01-2015 00:00:00	2015	January	14
15-01-2015 00:00:00	2015	January	15



Calculate the attrition rate for each department and visualize it using a heatmap.

```
Attritioncount = CALCULATE(count(general_data[EmployeeID]),general_data[Attrition]="Yes")
```

```
Attritionrate = DIVIDE([Attritioncount],[Total Employees in Department])
```



Combine the employee data with a different dataset using a left join and explain the potential pitfalls.

Combining two datasets using a left join is a common technique to maintain all records from the left dataset and add matching records from the right dataset, with nulls for non-matching cases. While this approach is useful for ensuring no loss of data from the primary dataset, there are potential pitfalls to be aware of.

Null Values in Right Dataset

When performing a left join, if there are records in the left dataset (Employee) without corresponding matches in the right dataset, the result will contain nulls for the non-matching fields.

Duplication of Rows

If the right dataset has multiple matches for a single record in the left dataset, a left join will create multiple rows, leading to duplication.

Incorrect or Inconsistent Keys

If there's inconsistency or incorrect data in the joining key, a left join might not behave as expected. This can happen due to typos, case sensitivity issues, or misaligned data.

Create a dynamic filter that allows users to filter employees based on both department and job role simultaneously.

Filters on this visual

...

Department

is (All)

Filter type ⓘ

Basic filtering

Search

☒ Select all

☒ Human Resources 9

☒ Research & Devel... 207

☒ Sales 90

☐ Require single selection

Filters on this visual

...

JobRole

is Healthcare Representative

Filter type ⓘ

Basic filtering

Search

☒ Select all

☒ Healthcare Representative 393

☐ Human Resources 156

☐ Laboratory Technician 777

☐ Manager 306

☐ Manufacturing Director 435

☐ Research Director 240

☐ Require single selection

Add data fields here

Apply conditional formatting to a table to highlight employees with the highest and lowest monthly incomes.

Background color - Background color

Format style

Gradient

Apply to

Values only

What field should we base this on?

Sum of MonthlyIncome

Summarization

Sum

How should we format empty values?

As zero

Minimum

Lowest value

Enter a value

Maximum

Highest value

Enter a value

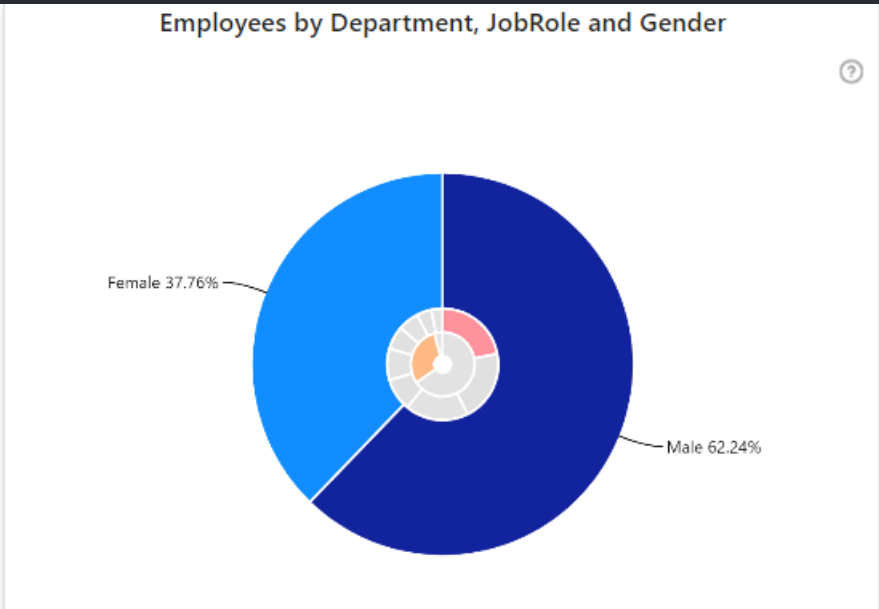
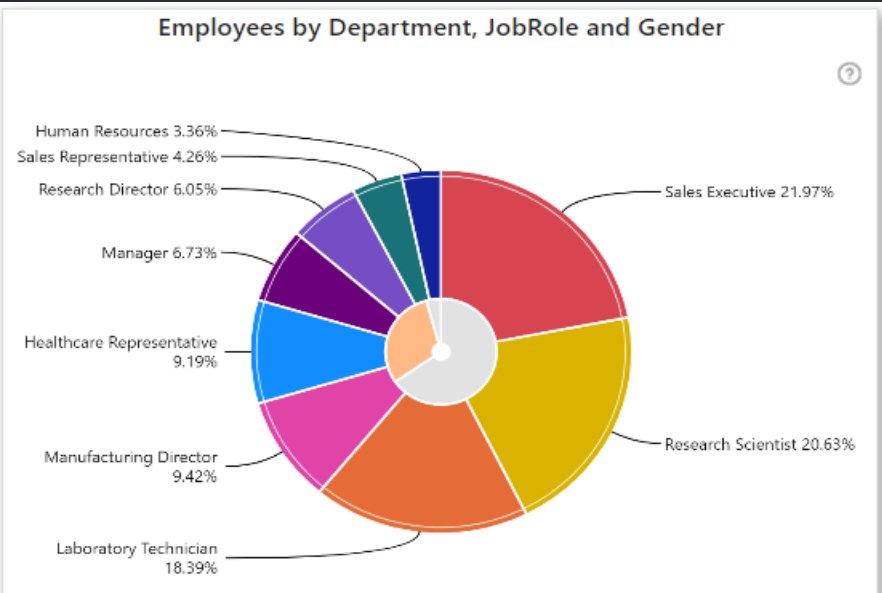
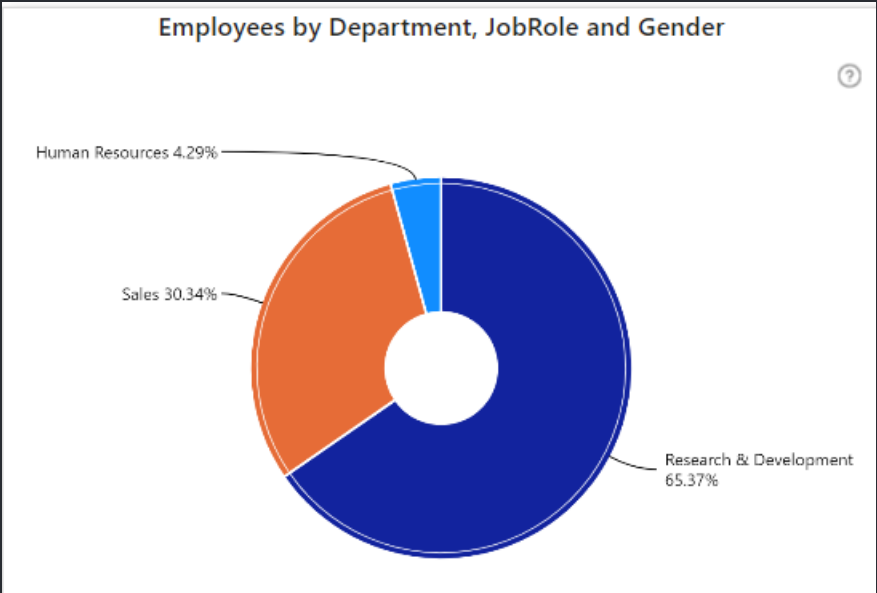
Add a middle color



EmployeeID	Sum of MonthlyIncome
1	131160
2	41890
3	193280
4	83210
5	23420
6	40710
7	58130
8	31430
9	20440
10	134640
11	79910
12	33770
13	55380
Total	286779270

Create a custom visual or use a third-party visual to present data in a way that's not available in the default Power BI visuals.

DRILL DOWN DONUT PRO VISUAL

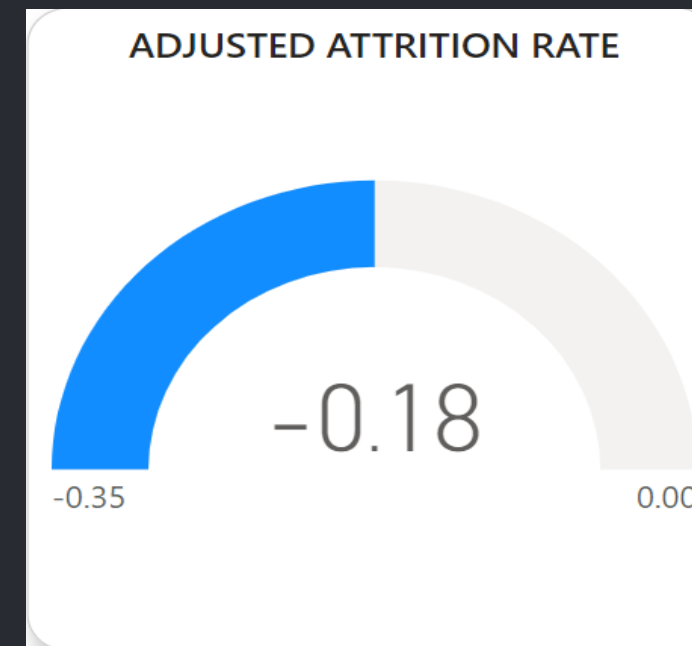
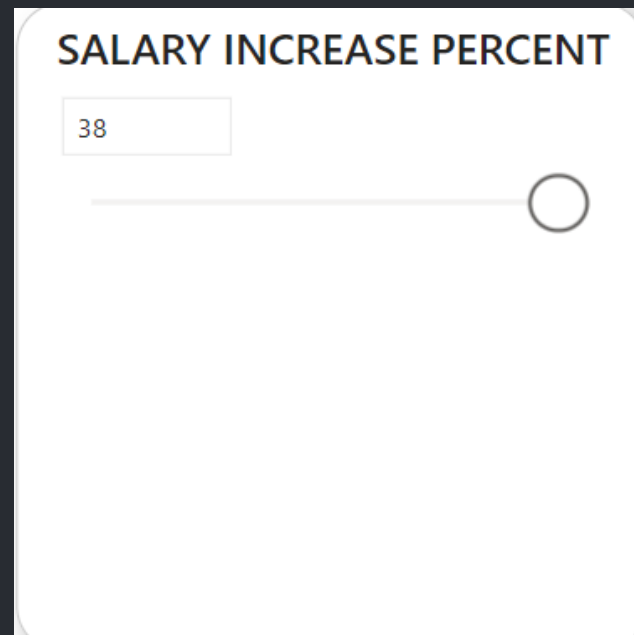


Explain how to optimize performance by creating aggregations for large datasets.

- Aggregations are pre-calculated summaries of data that can significantly improve query performance, especially when working with large datasets
- Creating Aggregations for Suitable Large Datasets that are Frequently queried, Complex calculations involved and has Multiple data sources
- Create Dedicated Aggregation Table, Define Aggregation Measures, Use DAX functions (SUM, COUNT, AVERAGE)
- Use Aggregations in Visualizations
 1. Reference aggregated measures
 2. Improve rendering speed
- Monitor and Optimize
 1. Performance Analyzer
 2. Query Diagnostics
- Benefits:
 - Significantly improved query performance
 - Faster rendering of visualizations
 - Balance between optimization and data freshness

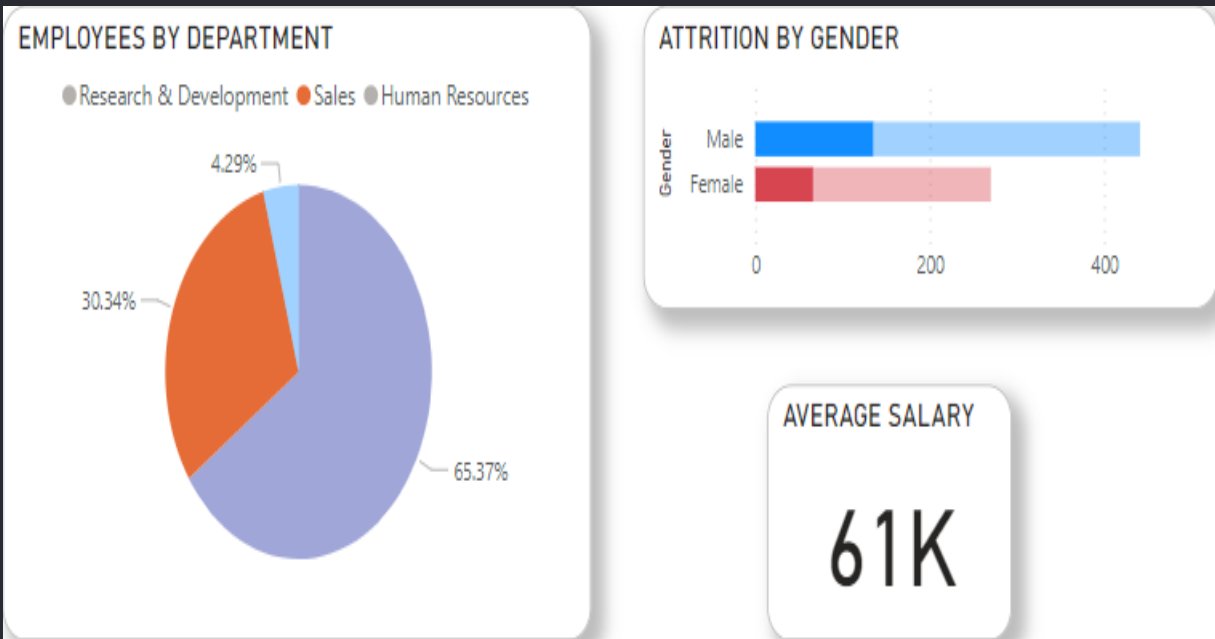
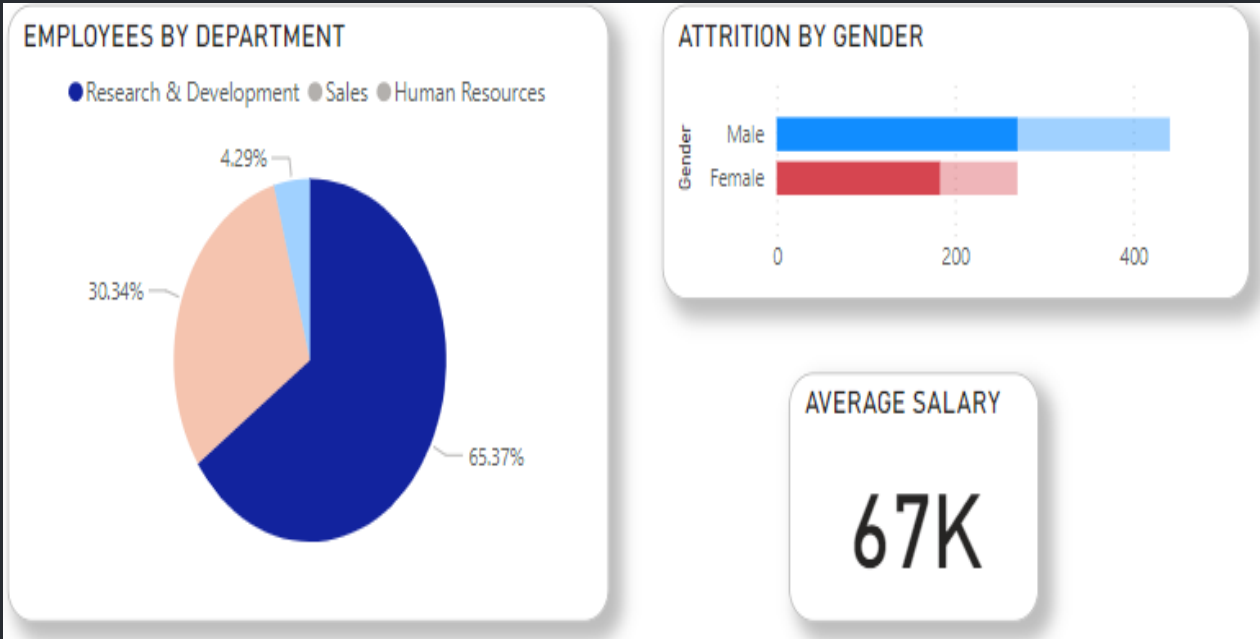
Use What-If parameters to show how attrition rates change when you adjust different factors (e.g., salary increase).

Created a parameter for salary increase in percentage points (from 0% to 40%, with 1% increments) and added it as a slicer to the report which allows users to adjust the parameter dynamically. This modifies the attrition rate based on the parameter value.



Demonstrate the use of cross-filtering between visuals to provide an interactive experience for users.

Cross-filtering between visuals is a powerful feature in Power BI that allows users to interact with one visual, and have other visuals on the same page update automatically based on the selection. This creates an interactive and engaging experience for users, enabling them to explore and analyze data more effectively.



CONCLUSION

In summary, the analysis of human resource data using Power BI has provided invaluable insights into various aspects of workforce. Through targeted data inquiries and the creation of a comprehensive dashboard, I've gained deeper insights into employee demographics, attrition rates, educational backgrounds, and performance metrics. Leveraging the visualization capabilities of Power BI, I've effectively represented and interpreted the data, empowering more informed decision-making.