



HR ANALYTICS DASHBOARD

Power BI Data Warehouse
& Business Intelligence Portfolio

MD Monajir Hussain | B1393756



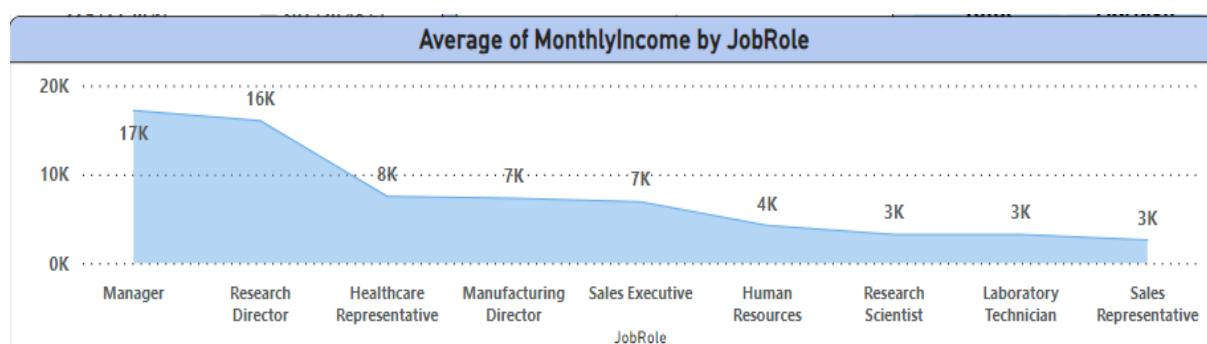
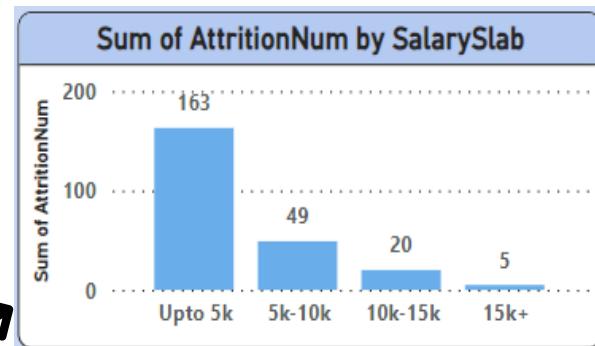
Table of Contents

1) Executive Summary.....	3
2) Introduction to the Power BI Project.....	4
3) Data Source Description and Business Questions	4
4) Data Pre-Processing and Data Cleaning.....	5
4.1) Dataset CSV View.....	6
4.2) Data Cleaning.....	8
Removing Nulls:	8
Removing Duplicates:	10
Spelling Error/ Find and Replace.....	12
Data Types.....	12
Removing Columns	13
Adding Columns	13
Adding a Table.....	15
New Measure.....	15
5) Data Modelling	15
6) Data Visualization	18
KPI 1:	19
KPI 2:	20
KPI 3:	20
KPI 4:	20
KPI 5:	21
KPI 6:	22
KPI 7:	22
KPI 8:	23
KPI 9:	24
KPI 10:	24
KPI 11:	25
KPI 12:	25
KPI 13:	26
KPI 14:	26
KPI 15:	27
7) The Final Dashboard	27
8) Conclusions and Recommendations.....	28
9) Reference	28

1) Executive Summary

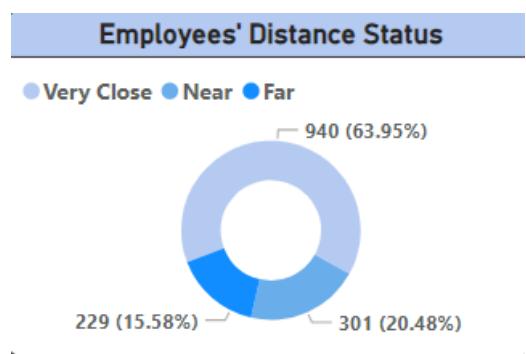
This HR Analytics Dashboard project is dedicated to improving employee performance and minimizing attrition by harnessing the capabilities of Microsoft Power BI. The dashboard focuses on delivering a comprehensive analysis of key performance indicators (KPIs) such as attrition rates, job satisfaction, overtime patterns, and monthly income distributions. These KPIs provide a holistic view of workforce dynamics, empowering stakeholders to make informed decisions for strategic human resource management.

The pivotal KPIs highlighted in the dashboard are Attrition Rates and Attrition by Salary Slab. In the bar chart of Attrition by Salary Slab, it is observed that employees with lower salary slabs, have a higher attrition rate, indicating a potential correlation between salary levels and employee turnover. However, it is crucial to consider the larger number of employees in the lower salary range, influencing the attrition numbers.



Moving to the next key indicator, the Average Monthly Income by Job Role sheds light on salary distributions across various roles. Managers exhibit the highest average monthly income, while sales representatives tend to have lower incomes on average. These insights are instrumental in formulating targeted strategies for salary adjustments, employee engagement, and overall workforce management.

Another significant KPI, Distance from Home to Work, adds a nuanced dimension to the dashboard. A new column was added to categorize distances as very close, near, and far, providing a detailed breakdown of how the distance factor contributes to employee dynamics. This categorization aids in identifying potential correlations between commuting distances and attrition rates, offering valuable insights for workplace planning and employee satisfaction initiatives.



2) Introduction to the Power BI Project

In the ever-evolving landscape of business, workforce optimization stands as a cornerstone for organizational success. This HR Analytics Dashboard project, leveraging Microsoft Power BI, is a strategic endeavour dedicated to improving employee performance and curbing attrition.

Referencing the comprehensive capabilities of Power BI, the project is set to transform raw data into actionable insights. By focusing on key performance indicators such as attrition rates, salary distributions, and commuting patterns, the dashboard aims to provide decision-makers with a comprehensive view of workforce dynamics. In doing so, it empowers stakeholders to make informed decisions, refine HR strategies, and foster an engaged organizational culture. With a commitment to data analytics and visualization, this project aspires to be a catalyst for aligning workforce management with strategic business objectives, ushering in a future of enhanced efficiency and resilience for the organization.

A1	EmployeeID	Age	AgeGroup	Attrition	BusinessTravel	Department	DistanceFromHome	Education	EmployeeCount	EnvironmentSatisfaction	HourlyRate	JobInvolvement	JobLevel	JobRole	JobSatisfaction	MaritalStatus	MonthlyIncome	MonthlyRate	NumCompaniesWorked	OverTime	PercentSalaryHike	PerformanceRating	RelationshipSatisfaction	TotalWorkingYears
1	RM297	18	18-25	Yes	Travel_Fre	200 Research	3	3 Life Scienc	1	400	1	Male	54	3	1 Laboratori	3 Sales Rep	5 Single	2500 Upto Sk	2500 Upto Sk	1 Y	13	3	3	10
2	RM402	18	18-25	No	Travel_Fre	813 Sales	10	3 Medical	1	411	1	Female	69	2	1 Sales Rep	3 Single	2200 Upto Sk	2724	1 Y	No	12	3	3	10
3	RM458	18	18-25	Yes	Travel_Fre	1306 Sales	5	3 Marketing	1	614	2	Male	69	3	1 Sales Rep	2 Single	1870 Upto Sk	8059	1 Y	Yes	14	3	4	10
4	RM728	18	18-25	No	Non-Trav	287 Research	5	2 Life Scien	1	1012	2	Male	73	3	1 Research	4 Single	1051 Upto Sk	13493	1 Y	No	15	3	4	10
5	RM829	18	18-25	Yes	Non-Trav	247 Research	8	1 Medical	1	1156	3	Male	80	3	1 Laboratori	3 Single	1904 Upto Sk	113556	1 Y	No	12	3	4	10
6	RM1073	18	18-25	No	Non-Trav	1124 Research	1	3 Life Scienc	1	1368	4	Female	97	3	1 Laboratori	4 Single	1611 Upto Sk	19305	1 Y	No	15	3	3	10
7	RM1154	18	18-25	Yes	Travel_Fre	544 Sales	3	2 Medical	1	1622	2	Female	70	3	1 Sales Rep	18420	1 Y	Yes	12	3	3	10		
8	RM1429	18	18-25	No	Non-Trav	526 Sales	14	3 Life Scienc	1	1831	2	Female	33	3	1 Research	3 Single	1321 Upto Sk	54	1 Y	No	16	3	3	10
9	RM1430	18	18-25	Yes	Travel_Fre	526 Sales	22	1 Marketing	1	167	4	Male	50	3	1 Sales Rep	3 Single	1675 Upto Sk	26800	1 Y	Yes	19	3	4	10
10	RM1510	19	18-25	Yes	Travel_Fre	1181 Research	3	1 Medical	1	201	2	Female	79	3	1 Laboratori	2 Single	1483 Upto Sk	14102	1 Y	No	14	3	4	10
11	RM1612	19	18-25	Yes	Travel_Fre	602 Sales	1	1 Technical	1	235	100	1	1 Sales Rep	1 Single	2325 Upto Sk	20989	0 Y	No	21	4	1	10		
12	RM1718	19	18-25	Yes	Travel_Fre	303 Research	2	3 Life Scien	1	243	2	Male	47	2	1 Laboratori	4 Single	1102 Upto Sk	9241	1 Y	No	22	4	3	10
13	RM4423	19	18-25	Yes	Travel_Fre	489 Human Re	2	2 Technical	1	566	1	Male	52	2	1 Human Re	2564 Upto Sk	18437	1 Y	No	12	3	3	10	
14	RM4849	19	18-25	Yes	Travel_Fre	54 Sales	21	3 Sales Rep	1	959	1	Male	37	2	1 Sales Rep	2 Single	1500 Upto Sk	9447	1 Y	Yes	13	3	2	10
15	RM5252	19	18-25	Yes	Travel_Fre	845 Research	9	3 Life Scien	1	1034	1	Male	54	3	1 Research	1 Single	2520 Upto Sk	1175	1 Y	No	25	4	3	10
16	RM5899	19	18-25	Yes	Non-Trav	504 Research	10	3 Medical	1	1248	1	Female	96	2	1 Research	2 Single	1859 Upto Sk	6148	1 Y	Yes	25	4	2	10
17	RM8910	19	18-25	Yes	Non-Trav	104 Research	10	3 Life Scien	1	1269	2	Female	57	4	1 Research	4 Single	2994 Upto Sk	21221	1 Y	Yes	12	3	4	10
18	RM1010	20	18-25	Yes	Travel_Fre	871 Research	6	3 Life Scien	1	137	4	Female	66	2	1 Laboratori	4 Single	2926 Upto Sk	19783	1 Y	Yes	18	3	2	10
19	RM4488	20	18-25	Yes	Travel_Fre	955 Research	1	3 Life Scien	1	657	4	Female	83	2	1 Research	2 Single	2838 Upto Sk	11757	1 Y	No	13	3	4	10
20	RM5134	20	18-25	Yes	Travel_Fre	1362 Research	10	1 Medical	1	701	4	Male	32	3	1 Research	3 Single	1009 Upto Sk	26999	1 Y	Yes	11	3	4	10
21	RM5302	20	18-25	Yes	Travel_Fre	40 Sales	2	3 Medical	1	922	3	Female	49	2	1 Sales Rep	3 Single	2044 Upto Sk	22052	1 Y	No	13	3	4	10
22	RM5599	20	18-25	Yes	Travel_Fre	129 Research	4	3 Sales Rep	1	960	1	Male	84	3	1 Laboratori	1 Single	2571 Upto Sk	1110	1 Y	No	19	3	2	10
23	RM5990	20	18-25	Yes	Travel_Fre	109 Research	11	3 Medical	1	1016	4	Female	98	2	1 Research	1 Single	2600 Upto Sk	18275	1 Y	Yes	15	3	1	10
24	RM7322	20	18-25	Yes	Travel_Fre	769 Sales	9	3 Marketing	1	1077	4	Female	54	3	1 Sales Rep	4 Single	2323 Upto Sk	17205	1 Y	Yes	14	3	2	10
25	RM7777	20	18-25	No	Travel_Fre	805 Research	3	3 Life Scien	1	1194	1	Male	87	2	1 Laboratori	3 Single	3031 Upto Sk	12828	1 Y	No	12	3	1	10
26	RM8577	20	18-25	No	Travel_Fre	654 Sales	21	3 Marketing	1	1226	3	Male	43	4	1 Sales Rep	4 Single	2678 Upto Sk	5050	1 Y	No	17	3	4	10
27	RM8777	20	18-25	No	Travel_Fre	1141 Sales	2	3 Medical	1	1657	3	Female	31	3	1 Sales Rep	3 Single	2783 Upto Sk	11251	1 Y	No	11	3	1	10
28	RM1179	20	18-25	No	Travel_Fre	727 Sales	9	3 Sales Rep	1	1680	4	Male	54	3	1 Sales Rep	1 Single	2720 Upto Sk	21082	1 Y	No	14	3	4	10
29	RM1198	20	18-25	No	Travel_Fre	391 Research	15	2 Life Scien	1	1680	3	Male	90	3	1 Research	4 Single	1232 Upto Sk	13180	1 Y	No	14	3	4	10
30	RM1255	21	18-25	No	Travel_Fre	396 Research	3	2 Medical	1	379	4	Male	100	2	1 Research	3 Single	3230 Upto Sk	16531	1 Y	No	17	3	1	10
31	RM2755	21	18-25	No	Travel_Fre	756 Sales	1	1 Technical	1	478	1	Female	99	2	1 Sales Rep	2 Single	2174 Upto Sk	9150	1 Y	Yes	11	3	3	10
32	RM3588	21	18-25	No	Non-Trav	895 Sales	9	2 Medical	1	484	1	Male	39	3	1 Sales Rep	4 Single	2620 Upto Sk	2851	1 Y	No	24	4	3	10
33	RM3683	21	18-25	No	Non-Trav	156 Sales	12	3 Life Scien	1	494	3	Female	90	4	1 Sales Rep	2 Single	2716 Upto Sk	25422	1 Y	No	15	3	4	10
34	RM3711	21	18-25	Yes	Travel_Fre	1343 Sales	22	1 Technical	1	669	3	Male	49	3	1 Sales Rep	3 Single	3447 Upto Sk	24444	1 Y	No	11	3	3	10
35	RM4497	21	18-25	Yes	Travel_Fre	4427 Research	18	1 Other	1	923	4	Female	65	3	1 Research	4 Single	2693 Upto Sk	2870	1 Y	No	19	3	1	10
36	RM6664	21	18-25	Yes	Travel_Fre	1334 Research	10	3 Life Scien	1	1079	3	Female	36	2	1 Laboratori	1 Single	1416 Upto Sk	17358	1 Y	No	13	3	1	10
37	RM7778	21	18-25	Yes	Travel_Fre	109 Sales	10	3 Sales Rep	1	1079	3	Female	36	2	1 Laboratori	1 Single	1416 Upto Sk	17358	1 Y	No	13	3	1	10

The dataset chosen for this project offers a comprehensive view of employee-related metrics, encompassing performance indicators and factors contributing to attrition. Its relevance to understanding employee behaviour makes it an ideal resource for driving insights. Before using Power BI for visualization and analytics, the dataset needs to be cleaned and transformed to some sort to tackle error in implementation.

This introduction lays the groundwork for HR Analytics endeavour, highlighting its significance in shaping the HR strategies and fostering a more engaged and retained workforce.

3) Data Source Description and Business Questions

The dataset selected for this HR Analytics project has been sourced from Kaggle, the preeminent platform in the data science community (Fares, 2022). Specifically, the dataset utilized is the "IBM HR Analytics Employee Attrition & Performance," accessible as a free-to-download CSV file.

The dataset provides a comprehensive snapshot of employee attributes, allowing a deep dive into various dimensions of workforce dynamics. It encapsulates essential details ranging from personal information to professional roles, providing a rich source for insightful analysis.

The screenshot displays the Power BI Data view for the 'HR_Analytics' dataset. On the left, a tree view shows all tables in the model, with 'HR_Analytics' expanded to show its columns: Age, AgeGroup, Attrition, BusinessTravel, DailyRate, Department, DistanceFromHome, DistanceToNearest, Education, EducationField, EmpID, EmployeeCount, EmployeeNumber, EnvironmentSatisfaction, Gender, HourlyRate, JobInvolvement, JobLevel, JobRole, JobSatisfaction, MaritalStatus, MonthlyIncome, MonthlyRate, NumCompaniesWorked, Over18, Overtime, PercentSalaryHike, PerformanceRating, RelationshipSatisfaction, Salary, SalaryDob, StandardHours, and WorkLifeBalance. On the right, the 'Properties' pane is open, showing the dataset's name, description, and a list of columns under the 'Data' tab. The 'Edit' button is visible at the bottom of the properties pane.

The selection of this dataset stems from its comprehensive coverage of pertinent employee-related attributes, making it instrumental in addressing critical business questions related to performance optimization and attrition reduction strategies.

Business Questions Addressed:

1. Why This Dataset?

The selection of this dataset is grounded in its richness, offering a diverse array of HR-related features crucial for a nuanced analysis of workforce dynamics.

2. Business Skills Development:

Beyond the dataset's content, its utilization is instrumental in honing essential business skills related to workforce analytics, strategic decision-making, and talent management.

3. BI Project Objectives:

The overarching goal is to unravel the intricate web of factors influencing employee performance and attrition. The dataset serves as the canvas for crafting targeted strategies and interventions.

4. Focused Features:

The project zeroes in on specific features like Attrition Rates, Job Satisfaction, Overtime, and Monthly Income, aligning with the key performance indicators critical for HR management.

5. Addressing Big Data:

While not categorized as "Big Data" in the traditional sense, the dataset mirrors real-world scenarios, allowing for a pragmatic exploration of HR analytics within a manageable scale.

4) Data Pre-Processing and Data Cleaning

The dataset includes a total of 38 columns. However, to create a clear and concise dashboard, unnecessary columns need to be filtered out, focusing on the most relevant information.

4.1) Dataset CSV View

It serves as the foundation for the dashboard, facilitating the proper selection of Key Performance Indicators (KPIs) tailored to the project's specific needs. Before integrating data into Power BI, a preprocessing phase is conducted to incorporate only the essential columns required to address the KPI-related inquiries.

EmpID	Age	Attrition	BusinessTravel	Department	DistanceFromHome	Education	EnvironmentSatisfaction	Gender	HourlyRate	JobInvolvement	JobLevel	JobRole	JobSatisfaction	MaritalStatus	MonthlyIncome	NumCompaniesOver18	Overtime	PercentSalaryHike	PerformanceRating	RelationshipSatisfaction	StandardHours	TotalWorkingYears	TrainingTimesLastYear	WorkLifeBalance	YearsAtCompany	
2	RM297	18-25	Yes	Travel_Rare	230 Research	3	3 Life Sci	Male	405	3	1	Laborator	3 Single	1420 Upto Sk	25223	1 Y	No	13	3	3						
3	RM302	18-25	No	Travel_Rare	812 Sales	10	3 Medical	1	411	4	Female	69	2	1 Sales Rep	3 Single	1200 Upto Sk	9724	1 Y	No	12	3	1				
4	RM458	18-25	Yes	Travel_Freq	100 Research	5	3 Marketing	1	614	3	2	Female	69	3	1 Sales Rep	2 Single	1370 Upto Sk	8460	1 Y	Yes	14	3	4			
5	RM728	18-25	No	Travel_Freq	261 Research	5	2 Life Sci	1	1202	2	2	Male	70	3	1 Research	4 Single	1051 Upto Sk	13493	1 Y	No	15	3	4			
6	RM629	18-25	Yes	Non-Trav	247 Research	8	1 Medical	1	1156	3	3	Male	80	3	1 Laborator	3 Single	1904 Upto Sk	13556	1 Y	No	12	3	4			
7	RM973	18-25	No	Non-Trav	1124 Research	1	3 Life Sci	1	1368	4	Female	97	3	1 Laborator	4 Single	1611 Upto Sk	19305	1 Y	No	15	3	3				
8	RM154	18-25	Yes	Travel_Freq	544 Sales	3	2 Medical	1	162	2	Female	70	3	1 Sales Rep	4 Single	1569 Upto Sk	18420	1 Y	Yes	12	3	3				
9	RM1132	18-25	No	Non-Trav	1431 Research	14	3 Medical	1	167	4	Male	50	3	1 Research	3 Single	1514 Upto Sk	8018	1 Y	No	16	3	3				
10	RM128	19-25	Yes	Travel_Rare	113 Research	3	2 Marketing	1	201	4	Male	79	3	1 Sales Rep	3 Single	1480 Upto Sk	26820	1 Y	Yes	19	3	4				
11	RM150	19-25	No	Travel_Rare	100 Research	3	2 Marketing	1	201	4	Female	79	3	1 Sales Rep	3 Single	1480 Upto Sk	18000	1 Y	No	14	3	4				
12	RM172	19-25	Yes	Travel_Freq	602 Sales	1	1 Technical	1	1235	3	Female	100	1	1 Sales Rep	1 Single	2325 Upto Sk	20980	0 Y	No	21	4	1				
13	RM178	19-25	Yes	Travel_Freq	303 Research	2	3 Life Sci	1	243	2	Male	47	2	1 Laborator	4 Single	1102 Upto Sk	9241	1 Y	No	22	4	3				
14	RM423	19-25	Yes	Travel_Freq	489 Human Re	2	2 Technical	1	566	1	Male	52	2	1 Human Re	4 Single	2564 Upto Sk	18437	1 Y	No	12	3	3				
15	RM689	19-25	Yes	Travel_Rare	419 Sales	21	3 Other	1	959	4	Male	37	2	1 Sales Rep	2 Single	2121 Upto Sk	9947	1 Y	Yes	13	3	2				
16	RM854	19-25	No	Travel_Rare	645 Research	9	2 Life Sci	1	1193	3	Male	54	3	1 Research	1 Single	2522 Upto Sk	7172	1 Y	No	25	4	3				
17	RM110	19-25	Yes	Travel_Rare	100 Research	10	3 Life Sci	1	1248	4	Female	96	2	1 Research	2 Single	2522 Upto Sk	11250	1 Y	Yes	25	4	2				
18	RM910	19-25	No	Travel_Rare	260 Research	25	3 Life Sci	1	1209	2	Male	57	4	1 Research	4 Single	2594 Upto Sk	22221	1 Y	Yes	12	3	4				
19	RM103	20-25	Yes	Travel_Freq	871 Research	6	3 Life Sci	1	137	4	Female	66	2	1 Laborator	4 Single	2926 Upto Sk	15783	1 Y	Yes	18	3	2				
20	RM448	20-25	No	Travel_Rare	595 Research	1	3 Life Sci	1	657	4	Female	83	2	1 Research	2 Single	2838 Upto Sk	11757	1 Y	No	13	3	4				
21	RM514	20-25	Yes	Travel_Rare	1362 Research	10	1 Medical	1	701	4	Male	32	3	1 Research	3 Single	1009 Upto Sk	26999	1 Y	Yes	11	3	4				
22	RM663	20-25	No	Travel_Rare	500 Sales	2	3 Medical	1	922	3	Female	49	2	1 Sales Rep	3 Single	2044 Upto Sk	22052	1 Y	No	13	3	4				
23	RM690	20-25	Yes	Travel_Rare	129 Research	4	3 Technical	1	961	1	Male	84	3	1 Laborator	1 Single	2971 Upto Sk	13008	1 Y	No	19	3	2				
24	RM732	20-25	Yes	Travel_Rare	100 Research	11	3 Marketing	1	1016	4	Female	98	2	1 Sales Rep	1 Single	2601 Upto Sk	18275	1 Y	Yes	15	3	1				
25	RM777	20-25	No	Travel_Freq	765 Sales	9	3 Marketing	1	1077	7	Female	54	3	1 Sales Rep	4 Single	2323 Upto Sk	13750	1 Y	Yes	14	3	2				
26	RM857	20-25	No	Travel_Rare	805 Research	3	3 Life Sci	1	1198	1	Male	87	2	1 Laborator	3 Single	3033 Upto Sk	15282	1 Y	No	12	3	1				
27	RM877	20-25	No	Travel_Rare	654 Sales	21	3 Marketing	1	1228	3	Male	43	4	1 Sales Rep	4 Single	2678 Upto Sk	5050	1 Y	No	17	3	4				
28	RM1179	20-25	No	Travel_Rare	1141 Sales	2	3 Medical	1	1657	3	Female	31	3	1 Sales Rep	3 Single	2783 Upto Sk	11251	1 Y	No	19	3	1				
29	RM1196	20-25	No	Travel_Rare	727 Sales	9	1 Life Sci	1	1680	4	Male	54	3	1 Sales Rep	1 Single	2728 Upto Sk	21082	1 Y	No	11	3	1				
30	RM024	21-25	No	Travel_Rare	391 Research	15	2 Life Sci	1	30	3	Male	96	3	1 Research	4 Single	1232 Upto Sk	15281	1 Y	No	14	3	4				
31	RM1275	21-25	Yes	Travel_Freq	996 Research	3	2 Marketing	1	379	4	Male	1000	2	1 Sales Rep	3 Single	3230 Upto Sk	16531	1 Y	No	17	3	1				
32	RM168	21-25	Yes	Travel_Freq	756 Sales	1	1 Technical	1	478	1	Female	99	2	1 Sales Rep	2 Single	2174 Upto Sk	9150	1 Y	Yes	11	3	1				
33	RM663	21-25	No	Non-Trav	895 Sales	9	2 Medical	1	484	1	Male	39	3	1 Sales Rep	4 Single	2630 Upto Sk	2851	1 Y	No	24	4	3				
34	RM371	21-25	Yes	Travel_Rare	156 Sales	12	3 Life Sci	1	494	3	Female	90	4	1 Research	2 Single	2716 Upto Sk	25422	1 Y	No	15	3	4				
35	RM497	21-25	No	Travel_Rare	1343 Sales	22	2 Technical	1	669	3	Male	49	3	1 Sales Rep	3 Single	3447 Upto Sk	24444	1 Y	No	11	3	3				
36	RM664	21-25	Yes	Travel_Rare	1427 Research	18	1 Other	1	923	4	Female	65	3	1 Research	4 Single	2693 Upto Sk	8870	1 Y	No	19	3	1				
37	RM778	21-25	Yes	Travel_Rare	1334 Research	10	3 Life Sci	1	1079	3	Female	36	2	1 Laborator	1 Single	1416 Upto Sk	17258	1 Y	No	13	3	1				

Now, this dataset is inserted into Power BI.

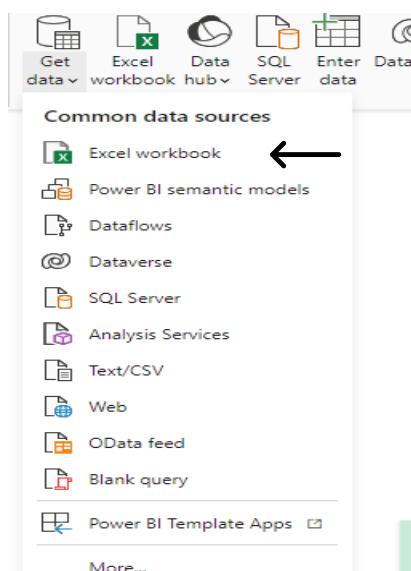
Loading Data:

Add data to your report
Once loaded, your data will appear in the Data pane.

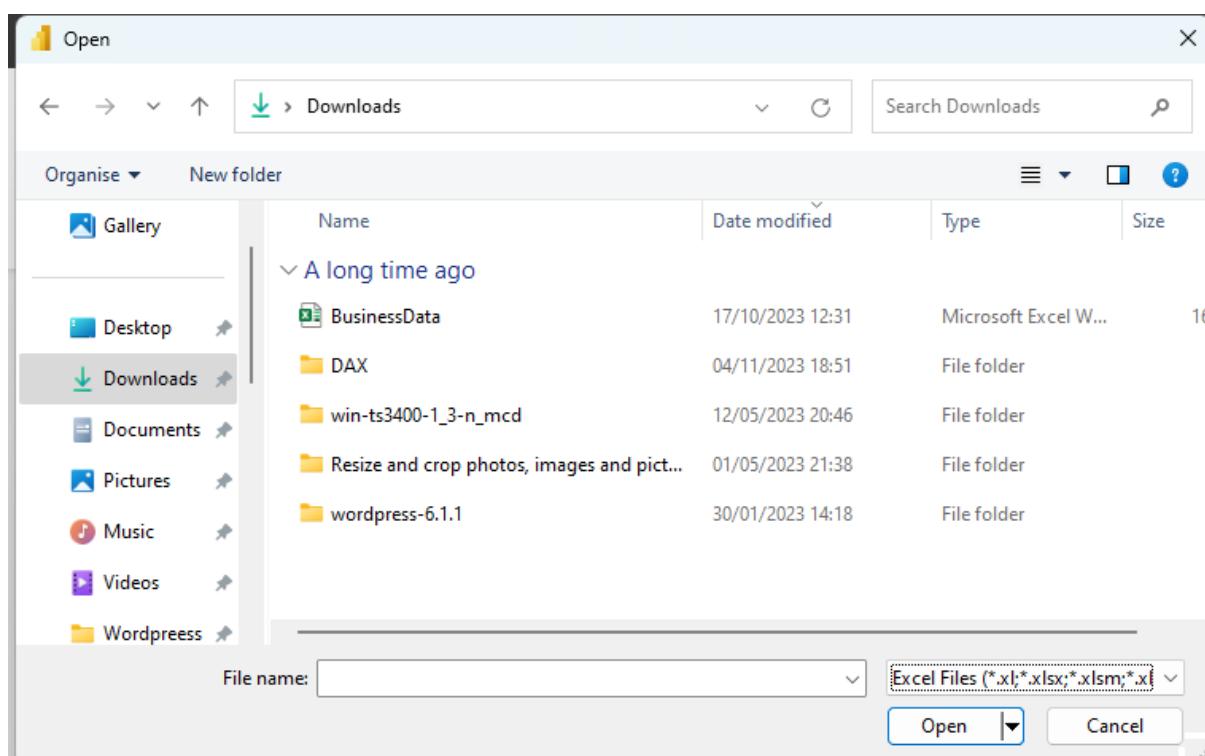
Get data from another source →

- Import data from Excel
- Import data from SQL Server
- Paste data into a blank table
- Try a sample semantic model

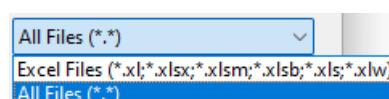
Open Power BI first on device. Then click on the Get data.



Now, for the dataset that is being used, the files are in excel format so from the drop-down list, Excel workbook option was selected.



Library will open, navigate to where the file is located.



One keynote is that most of the times user may not find the file that is because the file extension is wrong on the bottom right. Select All files extension.

Get the data by clicking open, after selecting your data.

The screenshot shows the 'Data Load' dialog in Power BI. At the top, there are three tabs: 'File Origin' (set to '65001: Unicode (UTF-8)'), 'Delimiter' (set to 'Comma'), and 'Data Type Detection' (set to 'Based on first 200 rows'). Below these are the first 200 rows of the 'HR_Analytics.csv' file. At the bottom right are four buttons: 'Extract Table Using Examples', 'Load' (highlighted with a black arrow), 'Transform Data', and 'Cancel'.

Before the dataset is loaded into Power BI, it be processed for cleaning by clicking on the **Transform Data** option next to the Load in the bottom right.

However, user can still transform the data after they have loaded it.

4.2) Data Cleaning

Soon as Transform Data is clicked, **Power Query Editor** will open.

The screenshot shows the Power Query Editor interface. The top ribbon has tabs like File, Home, Transform, Add Column, View, Tools, and Help. The main area shows a table with 29 rows of data from 'HR_Analytics.csv'. The 'APPLIED STEPS' pane on the right lists 'Promoted Headers' and 'Changed Type'. The 'Properties' pane on the far right shows the query name is 'HR_Analytics'. The status bar at the bottom says '29 rows'.

Now, the data is ready to be cleaned.

Removing Nulls:

To check if there are any Nulls or duplicates, click on **View** from top banner, then select **Column quality**.

There is only one Column that has Nulls.

Nulls can be removed in two ways.

Method 1 & 2

1. Simply click on the Remove Empty, this will remove all the null values.

2. Unselect null, this will remove the nulls.

1. Simply click on the Remove Empty, this will remove all the null values.
2. Unselect null, this will remove the nulls.

Method 3

Sort the data in the ascending order, this will push all the nulls to the top rows.

1 ² ₃ YearsSinceLastPromotion	1 ² ₃ YearsWithCurrManager
● Valid 100%	● Valid 94%
● Error 0%	● Error 0%
● Empty 0%	● Empty 6%
0	null
0	null
1	null
0	null
1	null
6	null
2	null
0	null
0	null

Then, go to Home>Remove Rows>Remove Top Rows.

Now, select up to where your rows are Null. And Remove Top rows that contains all the Nulls.

Remove Top Rows

Specify how many rows to remove from the top.

Number of rows

57

OK

Cancel

56	1	17	3	3	4	2	0	null
57	1	17	2	2	3	1	0	null
58	0	12	4	2	0	0	0	0

Applied Steps

It can be seen the applied steps in the right-side, under Query Settings

Removing Duplicates:

Before removing duplicates, we shall know that we can also identify duplicates. Select the desired column(s) containing potential duplicates. Navigate to **Group By** in **Home** tab.

Group By

Specify the column to group by and the desired output.

Basic Advanced

EmpID

New column name

Count

Operation

Count Rows

Column

OK

Cancel

This will count all the rows in the column.

A _B _C EmpID	I ₂ ₃ Count
Valid	100%
Error	0%
Empty	0%
RM297	1
RM302	1
RM458	1
RM728	1
RM829	1
RM973	1
RM1154	1
RM1312	1
RM128	1
RM150	1
RM172	1
RM178	1
RM423	1

Now, duplicates need to be identified. Click on Count drop-down>Number Filters>Does Not Equal.

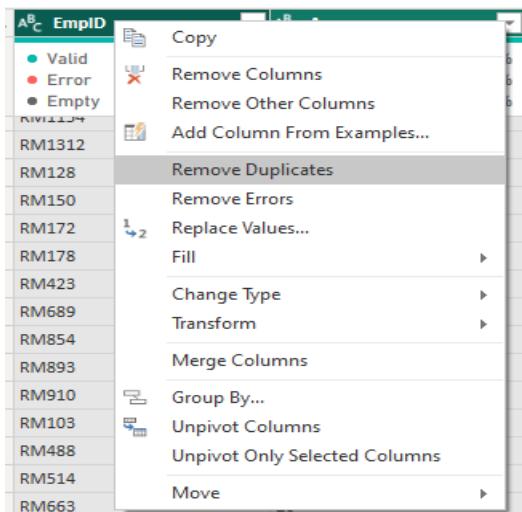
The screenshot shows the Power BI ribbon with the 'Number Filters' dropdown open. The 'Does Not Equal...' option is highlighted in grey, indicating it is selected. Other filter options like 'Equals...', 'Greater Than...', and 'Less Than...' are also visible in the list.

Now, put the value of Does not equal to 1, so it will give all the repeated data.

A _B _C EmpID	I ₂ ₃ Count
Valid	100%
Error	0%
Empty	0%
1	2
2	2
3	2
4	2
5	2
6	2
7	2
8	2
9	2
10	2

However, one cannot simply go about deleting the duplicates like that. Since a same Employee could have been recorded with respect to different attributes. This will only cause redundancy.

A _B _C EmpID	A _B _C Age	A _B _C AgeGroup	A _B _C Attrition	A _B _C BusinessTravel	A _B _C DailyRate	A _B _C Department	A _B _C DistanceFromHome
Valid	100%	Valid	100%	Valid	100%	Valid	100%
Error	0%	Error	0%	Error	0%	Error	0%
Empty	0%	Empty	0%	Empty	0%	Empty	0%



Select all the columns, then right click on the first Column, and select, Remove Duplicates option to remove all the duplicates.

Spelling Error/ Find and Replace

Sometimes there can be same data with different spellings, this leads to giving wrong insights. Although it is not always possible to find and replace such errors, but it is fairly important.

There is a such error in **BusinessTravel** Column.

The image consists of three screenshots labeled 1, 2, and 3. Screenshot 1 shows the 'BusinessTravel' column's context menu with 'Replace Values...' selected. Screenshot 2 shows the 'Replace Values' dialog box with 'TravelRarely' in the 'Value To Find' field and 'Travel_Rarely' in the 'Replace With' field. Screenshot 3 shows the Power BI interface with the 'BusinessTravel' column now correctly displaying 'Travel_Rarely' values.

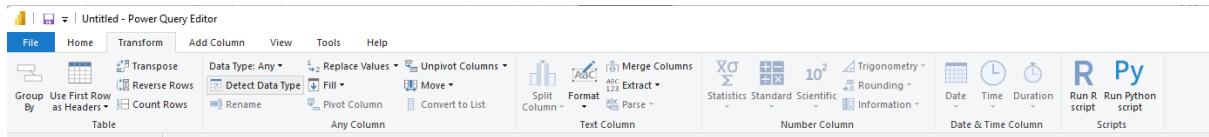
- Both **TravelRarely** and **Travel_Rarely** are same value but written differently. Unfortunately, Power BI cannot identify such mistakes.
- Right-click on one of them, then select **Replace Values...** option.
- Replace the value and click **OK**.

Data Types

Sometimes Power BI cannot detect the data type. As it can be seen, below all the columns are allocated as **text**.

EmpID	Age	AgeGroup	Attrition	BusinessTravel	DailyRate	Department	DistanceFromHome
RM297	18	18-25	Yes	Travel_Rarely	230	Research & Development	3
RM302	18	18-25	No	Travel_Rarely	812	Sales	10
RM458	18	18-25	Yes	Travel_Frequently	1306	Sales	5
RM778	18	18-25	No	Non-Travel	287	Research & Development	5

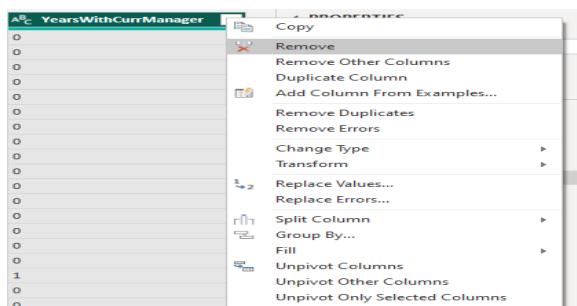
To solve this problem, Select **Detect Data Type** in the **Transform** tab.



The data is allocated right data type now.

Table.TransformColumnTypes(#"Replaced Value",{{"EmpID", type text}, {"Age", Int64.Type}, {"AgeGroup", type text}, {"Attrition", type text}, {"BusinessTravel", type text}, {"DailyRate", type number}, {"Department", type text}, {"DistanceFromHome", type number}, {"EmployeeCount", type number}, {"OverTime", type text}, {"OverTimeReason", type text}, {"TotalWorkingYears", type number}, {"YearsAtCompany", type number}, {"YearsWithCurrManager", type number}})								
	EmpID	Age	AgeGroup	Attrition	BusinessTravel	DailyRate	Department	DistanceFromHome
1	RM297	18	18-25	Yes	Travel_Rarely	230	Research & Development	
2	RM302	18	18-25	No	Travel_Rarely	812	Sales	
3	RM458	18	18-25	Yes	Travel_Frequently	1306	Sales	
4	RM728	18	18-25	No	Non-Travel	287	Research & Development	

Removing Columns



Not all the columns will be used in the dashboard.

Right- click on the column, then select **Remove** to remove that column.

Adding Columns

➤ Column 1

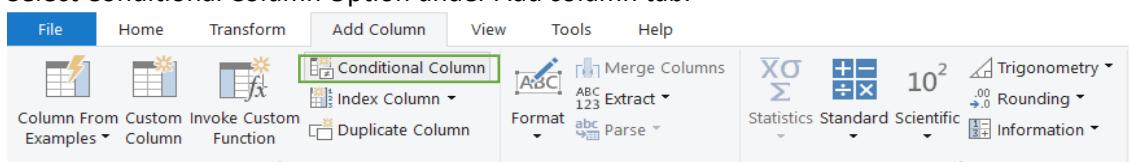
Attrition
No
Yes
Yes
No
No
Yes

Attrition is a Boolean data type, which is not useful in this visualization if not modified.

The numeric value of it will be computed from it, so calculations can be done to get a better insight in report.

To do so, a column will be added.

1. Go to **Power Query** by clicking on **Transform Data** under **Home** tab.
2. Select **Conditional Column** Option under **Add column** tab.



3. Adding a Conditional Column called **AttritionNum** that is computed from **Attrition** Column.

Add Conditional Column

Add a conditional column that is computed from the other columns or values.

New column name

AttritionNum

Column Name	Operator	Value	Output
If Attrition	equals	ABC 123	ABC 123 1
Else ABC 123 0			
<input type="button" value="Add Clause"/> <input type="button" value="OK"/> <input type="button" value="Cancel"/>			

This column will change all the values with **Yes** into **1** and **No** into **0**.

Click **OK**.

4. New Column is added.

The screenshot shows the Power Query Editor interface. On the left is a table preview with one column labeled 'AttritionNum'. The values in this column are binary: 1, 0, 1, 0, 1, 0, 1, 0, 1, 0. To the right of the table is a vertical pane titled 'APPLIED STEPS' which lists the following steps in sequence: Source, Promoted Headers, Changed Type, Removed Columns, Removed Duplicates, Replaced Value, Changed Type1, Removed Columns1, Reordered Columns, and finally 'Added Conditional Column' at the bottom.

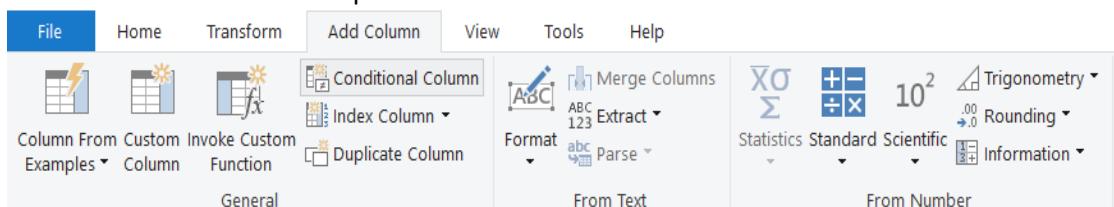
Note: Remember to update the Data Type.

➤ **Column 2**

The screenshot shows the Power Query Editor interface. On the left is a table preview with one column labeled 'DistanceFromHome'. The values in this column are numeric: 3, 9, 25, 4, 11, 12, 8. To the right of the table is a block of text that reads: "Here it can be seen that **DistanceFrom Home** is a numeric data type, which only gives distance. Our goal is to specify how much of a distance is close, near and far. To do so, a column will be added."

1. Go to **Power Query** by clicking on **Transform Data** under **Home** tab.

2. Select **Conditional Column** Option under **Add column** tab.



3. Adding a Conditional Column called **DistanceStatus** which is computed from **DistanceFromHome** column.

The screenshot shows the 'Add Conditional Column' dialog box. It has a header 'Add Conditional Column' and a sub-instruction 'Add a conditional column that is computed from the other columns or values.' Below this, there is a 'New column name' input field containing 'DistanceStatus'. The main area contains two conditional clauses:

- If **DistanceFromHome** is greater than or... Then **Far**
- Else If **DistanceFromHome** is greater than or... Then **Near**

At the bottom, there is a button 'Add Clause' and a row for 'Else' with the value 'Very Close'. At the very bottom are 'OK' and 'Cancel' buttons.

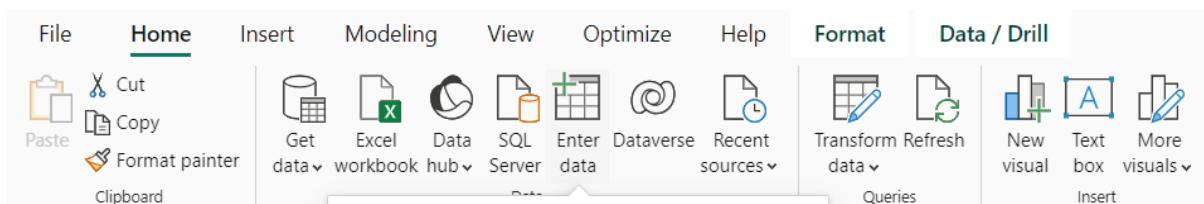
Here, using the values from the column, data is retrieved that gives categorical output.

4. New Column is added.

Note: Remember to update the Data Type.

Adding a Table

Adding a new table for all the measure that will be created.



Click on **Enter data** under **Home** tab. Give it a name called **New Measures**.

New Measure

It will be used quite often. To begin with, lets create a measure to count total Employees. Simple right-click on the new table that we created before called **New Measure**. All the measures will be saved there for convenience.

When Right-clicking, New Measures insight will open up.

```
1 TotalEmp = COUNTROWS('HR_Analytics')
```

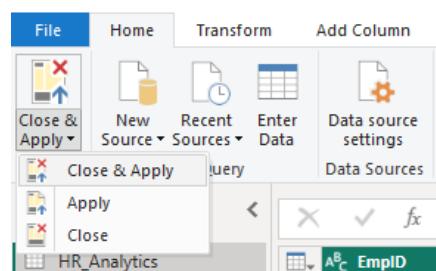
New Measures

Column1

TotalEmp

Put the formula, then it will create a new measure.

Always remember to click **Close & Apply** to save your changes.



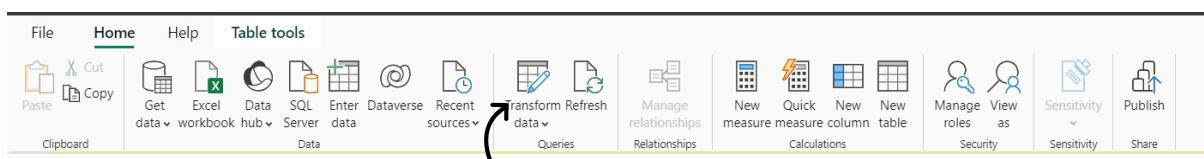
5) Data Modelling

Below, it is evident that the dataset encompasses a comprehensive set of columns. In establishing a data model for the HR Analytics project, the approach will likely involve restructuring this flat table

into a more organized model, such as a star schema, to enhance analytical efficiency.

The screenshot shows the Power BI Data view. On the left, there is a list of tables under the 'HR_Analytics' folder, including Age, AgeGroup, Attrition, BusinessTravel, DailyRate, Department, DistanceFromHome, Education, EducationField, EmpID, EmployeeCount, EmployeeNumber, EnvironmentSatisfaction, Gender, HourlyRate, JobInvolvement, JobLevel, JobRole, JobSatisfaction, MaritalStatus, MonthlyIncome, MonthlyRate, NumCompaniesWorked, Over18, Overtime, PercentSalaryHike, PerformanceRating, RelationshipSatisfaction, SalaryGlob, StandardHours, and WorkLifeBalance. On the right, the properties pane for the 'HR_Analytics' table is open, showing fields like Name (HR_Analytics), Description (Enter a description), and various columns listed under the 'Data' tab.

There are a lot of data redundancy that needs to be removed. To process data, go to **Power Query** from the top ribbon **Transform Data**.



The first table that will be created is the Combined table. This table will store the Age, Gender, Department and so on.

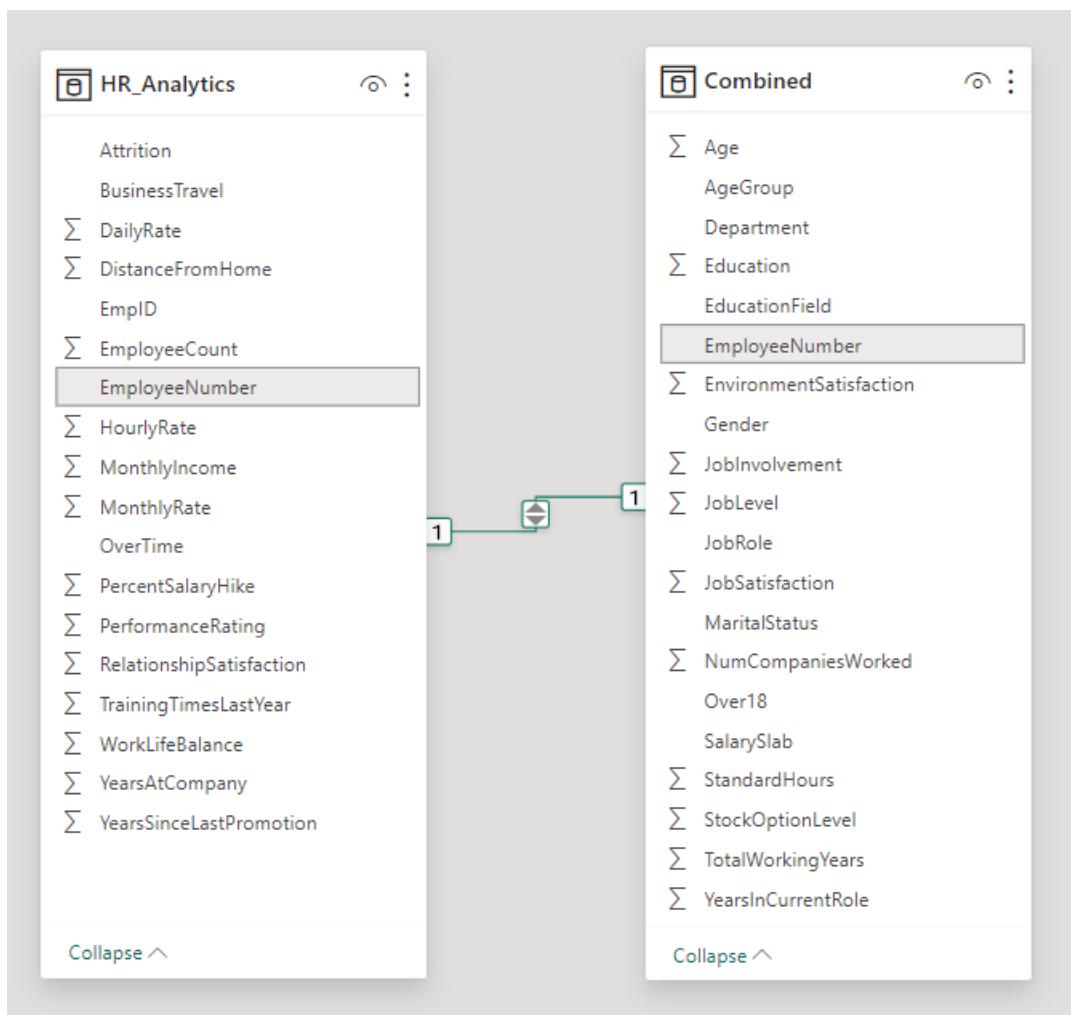
The screenshot shows the Power BI Power Query Editor. A red arrow points to the 'Transform' tab at the top. Below it, the 'HR_Analytics' table is displayed as a grid. The columns are EmpID, Age, AgeGroup, Attrition, BusinessTravel, DailyRate, Department, and DistanceFromHome. The data consists of 18 rows of employee information.

EmpID	Age	AgeGroup	Attrition	BusinessTravel	DailyRate	Department
RM297	18-25	Yes	Travel_Rarely	230	Research & Development	
RM302	18-25	No	Travel_Rarely	812	Sales	
RM458	18-25	Yes	Travel_Frequently	1306	Sales	
RM728	18-25	No	Non-Travel	287	Research & Development	
RM829	18-25	Yes	Non-Travel	247	Research & Development	
RM973	18-25	No	Non-Travel	1124	Research & Development	
RM1154	18-25	Yes	Travel_Frequently	544	Sales	
RM1312	18-25	No	Non-Travel	1431	Research & Development	
RM128	18-25	Yes	Travel_Rarely	528	Sales	
RM150	18-25	No	Travel_Rarely	1181	Research & Development	
RM172	19-25	Yes	Travel_Frequently	602	Sales	
RM178	19-25	Yes	Travel_Rarely	303	Research & Development	
RM423	19-25	Yes	Travel_Rarely	489	Human Resources	
RM689	19-25	Yes	Travel_Rarely	419	Sales	
RM854	19-25	No	Travel_Rarely	645	Research & Development	
RM893	19-25	Yes	Non-Travel	504	Research & Development	
RM910	19-25	No	Travel_Rarely	265	Research & Development	
RM103	20-25	Yes	Travel_Frequently	872	Research & Development	

- Right-Click on **HR_Analytics** table on the left and select **Duplicate**.
- A new table will appear called **HR_Analytics(2)**
- Rename this table as **Combined** by right-clicking on it and selecting **Rename**.
- Now remove all the columns that is not useful. To do so, click on the header of the any column that you want to keep, then press Ctrl, and click on all the columns that you need.
- Once the three columns are selected, right click on any of the headers of the three columns and select **Remove Other Columns**.
- The final result is shown below.

The screenshot shows the Power BI Data Editor interface. A query named "Combined" is open, displaying a table with the following columns: EmployeeNumber, Age, AgeGroup, Department, Education, EducationField, EnvironmentSatisfaction, and Gender. The "Reordered Columns" step is selected in the "APPLIED STEPS" section of the Query Settings pane.

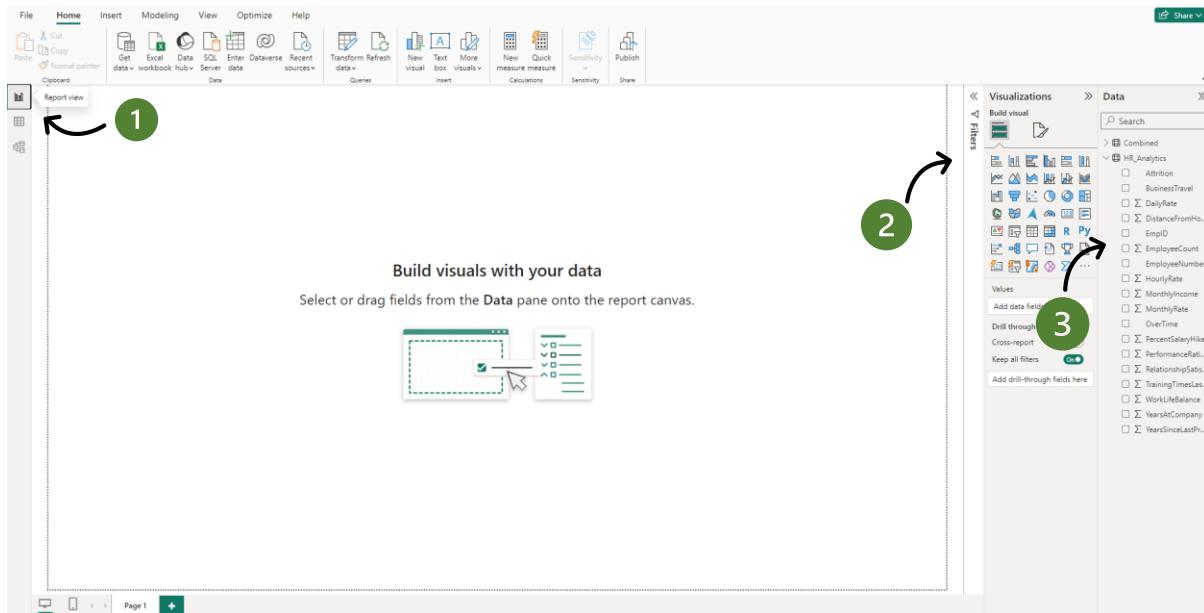
Now go to **HR_Analytics** and keep the required columns. This should have created two table that is organized.



EmployeeNumber is a primary key that creates a relationship between two table.

6) Data Visualization

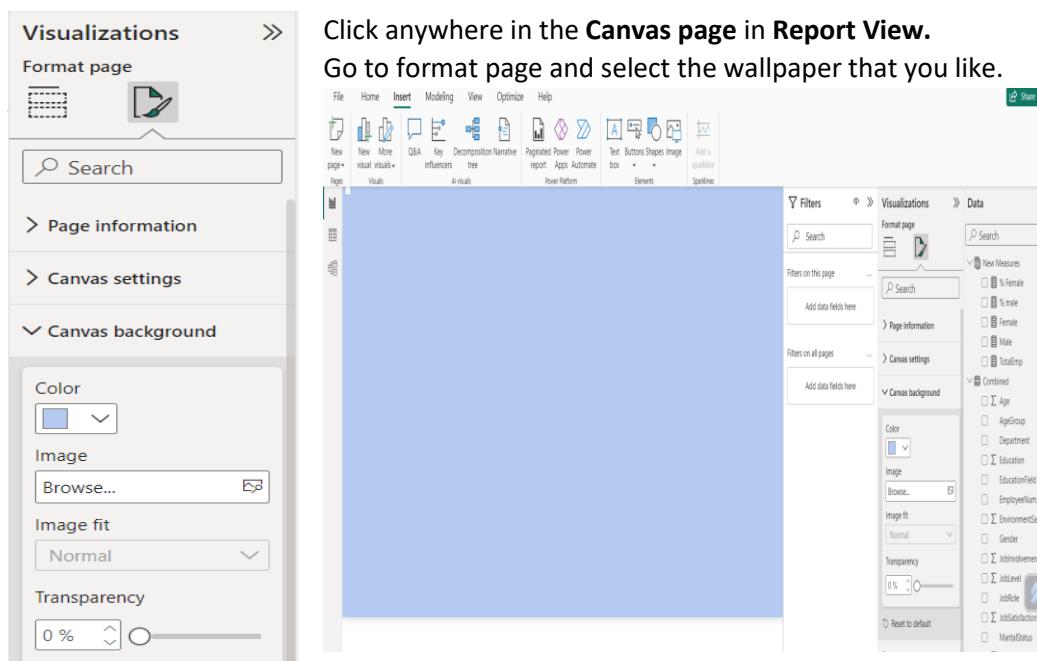
Now the data is good to go. To begin with visualization, go to the Report View.



1. Report View.
2. Visualization pane.
3. Fields and columns.

Background Poster

To keep the dashboard visually appealing, it is important to have a nice background poster. In this dashboard, Alphabet Blue colour background is used.



Page Title

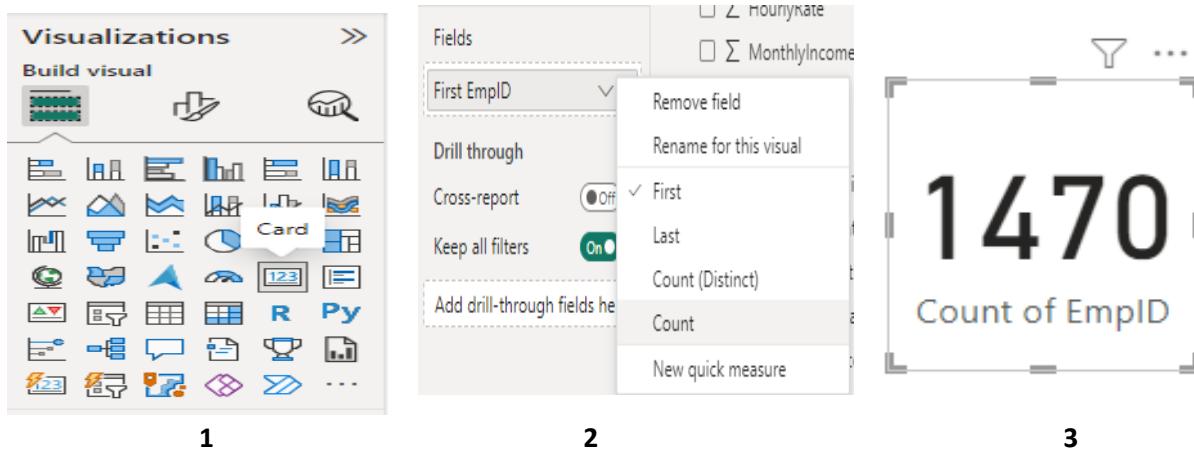
HR ANALYTICS DASHBOARD

Arial Black 24 **I** U

KPI 1:

Total count of employees

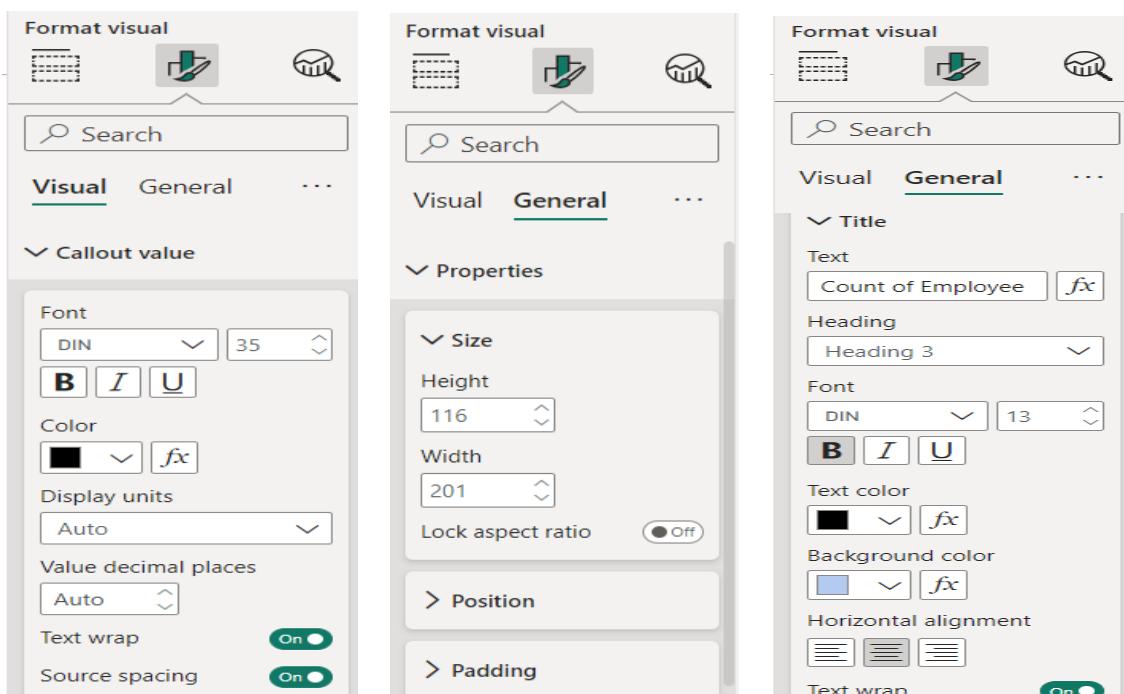
Here, a Card visual is used to find the Count of all the employees using EmpID and an aggregation.



1. Select the Card visual.
2. Drop **EmpID** from Data into Fields or simply drop it into the chart. Then, click drop down menu to get in-built **Count** filter, this will count all the EmpID for us.

Customization and Formatting:

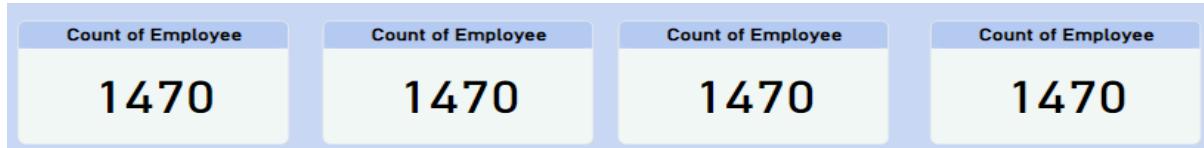
For better visualization and implementation, **Format visual** tool comes in handy.



The final visual:



A convenient aspect of Power BI is that there is no need to repeatedly design and implement the same graphs. Users can effortlessly copy and paste, subsequently adjusting the fields as needed, streamlining the graph creation process.

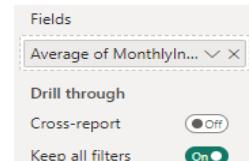


KPI 2:

Average Monthly Income



One of the copied cards visual from **KPI 1** is used, the formatting remains same except for the field and the title.

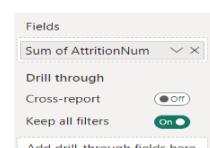


KPI 3:

Total count of Attritions



Card visual from **KPI 1** is copied and pasted, the formatting remains same except for the field and the title.

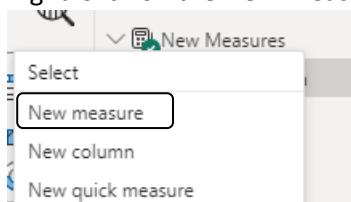


KPI 4:

Attrition Rate

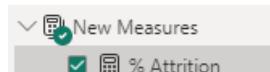
To calculate the percentage of attrition a **New Measure** will be used.

1. Right-Click on the New Measure table that was created earlier and select New Measure



2. Create a measure using the formula below.

```
1 % Attrition = SUM(HR_Analytics[AttritionNum])/SUM(HR_Analytics[EmployeeCount])
```



3. The final result is:

The image shows two side-by-side 'Format visual' dialog boxes for a 'Card' type visualization.

Left Card Visual:

- Title:** Attrition Rate
- Value:** 16.1%
- Fields:** % Attrition
- Drill through:** Off
- Cross-report:** Off

Right Format Visual:

- Title:** Attrition Rate
- Text:** Attrition Rate
- Font:** DIN, 20pt
- Color:** Black
- Display units:** Auto
- Value decimal places:** Auto
- Text wrap:** On
- Source spacing:** On

KPI 5:

Gender based information of employees

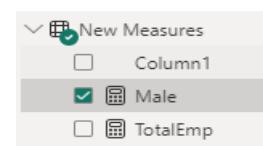
To get this insight, a **New Measure** will be created for each Male and Female

New measure for male:

```
1 Male = Calculate([TotalEmp], 'Combined'[Gender] = "male")
```

Using the formula above, a new measure for male is created.

To get the percentage: `1 % male = DIVIDE([Male], [TotalEmp], 0)`



Make sure to change the format and display on the top.



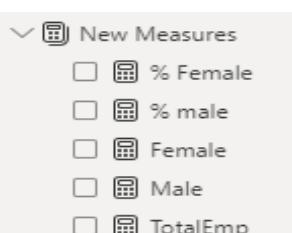
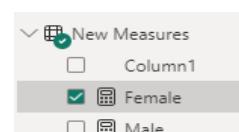
New measure for female:

```
1 Female = Calculate([TotalEmp], 'Combined'[Gender] = "female")
```

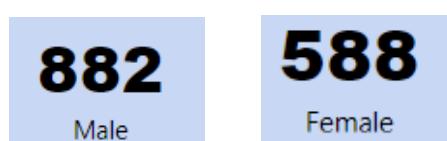
Using the formula above, a new measure for female is created.

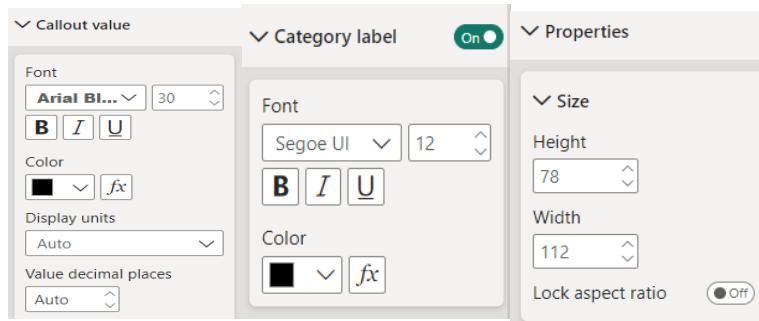
To get the percentage: `1 % Female = DIVIDE([Female], [TotalEmp], 0)`

Make sure to change the format and display on the top.



Now, they are ready for visualization. Both the count and percentage will be used in the Card visual.





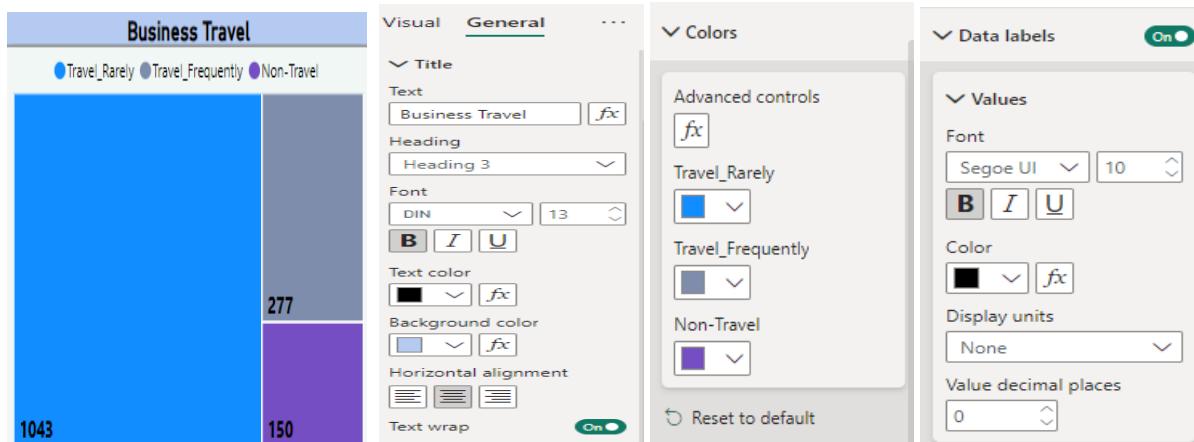
The final result is:



KPI 6:

Employees willing to business travel

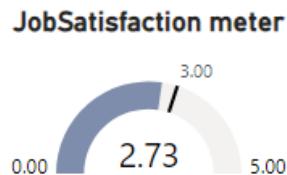
The **Treemap** will be used to count of employees willing to go for a business trip. There are three categories of employees willing to travel: Travel_Rarely, Travel_Frequently and Non-Travel.



KPI 7:

Average job satisfaction of employees

To get the average satisfaction rating of employees towards their job, **Gauge** will be used. The rating scale ranges from 0 to 5, with a predefined target set at 3.



The average rating is 2.73, which does not meet the organization's target.

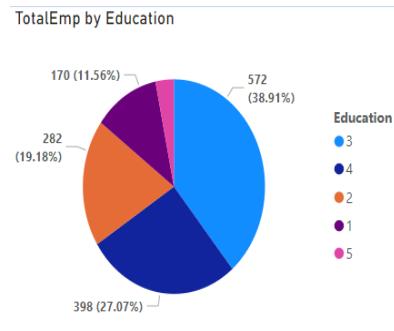
Format of the visual:

The screenshot shows the 'Format visual' interface for a gauge axis. It includes sections for 'Gauge axis', 'Colors', 'Data labels', 'Target label', and 'Callout value'. Each section contains various settings like font, color, and display units.

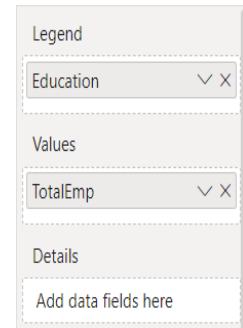
KPI 8:

Total Employees by education

Visualize the distribution of employees based on their education levels to gain insights into the educational background of the workforce.



The **Pie Chart** visualizing "Total Employees by Education" provides a quick and intuitive representation of the educational background of the workforce. This insight can be valuable for HR analytics and workforce planning. It can be seen that the greatest number of employees have third categorical education.

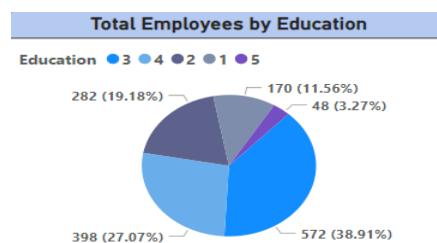


Customization and Formatting

For visually appealing and better insight, **Format visual** tool will be used.

The screenshot shows the 'Format visual' interface for a pie chart. It includes sections for 'Legend', 'Slices', and 'Detail labels'. The 'Detail labels' section is expanded, showing 'Options' (Position set to 'Outside') and 'Values' (Font set to 'Segoe UI' size 9, bold, italic, underline).

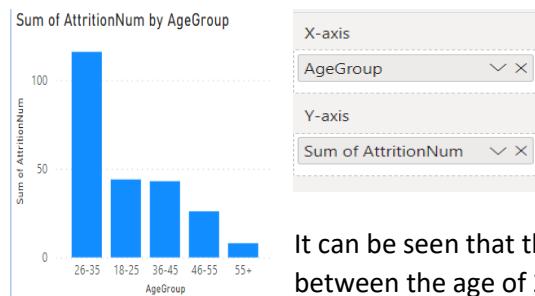
The final result is:



KPI 9:

Sum of attrition by age-group

Analysing the distribution of attrition across different age groups to identify any age-related patterns in employee turnover.

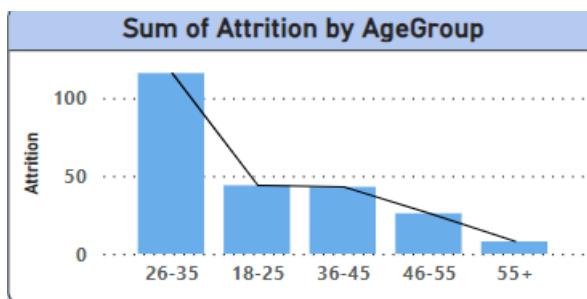


The **Bar Chart** visualizing the "Sum of Attrition by Age Group" provides insights into the age-related aspects of employee turnover. This analysis can contribute to understanding attrition patterns and informing HR strategies.

It can be seen that the greatest number of employees who leave are between the age of 26-35.

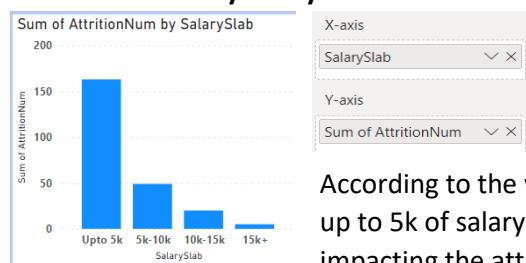
Customization and Formatting

The final result is:



KPI 10:

Sum of attrition by salary slab



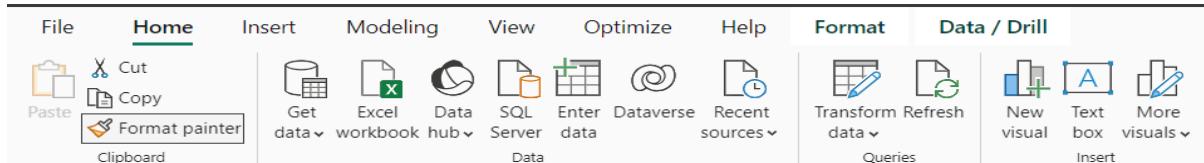
Analysing the distribution of attrition across different salary slabs to identify any patterns related to compensation levels and employee turnover.

According to the visualization, most of the employees that have left had up to 5k of salary. This shows how salary is one of the main factors impacting the attrition count.

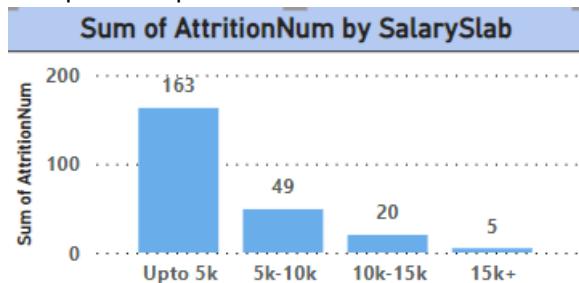
Customization and Formatting

In Power BI, users can streamline the formatting process by utilizing the **Format Painter** feature. Instead of manually formatting charts each time, users can copy the format from one chart and

easily apply it to others. This not only saves time but also ensures consistency in theme and format across different visualizations (Vivran, 2020).



The visual from **KPI 9** will be copied and pasted. The final result is:



KPI 11:

Count of attrition by job role

Analysing the count of attrition cases for different job roles to identify patterns and insights related to employee turnover across various positions.

A **matrix** table is used to plot the data.

Attrition by job role

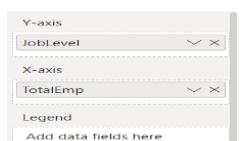
JobRole	No	Yes	Total
Sales Representative	50	33	83
Sales Executive	269	57	326
Research Scientist	245	47	292
Research Director	78	2	80
Manufacturing Director	135	10	145
Manager	97	5	102
Laboratory Technician	197	62	259
Total	1233	237	1470

This analysis can contribute to understanding the impact of job roles on employee turnover and inform targeted retention strategies. It can be seen that the greatest number of Laboratory Technicians are leaving jobs, followed by sales executives.

KPI 12:

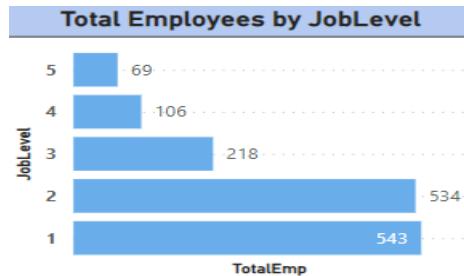
Total employee by job level

Analysing the distribution of total employees across different job levels to gain insights into the organizational structure and hierarchy.



A **clustered bar chart** is used for visualizing these data.

This analysis can inform workforce planning and organizational structure decisions.



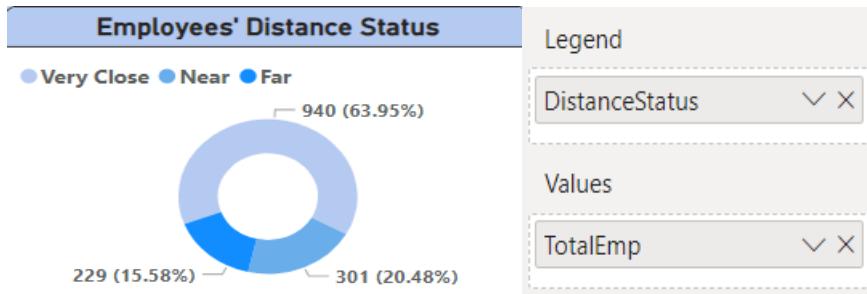
KPI 13:

Employees' distance status from home to work

Analysing the distribution of total employees based on their distance status from home to work to understand commuting patterns within the organization.

The column 2 that was made earlier to compute the numeric value of distance into specified status like: very close, near and far, a **Donut chart** will be made.

HR_Analytics
DistanceStatus



The chart shows that most of the employees live very close to their workplace while there are still some staffs who travel far to get to their workplace.

KPI 14:

Overtime by marital status

Analysing the distribution of overtime work among employees based on their marital status to understand if there are any patterns related to work hours and personal relationships.

A **Donut chart** will be used to visualize, while the formatting is copied from previous chart.



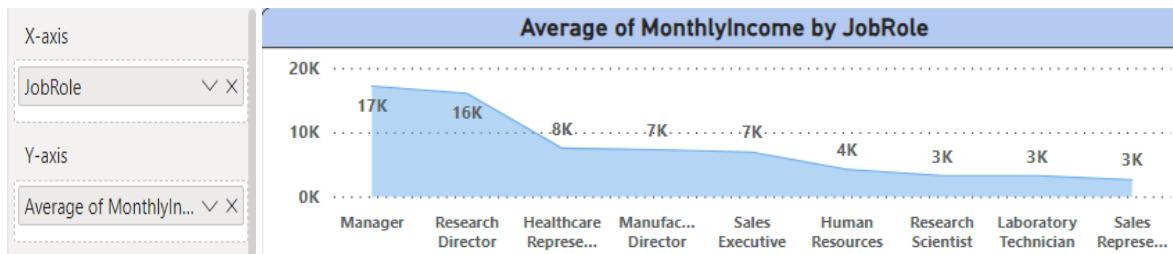
According to the chart, married people do more over-time work as compared to singles and divorced.

KPI 15:

Average monthly income by job role

Analysing the average monthly income across different job roles to understand the income distribution within the organization.

An area chart will be used to plot the insight, while the formatting is copied from previous bar charts.



It can be seen that managers have the highest monthly income of 17k followed by Research directors, whereas Sales representatives have the lowest income of 3k.

Slicer

A slicer will be incorporated into the dashboard to obtain precise and detailed insights based on the three distinct departments. The Department column will be plotted in the fields section for this purpose.



7) The Final Dashboard



8) Conclusions and Recommendations

In conclusion, the comprehensive analysis conducted through the HR Analytics Dashboard project yields valuable insights into critical facets of workforce dynamics. The refined dataset, carefully curated for relevant Key Performance Indicators (KPIs), has provided a nuanced understanding of factors influencing employee performance and attrition.

Summary of Findings:

The findings indicate noteworthy patterns in attrition rates, salary distributions, and commuting distances. Employees with lower salaries exhibit a higher attrition rate, necessitating a closer examination of compensation structures. Furthermore, the Average Monthly Income by Job Role KPI highlights significant variations, with managers earning the highest and sales representatives earning comparatively lower incomes.

Recommendations:

Based on these insights, recommendations are formulated to enhance talent retention and workforce satisfaction. Strategies addressing salary adjustments, targeted employee engagement initiatives, and flexible commuting arrangements are proposed. Moreover, fostering a culture of continuous feedback and professional development emerges as a key recommendation to boost overall job satisfaction and performance.

Personal Conclusions:

On a personal note, engaging in this project has underscored the pivotal role of data-driven decision-making in human resource management. The ability to translate raw data into actionable intelligence empowers organizations to proactively address challenges and optimize workforce strategies. This project reaffirms the significance of an integrated HR Analytics Dashboard in steering organizational success through informed and strategic human resource practices.

9) Reference

Fares Sayah (2022) Kaggle. IBM HR Analytics Employee Attrition & Performance. Available at: <https://www.kaggle.com/code/faressayah/ibm-hr-analytics-employee-attrition-performance> (Accessed on: 12/12/23)

Vivran (14/07/2020) 3 Useful Features you should know — Power BI Desktop. Available at: <https://medium.com/analytics-vidhya/3-useful-features-you-should-know-power-bi-desktop-292a9bc044> (Accessed on: 11/01/2024)