SMARTBRIDGE GUIDED PROJECT REPORT

(DATA ANALYTICS)

(JAIVIK CHAUHAN - 19BIT0039 – VIT VELLORE)

Employee Attrition Analytics using IBM Cognos

INTRODUCTION

Employee attrition is defined as the natural process by which employees leave the workforce – for example, through resignation for personal reasons or retirement – and are not immediately replaced. The key to success in any organization is attracting and retaining top talent. Attrition is an inevitable part of any business. There will come a time when an employee wants to leave your company – for either personal or professional reasons. Attriton is a huge concern of every organization . By knowing the proper attrition of employee , an organization can save monay and effort both by not investing in them . As the work dependency and other factors of leaving employees can lead organization to a huge loss.
Here, we will be using IBM Cognos software to do data analytics about employee attrition data and we prepare our different visualizations to perform data exploration to it . We are importing dataset from kaggle. Our aim to perform the identification of important factors that might be influential in determining which employee might leave the firm and who may not.

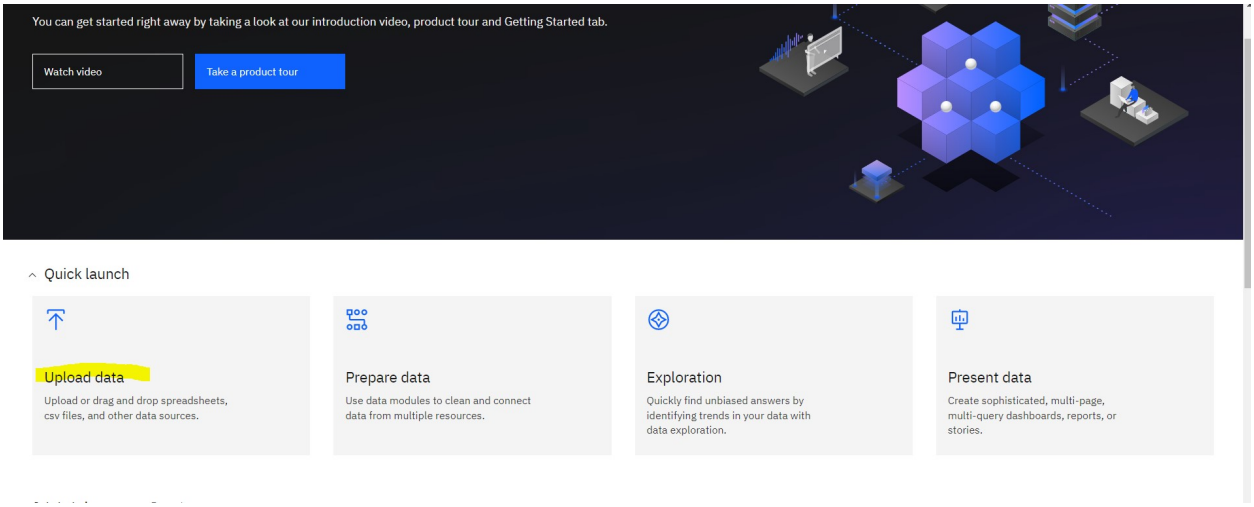
DATASET

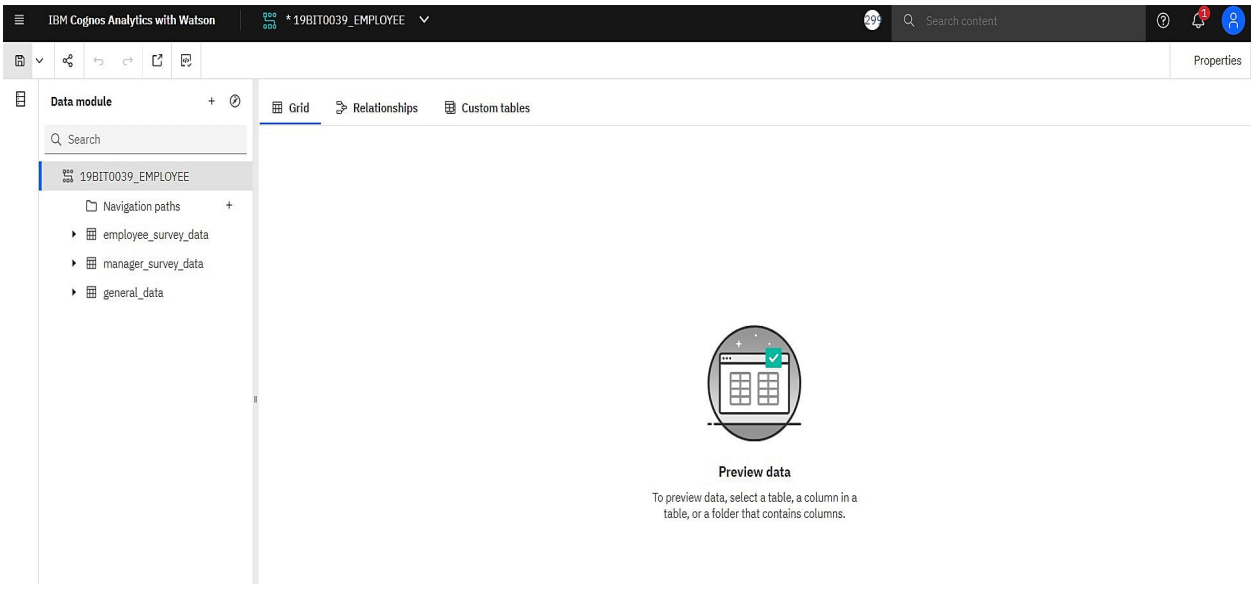
The data can be downloaded from KAGGLE dataset (https://www.kaggle.com/datasets/vjchoudhary7/hr-analytics-case-study).
We need to use only (3 files - General_data.csv, Employee_Survey_Data.csv, Manager_Survey_data.csv) for the current project.
Through this dataset, we are going to try to understand the factors to keep employees at the Company and which prompt others to leave.

IMPLEMENTATION

1. LOADING DATASET IN COGNOS:

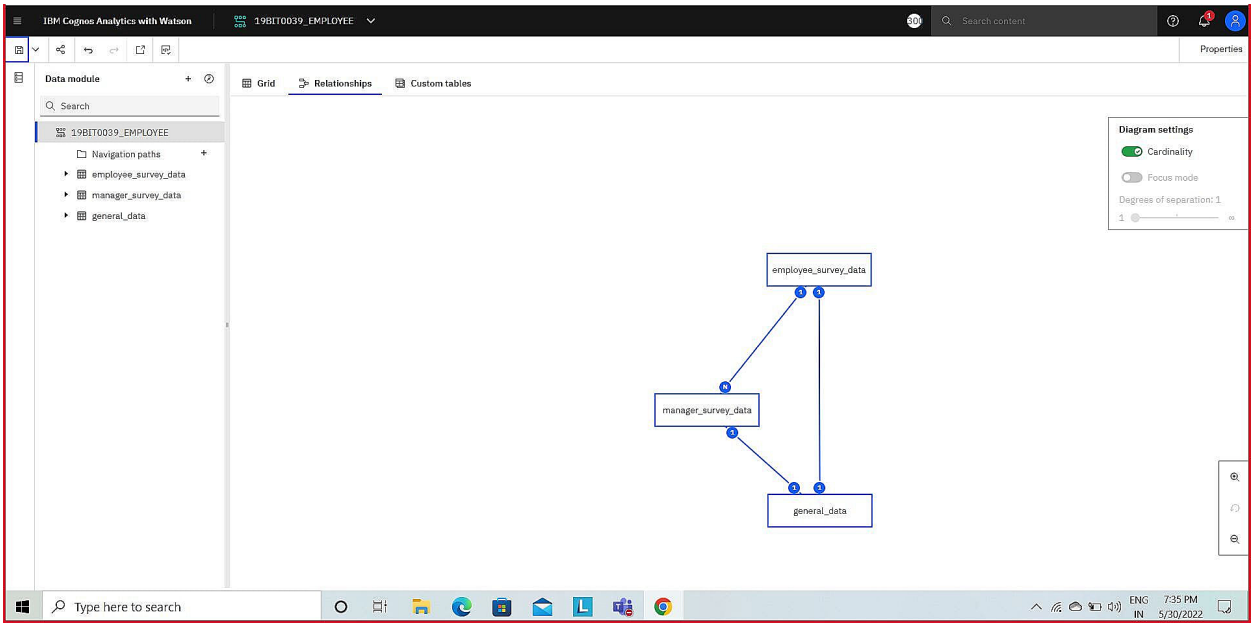
1. DOWNLOADING IT FROM Kaggle.com.
2. Choosing uploading option in cognos and selecting 3 files - General_data.csv, Employee_Survey_Data.csv, Manager_Survey_data.csv for the current project.



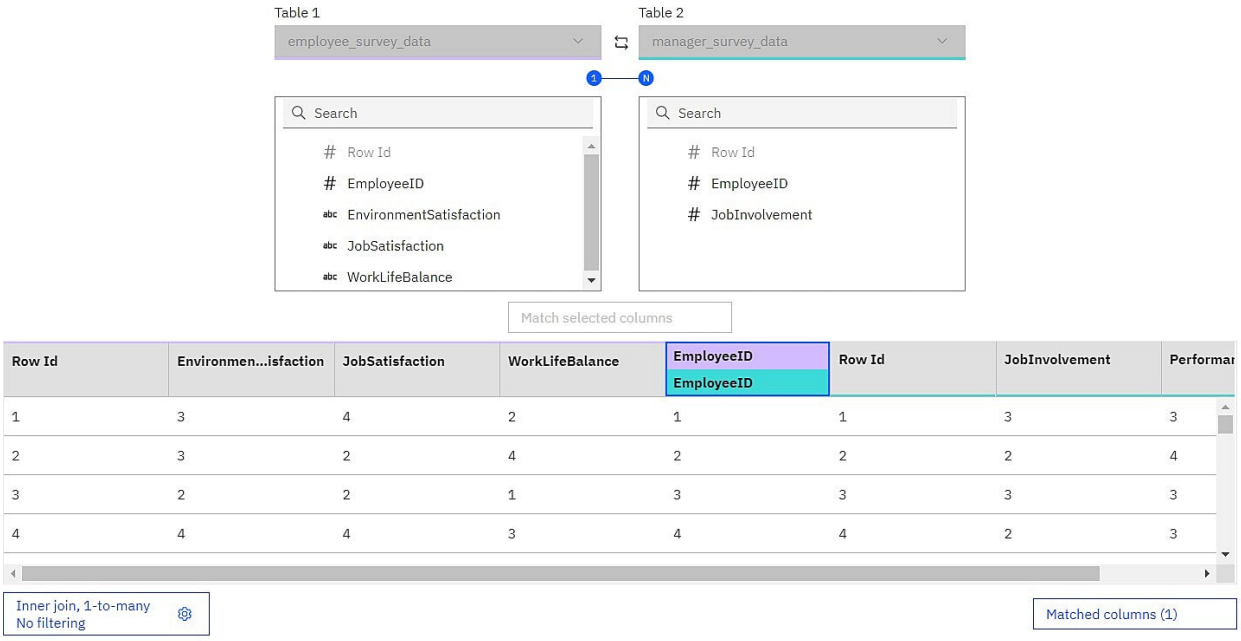


2. PREPARING DATASET:

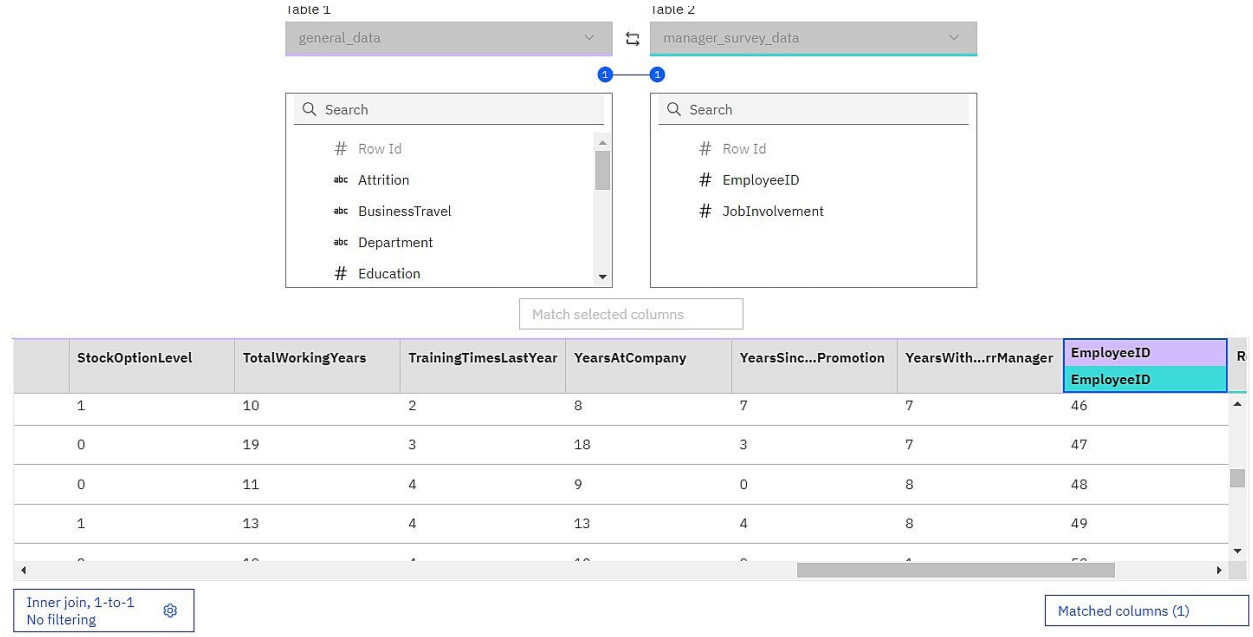
The preparation of the data is done by creating relationships between the general, manager_survey, employee_survey.



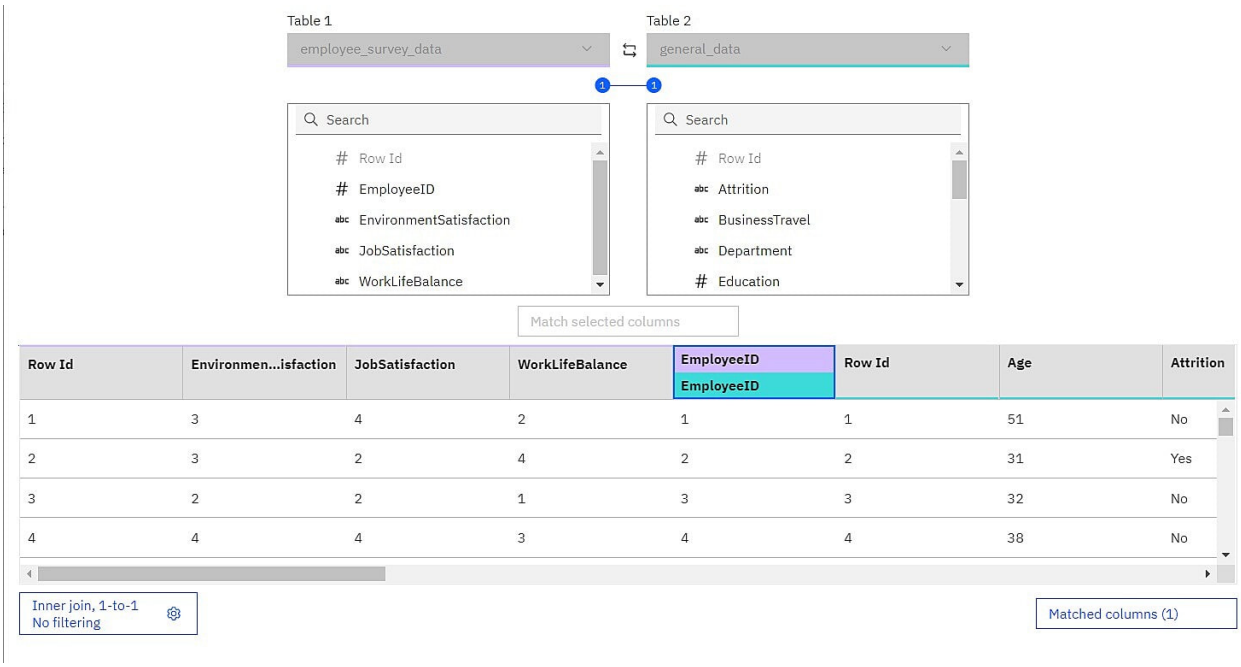
There is a 1 to many relationship between employee and manager data with the common field taken again as the EmployeeID.



There is a 1:1 relationship between general and manager data.



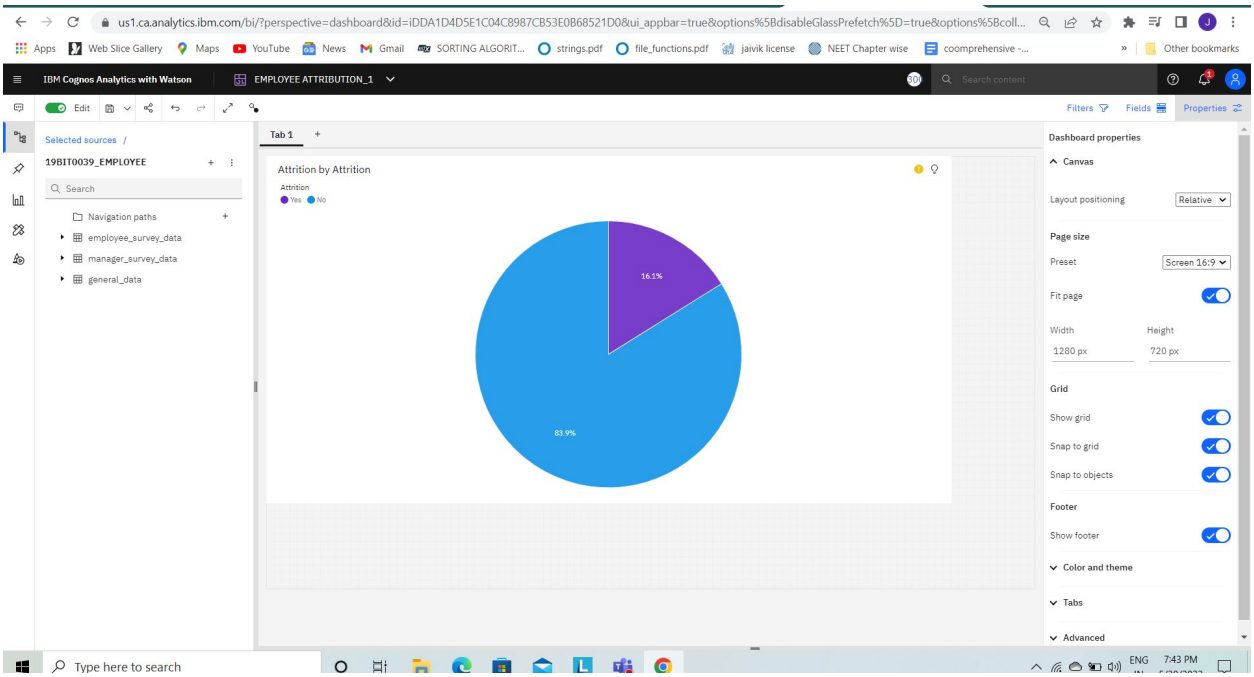
There is 1:1 relationship between general and employee data taking EmployeeID as the common column.



3. DATA VISULATION CHARTS:

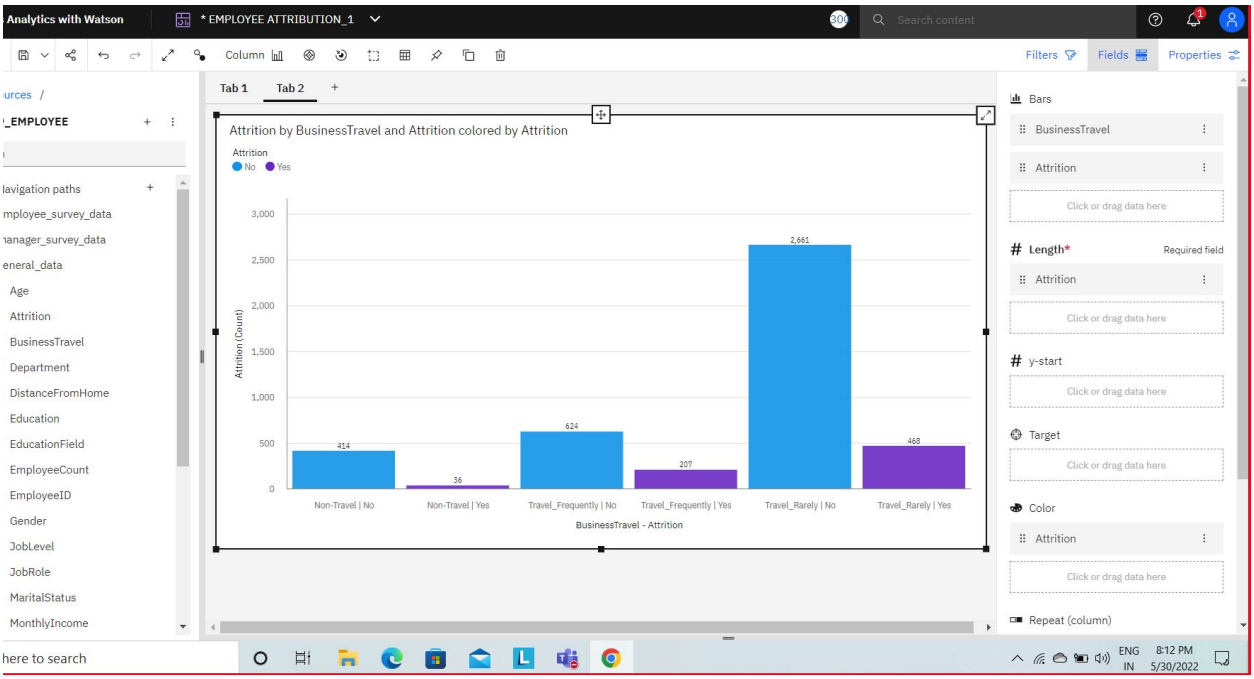
1. **ATTRITION STATUS:**

Visualization of the Attrition status is done on the basis of simple YES/NO on pie chart using the Cognos software dashboard.
The Size and the segment field is selected as Attrition.



2. ATTRITION BASED ON BUSINESS TRAVEL:

Here, we have used a bar chart. The bars include the fields of attrition and Business travels while the lengths denote values of attrition and color has data of attrition.

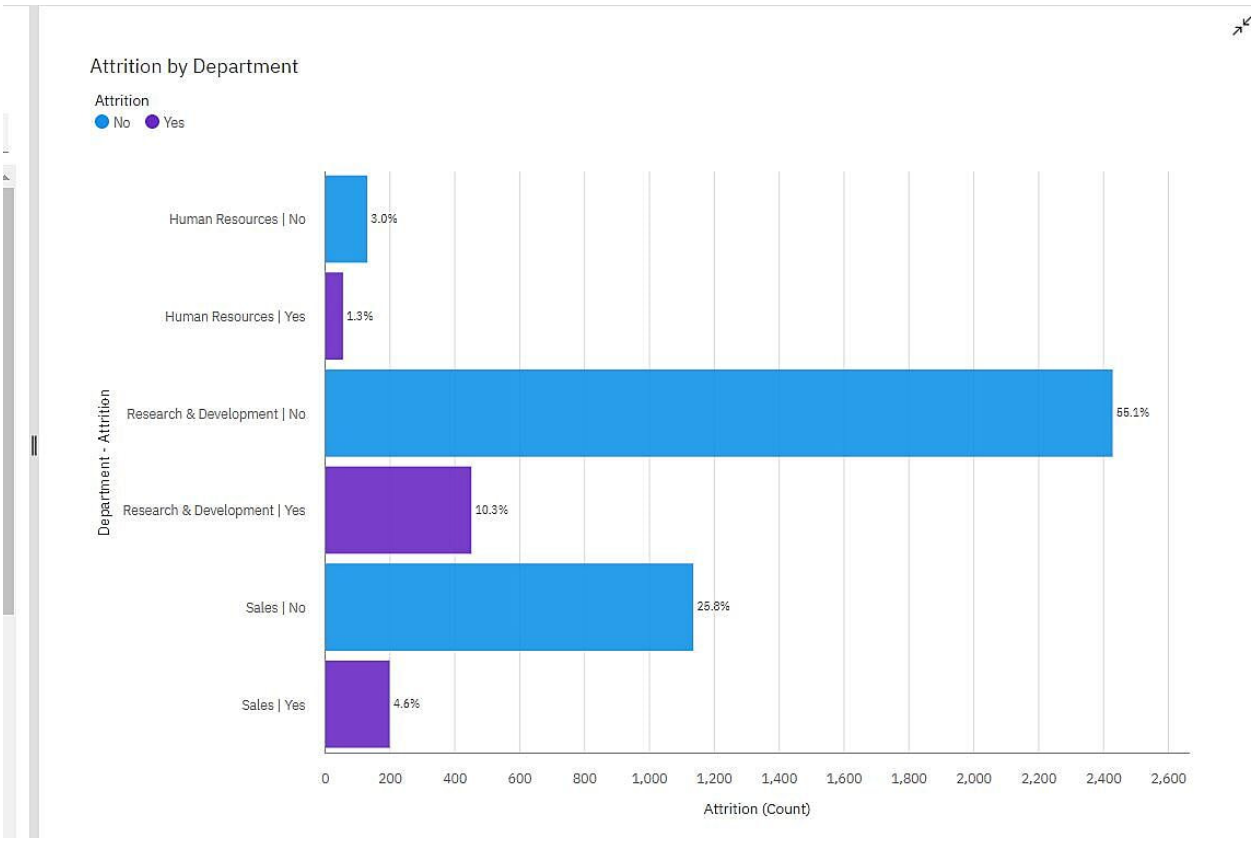


3. ATTRITION BASED ON DEPARTMENT, JOB ROLE, MARITAL STATUS AND EDUCATION:

Here, we have used a dashboard that can accommodate 4 charts. Each place on the dashboard has chart based on department, marital status, education and job role.

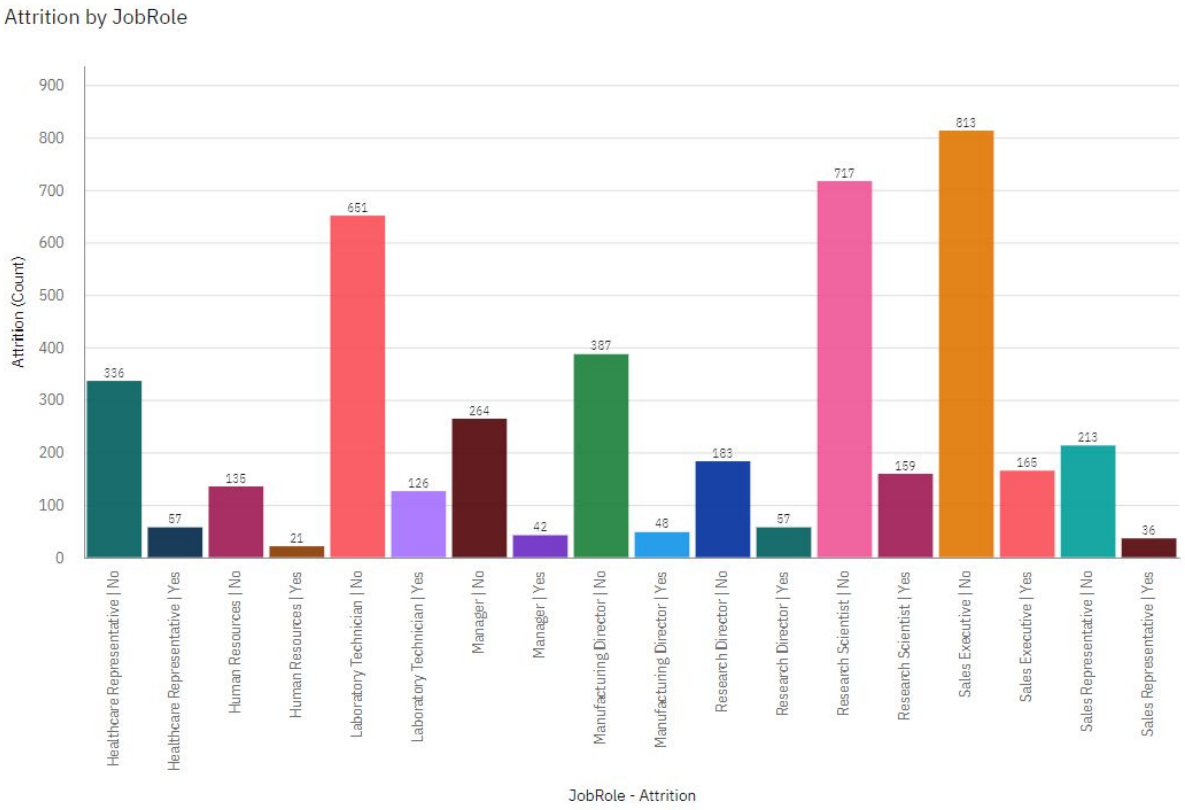
DEPARTMENT:

Bars have Department and Attrition values and length has data of attrition.



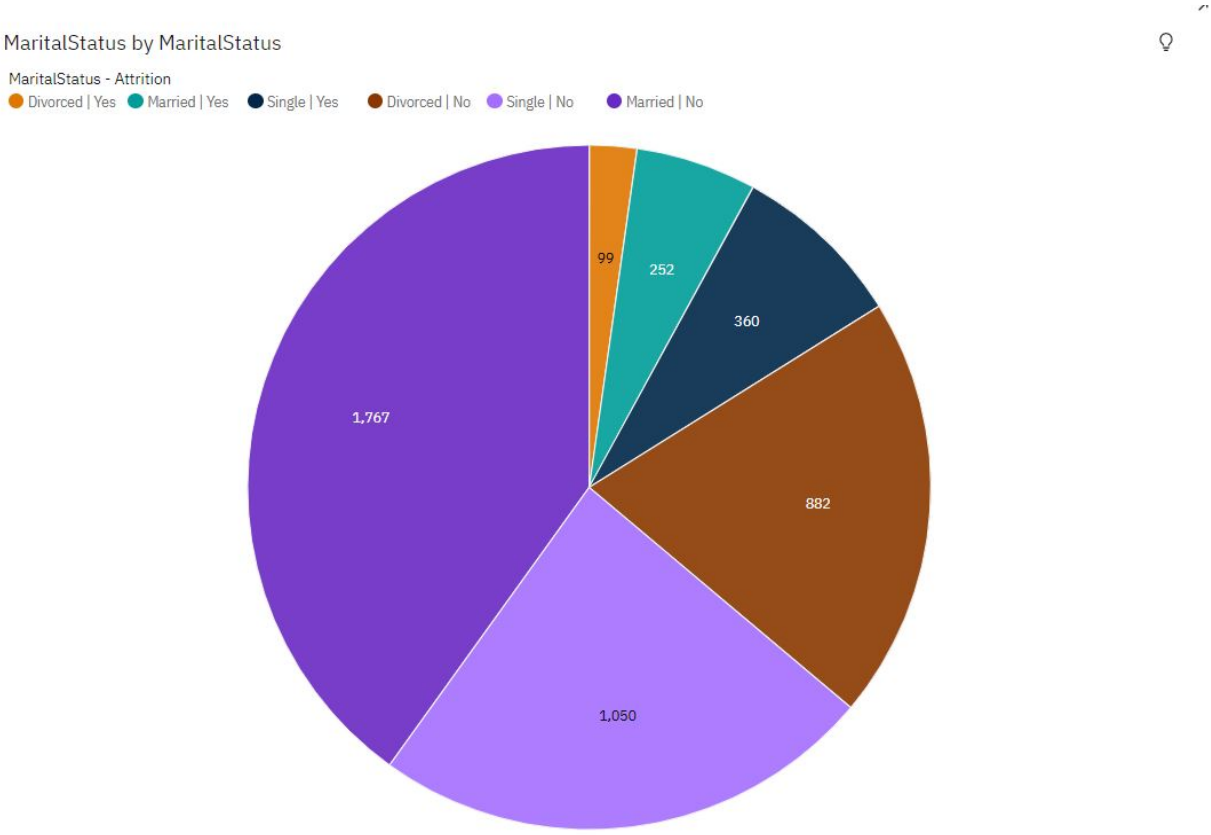
JOB ROLE:

Here, we have used column chart where Bars have the JobRole and Attrition values and length has data of attrition.



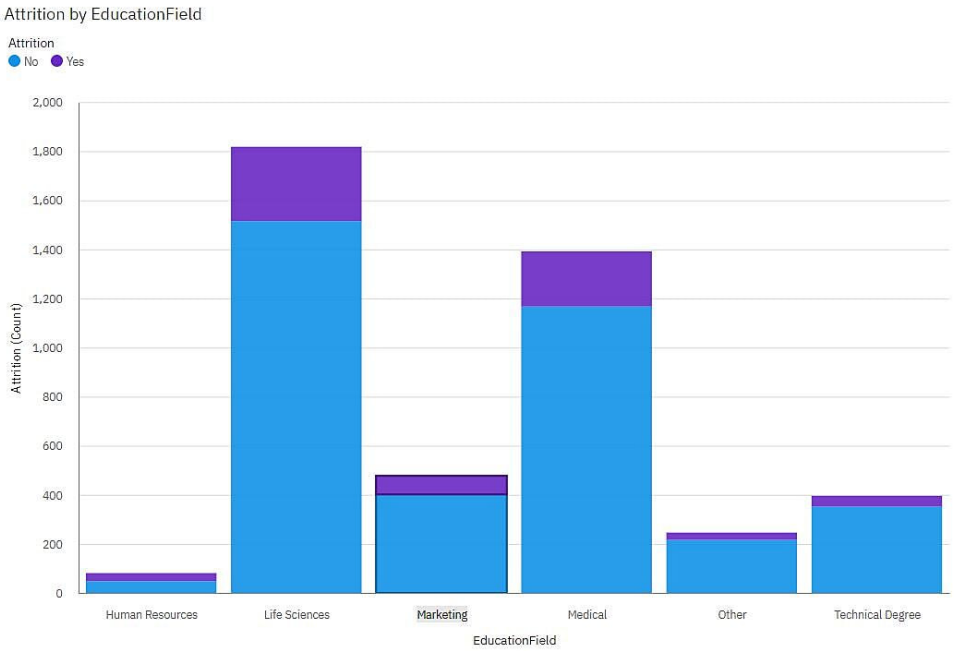
MARITAL STATUS:

Here a pie-chart is used. Segments denote the Marital Status and Attrition and size shows attrition.

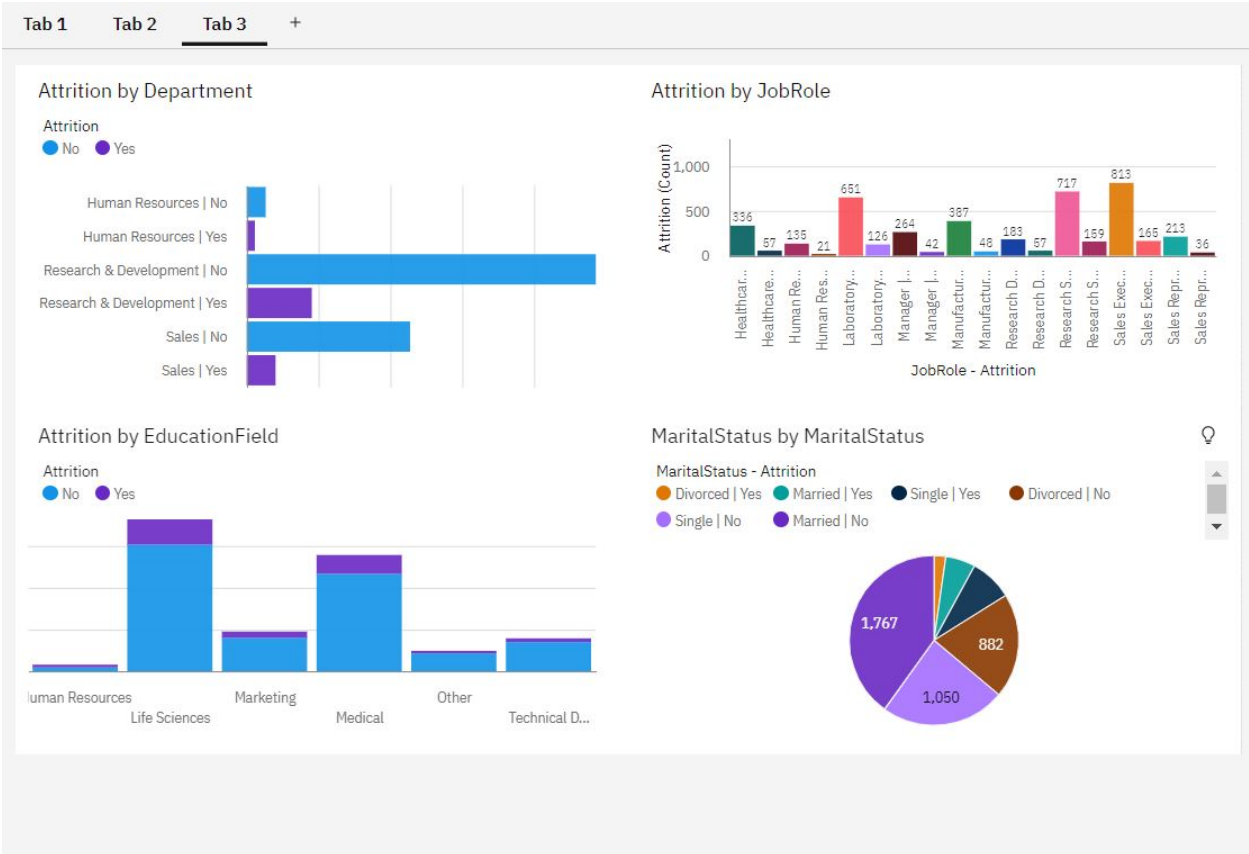


EDUCATION:

Here, stacked column graph is used. Bars of columns denote the EducationFields and length denotes attrition values.

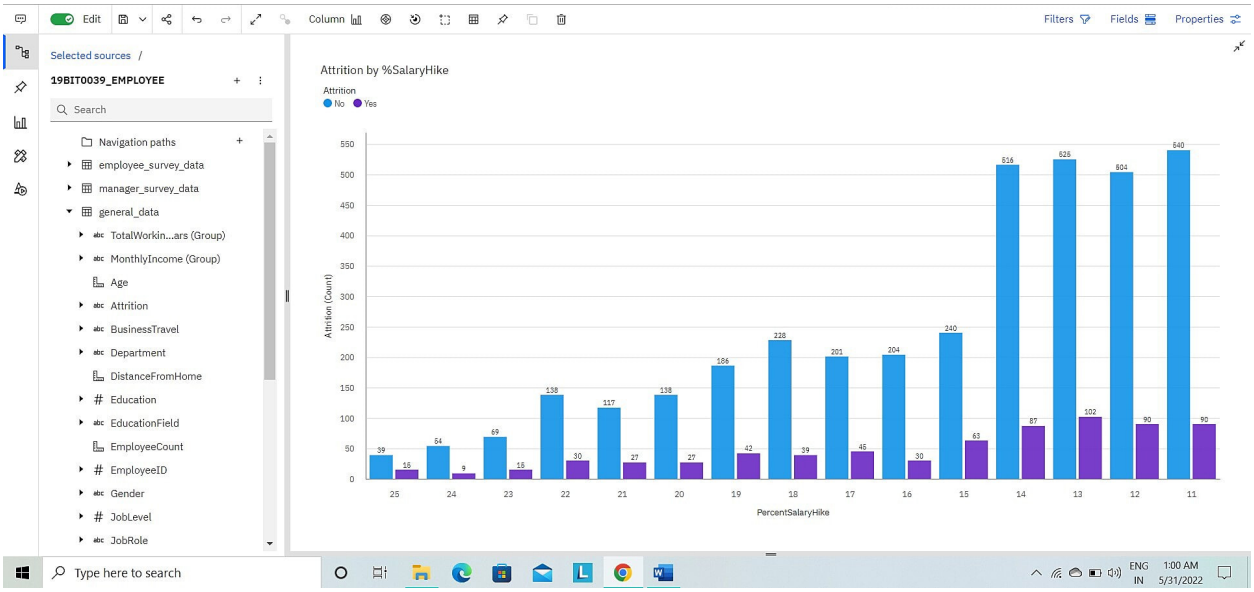


TOGETHER ALL 4 CHARTS:



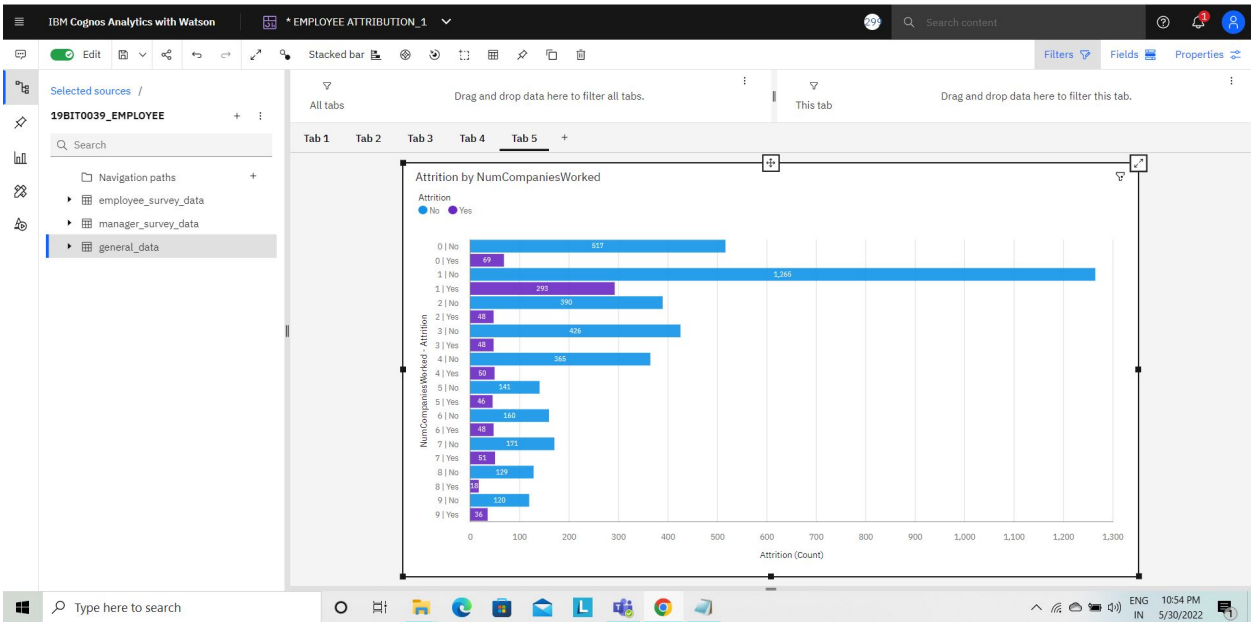
4. ATTRITION BASED ON SALARY HIKE % :

Here, column chart is used. Bars have SalaryHike and Attrition data and length has attrition values.



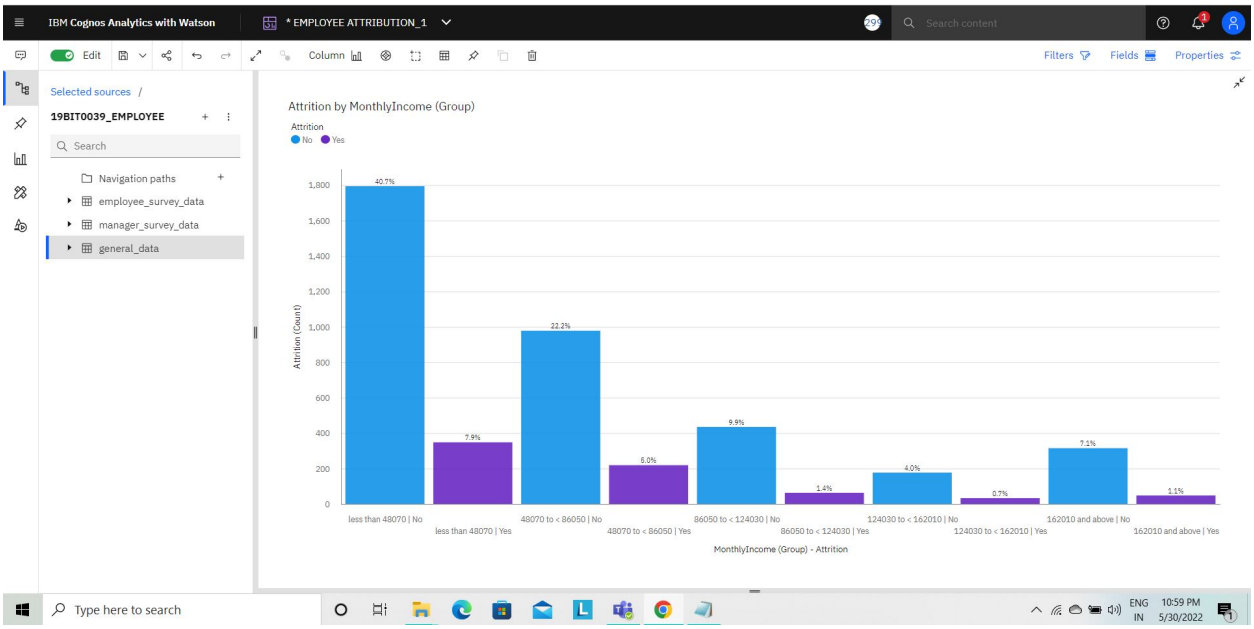
5. ATTRITION BASED ON NUMBER OF COMPANIES WORKED:

Here, stacked BAR graph is used. The bars denote the number of companies worked and Attrition values and the length of bars is determined by Attrition data and color is decided using Attrition.



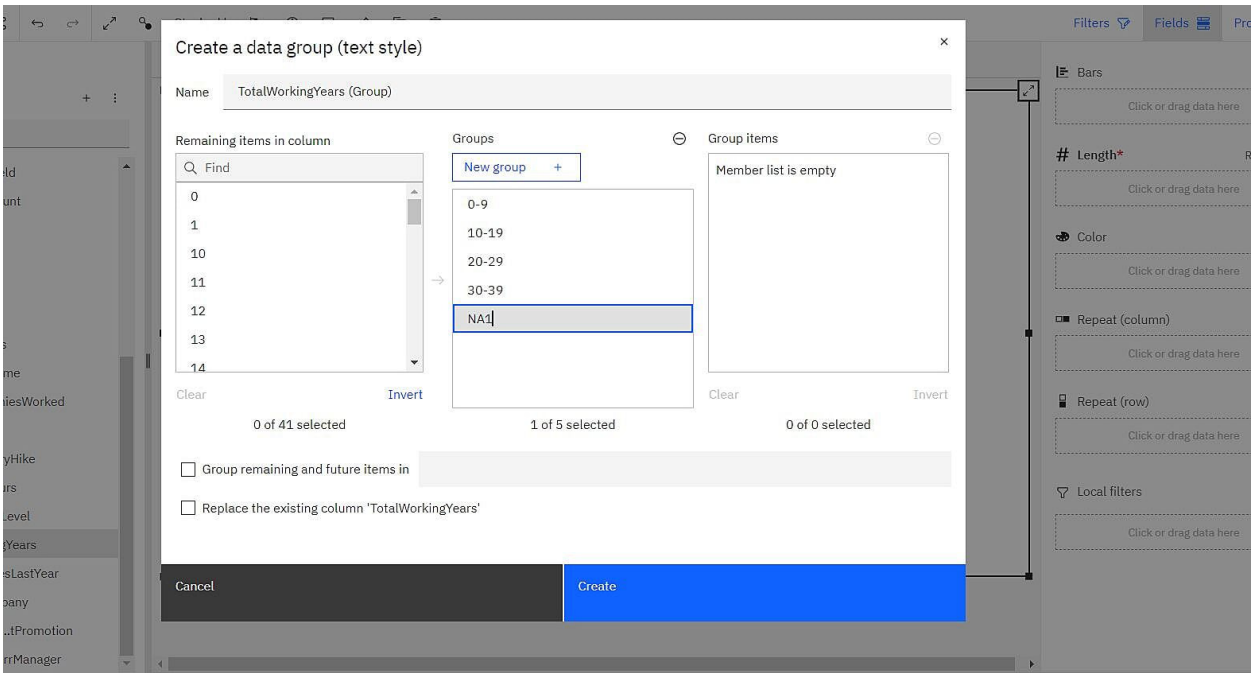
6. ATTRITION BASED ON MONTHLY INCOME:

Here, column chart is used. Bars of graph denote the Monthly Income and Attrition and length of bars is denoted using attrition values.

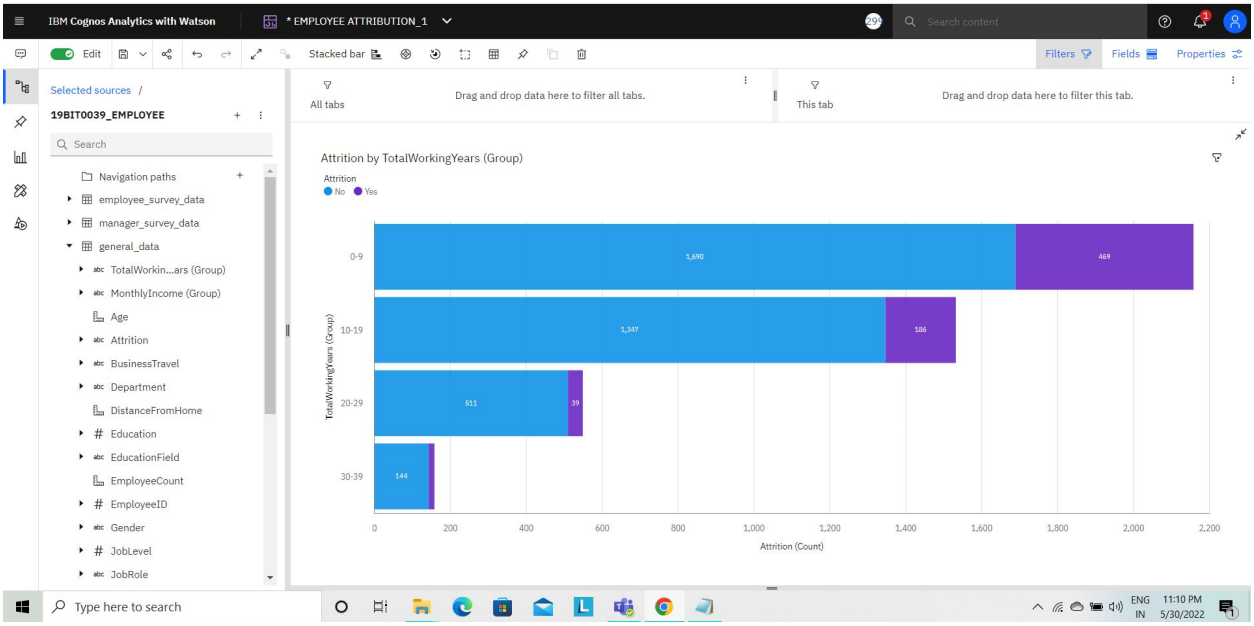


7. **ATTRITION BASED ON EXPERIENCE:**

Here, stacked bar graph is used.
It shows employees' working experience. As we had a large amount of data so, we grouped employees based on their years of experience i.e. 0-9 years of experience to 30-39 and more year of experience.

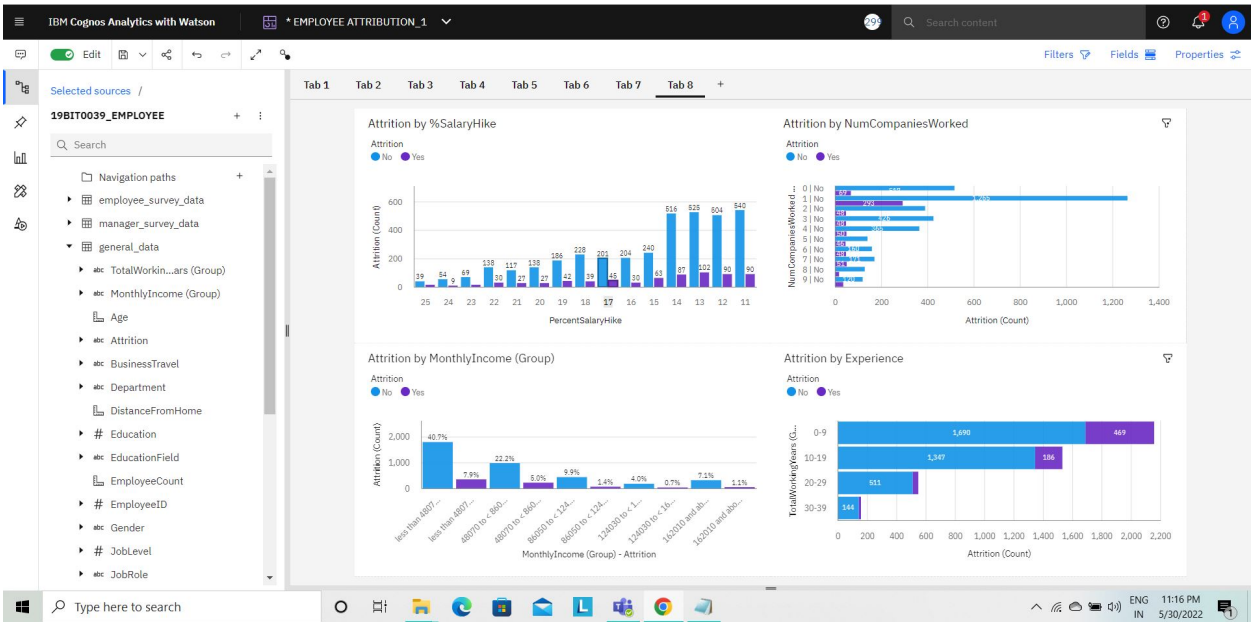


The Graph has bars denoting the number of years of working and their length is determined using attrition data.



8. **DASHBOARD:**

Here, a dashboard of 4 visualizations is made inside 1 tab that includes the last 4 data visualizations.



IBM COGNOS:

https://us1.ca.analytics.ibm.com/bi/?perspective=dashboard&pathRef=.my_folders%2FEMPLOYEE%2BATTRIBUTION_1&action=view&mode=dashboard&subView=model000001811584c55a_00000000

ADVANTAGES AND DISADVANTAGES OF CREATING DASHBOARD

ADVANTAGES:

- Dashboards provide greater visibility with information available whenever it is required to ensure businesses are better placed to respond to changing market conditions.
- IBM cognos allows wide range of file types of upload as data assests.
- It offers data connectors including SQL Server, CSV, Excel, and more.
- With dashboards, we are no longer wasting valuable time generating reports from multiple systems. Instead, data is drawn from a source and displayed as an easy to interpret visual overview.
- It doesn't matter who the user is; when they log in and run a report, they will only see the data they are permitted to see (Data security).
- Cognos can consistently run reports on multi-billion row tables without issue.
- For our project, It provide better Decision Making.
- Using cognos the visualization of data is also very easy and it provides a lot of charts and graph to make our customize dashboard very explainatory.

DISADVANTAGES:

- Flashy or cluttered design, with users attempting to incorporate too much information without understanding constraints or considering their specific needs from the range of different measurables detailed data analysis provides.

- Does not support multidimensional analysis.
- Data modules still can't match the flexibility of the framework manager. When dealing with large data models, or multi-grain, FM still wins.
- The business has no predetermined rules and hierarchies for how dashboard metrics are used. This means each employee can use the metrics in different ways, resulting in a diverse set of data being reported.
- Not a good IDE for Python/R.
- Search engine displaying data elements is still not available on version 11.1.7.

CONCLUSION:

With the help of diagrams, graphs, and maps we can understand given dataset and analyze important trends and findings in it. This understanding of data allows us to ask the right questions to reach our desired goals by optimizing methods. With this project, we learned how to upload and prepare data. We also learned statistical concepts which helped in calculations and plotting of graphs and maps to make a dashboard.
Through analysis of various charts and finding trends in it, I can conclude that the HR department of company should not give critical tasks to the employees of research department or sales department who are non married and getting less salary in monthly income and from medical or life science educational background and also should not give critical tasks to employees those who have less hiked in their salary. This will help lower employee attrition and indirectly increase productivity and trust of employee and finally benefit the organization

****