**NETFLIX FILMS DATA ANALYSIS PROJECT**

### **Project Overview:**

Analysing Films Data For Netflix Management To Gain Insights And Enhance Decision-Making.

### 

### **Project Objective:**

Generate A Improved Dynamic Dashboard’s For Netflix Management To Analyze And Visualize Films Data.

### **Project Requirements:**

* Problem Statement
* Data Source
* Software’s Required

**Problem Statements:**

**Key Performance Indicators (KPIs):**

1. **Number of Movies Added Over Time:**
   * Monthly or yearly count of movies added to the platform.
2. **Top Directors by Movie Count:**
   * Ranking of directors based on the number of movies they've directed.
3. **Top Countries Producing Movies:**
   * Ranking of countries by the number of movies produced.
4. **Average Duration of Movies:**
   * Overall average duration of movies in the dataset.
5. **Genre Distribution:**
   * Percentage distribution of movies across different genres.
6. **Correlation Between Variables:**
   * Correlation between variables like release year, duration.
7. Distribution of Movie types And Genres
8. Breakdown of the number of movies added on each date?
9. Which country has the highest number of movies in the dataset?
10. What is the average duration of movies in the dataset?
11. How many movies are categorized as Children & Family Movies?
12. Who is the director with the highest number of movies in the dataset?
13. Can you list the top 5 longest movies in the dataset?
14. What is the oldest movie in the dataset based on release year?
15. Is there any correlation between the release year and the duration of movies?
16. How many movies are rated TV-Y7?
17. Which movie has the largest cast?
18. What is the most common rating for movies in the dataset?
19. Can you provide a summary of the number of movies released each year?
20. How many movies have been directed by directors with more than one movie in the dataset?
21. What are the top 5 listed genres in the dataset?

**Data Source:**

Utilizing Customers Data Collected By HR’s At telecom Client For Comprehensive Data Analysis.

Data Collection Tool : MS excel

**Software’s Used:**

OS Tool : ChatGPT

BI Tool : Microsoft Power BI

### **Project Process:**

**Step By Step Process:**

1. Collecting the Data set
2. Importing Data set Into Power Bi
3. Data Transform And Cleaning
4. Data Processing(DAX)
5. Data Visualization
6. Final Dash Board

**1.Collecting Data**

The Data set collected From HR’s At Telecom Client In The Form Of Excel Sheets.

**2.Importing Data set Into Microsoft Power BI**

* For That, Open Power BI, Go to Get Data and Select Excel Workbook Then, Make a Connection With Excel File
* After Completion of Connection we can Load Or Transform Data Based On Requirement

**3,4.Data Cleaning &Processing**

* After Loading Data Into PowerBI By Using The Power Query Editor, We Perform DAX(Data Analysis Express) For Data cleaning and Processing
* These Are Some Of DAX Formulas written for Data Visualization

**DAX Formula’s:**

**1.Employees Data table**

1. Total Employee's = COUNT('Pharma Group AG'[Employee ID])
2. Voluntary\_Turnover\_Rate =DIVIDE(CALCULATE(COUNTROWS('Pharma Group AG'), 'Pharma Group AG'[FY20 leaver?] = "Yes" && 'Pharma Group AG'[In base group for turnover FY20] = "N"), COUNTROWS('Pharma Group AG'))
3. Overall\_Turnover\_Rate =DIVIDE(CALCULATE(COUNTROWS('Pharma Group AG'), 'Pharma Group AG'[FY20 leaver?] = "Yes")COUNTROWS('Pharma Group AG'))
4. Male Count = CALCULATE(COUNTROWS(FILTER('Pharma Group AG','Pharma Group AG'[Gender]="Male")))
5. Male % = DIVIDE([Male Count],[Total Employee's])
6. Involuntary\_Turnover\_Rate =DIVIDE(CALCULATE(COUNTROWS('Pharma Group AG'), 'Pharma Group AG'[FY20 leaver?] = "Y" && 'Pharma Group AG'[In base group for turnover FY20] = "Yes"),COUNTROWS('Pharma Group AG'))
7. FY21 Promotion Count = CALCULATE(COUNTROWS(FILTER('Pharma Group AG','Pharma Group AG'[Promotion in FY21?]="Yes")))
8. FY20 Promotion Count = CALCULATE(COUNTROWS(FILTER('Pharma Group AG','Pharma Group AG'[Promotion in FY20?]="Y")))
9. FY20 Leaver Count = CALCULATE(COUNTROWS(FILTER('Pharma Group AG','Pharma Group AG'[FY20 leaver?]="Yes")))
10. FeMale Count = CALCULATE(COUNTROWS(FILTER('Pharma Group AG','Pharma Group AG'[Gender]="Female")))
11. Female % = DIVIDE([FeMale Count],[Total Employee's])
12. Average Male Performance Rating =CALCULATEAVERAGE('Pharma Group AG'[FY20 Performance Rating]),'Pharma Group AG'[Gender] = "Male")
13. Average Female Performance Rating CALCULATE(AVERAGE('Pharma Group AG'[FY20 Performance Rating]),'Pharma Group AG'[Gender] = "Female")
14. % FY21 Promoted = DIVIDE([FY21 Promotion Count],[Total Employee's])
15. % FY21 Male Promoted =
16. VAR TotalFemaleEmployees =CALCULATE(COUNTROWS('Pharma Group AG'),'Pharma Group AG'[Gender] = "Male" )

VAR FemalePromoted =CALCULATE( COUNTROWS('Pharma Group AG'),'Pharma Group AG'[Gender] = "Male",'Pharma Group AG'[Promotion in FY21?] = "Yes") RETURN DIVIDE(FemalePromoted, TotalFemaleEmployees)

1. % FY21 Female Promoted VAR TotalFemaleEmployees CALCULATECOUNTROWS('Pharma Group AG'),'Pharma Group AG'[Gender] = "Female")VAR FemalePromoted =CALCULATE COUNTROWS('Pharma Group AG'),'Pharma Group AG'[Gender] = "Female" 'Pharma Group AG'[Promotion in FY21?] = "Yes")

RETURN DIVIDE(FemalePromoted, TotalFemaleEmployees)

1. % FY20 Promoted = DIVIDE([FY20 Promotion Count],[Total Employee's])
2. % FY20 Male Promoted =VAR TotalFemaleEmployees CALCULATECOUNTROWS('Pharma Group AG'),'Pharma Group AG'[Gender] = "Male)

VAR FemalePromoted CALCULATE(COUNTROWS('Pharma Group AG')'Pharma Group AG'[Gender] = "Male",'Pharma Group AG'[Promotion in FY20?] = "Yes)

RETURN DIVIDE(FemalePromoted, TotalFemaleEmployees)

1. % FY20 Male Hires =VAR TotalFemaleEmployees = CALCULATE( COUNTROWS('Pharma Group AG'),'Pharma Group AG'[Gender] = "Male")

VAR FemalePromoted =CALCULATE( COUNTROWS('Pharma Group AG'), 'Pharma Group AG'[Gender] = "Male",'Pharma Group AG'[New hire FY20?] = "Y" )

RETURN DIVIDE(FemalePromoted, TotalFemaleEmployees)

1. % FY20 Female Promoted = VAR TotalFemaleEmployees = CALCULATE(COUNTROWS('Pharma Group AG'),'Pharma Group AG'[Gender] = "Female" )

VAR FemalePromoted = CALCULATE( COUNTROWS('Pharma Group AG'),'Pharma Group AG'[Gender] = "Female", 'Pharma Group AG'[Promotion in FY20?] = "Yes" )

RETUR DIVIDE(FemalePromoted, TotalFemaleEmployees)

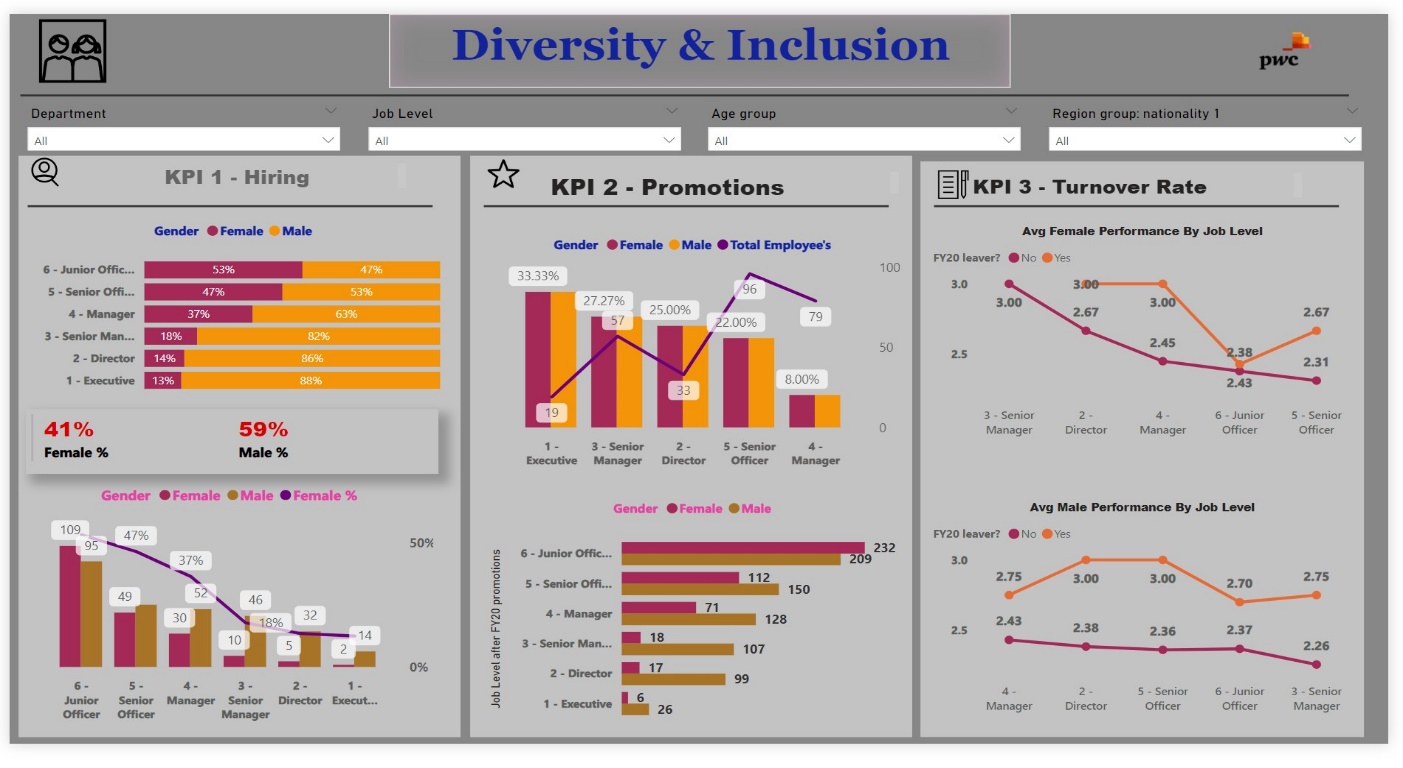
1. % FY20 Female Hires = VAR TotalFemaleEmployees = CALCULATE(  COUNTROWS('Pharma Group AG'), 'Pharma Group AG'[Gender] = "Female" )

VAR FemalePromoted = CALCULATE COUNTROWS('Pharma Group AG'), 'Pharma Group AG'[Gender] = "Female", 'Pharma Group AG'[New hire FY20?] = "Y")

RETURN DIVIDE(FemalePromoted, TotalFemaleEmployees)

**5.Data Visualization**

* After Cleaning And Processing The Data According To The Requirements of Human Resource At Telecom, Prepare Dashboards’ For A HR At Telecom To Get Insights And Improve Decision-Making





### **Conclusion:**

**For KPI’s**

**Gender Representation:**

* There is a noticeable disparity in gender representation, especially at different job levels, with a higher percentage of females hired at junior levels but males dominating senior management positions.

**Promotion Rates:**

* Females have a higher promotion rate at junior levels, but this trend reverses at senior positions, indicating potential barriers to advancement for females in higher roles.

**Turnover Rate:**

* The turnover rate for females is higher across all job levels except for senior management, suggesting possible issues with job satisfaction or work environment that need to be addressed.

**Performance Rating:**

* There appears to be a gender imbalance in performance ratings and executive positions, with males having a higher representation.

**Age Distribution:**

* The majority of employees fall within the 30-39 age group, which could have implications for succession planning and diversity efforts.

**Diversity Measures:**

* The data suggests that there may be opportunities to enhance diversity and inclusion, particularly by focusing on improving gender balance in leadership roles and addressing any disparities in performance ratings.

**Root causes of their slow progress**

**Gender Imbalance in Hiring:**

* A skewed gender ratio at the hiring stage can perpetuate a lack of diversity at higher levels.

**Promotion Disparities:**

* Unequal promotion rates between genders suggest systemic barriers that prevent equal advancement opportunities.

**Higher Turnover Rate for Females:**

* A higher turnover rate among females, especially in senior roles, may indicate issues with the work environment or culture that are not conducive to retention.

**Performance Rating Disparity:**

* If males are consistently receiving higher performance ratings, it could indicate a bias in evaluation processes or a lack of support for female employees’ development.

**Executive Gender Imbalance:**

* A low percentage of female executives suggests barriers to women reaching top leadership positions, which could be due to unconscious bias, lack of mentorship, or insufficient career development opportunities for women.

**Age Group Concentration:**

* A workforce concentrated in specific age groups may lack the benefits of intergenerational diversity, such as varied perspectives and experiences, which can hinder innovation and progress.