**3) HR Analytics (Diversion & Inclusion):**

**SQL Queries:**

-- DIVERSITY & INCLUSION ANALYSIS :

---------------------------------------------------------------------------

SELECT \* FROM PWC\_DataAnalytics.dbo.DiversityInclusion\_Dataset;

---------------------------------------------------------------------------

-- TRANSFORMATION : Last\_hrie\_date from varchar to date

SELECT \*

FROM PWC\_DataAnalytics.dbo.DiversityInclusion\_Dataset;

ALTER TABLE PWC\_DataAnalytics.dbo.DiversityInclusion\_Dataset

ALTER COLUMN Last\_hire\_date DATE;

UPDATE PWC\_DataAnalytics.dbo.DiversityInclusion\_Dataset

SET Last\_hire\_date = CAST(Last\_hire\_date AS DATE)

ALTER TABLE PWC\_DataAnalytics.dbo.DiversityInclusion\_Dataset

ALTER COLUMN FY19\_Performance\_Rating INT;

ALTER TABLE PWC\_DataAnalytics.dbo.DiversityInclusion\_Dataset

ALTER COLUMN FY20\_Performance\_Rating INT;

------------------------------------------------------------------------------------------------------------------------------------------------------

-- ANALYSIS:

SELECT \*

FROM PWC\_DataAnalytics.dbo.DiversityInclusion\_Dataset;

-- TOTAL EMPLOYEE FY20 :

SELECT

COUNT(Employee\_ID) AS Total\_Employee\_Count

FROM PWC\_DataAnalytics.dbo.DiversityInclusion\_Dataset;

---------------------------------------------------------------------------

-- TOTAL EMPLOYEE FY21 :

SELECT

SUM(CASE WHEN FY20\_leaver = 'No' THEN 1 ELSE 0 END) AS Total\_Employee\_FY21

FROM PWC\_DataAnalytics.dbo.DiversityInclusion\_Dataset;

---------------------------------------------------------------------------

-- EMPLOYEE FY20 BY GENDER :

SELECT

Gender,

COUNT(Employee\_ID) AS Total\_Employee\_Count

FROM PWC\_DataAnalytics.dbo.DiversityInclusion\_Dataset

GROUP BY Gender;

---------------------------------------------------------------------------

-- TOTAL EMPLOYEE FY21 BY GENDER :

SELECT

Gender,

SUM(CASE WHEN FY20\_leaver = 'No' THEN 1 ELSE 0 END) AS Total\_Employee\_FY21

FROM PWC\_DataAnalytics.dbo.DiversityInclusion\_Dataset

GROUP BY Gender;

---------------------------------------------------------------------------

-- EMPLOYEE FY20 PERCENTAGE BY GENDER :

SELECT

Gender,

CEILING(SUM(CASE WHEN FY20\_leaver = 'No' THEN 1 ELSE 0 END) \* 100 / (

SELECT SUM(CASE WHEN FY20\_leaver = 'No' THEN 1 ELSE 0 END) FROM PWC\_DataAnalytics.dbo.DiversityInclusion\_Dataset))

AS Employee\_FY21

FROM PWC\_DataAnalytics.dbo.DiversityInclusion\_Dataset

GROUP BY Gender;

---------------------------------------------------------------------------

-- EMPLOYEE FY21 PERCENTAGE BY GENDER :

SELECT

Gender,

CEILING(COUNT(DISTINCT Employee\_ID) \* 100 / (SELECT COUNT(DISTINCT Employee\_ID) FROM PWC\_DataAnalytics.dbo.DiversityInclusion\_Dataset)) AS Employee\_Count\_Rate

FROM PWC\_DataAnalytics.dbo.DiversityInclusion\_Dataset

GROUP BY Gender;

---------------------------------------------------------------------------

-- PROMOTION BY GENDER FY20 :

SELECT

Promotion\_in\_FY20,

Gender,

SUM(CASE WHEN New\_hire\_FY20 = 'N' THEN 1 ELSE 0 END) AS Promotion\_For\_FY20

FROM PWC\_DataAnalytics.dbo.DiversityInclusion\_Dataset

GROUP BY Gender, Promotion\_in\_FY20;

---------------------------------------------------------------------------

-- EMPLOYEE COUNT BY GENDER AFTER FY20 PROMOTION :

SELECT

Job\_Level\_after\_FY20\_promotions,

Gender,

COUNT(Employee\_ID) AS Employee\_FY20

FROM PWC\_DataAnalytics.dbo.DiversityInclusion\_Dataset

GROUP BY Gender, Job\_Level\_after\_FY20\_promotions;

---------------------------------------------------------------------------

-- EMPLOYEE COUNT BY GENDER AFTER FY21 PROMOTION :

SELECT

Job\_Level\_after\_FY21\_promotions,

Gender,

SUM(CASE WHEN FY20\_leaver = 'No' THEN 1 ELSE 0 END) AS Employee\_FY21

FROM PWC\_DataAnalytics.dbo.DiversityInclusion\_Dataset

GROUP BY Gender, Job\_Level\_after\_FY21\_promotions

HAVING SUM(CASE WHEN FY20\_leaver = 'No' THEN 1 ELSE 0 END) > 0;

---------------------------------------------------------------------------

-- PROMOTION BY GENDER IN FY20 :

SELECT

SUM(CASE WHEN New\_hire\_FY20 = 'N' THEN 1 ELSE 0 END) AS EmployeE\_Promotion\_FY20

FROM PWC\_DataAnalytics.dbo.DiversityInclusion\_Dataset

GROUP BY Promotion\_in\_FY20, Gender;

---------------------------------------------------------------------------

-- PROMOTION BY GENDER FY21 :

SELECT

COUNT(DISTINCT Employee\_ID) AS Employee\_Promotion\_FY21

FROM PWC\_DataAnalytics.dbo.DiversityInclusion\_Dataset

GROUP BY Promotion\_in\_FY21, Gender;

---------------------------------------------------------------------------

-- PROMOTION RATE IN FY20 :

SELECT

SUM(CASE WHEN New\_hire\_FY20 = 'N' AND Promotion\_in\_FY20 = 'Y' THEN 1 ELSE 0 END) \* 100 / (

SELECT SUM(CASE WHEN New\_hire\_FY20 = 'N' THEN 1 ELSE 0 END)) AS Promotion\_Rate\_FY20

FROM PWC\_DataAnalytics.dbo.DiversityInclusion\_Dataset;

---------------------------------------------------------------------------

-- PROMOTION RATE IN FY20 BY GENDER :

-- MALE

SELECT

SUM(CASE WHEN New\_hire\_FY20 = 'N' AND Promotion\_in\_FY20 = 'Y' AND Gender = 'Male' THEN 1 ELSE 0 END) \* 100 / (

SELECT SUM(CASE WHEN New\_hire\_FY20 = 'N' AND Promotion\_in\_FY20 = 'Y' THEN 1 ELSE 0 END)) AS Promotion\_Rate\_FY20\_Male

FROM PWC\_DataAnalytics.dbo.DiversityInclusion\_Dataset;

-- FEMALE :

SELECT

SUM(CASE WHEN New\_hire\_FY20 = 'N' AND Promotion\_in\_FY20 = 'Y' AND Gender = 'Female' THEN 1 ELSE 0 END) \* 100 / (

SELECT SUM(CASE WHEN New\_hire\_FY20 = 'N' AND Promotion\_in\_FY20 = 'Y' THEN 1 ELSE 0 END)) AS Promotion\_Rate\_FY20\_Female

FROM PWC\_DataAnalytics.dbo.DiversityInclusion\_Dataset;

---------------------------------------------------------------------------

-- PROMOTION RATE IN FY21 :

SELECT

SUM(CASE WHEN Promotion\_in\_FY21 = 'Yes' THEN 1 ELSE 0 END) \* 100 / (

SELECT

SUM(CASE WHEN In\_base\_group\_for\_Promotion\_FY21 = 'Yes' THEN 1 ELSE 0 END)

FROM PWC\_DataAnalytics.dbo.DiversityInclusion\_Dataset) AS Promotion\_Rate\_FY21

FROM PWC\_DataAnalytics.dbo.DiversityInclusion\_Dataset;

SELECT \* FROM PWC\_DataAnalytics.dbo.DiversityInclusion\_Dataset;

---------------------------------------------------------------------------

-- PROMOTION RATE IN FY21 BY GENDER :

-- MALE

SELECT

SUM(CASE WHEN Promotion\_in\_FY21 = 'Yes' AND Gender = 'Male' THEN 1 ELSE 0 END) \* 100 / (

SELECT SUM(CASE WHEN Promotion\_in\_FY21 = 'Yes' THEN 1 ELSE 0 END)) AS Promotion\_Rate\_FY21\_Male

FROM PWC\_DataAnalytics.dbo.DiversityInclusion\_Dataset;

-- FEMALE :

SELECT

SUM(CASE WHEN Promotion\_in\_FY21 = 'Yes' AND Gender = 'Female' THEN 1 ELSE 0 END) \* 100 / (

SELECT SUM(CASE WHEN Promotion\_in\_FY21 = 'Yes' THEN 1 ELSE 0 END)) AS Promotion\_Rate\_FY21\_Female

FROM PWC\_DataAnalytics.dbo.DiversityInclusion\_Dataset;

---------------------------------------------------------------------------

-- HIRING RATE FY20 :

WITH HiringRate AS(

SELECT

COUNT(DISTINCT Employee\_ID) AS Total\_Employee,

SUM(CASE WHEN New\_hire\_FY20 = 'Y' THEN 1 ELSE 0 END) AS New\_Hire\_Count,

SUM(CASE WHEN New\_hire\_FY20 = 'N' THEN 1 ELSE 0 END) AS Old\_Count

FROM PWC\_DataAnalytics.dbo.DiversityInclusion\_Dataset

)

SELECT

(New\_Hire\_Count \* 100) / ((Total\_Employee + Old\_Count)/2) AS Hiring\_Rate\_FY20

FROM HiringRate;

---------------------------------------------------------------------------

-- HIRING RATE FY20 BY GENDER :

WITH HiringRateGedner AS(

SELECT

Gender,

COUNT(DISTINCT Employee\_ID) AS Total\_Employee,

SUM(CASE WHEN New\_hire\_FY20 = 'Y' THEN 1 ELSE 0 END) AS New\_Hire\_Count,

SUM(CASE WHEN New\_hire\_FY20 = 'N' THEN 1 ELSE 0 END) AS Old\_Count

FROM PWC\_DataAnalytics.dbo.DiversityInclusion\_Dataset

GROUP BY Gender

)

SELECT

Gender,

CEILING((New\_Hire\_Count \* 100) / ((Total\_Employee + Old\_Count)/2)) AS Hiring\_Rate\_FY20

FROM HiringRateGedner;

---------------------------------------------------------------------------

-- EMPLOYEE EXIT RATE:

WITH ExitRate AS(

SELECT

COUNT(DISTINCT Employee\_ID) AS Total\_Employee,

SUM(CASE WHEN FY20\_leaver = 'Yes' THEN 1 ELSE 0 END) AS New\_Hire\_Count,

SUM(CASE WHEN New\_hire\_FY20 = 'N' THEN 1 ELSE 0 END) AS Old\_Count

FROM PWC\_DataAnalytics.dbo.DiversityInclusion\_Dataset

)

SELECT

(New\_Hire\_Count \* 100) / ((Total\_Employee + Old\_Count)/2) AS Hiring\_Rate\_FY20

FROM ExitRate;

---------------------------------------------------------------------------

-- EMPLOYEE EXIT RATE BY GENDER:

WITH ExitRateGender AS (

SELECT

Gender,

COUNT(DISTINCT Employee\_ID) AS Total\_Employee,

SUM(CASE WHEN FY20\_leaver = 'Yes' THEN 1 ELSE 0 END) AS Leaver\_Count,

SUM(CASE WHEN New\_hire\_FY20 = 'N' THEN 1 ELSE 0 END) AS Old\_Count

FROM PWC\_DataAnalytics.dbo.DiversityInclusion\_Dataset

GROUP BY Gender

)

SELECT

Gender,

CEILING((Leaver\_Count \* 100.0) / ((Total\_Employee + Old\_Count)/2)) AS Exit\_Rate\_FY20

FROM ExitRateGender;

---------------------------------------------------------------------------

-- PERFORMANCE RATING IN FY19 BY GENDER AND JOB LEVEL:

SELECT

Job\_Level\_before\_FY20\_promotions, Gender,

AVG(CASE WHEN New\_hire\_FY20 = 'N' THEN FY19\_Performance\_Rating ELSE 0 END) AS Performance\_Rating\_FY19

FROM PWC\_DataAnalytics.dbo.DiversityInclusion\_Dataset

GROUP BY Job\_Level\_before\_FY20\_promotions, Gender

HAVING AVG(CASE WHEN New\_hire\_FY20 = 'N' THEN FY19\_Performance\_Rating ELSE 0 END) > 0;

---------------------------------------------------------------------------

-- AVERAGE OF PERFORMANCE RATING IN FY19 :

SELECT

AVG(FY19\_Performance\_Rating) AS Performance\_Rating\_FY19

FROM PWC\_DataAnalytics.dbo.DiversityInclusion\_Dataset

WHERE New\_hire\_FY20 = 'N';

---------------------------------------------------------------------------

-- AVERAGE PERFORMANCE RATING BY GENDER IN FY19 :

SELECT

Gender,

AVG(FY19\_Performance\_Rating) AS Performance\_Rating\_FY19

FROM PWC\_DataAnalytics.dbo.DiversityInclusion\_Dataset

WHERE New\_hire\_FY20 = 'N'

GROUP BY Gender;

---------------------------------------------------------------------------

-- PERFORMANCE RATING IN FY20 BY GENDER AND JOB LEVEL :

SELECT

Job\_Level\_after\_FY20\_promotions,

Gender,

AVG(FY20\_Performance\_Rating) AS Performance\_Rating\_FY20

FROM PWC\_DataAnalytics.dbo.DiversityInclusion\_Dataset

GROUP BY Job\_Level\_after\_FY20\_promotions, Gender;

---------------------------------------------------------------------------

-- AVERAGE PERFORMANCE RATING IN FY20 :

SELECT

AVG(FY20\_Performance\_Rating) AS Performance\_Rating\_FY20

FROM PWC\_DataAnalytics.dbo.DiversityInclusion\_Dataset;

---------------------------------------------------------------------------

-- AVERAGE PERFORMANCE RATING BY GENDER IN FY20 :

SELECT

Gender,

AVG(FY20\_Performance\_Rating) AS Performance\_Rating\_FY20

FROM PWC\_DataAnalytics.dbo.DiversityInclusion\_Dataset

GROUP BY Gender;

---------------------------------------------------------------------------

-- AGE GROUP BY GENDER :

SELECT

Gender,

Age\_group,

COUNT(DISTINCT Employee\_ID) AS Employee\_Count

FROM PWC\_DataAnalytics.dbo.DiversityInclusion\_Dataset

GROUP BY Gender, Age\_group;

---------------------------------------------------------------------------

-- AGE GROUP BY EXECUTIVE LEVEL BY GENDER :

SELECT

Gender,

Age\_group,

COUNT(DISTINCT Employee\_ID) AS Employee\_Count\_at\_Executive\_Level

FROM PWC\_DataAnalytics.dbo.DiversityInclusion\_Dataset

WHERE Job\_Level\_after\_FY20\_promotions = '1 - Executive'

GROUP BY Gender, Age\_group;

---------------------------------------------------------------------------

-- REGION BY GENDER :

SELECT

Gender,

[Region group nationality 1],

COUNT(DISTINCT Employee\_ID) AS Employee\_Count

FROM PWC\_DataAnalytics.dbo.DiversityInclusion\_Dataset

GROUP BY Gender, [Region group nationality 1];

---------------------------------------------------------------------------

-- REGION BY EXECUTIVE LEVEL BY GENDER :

SELECT

[Region group nationality 1],

Age\_group,

COUNT(DISTINCT Employee\_ID) AS Employee\_Count\_at\_Executive\_Level

FROM PWC\_DataAnalytics.dbo.DiversityInclusion\_Dataset

WHERE Job\_Level\_after\_FY20\_promotions = '1 - Executive'

GROUP BY [Region group nationality 1], Age\_group;

---------------------------------------------------------------------------

-- JOB TYPE BY GENDER :

SELECT

Gender,

Time\_type,

COUNT(DISTINCT Employee\_ID) AS Employee\_Count

FROM PWC\_DataAnalytics.dbo.DiversityInclusion\_Dataset

GROUP BY Gender, Time\_type;

---------------------------------------------------------------------------

-- JOB TYPE BY EXECUTIVE LEVEL BY GENDER :

SELECT

Time\_type,

Gender,

COUNT(DISTINCT Employee\_ID) AS Employee\_Count\_at\_Executive\_Level

FROM PWC\_DataAnalytics.dbo.DiversityInclusion\_Dataset

WHERE Job\_Level\_after\_FY20\_promotions = '1 - Executive'

GROUP BY Time\_type, Gender;

---------------------------------------------------------------------------

SELECT \* FROM PWC\_DataAnalytics.dbo.DiversityInclusion\_Dataset;

**Power Bi DAX:**

1. # avg employee FY20 = DIVIDE((PharmaGroupAG[# before hire] + PharmaGroupAG[# employee in FY20]), 2)
2. # avg rating FY19 = CALCULATE(AVERAGE(PharmaGroupAG[FY19\_Performance\_Rating]), PharmaGroupAG[New\_hire\_FY20] = "N")
3. # avg rating FY19 Female = CALCULATE(AVERAGE(PharmaGroupAG[FY19\_Performance\_Rating]), PharmaGroupAG[New\_hire\_FY20] = "N", PharmaGroupAG[Gender] = "Female")
4. # avg rating FY19 Male = CALCULATE(AVERAGE(PharmaGroupAG[FY19\_Performance\_Rating]), PharmaGroupAG[New\_hire\_FY20] = "N", PharmaGroupAG[Gender] = "Male")
5. # avg rating FY20 = AVERAGE(PharmaGroupAG[FY20\_Performance\_Rating])
6. # avg rating FY20 Female = CALCULATE(AVERAGE(PharmaGroupAG[FY20\_Performance\_Rating]), PharmaGroupAG[Gender] = "Female")
7. # avg rating FY20 Male = CALCULATE(AVERAGE(PharmaGroupAG[FY20\_Performance\_Rating]), PharmaGroupAG[Gender] = "Male")
8. # before hire = CALCULATE(COUNT(PharmaGroupAG[Employee\_ID]), PharmaGroupAG[New\_hire\_FY20] = "N")
9. # employee in FY20 = DISTINCTCOUNT(PharmaGroupAG[Employee\_ID])
10. # employees in FY21 = CALCULATE(COUNT(PharmaGroupAG[Employee\_ID]), PharmaGroupAG[FY20\_leaver] = "No")
11. # hire\_FY20 = CALCULATE(COUNT(PharmaGroupAG[Employee\_ID]), PharmaGroupAG[New\_hire\_FY20] = "Y")
12. # leavers = CALCULATE(COUNT(PharmaGroupAG[Employee\_ID]), PharmaGroupAG[FY20\_leaver] = "Yes")
13. # men = CALCULATE(COUNT(PharmaGroupAG[Employee\_ID]), PharmaGroupAG[Gender] = "Male")
14. # women = CALCULATE(COUNT(PharmaGroupAG[Employee\_ID]), PharmaGroupAG[Gender] = "Female")
15. % exit rate = DIVIDE(PharmaGroupAG[# leavers], PharmaGroupAG[# avg employee FY20])
16. % hire rate = DIVIDE(PharmaGroupAG[# hire\_FY20], PharmaGroupAG[# avg employee FY20])
17. % promotion rate FY20 = DIVIDE(CALCULATE(COUNT(PharmaGroupAG[Employee\_ID]), PharmaGroupAG[Promotion\_in\_FY20] = "Y"), CALCULATE(COUNT(PharmaGroupAG[Employee\_ID]), PharmaGroupAG[New\_hire\_FY20] = "N" ))
18. % promotion rate FY20 Female = DIVIDE(CALCULATE(COUNT(PharmaGroupAG[Employee\_ID]), PharmaGroupAG[Promotion\_in\_FY20] = "Y", PharmaGroupAG[Gender] = "Female"), CALCULATE(COUNT(PharmaGroupAG[Employee\_ID]), PharmaGroupAG[Promotion\_in\_FY20] = "Y"))
19. % promotion rate FY20 Male = DIVIDE(CALCULATE(COUNT(PharmaGroupAG[Employee\_ID]), PharmaGroupAG[Promotion\_in\_FY20] = "Y", PharmaGroupAG[Gender] = "Male"), CALCULATE(COUNT(PharmaGroupAG[Employee\_ID]), PharmaGroupAG[Promotion\_in\_FY20] = "Y" ))
20. % promotion rate FY21 = DIVIDE(CALCULATE(COUNT(PharmaGroupAG[Employee\_ID]), PharmaGroupAG[Promotion\_in\_FY21] = "Yes"), CALCULATE(COUNT(PharmaGroupAG[Employee\_ID]), PharmaGroupAG[In\_base\_group\_for\_Promotion\_FY21] = "Yes"))
21. % promotion rate FY21 Female = DIVIDE(CALCULATE(COUNT(PharmaGroupAG[Employee\_ID]), PharmaGroupAG[Promotion\_in\_FY21] = "Yes", PharmaGroupAG[Gender] = "Female"), CALCULATE(COUNT(PharmaGroupAG[Employee\_ID]), PharmaGroupAG[Promotion\_in\_FY21] = "Yes"))
22. % promotion rate FY21 Male = DIVIDE(CALCULATE(COUNT(PharmaGroupAG[Employee\_ID]), PharmaGroupAG[Promotion\_in\_FY21] = "Yes", PharmaGroupAG[Gender] = "Male"), CALCULATE(COUNT(PharmaGroupAG[Employee\_ID]), PharmaGroupAG[Promotion\_in\_FY21] = "Yes"))