

“HR Data Project Documentation”

Project Overview:

Project Name: HR Data Project

Team Members: Omar Magdy, Ahmed Yasser, Menna Allah Abdul-Majeed, Samar Wagih and Kyrollos Nader

Main Objective: To analyze HR data to understand the drivers of employee attrition and provide actionable insights to decrease the attrition rate and increase employee tenure.

Importance: Reducing attrition is crucial for increasing employee tenure, which can lead to greater organizational stability, reduced recruitment and training costs, and improved knowledge retention and productivity.

Data Source: Graduation project dataset from the Digital Egypt Pioneers Initiative (DEPI).

Problem Statement:

The primary question this project aims to address is: **How can we decrease the employee attrition rate to increase tenure?** High attrition rates lead to loss of talent, increased costs, and reduced organizational efficiency. By analyzing employee data, we seek to identify factors contributing to attrition and provide actionable recommendations to improve retention.

Dataset Description:

Source:

The dataset is sourced from the Digital Egypt Pioneers Initiative (DEPI) as part of a graduation project. It contains HR-related data, including employee reviews, demographics, and employment details, with the primary dataset stored in files

Key Columns and Features:

The dataset includes several key columns and derived features critical to the analysis, as depicted in the data model:

- **Employee:** A detailed dataset capturing employee information, including demographics, education, job-related details, compensation, and retention metrics for HR and workforce analysis.
- **SatisfiedLevel:** A reference table mapping satisfaction level identifiers to descriptive categories, likely used to categorize employee satisfaction survey results.

- **Education Level:** A reference table mapping education level identifiers to descriptive categories.
- **PerformanceRating:** This dataset records employee performance evaluations, including satisfaction with work environment, job, and relationships, as well as training participation, work-life balance, and performance ratings from both employees and managers.
- **RatingLevel:** A reference table mapping performance rating identifiers to descriptive categories, likely used to categorize employee performance evaluations.

Data Preprocessing and Cleaning:

Cleaning Steps:

The dataset underwent extensive cleaning to ensure data quality and reliability:

- **Handling Missing Reviews:** Identified 190 employees who did not provide reviews because they were hired and left within the same year. These employees were excluded from review-based analyses, as a minimum tenure of one year was required for reviews.
- **Review Deduplication:** Reduced the initial 6,709 reviews to 4,982 by removing duplicate or irrelevant reviews (e.g., reviews made before hiring and after leaving, as seen in Performance.csv where EmployeeID "79F7-78EC" has a review on 1/2/2013 but was hired on 12/14/2013).
- **Handling Inconsistencies in Data:** We noticed some inconsistencies like an employee who is a sales executive but is in Technology department.
- **Data Validation:** Created a Validation column to validate the number of reviews against employee turnover. In Performance.csv, the Attrition column ("Yes"/"No") was cross-checked with YearsAtCompany.
- **Normalization:** Normalized tables to eliminate redundancy and improve data integrity, as reflected in the data model structure.

Feature Engineering: To enhance the dataset for analysis, the following features were engineered:

- **Leaving Date:** Calculated by adding YearsAtCompany to the HireDate.
- **Tenure Classification:** Categorized employees based on tenure to analyze retention patterns.
- **Satisfaction and Balance Columns:** Replaced numerical review results with text-based values (e.g., Very Satisfied, Neutral, Dissatisfied) for JobSatisfaction, RelationshipSatisfaction, and WorkLifeBalance to improve readability in visualizations.
- **Age Classification:** Grouped employees into categories (e.g., Young Adulthood, Early Middle Adulthood, Late Middle Adulthood).

- **Distance from Home Classification:** Categorized commuting distances (e.g., Short, Medium, Long, Very Long).
- **Talented Employees:** Identified employees who received a promotion in their first year (e.g., 1st Year Promoted Employees).

Tools and Technologies:

The project utilized the following tools for data processing, analysis, and visualization:

- **Power BI:** For creating interactive dashboards and visualizations, as shown in the provided screenshots.
- **Tableau:** For additional visualization and exploratory data analysis.

Applied Steps:

The image displays three screenshots of the Power BI Query Settings window for a query named 'Performance'. Each screenshot shows a different set of applied steps, illustrating the iterative process of data transformation.

Query Settings (Left): Shows the 'Name' field set to 'Performance'. The 'APPLIED STEPS' list includes: Source, Promoted Headers, Changed Type, Merged Columns, Changed Type1, Merged Queries, Expanded DimTraining, Trimmed Text, Removed Columns, Changed Type2, Merged Queries2, Expanded DimRole, Removed Columns1, Renamed Columns, Conditional "SalaryClassifi...", Reordered Columns, Changed Type3, Added Custom "LeavingD...", Conditional "TenureClassifi...", Changed Type6, Added Custom "'Validatio...", Filtered Rows, and Removed Columns2 "DimT...".

Properties (Middle): Shows the 'Name' field set to 'Performance'. The 'APPLIED STEPS' list includes: Removed Columns2 "DimT...", Changed Type4, Renamed Columns1, Conditional "SelfRating", Changed Type5, Removed Columns2, Renamed Columns2, Conditional "ManagerRati...", Removed Columns3, Renamed Columns3, Changed Type7, Reordered Columns1, Conditional "Environment...", Removed Columns4, Renamed Columns4, Conditional "JobSatisfactio...", Removed Columns5, Renamed Columns5, Conditional "RelationshipS...", Removed Columns6, Renamed Columns6, Conditional "WorkLifeBala...", Removed Columns7, and Removed Columns7.

Properties (Right): Shows the 'Name' field set to 'Performance'. The 'APPLIED STEPS' list includes: Conditional "ManagerRati...", Removed Columns3, Renamed Columns3, Changed Type7, Reordered Columns1, Conditional "Environment...", Removed Columns4, Renamed Columns4, Conditional "JobSatisfactio...", Removed Columns5, Renamed Columns5, Conditional "RelationshipS...", Removed Columns6, Renamed Columns6, Conditional "WorkLifeBala...", Removed Columns7, Renamed Columns7, Changed Type8, Reordered Columns2, Removed Columns8, Custom "JobSatisfactionSo...", Custom "WorkLifeBalance...", and Custom "RelationshipSatisf...".

PROPERTIES

Name

DimEmployee

All Properties

APPLIED STEPS

Promoted Headers

Changed Type

Conditional "AgeClassificat...

Conditional "DistanceFrom...

Removed Columns

Merged Columns

Changed Type1

Conditional "EducationLev...

Removed Columns1

Reordered Columns1

Merged Queries1

Expanded DimEducation

Removed Columns2

Renamed Columns1

Changed Type2

Conditional "DistanceFrom...

Custom "Department"

Merged Queries

Expanded DimJobRole

Removed Columns3

Renamed Columns

Changed Type3

Renamed Columns2

PROPERTIES

Name

DimJobRole

All Properties

APPLIED STEPS

Source

Promoted Headers

Changed Type

Removed Other Columns

Removed Duplicates

Added Index

Reordered Columns

Renamed Columns

Reordered Columns1

Changed Type1

Custom "Department"

Removed Columns

Reordered Columns2

Changed Type2

Renamed Columns1

Removed Duplicates1

Replaced Value

PROPERTIES

Name

DimTraining

All Properties

APPLIED STEPS

Source

Promoted Headers

Changed Type

Removed Other Columns

Removed Duplicates

Added Index

Reordered Columns

Changed Type1

Reordered Columns1

Renamed Columns

Changed Type2

PROPERTIES

Name

DimTime

All Properties

APPLIED STEPS

Source

Promoted Headers

Changed Type

Removed Other Columns

Reordered Columns

Conditional "TenureClassificat...

Reordered Columns1

Conditional "LeavingDate"

Reordered Columns2

Changed Type1

Custom "PromotionDate"

Changed Type2

Reordered Columns3

Renamed Columns

Conditional "1st YearPromoti...

Changed Type3

Renamed Columns1

PROPERTIES

Name

DimEducation

All Properties

APPLIED STEPS

Source

Promoted Headers

Changed Type

Removed Other Columns

Conditional "EducationLevel"

Removed Columns

Removed Duplicates

Changed Type1

Added Index

Reordered Columns

Changed Type2

Renamed Columns

M-Language:

Performance

Display Options

```

37 #Rename Columns2 = Table.RenameColumns(#Removed Columns2,{{"Self", "SelfRating"}}),
38 #Conditional "ManagerRating" = Table.AddColumn(#Renamed Columns2, "ManagerRating", each if [ManagerRating] = "1" then "Unacceptable" else if [ManagerRating] = "2" then "Needs Improvement" else if [ManagerRating] = "3" then "Satisfactory" else "Excellent"),
39 #Removed Columns3 = Table.RemoveColumns(#Conditional "ManagerRating",{"ManagerRating"}),
40 #Renamed Columns3 = Table.RenameColumns(#Removed Columns3,{{"ManagerRating", "ManagerRating"}}),
41 #Changed Type7 = Table.TransformColumnTypes(#Renamed Columns3,{"ManagerRating", type text}),
42 #Reordered Columns1 = Table.ReorderColumns(#Changed Type7,{"PerformanceID", "EmployeeID", "ReviewDate", "EnvironmentSatisfaction", "JobSatisfaction", "RelationshipSatisfaction", "WorkLifeBalance", "SelfRating"}),
43 #Conditional "EnvironmentSatisfaction" = Table.AddColumn(#Renamed Columns3, "EnvironmentSatisfaction", each if [EnvironmentSatisfaction] = "1" then "Very Dissatisfied" else if [EnvironmentSatisfaction] = "2" then "Dissatisfied" else if [EnvironmentSatisfaction] = "3" then "Satisfactory" else "Excellent"),
44 #Removed Columns4 = Table.RemoveColumns(#Conditional "EnvironmentSatisfaction",{"EnvironmentSatisfaction"}),
45 #Renamed Columns4 = Table.RenameColumns(#Removed Columns4,{{"EnvironmentSatisfaction", "EnvironmentSatisfaction"}}),
46 #Conditional "JobSatisfaction" = Table.AddColumn(#Renamed Columns4, "JobSatisfaction", each if [JobSatisfaction] = "1" then "Very Dissatisfied" else if [JobSatisfaction] = "2" then "Dissatisfied" else if [JobSatisfaction] = "3" then "Satisfactory" else "Excellent"),
47 #Removed Columns5 = Table.RemoveColumns(#Conditional "JobSatisfaction",{"JobSatisfaction"}),
48 #Renamed Columns5 = Table.RenameColumns(#Removed Columns5,{{"JobSatisfaction", "JobSatisfaction"}}),
49 #Conditional "RelationshipSatisfaction" = Table.AddColumn(#Renamed Columns5, "RelationshipSatisfaction", each if [RelationshipSatisfaction] = "1" then "Very Dissatisfied" else if [RelationshipSatisfaction] = "2" then "Dissatisfied" else if [RelationshipSatisfaction] = "3" then "Satisfactory" else "Excellent"),
50 #Removed Columns6 = Table.RemoveColumns(#Conditional "RelationshipSatisfaction",{"RelationshipSatisfaction"}),
51 #Renamed Columns6 = Table.RenameColumns(#Removed Columns6,{{"RelationshipSatisfaction", "RelationshipSatisfaction"}}),
52 #Conditional "WorkLifeBalance" = Table.AddColumn(#Renamed Columns6, "WorkLifeBalance", each if [WorkLifeBalance] = "1" then "Very Dissatisfied" else if [WorkLifeBalance] = "2" then "Dissatisfied" else if [WorkLifeBalance] = "3" then "Satisfactory" else "Excellent"),
53 #Removed Columns7 = Table.RemoveColumns(#Conditional "WorkLifeBalance",{"WorkLifeBalance"}),
54 #Renamed Columns7 = Table.RenameColumns(#Removed Columns7,{{"WorkLifeBalance", "WorkLifeBalance"}}),
55 #Changed Type8 = Table.TransformColumnTypes(#Renamed Columns7,{"ManagerRating", type text}, {"EnvironmentSatisfaction", type text}, {"JobSatisfaction", type text}, {"RelationshipSatisfaction", type text}),
56 #Reordered Columns2 = Table.ReorderColumns(#Changed Type8,{"PerformanceID", "EmployeeID", "ReviewDate", "WorkLifeBalance", "RelationshipSatisfaction", "JobSatisfaction", "EnvironmentSatisfaction", "ManagerRating", "SelfRating"}),
57 #Removed Columns8 = Table.RemoveColumns(#Reordered Columns2, {"SalaryClassification", "RoleID", "Employee BusinessTravel", "Employee Department", "Employee JobTitle", "Salary", "StockOptions"}),
58 #Custom "JobSatisfactionSort" = Table.AddColumn(#Removed Columns8, "JobSatisfactionSort", each if [JobSatisfaction] = "Very Dissatisfied" then 1 else if [JobSatisfaction] = "Dissatisfied" then 2 else if [JobSatisfaction] = "Satisfactory" then 3 else "Excellent"),
59 #Custom "WorkLifeBalanceSort" = Table.AddColumn(#Custom "JobSatisfactionSort", "WorkLifeBalanceSort", each if [WorkLifeBalance] = "Very Dissatisfied" then 1 else if [WorkLifeBalance] = "Dissatisfied" then 2 else if [WorkLifeBalance] = "Satisfactory" then 3 else "Excellent"),
60 else if [WorkLifeBalance] = "Dissatisfied" then 2
61 else if [WorkLifeBalance] = "Neutral" then 3
62 else if [WorkLifeBalance] = "Satisfied" then 4
63 else if [WorkLifeBalance] = "Very Satisfied" then 5
64 else null,
65 #Custom "RelationshipSatisfactionSort" = Table.AddColumn(#Custom "WorkLifeBalanceSort", "RelationshipSatisfactionSort", each if [RelationshipSatisfaction] = "Very Dissatisfied" then 1 else if [RelationshipSatisfaction] = "Dissatisfied" then 2 else if [RelationshipSatisfaction] = "Satisfactory" then 3 else "Excellent"),
66 else if [RelationshipSatisfaction] = "Dissatisfied" then 2
67 else if [RelationshipSatisfaction] = "Neutral" then 3
68 else if [RelationshipSatisfaction] = "Satisfied" then 4
69 else if [RelationshipSatisfaction] = "Very Satisfied" then 5
70 else null)
71
72 --#Custom "RelationshipSatisfactionSort"

```

Performance

Display Options ?

```

1 let
2   Source = Csv.Document(File.Contents("C:\Data Science\Projects\DEPI Project (HR)\Datasets\Powe-BI Dataset\Performance.csv"),[Delimiter=",", Columns=33, Encoding=65001, QuoteStyle=QuoteStyle.None]),
3   #"Promoted Headers" = Table.PromoteHeaders(Source, [PromotableCells=true]),
4   #"Changed Type" = Table.TransformColumnTypes(#"Promoted Headers",{{"PerformanceID", type text}, {"EmployeeID", type text}, {"ReviewDate", type date}, {"EnvironmentSatisfaction", type text}, {"JobSatisfaction", type text}, {"Expanded Disincline", type text}, {"Expanded Disincline", type text}, {"Expanded Disincline", type text}, {"Expanded Disincline", type text}, {"Expanded Disincline", type text}}, {"Employee Last Name", type text}, {"Employee First Name", type text}, {"CombinedYearByDisincline", type text}, {"QuoteStyle=None", "None"}),
5   #"Changed Type2" = Table.TransformColumnTypes(#"Changed Columns",{{"Employee_Age", type text}, {"Employee_Education", type text}, {"Employee_YearStartCompany", type text}, {"Employee_YearsInMostRecentRole", type text}, {"Employee_YearsInLeastRecentRole", type text}}, {"Employee_Hiredate", type date}),
6   #"Merged Queries" = Table.NestedJoin(#"Changed Type2",{ "TrainingOpportunitiesWithinYear", "TrainingOpportunitiesTaken"},Disincline,{ "TrainingOpportunitiesWithinYear", "TrainingOpportunitiesTaken"}, "Disincline", JoinKind.LeftOuter),
7   #"Expanded Disincline" = Table.ExpandTableColumn(#"Merged Queries","Disincline","TrainingID","Disincline.TrainingID")),
8   #"Renamed Text" = Table.TransformColumns(#"Expanded Disincline",{{"Employee_EducationField", Text.From, type text}}),
9   #"Removed Columns" = Table.RemoveColumns(#"Trained Text",{ "TrainingOpportunitiesWithinYear", "TrainingOpportunitiesTaken"}),
10  #"Changed Type2" = Table.TransformColumnTypes(#"Removed Columns",{{"Employee_StockOptionLevel", type text}}),
11  #"Merged Query2" = Table.NestedJoin(#"Changed Type2",{ "Employee_JobRole", "Employee_Department", "Employee_BusinessTravel"},DimJobRole,{"JobRole", "Department", "BusinessTravel", "DimJobRole", JoinKind.LeftOuter),
12  #"Expanded DimJobRole" = Table.ExpandTableColumn(#"Merged Query2","DimJobRole","JobRoleID","DimJobRole.JobRoleID"),
13  #"Removed Columns1" = Table.RemoveColumns(#"Expanded DimJobRole",{ "Name", "Employee_Gender", "Employee_Age", "Employee_DistanceFromHome (KM)", "Employee_State", "Employee_Ethnicity", "Employee_Education", "Employee_YearStartCompany", "Employee_YearsInMostRecentRole", "Employee_YearsInLeastRecentRole"}),
14  #"Renamed Columns" = Table.RenameColumns(#"Removed Columns1",{ "Disincline.TrainingID", "TrainingID", "DimJobRole.JobRoleID", "JobRoleID", "Employee_YearStartCompany", "YearstartCompany", "Employee_YearsInMostRecentRole", "YearsinceLastPromotion"}),
15  // Salary Classification
16  #"SalaryClassification" = Table.AddColumn(#"Renamed Columns", "SalaryClassification", each if [Salary] <= 50000 then "Low Salary" else if [Salary] <= 30000 then "Middle Salary" else "High Salary"),
17  #"Reordered Columns" = Table.ReorderColumns(#"Conditional_SalaryClassification",{ "PerformanceID", "EmployeeID", "ReviewDate", "EnvironmentSatisfaction", "JobSatisfaction", "RelationshipSatisfaction", "WorkLifeBalance", "TrainingID", "JobRoleID", "Department", "BusinessTravel", "DimJobRole", "JobRoleID", "DistanceFromHome (KM)", "State", "Ethnicity", "EducationField", "YearStartCompany", "YearsInMostRecentRole", "YearsInLeastRecentRole", "SalaryClassification"}),
18  #"LastDayDate" = Table.TransformColumnTypes(#"Reordered Columns",{{"SalaryClassification", type text}, {"YearstartCompany", Int64.Type}}),
19  // Last Day Date
20  #?Add Custom "LeavingDate" = Table.AddColumn(#"Changed Type2", "LeavingDate", each if [Attrition] = "No" then Edatetime(2099, 12, 5)
21  else if [Attrition] = "Yes" then Date.AddYears([Hiredate], [YearstartCompany])
22  else null),
23  // Experience Level Classification
24  #?Conditional "TenureClassification" = Table.AddColumn(#"Added Custom "LeavingDate"", "ExperiencelevelClassification", each if [YearsatCompany] <= 1 then "Entry-Level" else if [YearsatCompany] <= 4 then "Beginner" else if [YearsatCompany] <= 7 then "Intermediate" else "Senior"),
25  #"Changed Type2" = Table.TransformColumnTypes(#"Conditional_TenureClassification",{{"ExperiencelevelClassification", type text}}),
26  // Validation Column
27  #?Add Custom "ValidationColumn" = Table.AddColumn(#"Changed Type2", "ValidationColumn", each if [ReviewDate] > [Hiredate] and Date.Year([ReviewDate]) <= Date.Year([LeavingDate]) then "Valid" else "Invalid"),
28  #"Filtered Rows" = Table.SelectRows(#"Added Custom "ValidationColumn"" , each ([ValidationColumn] = "Valid")),
29  // Disincline
30  #?Remove Columns2 "DimTime" = Table.RemoveColumns(#"Filtered Rows",{ "Hiredate", "Attrition", "LeavingDate", "YearstartCompany", "ExperiencelevelClassification", "YearsInMostRecentRole", "YearsInLeastRecentRole", "SalaryClassification"}),
31  #"Changed Type2" = Table.TransformColumnTypes(#"Removed Columns2",{{"DimTime"="", {"ValidationColumn", type text}}),
32  #"Renamed Columns1" = Table.RenameColumns(#"Changed Type2",{{"Employee_StockOptionLevel", "StockOptionLevel", "Employee_EducationField", "Educationfield"}),
33  #"Conditional_SelfRating" = Table.AddColumn(#"Renamed Columns1", "SelfRating", each if [SelfRating] = "1" then "Unacceptable" else if [SelfRating] = "2" then "Needs Improvement" else if [SelfRating] = "3" then "Meets Expectations" else "Exceeds Expectations"},
34  #"Changed Types" = Table.TransformColumnTypes(#"Conditional_SelfRating",{{"SelfRating", type text}}, {"SelfRating", type text}),
35  #"Removed Columns2" = Table.RemoveColumns(#"Changed Types",{ "SelfRating"}),
36  #"Renamed Columns2" = Table.RenameColumns(#"Removed Columns2",{{"Self", "SelfRating"}}),
37  #"Conditional_ManagementRating" = Table.AddColumn(#"Renamed Columns2", "ManagementRating", each if [ManagerRating] = "1" then "Unacceptable" else if [ManagerRating] = "2" then "Needs Improvement" else if [ManagerRating] = "3" then "Meets Expectations" else "Exceeds Expectations"},
38  #"Changed Types" = Table.TransformColumnTypes(#"Conditional_ManagementRating",{{"ManagementRating", type text}}, {"ManagementRating", type text})

```

✓ No syntax errors have been detected.

DimEmployee

Display Options ?

```

1 let
2     Source = Csv.Document(File.Contents("C:\Data Science\Projects\DEPI Project (HR)\Datasets\Power BI Dataset\OtherTables.csv"),[Delimiter=";", Columns=23, Encoding=65001, QuoteStyle=QuoteStyle.None]),
3     *Promoted Headers = Table.PromoteHeaders(Source,[PromotableScalars=true]),
4     *Changed Type1 = Table.TransformColumnTypes(*Promoted Headers,{"First Name", type text}, {"Last Name", type text}, {"Gender", type text}, {"Age", Int64.Type}, {"Business Travel", type text}),
5     *Conditional "Age Classification" = Table.AddColumn(*Changed Type1,"AgeClassification", each if [Age] < 18 then "Emerging Adulthood" else if [Age] < 25 then "Young Adulthood" else if [Age] < 35 then "Early Career" else if [Age] < 45 then "Midlife" else "Seniority"},{type text}),
6     *Conditional "Distance From Home Classification" = Table.AddColumn(*Conditional "Age Classification","DistanceFromHomeClassification", each if [*DistanceFromHome (KM)] < 10 then "Short" else if [*DistanceFromHome (KM)] > 10 then "Long"},{type text}),
7     *Removed Columns = Table.RemoveColumns(*Conditional "Distance From Home Classification",{"HireDate","Attrition","YearAtCompany","YearsInCurrentRole","YearsSinceLast Promotion","YearsWithCurrent Manager"}),
8     *Merged Columns = Table.Combine([Table.RemoveColumns(*Changed Type1,"Last Name"),Source.CombineTextByDelimiter(", ",QuoteStyle.None),"Name"]),
9     *Changed Type2 = Table.TransformColumnTypes(*Merged Columns,{{"AgeClassification", type text}, {"DistanceFromHomeClassification", type text}}),
10    *Conditional "Education Level" = Table.AddColumn(*Changed Type2,"EducationLevel", each if [Education] = 1 then "No Formal Qualifications" else if [Education] = 2 then "High School" else if [Education] = 3 then "College" else if [Education] = 4 then "Postgraduate"},{type text}),
11    *Reordered Columns = Table.ReorderColumns(*Removed Columns,{"EmploymentID","Name","Gender","Age","AgeClassification","DistanceFromHome (KM)","DistanceFromHomeClassification","State","Ethnicity","EducationLevel","EducationField","EducationLevel","EducationField","EducationLevel","EducationField","EducationLevel","EducationField","EducationLevel","EducationField","EducationLevel","EducationField","EducationLevel","EducationField","EducationLevel","EducationField"}),
12    *Merged Queries1 = Table.NestedJoin(*Reordered Columns, {"EducationField","EducationLevel"}, EducationField, "EducationLevel", JoinKind.LeftOuter),
13    *Expanded Dinedication = Table.ExpandTableColumn(*Merged Queries1,"Dinedication",{"EducationID"}, {"Dinedication.EducationIDID"}),
14    *Removed Columns1 = Table.RemoveColumns(*Expanded Dinedication,{"EducationField","EducationLevel","EducationField","EducationLevel","EducationField","EducationLevel","EducationField","EducationLevel","EducationField","EducationLevel","EducationField","EducationLevel","EducationField","EducationLevel","EducationField","EducationLevel","EducationField","EducationLevel","EducationField","EducationLevel","EducationField","EducationLevel","EducationField","EducationLevel"}),
15    *Renamed Columns = Table.RenameColumns(*Removed Columns1,{"Dinedication.EducationIDID","EducationIDID"}),
16    *Changed Type3 = Table.TransformColumnTypes(*Renamed Columns1,({"Age", type text}, {"DistanceFromHome (KM)", type text}, {"StockOptionLevel", type text})),
17    *Conditional "Distance From Home Classification" = Table.AddColumn(*Changed Type3,"DistanceFromHomeClassification", each if [*DistanceFromHome Classification] = "Very Long" then 4 else if [*DistanceFromHome Classification] = "Long" then 3 else if [*DistanceFromHome Classification] = "Medium" then 2 else if [*DistanceFromHome Classification] = "Short" then 1 else 0},{type integer}),
18    *Custom "Department" = Table.AddColumn(*Conditional "Distance From Home Classification","Custom", each if [EmployeeID] = "9758 0025" and [Department] = "Technology" then "Sales" else [Department]),
19    *Merged Queries2 = Table.NestedJoin(*Custom "Department",{"JobRole","Custom","BusinessTravel"},DimJobRole,{"JobRole","Department","BusinessTravel"},"DimJobRole",JoinKind.LeftOuter),
20    *Expanded DimJobRole = Table.ExpandTableColumn(*Merged Queries2,"DimJobRole","DimJobRoleID",{type text}),
21    *Removed Columns2 = Table.RemoveColumns(*Expanded DimJobRole,{"Department","BusinessTravel"}),
22    *Renamed Columns = Table.RenameColumns(*Removed Columns2,{{"Custom", "Department"}}),
23    *Changed Type4 = Table.TransformColumnTypes(*Renamed Columns,{{"Department", type text}}),
24    *Renamed Columns2 = Table.RenameColumns(*Changed Type4,{{"DimJobRole.RoleID","RoleID"}})
25 in
26     *Renamed Columns2

```

27 | **W** Retained Columns2


✓ No syntax errors have been detected.

DimJobRole

Display Options 

```
1 let
2     Source = Csv.Document(File.Contents("C:\Data Science\Projects\DEPI Project (HR)\Datasets\Power BI Dataset\OtherTables.csv"),[Delimiter=";", Columns=23, Encoding=65001, QuoteStyle=QuoteStyle.None]),
3     #Promoted Headers = Table.PromoteHeaders(Source, [PromoteAllScalars=true]),
4     #Changed Type = Table.TransformColumnTypes(#Promoted Headers,{{"EmployeeID", type text}, {"FirstName", type text}, {"LastName", type text}, {"Gender", type text}, {"Age", type text}, {"BusinessTravel", type text}, {"Department", type text}, {"JobRole", type text}}),
5     #Removed Other Columns = Table.SelectColumns(#Changed Type,{"BusinessTravel", "Department", "JobRole"}),
6     #Removed Duplicates = Table.Distinct(#Removed Other Columns),
7     #Added Index = Table.AddIndexColumn(#Removed Duplicates, "Index", 1, 1, Int64.Type),
8     #Reordered Columns = Table.ReorderColumns(#Added Index,{"Index", "BusinessTravel", "Department", "JobRole"}),
9     #Renamed Columns = Table.RenameColumns(#Reordered Columns,{{"Index", "RoleID"}}),
10    #Reordered Columns1 = Table.ReorderColumns(#Renamed Columns,{"RoleID", "JobRole", "Department", "BusinessTravel"}),
11    #Changed Type1 = Table.TransformColumnTypes(#Reordered Columns1,{{"RoleID", type text}}),
12    #Custom "Department" = Table.AddColumn(#Changed Type1, "Custom", each if [RoleID] = "38" then "Sales" else [Department]),
13    #Removed Columns = Table.RemoveColumns(#Custom "Department",{"Department"}),
14    #Reordered Columns2 = Table.ReorderColumns(#Removed Columns,{"RoleID", "JobRole", "Custom", "BusinessTravel"}),
15    #Changed Type2 = Table.TransformColumnTypes(#Reordered Columns2,{{"Custom", type text}}),
16    #Renamed Columns1 = Table.RenameColumns(#Changed Type2,{{"Custom", "Department"}}),
17    #Removed Duplicates1 = Table.Distinct(#Renamed Columns1, {"JobRole", "Department", "BusinessTravel"}),
18    #Replaced Value = Table.ReplaceValue(#Removed Duplicates1,"39","38",Replacer.ReplaceText,{"RoleID"}),
19 in
20    #Replaced Value"
```

DimTraining

Display Options 

```
1 let
2     Source = Csv.Document(File.Contents("C:\Data Science\Projects\DEPI Project (HR)\Datasets\Power BI Dataset\DimTraining.csv"),[Delimiter=";", Columns=11, Encoding=65001, QuoteStyle=QuoteStyle.None]),
3     #Promoted Headers = Table.PromoteHeaders(Source, [PromoteAllScalars=true]),
4     #Changed Type = Table.TransformColumnTypes(#Promoted Headers,{{"PerformanceID", type text}, {"EmployeeID", type text}, {"ReviewDate", type date}, {"EnvironmentSatisfaction", Int64.Type}, {"JobSatisfaction", Int64.Type}, {"TrainingID", type text}, {"TrainingOpportunitiesTaken", Int64.Type}, {"TrainingOpportunitiesWithinYear", Int64.Type}}),
5     #Removed Other Columns = Table.SelectColumns(#Changed Type,{"TrainingOpportunitiesTaken", "TrainingOpportunitiesWithinYear"}),
6     #Removed Duplicates = Table.Distinct(#Removed Other Columns),
7     #Added Index = Table.AddIndexColumn(#Removed Duplicates, "Index", 1, 1, Int64.Type),
8     #Reordered Columns = Table.ReorderColumns(#Added Index,{"Index", "TrainingOpportunitiesTaken", "TrainingOpportunitiesWithinYear"}),
9     #Changed Type1 = Table.TransformColumnTypes(#Reordered Columns,{{"Index", type text}}),
10    #Reordered Columns1 = Table.ReorderColumns(#Changed Type1,{"Index", "TrainingOpportunitiesWithinYear", "TrainingOpportunitiesTaken"}),
11    #Renamed Columns = Table.RenameColumns(#Reordered Columns1,{{"Index", "TrainingID"}}),
12    #Changed Type2 = Table.TransformColumnTypes(#Renamed Columns,{{"TrainingOpportunitiesWithinYear", Int64.Type}, {"TrainingOpportunitiesTaken", Int64.Type}}),
13 in
14    #Changed Type2"
```


DimTime

```
1 let
2     Source = Csv.Document(File.Contents("C:\Data Science\Projects\DEPI Project (HR)\Datasets\Power BI Dataset\OtherTables.csv"),[Delimiter=";", Columns=23, Encoding=65001, QuoteStyle=QuoteStyle.None]),
3     #Promoted Headers = Table.PromoteHeaders(Source, [PromoteAllScalars=true]),
4     #Changed Type = Table.TransformColumnTypes(#Promoted Headers,{{"EmployeeID", type text}, {"FirstName", type text}, {"LastName", type text}, {"Gender", type text}, {"Age", Int64.Type}, {"BusinessTravel", type
5     #Removed Other Columns = Table.SelectColumns(#Changed Type,{"YearsAtCompany", "YearsInMostRecentRole", "YearsSinceLastPromotion", "YearsWithCurrManager", "Attrition", "HireDate", "EmployeeID"}),
6     #Reordered Columns = Table.ReorderColumns(#Removed Other Columns,{"EmployeeID", "HireDate", "Attrition", "YearsAtCompany", "YearsInMostRecentRole", "YearsSinceLastPromotion", "YearsWithCurrManager"}),
7     #Conditional ""TenureClassification"" = Table.AddColumn(#Reordered Columns, "ExperienceLevelClassification", each if [YearsAtCompany] <= 1 then "Entry-Level" else if [YearsAtCompany] <= 4 then "Beginner" else
8     #Reordered Columns1 = Table.ReorderColumns(#Conditional ""TenureClassification"", {"EmployeeID", "HireDate", "Attrition", "YearsAtCompany", "ExperienceLevelClassification", "YearsInMostRecentRole", "YearsSince
9     #Conditional "LeavingDate"" = Table.AddColumn(#Reordered Columns1, "LeavingDate", each if [Attrition] = "No" then #Date(2099, 12, 5)
10    else if [Attrition] = "Yes" then Date.AddYears([HireDate], [YearsAtCompany])
11    else null),
12    #Reordered Columns2 = Table.ReorderColumns(#Conditional "LeavingDate", {"EmployeeID", "HireDate", "Attrition", "LeavingDate", "YearsAtCompany", "ExperienceLevelClassification", "YearsInMostRecentRole", "Yea
13    #Changed Type1 = Table.TransformColumnTypes(#Reordered Columns2,{{"LeavingDate", type date}, {"ExperienceLevelClassification", type text}}),
14    #Custom "PromotionDate"" = Table.AddColumn(#Changed Type1, "PromotionDate", each if [YearsAtCompany] = [YearsSinceLastPromotion] then
15    Date.AddYears([HireDate], [YearsAtCompany] - [YearsSinceLastPromotion] + 1)
16    else
17    Date.AddYears([HireDate], [YearsAtCompany] - [YearsSinceLastPromotion])),
18    #Changed Type2 = Table.TransformColumnTypes(#Custom "PromotionDate",{{"PromotionDate", type date}}),
19    #Reordered Columns3 = Table.ReorderColumns(#Changed Type2,{"EmployeeID", "HireDate", "Attrition", "LeavingDate", "YearsAtCompany", "ExperienceLevelClassification", "YearsInMostRecentRole", "YearsSinceLastPro
20    #Renamed Columns = Table.RenameColumns(#Reordered Columns3,{{"ExperienceLevelClassification", "TenureClassification"})),
21    #Conditional "1st YearPromotion"" = Table.AddColumn(#Renamed Columns, "TalentedEmployees", each if [YearsAtCompany] = [YearsSinceLastPromotion] then "Talented" else "Not Talented"),
22    #Changed Type3 = Table.TransformColumnTypes(#Conditional "1st YearPromotion",{{"TalentedEmployees", type text}}),
23    #Renamed Columns1 = Table.RenameColumns(#Changed Type3,{{"TalentedEmployees", "1st YearPromotion"}}),
24    in
25    #Renamed Columns1"
```

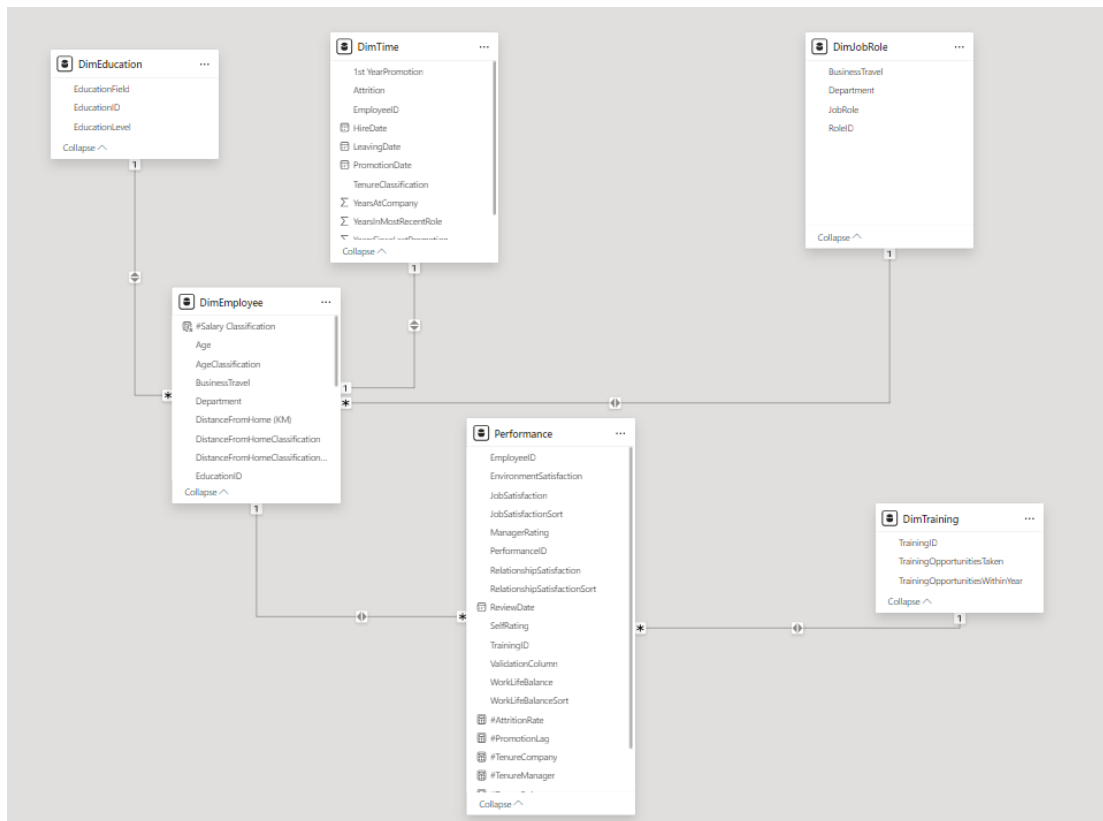
✓ No syntax errors have been detected.

DimEducation

```
1 let
2     Source = Csv.Document(File.Contents("C:\Data Science\Projects\DEPI Project (HR)\Datasets\Power BI Dataset\OtherTables.csv"),[Delimiter=";", Columns=23, Encoding=65001, QuoteStyle=QuoteStyle.None]),
3     #Promoted Headers = Table.PromoteHeaders(Source, [PromoteAllScalars=true]),
4     #Changed Type = Table.TransformColumnTypes(#Promoted Headers,{{"EmployeeID", type text}, {"FirstName", type text}, {"LastName", type text}, {"Gender", type text}, {"Age", Int64.Type}, {"BusinessTravel", type
5     #Removed Other Columns = Table.SelectColumns(#Changed Type,{"EducationField", "EducationLevel"}),
6     #Conditional ""EducationLevel"" = Table.AddColumn(#Removed Other Columns, "EducationLevel", each if [Education] = 1 then "No Formal Qualifications" else if [Education] = 2 then "High School" else if [Educati
7     #Removed Columns = Table.RemoveColumns(#Conditional ""EducationLevel"", {"EducationLevel"}),
8     #Removed Duplicates = Table.Distinct(#Removed Columns),
9     #Changed Type1 = Table.TransformColumnTypes(#Removed Duplicates,{{"EducationLevel", type text}}),
10    #Added Index = Table.AddIndexColumn(#Changed Type1, "Index", 1, 1, Int64.Type),
11    #Reordered Columns = Table.ReorderColumns(#Added Index, {"Index", "EducationField", "EducationLevel"}),
12    #Changed Type2 = Table.TransformColumnTypes(#Reordered Columns,{{"Index", type text}}),
13    #Renamed Columns = Table.RenameColumns(#Changed Type2,{{"Index", "EducationID"}}),
14    in
15    #Renamed Columns"
```

Data Modeling:

- **DimEmployee** has a one-to-many relationship with **DimTime** based on EmployeeID. This means one employee can have multiple records in the DimTime table (though based on the DimTime fields, this might represent different events or periods related to their employment).
- **DimEmployee** has a one-to-many relationship with **DimEducation** based on EducationID. One employee can have one education level recorded in DimEducation.
- **DimEmployee** has a one-to-many relationship with **DimJobRole** based on RoleID. One employee can have one job role recorded in DimJobRole.
- **Performance** has a one-to-many relationship with **DimEmployee** based on EmployeeID. One employee can have multiple performance records over time.
- **Performance** has a one-to-many relationship with **DimTraining** based on TrainingID. One training record can be associated with multiple performance evaluations.



<input type="checkbox"/>	DimEmployee (EducationID)		DimEducation (EducationID)	Active	...
<input type="checkbox"/>	DimEmployee (EmployeeID)		DimTime (EmployeeID)	Active	...
<input type="checkbox"/>	DimEmployee (RoleID)		DimJobRole (RoleID)	Active	...
<input type="checkbox"/>	Performance (EmployeeID)		DimEmployee (EmployeeID)	Active	...
<input type="checkbox"/>	Performance (TrainingID)		DimTraining (TrainingID)	Active	...

Exploratory Data Analysis (EDA):

The EDA phase revealed several key insights about the dataset:

- **Review Data:** The dataset initially contained 6,709 reviews, reduced to 4,982 after cleaning. Many employees provided reviews both before hiring and after leaving, which required filtering to ensure accuracy.
- **Non-Reviewing Employees:** 190 employees did not provide reviews because they left within their first year, leading to the conclusion that employees must stay at least one year to contribute meaningful review data.
- **Talent Identification:** 556 employees received promotions in their first year, flagged as talented.

Question Analysis:

EDA Questions and Answers for HR Analytics

Q: What percentage of employees work overtime vs. regular hours? A: 416 employees work overtime out of the total workforce (28.3%).

Q: What is the overall attrition rate across the organization? A: 237 total employees left the organization.

Q: Is there a relationship between overtime work and employee attrition? A: Yes, there is a direct relationship between overtime and employee departures.

Q: What is the attrition rate for overtime workers vs. non-overtime workers? A:

- Overtime workers: 53.6% attrition rate
- Non-overtime workers: 23.4% attrition rate
- Overtime workers are 2.3 times more likely to leave

Q: What proportion of departing employees were working overtime? A: Overtime workers represent 53.5% of all employees who left (127 out of 237 total departures).

Q: Among overtime workers who left, what was their salary distribution? A: Of the 127 overtime workers who left, 68 (53.5%) had below-average salaries.

Q: How does the combination of overtime + below-average salary affect attrition? A: The combination of overtime work and lower compensation appears to be a significant driver of attrition.

Q: How are employees distributed across different stock option levels (0-3)? A: Employees are distributed across 4 levels (0-3), with specific counts not provided but attrition rates available.

Q: What is the attrition rate for each stock option level? A:

- Level 0 (No stock options): 24.4% attrition rate (highest)
- Level 1: 9.4% attrition rate
- Level 2: 7.6% attrition rate (lowest)
- Level 3 (Highest stock options): 17.6% attrition rate

Q: How does overtime frequency vary by stock option level? A:

- Level 0: 29.1% overtime rate
- Level 3: 34.1% overall overtime rate

Q: Why does Level 3 (highest stock options) have higher attrition than Levels 1 and 2? A: Level 3 employees show concerning attrition when combined with high overtime - departing Level 3 employees had a 66.67% overtime rate.

Q: What percentage of departing Level 3 employees were working overtime? A: 66.67% of departing Level 3 employees were working overtime.

Q: Are there other factors affecting Level 3 retention beyond stock options? A: Yes, the combination of highest stock options with high overtime work appears to create an unexpected attrition risk.

Q: What are the attrition rates by tenure groups? A:

- Entry-level employees: 33% attrition rate
- Beginner employees: 16% attrition rate
- Experienced employees: 7.4% attrition rate

Q: Which tenure group has the highest attrition risk? A: Entry-level employees have the highest attrition risk at 33%.

Q: How much higher is entry-level attrition compared to experienced employees? A: Entry-level employees have more than 4 times the attrition rate of experienced employees (33% vs 7.4%).

Q: Is there a "critical period" where employees are most likely to leave? A: Yes, employees who leave typically have shorter tenure with the company, with entry-level being the most vulnerable period.

Q: What is the attrition rate for each travel frequency group? A:

- No travel: 8% attrition rate (baseline)
- Some travel: 15% attrition rate (1.9× higher than no travel)
- Frequent travel: 24.9% attrition rate (3.1× higher than no travel)

Q: Is there a linear relationship between travel frequency and attrition? A: Yes, there is a clear correlation between increased travel frequency and higher attrition rates.

Q: How much does travel increase attrition risk compared to baseline? A:

- Some travel increases risk by 1.9×
- Frequent travel increases risk by 3.1×

Q: How has the age distribution of new hires changed over time? A: In 2022, there was a significant shift: 111 out of 155 new hires (71.6%) were young adults.

Q: How have average salaries changed year-over-year? A:

- 2018: \$103K average salary
- 2019: \$113K (9.7% increase)
- 2020-2021: \$127K peak (12.4% increase from 2019)
- 2022: \$97K (23.6% decrease from 2021)

Q: How many young adults have been promoted to management roles? A: 48 managers were hired from 2020-2022, but none were young adults, indicating a potential career advancement ceiling.

Q: Are there barriers to advancement for younger employees? A: Yes, the lack of young adult representation in management roles may signal limited advancement opportunities.

Q: How has diversity in hiring changed from 2018-2022? A:

- 2020-2021: 100% of new hires were white
- 2022: 129 out of 155 hires (83.2%) were white

Q: What does this trend indicate? A: There is a concerning trend showing significant lack of diversity in recent hiring, which may create inclusion challenges and could contribute to higher attrition among minority employees.

Q: What caused the significant salary drop in 2022? A: The 24% decrease in average salary coincided with the shift to hiring predominantly younger employees (71.6% young adults in 2022).

Q: Are there salary disparities based on hire date for similar roles? A: Yes, this may create salary disparities among similar roles based on hire date, with employees hired during higher salary periods potentially perceiving inequitable treatment.

Q: How do current salary levels compare to previous years? A: 2022 salary levels (\$97K) may be less competitive for attracting and retaining talent compared to the 2020-2021 peak of \$127K.

Q: What percentage of employees receive first-year promotions? A: 300 out of 545 hires (55%) received first-year promotions.

Q: How has the first-year promotion rate changed over time? A: There has been an increasing pattern of first-year promotions from 2018 to 2021.

Q: Do employees who receive first-year promotions have lower attrition? A: Early promotion likely increases engagement and reduces attrition, and may offset other attrition risk factors.

Data Analysis Expressions (DAX):

```
1 #PromotionLag = AVERAGE(DimTime[YearsSinceLastPromotion])

1 #AttritionRate = CALCULATE(COUNT(DimTime[Attrition]),DimTime[Attrition]="Yes")/COUNTROWS(DimTime)
|
1 #TenureCompany = AVERAGE(DimTime[YearsAtCompany])

1 #TenureManager = AVERAGE(DimTime[YearsWithCurrManager])

1 #TenureRole = AVERAGE(DimTime[YearsInMostRecentRole])

1 #Total Employees = DISTINCTCOUNT(DimTime[EmployeeID])
```

```

1 #Training Utilization Rate =
2 DIVIDE(
3     SUMX('Performance', RELATED(DimTraining[TrainingOpportunitiesTaken])),
4     SUMX('Performance', RELATED(DimTraining[TrainingOpportunitiesWithinYear])),
5     0
6 )
7

```

```

1 #AverageSalary = AVERAGE(DimEmployee[Salary])

```

```

1 #OvertimeRate =
2 DIVIDE(
3     CALCULATE(
4         DISTINCTCOUNT(DimEmployee[EmployeeID]),
5         DimEmployee[OverTime]= "Yes"
6     ),
7     DISTINCTCOUNT(DimEmployee[EmployeeID]),
8     0
9 )

```

```

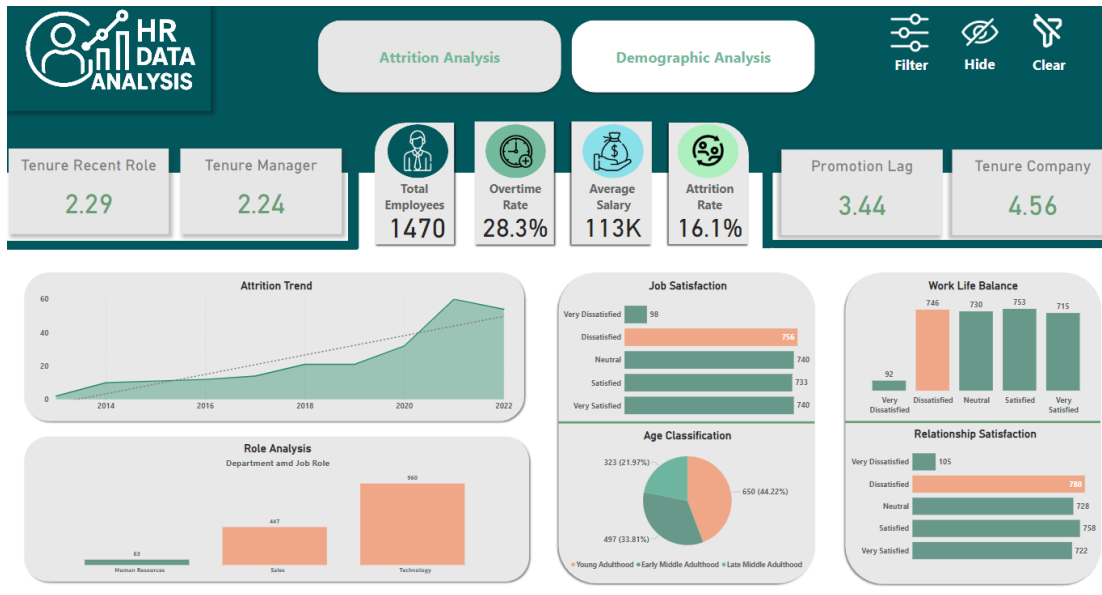
1 #Salary Classification =
2 VAR CurrDept = RELATED(DimJobRole[Department])
3 VAR CurrSal = DimEmployee[Salary]
4 VAR DeptTable =
5     FILTER(
6         ALL(DimEmployee),
7         RELATED(DimJobRole[Department]) = CurrDept
8     )
9 VAR P33 = PERCENTILEX.INC( DeptTable, DimEmployee[Salary], 0.33 )
10 VAR P66 = PERCENTILEX.INC( DeptTable, DimEmployee[Salary], 0.66 )
11 RETURN
12 IF( CurrSal <= P33, "Low Salary",
13     IF( CurrSal <= P66, "Medium Salary", "High Salary" )
14 )

```

Dashboards:

Power BI:

- Attrition Analysis



- Demographic Analysis

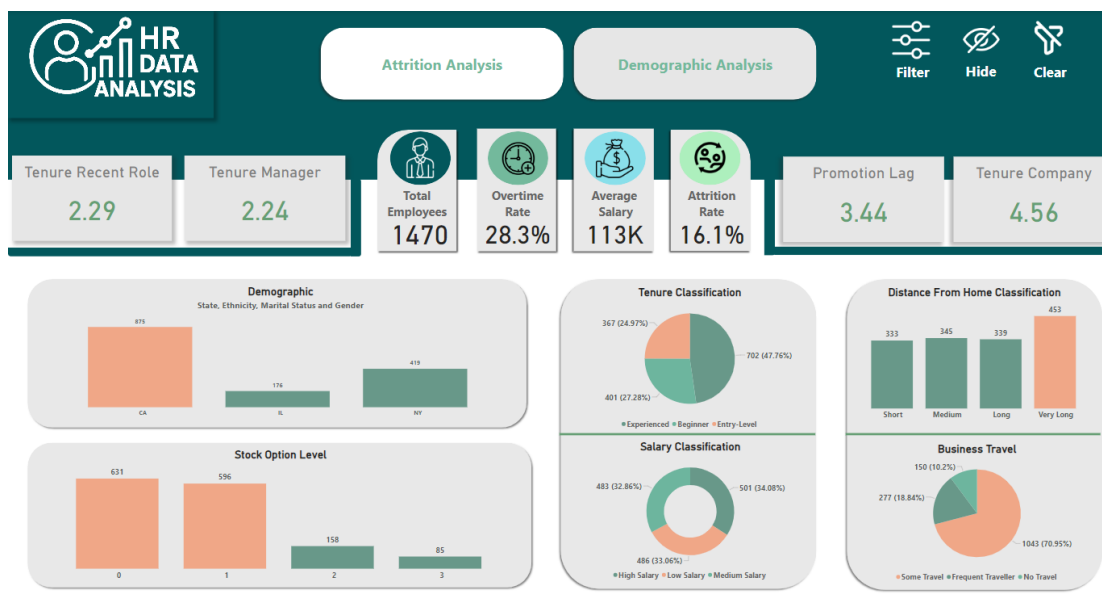
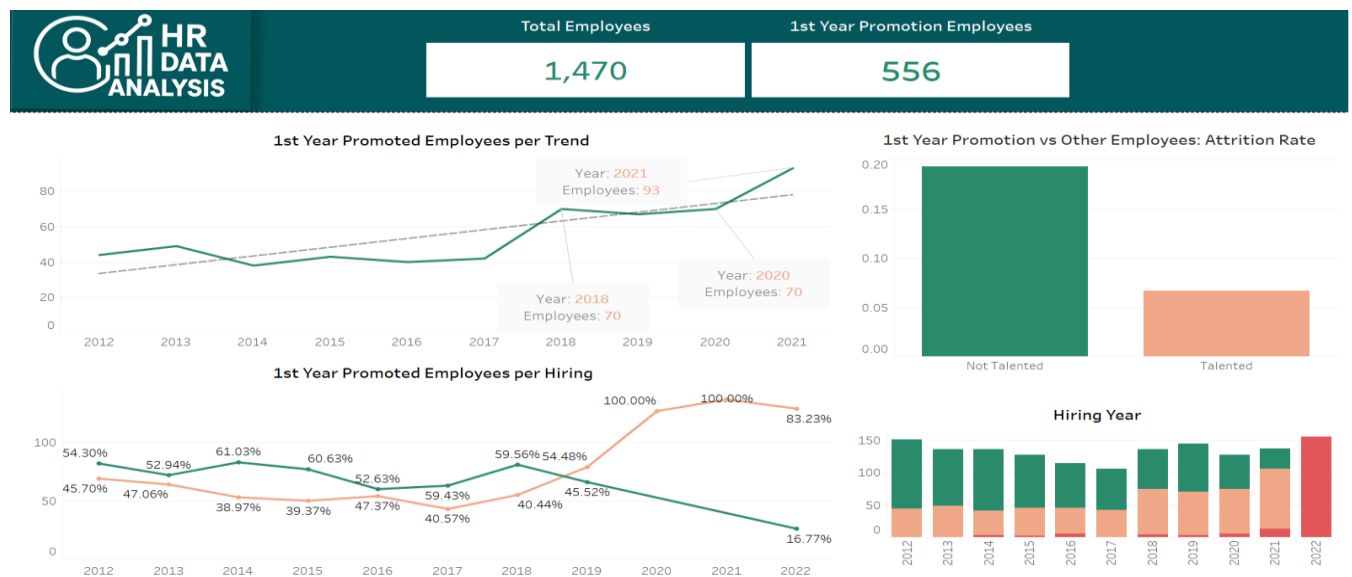


Tableau:

1st Year Promotion Analysis



Insights and Recommendations:

Insights:

1. Overtime and Attrition

- **Direct relationship: Strong positive correlation between overtime and employee departures**
- **Overtime workers who left: 127 out of 416 (30.5%)**
- **Proportion of all departures: Overtime workers represent 53.5% of all employees who left**
- **Comparative attrition rates:**
 - 53.6% for employees working overtime
 - 23.4% for employees not working overtime
 - Overtime workers are 2.3 times more likely to leave
- **Salary impact: Of the 127 overtime workers who left, 68 (53.5%) had below-average salaries**

- **Key insight: The combination of overtime work and lower compensation appears to be a significant driver of attrition**

2. Stock Option Levels and Attrition

- **Level 0 (No stock options)**
 - **Highest attrition rate: 24.4%**
 - **Overtime rate: 29.1%**
 - **Most vulnerable employee group**
- **Level 1**
 - **Low attrition rate: 9.4%**
 - **Significantly better retention than Level 0**
- **Level 2**
 - **Lowest attrition rate: 7.6%**
 - **Most effective stock option level for retention**
- **Level 3 (Highest stock options)**
 - **Unexpectedly high attrition rate: 17.6%**
 - **Overall overtime rate: 34.1%**
 - **Departing Level 3 employees had 66.67% overtime rate**
 - **Anomaly: Despite highest stock options, this group shows concerning attrition when combined with high overtime**

3. Tenure and Attrition

- **Strong inverse relationship: As employee tenure increases, attrition rate decreases significantly**
- **Entry-level employees: 33% attrition rate**
 - **Highest vulnerability in the organization**
 - **More than 4 times the attrition rate of experienced employees**
- **Beginner employees: 16% attrition rate**
 - **Moderate risk, but still more than double the rate of experienced employees**
- **Experienced employees: 7.4% attrition rate**

- Lowest attrition group
- Demonstrates the value of employee retention over time
- Key observation: Employees who leave typically have shorter tenure with the company
- Retention priority: Critical need to focus on entry-level employee engagement and development

4. Business Travel and Attrition

- Direct relationship: Clear correlation between increased travel frequency and higher attrition rates
- No travel: 8% attrition rate (baseline)
- Some travel: 15% attrition rate (1.9× higher than no travel)
- Frequent travel: 24.9% attrition rate (3.1× higher than no travel)
- Pattern implication: Business travel appears to be a significant stressor contributing to employee departures
- Consideration: Travel requirements may need reassessment, especially for roles with already high attrition risk factors

5. Age, Hiring Patterns and Attrition

- 2022 Hiring shift: 111 out of 155 new hires (71.6%) were young adults
 - Coincided with drop in average salary to \$97K from \$127K in 2020-2021
- Management hiring bias: 46 managers hired from 2020-2022, none were young adults
 - Potential career advancement ceiling for younger employees
 - May contribute to higher attrition among younger demographic
- Salary trend correlation with age:
 - 2018-2021: Decreasing young employee hires coincided with rising average salaries
 - 2022: Shift to predominantly young hires coincided with 24% drop in average salary
- Career progression concern: Lack of young adult representation in management roles may signal limited advancement opportunities

6. Diversity and Potential Impact on Attrition

- **Concerning trend: Significant lack of diversity in recent hiring**
 - **2020-2021: 100% of new hires were white**
 - **2022: 129 out of 155 hires (83.2%) were white**
- **Organizational culture implications:**
 - **Potential for homogeneous workplace culture**
 - **May create inclusion challenges for non-white employees**
 - **Could contribute to higher attrition among minority employees**
- **Broader impact: Lack of diversity can limit perspectives, innovation, and company growth**
- **Regulatory and reputation risks: Increasingly important to stakeholders and potential recruits**

7. Salary Trends and Potential Attrition Impact

- **5-year salary progression:**
 - **2018: Average salary \$103K**
 - **2019: Average salary \$113K (9.7% increase)**
 - **2020-2021: Peak average salary \$127K (12.4% increase from 2019)**
 - **2022: Drop to \$97K (23.6% decrease from 2021)**
- **Correlation with hiring practices:**
 - **Significant 2022 decrease coincides with shift to hiring predominantly younger employees**
 - **This may create salary disparities among similar roles based on hire date**
- **Attrition risk: Employees hired during higher salary periods may perceive inequitable treatment**
- **Market competitiveness: 2022 salary levels may be less competitive for attracting and retaining talent**

8. First-Year Promotion and Retention

- **Positive trend: Increasing pattern of first-year promotions from 2018 to 2021**
- **Promotion rate: 300 out of 545 hires (55%) received first-year promotions**

- **Potential retention impact:**
 - Early promotion likely increases engagement and reduces attrition
 - May offset other attrition risk factors
- **Opportunity:** Further analysis needed to correlate first-year promotion with long-term retention

Recommendations:

- Review Overtime Policies

- Minimize excessive overtime, especially for lower-paid employees, to reduce burnout and turnover.

- Improve Compensation and Benefits

- Adjust salary structures and stock option plans to enhance retention across all employee levels.

- Strengthen Onboarding and Entry-Level Support

- Implement mentorship and development programs to increase engagement among new hires.

- Reevaluate Business Travel Requirements

- Reduce frequent travel for roles with high attrition risk to improve work-life balance.

- Support Early Career Advancement

- Continue promoting high-performing employees in their first year to boost motivation and loyalty.

- Create Clear Growth Paths for Young Talent

- Establish transparent promotion opportunities to retain younger employees and avoid career stagnation.

- Promote Diversity in Hiring

- Adopt inclusive recruitment strategies to build a more diverse and representative workforce.

- Ensure Pay Equity Over Time

- Monitor internal salary fairness to prevent dissatisfaction among employees hired during different periods

Conclusion:

This HR data analysis project provided critical insights into the underlying drivers of employee attrition and revealed actionable strategies to enhance retention and increase tenure. Through extensive data cleaning, feature engineering, and exploration using Power BI and Tableau, we identified key factors influencing employee turnover.

The analysis uncovered that overtime work, especially when coupled with below-average compensation, is a major contributor to attrition. Employees working overtime were found to be 2.3 times more likely to leave, and more than half of those who left earned less than the average salary. Similarly, lack of stock options, particularly at Level 0, showed the highest attrition rates, highlighting the importance of financial incentives.

Tenure analysis confirmed that new and entry-level employees face the greatest risk of leaving, with a 33% attrition rate, over 4 times higher than that of experienced employees. Moreover, frequent business travel and limited career progression, especially for young adults, emerged as additional stressors that may drive departures.

Diversity issues also surfaced, as recent hiring trends showed a lack of ethnic variety, raising potential concerns around inclusion and long-term engagement. Furthermore, the significant drop in average salaries in 2022 likely contributed to perceived inequities and higher turnover among newer employees.

On a positive note, the study revealed that early promotions (within the first year) strongly correlate with improved retention, offering a promising strategy to boost morale and reduce attrition.

In summary, this project emphasizes that a holistic, data-driven HR strategy—one that balances compensation, promotion opportunities, work-life balance, and inclusive hiring—is essential for retaining top talent and building a stable, productive workforce.