SQL-Mongo Project –
IBM HR Analytics Employee Attrition & Performance

BUAN 6320

GROUP 8

Venkat Sai Kiran Bachu (VXB220061)

Nikhil Inturi (NXI220006)

Shiva Kumar Reddy Koppula (SXK230064)

Rupa Reddy Nandikonda (RXR230021)

Varnika Reddy Babireddy (VXB230010)

Akshara Reddy Mutyala (ARM230002)

Contents

Relational Data Model

Assumptions/Notes About Data Entities and Relationships

The following 20 database design assumptions are derived from the available data:

1. **Employee Uniqueness:** EmployeeNumber serves as a unique means of identifying each employee..

2. **Attrition Values:** The EmployeeDetails table's Attrition field can only have the values "Yes" or "No."

3. **Business Travel Categories:** Only some categories, such "Travel_Rarely," "Travel Frequently," or "Non-Travel," are included in the EmployeeDetails table's BusinessTravel column.

4. **Education Levels:** Employees' educational attainment is indicated in the EmployeeDetails table's Education column, which is restricted to the specified categories.

5. **EnvironmentSatisfactionLevels:**The EmployeeDetails table's EnvironmentSatisfaction column only includes the categories that are given and relates to the degree of satisfaction.

6. **Gender Values:** Within the EmployeeDetails database, the Gender field can only contain the values 'Male' or 'Female'.

7. **Job Involvement Levels:** The JobInvolvement column in the EmployeeDetails table corresponds to involvement levels and is limited to the provided categories.

8. **Job Roles:** The JobRole column in the EmployeeDetails table corresponds to specific job roles and is limited to the provided categories.

9. **Marital Status Values:** The EmployeeDetails table's MaritalStatus column can only have the values "Single," "Married," or "Divorced."

10. **Performance Ratings:** Performance ratings are reflected in the EmployeeDetails table's PerformanceRating column, which is restricted to the specified categories.

11. **Relationship Satisfaction Levels:** The RelationshipSatisfaction column in the EmployeeDetails table corresponds to satisfaction levels and is limited to the provided categories.

12. **Work-Life Balance Levels:** e WorkLifeBalance column in the EmployeeDetails table corresponds to balance levels and is limited to the provided categories.

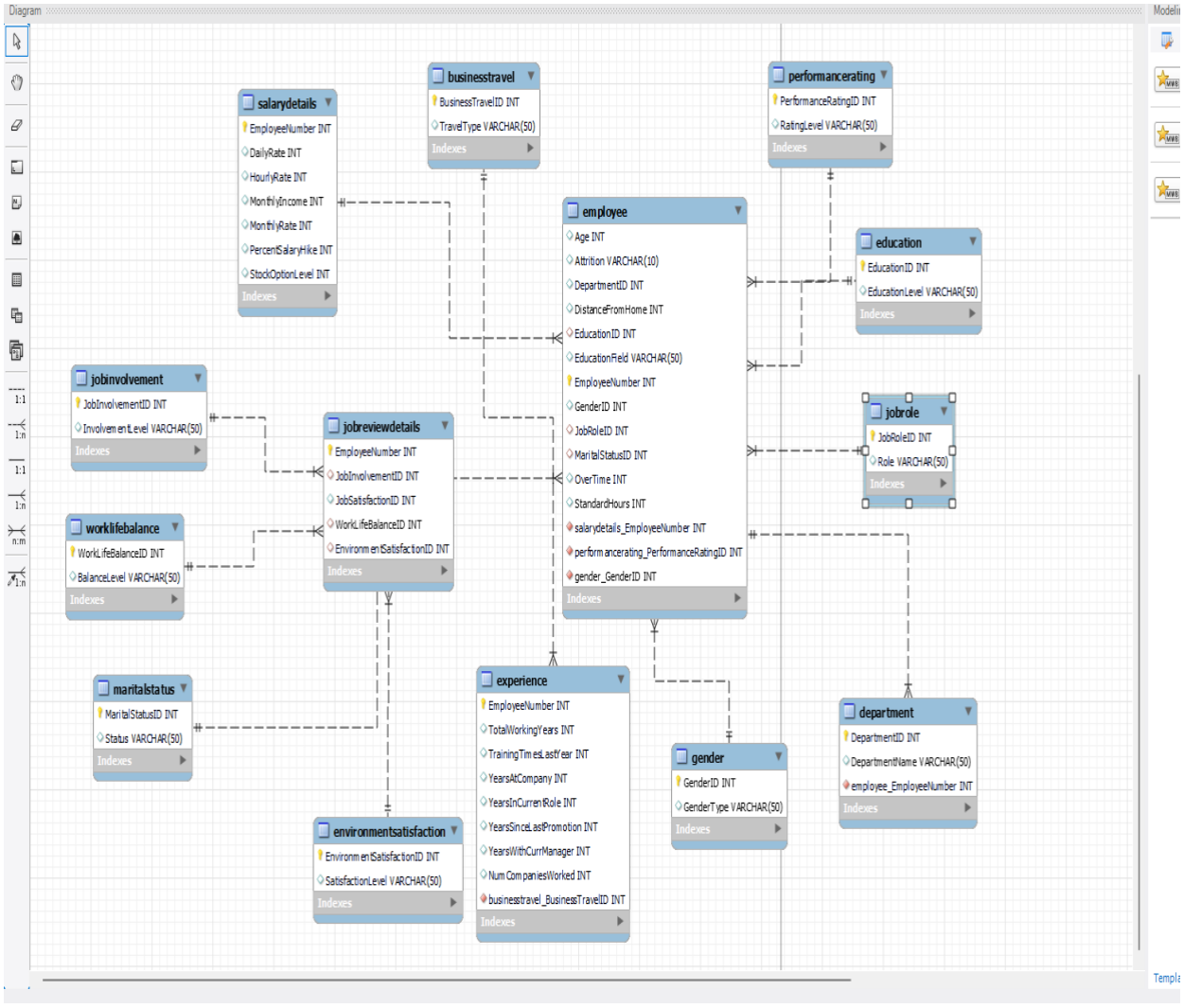
13. **Department Uniqueness:** Each department in the Department table is uniquely identified by its DepartmentID.

14. **Education Levels Uniqueness:** Each education level in the Education table is uniquely identified by its EducationID.

15. **Job Roles Uniqueness:** Each job role in the JobRole table is uniquely identified by its JobRoleID.

16. **No Null Values:** All required fields in the database tables are presumed to have non-null values, particularly those associated with foreign keys

Entity-Relationship Diagram



Original Datase

Assumptions/Notes About Data Set

Revised Table Structure:

1. We used Label Encoding for the dataset to convert the string values (VARCHAR) to numeric (INT). After normalization, these numeric values are used as Primary Keys in the Lookup Tables. For the 'Department' table, we have encoded the three categories: 'Sales,' 'Research & Development,' and 'Human Resources' into 1, 2, and 3, respectively.

2. For the 'Education' table, we have encoded the five categories: 'Below College', 'College,' 'Bachelor,' 'Master,' and 'Doctor' into 1, 2, 3, 4, and 5, respectively.

3. For the 'EnvironmentSatisfaction' table, we have encoded the four categories: 'Low', 'Medium', 'High', and 'Very High' into 1, 2, 3, and 4, respectively.

4. For the 'gender' table, we have encoded the categories 'Male' and 'Female' into M and F, respectively.

5. For the 'JobInvolvment' table, we have encoded the four categories: 'Low,' 'Medium,' 'High,' and 'Very High' into 1, 2, 3, and 4, respectively.

6. For the 'JobLevel' table, we encoded five categories: 'Entry Level', 'Intermediate', 'Senior', 'Manager', and 'Executive' as 1, 2, 3, 4, and 5.

7. We have encoded the three categories in the 'MaritalStatus' table: 'Single' is encoded as 1, 'Married' as 2, and 'Divorced' as 3.

8. We have encoded the categories 'Low', 'Good', 'Excellent', and 'Outstanding' into 1, 2, 3, and 4 for the 'PerformanceRating' table.

9. For the 'RelationshipSatisfaction' table, we have coded the categories 'Low', 'Medium', 'High', and 'Very High' as 1, 2, 3, and 4, respectively.

10. For the 'WorkLifeBalance' table, we have mapped 'Bad', 'Good', 'Better', and 'Best' to 1, 2, 3, and 4.

11. For the ‘JobSatisfaction’ table, we have coded the categories ‘Low’ ,’Medium’,’High’ and ‘Very High’ as 1,2,3,4 respectively

12. For the ‘BusinessTravel’ table, we have coded the categories ‘Non-Travel’ ,’TravelRarely’,and ‘TravelFrequently’ as 1,2,3 respectively

Data in the Database

Education

EducationID INT

EducationLevel

EnvironmentSatisfaction

EnvironmentSatisfactionID

SatisfactionLevel

JobInvolvement

JobInvolvementID

InvolvementLevel

JobRole

JobRoleID

Role

JobSatisfaction

JobSatisfactionID

SatisfactionLevel

MaritalStatus

MaritalStatusID

Status

BusinessTravel

BusinessTravelID

TravelType

Gender

GenderID

GenderType

PerformanceRating

PerformanceRatingID

RatingLevel

RelationshipSatisfaction

RelationshipSatisfactionID

SatisfactionLevel

WorkLifeBalance

WorkLifeBalanceID

BalanceLevel

Department

DepartmentID

DepartmentName

Experience

EmployeeNumber

TotalWorkingYears

TrainingTimesLastYear

YearsAtCompany

YearsInCurrentRole

YearsSinceLastPromotion

YearsWithCurrManager

BusinessTravelID

NumCompaniesWorked

JobReviewDetails

EmployeeNumber

JobInvolvementID

JobSatisfactionID

WorkLifeBalanceID

EnvironmentSatisfactionID

SalaryDetails

EmployeeNumber

DailyRate

HourlyRate

MonthlyIncome

MonthlyRate

PercentSalaryHike

StockOptionLevel

Employee

Age

Attrition

DepartmentID

DistanceFromHome

EducationID

EducationField

EmployeeNumber

GenderID

JobLevelID

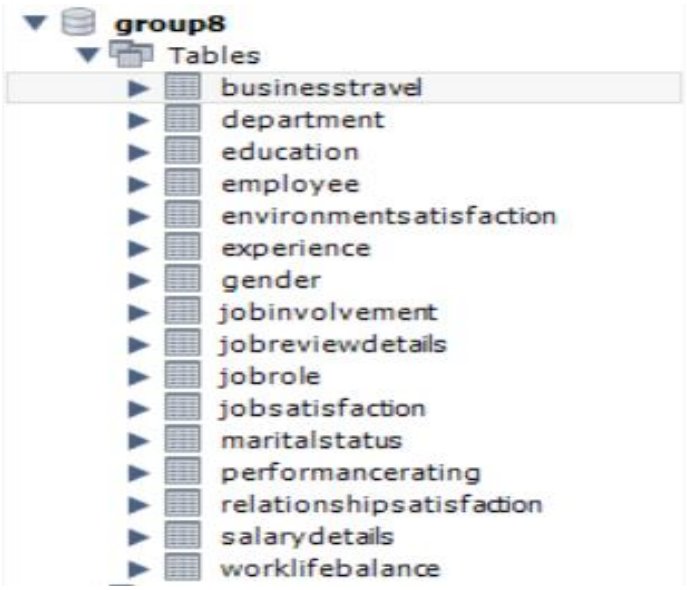
JobRoleID

MaritalStatusID

OverTime

StandardHours

Data Objects

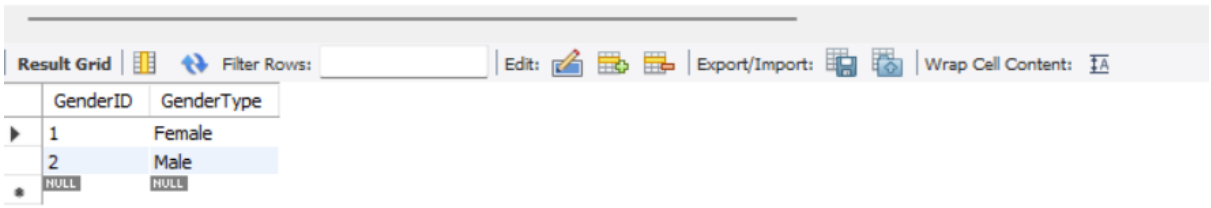


SQL Queries

SQL Query 1

```
CREATE TABLE Gender (
  GenderID INT PRIMARY KEY,
  GenderType VARCHAR(50)
);
```

```
123 •   SELECT * FROM Gender;
```



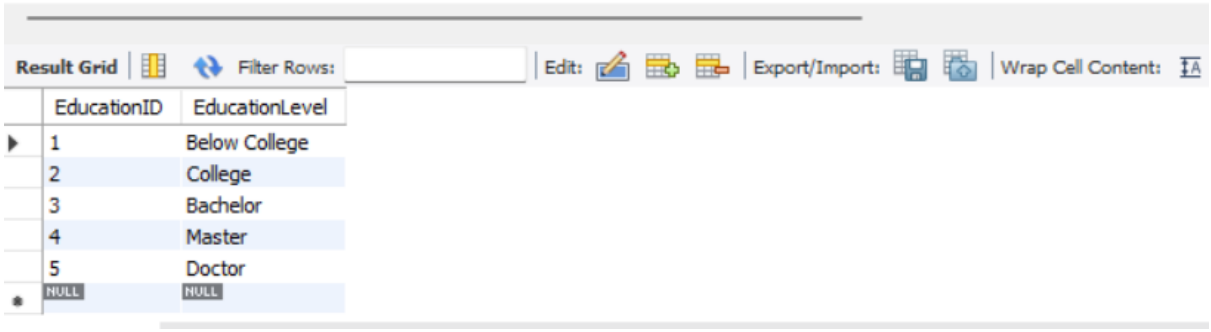
Result Grid | Filter Rows: | Edit: | Export/Import: | Wrap Cell Content:

GenderID	GenderType
1	Female
2	Male
NULL	NULL

SQL Query 2

```
CREATE TABLE Education (
  EducationID INT PRIMARY KEY,
  EducationLevel VARCHAR(50)
);
```

```
128 •   SELECT * FROM Education;
```

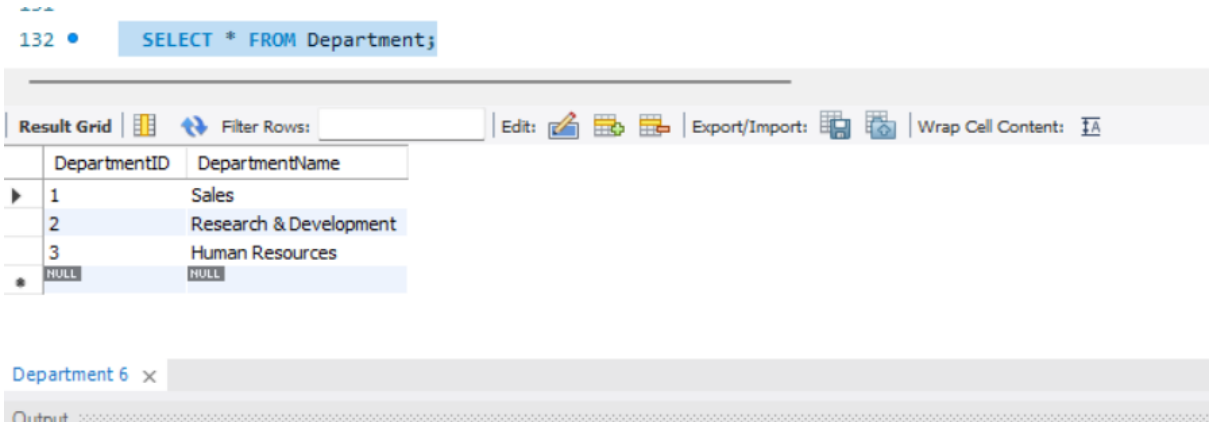


Result Grid | Filter Rows: | Edit: | Export/Import: | Wrap Cell Content:

EducationID	EducationLevel
1	Below College
2	College
3	Bachelor
4	Master
5	Doctor
NULL	NULL

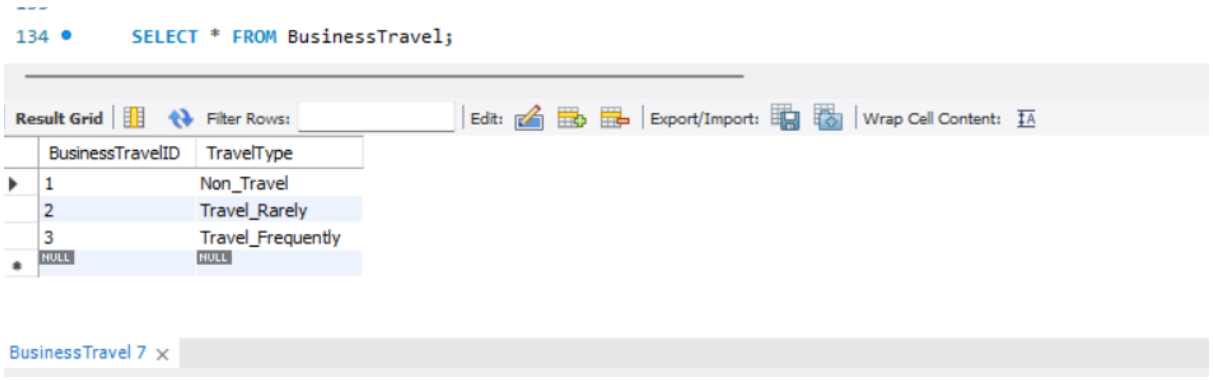
SQL Query 3

```
CREATE TABLE Department (
  DepartmentID INT PRIMARY KEY,
  DepartmentName VARCHAR(50)
);
```



SQL Query 4

```
CREATE TABLE BusinessTravel (
  BusinessTravelID INT PRIMARY KEY,
  TravelType VARCHAR(50)
);
```



SQL Query 5

```
Select * from worklifebalance;
```

SQL Query 6

```
select * from maritalstatus;
```

SQL Query 7

```
select * from group8.jobinvolvement;
```

SQL Query 8

```
select * from group8.relationshipsatisfaction;
```

SQL Query 9

```
select * from group8.performancerating;
```

SQL Query 10

```
select * from group8.environmentsatisfaction;
```

SQL Query 11

```
select * from group8.relationshipsatisfaction;
```

SQL Query 12

```
Create Database Group8;
use Group8;
-- Create Education Table
CREATE TABLE Education (
  EducationID INT PRIMARY KEY,
  EducationLevel VARCHAR(50)
);
-- Create EnvironmentSatisfaction Table
CREATE TABLE EnvironmentSatisfaction (
  EnvironmentSatisfactionID INT PRIMARY KEY,
```

```
  SatisfactionLevel VARCHAR(50)
);
-- Create JobInvolvement Table
CREATE TABLE JobInvolvement (
  JobInvolvementID INT PRIMARY KEY,
  InvolvementLevel VARCHAR(50)
);

-- Create JobRole Table
CREATE TABLE JobRole (
  JobRoleID INT PRIMARY KEY,
  Role VARCHAR(50)
);
-- Create JobSatisfaction Table
CREATE TABLE JobSatisfaction (
  JobSatisfactionID INT PRIMARY KEY,
  SatisfactionLevel VARCHAR(50)
);
-- Create MaritalStatus Table
CREATE TABLE MaritalStatus (
  MaritalStatusID INT PRIMARY KEY,
  Status VARCHAR(50)
);
-- Create BusinessTravel Table
CREATE TABLE BusinessTravel (
  BusinessTravelID INT PRIMARY KEY,
  TravelType VARCHAR(50)
);
-- Create Gender Table
CREATE TABLE Gender (
```

```
  GenderID INT PRIMARY KEY,
  GenderType VARCHAR(50)
);
-- Create PerformanceRating Table
CREATE TABLE PerformanceRating (
  PerformanceRatingID INT PRIMARY KEY,
  RatingLevel VARCHAR(50)
);
-- Create RelationshipSatisfaction Table
CREATE TABLE RelationshipSatisfaction (
  RelationshipSatisfactionID INT PRIMARY KEY,
  SatisfactionLevel VARCHAR(50)
);
-- Create WorkLifeBalance Table
CREATE TABLE WorkLifeBalance (
  WorkLifeBalanceID INT PRIMARY KEY,
  BalanceLevel VARCHAR(50)
);
-- Create Department Table
CREATE TABLE Department (
  DepartmentID INT PRIMARY KEY,
  DepartmentName VARCHAR(50)
);


-- EXPERIENCE
CREATE TABLE Experience(
EmployeeNumber INT PRIMARY KEY,
   TotalWorkingYears INT,
   TrainingTimesLastYear INT,
```

```
   YearsAtCompany INT,
   YearsInCurrentRole INT,
   YearsSinceLastPromotion INT,
   YearsWithCurrManager INT,
BusinessTravelID INT,
NumCompaniesWorked INT,
  FOREIGN KEY (BusinessTravelID) REFERENCES BusinessTravel(BusinessTravelID)
);
-- job review
CREATE TABLE JobReviewDetails(
   EmployeeNumber INT PRIMARY KEY,
   JobInvolvementID INT,
   JobSatisfactionID INT ,
  WorkLifeBalanceID INT,
  EnvironmentSatisfactionID INT,
  FOREIGN KEY (EnvironmentSatisfactionID) REFERENCES
EnvironmentSatisfaction(EnvironmentSatisfactionID),
  FOREIGN KEY (WorkLifeBalanceID) REFERENCES
WorkLifeBalance(WorkLifeBalanceID),
  FOREIGN KEY (JobInvolvementID) REFERENCES JobInvolvement(JobInvolvementID)
);


CREATE TABLE SalaryDetails (
EmployeeNumber INT PRIMARY KEY,
DailyRate INT,
HourlyRate INT,
MonthlyIncome INT,
MonthlyRate INT,
PercentSalaryHike INT,
StockOptionLevel INT
);
```

```
-- Create Employee Table


CREATE TABLE Employee (
  Age INT,
  Attrition VARCHAR(10),
  DepartmentID INT,
  DistanceFromHome INT,
  EducationID INT,
  EducationField VARCHAR(50),
  EmployeeNumber INT PRIMARY KEY,
  GenderID INT,
  JobLevelID INT,
  JobRoleID INT,
  MaritalStatusID INT,
  OverTime VARCHAR(50),
  StandardHours INT,
  FOREIGN KEY (EducationID) REFERENCES Education(EducationID),
  FOREIGN KEY (JobRoleID) REFERENCES JobRole(JobRoleID),
  FOREIGN KEY (MaritalStatusID) REFERENCES MaritalStatus(MaritalStatusID));
```