Day 1 – Power BI Data Modeling (HR Dataset) – 20 Exercises

1. Employee–Department Relationship

You have an **Employees** table and a **Departments** table.
One department can have multiple employees, but some employees are not assigned a department.
How would you model this relationship, and what cross-filter direction would you choose?

2. Historical Salary Changes

Salary changes are stored in a **SalaryHistory** table with EffectiveDate.
Employees can have multiple salary records over time.
How would you model this to correctly calculate *current salary* and *average salary over time*?

3. Active vs Inactive Employees

The **Employees** table has HireDate and TerminationDate.
You want to calculate headcount as of any selected date using a Date slicer.
How should the Date table be related, and what modeling challenge arises?

4. Multiple Date Relationships

An HR table contains HireDate, TerminationDate, and PromotionDate.
You need analytics based on all three dates.
How would you handle relationships without breaking the model?

5. Many-to-Many Scenario

Employees can work on multiple projects, and projects can have multiple employees.
You have **Employees**, **Projects**, and **EmployeeProjects** tables.
How would you model this relationship and avoid ambiguous filtering?

6. Factless Fact Table

You have a table that only records employee training attendance (EmployeeID, TrainingID, Date).

There are no numeric measures.
How would you model and analyze participation rates?

7. Role Changes Over Time

An employee can change job roles multiple times.
Role history is stored in a **JobHistory** table.
How would you model this to analyze role distribution at any point in time?

8. Gender Ratio by Department

Gender is stored in Employees, department in Departments.
Some departments have no employees.
How would you ensure departments with zero employees still appear in visuals?

9. Duplicate Employee IDs

Your HR system has duplicate EmployeeIDs due to re-hiring.
A separate **EmployeeKey** exists as a surrogate key.
How should your model be structured to avoid incorrect aggregation?

10. Attrition Analysis

You need to calculate attrition rate by month.
Terminated employees remain in the Employees table.
How would your data model support correct attrition calculation?

11. Hierarchical Organization Structure

Employees have ManagerID (self-referencing).
You need to analyze headcount by organizational hierarchy.
How would you model and manage this self-relationship?

12. Performance Ratings

Performance ratings are given annually and stored in a separate table.
Some employees miss a year.
How would you model to ensure correct averages without skewing results?

13. Compensation Components

Salary, bonus, and allowances are in separate tables.
You need total compensation analysis by employee and department.
How would you model these tables efficiently?

14. Training Cost Allocation

Training costs are stored at training level, not employee level.
Multiple employees attend the same training.
How would you model cost per employee correctly?

15. Time Intelligence Trap

Your Date table is marked as a Date table.
However, headcount trends show incorrect values.
What modeling mistake could cause this in an HR dataset?

16. Inactive Relationships

An employee belongs to one department at a time, but department changes are tracked.
Only one relationship can be active.
How would you design the model to analyze historical department data?

17. Security-Driven Modeling

HR data requires row-level security by department.
How does your data model design affect RLS performance and accuracy?

18. Star vs Snowflake

You receive HR data in a snowflake schema.
Performance is slow and visuals are complex.
What changes would you make to optimize the model?

19. Contract vs Full-Time Employees

Employment type is stored in a lookup table.
Some employees change employment type over time.
How would you model this to avoid misleading headcount results?

20. Data Granularity Conflict

Attendance data is daily, salary data is monthly, and performance data is yearly.
How would you design the model to avoid incorrect aggregations across visuals?

🔹 **Day 2 – Power BI Data Modeling (HR Dataset) – 20 Exercises**

1. Slowly Changing Dimensions (Type 2)

Employee department changes are stored with StartDate and EndDate.
You need to analyze headcount by department historically and currently.
How would you model this to avoid double-counting employees?

2. Headcount Snapshot Table

HR provides a monthly snapshot table instead of transactional data.
Snapshots overwrite previous records for the same employee.
How would you design relationships to enable time-based analysis?

3. Multiple HR Systems

Employee data comes from two HR systems with different EmployeeIDs.
A master mapping table exists.
How would you model this to maintain a single version of truth?

4. Bridge Table Filtering

You use a bridge table for employee–skill mapping.
Filtering by skill shows incorrect department totals.
What modeling issue could cause this, and how would you fix it?

5. Attrition vs Transfers

Employees who transfer departments should not be counted as attrition.
Transfers are logged in a separate table.
How would you design the model to correctly distinguish both?

6. Dual Grain Conflict

Absence data is at daily grain, payroll data is monthly.
A single Date table is used.
How would you prevent misleading time intelligence calculations?

7. Self-Service Reporting Risk

Business users create reports using the same dataset.
Incorrect joins lead to inflated headcount.
What modeling strategies would you use to protect data integrity?

8. Role-Based Slicer Behavior

Users want a slicer that switches analysis between current role and historical role.
Role history is stored in a fact table.
How would you model this behavior without duplicating tables?

9. Many-to-Many with Time Dependency

Employees can belong to multiple cost centers during overlapping periods.
You must analyze cost allocation accurately.
How would you model this temporal many-to-many relationship?

10. Missing Dimension Records

Some EmployeeIDs in attendance data do not exist in the Employees table.
Deleting records is not allowed.
How would you model this to avoid broken visuals?

11. HR KPI Consistency

Different HR KPIs require different “headcount” definitions.
Some use start-of-period, others average headcount.
How would your data model support multiple definitions?

12. Calendar vs Fiscal Year

HR operates on a fiscal year different from the calendar year.
Both views are required in reports.
How would you model the Date table and relationships?

13. Performance Optimization Scenario

A large HR model has slow visuals due to high-cardinality columns.
What modeling changes would improve performance without losing detail?

14. Security and History

Row-level security must apply to both current and historical employee data.
Managers should only see employees who ever reported to them.
How would you design the model to handle this?

15. Rehire Logic

An employee leaves and rejoins after one year.
They must be counted as one person historically, but two employment periods.
How would you model this distinction?

16. Derived Dimensions

Tenure group (0–1, 1–3, 3–5 years) changes daily.
Should this be a calculated column or a dimension table?
Explain the modeling implications.

17. Composite Keys

A fact table uses EmployeeID + Date as a natural key.
Power BI does not support composite relationships directly.
How would you design around this limitation?

18. HR Analytics at Scale

HR wants analytics across regions with different labor laws.
Some measures apply only to specific countries.
How would you model this to avoid incorrect aggregations?

19. Star Schema Violation

A dimension table is connected to multiple fact tables incorrectly.
This causes circular dependency warnings.
How would you restructure the model?

20. Future-Proofing the Model

New HR modules will be added every quarter.
You want minimal changes to the existing model.
What modeling principles would you apply to make it scalable?