Lesson 5: Data Modeling Basics

1. What is a primary key in a table?

The PK ensures that every row in the table is uniquely identified and one key cannot be repeated in another in terms of PK.

2. Name the two types of table relationships in Power BI.

Many-to one -> this is the most common type. For example, one customer, many orders.

Many-to-many -> this type is also common. For example, many courses and many students

3. How do you create a relationship between two tables in Power BI?

With the help of PK and FK. Also,
Steps to Create a Relationship in Power BI

1. **Open Model View:**
    - Go to **Model View** (left-side panel with a diagram icon).
2. **Drag and Drop Columns:**
    - Drag a column (usually the **foreign key**) from one table onto the related column (the **primary key**) of the other table.
    - Example: Drag CustomerID from **Orders** → CustomerID in **Customers**.
3. **Configure the Relationship:**
    - The **Create Relationship** dialog opens.
    - Verify or adjust:
        - **Cardinality:** (One-to-many, Many-to-many, or One-to-one).
        - **Cross-filter direction:** (Single or Both).
        - **Make this relationship active** (if multiple relationships exist).
4. **Click OK** to establish the relationship.

4. What is a "star schema"?

Star schema's structure is fact tables are surrounded by dimension tables.

5. Which table is typically the fact table in a sales dataset?

Sales table in the sales data set is considered as fact table.

6. Link Sales.csv to Customers.csv using CustomerID (one-to-many).

I did it on PowerBI

7. Why is ProductID in Sales.csv a foreign key? Because, Sales table is dimension table and therefore ProductID from Products is linked as a foreign key. Also in sales table ProductID can be used more than once, because, the product with the same id might be sold many times.
8. Fix a relationship error where ProductID has mismatched data types.

Fixed in PowerBI

9. Explain why a star schema improves performance.

Star schema is more beneficial with its simplicity and denormalized structure with central fact tables surrounded by dimension tables improves query performance by reducing the number of joins required.

10. Add a new column TotalSales in Sales (Quantity * Price from Products).

Done in powerBI

11. Optimize a model with circular relationships—how would you resolve it?
    Simplify the model data, change your model to star schema where in the center fact tables surrounded by dimension tables.
    Break the loop -> redesign the model which is causing looping
12. Create a role-playing dimension for OrderDate and ShipDate.

It is done in PowerBI

13.Handle a many-to-many relationship between Customers and Products.

There is a bridge table between Customers and Products and with that many to many relationship is handled

14.Use bidirectional filtering sparingly—when is it appropriate?

Useful cases:
Let’s say students and courses are linked with Enrollment table and we want courses to filter students and students to filter courses.

Another use case is two fact tables are sharing one dimension. Let’s say Sales and Budget both tables are linked to Date and we want a slice from Date to affect both tables.

15. Write DAX to enforce referential integrity if a CustomerID is deleted.
There is not exact way to enforce referential integrity if cusomerID is deleted. But, it can be done with writing code if referential integrity happens it raises warning.