

## Data Modeling Basics - Homework Answers

### What is a primary key in a table?

A primary key is a column (or combination of columns) that uniquely identifies each row in a table. For example, CustomerID in Customers.csv is a primary key.

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

1. One-to-Many (1:\*) : One record in one table matches multiple records in another.
2. Many-to-Many (\*:\*) : Multiple records in one table match multiple records in another.

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

Go to Model view, drag a column from one table to the matching column in another table, check the cardinality, and confirm.

### What is a "star schema"?

A star schema is a data model with a central fact table connected to multiple dimension tables, forming a star-like layout. It separates measurable data (facts) from descriptive data (dimensions).

### Which table is typically the fact table in a sales dataset?

The Sales table is the fact table because it contains measurable transaction data such as Quantity and OrderDate, along with foreign keys to dimensions.

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

In Power BI, create a relationship: Sales[CustomerID] → Customers[CustomerID] with cardinality One-to-Many, where Customers is the 'one' side.

### Why is ProductID in Sales.csv a foreign key?

Because it refers to a record in another table (Products.csv) and connects sales transactions to product details without uniquely identifying a sales record.

### Fix a relationship error where ProductID has mismatched data types.

In Power Query, change ProductID column in both Sales and Products tables to the same data type (Whole Number), then re-establish the relationship.

### Explain why a star schema improves performance.

A star schema reduces join complexity, uses smaller dimension tables for better compression, and simplifies query logic, resulting in faster performance.

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

DAX formula:

TotalSales = Sales[Quantity] \* RELATED(Products[Price])

### **Optimize a model with circular relationships—how would you resolve it?**

Remove unnecessary relationships, use bridge tables, or set one relationship to inactive and activate it in DAX with USERELATIONSHIP().

### **Create a role-playing dimension for OrderDate and ShipDate.**

Duplicate the Dates table: one linked to Sales[OrderDate] and another linked to Sales[ShipDate]. Both reference the same underlying date values but serve different purposes.

### **Handle a many-to-many relationship between Customers and Products.**

Create a bridge table containing unique CustomerID–ProductID combinations and link Customers → Bridge → Products.

### **Use bidirectional filtering sparingly—when is it appropriate?**

Only when necessary, such as for security roles or when filters must flow both ways for correct results. Overuse can cause performance issues and ambiguity.

### **Write DAX to enforce referential integrity if a CustomerID is deleted.**

```
ValidSales = CALCULATE(  
    COUNTROWS(Sales),  
    NOT(ISBLANK(RELATED(Customers[CustomerID])))  
)
```