# Homework

1. What is a primary key in a table?

A primary key is a column (or a set of columns) that uniquely identifies each row in a table. It ensures there are no duplicate or null values. When a table has a primary key, it can be used to join with other tables through foreign keys, allowing data to be related across table

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

The two types of table relationships in Power BI are one-to-many (1:\*) and **many-to-one (\*:1)**. There's also **many-to-many (\*:\*)** but it's less commonly used and requires special attention

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

To create a relationship between two tables in Power BI, go to the **Model view,** drag a column (usually a key) from one table to a matching column in another table

1. What is a "star schema"?

A star schema is a data model where a central **fact table** is connected to multiple **dimension tables.** The structure looks like a star, with the fact table in the center and dimension tables around it. It is simple, fast, and ideal for reporting and analysis

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

In a sales dataset, the **Sales** table is usually the fact table. It stores measurable transactions like quantity sold, order dates, and links to related dimensions such as Customers and Products

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

**Model view**, then drag CustomerID from the **Customers** table to CustomerID in the **Sales** table. This creates a **one-to-many** relationship: one customer can have many sales

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

In Sales.csv, **ProductID** is a foreign key because it refers to the **ProductID** in the Products.csv table. It creates a link between each sale and the product being sold, allowing us to pull product details like name or price

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

**Power Query Editor,** checkthe **ProductID** columninboththe **Sales** and **Products** tables, and make sure they are the **same data type -** usually **Whole Number.** After correcting, click **Close & Apply,** then re-create therelationship

1. Explain why a star schema improves performance

A star schema improves performance because it uses a simple structure with a central fact table linked to dimension tables. This reduces the number of joins, keeps relationships straightforward (one-to-many), and allows Power BI’s engine to compress and query data more efficiently

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

To add a **TotalSales** column in the **Sales** table, multiply Quantity by the related Price from the **Products** table

1. Optimize a model with circular relationships—how would you resolve it?

Break the loop by deleting one relationship. Use a **bridge table** or DAX like TREATAS if needed. Keep relationships **single-directional** for better performance

1. Create a role-playing dimension for OrderDate and ShipDate

Create one Date table. Duplicate it in the model and rename as **OrderDate** and **ShipDate**. Link each to the corresponding column in the Sales table. Now both dates use the same data but act as separate dimensions

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

Create CustomerProduct with CustomerID and ProductID. Link both to their main tables with **one-to-many** relationships. Avoid bidirectional filters unless absolutely necessary

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

Use bidirectional filtering only when two dimension tables need to filter each other, like in many-to-many relationships or complex slicer scenarios. Avoid it if performance or ambiguity is a conce

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

(MissingCustomerCheck :=

IF (

NOT ISBLANK(Sales[CustomerID]) &&

ISBLANK(RELATED(Customer[CustomerID])),

1,

0

))