**Statistics**

1. A complain was raised that the boys in the school were underfed.. Average weight of boys is 32kg with standard deviation 9kg.

A sample of 25 boys were taken from the school and the average was found to be 29.5kg.

With alpha=0.05 we need to check whether the complaint is true or not.

1. You have two coins in your hand. Out of the two coins, one is a real coin (heads and tails) and the other is a faulty coin with **tails on both sides.**

You are blindfolded and forced to choose a random coin and then toss it in the air. The coin lands with tails facing upwards. **Find the probability that this is the faulty coin.**

* Print the probability of the coin landing tails.
* Print the probability of the coin being faulty and landing tails.
* Print and solve for the probability that the coin is faulty, given it came down on tails.

1. A national consumer magazine reported the following correlations.

The correlation between car weight and car reliability is -0.30.

The correlation between car weight and annual maintenance cost is 0.20.

Which of the following statements are true?

I. Heavier cars tend to be less reliable.   
II. Heavier cars tend to cost more to maintain.   
III. Car weight is related more strongly to reliability than to maintenance cost.

(A) I only   
(B) II only   
(C) III only   
(D) I and II only   
(E) I, II, and III

1. Molly earned a score of 940 on a national achievement test. The mean test score was 850 with a standard deviation of 100. What proportion of students had a higher score than Molly? (Assume that test scores are normally distributed.). Use Z table from <http://www.z-table.com/>
2. What is central limit theorem?
3. What is p-value?
4. What is Normal Distribution. How would you detect an outlier using Normal Distribution?
5. How is correlation different from covariance.

Python.

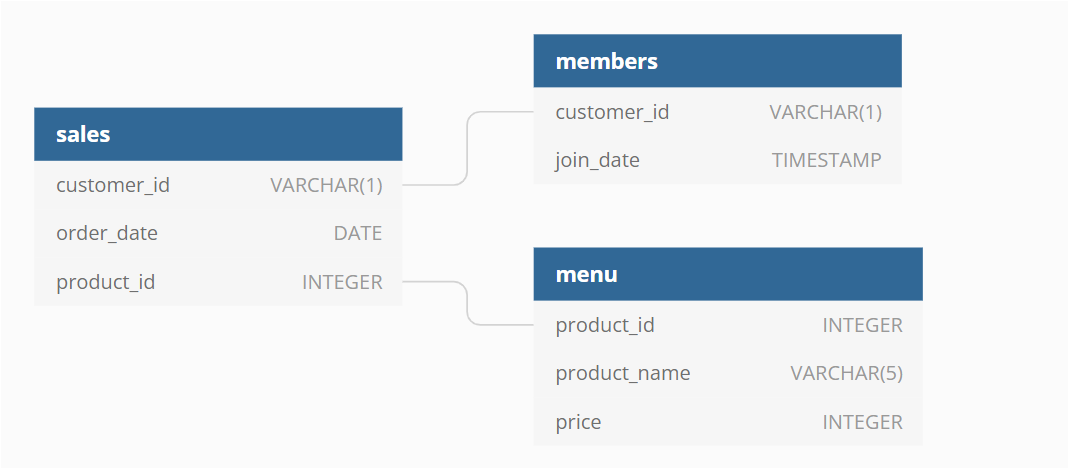
Load the dataset - <https://raw.githubusercontent.com/SonerYldrm/datasets/main/groceries.csv>

* The price column contains some missing values. How can we replace these missing values with the average price of the product? Please note that we should not fill the missing values with the average value of the price column. The missing values for apples need to be filled with the average price of apples, and so on.
* How can we find the difference between the maximum and minimum price of each product?
* Find the total sales\_quantity for each product group.
* When unit = "pcs”. Look at the product description column. It’s a concatenation of product name and weight.

Create a new table where unit = pcs and add a new column called weight.

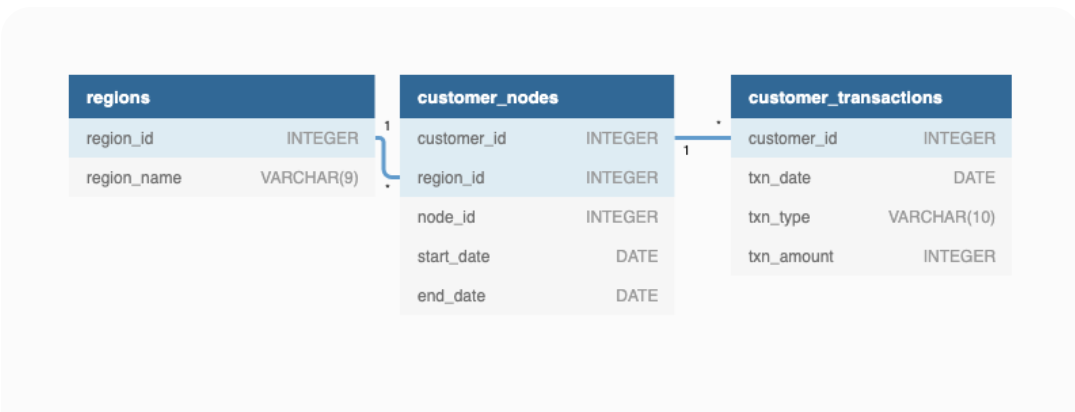
Example the “butter-0.25” weighs 0.25 kg.

**SQL 1**



1. What is the total amount each customer spent at the restaurant?
2. How many days has each customer visited the restaurant?
3. What was the first item from the menu purchased by each customer?

**SQL 2**



1. How many unique nodes are there on the Data Bank system?
2. What is the number of nodes per region?
3. How many customers are allocated to each region?