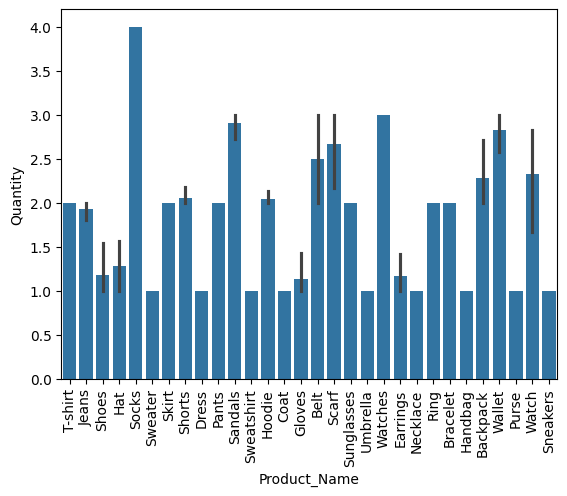
Data preprocessing –

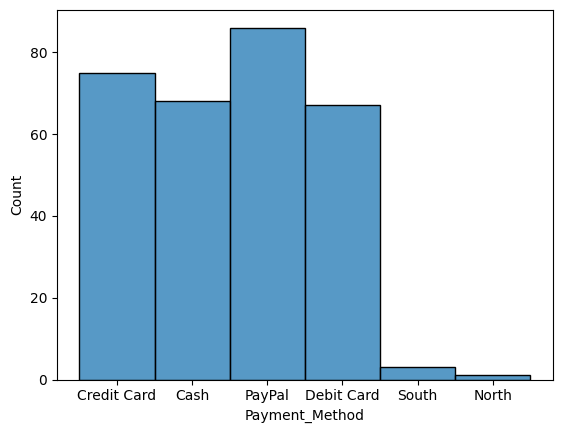
The column “Quantity” has 35 missing values. The columns “Customer\_Satisfaction” and “Delivery\_Time” have 1 missing value. The column “Region” has 7 missing values. Thus, we need to use the imputation technique to replace these null values. “Delivery\_Time” and “Region” are categorical variables. Thus, we will use the mode to replace the null values in these columns. “Quantity” and “Customer\_Satisfaction” are numerical variables. There is no particular range in which the data is concentrated in for both the columns. Thus, we can use median to replace the null values in these columns.

Answer 1 –



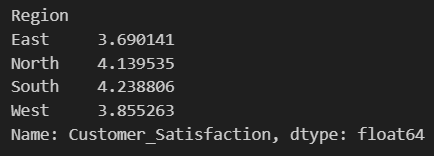
Thus, “Socks” is the highest-selling product quantity-wise in the given dataset. “Watches” is the second highest-selling product quantity-wise in the given dataset.

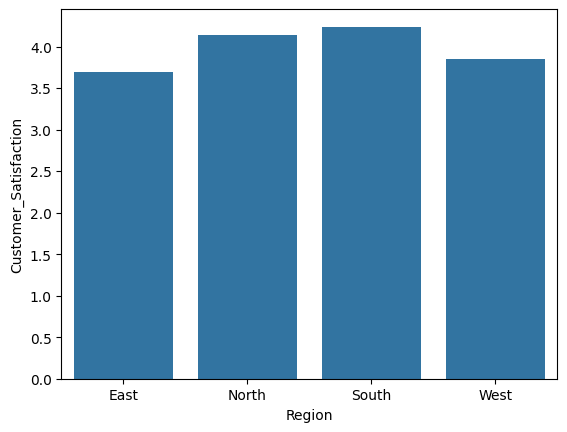
Answer 2 –



PayPal is the most commonly used payment method in the given dataset.

Answer 3 –





The above image and graph show the average customer satisfaction rating for different regions in the given dataset.