**Case Studies**

1. In a Sound and stereo store in Chicago the manager wants to know is there any relationship between the number of weekend television commercials shown and sales at the store at the following week. The sample data of the store is provided in the spread sheet tab. Use that to solve the below problem
2. Find the covariance between the variables and tell us is there any relationship between them
3. Is the Sales and number of commercials correlated? If yes - Positive/ Negative
4. In the above problem the Manager has been provided the data with the sales amount measured in Kilo $ (1000$). But now the sales team provides a new data in Tab 2 with the amount in $. Does this unit of measurement impact the covariance and correlation calculated in problem1. The data is in Tab2

Calculate the new covariance and correlation and confirm

1. The sales report about the pharmaceutical company in million $ for the 21 states in US has been provided in the spread sheet
2. Provide the five number summary of the box plot
3. Compute the lower and upper limits
4. Does the data contain any outlier ?
5. Ohio state has the highest sales at $14,138 million. Suppose a data entry error has been made as $ 41,138 million would this been identified as an outlier and corrected
6. Out of 21 states provided what percentile states are above the Washington
7. What is the 85th percentile ?
8. In an IPL auction a team wants to buy a foreign player who is an all-rounder to balance their side. The team has the budget to spend on only one player and they are looking for a player who could help them in their batting department. The franchise now has the data with the runs scored by player A and player B in their last 20 matches and are confused on their selection

Can you help them with your analysis? You know what needs to be done

1. With the same data as above - Consider the data provide for the 20 matches as a population data, pick random samples from the population with sample size (n) as 7 and let the number of samples be 10. Compute mean of individual samples and consider this a new data in a table. Now compute the Mean and standard deviation of this tabular data
2. Do you find any difference between the population mean and the mean of sample means.
3. Is the population standard deviation higher or lower than the new standard deviation you calculated

**Example for Problem 5 – For reference workout**

Lets assume a population as (Test data) - 15,5,20,45,85,65,30,10,20

**Picking Random 3 samples of size (n) 5 means**

5, 20, 10, 85, 45

10, 65, 20, 30, 45

15, 5, 20, 45, 85

**Compute their individual mean**

5+20+10+85+45 = 33

10+65+20+30+45= 34

15+5+20+45+85 = 34

**New data in a new Table**

33

34

34

Now compute the Mean and Standard deviation of the above final data and compare this with the actual population data