THINK PAIR SHARE

PGP – DSBA

1. **Megha Bhan**
2. **Jishwanth S V**
3. **Arnab Ghosal**

Case Snippet 1: 10 Marks

1b) BeautyPlus has launched a new face wash in the market. The brand has planned to target the customers of the rival brand Dove. It has been found through several research studies that 8% of customers in the market are Dove buyers and that a high proportion, 30% of them shop at Target which is a personal care retailer. It is also known that 25% of customers of personal care products shop at Target. The brand manager of BeautyPlus is planning to conduct a sampling programme of its product at Target. However, she thinks that doing this will only be viable if at least 15% of Target shoppers are Dove buyers. Should the brand manager go ahead with the programme?

Solution:

We know the following data from the question

Probability that a customer gets Dove face wash

P(Dove) = P(D) = 8% = 0.08

Probability that a Dove buyer shops at Target

P(Target/Dove) = P(T/D) = 30% = 0.3

Probability that a customer of personal care products shops at Target.

P(Target) = 25% = 0.25

We need to find the probability that a Target shopper is a Dove buyer So we have a great tool i.e. Bayes' Theorem

$$P\left(\frac{Dove}{Target}\right) = P\left(\frac{D}{T}\right) = \frac{P\left(\frac{T}{D}\right) * P(D)}{P(T)}$$

= (0.3) x (0.08) / (0.25)

= 0.096

Conclusion: -
By using Bayes' Theorem, we found that only 9.6% of the Target shoppers are Dove buyers. So statistically we can conclude that the programme may not be viable since required percentage should be at least 15% (as stated in the Question)

Case Snippet 2: 10 Marks

2b) Two Fast Moving Consumer Goods (FMCG) companies are in close competition these days to become the undisputed leader of the industry. Both the companies produce and sell multiple lines of products. The sales performance of both the companies in the context of selling a non-alcoholic beverage is given below:

Company A

*Mean Sales (One Year Average) – **5000** units*

*Standard Deviation – **500** units*

Company B

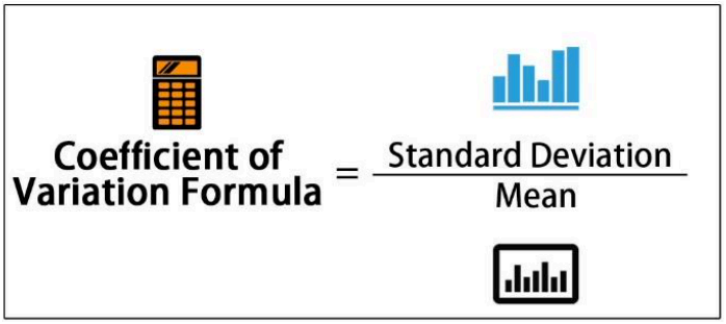
*Mean Sales (One Year Average) – **7500** units*

*Standard Deviation – **1000** units*

Compute Coefficient of Variation (CV) for both the companies and comment on the result.

Solution:

The Coefficient of variation (CV) is the ratio of the standard deviation to the mean. It is generally expressed as a percentage.



For Company A

Coefficient of Variation = (*Standard deviation*/*Mean*) ∗ 100%

$$CV = \frac{\sigma}{\mu} * 100$$

$$= \frac{500}{5000} * 100\ \%$$

$$= 10\%$$

For Company B

Coefficient of Variation = (*Standard deviation*/*Mean*) ∗ 100%

$$= \frac{1000}{7500} * 100\ \%$$

$$= 13.33\%$$

Conclusion: -
The higher the coefficient of variation, the greater the level of dispersion around the mean. The lower the value of the coefficient of variation, the more precise the estimate. Based on the above results, statistically we can conclude that the estimate for Company A is more precise than that of Company B.

Case Snippet 3: 10 Marks

3a) The table shows the annual return of stocks under different sectors. Is there any evidence that the stock return depends on sector like industrial or consumer or service? (Ref attachment stocks.csv)

Solution:

Let't do some EDA to understand the Stocks Data Head Of the Data:-

First 5 rows are below

Sector	Stock_Return
Consumer	0.0632
Consumer	0.1473
Consumer	0.1195
Consumer	0.1236
Consumer	0.1028

Information about Data :-

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 90 entries, 0 to 89
Data columns (total 2 columns):
 #   Column        Non-Null Count  Dtype  
---  ------        --------------  -----  
 0   Sector        90 non-null     object 
 1   Stock_Return  90 non-null     float64
dtypes: float64(1), object(1)
memory usage: 1.5+ KB
```

Dataset is having 90 rows & 2 columns, there are no null values present

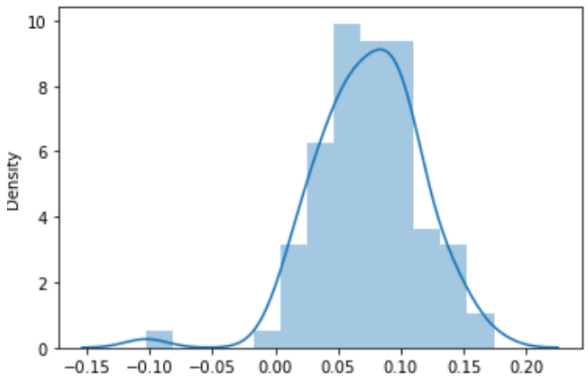
Description of the dataset: -

	Sector	Stock_Return
count	90	90
unique	3	NaN
top	Service	NaN
freq	30	NaN
mean	NaN	0.07
std	NaN	0.04
min	NaN	-0.1
25%	NaN	0.05
50%	NaN	0.08
75%	NaN	0.1
max	NaN	0.17

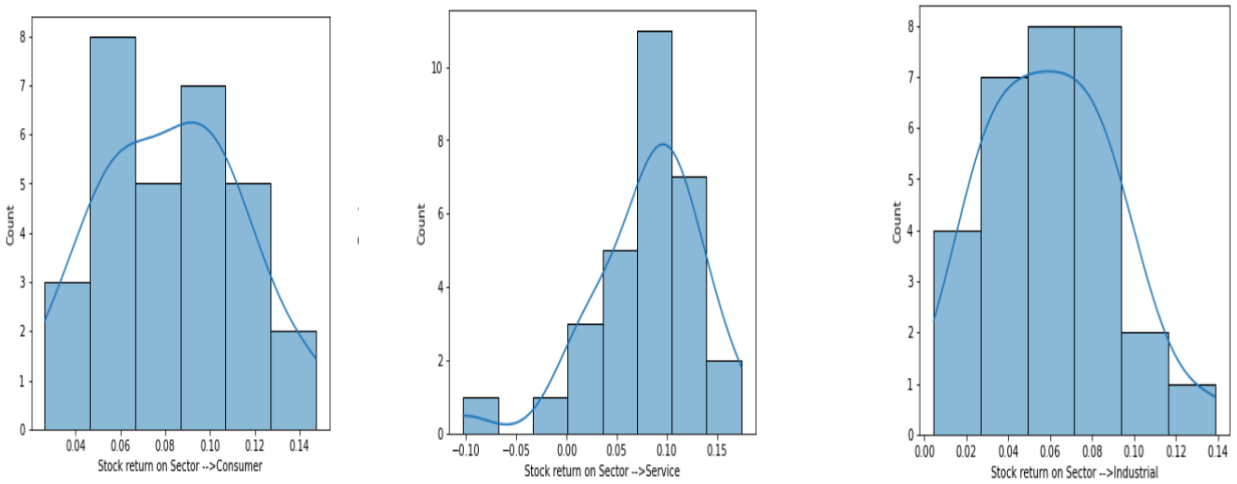
Above table is five-point summary of the give data. There are three different sectors i.e. , Consumer, Service and Industrial. The median of stock return is 0.08 and its mean and standard deviation is 0.07 and 0.04 respectively.

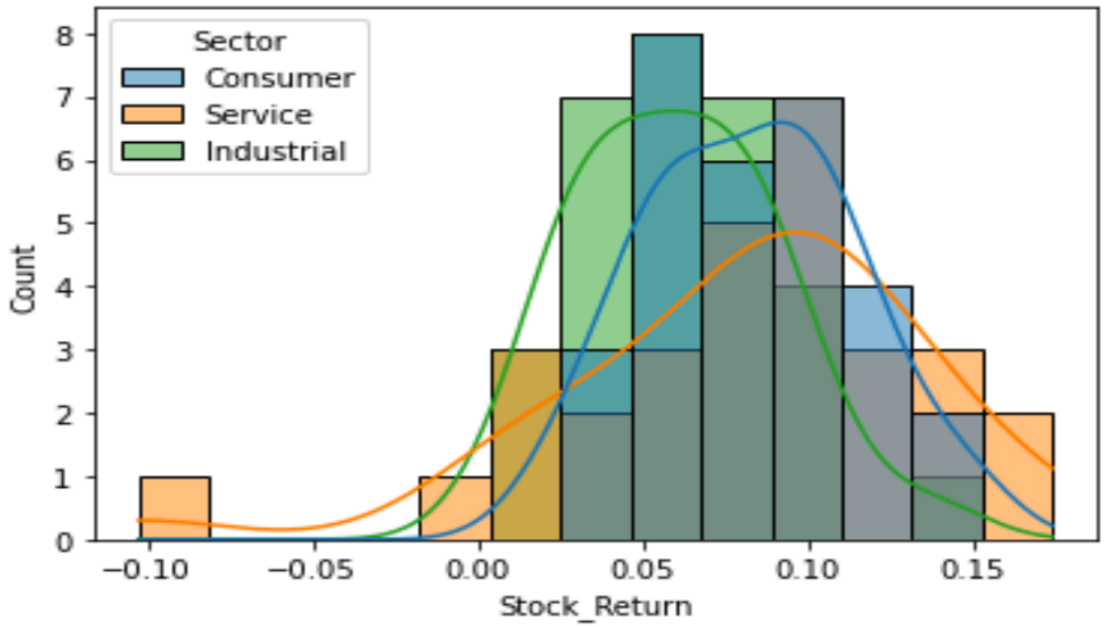
Various graphs had been studies before interpretation.

• Whole Stocks Distribution

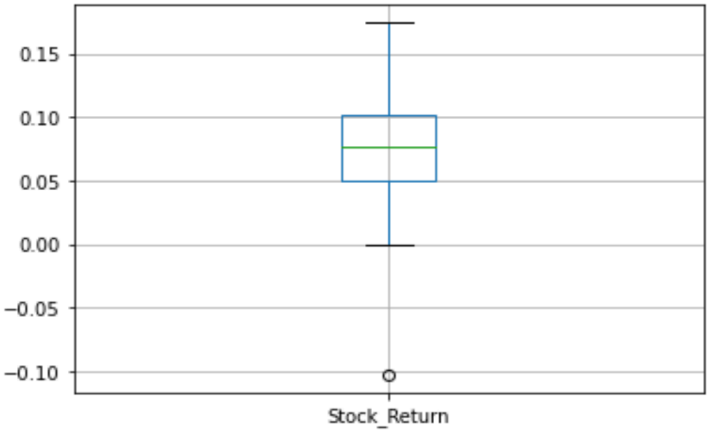


• **Stocks Distribution Sector wise**





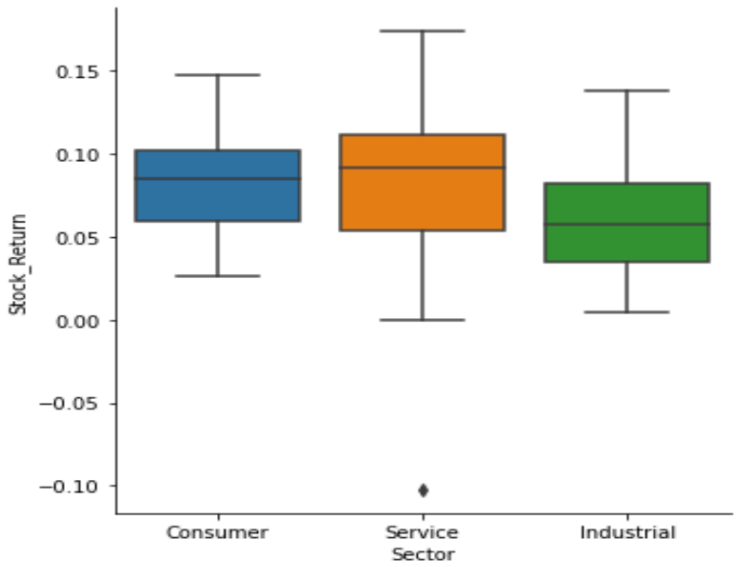
A box plot and whisker plot are a way of summarising a set of data measured on an interval scale. It is often used in explanatory data analysis.



Box plot is drawn on sector wise so that we can analyse data from every corner.

Note:

Box plots divide the data into sections that each contain approximately 25% of the data in that set. Box plots are useful as they provide a visual summary of the data enabling researchers to quickly identify mean values, the dispersion of the data set, and signs of skewness.



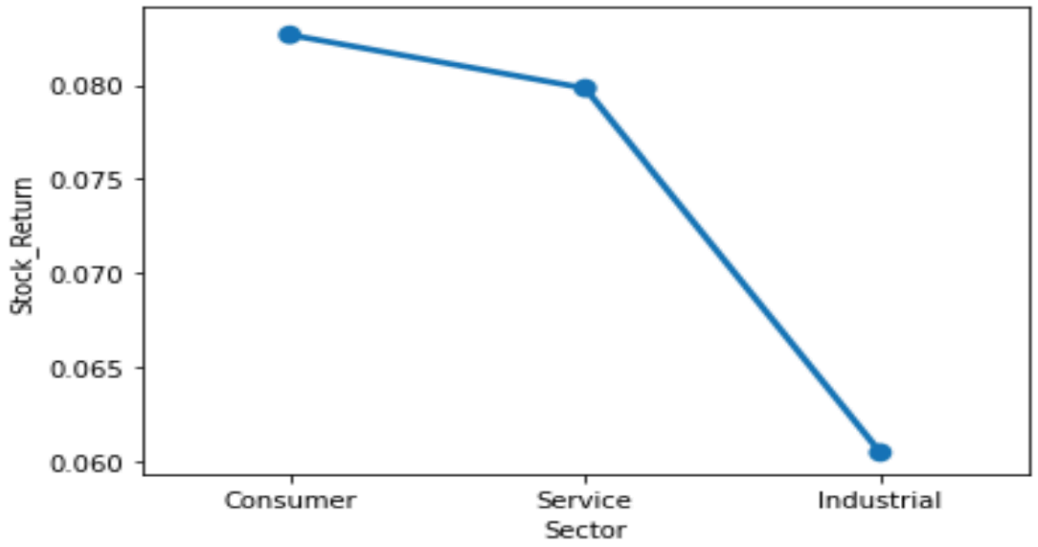
- **Point Plot Graph**

A point plot represents an estimate of central tendency for a numeric variable by the position of scatter plot points and provides some indication of the

Note:

Point plots can be more useful than bar plots for focusing comparisons between different levels of one or more categorical variables. They are particularly adept at showing interactions: how the relationship between levels of one categorical variable change across levels of a second categorical variable. The lines that join each point from the same hue level allow interactions to be judged by differences in slope, which is easier for the eyes

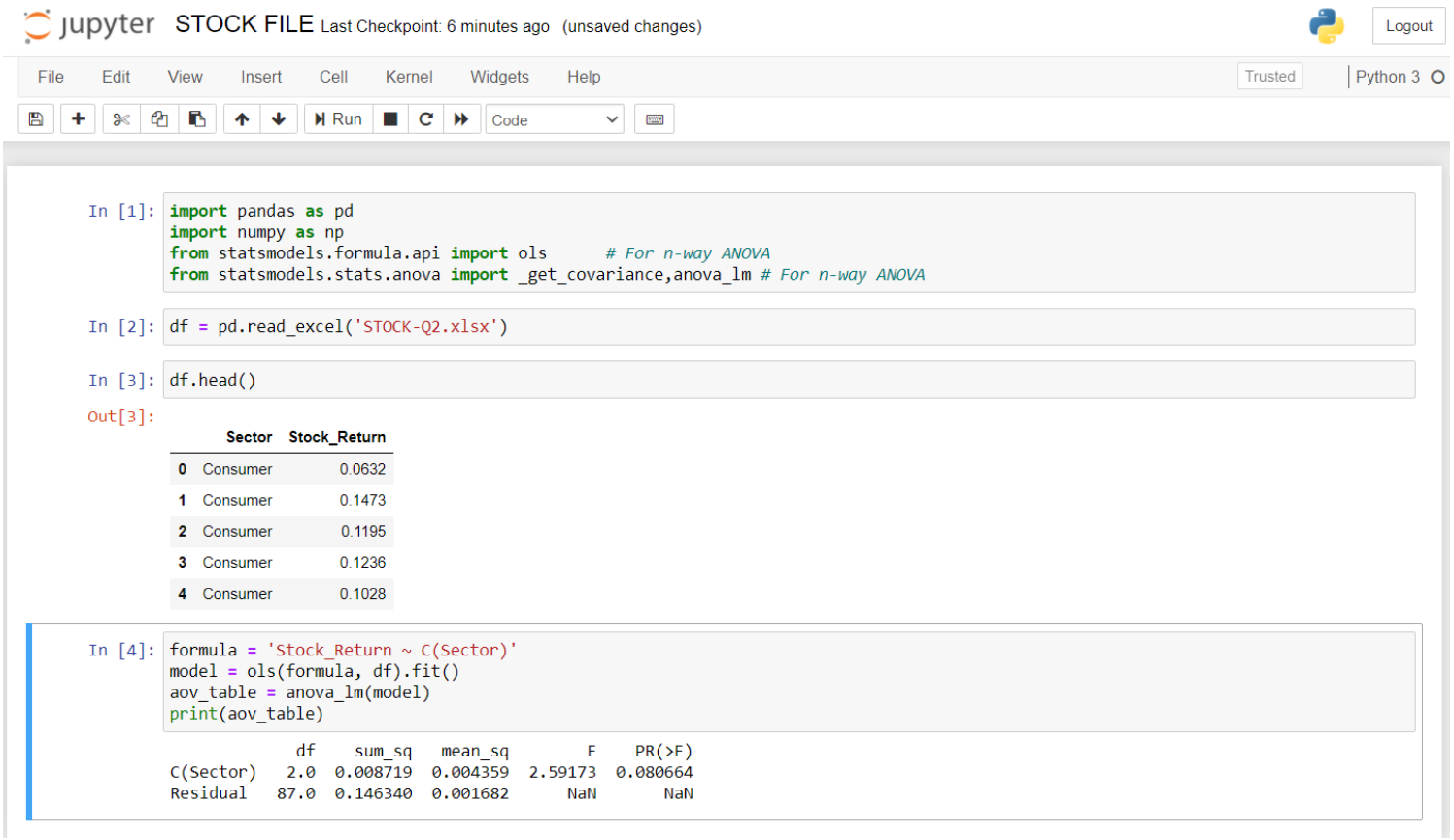
The Stocks Mean Sector wise using Point plot: -



From the above point plot, we see that stock return depends on sector like industrial or consumer or service. So to check this inference we have to use ANOVA statistical method. An ANOVA test is a way to find out if survey or experiment results are significant. In other words, they help us to find out if we need to reject the null hypothesis or accept the alternate hypothesis.

Steps to solve the problem in Jupyter Notebook

- *Import necessary libraries*
- *Load & Read the excel or csv file*
- *Pandas Head() function return the top n (by Default 5) rows*
- *Apply the ANOVA formula & RUN to obtain the ANOVA table :*



Formulation of the hypothesis :

H0*: stock return does not depend on sector like industrial or consumer or service*

H1*: stock return depends on sector like industrial or consumer or service*

Confidence level (α) = 0.05

From the above AOV table we found that p value is greater than ***α*** *i.e.,* ***0.0806 > 0.05***

Hence, we do not reject null hypothesis.

Conclusion:
There is no evidence that stock return depends on sector i.e. industrial or consumer or service.