[Frequently Asked Questions]

[Use of Spreadsheets for Data Analysis]

Subject:	Business Economics
Course:	B. A. (Hons.), 5th Semester, Undergraduate
Paper No. & Title:	Paper – 502 Computational Techniques Business Economics
Unit No. & Title:	Unit – 2 Use of Spreadsheets for Data Analysis
Lecture No. & Title:	1 (One) Use of Spreadsheets for Data Analysis

Frequently Asked Questions

1) What is Excel statistical package?

Ans - Excel is the widely used statistical package, which serves as a tool to understand statistical concepts and computation to check hand-worked calculation in solving homework problems. Most of Excels statistical procedures are part of the Data Analysis tool pack. It includes a variety of choices including simple descriptive statistics, t-tests, correlations, one or two way analysis of variance, regression, etc. If the Excel spreadsheet does not have a Data Analysis item on the Tools menu, we need to install the **Data Analysis ToolPak add-in**. Pivot Table in the Data menu can be used to generate summary tables of means, standard deviations, counts, etc. Even Excel functions can be used to generate some statistical measures such as average, standard deviation and correlation coefficient.

2) What are Pivot table?

Ans - Pivot table provides a way of generating summaries and organising data in ways that are more useful for particular tasks. These are extremely useful for creating contingency tables, cross-tabulations and tables of means and other summary statistics. It is also presenting concise, attractive, and annotated online or printed reports.

3) Explain Data Analysis ToolPak of Excel.

Ans - The Data Analysis ToolPak is an Excel add-in, gives access to a number of helpful tools for running statistical analysis in the workbook, more specifically, undertaking a variation analysis. It is installed when we install Excel but needs to be loaded prior to use. To do so: Select File > Excel Options > Add-Ins > Analysis ToolPak.

4) Introduce ANOVA.

Ans - The ANOVA function in Excel is used for variance analysis. A form of hypothesis testing, it will determine whether two or more factors have the same mean. Currently, it has three different variations depending on the test that can be performed: Single factor, two-factor with replication and two factors without replication.

5) What is ANOVA: Two-Factor with Replication? Describe with suitable examples

Ans - A two way ANOVA with replication is performed when we have two groups and individuals within that group are doing more than one thing (i.e. taking two tests). For example, there are 12 individuals in a group having scores in two tests. Here, we want to know, is there any significant difference between results and scores in Group A and Group B. We will figure out if we are going to reject the null hypothesis or not, we'll basically be looking at two factors:

- If the F-value (f)is larger than the f critical value (f crit)
- If the p-value is smaller than the chosen alpha level (0.05).

6) What is Correlation?

Ans - The Correlation coefficient (a value between -1 and +1) tells how strongly two variables are related to each other. There are two ways to find the correlation coefficient between two variables: CORREL function and the Analysis ToolPak add-in.

In Correlation, when values of one variable increases with the increase in another variable, it is supposed to be a Positive Correlation. On the other hand, if the values of one variable decrease with the decrease in another variable, then it would be a Negative Correlation. There might be the case when there is no change in a variable with any change in another variable. In this case, it is defined as **no correlation** between the two.

7) Differentiate between Linear Correlation and Non-Linear Correlation?

Ans - When the change in one variable results in the constant change in the other variable, then we say the correlation is linear. When the amount of change in one variable is not in a constant ratio to the change in the other variable, we say that the correlation is non linear. For example:

X:	10	20	30	40	50
Y:	20	40	60	80	100

Linear Correlation

X:	10	20	30	40	50
Y:	10	30	70	90	120

Non Linear Correlation

8) What is Covariance?

Ans - Covariance is the measure of how much two sets of data varies. It determines the degree to which the two variables are related or how they vary together. The Covariance is the average of the product of deviations of data points from their respective means.

9) State the difference between Variance and Covariance?

Ans - Variance is the measure of how much a single variable changes, whereas the Covariance is the measure of how much two variables change together. Variance is particularly useful when calculating the probability of future events or performance. A covariance is used to calculate the correlation between variables.

10) What is Descriptive Statistics?

Ans - The Descriptive Statistics analysis tool generates a report of univariate statistics for data in the input range, providing information about the central tendency and variability of data including:

- The mean, mode, median and range.
- Variance and standard deviation.
- Skewness and Kurtosis
- Count, maximum and minimum.