



**INTERNATIONAL JOURNAL OF
PHARMACEUTICAL SCIENCES**
[ISSN: 0975-4725; CODEN(USA): IJPS00]
Journal Homepage: <https://www.ijpsjournal.com>



Review Article

Optimizing Drug Development Using Statistical Software: Key Components of Pharma Industrial Trials

Yash Srivastav¹, Aditya Srivastav², Jaya Singh³

¹*Azad Institute of Pharmacy & Research, Lucknow, U.P, India.*

²*Institute of Pharmacy Bundelkhand University, Kanpur Road Jhansi, U.P, India.*

³*Tahira Institute of medical sciences, Gorakhpur, U.P, India.*

ARTICLE INFO

Received: 24 July 2024

Accepted: 24 July 2024

Published: 26 July 2024

Keywords:

Clinical trial, MS Excel, SPSS, Mini-Tab, R Software.

DOI:

10.5281/zenodo.12910498

ABSTRACT

The clinical trials industry has seen a perceptible increase in recent years and India is one of the global destinations for clinical trials. The safety and efficacy of all new treatments as a drug, vaccines, medical devices, and dietary supplements analyzed through clinical trials. Clinical trials are the most important part of drug discovery for the treatment of diseases such as COVID-19, swine flu, cancer etc. Clinical trials are experiments or observations in clinical research to generate data on the safety and efficacy of new molecules. Researchers or investigators initially select a small group of volunteers or patients and subsequently conduct larger comparative studies based on the type of new molecules. Clinical study design aims to ensure the scientific validity and reproducibility of the results. In clinical trials, the number of subjects (sample size) greatly impacts the ability to measure the effects of the interference. This ability is described as its 'power' and a larger sample size increases the statistical power. The statistical power estimates the ability of a trial to detect a difference of a particular size between the treatment and control groups.

INTRODUCTION

Clinical trial management systems are often used by research sponsors or contract research organizations (CROs) for the planning and management of clinical trials. An interactive voice response system is also used by sites to register the enrollment of patients using a phone and to allocate patients to a particular treatment

arm. Statistical software is used to analyze the collected data and prepare the outcomes for regulatory submission. The most commonly used packages for statistical analysis are Excel, SPSS, MINITAB, SAS JMP, STATA, S-PLUS, R, and SOLAS. Graph Pad Prism is most useful for researchers to perform laboratory studies and clinical trial tests using t-tests, one-way ANOVA,

***Corresponding Author:** Yash Srivastav*

Address: Azad Institute of Pharmacy & Research, Lucknow, U.P, India.

Email ✉: neelashsrv76@gmail.com

Relevant conflicts of interest/financial disclosures: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

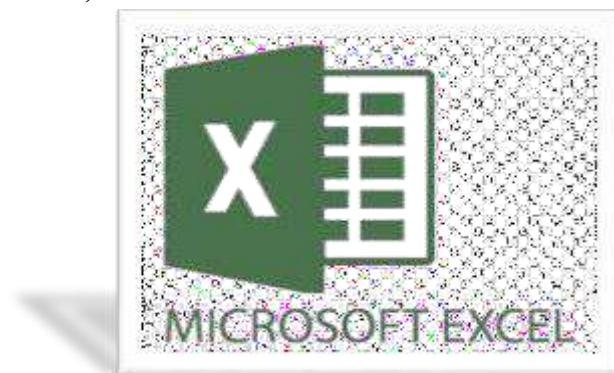


survival analysis and probability models like the logistic regression model [1]. SAS (Statistical Analysis Software) is available only for Windows operating systems and is commonly used for statistical analysis and data visualization. Statistical software is a specialized program designed to allow users to perform complex statistical analysis. These software are tools for the organization, interpretation and presentation of selected data sets. There are many challenges or problems in performing industrial and clinical trials in a developing country like India. The main problems are a lack of experts with formal training in bioethics, limited experience with regulatory trials, infrequent meetings of Institute Ethics Committees (IECs) and a lack of clearly defined roles and responsibilities of its members. Clinical investigators may face many challenges such as funding, responding to multiple review cycles, recruiting patients, establishing clinical trial and material transfer agreements with sponsors and medical centres, time-bound research studies and completing associated paperwork [2]. The cost of clinical trials is very high and it may be billions of dollars per approved drug. Clinical trials may be sponsored by the pharmaceutical industry or any other government organization. A clinical trial might also include an extended post-study follow-up period from months to years for patients or people who have participated in the trial. The main barrier to completing studies is the shortage of patients or people who take part in the trials. Some drug trials require patients to have unusual combinations of disease characteristics. It is a challenge to find the appropriate patients and obtain their consent. The studies were perhaps performed in selected months to avoid seasonal influences such as airborne allergies, influenza, skin diseases, seasonal affective disorder etc. Costs for clinical trials can range into the billions of dollars per approved drug. The sponsor may be a governmental organization or a pharmaceutical,

biotechnology or medical device company. Certain functions necessary to the trial, such as monitoring and laboratory work, may be managed by an outsourced partner, such as a contract research organization or a central laboratory Only 10 per cent of all drugs started in human clinical trials become approved drugs. Some clinical trials involve healthy subjects with no pre-existing medical conditions. Other clinical trials pertain to people with specific health conditions who are willing to try an experimental treatment [3].

2. COMPUTER SOFTWARE

Introduction to Microsoft Excel (Spread Sheets)



Microsoft Excel provides a system in which available data can be presented in an organized and systematic manner. A spreadsheet is a computer program sheet or document that we can use for arithmetic computations in columns and rows. There are several spreadsheet programs but of all of them, Excel is the most widely used. People have been using it for the last 30 years and throughout these years, it has been upgraded with more and more features [4]. The best part about Excel is, that it can apply to many business tasks, including statistics, finance, data management, forecasting, analysis, inventory, billing and business intelligence.

The following are a few things.

- Number Crunching
- Charts and Graphs
- Store and Import Data
- Manipulating Text

- Templates/Dashboards
- Automation of Tasks

There are three most important components of Excel:

1. Cell

A cell is the smallest but most powerful part of a spreadsheet. Data can be entered into a cell either by typing or by copy-paste. Data can be a text, a number, or a date. It can also be customised by changing its size, font colour, background colour, and borders [5]. Every cell is identified by its cell address, cell address contains its column number and row number (If a cell is on the 11th row and column AB, then its address will be AB11).

2. Worksheet

A worksheet is made up of individual cells which can contain a value, a formula, or text. It also has an invisible draw layer, which holds charts, images, and diagrams. Each worksheet in a workbook is accessible by

clicking the tab at the bottom of the workbook window. In addition, a workbook can store a chart sheet; a chart sheet displays a single chart and is accessible by clicking a tab.

3. Workbook

A workbook is a separate file just like every other application. Each workbook contains one or more worksheets. You can also say that a workbook is a collection of multiple worksheets or can be a single worksheet. You can add or delete worksheets, hide them within the workbook without deleting them, and change the order of your worksheets within the workbook [6].

Microsoft Excel Window Components

Before you start using it, it's really important to understand what's where in its window. So ahead we have all the major components which you need to know before entering the world of Microsoft Excel.

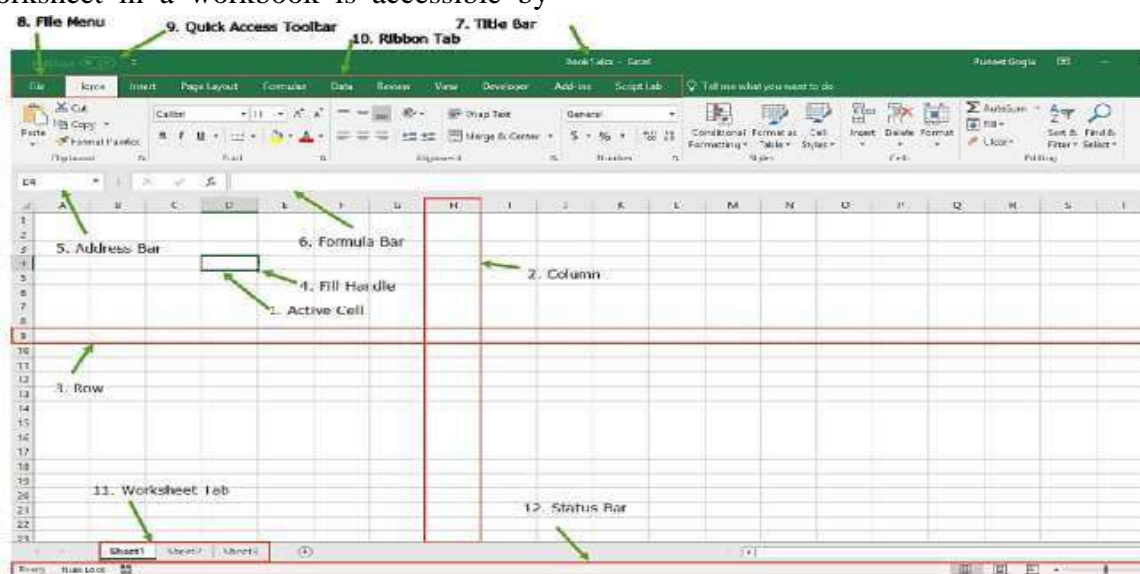


Figure 1: Excel Window Components Outlook

- **Active Cell**

A cell that is currently selected. It will be highlighted by a rectangular box and its address will be shown in the address bar. You can activate a cell by clicking on it or

by using your arrow buttons. To edit a cell, you double-click on it or use F2 as well.

- **Columns**

A column is a vertical set of cells. A single worksheet contains 16384 total columns. Every column has its alphabet for identity,

from A to XFD. You can select a column by clicking on its header [7].

- **Rows**
A row is a horizontal set of cells. A single worksheet contains 1048576 total rows. Every row has its number for identity, starting from 1 to 1048576. You can select a row by clicking on the row number marked on the left side of the window.
- **Fill Handle**
It's a small dot present on the lower right corner of the active cell. It helps you to fill numeric values, text series, insert ranges, insert serial numbers, etc.
- **Address Bar**
It shows the address of the active cell. If you have selected more than one cell, then it will show the address of the first cell in the range.
- **Formula Bar**
The formula bar is an input bar, below the ribbon. It shows the content of the active cell and you can also use it to enter a formula in a cell.
- **Title Bar**
The title bar will show the name of your workbook, followed by the application

name ("Microsoft Excel").

- **File Menu**
The file menu is a simple menu like all other applications. It contains options like (Save, Save As, Open, New, Print, Excel Options, and Share etc).
- **Quick Access Toolbar**
A toolbar to quickly access the options that you frequently use. You can add your favourite options by adding new options to the quick access toolbar.
- **Ribbon Tab**
Starting from Microsoft Excel 2007, all the options menus are replaced with ribbons. Ribbon tabs are a bunch of specific option group that further contains the option.
- **Worksheet Tab**
This tab shows all the worksheets that are present in the workbook. By default you will see, three worksheets in your new workbook with the names Sheet1, Sheet2, and Sheet3 respectively.
- **Status Bar**
It is a thin bar at the bottom of the Excel window. It will give you instant help once you start working in Excel [8].

Data Collection in Microsoft Excel



Participant	Age	Gender	Country of origin	Time in Australia	Like:Nutella	TriedBefore:Nutella	Recommend:Nutella	Like:Vegemite	TriedBefore:Vegemite	Recommend:Vegemite
1	23	F	INDIA	3 days	Y	N	Y	N	N	N
2	25	F	NEPAL	10 hours	Y	N	Y	N	N	N
3	27	M	BANGLADESH	4 months	N	N	N	N	N	N
4	22	F	CHINA	5 months	Y	Y	N	N	N	N
5	26	M	HONG KONG	3 months	Y	Y	N	N	N	N
6	31	F	SINGAPORE	2 weeks	Y	N	Y	N	N	Y
7	24	M	INDIA	3 weeks	Y	N	Y	N	N	N
8	24	M	CHINA	5 weeks	N	N	Y	N	N	N
9	25	F	NEPAL	1 month	Y	N	N	N	N	N
10	23	F	INDIA	15 days	Y	N	Y	N	N	N
11	52	F	HONG KONG	8 days	N	N	N	N	N	N
12	30	M	HONG KONG	3 weeks	Y	N	Y	Y	N	N
13	28	F	NEPAL	5 days	Y	N	Y	N	N	N
14	29	M	INDIA	2 months	Y	Y	N	N	N	N
15	27	F	NEPAL	10 days	Y	N	N	N	N	N

Figure 2: Data Collection in Excel (Spread Sheets)

Data collection is any process in which information is gathered and expressed in a summary form, for purposes such as statistical analysis. A common collection purpose is to get more information about particular groups based on specific variables such as age, profession, or income. The information about such groups can then be used for website personalization to choose content and advertising likely to appeal to an individual belonging to one or more groups for which data has been collected. For example, a site that sells music CDs might advertise certain CDs based on the age of the user and the data collection for their age groups [9]. Collection or grouping is a common task in database management. It allows obtaining summary statistics (e.g. mean, sum, max, etc.) of one or more quantitative variables. The data can be collected in Excel using the XLSTAT statistical software. An Excel sheet with both the data and the results can be downloaded by clicking on the given button. XLSTAT is a user-friendly software that is used to analyse the

data of statistics [10]. It was developed by Thierry Fahmy in 1993. This software is used to summarize data using simple statistics like mean, median, standard deviation etc. It easily extracts information from a large data set form. It helps to accept or reject a very precise hypothesis assuming error risks. It is used in modelling by way of phenomenon according to the set of parameters. Further use for XLSTAT software includes using Simple Linear Regression and Linear Regression XLSTAT to develop a statistical model to solve problems [11].

3. STATISTICAL PACKAGE FOR THE SOCIAL SCIENCES (SPSS)

SPSS means “Statistical Package for the Social Sciences” and was first launched in 1968. Since SPSS was acquired by IBM in 2009, it's officially known as IBM SPSS Statistics but most users still just refer to it as “SPSS”[10,11]. SPSS is a widely used program for statistical analysis in social science. It is also used by market researchers,

health researchers, survey companies, government, education researchers, marketing organizations, data miners and others. The original SPSS manual by Nie, Bent and Hull, 1970 has been described as one of “sociology’s

most influential books” for allowing ordinary researchers to do their statistical analysis [12].

	id	first_name	gender	dob	educ	mr
1	1	Kevin	1	03-Oct-1993	.	
2	2	Ayden	1	31-Oct-1996	4	
3	3	Madelyn	0	13-Dec-1985	5	
4	4	Madelyn	0	10-Jun-1988	1	
5	5	Tristan	1	23-Dec-1995	3	

Figure 3:- IBM SPSS

SPSS is a trial ware that you can use to record and then analyze data. (Repetition)

SPSS has a user interface that looks like Microsoft Excel as the UI is set up as a spreadsheet. MATLAB, MINITAB, POWER BI, Stata, MYSQL, and Tableau are data management alternatives to SPSS. All programs are suitable for professional use.

SPSS – Quick Overview Main Features

SPSS is software for editing and analyzing all sorts of data. These data may come from basically

any source: scientific research, a customer database, Google Analytics or even the server log files of a website. SPSS can open all file formats that are commonly used for structured data such as

- Spreadsheets from MS Excel
- Plain text files (.txt or .csv)
- Relational (SQL) databases
- Stata and SAS

SPSS data view

After opening data, SPSS displays them in a spreadsheet.

	gender	birthday	source_2010	income_2010	in
1	0	17-Apr-1979	0	\$6,072.40	
2	1	10-Apr-1985	0	\$12,706.65	
3	0	27-Jun-1986	0	\$14,912.82	
4	1	27-Dec-1983	1	\$16,338.36	
5	0	26-Apr-1979	0	\$22,606.99	
6	1	20-Feb-1971	0	\$23,544.95	
7	0	25-Apr-1987	1	\$24,985.21	

Figure 4:- Overview SPSS Display: Data View

This sheet called data view - always displays our data values. For instance, our first record seems to contain a male respondent from 1979 and so on. A more detailed explanation of the exact meaning

of our variables and data values is found in a second sheet shown next page [8].

SPSS Variable View

	Name	Type	Label	Val
1	id	Numeric	Unique person identifier	None
2	gender	Numeric		{0, Male
3	birthday	Date		None
4	source_2010	Numeric	Primary source of income in 2010	{0, Cor
5	source_2011	Numeric	Primary source of income in 2011	{0, Cor
6	income_2010	Dollar	Gross total income over 2010	None
7	income_2011	Dollar	Gross total income over 2011	None
8	income_2012	Dollar	Gross total income over 2012	None

Figure 5:- SPSS Variable View

An SPSS data file always has a second sheet called variable view. It shows the metadata associated with the data. Metadata is information about the meaning of variables and data values. This is generally known as the “codebook” but in SPSS it's called the dictionary. For non-SPSS users, the look and feel of SPSS’ Data Editor Window probably comes closest to an Excel workbook containing two different but strongly related sheets.

Data Analysis

Right, so SPSS can open all sorts of data and display them and their Metadata in two sheets in its Data Editor window. So how to analyze your data in SPSS? Well, one option is using SPSS's elaborate menu options. For instance, if our data contain a variable holding respondents’ incomes over 2010, we can compute the average income by navigating to Descriptive Statistics as shown below [13].

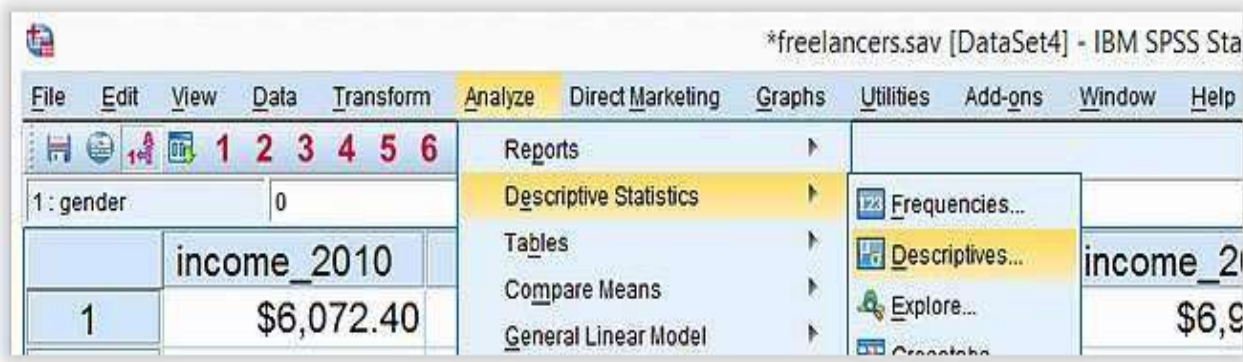


Figure 6:-Data Analysis

Doing so opens a dialogue box in which we select one or many variables and one or several statistics

we'd like to inspect.

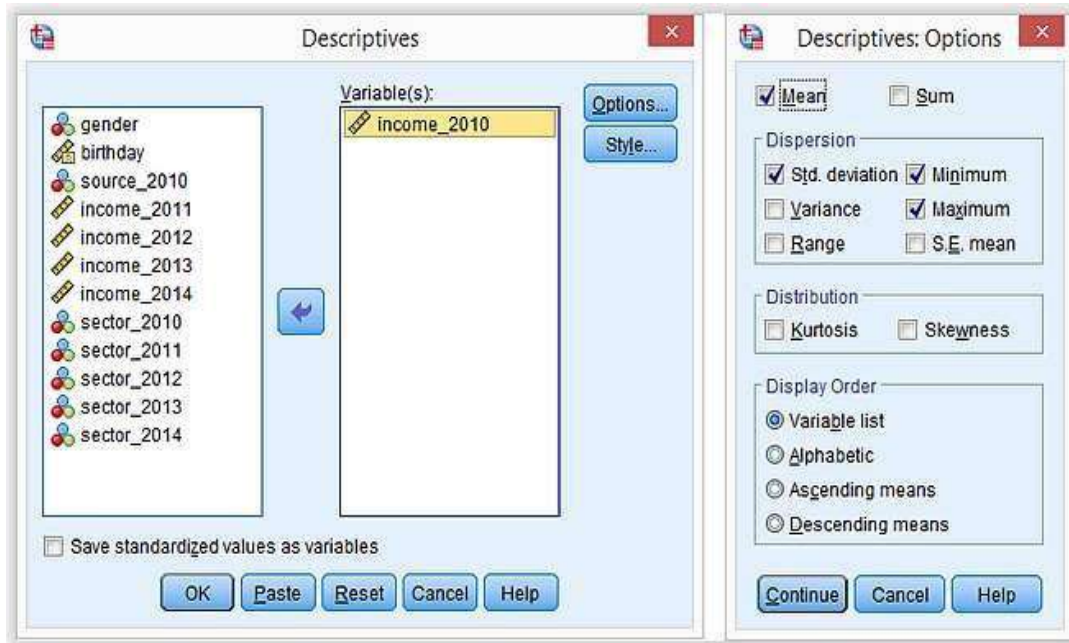


Figure 7:- Dialog box

SPSS Output Window

After clicking OK, a new window opens up: SPSS’ output viewer window. It holds a nice table

with all the statistics on all the variables we chose. The screenshot below shows what it looks like.

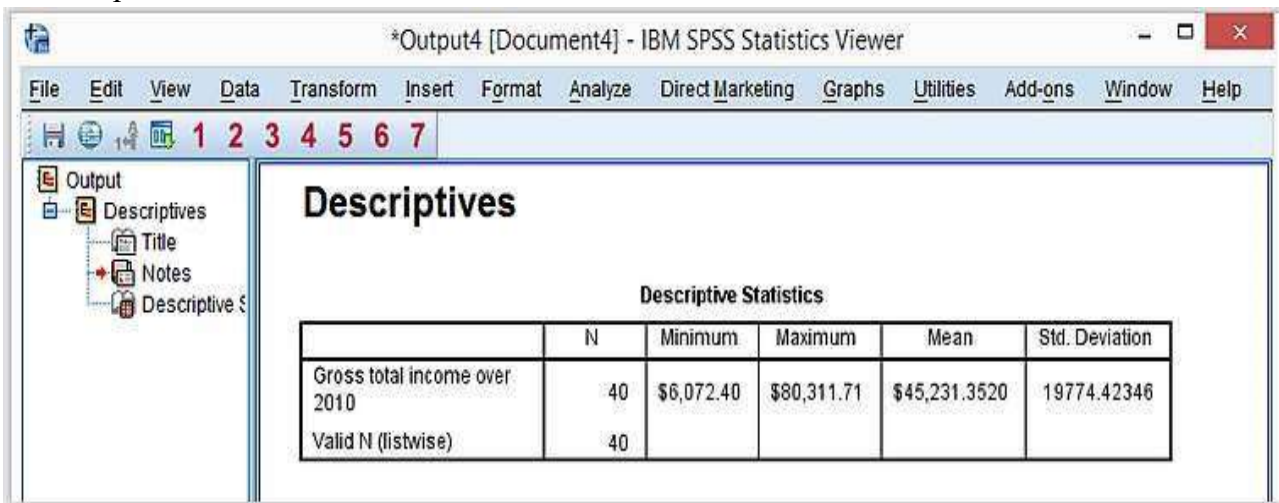


Figure 8:- SPSS Output Window

As we see, the Output Viewer window has a different layout and structure than the Data Editor window we saw earlier. Creating output in SPSS does not change our data in any way; unlike Excel, SPSS uses different windows for data and research outcomes based on those data. For non-SPSS users, the look and feel of SPSS’ Output Viewer window probably comes closest to a

PowerPoint slide holding items such as blocks of text, tables and charts.

SPSS Reporting

SPSS Output items, typically tables and charts, are easily copy-pasted into other programs. For instance, many SPSS users use a word processor such as MS Word, OpenOffice or Google Docs for reporting. Tables are usually copied in rich

text format, which means they'll retain their styling such as fonts and borders. The screenshot below illustrates the result [14].

The screenshot shows a window titled 'Untitled 1 - OpenOffice Writer'. The main content is a table titled 'Descriptive Statistics'. The table has six columns: N, Minimum, Maximum, Mean, and Std. Deviation. The first row of data is for 'Gross total income over 2010' with a value of 40 for N, \$6,072 for Minimum, \$80,312 for Maximum, \$45,231 for Mean, and \$19,774 for Std. Deviation. The second row is 'Valid N (listwise)' with a value of 40 for N.

Descriptive Statistics					
	N	Minimum	Maximum	Mean	Std. Deviation
Gross total income over 2010	40	\$6,072	\$80,312	\$45,231	\$19,774
Valid N (listwise)	40				

Figure 9:- SPSS Reporting

SPSS Syntax Editor Window

The output table we showed was created by running Descriptive Statistics from the SPSS menu. Now, SPSS has a second option for running this (or any other) command: we can open a third window, known as the syntax editor window.

Here we can type and run SPSS code known as SPSS syntax.

```
descriptives income_2010.
```

For instance, running descriptive income 2010. Has the same result as running this command from the SPSS menu as we did earlier.

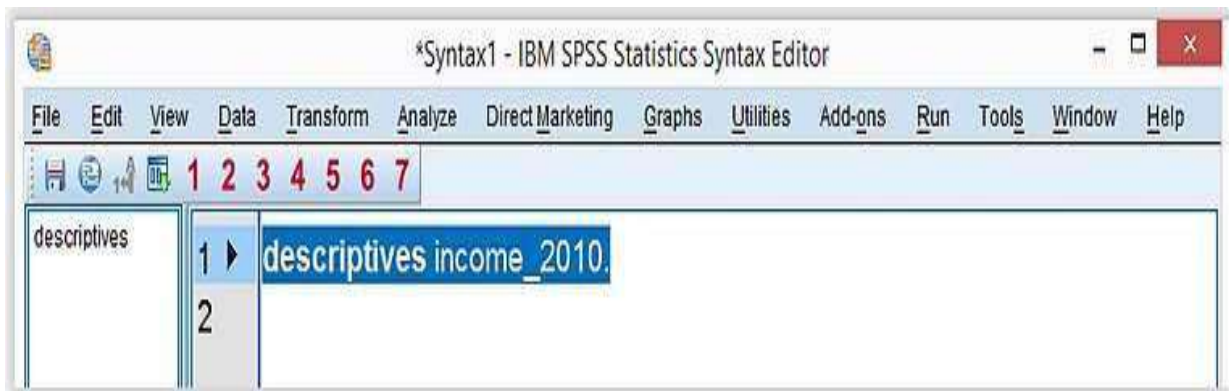


Figure 10:- SPSS Syntax Window

Besides typing commands into the Syntax Editor window, most of them can also be pasted into it by clicking through the SPSS menu options. As so, SPSS users unfamiliar with syntax can still use it. The basic point is that syntax can be saved,

corrected, rerun and shared between projects or users. Your syntax makes your SPSS work replicable. If anybody raises any doubts regarding your outcomes, you can show exactly what you did and -if correct rerun it in seconds.

For non-SPSS users, the look and feel of SPSS' Syntax Editor Window probably comes closest to Notepad: a single window just containing plain text [15].

SPSS – Overview Main Features

Now that we have a basic idea of how SPSS works, let's take a look at what it can do. Following a typical project workflow, SPSS is great for [9].

- Opening data files, either in SPSS file format or many others;
- Editing data such as computing sums and means over columns or rows of data. SPSS has outstanding options for more complex operations as well.

- Creating tables and charts containing frequency counts or summary statistics over groups of cases and variables.
- Running inferential statistics such as ANOVA, regression and factor analysis.
- Saving data and output in a wide variety of file formats.

Opening Data Files

SPSS has its own data file format. Other file formats it easily deals with include MS Excel, plain text files, SQL, Stata and SAS.

	A	B	C	D	E
1	Page	Month of the year	Pageviews	Unique Pageviews	Avg. Time on Page
2	/basics/	12	837	526	00:00:31
3	/mean-center-many-variables/	2	823	680	00:04:04
4	/statistical-tests/	1	819	508	00:00:43

Web analytics data -often downloaded as Excel files- can easily be opened and further analyzed in SPSS

Figure 11:- Data Files

Editing Data

In real-world research, raw data usually need some editing before they can be properly analyzed. Typical examples are creating means or sums as new variables, restructuring data or detecting and removing unlikely observations. SPSS performs such tasks and more complex ones with amazing efficiency. For getting things done fast, SPSS contains many numeric functions, string functions, date functions and other handy routines.

Tables and Charts

All basic tables and charts can be created easily and fast in SPSS. Typical examples are demonstrated under Data Analysis. A real weakness of SPSS is that its charts tend to be ugly and often have a clumsy layout. A great way to overcome this problem is by developing and applying SPSS chart templates. Doing so, however, requires a fair amount of effort and expertise [10].

Inferential Statistics

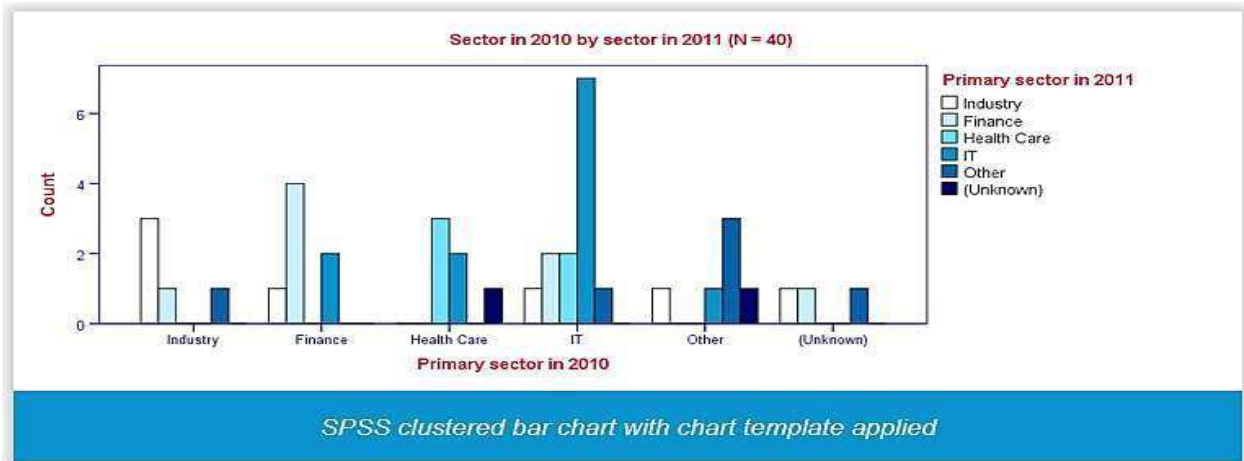


Figure 12:-Inferential Statistics

SPSS contains all basic statistical tests and multivariate analyses such as:-

- t-tests
- chi-square tests
- ANOVA
- correlations and other association measures
- regression
- nonparametric tests
- factor analysis
- cluster analysis

Some analyses are available only after purchasing additional SPSS options on top of the main program. An overview of all commands and the options to which they belong is presented in Overview All SPSS Commands.

	Test Value = 400					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
body_weight	(1) -2.428	(2) 39	(3) .020	(4) -30.450	-55.81	-5.09

SPSS One Sample T-Test Output Example

Figure 13:- Sample Test

Saving Data and Output

SPSS data can be saved in a variety of file formats, including:-

- MS Excel
- plain text (.txt or .csv)
- Stata
- SAS

The options for output are even more elaborate: charts are often copy pasted as images in .png format. For tables, rich text format is often used

because it retains the tables, layout, fonts and borders. Besides copy-pasting individual output items, all output items can be exported in one go to .pdf, HTML, MS Word and many other file formats. A terrific strategy for writing a report is creating an SPSS output file with nicely styled tables and charts. Then export the entire document to Word and insert explanatory text and titles between the output items [8,7].

4. MINITAB



Minitab is statistical analysis software developed in 1972 at the Pennsylvania State University. This is designed essentially for the Six Sigma professionals to give effective solutions for statistical analysis in most Six Sigma projects. (Reference) Minitab is simple, and effective for inputting or manipulating statistical data, identifying trends and patterns and easily finding solutions to the current issues. Minitab is an easy-to-use statistical software package that was designed especially for the teaching of introductory statistics courses as well as statistical research. Statistical analysis computer applications have the advantage of being accurate,

reliable, and generally faster than computing statistics and drawing graphs by hand [13]. Minitab also produces other software that can be used in conjunction with Minitab; Quality Trainer is an eLearning package that teaches statistical tools and concepts in the context of quality improvement and Quality Companion is a tool for managing Six Sigma and Lean manufacturing. Minitab is a general-purpose statistical software package designed for easy interactive use. Minitab is well suited for instructional applications but is also powerful enough to be used as a primary tool for analyzing research data.

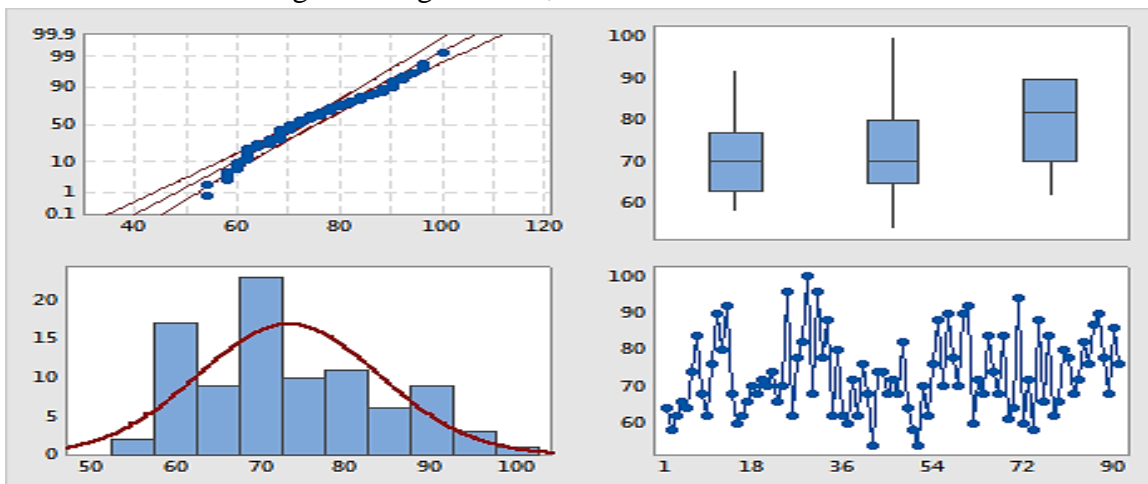


Figure 14:- Graphical Work in Minitab

There are many versions of Minitab such as Minitab Student Version 14, and Minitab Version 13 running under Windows. The text is based on

Minitab Student Version 14 and Minitab Version 14. The core of the manual is a discussion of the menu commands while not neglecting to refer to

the session commands, as these are needed for certain problems. The material on session commands is always at the end of each section and can be skipped if the reader will not be using them. In the author's experience, ease of learning and use are the salient features of the package, with obvious benefits to the student and to the instructor, who can relegate many details to the software. While more sophisticated packages are necessary for higher-level professional work, it is our experience that attempting to teach one of these in a course forces too much attention to technical aspects. The time students need to spend to learn Minitab is relatively small and it is a great virtue. Further Minitab will serve as a perfectly adequate tool for many of the statistical problems that students will encounter in their undergraduate education. The following applications are performed using Minitab Software for statistics [14].

Minitab for Data Management

Entering Data into a Worksheet, Importing Data, Patterned Data, Printing Data in the Session Window, Assigning Constants, Naming Variables and Constants, Information about a Worksheet,

Editing a Worksheet, Saving, Retrieving, and Printing, Mathematical Operations Arithmetical Operations, Mathematical Functions, Comparisons and Logical Operations Column and Row Statistics, Sorting Data, Computing Ranks.

Minitab for Data Analysis

Looking at Data. Distributions, Plotting Data, The Normal Distribution, Looking at Data Relationships, Producing Data, Probability: The study of Randomness, Sampling Distributions, Introduction to Inference, Inference for Distributions, Inference for Proportions, Inference for Two Way Tables, Inference for Regression, Multiple Regression, One Way Analysis of Variance, Two Way Analysis of Variance, Bootstrap methods and Permutation Tests, Nonparametric Tests, Logistic Regression, Statistics for Quality: Control and Capability, Functions in Minitab, Mathematical Functions, Column Statistics, Row Statistics, More Minitab Commands, Programming in Minitab, Matrix Algebra in Minitab. Apart from the above listed programs other applications of Minitab can be learned during the use of software.

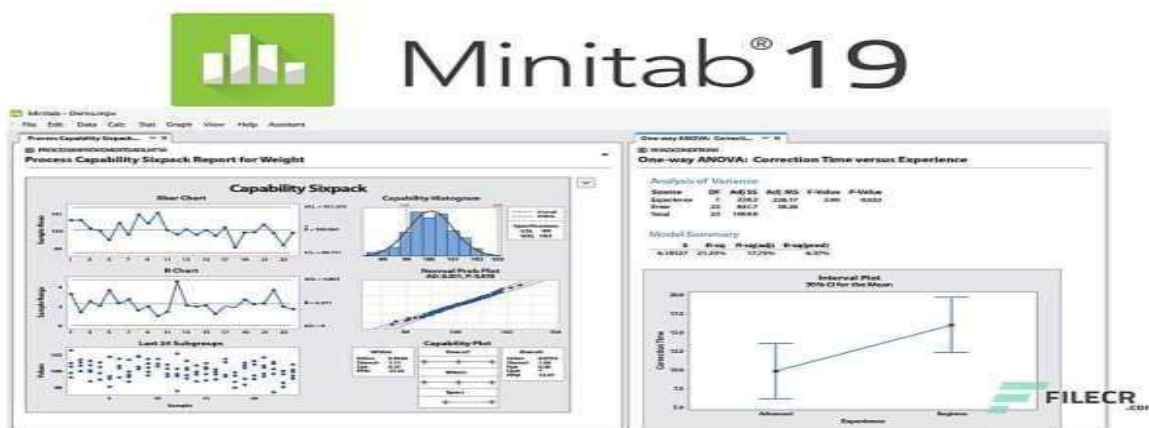


Figure 15 :- Minitab 19

5. R-ONLINE STATISTICAL SOFTWARE TO INDUSTRIAL AND CLINICAL TRIAL APPROACH

R is a programming language for statistical computing and graphics supported by the R Core Team and the R Foundation for Statistical Computing. Created by statisticians Ross Ihaka

and Robert Gentleman. R is used among data miners and statisticians for data analysis and developing statistical software. Users have created packages to augment the functions of the R language [15]. According to surveys like Rexter's Annual Data Miner Survey and studies of scholarly literature databases-R is one of the most commonly used programming languages in data mining. As of February 2022, R ranks 13th in the TIOBE index, a measure of programming language popularity. The official R software environment is an open-source free software environment within the GNU package, available under the GNU General Public License. It is written primarily in C, FORTRAN, and R itself. Precompiled executables are provided for various operating systems. R has a command line interface. Multiple third-party graphical user interfaces are also available, such as R-Studio an integrated development environment, and Jupiter, a notebook interface. R and its libraries implement various statistical and graphical techniques, including linear and nonlinear modelling, classical statistical tests, spatial and time-series analysis, classification, clustering, and others. R is easily extensible through functions and extensions, and its community is noted for contributing packages. Many of R's standard functions are written in R which makes it easy for users to follow the algorithmic choices made. For computationally intensive tasks C, C++, and FORTRAN code can be linked and called at run time. Advanced users can write C, C++, Java, .NET or Python code to manipulate R objects directly [5,11]. R is highly extensible through the use of packages for specific functions and specific applications. Due to its S heritage, R has stronger object-oriented programming facilities than most statistical computing languages. Extending it is facilitated by its lexical scoping rules [4,15]. Another of R's strengths is static graphics; it can produce publication-quality graphs that include

mathematical symbols. Dynamic and interactive graphics are available through additional packages.

The R environment

R is an integrated suite of software facilities for data manipulation, calculation and graphical display. It includes:-

An effective data handling and storage facility. A suite of operators for calculations on arrays in particular matrices. A large, coherent, integrated collection of intermediate tools for data analysis.

Graphical facilities for data analysis and display either on screen or on hardcopy and a developed well-developed, simple and effective programming language that includes conditionals, loops, defined user-defined recursive functions and input and output facilities. The term "environment" is intended to characterize it as a fully planned and coherent system rather than an incremental accretion of very specific and inflexible tools as is frequently the case with other data analysis software [7,15]. R like S is designed around a true computer language and it allows users to add additional functionality by defining new functions. Much of the system is itself written in the R dialect of S which makes it easy for users to follow the algorithmic choices made. For computationally intensive tasks C, C++ and FORTRAN code can be linked and called at run time. Advanced users can write C code to manipulate R objects directly. Many users think of R as a statistics system. We prefer to think of it as an environment within which statistical techniques are implemented. R can be extended via packages. There are about eight packages supplied with the R distribution and many more are available through the CRAN family of Internet sites covering a very wide range of modern statistics. R has its own Latex documentation format which is used to supply comprehensive documentation both online in several formats and in hardcopy [12].



CONCLUSION

Practical Components of Industrial and Clinical trial problems are important for drug discovery. This project described Computer Software, in which Microsoft Excel (Spread Sheet) and Data collection and also described, Statistical Package for the Social Science (SPSS) is a widely used program for statistical analysis in social science. It is also used by market researchers, health researchers, survey companies, government, education, researchers, marketing organizations, data miners and others Minitab is simple, and effective for inputting or manipulating statistical data, identifying trends and patterns and easily finding solutions to the current issues. It is very important for designed and graphical work. The official R software environment is an open-source free software environment within the GNU package, available under the General public license. Observations in clinical research need to generate data on the safety and efficacy of new molecules.

ETHICAL STATEMENT

A pharmacist ought to act with integrity and sincerity. A pharmacist abstains from behaviours that could undermine their commitment to acting in their patient's best interests, such as prejudiced acts or behaviours and unfavourable working environments that impair their judgment. A pharmacist upholds their reputation in the industry.

INFORMED CONSENT

Using websites, review articles, and other sources to produce research content.

DISCLAIMER(ARTIFICIAL INTELLIGENCE)

Author(s) at this moment declare that NO generative AI technologies such as Large Language Models (ChatGPT, COPILOT, etc) and text-to-image generators have been used during the writing or editing of manuscripts.

REFERENCE

1. Chandrakant, kokare. P. (2004). Biostatistics and Research Methodology (Vol. 1st). Bangkok .
2. Berry, D. A. Statistics: A Bayesian Perspective (Duxbury, California, 1996).
3. Berger, J. O. & Berry, D. A. Statistical analysis and the illusion of objectivity. *Am. Scientist* 159–165 (1988).
4. DHA_Dr.Vashistha Shikha, M.S Excel: Spreadsheets and their uses in Business, Excel Worksheets, Excel Formatting techniques,(website).
5. Dr. Kumar, Sanjay. & Kumar Deepak. Biostatistics and Research Methodology (vol. 1st). PV Books, 185-186.
6. Afifi, A. A., Clark, V., & May, S. (2003). Computer-aided multivariate analysis (4th Ed.). London.
7. Beasley, W. H., Deshea, L., Toothaker, L. E., Mendoza, J. L., Bard, D. E., & Rodgers, J. L. (2007).
8. Fleiss, J. L. (1993). The statistical basis of meta-analysis. *Statistical Methods in Medical Research*, 2, 121–145.
9. Patil VM, Noronha V, Joshi A, Choughule AB, Bhattacharjee A, Kumar R, et al. Phase III study of gefitinib or pemetrexed with carboplatin in EGFR-mutated advanced lung adenocarcinoma.
10. <https://www.spss-tutorials.com/spss-what-is-it/website>.
11. Cox DR. Analysis of Survival Data. Chapman and Hall/CRC: Routledge; 2018.
12. <https://en.wikipedia.org/wiki/SPSS/> SPSS-Wikipedia,website.
13. Fisher, R.A. (1925). *Statistical Methods for Research Workers*. Oliver & Boyd, Edinburgh.
14. Hogg, R.V., Ledolter, J. (1987). *Applied Statistics for Engineers and Physical Scientists*. Macmillan Publishing Company,



NY.

15. <http://www.r-project.org/about.html>.
Language, website.

R-

HOW TO CITE: Yash Srivastav*, Aditya Srivastav, Jaya Singh, Optimizing Drug Development Using Statistical Software: Key Components of Pharma Industrial Trials, *Int. J. of Pharm. Sci.*, 2024, Vol 2, Issue 7, 1896-1911. <https://doi.org/10.5281/zenodo.12910498>

