Report on Introduction to Data Management

World Population Data

Submitted on partial fulfillment of the requirements for the award of degree of

Bachelor of Technology

In

Computer Science & Engineering

Submitted to

LOVELY PROFESSIONAL UNIVERSITY

PHAGWARA, PUNJAB

FROM 04/10/22 TO 04/11/22

SUBMITTED BY -

Name of Student: ABHINAV SRIVASTAVA

Registration Number: 12003236

Signature of the Student:

To whom so ever it may concern

I, ABHINAV SRIVASTAVA,12003236, hereby declare that the work done by me on **"World Population Data"**, is a record of original work for the partial fulfillment of the requirements for the award of the degree, **B. Tech in Computer Science & Engineering.**

Name of Student: ABHINAV SRIVASTAVA

Registration Number: 12003236

Signature of the Student:

Date: 05-11-2022

ACKNOWLEDGEMENT

With the blessings of Almighty God, I have been able to complete my project work of Introduction to Data Management (INT-217) of B.Tech. in Computer Science & Engineering. I am indebted to all who have contributed in taking up the project work as part of the B.Tech. program and I express my gratitude to all of them from core of my heart. I express my deepest thanks to **Sameeksha Khare**, instructor of course Introduction to Data Management (INT-217) for giving necessary advices, guidance and arrange all resources to make my project work easier. I take this moment to acknowledge her contribution gratefully. I also do not like to miss the opportunity to acknowledge the contribution of all faculty members of the department for their kind assistance and cooperation during the development of my project.

ABSTRACT

The purpose of this project is to develop a computerized and interactive dashboard to analyze the data in a scientific manner. In this project, the dataset is chosen from a website kaggle.com which is about to predict the population of various country. Previously there were no such tools to analyze whole country population. This data analysis will help the people to analyze the world population in different criteria. In this project, the data set used are very scientifically taken.

System Requirements

A. Project Profile:

Project Title:	World Population Data
Organization:	Lovely Professional University
Developed By:	Abhinav Srivastava
Internal Guide:	Sameeksha Khare

B. Project Tools:

Platform tools	MS Excel 2019

C. Hardware Requirement Recommended:

Internet Connection	Not Required
Version	Windows 7 and above.

Table of contents

S.NO	ITEMS	PAGE NO.
1.	DECLARATION	**2**
2.	ACKNOWLEDGEMENT	**3**
3.	ABSTRACT & SYSTEM REQUIREMENTS	**4-5**
4.	CHAPTER 1-INTRODUCTION	**7-12**
5.	CHAPTER 2-OVERVIEW	**13-21**
6.	CHAPTER 3-IMPLEMENTATION AND RESULTS	**22-29**
7.	CHAPTER 4- CONCLUSION	**30**

CHAPTER 1-INTRODUCTION

1.1 ABOUT DATA MANAGEMENT:

The ability to analyse data is a powerful skill that helps you make better decisions. Microsoft Excel is one of the top tools for data analysis and the built-in pivot tables are arguably the most popular analytic tool.

In this course, you will learn how to perform data analysis using Excel's most popular features. You will learn how to create pivot tables from a range with rows and columns in Excel. You will see the power of Excel pivots in action and their ability to summarize data in flexible ways, enabling quick exploration of data and producing valuable insights from the accumulated data.

Pivots are used in many different industries by millions of users who share the goal of reporting the performance of companies and organizations. In addition, Excel formulas can be used to aggregate data to create meaningful reports. To complement, pivot charts and slicers can be used together to visualize data and create easy to use dashboards.

Ms Excel:-

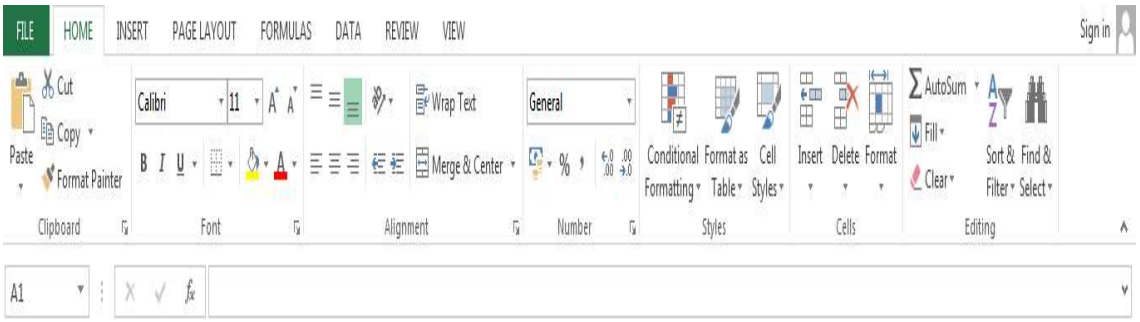
Microsoft Excel is a spreadsheet program used to record and analyse numerical and statistical data. Microsoft Excel provides multiple features to perform various operations like calculations, pivot tables, graph tools, macro programming, etc. It is compatible with multiple OS like Windows, macOS, Android and iOS.

A Excel spreadsheet can be understood as a collection of columns and rows that form a table. Alphabetical letters are usually assigned to columns, and numbers are usually assigned to rows. The point where a column and a row meet is called a cell. The address of a cell is given by the letter representing the column and the number representing a row.

1.2 FEATURES OF EXCEL:-

Various editing and formatting can be done on an Excel spreadsheet.

The image below shows the composition of features in MS Excel: -



- **Home**
  - Comprises options like font size, font styles, font colour, background colour, alignment, formatting options and styles, insertion and deletion of cells and editing options
  -
- **Insert**
  - Comprises options like table format and style, inserting images and figures, adding graphs, charts and sparklines, header and footer option, equation and symbols

- **Page Layout**
  - Themes, orientation and page setup options are available under the page layout option

- **Formulas**

- Since tables with a large amount of data can be created in MS excel, under this feature, you can add formulas to your table and get quicker solutions

- **Data**
  - Adding external data (from the web), filtering options and data tools are available under this category
  -
- **Review**
  - Proofreading can be done for an excel sheet (like spell check) in the review category and a reader can add comments in this part

- **View**
  - Different views in which we want the spreadsheet to be displayed can be edited here. Options to zoom in and out and pane arrangement are available under this category.

Benefits of Using MS Excel: -

MS Excel is widely used for various purposes because the data is easy to save, and information can be added and removed without any discomfort and less hard work.

Given below are a few important benefits of using MS Excel:

- **Easy To Store Data:** Since there is no limit to the amount of information that can be saved in a spreadsheet, MS Excel is widely used to save data or to analyse data. Filtering information in Excel is easy and convenient.
- **Easy To Recover Data:** If the information is written on a piece of paper, finding it may take longer, however, this is not the case with excel spreadsheets. Finding and recovering data is easy.
- **Application of Mathematical Formulas:** Doing calculations has become easier and less time-taking with the formulas option in MS excel
- **More Secure:** These spreadsheets can be password secured in a laptop or personal computer and the probability of losing them is way lesser in comparison to data written in registers or piece of paper.
- **Data at One Place:** Earlier, data was to be kept in different files and registers when the paperwork was done. Now, this has become convenient as more than one worksheet can be added in a single MS Excel file
- **Neater and Clearer Visibility of Information:** When the data is saved in the form of a table, analyzing it becomes easier. Thus, information is a spreadsheet that is

MS Excel – Points to Remember

There are certain things which one must know with respect to MS Excel, its applications and usage:

- An MS Excel file is saved with an extension of .xls
- Companies with large staff and workers use MS Excel as saving employee information becomes easier
- Excel spreadsheets are also used in hospitals where the information of patients can be saved more easily and can be removed conveniently once their medical history is cleared
- The sheet on which you work is called a Worksheet
- Multiple worksheets can be added in a single Excel file
- This is a data processing application.

CHAPTER-2: - OVERREVIEW

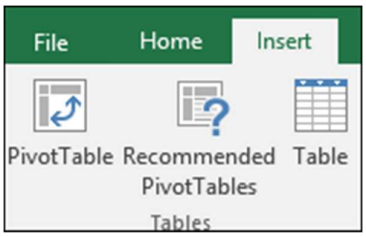
A PivotTable is a powerful tool to calculate, summarize, and analyze data that lets you see comparisons, patterns, and trends in your data.

Create a PivotTable in Excel

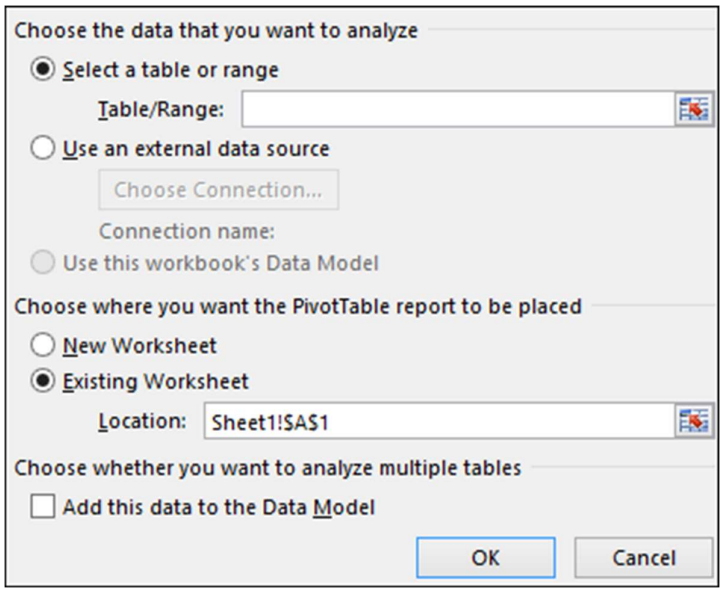
1.Select the cells you want to create a PivotTable from.

> **Note:** Your data should not have any empty rows or columns. It must have only a single-row heading.

2.Select **Insert** > **PivotTable**.



3.Under **Choose the data that you want to analyze**, select **Select a table or range**.



4.In **Table/Range**, verify the cell range.

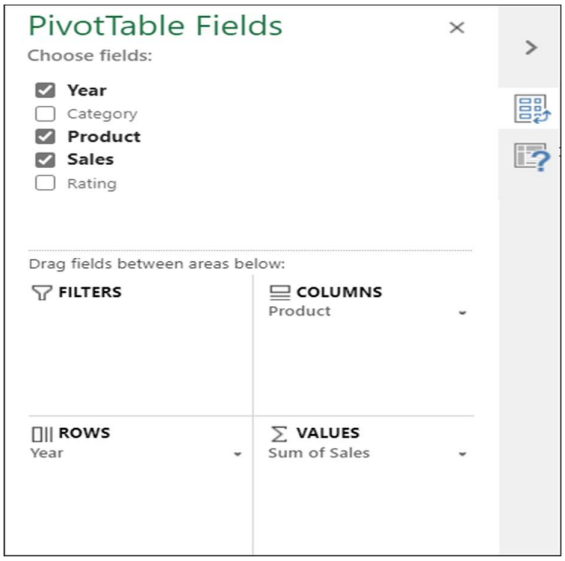
5.Under **Choose where you want the PivotTable report to be placed**, select **New worksheet** to place the PivotTable in a new worksheet or **Existing worksheet** and then select the location you want the PivotTable to appear.

6.Select **OK**.

Building out PivotTable

1.To add a field to your PivotTable, select the field name checkbox in the **PivotTables Fields** pane.

Note: Selected fields are added to their default areas: non-numeric fields are added to **Rows**, date and time hierarchies are added to **Columns**, and numeric fields are added to **Values**.



2.To move a field from one area to another, drag the field to the target area.

Excel provides you different types of charts that suit your purpose. Based on the type of data, you can create a chart. You can also change the chart type later.

Excel offers the following major chart types −

- Column Chart
- Line Chart
- Pie Chart
- Doughnut Chart
- Bar Chart
- Area Chart
- XY (Scatter) Chart
- Bubble Chart
- Stock Chart
- Surface Chart
- Radar Chart
- Combo Chart

Each of these chart types have sub-types. In this chapter, you will have an overview of the different chart types and get to know the sub-types for each chart type.

Column Chart

A Column Chart typically displays the categories along the horizontal (category) axis and values along the vertical (value) axis. To create a column chart, arrange the data in columns or rows on the worksheet.

A column chart has the following sub-types −

- Clustered Column.
- Stacked Column.
- 100% Stacked Column.
- 3-D Clustered Column.
- 3-D Stacked Column.
- 3-D 100% Stacked Column.
- 3-D Column.

Line Chart

Line charts can show continuous data over time on an evenly scaled Axis. Therefore, they are ideal for showing trends in data at equal intervals, such as months, quarters or years.

In a Line chart −

- Category data is distributed evenly along the horizontal axis.
- Value data is distributed evenly along the vertical axis.

To create a Line chart, arrange the data in columns or rows on the worksheet.

A Line chart has the following sub-types −

- Line

- Stacked Line
- 100% Stacked Line
- Line with Markers
- Stacked Line with Markers
- 100% Stacked Line with Markers
- 3-D Line

Pie Chart

Pie charts show the size of items in one data series, proportional to the sum of the items. The data points in a pie chart are shown as a percentage of the whole pie. To create a Pie Chart, arrange the data in one column or row on the worksheet.

A Pie Chart has the following sub-types −

- Pie
- 3-D Pie
- Pie of Pie
- Bar of Pie

Doughnut Chart

A Doughnut chart shows the relationship of parts to a whole. It is similar to a Pie Chart with the only difference that a Doughnut Chart can contain more than one data series, whereas, a Pie Chart can contain only one data series.

A Doughnut Chart contains rings and each ring representing one data series. To create a Doughnut Chart, arrange the data in columns or rows on a worksheet.

Bar Chart

Bar Charts illustrate comparisons among individual items. In a Bar Chart, the categories are organized along the vertical axis and the values are organized along the horizontal axis. To create a Bar Chart, arrange the data in columns or rows on the Worksheet.

A Bar Chart has the following sub-types −

- Clustered Bar

- Stacked Bar
- 100% Stacked Bar
- 3-D Clustered Bar
- 3-D Stacked Bar
- 3-D 100% Stacked Bar

Area Chart

Area Charts can be used to plot the change over time and draw attention to the total value across a trend. By showing the sum of the plotted values, an area chart also shows the relationship of parts to a whole. To create an Area Chart, arrange the data in columns or rows on the worksheet.

An Area Chart has the following sub-types −

- Area
- Stacked Area
- 100% Stacked Area
- 3-D Area
- 3-D Stacked Area
- 3-D 100% Stacked Area

XY (Scatter) Chart

XY (Scatter) charts are typically used for showing and comparing numeric values, like scientific, statistical, and engineering data.

A Scatter chart has two Value Axes −

- Horizontal (x) Value Axis
- Vertical (y) Value Axis

It combines x and y values into single data points and displays them in irregular intervals, or clusters. To create a Scatter chart, arrange the data in columns and rows on the worksheet.

Place the x values in one row or column, and then enter the corresponding y values in the adjacent rows or columns.

Consider using a Scatter chart when −

- You want to change the scale of the horizontal axis.
- You want to make that axis a logarithmic scale.
- Values for horizontal axis are not evenly spaced.
- There are many data points on the horizontal axis.
- You want to adjust the independent axis scales of a scatter chart to reveal more information about data that includes pairs or grouped sets of values.
- You want to show similarities between large sets of data instead of differences between data points.
- You want to compare many data points regardless of the time.

The more data that you include in a scatter chart, the better the comparisons you can make.

A Scatter chart has the following sub-types −

- Scatter
- Scatter with Smooth Lines and Markers
- Scatter with Smooth Lines
- Scatter with Straight Lines and Markers
- Scatter with Straight Lines

Bubble Chart

A Bubble chart is like a Scatter chart with an additional third column to specify the size of the bubbles it shows to represent the data points in the data series.

A Bubble chart has the following sub-types −

- Bubble
- Bubble with 3-D effect

Stock Chart

As the name implies, Stock charts can show fluctuations in stock prices. However, a Stock chart can also be used to show fluctuations in other data, such as daily rainfall or annual temperatures.

To create a Stock chart, arrange the data in columns or rows in a specific order on the worksheet. For example, to create a simple high-low-close Stock chart, arrange your data with High, Low, and Close entered as Column headings, in that order.

A Stock chart has the following sub-types −

- High-Low-Close
- Open-High-Low-Close
- Volume-High-Low-Close
- Volume-Open-High-Low-Close

Surface Chart

A Surface chart is useful when you want to find the optimum combinations between two sets of data. As in a topographic map, colors and patterns indicate areas that are in the same range of values.

To create a Surface chart −

- Ensure that both the categories and the data series are numeric values.
- Arrange the data in columns or rows on the worksheet.

A Surface chart has the following sub-types −

- 3-D Surface

- Wireframe 3-D Surface
- Contour
- Wireframe Contour

Radar Chart

Radar charts compare the aggregate values of several data series. To create a Radar chart, arrange the data in columns or rows on the worksheet.

A Radar chart has the following sub-types −

- Radar
- Radar with Markers
- Filled Radar

Combo Chart

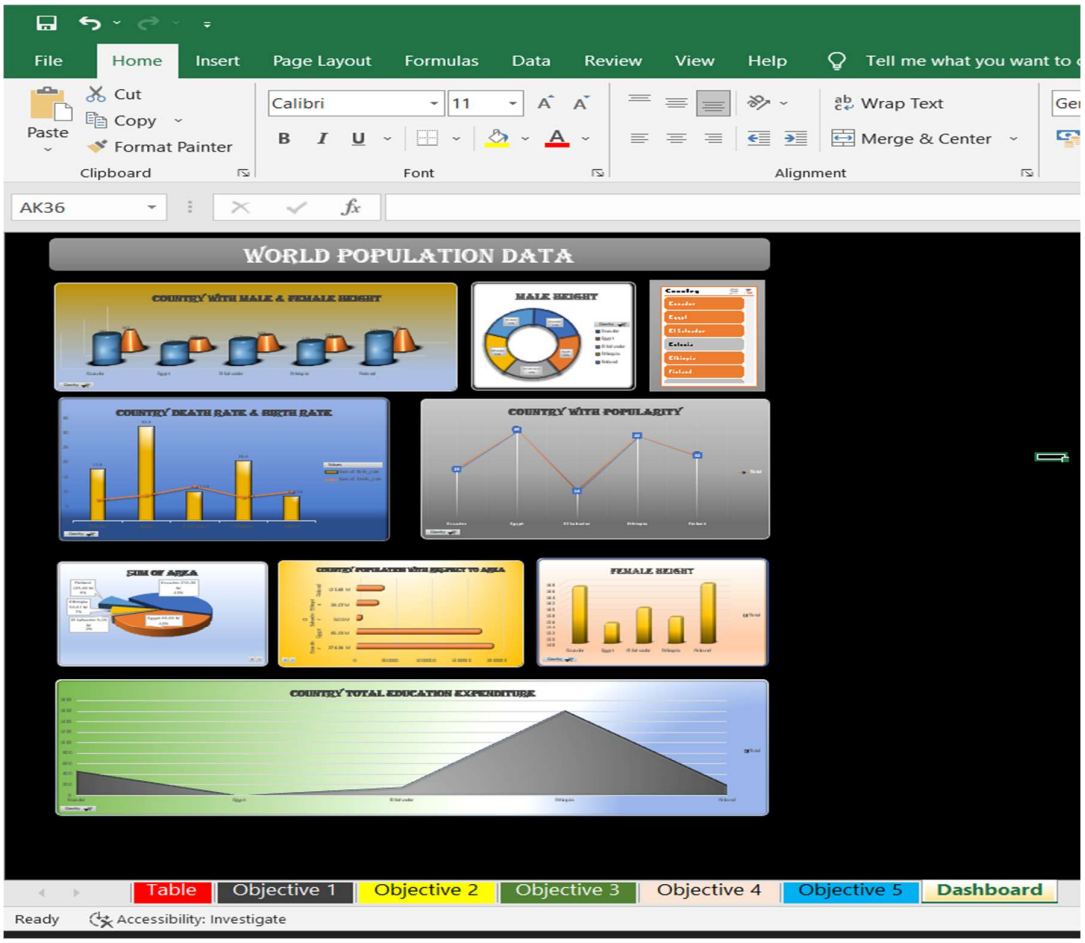
Combo charts combine two or more chart types to make the data easy to understand, especially when the data is widely varied. It is shown with a secondary axis and is even easier to read. To create a Combo chart, arrange the data in columns and rows on the worksheet.

A Combo chart has the following sub-types −

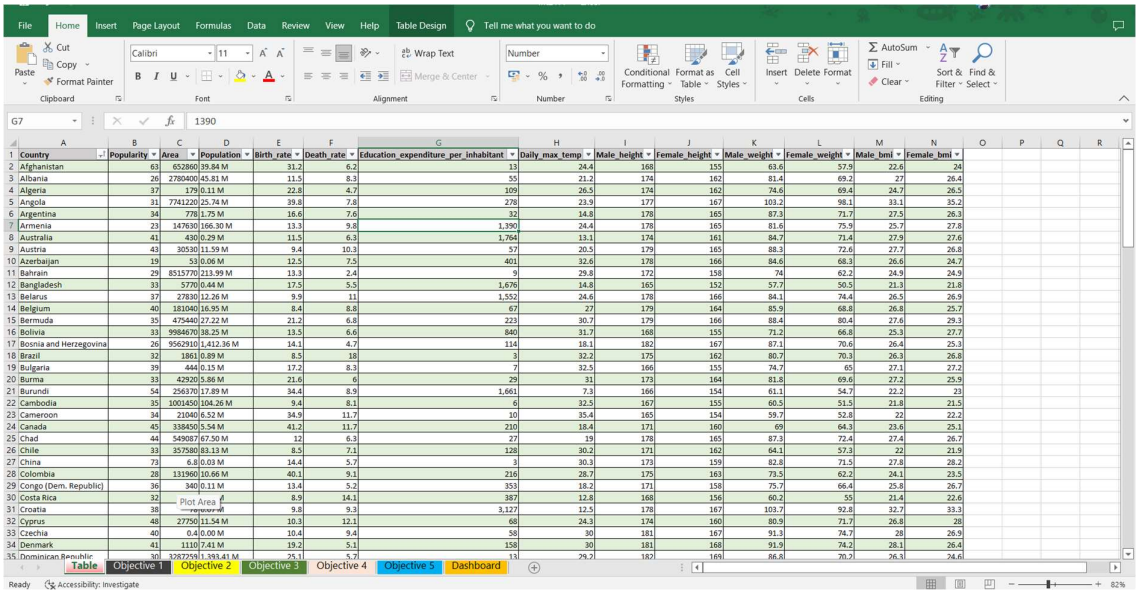
- Clustered Column – Line
- Clustered Column – Line on Secondary Axis
- Stacked Area – Clustered Column
- Custom Combination

CHAPTER-3 IMPLEMENTATION & RESULTS

Dashboard

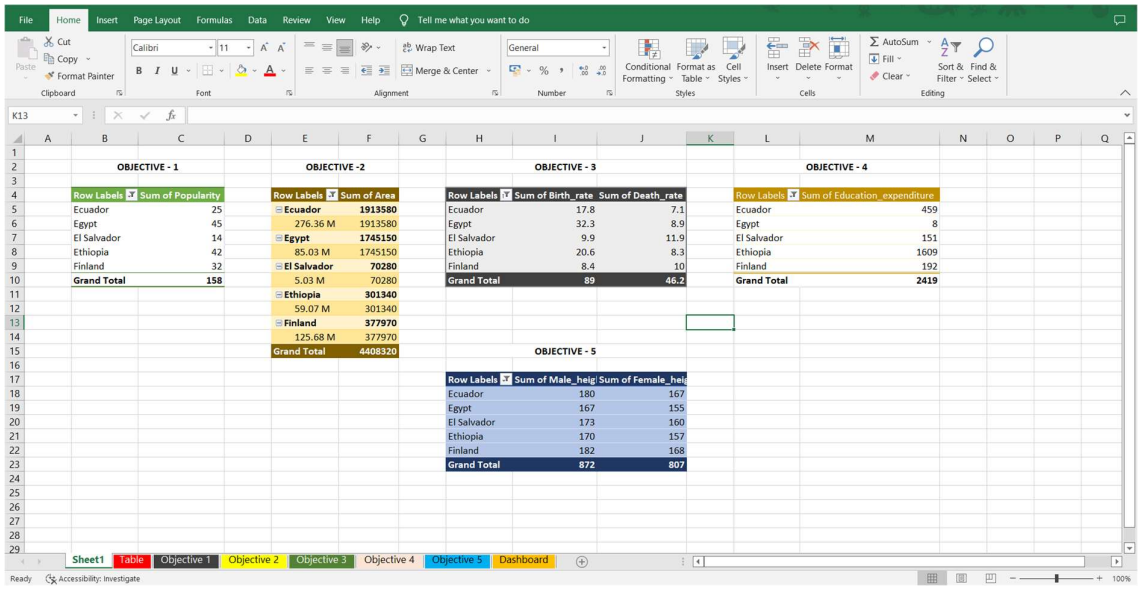


Dataset Table



Country	Popularity	Area	Population	Birth_rate	Death_rate	Education_expenditure_per_inhabitant	Daily_max_temp	Male_height	Female_height	Male_weight	Female_weight	Male_bmi	Female_bmi
Afghanistan	63	652860	39.84 M	31.2	6.2	13	24.4	168	155	63.6	57.9	22.6	24
Albania	26	2780400	45.81 M	11.5	8.3	55	21.2	174	162	81.4	69.2	27	26.4
Algeria	37	179	0.11 M	22.8	4.7	109	26.5	174	162	74.6	69.4	24.7	26.5
Angola	31	7741220	25.74 M	39.8	7.8	278	23.9	177	167	103.2	98.1	33.1	35.2
Argentina	34	778	1.75 M	16.6	7.6	32	14.8	178	165	87.3	71.7	27.5	26.3
Armenia	23	147630	166.30 M	13.3	9.8	1,390	24.4	178	165	81.6	75.9	25.7	27.8
Australia	41	430	0.29 M	11.5	6.3	1,764	13.1	174	161	84.7	71.4	27.9	27.6
Austria	43	30530	11.59 M	9.4	10.3	57	20.5	179	165	88.3	72.6	27.7	26.8
Azerbaijan	19	53	0.06 M	12.5	7.5	401	32.6	178	166	84.6	68.3	26.6	24.7
Bahrain	29	8515770	213.99 M	13.3	2.4	9	29.8	172	158	74	62.2	24.9	24.9
Bangladesh	33	5770	0.44 M	17.5	5.5	1,676	14.8	165	152	57.7	50.5	21.3	21.8
Belarus	37	27830	12.26 M	9.9	11	1,552	24.6	178	166	84.1	74.4	26.5	26.9
Belgium	40	181040	16.95 M	8.4	8.8	67	27	179	164	85.9	68.8	26.8	25.7
Bermuda	35	475440	27.22 M	21.2	6.8	223	30.7	179	166	88.4	80.4	27.6	29.3
Bolivia	33	9984670	38.25 M	13.5	6.6	840	31.7	168	155	71.2	66.8	25.3	27.7
Bosnia and Herzegovina	26	9562910	1,412.36 M	14.1	4.7	114	18.1	182	167	87.1	70.6	26.4	25.3
Brazil	32	1861	0.89 M	8.5	18	3	32.2	175	162	80.7	70.3	26.3	26.8
Bulgaria	39	444	0.15 M	17.2	8.3	7	32.5	166	155	74.7	65	27.1	27.2
Burma	33	42920	5.86 M	21.6	6	29	31	173	164	81.8	69.6	27.2	25.9
Burundi	54	256370	17.89 M	34.4	8.9	1,661	7.3	166	154	61.1	54.7	22.2	23
Cambodia	35	1001450	104.26 M	9.4	8.1	6	32.5	167	155	60.5	51.5	21.8	21.5
Cameroon	34	21040	6.52 M	34.9	11.7	10	35.4	165	154	59.7	52.8	22	22.2
Canada	45	338450	5.54 M	41.2	11.7	210	18.4	171	160	69	64.3	23.6	25.1
Chad	44	549087	67.50 M	12	6.3	27	19	178	165	87.3	72.4	27.4	26.7
Chile	33	357580	83.13 M	8.5	7.1	128	30.2	171	162	64.1	57.3	22	21.9
China	73	6.8	0.03 M	14.4	5.7	3	30.3	173	159	82.8	71.5	27.8	28.2
Colombia	28	131960	10.66 M	40.1	9.1	216	28.7	175	163	73.5	62.2	24.1	23.5
Congo (Dem. Republic)	36	340	0.11 M	13.4	5.2	353	18.2	171	158	75.7	66.4	25.8	26.7
Costa Rica	32	[illegible]	[illegible]	8.9	14.1	387	12.8	168	156	60.2	55	21.4	22.6
Croatia	38	[illegible]	[illegible]	9.8	9.3	3,127	12.5	178	167	103.7	92.8	32.7	33.3
Cyprus	48	27750	11.54 M	10.3	12.1	68	24.3	174	160	80.9	71.7	26.8	28
Czechia	40	0.4	0.00 M	10.4	9.4	58	30	181	167	91.3	74.7	28	26.9
Denmark	41	1110	7.41 M	19.2	5.1	158	30	181	168	91.9	74.2	28.1	26.4
Dominican Republic	30	3287290	1,393.41 M	25.1	5.7	13	29.2	187	169	86.8	70.2	26.3	24.6

PIVOT TABLE



OBJECTIVE - 1

Row Labels	Sum of Popularity
Ecuador	25
Egypt	45
El Salvador	14
Ethiopia	42
Finland	32
Grand Total	**158**

OBJECTIVE -2

Row Labels	Sum of Area
Ecuador	**1913580**
276.36 M	1913580
Egypt	**1745150**
85.03 M	1745150
El Salvador	**70280**
5.03 M	70280
Ethiopia	**301340**
59.07 M	301340
Finland	**377970**
125.68 M	377970
Grand Total	**4408320**

OBJECTIVE - 3

Row Labels	Sum of Birth_rate	Sum of Death_rate
Ecuador	17.8	7.1
Egypt	32.3	8.9
El Salvador	9.9	11.9
Ethiopia	20.6	8.3
Finland	8.4	10
Grand Total	**89**	**46.2**

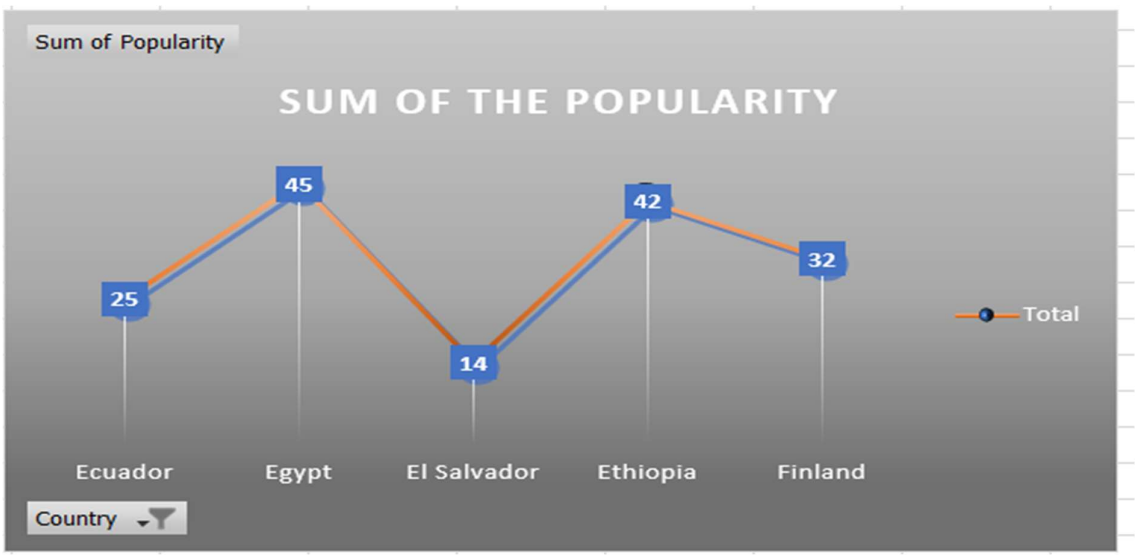
OBJECTIVE - 4

Row Labels	Sum of Education_expenditure
Ecuador	459
Egypt	8
El Salvador	151
Ethiopia	1609
Finland	192
Grand Total	**2419**

OBJECTIVE - 5

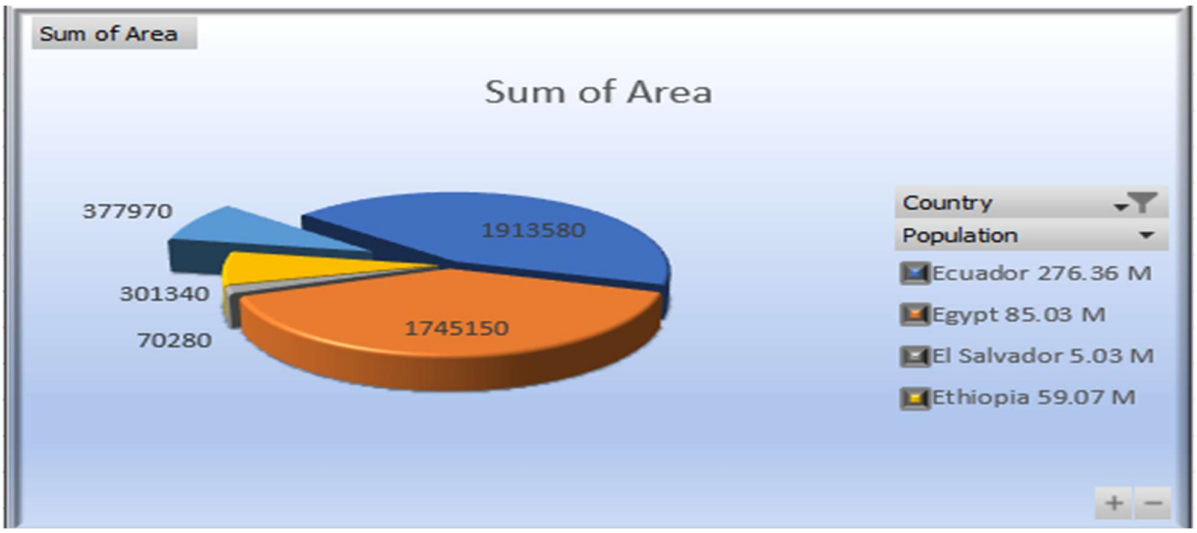
Row Labels	Sum of Male_heig	Sum of Female_heig
Ecuador	180	167
Egypt	167	155
El Salvador	173	160
Ethiopia	170	157
Finland	182	168
Grand Total	**872**	**807**

OBJECTIVE 1: - Line Chart



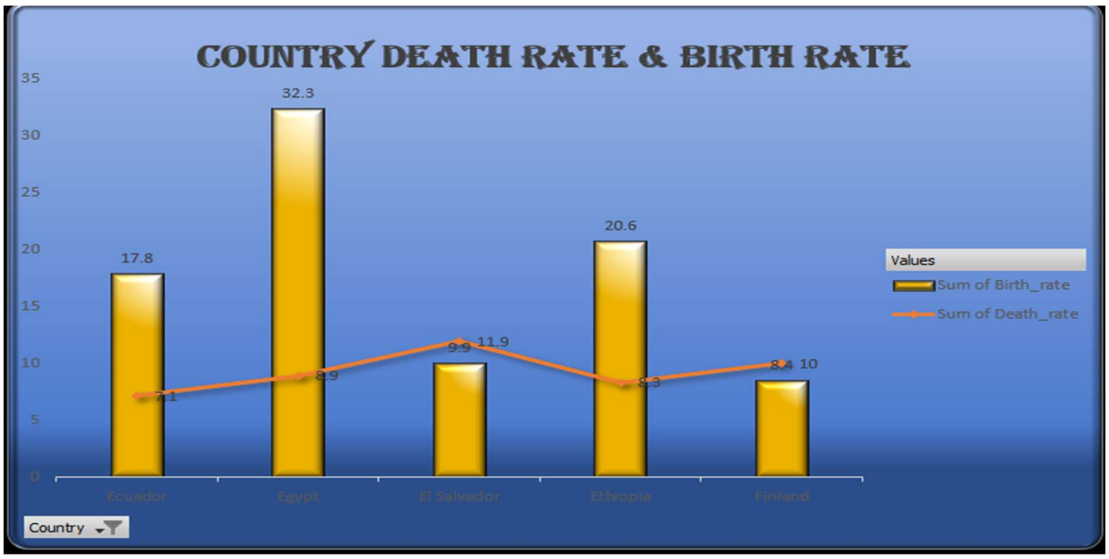
This chart shows the popularity with respect to different countries.

OBJECTIVE 2: - 3-D PIE CHART



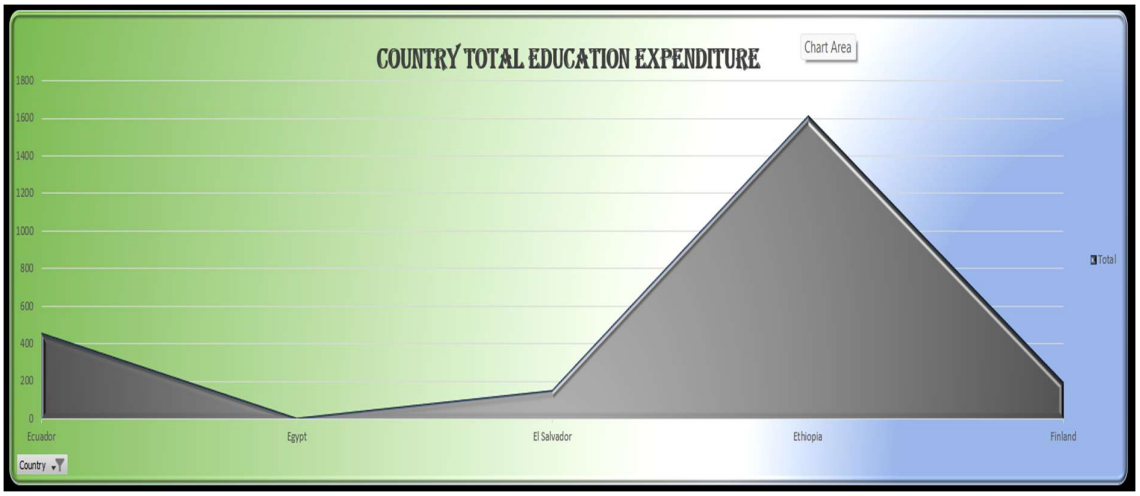
This 3-D pie chart shows the total area of different countries.

OBJECTIVE 3: -CLUSTERED BAR CHART



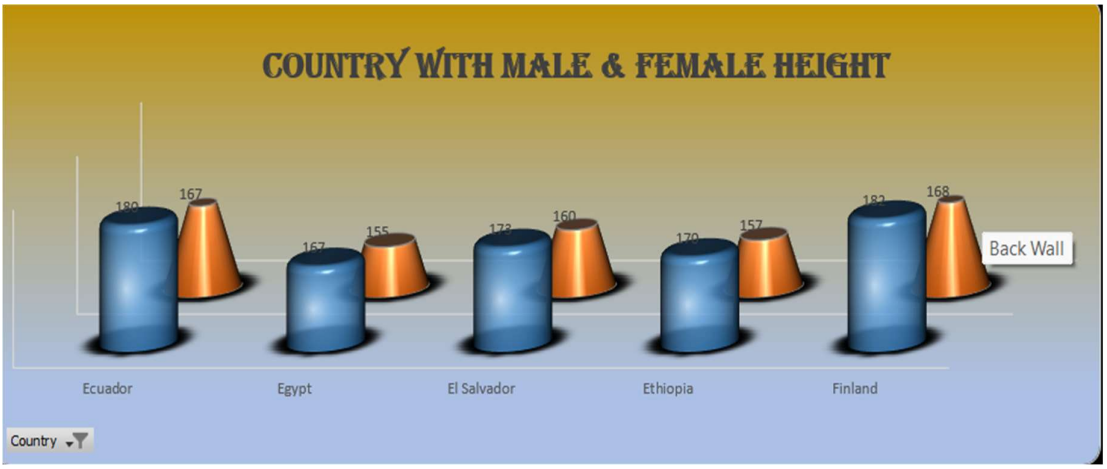
This bar chart shows the death rate and birth rate of different countries. It seems from the above graph Egypt is having a highest birth rate compared to other countries.

OBJECTIVE 4: LINE WITH MARKER CHART



This chart shows that the people having resting blood pressure high are more prone to heart failure.

OBJECTIVE 5: - 3D Cluster Column Chart



This column chart shows the male & female height of different countries.

CHAPTER-4 CONCLUSION

At present a population is increasing day by day. Previously there were no such tools to analyze whole country population. But this project provides a way to analyze population of different countries on the basis of different factors like country with popularity, death rate, birth rate, height, etc.