FINAL ASSESMENT – DATA SCIENCE WITH EXCEL

Q1) What are charts and different types of charts ?

Ans A chart is a poweful tool that allows you to visually display data in a variety of diff chart formats such as Bar , Coloum ,Pie , Line , Area , Surface , Scatter , Water fall , Trend Arrow , Guage , Themometer , Milestone etc A)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.

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

B) 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.

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

D) 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.

E) Gauge charts, also known as dial charts or speedometer charts, uses needles to show information as a reading on a dial. On a gauge chart, the value for each needle is read against the colored data range or chart axis. This chart type is often used in executive dashboard reports to show key business indicators. Gauge charts are useful for comparing values between a small numbers of variables either by using multiple needles on the same gauge or by using multiple gauges.

F) Thermometer chart is a visualization of the actual value of well-defined measure, for example, task status as compared to a target value.

G) A commitment to delivering is as important as the project itself. A milestone chart is an effective tool to depict project scope and timelines.

Q2) What are pivot tables in excel and their implementation ?

Ans Pivot Table is used to summarise, sort, reorganise, group, count, total or average data stored in a table. It allows us to transform columns into rows and rows into columns. It allows grouping by any field (column), and using advanced calculations on them.

Embedding frequently used lists, such as employees, departments or key customers into Excel's interface for use with any spreadsheet Filtering pivot table data based on date ranges by way of the Timeline feature in Excel 2013 and later

Q3) What is conditional Formatting ? Distinguish 5 types of conditional Formatting .

Ans Conditional formatting can help make patterns and trends in your data more apparent. To use it, you create rules that determine the format of cells based on their values, such as the following monthly temperature data with cell colors tied to cell values.

.**Background Color Shading** (**of cells**) -Color Scales change the color of each cell based on its value. Each color scale uses a two- or three-color gradient. For example, in the Green-Yellow-Red color scale, the highest values are green, the average values are yellow, and the lowest values are red.

.Foreground Color Shading (of fonts)

.**Data Bars** - Data Bars are horizontal bars added to each cell, much like a bar graph
.Icons - Icon Sets add a specific icon to each cell based on its value.
.Values

Q4) How to clear Fomatting in Excel without actually removing the cell content?

Ans First select all the cells by clicking on the arrow sign from the intersect point of the row and column number. After that, go to Home > Editing > clear and select Clear Formats. As a result, all the formatting of your entire dataset will be removed.

Q7) Define Statistics - Statistical functions are used when a mathematical process is required for a range of cells, such as summing the values in several cell locations. For these computations, functions are preferable to formulas because adding many cell locations one at a time to a formula can be very time-consuming. Statistical functions can be created using cell ranges or selected cell locations separated by commas. Make sure you use a cell range (two cell locations separated by a colon) when applying a statistical function to a contiguous range of cells. The Paste Formulas option is used when you need to paste formulas without any formatting treatments into cell locations that have already been formatted. You need to set multiple levels, or columns, in the Sort dialog box when sorting data that contains several duplicate values.

Q8) Explain Any two Data Analysis Tool Pack - The data analysis functions can be used on only one worksheet at a time. When you perform data analysis on grouped worksheets, results will appear on the first worksheet and empty formatted tables will appear on the remaining worksheets. To perform data analysis on the remainder of the worksheets, recalculate the analysis tool for each worksheet. If the Data Analysis command is not available, you need to load the Analysis ToolPak add-in program

B) F test - The F-Test Two-Sample for Variances analysis tool performs a two-sample F-test to compare two population variances. For example, you can use the F-Test tool on samples of times in a swim meet for each of two teams.

The tool provides the result of a test of the null hypothesis that these two samples come from distributions with equal variances, against the alternative that the variances are not equal in the underlying distributions.

The tool calculates the value f of an F-statistic (or F-ratio). A value of f close to 1 provides evidence that the underlying population variances are equal. In the output table, if f < 1 "P(F <= f) one-tail" gives the probability of observing a value of the F-statistic less than f when population variances are equal, and "F Critical one-tail" gives the critical value less than 1 for the chosen significance level, Alpha. If f > 1, "P(F <= f) one-tail" gives the probability of observing a value of the F-statistic greater than f when population variances are equal, and "F Critical one-tail" gives the critical value greater than 1 for Alpha.

B) Expotential - The Exponential Smoothing analysis tool predicts a value that is based on the forecast for the prior period, adjusted for the error in that prior forecast. The tool uses the smoothing constant *α*, the magnitude of which determines how strongly the forecasts respond to errors in the prior forecast.

Q9) Explain Histogram –

A) A Histogram is a column chart that shows the frequency of the occurrences of a variable in the specified range.

B) It is useful for producing data distributions and charts. You can analyze and interpret the distribution of data quickly and easily

To Add Histogam chart

- Go to the Data tab and click Data Analysis.
- In the Data Analysis dialog box, select Histogram.
- In the Histogram dialogue box, select the Input range, Bin range, and Output range. Checkmark the chart output. Click ok
- The first bin shows all the values below it. In the above case, 20 shows 0 values, which shows that there are 0 employees that are less than age 20.

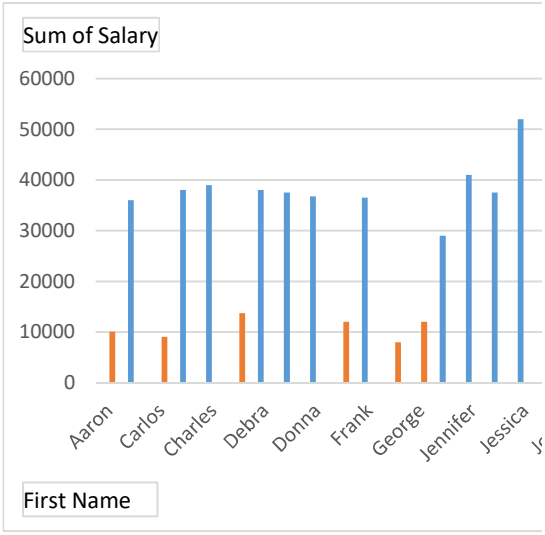
First Name	(All)

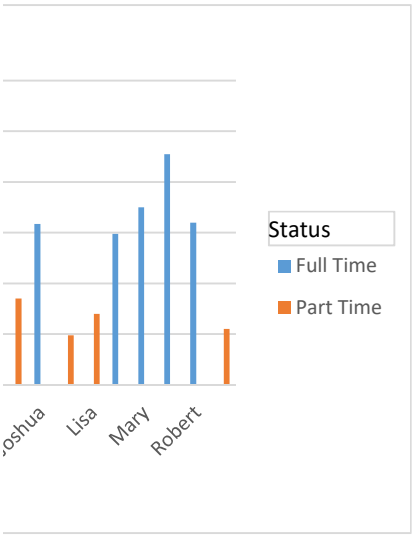
Sum of Salary	Column Labels		
Row Labels	**Full Time**	**Part Time**	**Grand Total**
Baxter	36000		36000
Boyle	38050		38050
Brown	71000	8000	79000
Buckner	37500		37500
Chandler	29000		29000
Coleman	37500		37500
Grimes		17000	17000
Hensley	52000		52000
Huber	31750		31750
Knowles		10050	10050
Moore	41000		41000
Morin	36500		36500
Munoz		11000	11000
Perry		12050	12050
Price		14000	14000
Schultz	38050		38050
Small	45500		45500
Steele	36750		36750
Tillman		9750	9750
Vang	29750		29750
Velazquez		9075	9075
Williams	35000		35000
Wilson		12000	12000
Wood		13750	13750
Grand Total	**595350**	**116675**	**712025**

Last Name	First Name	Status	Salary
Williams	Mary	Full Time	$ 35,000
Brown	Robert	Full Time	$ 32,000
Wilson	Elizabeth	Part Time	$ 12,000
Moore	Jennifer	Full Time	$ 41,000
Brown	Charles	Full Time	$ 39,000
Price	Lisa	Part Time	$ 14,000
Wood	Daniel	Part Time	$ 13,750
Coleman	Donald	Full Time	$ 37,500
Perry	George	Part Time	$ 12,050
Steele	Donna	Full Time	$ 36,750
Schultz	Carol	Full Time	$ 38,050
Munoz	Ruth	Part Time	$ 11,000
Chandler	Jason	Full Time	$ 29,000
Small	Matthew	Full Time	$ 45,500
Hensley	Jessica	Full Time	$ 52,000
Brown	Gary	Part Time	$ 8,000
Grimes	Jose	Part Time	$ 17,000
Baxter	Brenda	Full Time	$ 36,000
Morin	Frank	Full Time	$ 36,500
Tillman	Kathleen	Part Time	$ 9,750
Huber	Joshua	Full Time	$ 31,750
Boyle	Debra	Full Time	$ 38,050
Buckner	Jerry	Full Time	$ 37,500
Knowles	Aaron	Part Time	$ 10,050
Velazquez	Carlos	Part Time	$ 9,075
Vang	Marilyn	Full Time	$ 29,750

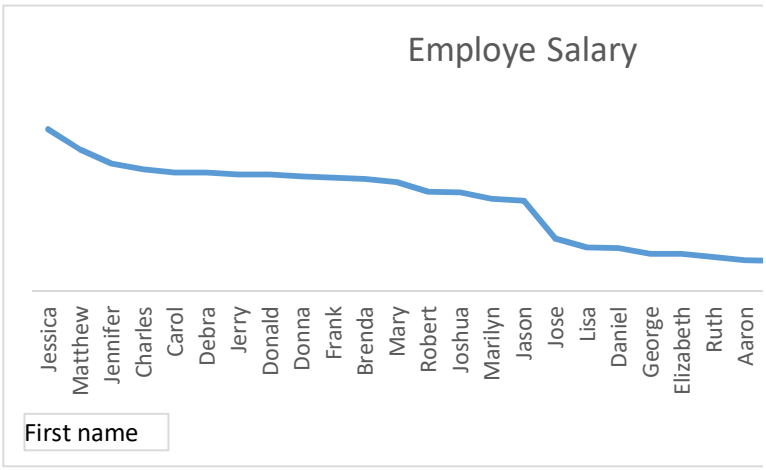
Sum of Salary
Row Labels
Aaron
Brenda
Carlos
Carol
Charles
Daniel
Debra
Donald
Donna
Elizabeth
Frank
Gary
George
Jason
Jennifer
Jerry
Jessica
Jose
Joshua
Kathleen
Lisa
Marilyn
Mary
Matthew
Robert
Ruth
Grand Total

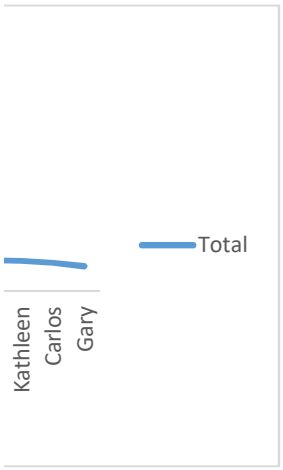
Column Labels		
Full Time	**Part Time**	**Grand Total**
	10050	10050
36000		36000
	9075	9075
38050		38050
39000		39000
	13750	13750
38050		38050
37500		37500
36750		36750
	12000	12000
36500		36500
	8000	8000
	12050	12050
29000		29000
41000		41000
37500		37500
52000		52000
	17000	17000
31750		31750
	9750	9750
	14000	14000
29750		29750
35000		35000
45500		45500
32000		32000
	11000	11000
595350	**116675**	**712025**

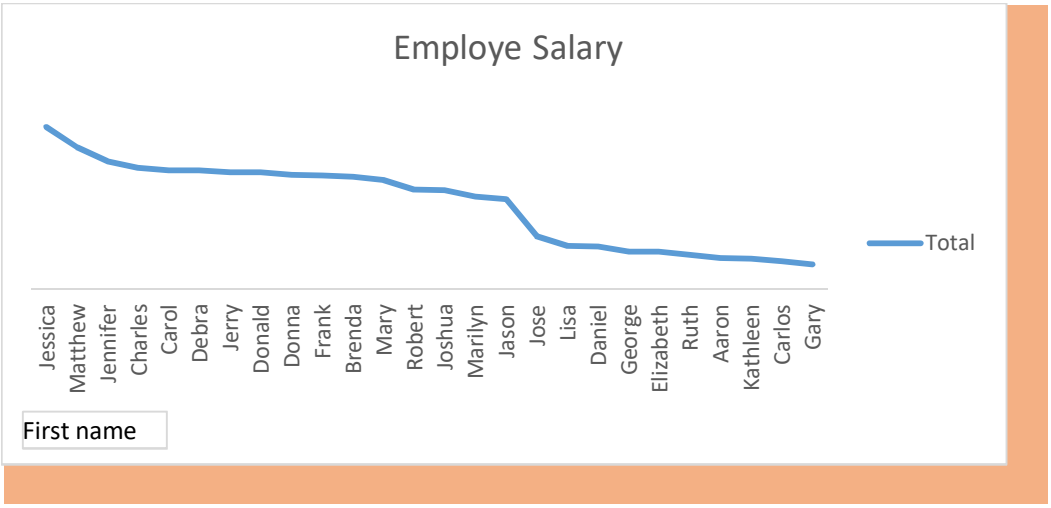



oshua
Lisa
Mary
Robert
Status
Full Time
Part Time

Row Labels	Sum of Salary
Jessica	52000
Matthew	45500
Jennifer	41000
Charles	39000
Carol	38050
Debra	38050
Jerry	37500
Donald	37500
Donna	36750
Frank	36500
Brenda	36000
Mary	35000
Robert	32000
Joshua	31750
Marilyn	29750
Jason	29000
Jose	17000
Lisa	14000
Daniel	13750
George	12050
Elizabeth	12000
Ruth	11000
Aaron	10050
Kathleen	9750
Carlos	9075
Gary	8000
Grand Total	**712025**




Total
Kathleen
Carlos
Gary


Employe Salary
Total
Jessica
Matthew
Jennifer
Charles
Carol
Debra
Jerry
Donald
Donna
Frank
Brenda
Mary
Robert
Joshua
Marilyn
Jason
Jose
Lisa
Daniel
George
Elizabeth
Ruth
Aaron
Kathleen
Carlos
Gary
First name