Analytics is scientific process to examine raw data to draw meaningful and logical insights.

4 distinct Types of Analytics Based on current and future needs of business.

* Descriptive—Explains what has happened.
* Diagnostic- States why it happened.
* Predictive--- depicts what could happen
* Prescriptive—talks what should happen.

Descriptive

Diagnostic—Digging deeper

Predictive—To condense data—statisticals, Data modeling, Data mining—latest and past trends

Prescriptive—Desc+Predictive—optimizes best solutions among different ones.

Analytics is divided bases on use and functionality of analytics across business divisions below examples.

Examples of areas of analytics

* Customer Analytics
* Industry focused analytics
* Financial Analytics
* Performance Analytics
* Risk Analytics

Customer Analytics🡪Organizations🡪Critical decisions🡪Offers🡪Customers analytics

Techniques:

1. Market Segmentation
2. Predictive analytics--- customer behavior—Customer Acquisition and Customer Retention in telecom companies.
3. Data modeling
4. Data Visualization

Financial Analytics:

Financial Analytics🡪Financial Executives-🡪Finance related business questions🡪forecast

It provides multiple views and derives insights from them to the necessary actions.

Ex. Consolidation of financial statements

Performance analytics:

🡪Data and Technology🡪Business Performance🡪Improve Performance

Ex. Managing daily operations, Budgeting, Planning strategies, meeting deadlines, pinpoint areas of improvement

Risk Analytics:

🡪try to foresee uncertainties🡪 Helps to evaluate project’s success or failure.

Define and understand-🡪Risk analytics🡪Manage and mitigate risks-🡪Wise decisions

Quantitative, Qualitative types of rsk

Quantifies the possible project results🡪Numerically evaluate the possibilities🡪Predict any of these possibilities are true.

Qualitative—Performed on all risks🡪Define various project related threats and risks🡪Avoid risks

Credit score in banking industry, corruption

**Formatting and Functions:**

Custom Formatting

Conditional formatting

Statistical and mathematical

Logical functions

|  |  |
| --- | --- |
| Format | Conditional formatting |
| Format the data  Make data easy to read  Present the data efficiently | Explore and visualize data  Detect critical issues  Identify patterns and trends  Highlights information by using colors, icons, data bars.  Changes appearance of one or more cells when cell value meets certain conditions |

Excel-Identifies Patterns& Trends and performs mathematical operations

|  |  |  |
| --- | --- | --- |
| Excel | Logical | Vlookup, Hlookup  If, And, Not  Rank  Quartile |
| Statistical |
| Mathematical |

These above functions will help to manage data and perform descriptive statistical analysis.

Formulae: Basic and complex operations

Autosum

Lookup & Reference

Logical

Text

Financial

Date & Time

Conditional Formatting--- Applying rule of greater than---2.008

Logical functions:

If, And, Or, Not, true, False

=IF(G5>10000,"5%",IF(AND(G5>=5000,G5<=10000),"2.5%","1.5%"))

=IF(OR(G9>10000,H9>=5000),"Good","NI")

=(True)

=(False)

=not(G5)

Lookup functions:

Vlookup,Hlookup,Index, offset, Match

Vertical Lookup

Horizantal lookup

=VLOOKUP(B3,B2:E40,4,0) Value B3 should be in left most column

=VLOOKUP(1149,B2:E40,4,FALSE)

=HLOOKUP(H4,H8:AS11,4,0) Value H4 should be in top most row

Index, Match, Offset:

Index--- **Excel INDEX function** returns the value at a given position in a range or array

=MATCH(G5,$B$17:$B$1342,0) === lookup\_Array must be the column in which value specified or to search

$ is for fixed array--- Absolute reference fixed array and fixed cell position

Without $ it is relative reference--- Cell position is subject to change

<https://www.ablebits.com/office-addins-blog/2015/11/25/relative-absolute-reference-excel/>

=-INDEX($B$17:$I$1342,J5,$J$3)

=OFFSET($B$17,E7-1,$J$3-1)

??????

Statistical Functions:

SUMIFS ---- adding the cells defined by specified set of conditions while SUMIF is adding cells given by criteria.

COUNTIFS

PERCENTILE

QUARTILE

STDEV

Median

=PERCENTILE(H4:H1328,0.6)====7308

i.e.. 60 % of sales is less than or equal to 7308

=PERCENTILE(H4:H1328,0.25)

=QUARTILE(H4:H1328,1)== Gives same result as percentile but it can give values only 0, 25th. 50th, 75th, maximum percentile values

**PIVOT Table**

Used for

Calculate and Summarize

Analyze

Present

Explore

Creating a pivot table

Grouping – Grouping by the header values, group the data for any field added as row or column

Includes date and time

Custom calculation

New field creation in pivot table???? Use Calculated field

Calculated Item

Slicers easy to use filtering data in pivot table.

**Dash boarding:**

Interactive dashboards and its design and format

Form controls

Charts—thermometer ---- How much have we achieved from our target.

Pareto--- Secondary axis

Radio buttons, combo box, check box

Combo box--- drop down box

1. Who is my audience?

For managerial level: Associate level analysis

For CEO level: Business level analysis

1. Value of dashboard add

KPIs

1. What type of dashboard--- Operational, Analytical, strategical
2. No Scrolling--- Avoid fragmenting data
3. Add form controls
4. Actionable information

Charts:

Line,----trends in data

Bar,---cmparing data by isplaying values in horzantanl columns

Pie, -----Slicing of each portion

Column--- Variety of data and comparisons btwn them

Form controls: Gives us access to interact with our data

|  |  |
| --- | --- |
| Labels | Spin button |
| Group Box | List box |
| Button | Option Button |
| Check box | Scroll bar |
| Combo box |  |

**Analytics with excel**

What-If analysis----- Used to find out missing data in a table or in incomplete data

Creation of Histogram---Graphical representation of distribution of a variable

What-If Analysis Tools:

* Goal seek------ Describe Goal Seek to get desired result
* Data Tables---- Analyze results for various input values
* Scenario Manager----Comparing and creating results of different scenarios using scenario manager

Descriptive statistics provides summary of key statistics of dataset. Like Mean, Median. Average etc..

Solver--------Solving linear problems using solver Add-in--- Finds optimal solution for formula

Solver add in:

File🡪Otions🡪SSOlver addin

Data tab🡪Solver add in

**Data Analysis using statistics:**

* Many organizations depend on statistical analysis to

Organize data

Evaluate their performance

Predict future trends

* Statistical analysis used to

1. Collection
2. Examination
3. Summarization
4. Manipulation
5. Interpretation

Of Quantitative data to identify

* Causes
* Patterns
* Trends and
* Relationships

Learnings:

Creating Moving Average charts

Hypothesis testing

ANOVA

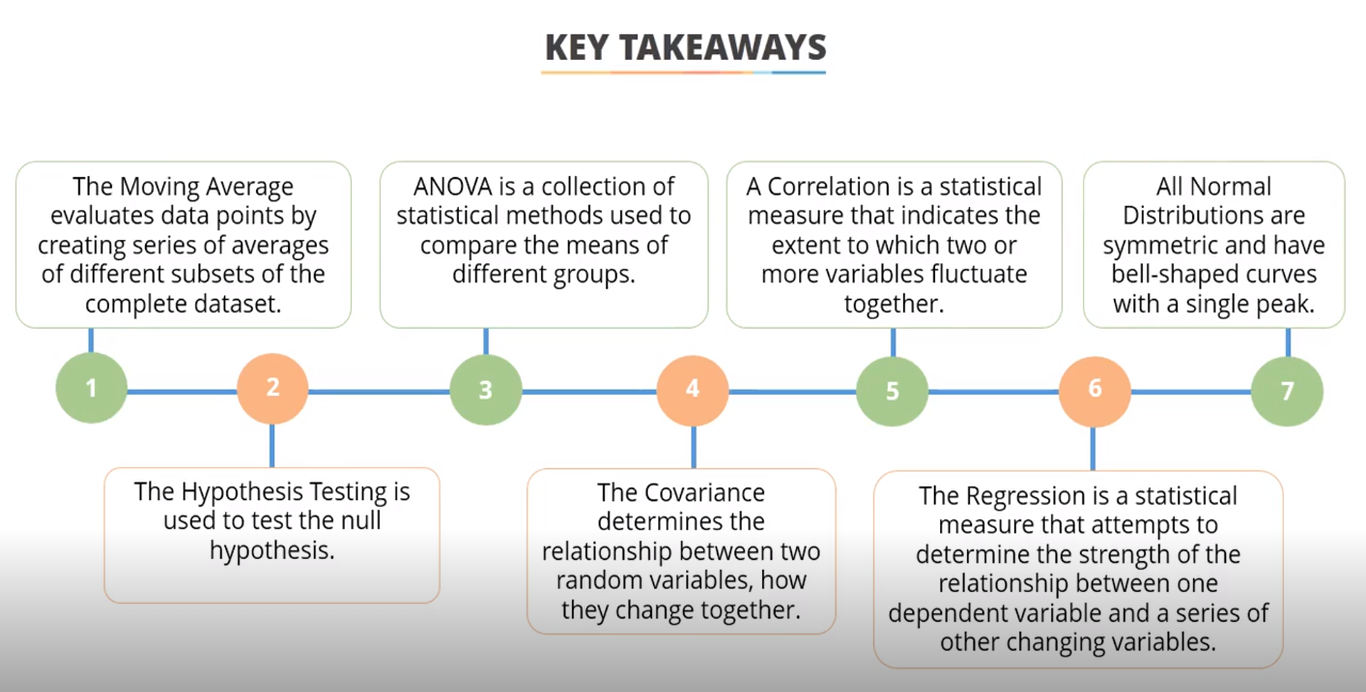
Covariance

Correlation

Regression

Normal Distribution

In Data Analysis tab



Hypothesis testing:

Ex. 1. Salary of men and women are equal.--- Null hypothesis until alternate hypothesis comes.

2. Salary of women is higher than men----- alternate hypothesis is true until it is proven false.

Ex2.Scores of student before and after taking course.----- t-test paired sample for mean.

The Null Hypothesis(H0) means that the mean/average of two populations is equal.

t-Test Two sample assuming unequal variances

The Null hypothesis is rejected, if

t Stat < -t Critical two-tail or

t Stat > t Critical two-tail

ANOVA- Statistical Method--- ANalysis Of Variance

Used to compare means of different groups.

|  |  |
| --- | --- |
| t-Test allows to analyze between 2 groups only | ANOVA allows to analyze between two or more groups. |

ANOVA Single factor.

F stat(observed) < F Critical -🡪Null hypothesis is not rejected.

F stat(observed) > F Critical -🡪Null hypothesis is rejected.

**Covarince:**

If X increases, Y increases or vice versa.

Positive Covariance if It increases.

Negative covariance if it decreases.

IF covariance is 0 then dataset doesn’t vary together.

**Correlation:**

Indicates the extent to which two or more variables fluctuate together

Correlation coefficient--🡪 -1<Coefficient value<1

If value is +1 then perfect correlation between variables🡪Positive correlation

If value is -1 then correlation between variables🡪Perfect negative correlation

Value = 0 indicates no correlation

=correl(array1, array2)

Scatter chart can also be created.

|  |  |
| --- | --- |
| Correlation | Covariance |
| 1. Values fall between -1 and +1 2. Correlation coefficient doesn’t depend on units of measure. 3. Used to compare similarities of multiple datasets that uses different units of measure or scale. 4. Change in one item results similar change in other. | 1. No specific measure of variation. 2. Values depends on Unites of measure of X and Y. 3. Covariance values of data sets using different set of measure are not comparable. |

**Regression**

Used to determine the relation between 1 dependent variable (Usually denoted by Y) and other series of independent variables.

**Normal distribution:**

Symmetric, Bell shaped curve with single peak.

Helps to define probability distribution of

* Rainfall
* Height
* Weight
* Manufacturing error
* Weight error
* Test scores

Noraml distribution curve= Mean + STD DEV

Normal density curves satisfy Empherical rule or 68-95-99.7% rule in statistics.

**POWER BI**

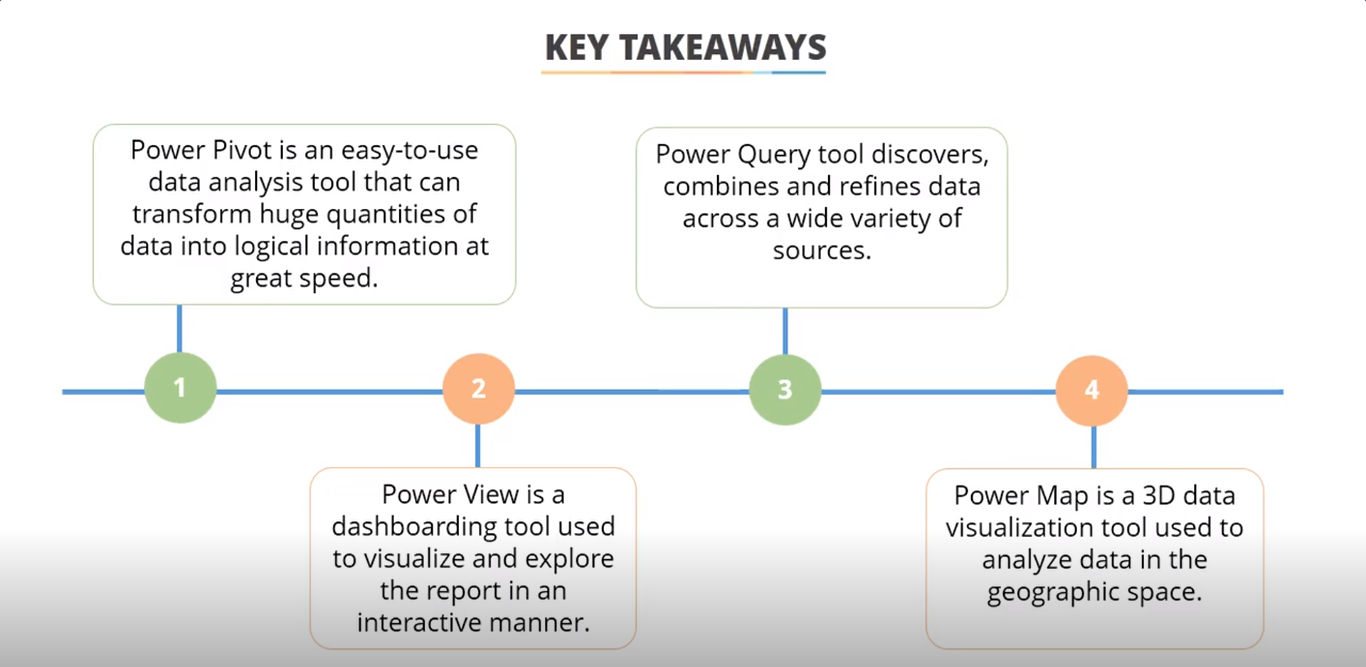
Features of Excel—Add Ins

Power Pivot

Power view

Power Query

Power Map



Learnings:

Using power pivot to create pivot table from multiple locations and different sources

Create interactive reports using power view

Use of Power Query to combine data from multiple sources

3D charts using power map.

Power View:

* Dash boarding tool
* Visualize and explore in interactive manner
* Provides intuitive ad adhoc reporting
* Improved versions of tools---charts, maps, slicer, hyperlink etc.

Power Query:

We can combine data from multiple worksheet into one with the help of Power Query.

Import manipulate transform append merge arrange consolidate automate--- of data process.

Discover, reshape and combine data across variety of sources like

1. Web
2. Hadoop
3. Azure Market place
4. OData
5. Relational
6. Structured and semi structured

It can search in public sources such as Wikipedia.

Data🡪From table🡪It opens Query Editor🡪close load to🡪only create connection

Power Map: