Power bi vs tableau

Q1)what is power bi ? and why we use it?and most important features of power bi ?

|  |  |  |
| --- | --- | --- |
| Excel | Power bi | Tableau |
| Excel is better at handling tabular style reports. | Power bi is a bi tool introduced by Microsoft which will help us to analyze, clean and convert that data into visual format. so that we can take better decisions related to business. | Tableau is more expensive than power bi.  It also can extract data from DB like pdf,excel,text,documents,Pyhton,hadoop etc. |
| **Allows displaying duplicate tables.**  **Does not offer advanced cross filtering between charts** | **1) Most preferable tool in India .search volume is increasing.**  **2)cost of power bi :cheapest tool**  **3)maximum numbers of features -1200 functions, Dax functions and 280 +graphs approx**  **4) Data connectivity: 100 different sources from diff sources.**  **5) Recognized by Gartner.** | Highly secure  Multiple data sources connection/  Tableau can run on windows and mac os. |
| Offers hig-level analytics | It performs faster and better when data volume is limited but becomes slow while handling bulk data. | Suitable for handling large volumes of data easily. |
| Excel does not have some new charts and they cannot connect to the data models. | It supports R language based visualizations.  Once report is generated than we can share on multiple system same report . | It provides full integrated support R and python. |
|  | It is suitable for small and medium sized organization.  Does not run on mac ,can run only on windows .most of the people prefer PBI because it coming from microsoft family.many companies who are going into a digital transformation they are choosing Microsoft azure as a cloud service. | It is suitable for medium and large sized organizations. |
|  | Exp:tableau,pbi,qlick,altyrex,looker |  |

**Q2)difference between managed enterprise and self-service enterprise?**

|  |  |
| --- | --- |
| Managed enterprise | Self-service enterprise(SSBI) |
| Complex programming skills are required to generate the reports.  It is the deployment OF BI throught a large corporation. | It is a approach to data analytics ,reporting and visualization that enables user to generate easy to understand, intuitive, and actionable dashboards without put so much efforts. |
| Unordered data flow, so efficiency will decrease, take lot of time, lots of efforts, lots of preprocessing. | Ordered data flow.  Efficient and take less time ,etc.  Analyze the data is easy.there is no need to expertise in bi tools . |
|  | 2 parts to Microsoft SSBI-excel bi toolkit and power bi |

**Q3)why are you interested in power bi ?what makes you think that power bi would be a great career option?**

It is a cloud based data reporting and visualization tool,power bi makes it possible for users to generate reports online. Once we generate any reports we can share with our colleagues in their workplace.Now a days power bi is most popular tool.Q&A facilities .less work,more efficiency .

**Power bi desktop**

**Get data –easily connect ,clean and mashup data**

Power bi 80+ sources ,both on –premises and cloud.

Shape,transform and clean data for analysis .

Live connectivity to on –premises and cloud data sources.

Extend with custom data connectors for any data science.

Prep your data using the familiar power query (transform data then open power query editor /is the heart of power bi)experience on the web .

Get started quickly with a common data model .

Extract self-service prep to azure data lake storage.

2)analyze-build powerful models and flexible measures.

Automatically create model when connecting to data .

High performance ,in memory engine

Point and click analysis with quick measures ,clustering &binning.

Create powerful measures with familiar DAX formulas.

Visualize :create stunning interactive reports.

**Q4). What are the versions of Power BI?/product portfolio.**

1)Power Bi Desktop /free

2)Microsoft power bi pro: Full version of BI. It allows users unlimited reporting ,sharing and viewing reports. And pro can publish their reports n number of times .

3)premium[need license for all user in an organization]

**Q5)What are the major components of Power BI?**

### ****Components of Power BI****

1)power query:Etl tool which will extract ,transform and load the data .it represents the user interface.

2)power pivot-use for data modeling when we have multiple data sources and want to connect them and create relationship between them.

3)power view:can help us to create 250+charts ,all use to create reports,presentation and dashbords.

4)power bi service:want to share reports with different peoples ./It is not nothing but it is just app.powerbi.com

5)power bi desktop:power bi desktop allow user to connect data+transform data+and also clean dataset.and click on apply and close .

Get data –analyze-visualize-publish-collaborate.

Note:::🡪power view is help us to create charts ,reports,presentation but power query is a etl tool which is use to filter the data,merge the data and calculate etc.

**Now just imagine A –EMPLOYEE ,B –MANAGER , A publish it and Manager spotted some issue and edit it.**

**Now A wants to do changes and A don’t have latest power bi . files. go to workspace and go to dashboard and download and you will get latest pbix file .**

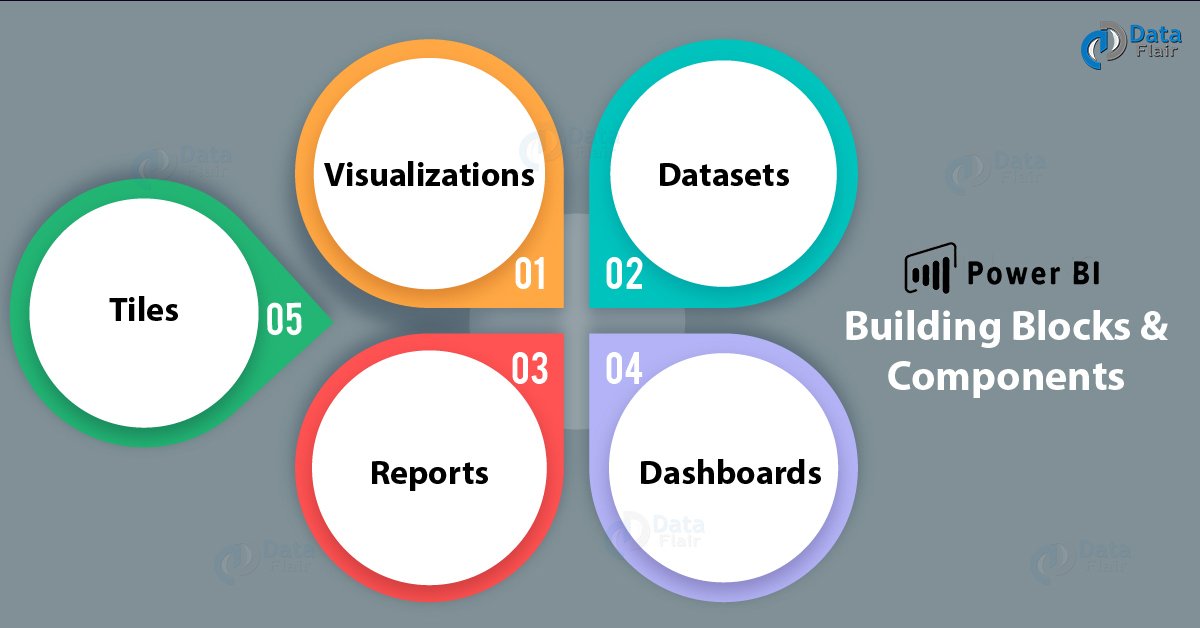
**Q6)Common workflow:**

1)begins by connecting to data sources and building a report in power bi desktop .

2)publish report from power bi desktop to the power bi service.

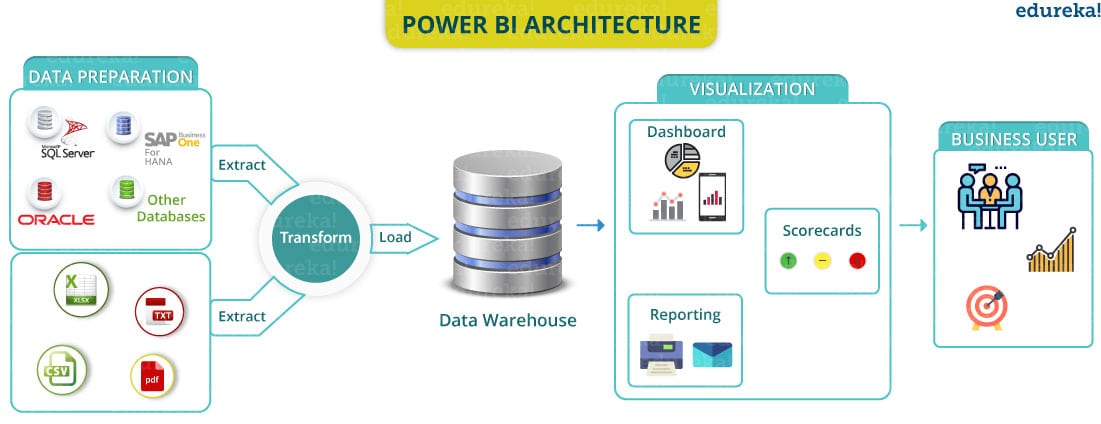
3)share it to end users with the **power bi service.**

4)mobile devices can view and interact with the reports.



VDA DRT

**Q7)Architecture :**



Common data sources:excel,csv,power bi datasets,sql server,analysis services ,text/csv

### Q8) ****What is the purpose of the ‘Get ’ icon in Power BI?****

Load the data and we can load multiple files also .

Common data sources:excel,csv,power bi datasets,sql server,analysis services ,text/csv

### Q9) ****How can we filter data in Power BI?****

In filter span we have only 3 types available 1)filters on this visual 2)filter on this pages 3)filter on all pages .

There are four standard types of filters:

Visual filter :we applies to a single visual on a report page .when we select a visual on the report canvas ,we see visual level filters. Suppose we have many charts I just only want to apply filter only any 1 chart –we apply filter on this visual.

Page filter :apply to all the visuals on the 1 page

Report filter :applies to all pages in the report .

**Drill through filter** :it use to drilldown and navigate through different pages. Create hierarchy through drill through.

Exp: suppose we make p4,p5 and p6 pages ..

Add drill through add segment or colname and then got the p4 page and right click you will see drill through option with page number .

**TOOL TIPS AND DRILL Through:**

Whenever we are working on dashboards .There could be multiple scenarios there we have might end up with lots of visualizations, reports.

When we have lots of visualizations on our report it will look like very clumsy or heavy. When someone see that clumsy report there were multiple things so other person will get confused.

Tool tips: Something which is visible when we hover a mouse on a particular graph. Which shows on you top of the visualization when we hover our mouse top on it. We are learning tool tip because showing a graph in tool tip .just like when I will hover in male then it will show age distribution of male.

It is showing gender and count .

Go to second page and select format page and tooltip and yes .

After that go to first page again and format then add tooltip and check out this tooltip.

Always show small graph .when we publish this report and check out .user can see only home page +tool tip. He is not able to see the second page.

Unhide and publish it again-in this case hide and publish it .

**Drill through** : I make a graph and I want to add Gender as drillthrough field .whenever we see a gender column in home page .User wants to know more about particular section. we can use drill through .

Add drill through: if we have more detailed explanation then we can use drill through and if we have 1 graph then use tooltip.

Just like we have p1,p2,p3 file and I want to navigate all those files so we use drill through.

**Q10) Difference between graph and chart?**

A graph is a chart used to show the mathematical relationship between various data sets by plotting horizontal (x-axis) and vertical (y-axis).x-y both are numbers.

Exp: bar graph , line graph

A chart represents information that can be in the form of a diagram, picture or graph. Which makes the user to understand the same better.

Types of charts :1)vertical bar chart 2)historical bar chart3)stacked bar charts

All graphs are charts but all charts not graphs.

### ****Q11)What are the essential applications of Power BI?****

**1) Project Management Office: Identify the current situation of various business unit .**

**2) Business and Data Analysts.**

**3) IT Professionals.**

**4) Reports the consumer.**

**Draw back/ limitations: It does not accept file size larger than 1 gb.It only shares dashboards and reports with users logged in with same email address.** Dashboard doesn't accept or pass user, account, or other entity parameters.

**Q12)WHAT ARE CONTENT PACKS?in workplace**

A content pack is **a feature that you can access if you have either the pro or premium version of Power BI**. With Power BI in general, you have shareable reporting options, such as dashboards. A content pack lets you boost your collaboration and organization within your reporting efforts.

Just like where you want to share dashboards with specific groups and my entire organization.

Title and description ,image upload as well ,and publish now content packs will create.

**Q13)WHAT are the different views in power bi ?**

In the power bi desktop in right side there are 3 things:1)report 2)data 3)view /model

**Q14)Security in power bi and row level security?**

RLS(ROW LEVEL SECURITY]we are working in any big company .our company have different centers –India ,parag ,new jursy.

In my dashboard we some graphs and all the information’s is here publicly. if hong kong user have raised a complained that someone is using my information .-called DATA BREACH-SOMEONE ELSE LOOK their data.

So what we will do we will create 4 dashobards ,it will take more space,more time ,and efficiency will decrease.

ROW LEVEL SECURITY :we can restrict the user just like Philippines user can see only their data. No other one can see their data .

Implement rls:modeling- manager roles

How to give access to user ?

Go to workplace[where reports are shared] –dataset—security-rowlevel security –click on france-give name and only those can see this report.

**Administration options**:how to administrator power bi dashboard?Which have permissions ?which can edit the report?

Admin[can have edit,delete]----

Member-

contributer[data analyst,edit access]

viewer[only view the reports].

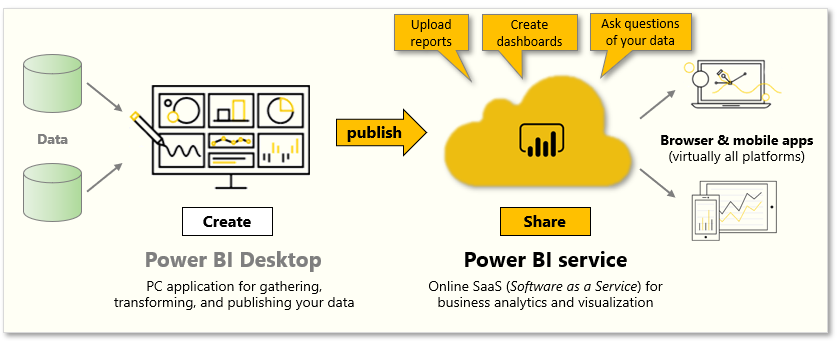
**Is it possible to refresh power bi reports after they are published to the cloud?**

### Power bi gateway: gateway is very helpful when we need to connect our power bi files on premises data sources. Software that is required to access data situated in an on-premises network. gateway act as a gatekeeper in an on-premises network.

Suppose I have on premises sql server I need to connect my power bi file .and then on power bi services I want to refresh my data on daily,monthly,weekly .according to this I want to refresh our data .

**Q15)Power bi services and integration with other apps.**

Power bi services:nothing but api.powerbi.com



### **Integration with power bi services:power bi can be integrated with multiple things .we can integrate power bi with Jupiter notebook.,teams,cottona.**

### ****EXCEL and power bi :**if we know excel then we can do easily work on power bi.**

### ****Q16)bidirectional /cross filtering?****

### **Ans:relationship between tables from both the direction and we use this bidirectional relationship to have the filter both side from the filter..**

### **1)Sales table**

### **2)manager table**

### **3)region data-multiple manager assign to multiple region ,multiple region have multiple manager ,means it have M-TO-M relationship**

### **How to resolve this ?**

### **1 manager have different region under her and have to pick up the sales from different region combine together for that particular manager.**

### **Between region and sales 1 to many relationship between tables .**

### **1-TO-M relationship is establish between but it is bidirectional region to manager .**

### ****Q17)M-code /language?****

### **Just like we load any data and transfer that data ,then it will open in power query editor .m language is written in power query.on the right hand side query setting.applied any steps it is generated.**

### **1)just like I go to add column and I applied format and upper case**

### **Steps is generated .**

### **TableAddColumn comes in formula bar ,**

### **SHIP MODE IN previous column name.all these steps are written in backside . view-advacned editor.**

### ****Q18)why should general formatting be applied to data in power bi?****

### ****Make a tiles –format painter-properties,title,effects,eader icons,tooltips,alt text.****

Formatting a table:right click+copy+copy visual and go to next page and paste

Increase the heading size:header column+change the size and colors

Add a title:general+title

Conditional formatting :make a table +conditional formatting applied to a numerical column .

Right click on columns lables+conditional formatting and performs operations.

We can customize the icons also just like add increase sign ,decrease sign etc.

### ****Q19)which in-memory analytics engine is used in power pivot?****

### in-memory analytics engine used in Power Pivot is called xVelocity, but it is commonly referred to by its original name, Vertipaq. [introduced in sql server 2008 r2].

**Q20)Where is data stored in Power BI?**

**When data is ingested into Power BI, it is basically stored in Fact and Dimension tables.**

### One Dimension Filters Another Dimension in Power BI - RADACAD

### ****FACT TABLES:**IT Contains measurement of business processes and it contain foreign keys for the dimension table**.in fact table it is always store numerical values.and foreign keys .****

### **Just like any sales tables is there it contains foreign keys of all other columns.foreign key-brach\_id**

### ****Dimension table:**is a table which contains attribute of measurements stored in the fact tables.it always contains text,descriptive,alphanumeric values .**

### **Load dim table first because without any primary key without any value .**

### **Fact table contains foreign key-Referential integrity .**

### ****Q21)Data profiling tools** ::provide new and intuitive ways to clean ,transform and understand data .**

### ****Includes:column quality,column distribution ,column profile****

### **Power query editor-view-column distribution, tells the distribution ,how many number are theirs .**

### **column quality[valid ratio,empty data ,errors] i-just identified it .**

### **column profile .tells about statistics about any particular columns and value distribution ..**

### ****Q22)Merging:**1 age column and after some time we have got another age same column names file we need to concatenate nothing but joins if we have common column then we can join it ..**

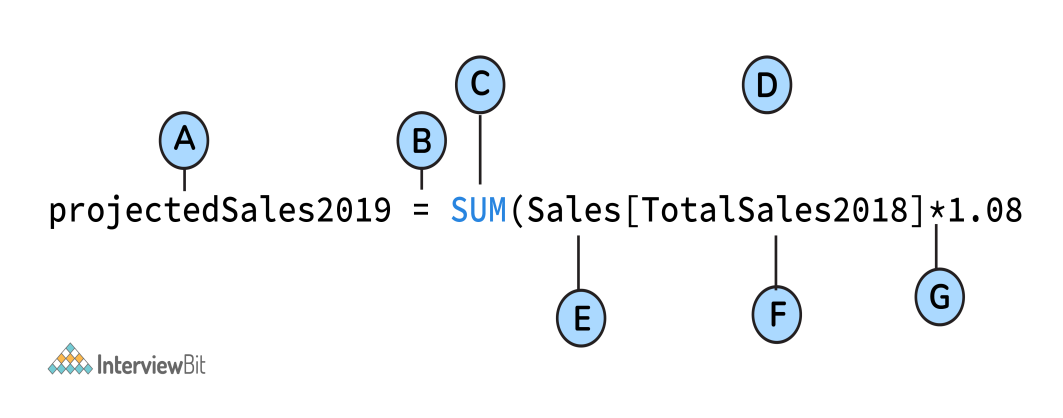
### **Exp:name+address-nameaddress**

### ****Appending**:when we have additional rows of data that we would like to add to an exisiting query.**

### **Exp:[1,2,3] [4,5] append [1,2,3,[4,5]]**

**Q23)Understanding DATA Analysis Expression**:

we have to express something in that.it’s come through Microsoft family.creating calculated columns,measures and custom tables. It is a collection of functions, operators, and constants that can be used in a formula, or expression, to calculate and return one or more values.



A: Name of the project  
B: Start of the DAX formula  
C: DAX function (to add)  
D: Parentheses defining arguments  
E:  Name of the table  
F: Name of the field  
G: Operator

**What do you need to know?**

Contexts:2 types of contexts 1)row context :we are adding a new column in existing file .rollno,each and every row have something different value.something changing row by row.

2)filter context:it is on overall table basis.mean value of age,just like a single value it will give.filtering out the values from the complete tables.

DAX:2 business logics:

1) **Measure** :It is nothing but it is an attribute which is calculated for the whole table.just like total sum of a column .it is done on the tabular basis.

Select sum(coln) from tablname;-gives a single value ./single number these values are called measures.

Exp:max age,total sum of age,total age .

2)calculated columns: newely created or derived columns .

We have age column 20,30,40 we are creating new column agebins 0-20 ,30 comes under 20-30 ,when we are creating new values based on the previous values .

Difference:context of evaluation-how they are evaluated .

|  |  |
| --- | --- |
| Calculated columns | Measure |
| Just evaluate row by row .thats why we called it row context | This is basically called filter context.and give single value . |
| Computed at the row level within the table it belong to. | Evaluated in the context of the cell evaluated in a report or in a DAX query. |
| Always computed at compile time. | Always computed at run time . |
| dynamic results ,based on rows. | Dynamic[means change in report,also change in result] results,based on filters |
|  | Not attached to any specific table/we can combine multiple tables and we can create a measure . |

**Implicit measures:**

If we use a calculated column or a measure to find another measure. Called implicit measure.

Exp:how old are you?50

Another tenure\_years -3,3,3,4,5—this is newly created column

X[special type of measure]=SUM(tenure\_years)

If I change in any single row ,then result will also changed.

DAX is great in 2 things1)aggregations 2)filtering

aggregations :combining a group of values into 1 value .

Exp:SUM,MAX,MIN,distinct count

Use card from visualizations –SUM,AVERAGE

Functions :sum,max,min,count,distinct count,countrows,datediff,dateadd

Dateiff:it will give the difference in the date.

Dateadd-add some days and create a new column .

FORMAT-CONVERT TEXT INTO number format.

FORMAT(TODAY(),”MMM”)-GIVE MONTH

TODAY()-TODAY DATE

CALCULATED TABLE

DATES=CALENDAR(range)-create a dates tables with a date per day between specified range .Also creates a date hierarchy.

Imagine we have time series data[where is 1 column is date column .and that date column always are in same time difference –jan,feb,mar,apr]

If there is any type of data [jan,feb,marc,may,june,augu] so we have to create another column and create in series .if we filter the data and it will give null ,here we create a calendar () and we combine data ,it will show month name but null value.

Create new table :create a new table

CALENDER();it returns a table with 1 column of all dates between start and end date.

Jan 1949—dec 1960 data u have.

Manager –create last 5 year data only.

Tablename=CALENDAR(DATE(1998,1,1),DATE(2000,1,31))

We can create new more columns year,month,date

MONTHNAME = FORMAT(newtable[Date],"MMMM")

**WHAT ARE THE BENEFITS OF CREATING A NEW QUERY FILE ?**

Suppose we are working in multiple files .all data in all different files .

we have 4 files .close and apply .when we start creating measure .

SUPPOSE YOU ARE ON LEAVE .and another person are finding measure so create a seprate query .in power desktop query.APPLY AND Close

In that new measure table create new measure.THIS is the best way to work.

Best practice organize your code

1)create seprate table for measures.

2) add drill though and tool tip

3)limit visuals

4)process as much as data as required in the original source.[work with smaller data first then change large dataset]

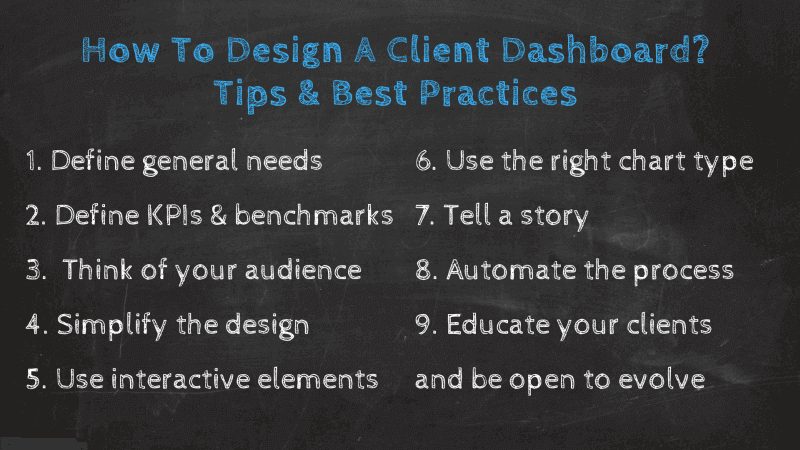
5)certified visuals are recommended

6)use a lighter background

Relationship in power bi:create primary key,unique key and reference key.

In model drag any 1 column in second table

Cardinality:



EDA: Exploratory data analysis.deals with the process of performing initial investigations on data with the help of summary statistics and graphical representation.

EDA :we have to find the information related to data .

Outline of performing eda :

Q1)what is the business problem?

Q2)what kind of data do you have and how do you treat different types?

Q3)what is missing fro the data and how to deal with them.?

Q4)where are the outliers and why should we care about them?

Q5)how can we add,change or remove features to get more out of your data?

Discover patterns

Spot anomalies [

Test hypothesis

Steps involved in eda

Data sourcing-data cleaning –categorical data/ features[1)univariate

2)bivariant 3)multi variant

–numerical data/-[co relation] --

derived metrices-derived a new variable from the existing variables

feature binning-feature encoding-from domain knowledge-calculated from data.

Check assumptions

Data cleaning :handling missing values .

**A)Delete rows /columns**

1)rows:can be deleted if it has an insignificant no of missing values

2)columns can be deleted if it>75%

**B)replace with mean,mode,median**

Can be used on an independent variable when it has numerical variables

Catgegorical features:apply mode method

C**)algorithm imputation**

machine learning algorithms exp:knn,navies bayes ,random forest

**D)predicting the missing values**

training set: data set with no missing values

testing set: data set with missing values

target variables: missing values

TYPES OF DATA

1)Qualitative:a variable to describe the quality of the population /categorical data [1)nominal –represent qualitative infor without order /value represents discrete units exp:m/f,this or that

2)ordinal:represemts qualitative info with order –economics status:low high and medium ,students grade

2)Quantitative :a variable to quantify the population /numerical data

1)discrete –count full values not decimal data-student number,total score in football.

Continuous-number within a range of values height