

Business Intelligence & Analytics

3: Credits TAE: 10M CAE: 15M ESE: 50M

UNIT 1: Introduction

Buzzwords . Analysis vs Analytics.
Business Analytics . Data Analytics and
Data Science . Adding BI and ML.
Info graphic . Data Analytics life cycle.

Buzzwords:

Business Intelligence and Analytics is a discipline that uses digital tools and strategies to analyse business data. The gathered information is used to improve business operations make future predictions and adapt strategies to fit observed trends.

4 Types of Business Analytics:

- 1) Descriptive , predictive and prescriptive analytics
- 2) Descriptive Analytics
- 3) Predictive Analytics
- 4) Prescriptive Analytics
- 5) A data - led future.

Difference Between Business Intelligence (BI)

&

Business Analytics (BA) -

BI → Deals with what happened in the past and how it happened leading up to the present moment. It identifies big trends and patterns without digging too much into the whys or predicting what will happen next

Business Analytics (BA)

Deals with the why's of what happened in the past. It breaks down contributing factors and causality. It also uses these why's to make predictions of what will happen in the future.

BI and BA have different applications. On one side, BI tools can help organizations in identifying relevant trends as well as the explanation to various past scenarios. On the other hand, BA is concerned with more advanced applications such as predictive analytics and statistic modeling.

Business Intelligence

Definition: What happened and why it happened
(descriptive Analytics)

Usage: Uses past and current data to track present performance and drive progress

Applications: Performance Management, visual insights, flexible reporting

Business Analytics

What will happen in the future?
(Predictive Analytics)

Uses historical data to generate predictions for future strategic decisions

Statistical Analysis, data modeling, financial forecasting.

1. Description vs prediction

A fundamental differentiation factor is in the method each of them uses as a base.

while BI tells you what has happened in the past and what is happening now (descriptive Analytics)

Descriptive Analytics:

As its name suggests this analysis method is used to describe and summarize the main characteristics found on a dataset. By identifying patterns and relationships, this method can communicate change over time and use trends and insights to drive the decision-making process.

In a business context it is used to compare performance to previous months, progress towards goals, gather client feedback, just to name a few.

Predictive Analytics:

This method uses advanced statistical techniques coming from data mining and machine learning technologies to analyze current and historical data and generate accurate predictions.

In a business context this method identifies patterns and trends and can

forecast inventory, predict customer responses to new products, assess risks among others.

Business Analytics

Data Analytics

Career Overview

The focus of business analytics is data and reporting - analyzing past business performance and predicting future business performance.

Data Analysts:

- Look at sales figures, competitive market research, and logistics data to interpret meaningful trends, gaps, and opportunities

- Communicate findings to stakeholders and senior management through summaries, graphs, and forecasting models.

Business Analysis

The focus of business analysis is functions and processes - identifying business needs and recommending solutions.

Business Analysts:

- Evaluate an organization's business model including documents, processes, systems, and technologies
- Gather and translate business problems and requirements into language understandable by the technical teams

- Skills Required** an organization
- Programming skills in the languages used at the organization
 - Data visualization and presentation skills
 - foundation in statistics (and potentially more advanced methods and algorithms, like machine learning NLP neural networks or "deep learning")
 - Ability to translate mathematical and statistical methods and results into business language to present to stakeholders and executives.
 - Technical and business knowledge
 - Understanding of the specific technologies that the organization is using.
 - Ability to speak in the technical terms used by internal departments
 - Ability to work well with others

Related Job Titles

Data Analyst	Data Manager	Data Technician
--------------	--------------	-----------------

Business Analyst
Operations Analyst
Functional Analyst
Strategy Consultant.

10 Analytics & Business Intelligence Buzzwords

- 1) Decision Intelligence
- 2) Data Science
- 3) X Analytics
- 4) Augmented Data Management
- 5) Predictive Analytics
- 6) Self - Service Business Intelligence
- 7) Mobile BI
- 8) Embedded Analytics
- 9) Collaborative BI
- 10) Cognitive Computing.

1) Decision Intelligence:

Machine learning algorithms being integrated into decision-making processes have created a field of decision models decision intelligence. It encompasses a range of decision-making techniques to design, model, align, execute and track decision models and processes. It observes, investigates models, contextualizes and executes decision models. As the human would not be able to process large volumes of data, decision intelligence will take care of such data volumes using ML algorithms.

2) Data Science:

Data science enables better decision-making, predictive analysis

and pattern discovery by using modern tools and techniques. Most business today hire data scientists to examine and interpret their data. However when BI Software will advance in the coming years, data science will be automated making it far more accessible and easier to interpret.

3) X Analytics:

X Analytics, a term coined by Gartner where X is the data variable for a range of different structured and unstructured content such as text analytics, video analytics, audio analytics etc. It is the ability to run any type of analytics on all of an organization's structured and unstructured data, regardless of where data resides and in what format. Combined with AI & other techniques such as graph analytics, X analytics will play a key role in recognizing, predicting and planning for natural disasters and other business crises & opportunities in the future.

Augmented Data Management:

It utilizes machine learning tools and AI techniques to optimize and improve operations. Augmented data management helps businesses secure high-quality data for real time analytics that converts into business decisions.

Traditionally, businesses' data analytics processes are dependent on the data team for traversing the organization hierarchy to scout the right data, cleanse it, model it, analyse it and derive insights. With augmented data management, companies can harness data through cross-department collaboration and accomplish various tasks.

Predictive Analytics

Predictive analytics defines as the practice of extracting information from existing data sets to predict future probabilities. It has already been used by a wide range of organizations, regardless of types and sizes. And its market is projected to

Role of Business Intelligence Analyst

Assesses market strategies by analyzing related products, markets, or trends. Uses business intelligence data and tools to identify and monitor current and potential customers. Identifies markets for future product development and methods to improve sales of existing products by analyzing technology trends.

The BI Analyst will gather, analyze and compile data needed to identify trends and patterns, then make recommendations for business actions.

Duties and Responsibilities:

- Gathers business intelligence from a variety of sources including company data, industry and field reports, public information, or purchased sources
- Assesses & Compiles business intelligence or trends to support actionable recommendations.
- Maintains current knowledge of industry and business trends through communication with professional organizations, suppliers, customers, competitors, and other informed individuals.
- Summarizes financial and economic data reports for review by executives, managers, clients, and stakeholders.
- Performs other related duties as assigned

Required skills / Abilities:

- Excellent verbal and written communication skills.
- Proficient in Microsoft Office Suite or related software.
- Ability to collect and compile relevant data.
- Extremely organized with great attention to detail.
- Excellent ability to analyse information and think systematically.
- Strong business analysis skills.
- Works well independently and as part of a team.
- Thorough understanding of the company's business processes and the industry at large.
- Ability to handle database and understand technology -driven business intelligence tools.

Education and Experience:

- Bachelor's degree in Business Management, Economics, Accounting, finance, or computer and Information Science required, Master's degree preferred.

Technical Skills Required:

- knowledge of data warehouse
- Data Modeling
- Data Analysis
- Data Mining
- Programming languages such as Python & C#
- Communication skills & Critical thinking

Database Management & reporting.

a. Does business intelligence need coding?

BI requires coding for processing data to produce useful insights. Coding is used in the data modeling and warehousing stages of the BI project lifecycle.

b. Business Intelligence Examples

- 1) Comprehensive platforms
- 2) Data Visualization
- 3) Embedded Software Applications.
- 4) Location Intelligence Software
- 5) self-service software built for non-tech users

Difference between business analyst and Business Intelligence

- Business Intelligence involves using historical data to make decisions about how a company should run in the present Day
- Business analysis may use historical data to predict what might happen in the future or how an organization can move forward.

what does a business intelligence analyst do?

→ BI Analyst uses data and other information to help organizations make sound business decisions.

- Breaking down key business data:

A business Intelligence analyst might gather, clean, and analyze data like revenue, sales, market information, or customer engagement metrics of a business. BI analysts can also be asked to program tools and data models to help visualize or monitor data.

- Interpreting the Data:

finding patterns or seeing areas in the data that signal a potential for improvement in business practices is a key part of a BI analyst's job.

Ex: A BI Analyst might analyze market trends to understand how a company might need to adapt its product.

- sharing findings: Sharing findings can include anything from visualizing data in graphs and charts, to putting reports together and presenting in front of other teams or clients. Business Intelligence analyst will also make recommendations to improve or grow the business based on their findings.

Data Analytics & Data Science:

Data Analytics:

Data analytics focuses on processing and performing statistical analysis of existing datasets. Analysts concentrate on creating methods to capture, process, and organize data to uncover actionable insights for current problems, and establishing the best way to present this data. More simply, the field of data and analytics is directed toward solving problems for questions.

What is Data Science?

Data Science is multidisciplinary field focused on finding actionable insights from large sets of raw and structured data. The field primarily fixates on unearthing answers to the things we don't know we don't know. Data science experts use several different techniques to obtain answers, incorporating Computer Science, predictive analytics, statistics, and machine learning to parse through massive datasets in an effort to establish solutions to problems that haven't been thought of yet.

Difference between Data Analytics & Data Science

Data Science is an umbrella term for a group of fields that are used to mine large datasets. Data analytics software is a

a more focused version of this and can even be considered part of the larger process-

- Analytics is devoted to realizing actionable insights that can be applied immediately based on existing queries
- Data science isn't concerned with answering specific queries; instead parsing through massive datasets in sometimes, unstructured ways to expose insights -
- Data Analysis works better when it is focused having questions in mind that need answers based on existing Data

Adding Business Intelligence & Machine learning

BI is an umbrella term that includes the applications, infrastructure and tools, and best practices that enable access to and analysis of information to improve and optimize decisions and performance.

The general consensus that BI generally covers these five areas:

1. Provides historical, current, and predictive views of business operations
2. leverages data that has been gathered into a data warehouse, data mart or operational data.

3. provides reporting , dashboards , visualization , queries , analysis and discovery.
4. Uses sales , marketing , operations , finance and many other sources of data
5. Provides on-premise software installed or SaaS - based which is hosted by the application service provider (ASP) or a mix of both on-premise and SaaS.

Now Adding BI & ML

Increasingly , BI is embracing features and capabilities that fuse machine learning (ML) and AI with traditional BI offerings . Advanced , predictive analytics is about calculating trends and future possibilities . Predicting analytics is about calculating trends and future possibilities . Predicting potential outcomes and making recommendation . That goes well beyond the traditional queries and reports that have defined (rightly or wrongly) BI in the past .

This shifts the traditional role of BI from "what happened," to an AI - driven model : which layers in answers to "what will happen next ?"

This is a significant shift , because it not only impacts the software but the data , people , and processes that

that need to support this evolution.
for ex: HBR recently published an article
"If your Data is Bad, Your machine
learning Tools are useless." and stated
that "Poor data quality is enemy
number one to the widespread,
profitable use of machine learning
while the caustic observation,
"garbage-in, garbage-out" has
plagued analytics and decision-
making for generations. it carries
a special warning for machine
learning.

"The garbage in, garbage-out" for
data is true not just for ML, but
for any use of data. Including BI.

What are the AI & ML Trends in
the BI Industry.

Tableau says it plans to add a new
machine learning recommendations engine
to its platform that will help algorithms
surface contextually relevant data.

Tableau acquired clearGraph which they
have stated will help enhance their
product with smart data discovery
and data analysis through Natural
language Processing (NLP). Tableau
will likely continue to acquire additional
expertise and products, as the needs

for AI and ML mature

Infographic:

Infographics are graphic visual representations of information, data or knowledge intended to present information quickly and clearly. They can improve cognition by utilizing graphics to enhance the human visual system's ability to see patterns and trends.

- An infographic is a collection of imagery, data visualizations like pie charts and bar graphs, and minimal text that gives an easy-to-understand overview of a topic.
- Infographics are a valuable tool for visual communication. The most visually unique creative infographics are often the most effective because they grab our attention and don't let go.

Infographic Templates

- 1> Design a simple productivity infographic
How to Build Good Habits to Boost productivity
- 2> Design a simple infographic on topic TeamWork
- 3> Design a simple Infographic on topic leadership

4) Design template using Infographic on a topic Engineer's Day.

5) Design creative template for Web Development

6) Tree Diagram for Business Management process

7) Tree Diagram for Digital / Internet Marketing Strategy.

Why should you use Infographics

- 1) Provide a quick overview of a topic
- 2) Explain a complex process
- 3) Display Research findings or Survey data
- 4) Summarize a long blog post or report
- 5) Compare and contrast multiple options
- 6) Raise awareness about an issue or cause.

How to create an Infographic

Infographic outline

1. Organize your information with an infographic outline

2. Pick an Infographic template

3. Customize your Infographic

1. Organize your information with an Infographic outline

The process of creating an outline will help you organize your thoughts and ensure that your content will work in an infographic.

Create an Infographic outline from existing content using these 4 steps

1. Determine the key takeaways of your content
2. Determine the title, headers, sub-headers, and facts

3. Consider the length of paragraphs and points

4. Include notes for the designer

1. Content Takeaways

what are the most important points that you want your audience to know? You can summarize your entire piece of content, or focus in on one key section.

2. Organize with sections & headers

Your title and subheaders should be short and descriptive. Each section of your infographic should only be a few points or a couple sentences.

3. Keep your Text Brief and to the point

To remain readable infographics typically don't have too much text.

keep your points concise and look for opportunities to use icons & charts to communicate information

4. Include Notes for the Designers
clearly lay out any design preferences and branding requirements that your designers should keep in mind. Including some inspirational images can help too!

Reasons why Infographics must be part of your web design.

1. Attention-Grabbing and visually Appealing
Research shows that our attention span is 8 seconds. The use of Infographics allows for easier consumption of content.
90% of the information that goes into your brain is visual.
2. Allows for Better Communication of the key Message.

If you're focused on improving your nonprofit website's design, you will have to include a lot of information on the site. Such could include case studies reports on the use of donor funds, and so much more. Audience never engage with text only. However, the perfect combination of text &

visualization will solve the problems

3. They Are fantastic for SEO:

Search Engine Optimization is critical for good rankings. Google uses specific criteria to determine where they place your site on the results pages. The higher the ranking, the better your visibility. It is critical for brand building and that is why landing on the first page is a significant coup.

Think about how you will present the infographics certain things help with SEO. Such as:

- keyword use - both long tail & short tail
- The quality of the information
- Image alt text
- Meta descriptions
- Titles & headers

4. The Role of Infographics & Link Building.

Link building is a powerful and very effective SEO technique. It is the process of creating linkages with other websites.

Your aim with link building should be to land on high authority sites. Data makes the information trustworthy. It makes your content relevant & most importantly factual.

5. Infographic will Increase Engagement
When audiences trust what you share, you become the go-to source for information. Several things will start happening. These will be better engagement with your content. The online visitor spends time on your website because the information is relevant. It reduces bounce rates and increases engagement.

Data Analytics Life Cycle

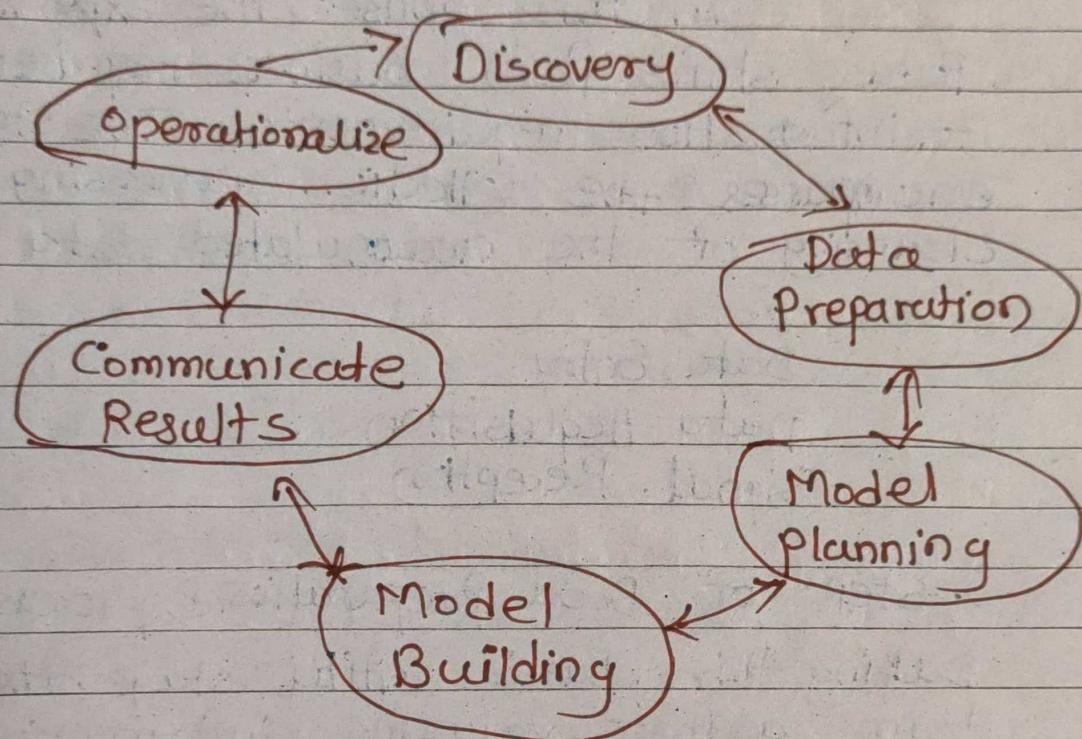
Importance of Data Analytics Lifecycle

Data Analytics Lifecycle defines the roadmap of how data is generated, collected, processed, used, and analyzed to achieve business goals. It offers a systematic way to manage data for converting it into information that can be used to fulfill organizational and project goals. The process provides the direction and methods to extract information from the data and proceed in the right direction to accomplish business goals.

Data Analysis

Data analysis is

modifying, processing, and cleaning raw data to obtain useful, significant information that supports business decision-making



1. Phase I - Data Discovery & formation

This phase is all about defining the data's purpose and how to achieve it by the end of the data analytics lifecycle. The stage consists of identifying critical objectives a business is trying to discover by mapping out the data.

- 1) The data Science team learn and investigate the problem.
- 2) Develop Context & Understanding
- 3) Come to know about data Sources needed & available for the project

- The team formulates initial hypotheses that can be later tested with data

phase 2: Data Preparation

In this phase, the experts focus shifts from business requirements to information requirements. The stage encompasses the collection, processing & cleansing of the accumulated data.

Data Entry

Data Acquisition

Signal Reception

• Steps of Data Preparation & processing

During this phase's initial stage, the team gathers valuable information & proceeds with the business ecosystem lifecycle. Various data collection methods are used for this purpose such as

Data Entry: Collecting recent data using manual data entry techniques or digital systems within the organization.

Data Acquisition: Gathering data from external sources

Signal Reception:

Capturing data from digital devices, including the Internet of

Things & Control Systems

- Steps to explore, preprocess, and condition data prior to modeling & analysis
- It requires the presence of an analytic sandbox, the team execute, load & transform, to get data into the sandbox.
- Data preparation tasks are likely to be performed multiple times & not in predefined order.
- Several tools commonly used for this phase are - Hadoop, Alpine Miner, Open Refine, etc

phase 3 : Model planning:

This phase needs the availability of an analytic sandbox for the team to work with data & perform analytics throughout the project duration. The team can load data in several ways.

Extract, Transform, Load (ETL)

It transforms the data based on a set of business rules before loading it into the sandbox

Extract, Load, Transform (ELT) -

It loads the data into the sandbox and then transforms it based on a set of business rules.

Extract, Transform, Load, Transform (ETLT)

It's the combination of ETL & has two transformation levels

- Team explores data to learn about relationships between variables & subsequently selects key variables & the most suitable models.
- In this phase, data science team develop data sets for training, testing & production purposes.
- Team builds & executes models based on the work done in the model planning phase.
- Several tools commonly used for this phase are - Matlab, Statistica.

Phase 4: Model Building:

- Team develops datasets for testing, training, and production purposes.
- Team also considers whether its existing tools will suffice for running the models or if they need more robust environment for executing models.
- Free or open-source tools - R and PL/I/R, Octave, WEKA.
- Commercial tools - Matlab, STATISTICA.

Phase 5: Communication Results

- After executing model team need to compare outcomes of modeling to criteria established

for success and failure.

- Team considers how best to articulate findings and outcomes to various team members and stakeholders, taking into account warning, assumptions.
- Team should identify key findings, quantify business value, and develop narrative to summarize and convey findings to stakeholders.

Phase 6: Operationalize

- The team communicates benefits of project more broadly and sets up pilot project to deploy work in controlled way before broadening the work to full enterprise of users.
- This approach enables team to learn about performance and related constraints of the model in production environment on small scale , and make adjustments before full deployment.
- The team ~~delivers~~ delivers final reports, briefings, codes.
- Free or open source tools - octave, WEKA, SQL, MADlib.

UNIT 2 - Business Intelligence Analysis Descriptive and Inferential Statistics

Descriptive Statistics:

population and Sample, Types of Data, Measurement Levels, Representation of categorical variables, Measures of Central Tendency (Mean, Median, Mode), Skewness, Variance, Standard Deviation, Coefficient of variation, Covariance, Correlation.

Population and Sample:

A population is the entire group that you want to draw conclusions about.

A sample is the specific group that you will collect data from. The size of the sample is always less than the total size of the population.

Population

- Population Parameter
- Mean of population, Median, Mode, Range
- Standard Deviation σ (sigma)

Sample

- Sample characteristics as sample statistics
- Mean of Sample, Median, Mode
- Standard Deviations s

Descriptive statistics are brief informational coefficients that summarize a given dataset, which can be either a representation of the entire population or a sample of a population

• population doesn't always refer to people
It can mean a group containing elements of anything you want to study such as objects, events, organizations, countries, species, organisms, etc.

Examples

population: Advertisements for IT jobs in the UK

sample: The top 50 search results for advertisements for IT jobs in the UK on 1 Aug 22

population: Songs from Indian Idle Song contest

sample: winning songs from the Indian Idle contest

population: Undergraduate students of the Indica

sample: 300 undergraduate students from SPPU who volunteer for NSS camps

Types of Data:

- 1) Nominal Data - Data can only be categorized
- 2) Ordinal Data - Data can be categorized & ranked
- 3) Interval Data: - , -- and evenly spaced
- 4) Ratio Data - Data can be categorized, ranked, evenly spaced and has a Natural zero.

Qualitative Data Type

Qualitative or Categorical Data Describes the object under consideration using a finite set of discrete classes. It means that this type of data can't be counted or measured easily using numbers & therefore divided into categories.

Ex: The gender of a person (male or female)

- These are usually extracted from audio images, or text medium. Another example can be of a smartphone brand that provides information about the current rating, the color of the phone, category of the phone & so on.

Subcategories

- 1) Nominal
- 2) Ordinal

1) Nominal:

These are the set of values that don't possess a natural ordering. phone color can not be compare as we can't say red color is greater than Blue. same for gender

Mobile phone categories whether it is midrange, Budget Segment or premium Smartphone's also nominal Data Type

Ordinal :

These types of values have a natural ordering while maintaining their class of values. If we consider the size of a clothing brand then we can easily sort them according to their ~~name tag~~ in the order of small < medium ~~>~~ large.

Grading system A+ < B+ < A+

Quantitative Data Types:

This data type tries to quantify things and it does by considering numerical values that make it countable in nature.

The price of a smartphone, discount offered, number of ratings on a product, the frequency of processor of a smartphone, or ram of that particular phone. all these things fall under the category of quantitative data types.

The key thing is that there can be an infinite number of values a feature can take. for instance, the price of a smartphone can vary from ~~X~~ amount to any value and it can be further broken down based on fractional values.

Sub categories

Discrete: The numerical values which fall under are integers or whole numbers are

placed under this category. The number of speakers in the phone, cameras, cores in the processor, the number of sins supported all these are some of the examples of the discrete data type.

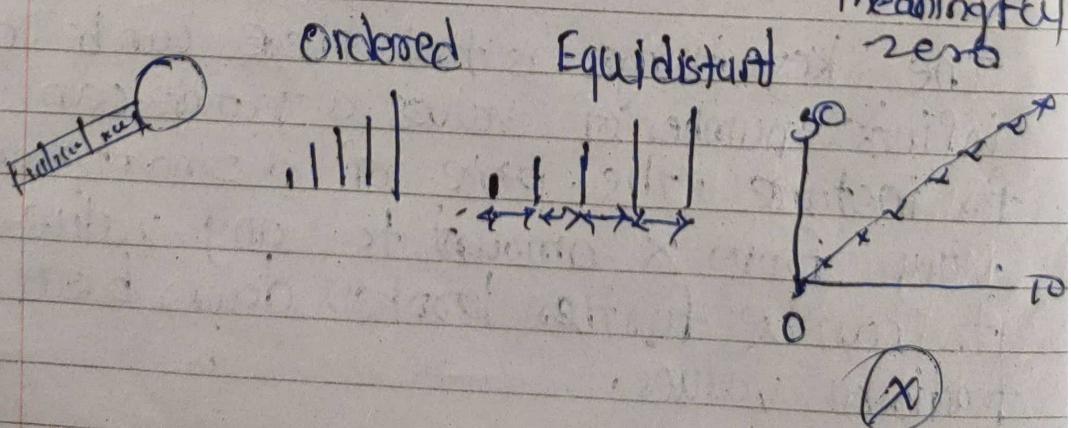
Continuous

The fractional numbers are considered as continuous values. These can take the form of the operating frequency of the processors, the android version of the phone, wifi frequency, temperature of the cores & so on.

Interval Data Definition:

Interval data is measured numerical data that has equal distances between adjacent values but no meaningful zero.

Measured



Interval Data are measured and ordered with equidistant items, but have no meaningful Zero. Interval data are ordered.

can be continuous (have an infinite number of steps) or discrete (organised into categories) and the degree of difference between items is meaningful (their intervals are equal but not their ratio)