**RESEARCH ON DATA ANALYSIS AND DATA ANALYTICS**

**1. DIFFERENCE BETWEEN DATA AND INFORMATION**

**Data:** Data is a collection of raw, unorganised facts and details like text, observations, figures, symbols and description of things etc. In other words, data does not carry any specific purpose and has no significance by itself. Moreover, data is measured in terms of bits and bytes which are basic units of information in the context of computer storage and processing.

**Information:** Information is the processed, organised and structured data. It provides context for data and enables decision making.

|  |  |
| --- | --- |
| **Data** | **Information** |
| Data is unorganised and unrefined facts | Information comprises processed, organised data presented in a meaningful context |
| Data is an individual unit that contains raw materials which do not carry any specific meaning | Information is a group of data that collectively carries a logical meaning. |
| Data doesn’t depend on information. | Information depends on data. |
| It is measured in bits and bytes. | Information is measured in meaningful units like time, quantity, etc. |
| Raw data alone is insufficient for decision making | Information is sufficient for decision making |
| An example of data is a student’s test score | The average score of a class is the information derived from the given data. |

**2. TYPES OF DATA**

There are two types of data.

* Qualitative data
* Quantitative data

**QUALITATIVE DATA:** It is non-numerical data. Eg: the texture of the skin, the color of the eyes, etc.

**QUANTITATIVE DATA:** Quantitative data is given in numbers. Data in the form of questions such as “how much”, “how many”, gives the quantitative data.

**3. ROLES AND RESPONEBILITIES OF DATA ANALYST**

* Data analysts work with data to help their organizations make better business decisions.
* Using techniques from a range of disciplines, including computer programming, mathematics, and statistics, data analysts draw conclusions from data to describe, predict, and improve business performance.
* Using automated tools to extract data from primary and secondary sources
* Removing corrupted data and fixing coding errors and related problems
* Developing and maintaining databases, data systems – reorganizing data in a readable format
* Performing analysis to assess quality and meaning of data
* Filter Data by reviewing reports and performance indicators to identify and correct code problems
* Using statistical tools to identify, analyze, and interpret patterns and trends in complex data sets that could be helpful for the diagnosis and prediction
* Assigning numerical value to essential business functions so that business performance can be assessed and compared over periods of time.
* Analyzing local, national, and global trends that impact both the organization and the industry
* Preparing reports for the management stating trends, patterns, and predictions using relevant data
* Working with programmers, engineers, and management heads to identify process improvement opportunities, propose system modifications, and devise data governance strategies.
* Preparing final analysis reports for the stakeholders to understand the data-analysis steps, enabling them to take important decisions based on various facts and trends.

**4. DIFFERENCE BETWEEN DATA ANALYSIS AND DATA ANALYTICS**

**Data Analytics:** Data analytics is the science of analysing raw data to make conclusions about that information. Many of the techniques and processes of data analytics have been automated into mechanical processes and algorithms that work over raw data for human consumption.

**Data Analysis: It** is defined as a process of cleaning, transforming, and modelling data to discover useful information for business decision-making. The purpose of Data Analysis is to extract useful information from data and taking the decision based upon the data analysis.

|  |  |  |
| --- | --- | --- |
| **Basis for Comparison** | **Data Analytics** | **Data Analysis** |
| **Form** | Data analytics is general form of analytics which is used in businesses to make decisions from data which are data-driven | Data analysis is a specialized form of data analytics used in businesses to analyse data and take some insights of it. |
| **Structure** | Data analytics consist of data collection and inspect in general and it has one or more users. | Data analysis consisted of defining a data, investigation, cleaning, transforming the data to give a meaningful outcome. |
| **Tools** | There are many analytics tools in a market but mainly R, Tableau Public, Python, SAS, Apache Spark, Excel are used. | For analysing 555555555555566 the data Open Refine, KNIME, RapidMiner, Google Fusion Tables, Tableau Public, NodeXL, Wolfram Alpha tools are used. |
| **Sequence** | Data analytics life cycle consist of Business Case Evaluation, Data Identification, Data Acquisition & Filtering, Data Extraction, Data Validation & Cleansing, Data Aggregation & Representation, Data Analysis, Data Visualization, Utilization of Analysis Results. | The sequence followed in data analysis are data gathering, data scrubbing, analysis of data and interpret the data precisely so that you can understand what your data want to say. |
| **Usage** | Data Analytics, in general, can be used to find masked patterns, anonymous correlations, customer preferences, market trends and other necessary information that can help to make more notify decisions for business purpose. | Data analysis can be used in various ways like one can perform analysis like descriptive analysis, exploratory analysis, inferential analysis, predictive analysis and take useful insights from the data. |

**5. TYPES OF DATA ANALYSIS APPLICATION**

There are two types of data analysis applications.

* Exploratory Data Analysis
* Confirmatory Data Analysis

**EXPLORATORY DATA ANALYSIS:**

* Exploratory is the method used to explore the big data set that will yield conclusions or predictions.
* The process entails “figuring out what to make of the data, establishing the questions you want to ask and how you’re going to frame them, and coming up with the best way to present and manipulate the data you have to draw out those important insights.
* Using exploratory analysis, data analysts are looking for clues and trends that will help them come to a conclusion.
* The processes of EDA involve a myriad of tasks, including spotting mistakes and missing data
* Identifying important variables in the data set; testing a hypothesis related to a specific model; and establishing a model that can explain the data in the most succinct way possible.
* It also involves determining the best way to present the final assessment.

**CONFIRMATORY DATA ANALYSIS:**

* CDA is the process used to evaluate evidence by challenging their assumptions about the data.
* This part of the process is where they work backward from their conclusions and weigh the merits of the results of their work.
* It’s like examining evidence and questioning witnesses in a trial, trying to determine the guilt or innocence of the defendant.
* CDA involves processes like testing hypotheses, producing estimates, regression analysis (estimating the relationship between variables) and variance analysis (evaluating the difference between the planned and actual outcome).

**6. APPLICATION AREAS OF DATA ANALYTICS**

**Transportation:** Data analytics can be applied to help in improving Transportation Systems and intelligence around them. The predictive method of the analysis helps find transport problems like Traffic or network congestions.

**Logistics and Delivery:** There are different logistic companies like DHL, FedEx, etc that uses data analytics to manage their overall operations. Using the applications of data analytics, they can figure out the best shipping routes, approximate delivery times, and also can track the real-time status of goods that are dispatched using GPS trackers.

**Web Search or Internet Web Results:** The web search engines like Yahoo, Bing, Duckduckgo, Google uses a set of data to give you when you search a data. Whenever you hit on the search button, the search engines use algorithms of data analytics to deliver the best-searched results within a limited time frame.

**Manufacturing:** Data analytics helps the manufacturing industries maintain their overall working through certain tools like prediction analysis, regression analysis, budgeting, etc. The unit can figure out the number of products needed to be manufactured according to the data collected and analyzed.

**Security:** Data analyst provides utmost security to the organization; Security Analytics is a way to deal with online protection zeroed in on the examination of information to deliver proactive safety efforts

**Education:** It is mostly used in adaptive learning, new innovations, adaptive content, etc. Is the estimation, assortment, investigation, and detailing of information about students and their specific circumstance, for reasons for comprehension and streamlining learning and conditions in which it happens.

**Healthcare:** Applications of data analytics in healthcare can be utilized to channel enormous measures of information in seconds to discover treatment choices or answers for various illnesses.

**Military:** Military applications of data analytics bring together an assortment of specialized and application-situated use cases. It empowers chiefs and technologists to make associations between information investigation and such fields as augmented reality and psychological science that are driving military associations around the globe forward.

**7.TYPES OF DATA ANALYTICS**

There are four types of data analytics.

* Descriptive Analysis
* Diagnostic Analysis
* Predictive Analysis
* Prescriptive Analysis

**DESCRIPTIVE ANALYSIS**

* Descriptive analytics are the backbone of reporting—it’s impossible to have business intelligence tools and dashboards without it. It addresses basic questions of “how many, when, where, and what.”
* Once again, descriptive analytics can be further separated into two categories: ad hoc reporting and canned reports.
* A **canned report** is one that has been designed previously and contains information around a given subject. An example of this is a monthly report sent by your ad agency or ad team that details performance metrics on your latest ad efforts.
* **Ad hoc reports**, on the other hand, are designed by you and usually aren’t scheduled. They are generated when there is a need to answer a specific business question.
* These reports are useful for obtaining more in-depth information about a specific query.
* An ad hoc report could focus on your corporate social media profile, examining the types of people who’ve liked your page and other industry pages, as well as other engagement and demographic information.
* Its hyper specificity helps give a more complete picture of your social media audience. Chances are you won’t need to view this type of report a second time (unless there’s a major change to your audience).

**DIAGNOSTIC ANALYSIS**

* While not as exciting as predicting the future, analyzing data from the past can serve an important purpose in guiding your business. Diagnostic data analytics is the process of examining data to understand cause and event or why something happened.
* Techniques such as drill down, data discovery, data mining, and correlations are often employed.
* Diagnostic data analytics help answer why something occurred. Like the other categories, it too is broken down into two more specific categories: discover and alerts and query and drill downs.
* Query and drill downs are used to get more detail from a report.
* For example, a sales rep that closed significantly fewer deals one month. A drill down could show fewer workdays, due to a two-week vacation.

**PREDICTIVE ANALYSIS**

* It may be the most commonly used category of data analytics. Businesses use predictive analytics to identify trends, correlations, and causation.
* The category can be further broken down into predictive modelling and statistical modelling; however, it’s important to know that the two go hand in hand.
* For example, an advertising campaign for t-shirts on Facebook could apply predictive analytics to determine how closely conversion rate correlates with a target audience’s geographic area, income bracket, and interests. From there, predictive modelling could be used to analyse the statistics for two or more target audiences, and provide possible revenue values for each demographic.

**PRESCRIPTIVE ANALYSIS**

* Prescriptive analytics is where AI and big data combine to help predict outcomes and identify what actions to take.
* This category of analytics can be further broken down into optimization and random testing. Using advancements in ML, prescriptive analytics can help answer questions such as “What if we try this?” and “What is the best action?” You can test the correct variables and even suggest new variables that offer a higher chance of generating a positive outcome.