

## UNIT II

### 4 Types of Data - Nominal, Ordinal, Discrete, Continuous -

## **4 Types Of Data – Nominal, Ordinal, Discrete and Continuous**

### Introduction - Importance of Data

“Data is the new oil.” Today data is everywhere in every field. Whether you are a data scientist, marketer, businessman, data analyst, researcher, or you are in any other profession, you need to play or experiment with raw or structured data. This data is so important for us that it becomes important to handle and store it properly, without any error. While working on these data, it is important to know the types of data to process them and get the right results. **There are two types of data: Qualitative and Quantitative data**, which are further classified into:

The data is classified into four categories:

- Nominal data.
- Ordinal data.
- Discrete data.
- Continuous data.

Now business runs on data, and most companies use data for their insights to create and launch campaigns, [design strategies](#), launch products and services or try out different things. According to a report, today, at least 2.5 quintillion bytes of data are produced per day.

### **Types of Data**

#### Qualitative or Categorical Data

Qualitative or Categorical Data is data that can't be measured or counted in the form of numbers. These types of data are sorted by category, not by number. That's why it is also known as Categorical Data. These data consist of audio, images, symbols, or text. The gender of a person, i.e., male, female, or others, is qualitative data.

Qualitative data tells about the perception of people. This data helps market researchers understand the customers' tastes and then design their ideas and strategies accordingly.

The other examples of qualitative data are :

- What language do you speak
- Favorite holiday destination
- Opinion on something (agree, disagree, or neutral)
- Colors

The Qualitative data are further classified into two parts :

### Nominal Data

Nominal Data is used to label variables without any order or quantitative value. The color of hair can be considered nominal data, as one color can't be compared with another color.

The name "nominal" comes from the Latin name "nomen," which means "name." With the help of nominal data, we can't do any numerical tasks or can't give any order to sort the data. These data don't have any meaningful order; their values are distributed into distinct categories.

#### **Examples of Nominal Data :**

- Colour of hair (Blonde, red, Brown, Black, etc.)
- Marital status (Single, Widowed, Married)
- Nationality (Indian, German, American)
- Gender (Male, Female, Others)
- Eye Color (Black, Brown, etc.)

### Ordinal Data

Ordinal data have natural ordering where a number is present in some kind of order by their position on the scale. These data are used for observation like customer satisfaction, happiness, etc., but we can't do any arithmetical tasks on them.

Ordinal data is qualitative data for which their values have some kind of relative position. These kinds of data can be considered “in-between” qualitative and quantitative data. The ordinal data only shows the sequences and cannot use for statistical analysis. Compared to nominal data, ordinal data have some kind of order that is not present in nominal data.

#### Examples of Ordinal Data :

- When companies ask for feedback, experience, or satisfaction on a scale of 1 to 10
- Letter grades in the exam (A, B, C, D, etc.)
- Ranking of people in a competition (First, Second, Third, etc.)
- Economic Status (High, Medium, and Low)
- Education Level (Higher, Secondary, Primary)

#### Difference between Nominal and Ordinal Data

Nominal Data	Ordinal Data
Nominal data can't be quantified, neither they have any intrinsic ordering	Ordinal data gives some kind of sequential order by their position on the scale
Nominal data is qualitative data or categorical data	Ordinal data is said to be “in-between” qualitative data and quantitative data
They don't provide any quantitative value, neither can we perform any arithmetical operation	They provide sequence and can assign numbers to ordinal data but cannot perform the arithmetical operation
Nominal data cannot be used to compare with one another	Ordinal data can help to compare one item with another by ranking or ordering
<b>Examples:</b> Eye color, housing style, gender, hair color, religion, marital status, ethnicity, etc	<b>Examples:</b> Economic status, customer satisfaction, education level, letter grades, etc

#### Quick Check – [Introduction to Data Science](#)

#### Quantitative Data

Quantitative data can be expressed in numerical values, making it countable and including statistical data analysis. These kinds of data are also known as Numerical data. It answers the questions like “how much,” “how many,” and “how often.” For example, the price of a phone, the computer’s ram, the height or weight of a person, etc., falls under quantitative data.

Quantitative data can be used for statistical manipulation. These data can be represented on a wide variety of graphs and charts, such as bar graphs, histograms, scatter plots, boxplots, pie charts, line graphs, etc.

#### Examples of Quantitative Data :

- Height or weight of a person or object
- Room Temperature
- Scores and Marks (Ex: 59, 80, 60, etc.)
- Time

The Quantitative data are further classified into two parts :

#### Discrete Data

The term discrete means distinct or separate. The discrete data contain the values that fall under integers or whole numbers. The total number of students in a class is an example of discrete data. These data can’t be broken into decimal or fraction values.

The discrete data are countable and have finite values; their subdivision is not possible. These data are represented mainly by a bar graph, number line, or frequency table.

#### Examples of Discrete Data :

- Total numbers of students present in a class
- Cost of a cell phone
- Numbers of employees in a company
- The total number of players who participated in a competition
- Days in a week

#### Continuous Data

Continuous data are in the form of fractional numbers. It can be the version of an android phone, the height of a person, the length of an object, etc. Continuous data

represents information that can be divided into smaller levels. The continuous variable can take any value within a range.

The key difference between discrete and continuous data is that discrete data contains the integer or whole number. Still, continuous data stores the fractional numbers to record different types of data such as temperature, height, width, time, speed, etc.

#### Examples of Continuous Data :

- Height of a person
- Speed of a vehicle
- “Time-taken” to finish the work
- Wi-Fi Frequency
- Market share price

### Difference between Discrete and Continuous Data

#### Discrete Data

Discrete data are countable and finite; they are whole numbers or integers

Discrete data are represented mainly by bar graphs

The values cannot be divided into subdivisions into smaller pieces

Discrete data have spaces between the values

**Examples:** Total students in a class, number of days in a week, size of a shoe, etc

#### Continuous Data

Continuous data are measurable; they are in the form of fractions or decimal

Continuous data are represented in the form of a histogram

The values can be divided into subdivisions into smaller pieces

Continuous data are in the form of a continuous sequence

**Example:** Temperature of room, the weight of a person, length of an object, etc

## Conclusion

In this article, we have discussed the data types and their differences. Working on data is crucial because we need to figure out what kind of data it is and how to use it to get valuable output out of it. It is also important to know what kind of plot is suitable for which data category; it helps in data analysis and visualization. Working with data

requires good data science skills and a deep understanding of different types of data and how to work with them.

Different types of data are used in research, analysis, [statistical analysis](#), [data visualization](#), and data science. This data helps a company analyze its business, design its strategies, and help build a successful data-driven decision-making process.

## Types of data FAQs

What are the 4 types of data?

Data is classified into 4 types.

Nominal data.

Ordinal data.

Discrete data.

Continuous data.

Are nominal and ordinal discrete?

Nominal and ordinal data can act as both discrete and continuous.

What are nominal data and ordinal data?

Nominal Data is used to label variables without any order or quantitative value.

Ordinal data have natural ordering where a number is present in some kind of order by their position on the scale.

What is discrete data?

Discrete is data that can take only certain values.

What is the difference between continuous and discrete?

Discrete data is a sort of numerical data that consists of full, concrete numbers that have particular and fixed data values that are determined by counting. Complex numbers and fluctuating data values that be measured over a defined time frame are referred to as continuous data.

## Importance of learning Linear Algebra in Data Science

Our technology has advanced to the point where it is today thanks to a number of fields, including data science, AI and ML, robotics, and computer vision. As you begin to learn more about these technologies, you will run into a number of jargon or particular words that are used frequently in connection with them.

Terms such as support vector machines, Lagrange multipliers, and ridge regression are some examples that describe this. These terms may have been made up to keep outsiders at bay, but they reveal a lot about their connection to the standard linear algebra that we are all familiar with from our time in school. Therefore, understanding linear algebra and optimization theory is crucial for anyone learning data science. When using ML to solve problems or when attempting to make sense of the vast amounts of available data, you also need to learn data science and be familiar with its applications.

## **Linear Algebra: A Brief Introduction**

An area of mathematics called linear algebra has many uses in both science and engineering. Our scientists are lagging when it comes to having a deeper understanding of it, despite the fact that it is so important and has uses that are unimaginably broad. This is primarily due to the fact that few scientists regularly use discrete mathematics, which is why we find them in their work. For researchers and workers involved in the tech field, the fact that it is a part of mathematics' continuous branch, makes it less interesting. Let's be crystal clear about one thing. You will find it extremely difficult to learn and use several ML algorithms, including the deep learning ones if you don't even have a fundamental understanding of how linear algebra functions.

## **Why is Linear Algebra So Important in Data Science?**

Let's understand in the following points why it is essential to have a thorough understanding of linear algebra in the data science field.

### **1. Linear algebra can help you understand statistics better.**

For machine learning to effectively organize and integrate data, statistics are crucial. You must first understand how linear algebra functions if you want to better comprehend statistical concepts. Powerful statistical topics like multivariate analysis can be integrated into your project using the methods, operations, and notations of linear algebra. Take into account that you are working with patient information which includes critical data like heart rate, weight, blood pressure, etc. These are just a few of the many components that are responsible for the data set created that you are analyzing. For the sake of this conversation, let's assume that gaining weight will raise blood pressure. The existence of a linear relationship in this situation may not be too difficult to grasp. Therefore, you must have a solid grasp of linear algebra in order to better understand how changing one variable affects another.

### **2. Data science success is largely dependent on linear algebra.**

In terms of sophisticated mathematics, there seems to be no doubt that calculus triumphs over linear algebra. More than just assisting you with integration, differentiation, and limits, integral and differential

calculus also provides you with the foundational knowledge needed for applications like tensors and vectors. Your understanding of linear functions and equations will improve as a result of learning these things, among other things. The Simplex method and spatial vectors are two additional more advanced concepts that you will be familiar with. If you want to understand these ideas better, learning linear algebra is what you can start with.

### **3. Improved machine learning algorithms are made possible by the use of linear algebra.**

Building more effective supervised and unsupervised machine learning algorithms is possible using the knowledge you gain from linear algebra. With the help of linear algebra, you can build supervised learning algorithms from scratch, such as support vector machines, decision trees, logistic regression, and linear regression (SVM). You can even leverage it along with some popular unsupervised ML algorithms such as SVD, clustering, and components analysis. The ability to modify different parameters will be made possible by linear algebra, which will aid you in developing a deeper understanding of the machine learning project you are working on.

### **4. Data Science and ML prediction**

Learning linear algebra helps you develop the intuition or awareness that is so crucial to machine learning and data science. More viewpoints will be available from you now. Your ability to think broadly and unwaveringly will improve as a result of your study of matrices and vectors. There are numerous options. You might begin acting in ways that are very difficult for people to comprehend around you. You could start arranging and visualizing various graphs. For various data science and machine learning components, you could begin using more parameters.

Data science is a multi-faceted, interdisciplinary field of study. It's not just dominating the digital world. It's integral to some of the most basic functions - internet searches, social media feeds, political campaigns, grocery store stocking, airline routes, hospital appointments, and more. It's everywhere. What makes data science so applicable to the human experience? Among other disciplines, statistics is one of the most important disciplines for data scientists.

Josh Wills, a former head of data engineering at Slack, said "A data scientist is a person who is better at statistics than any programmer and better at programming than any statistician."

In other words, statistics is an inherently necessary component of data science. We'll explore more on this concept below, in addition to the best ways for learners to gain statistical knowledge for a data science position.

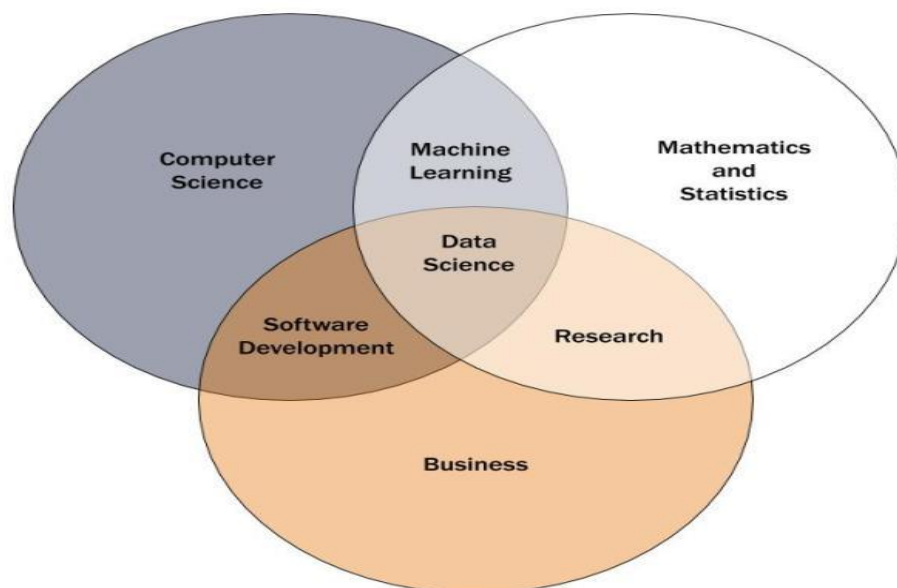


# Introduction to statistics for data science

Statistical analysis and probability influence our lives on a daily basis. Statistics is used to predict the weather, restock retail shelves, estimate the condition of the economy, and much more. Used in a variety of professional fields, statistics has the power to derive valuable insights and solve complex problems in business, science, and society. Without hard science, decision making relies on emotions and gut reactions. Statistics and data override intuition, inform decisions, and minimize risk and uncertainty.

In data science, statistics is at the core of sophisticated machine learning algorithms, capturing and translating data patterns into actionable evidence. Data scientists use statistics to gather, review, analyze, and draw conclusions from data, as well as apply quantified mathematical models to appropriate variables. Data scientists work as programmers, researchers, business executives, and more. However, what all of these areas have in common is a basis of statistics. Thus, statistics in data science is as necessary as understanding programming languages.

Towards Data Science, a website which shares concepts, ideas, and codes, supports that **data science knowledge is grouped into three main areas: computer science; statistics and mathematics; and business or field expertise**. These areas separately result in a variety of careers, as displayed in the diagram below. Combining computer science and statistics without business knowledge enables professionals to perform an array of machine learning functions. Computer science and business expertise leads to software development skills. Mathematics and statistics (combined with business expertise) result in some of the most talented researchers. It is only with all three areas combined that data scientists can maximize their performance, interpret data, recommend innovative solutions, and create a mechanism to achieve improvements.



Source: [Towards Data Science](#)

Statistical functions are used in data science to analyze raw data, build data models, and infer results. Below is a list of the key statistical terms:

- Population: the source of data to be collected.
- Sample: a portion of the population.
- Variable: any data item that can be measured or counted.
- Quantitative analysis (statistical): collecting and interpreting data with patterns and data visualization.
- Qualitative analysis (non-statistical): producing generic information from other non-data forms of media.
- Descriptive statistics: characteristics of a population.
- Inferential statistics: predictions for a population.
- Central tendency (measures of the center): mean (average of all values), median (central value of a data set), and mode (the most recurrent value in a data set).
- Measures of the spread:
  - Range: the distance between each value in a data set.
  - Variance: the distance between a variable and its expected value.
  - Standard deviation: the dispersion of a data set from the mean.

Statistical techniques data scientists need to master

Data scientists go beyond basic data visualization and provide enterprises with information-driven, targeted data. Advanced mathematics in statistics tightens this process and cultivates concrete conclusions.

Statistical techniques for data scientists

There are a number of statistical techniques that data scientists need to master. When just starting out, it is important to grasp a comprehensive understanding of these principles, as any holes in knowledge will result in compromised data or false conclusions.

General statistics: The most basic concepts in statistics include bias, variance, mean, median, mode, and percentiles.

**Probability distributions:** Probability is defined as the chance that something will occur, characterized as a simple “yes” or “no” percentage. For instance, when weather reporting indicates a 30 percent chance of rain, it also means there is a 70 percent chance it will not rain. Determining the distribution calculates the probability that all those potential values in the study will occur. For example, calculating the probability that the 30 percent chance for rain will change over the next two days is an example of probability distribution.

**Dimension reduction:** Data scientists reduce the number of random variables under consideration through feature selection (choosing a subset of relevant features) and feature extraction (creating new features from functions of the original features). This simplifies data models and streamlines the process of entering data into algorithms.

**Over and under sampling:** Sampling techniques are implemented when data scientists have too much or too little of a sample size for a classification. Depending on the balance between two sample groups, data scientists will either limit the selection of a majority class or create copies of a minority class in order to maintain equal distribution.

**Bayesian statistics:** Frequency statistics uses existing data to determine the probability of a future event. Bayesian statistics, however, takes this concept a step further by accounting for factors we predict will be true in the future. For example, imagine trying to predict whether at least 100 customers will visit your coffee shop each Saturday over the next year. Frequency statistics will determine probability by analyzing data from past Saturday visits. But Bayesian statistics will determine probability by also factoring for a nearby art show that will start in the summer and take place every Saturday afternoon. This allows the Bayesian statistical model to provide a much more accurate figure.

Statistical skills needed to perform data science jobs

Data science requires a mixture of technical skills, such as R and Python programming languages, as well as “soft skills,” including communication and attention to detail. Here are several of the most important skills data scientists need to hone in order to strengthen statistical abilities.

**Data manipulation:** Using Excel, R, SAS, Stata, and other programs, data scientists have the ability to clean and organize large data sets.

**Critical thinking and attention to detail:** Using linear regression, data scientists extract and model relationships between dependent and independent variables. Data scientists choose methods with built-in assumptions which are considered during their application. Violating or inappropriately choosing assumptions will lead to flawed results.

**Curiosity:** The desire to solve complex puzzles drives data scientists to design data plots and explore assumptions. They also discover patterns and sequences by using advanced data visualizations.

**Organization:** Data scientists are inundated with information from various sources and ongoing project opportunities. With budget and time constraints, data scientists perform efficiently when they are well-versed in statistical functions. In addition, having routinized processes helps ensure data is not compromised.

**Innovation and problem solving:** Above and beyond pure computations and basic data analysis, data scientists use applied statistics to pair abstract findings to real-world problems. Data scientists also use predictive analytics to determine future courses of action. All of this requires careful consideration, using both logical and innovative approaches to analyze issues and solve problems.

**Communication:** All of the work a data scientist does must be translated into a captivating story that industry leaders and executives can appreciate. Data scientists fill the gap between technology and operations. They translate findings into text and data visualizations that executives and clients can easily understand: an essential skill for a data scientist.

**Statistics:** Data scientists should consider learning statistics, because statistics connects data to the questions businesses are asking across all disciplines. Questions including:

- How can we create efficiencies?
- How can we limit spending and increase revenue?
- How can we maximize communications with our target audience?

How to learn statistics for data science

Data scientist shortages have pushed enterprises to [get creative while trying to fill the data talent gap](#). Some companies retrain existing staff in-house or arrange for graduate study in data science. Regardless of the method, education is the central force driving these efforts. Three popular educational paths are massive open online courses (MOOCs), bootcamps, or master's programs. While the data science education options leave employers wondering which path is best, masters degree programs have traditionally been the most valued among the three.

The best education in data science depends upon matching a student's needs with the most appropriate training resources. The process of learning statistics in data science, for instance, will look different depending on a person's educational and professional background. It is reasonable for a data science professional who has already acquired a data science foundation to sharpen their probability techniques through a variety of learning options. However, a recent college graduate, however, will find the deepest comprehensive data science training through a data science master's program.

Here is a quick glance at the pros and cons of learning data science through MOOCs, bootcamps, and master's degrees.

## **5 Structured Thinking Techniques for Data Science Problem-Solving**

Regardless of your experience level in data science, when faced with a new problem or challenge, we all need to take a step back and look at the bigger picture so we can solve the problem efficiently without wasting any time.

With different levels of expertise, every data scientist will address problems slightly differently. After all, the more projects you build, the more knowledge you will have on solving future problems. But, whether you're a beginner or not and looking for a more efficient way to solve problems, then this article is for you.

### **10 Skills of A Successful Data Scientist**

This article will discuss different structured thinking techniques that you can use to solve problems in data science or real life. Structured thinking is a problem-solving set of steps and techniques that work by dividing a large problem into smaller ones to solve the big problem faster and more efficiently.

#### **What is structured thinking?**

Structured thinking is a framework for solving unstructured problems — which covers almost all data science problems. Using a structured approach to solve problems doesn't only help with solving the problem faster but also identifies the parts of the problem that may need some extra attention. Think of structured thinking like the map of a new city you're visiting.

Without a map, you probably will find it difficult to reach your destination. Even if you did eventually reach your destination, it would probably have taken you to double the time you might have needed if you did have a map.

Structured thinking is a framework and not a fixed mindset; it can be modified to match any problem you need to solve. This article will go through five structured-thinking techniques to use in your next data science project.

### **№1: Six-Step Problem Solving Model**

We will kick off our list of techniques with the simplest and easiest-to-use one: the 6-step problem-solving model. This technique uses an analytical approach to solve any given problem. As the name suggests, this technique uses 6 steps to solve a problem, which are:

1. Have a clear and concise problem definition.
2. Study the roots of the problem.
3. Brainstorm possible solutions to the problem.
4. Examine the possible solution and chose the best one.
5. Implement the solution effectively.
6. Evaluate the results.

This model follows the mindset of continuous development and improvement. So, on step 6, if your results didn't turn out the way you wanted, you can go back to stem 4 and choose another solution or to step 1 and try to define the problem differently.

My favorite part about this simple technique is how easy it is to alter it based on the specific problem you are attempting to solve. You can read more about this technique [here](#).

### **№2: The Drill-Down Technique**

Our next technique is the drill-down technique. This technique is more suitable for complex and larger problems that multiple people will be working on. The whole purpose of using this technique is to break down a problem to its roots to ease up finding solutions for it.

To use the drill-down technique, you first need to start by creating a table. The first column of the table will contain the outlined definition of the problem, followed by a second column containing the factors causing this problem. Finally, the third column will contain the cause of the second column's contents, and you will continue to drill down on each column until you reach the root of the problem.

Once you reach the root causes of the problem, you can then use these root causes to develop solutions for the bigger problem. You can read more about the drill-down technique [here](#).

### **№3: Eight Disciplines of Problem Solving**

Next up are the eight disciplines of problem-solving. This technique offers a practical plan to solve a problem using an eight-step process. You can think of this technique as an extended, more detailed version of the 6 steps problem-solving model discussed earlier in the article.

Each of the eight disciplines in this process is designed to move you a step closer to finding the perfect solution to your problem. So, after you have got the prerequisites of your problem, you can follow these eight disciplines D1~D8.

1.     **D1: Put together your team.** Having a team with the set of skills needed to solve the project can make moving forward much easier.
2.     **D2: Define the problem.** Describe the problem using quantifiable terms the who, what, where, when, why, and how.
3.     **D3: Develop a working plan.**
4.     **D4: Determine and identify root causes.** Identify the root causes of the problem using the [cause and effect diagrams](#) to map causes against their effects.

5. **D5: Choose and verify permanent corrections.** Based on the root causes, assess the work plan you developed earlier and edit it if needed.
6. **D6: Implement the corrected action plan.**
7. **D7: Assess your results.**
8. **D8: Congratulate your team.** After the end of a project, it's essential to take a step back and appreciate the work that has been done before jumping into a new project.

#### **№4: The Cynefin Framework**

The Cynefin framework technique, like the rest of the techniques, works by breaking down a problem into its root causes to reach an efficient solution. The Cynefin framework works by approaching the problem from one of 5 different perspectives.

The Cynefin framework can be considered a higher-level approach because it requires you to place your problem into one of the 5 contexts.

- 1. Obvious Contexts.** In this context, your options are clear, and the cause-and-effect relationships are apparent and easy to point out.
- 2. Complicated Contexts.** In this context, the problem might have several correct solutions. In this case, a clear relationship between cause and effect may exist, but it's not apparent to everyone.
- 3. Complex Contexts.** If your problem is impossible to find a direct answer for, then it falls into the complex context. Complex contexts are problems that have unpredictable answers. The best approach here is to follow a trial and error approach to solve it.
- 4. Chaotic Contexts.** In this context, there is no relationship between cause and effect. So, your main goal will be to establish a correlation between the causes and effects to solve the problem.



**5. Disorder.** The final context is disorder, which is the most difficult of the context to categorize. And the only way to do it is to eliminate the other contexts and gather further information.

You can read further in-depth knowledge about the Cynefin approach [here](#).

### **№5: The 5-Whys Technique**

Our final technique in this article is the 5-whys technique, or how I like to call it, the curious child approach. I think this is the most well-known and natural approach to solve a problem.

This technique just follows a simple approach of asking why 5 times. First, you start with the main problem and ask why it occurred and then keep asking why until you reach the root cause of the said problem. Now, you may need to ask more or less than 5 whys to reach your answer.

Regardless of that, you can use this method to finally reach the root cause of the problem—more on the 5 whys problem and how to use it efficiently in [this](#) article.

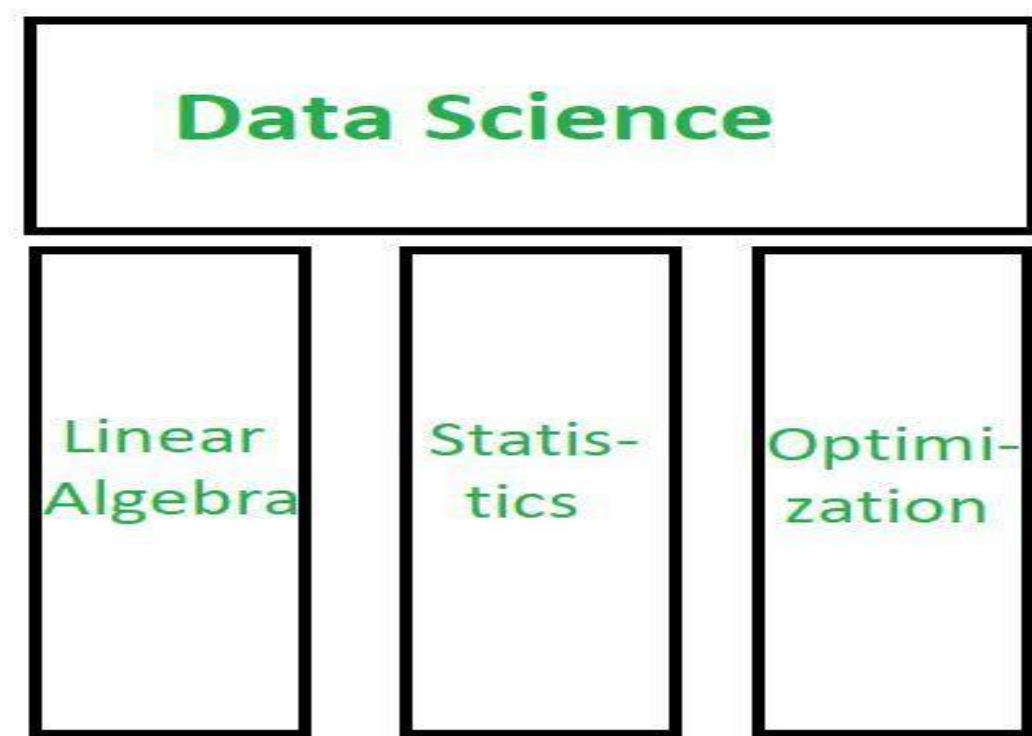
### **Takeaways**

Each one of us has our own way of approaching a new problem. If we are working as part of a team, which is probably the most probable case, you will need to divide the problem into smaller problems and assign those to different team members.

But, to be able to divide a problem, you need to understand it fully and know the expected end results of it. One of the best and most efficient ways to that is through structured thinking. Moreover, structured thinking is a flexible framework that anyone can use to make solving unstructured problems a much smoother process.

# Optimization for Data Science

From a mathematical foundation viewpoint, it can be said that the three pillars for data science that we need to understand quite well are **Linear Algebra**, **Statistics** and the third pillar is **Optimization** which is used pretty much in all data science algorithms. And to understand the optimization concepts one needs a good fundamental understanding of linear algebra.



## What's Optimization?

Wikipedia defines optimization as a problem where you maximize or minimize a real function by systematically choosing input values from an allowed set and computing the value of the function. That means when we talk about optimization we are always interested in finding the best solution. So, let say that one has some functional form(e.g in the form of  $f(x)$ ) that he is interested in and he is trying to find the best solution for this functional form. Now, what does best mean? One could either say he is interested in minimizing this functional form or maximizing this functional form.

## Why Optimization for Machine Learning?

- Almost all machine learning algorithms can be viewed as solutions to optimization problems and it is interesting that even in cases, where the original machine learning technique has a basis derived from other fields for example, from

biology and so on one could still interpret all of these machine learning algorithms as some solution to an optimization problem.

- A basic understanding of optimization will help in:
- More deeply understand the working of machine learning algorithms.
- Rationalize the working of the algorithm. That means if you get a result and you want to interpret it, and if you had a very deep understanding of optimization you will be able to see why you got the result.
- And at an even higher level of understanding, you might be able to develop new algorithms yourselves.

## Components of an Optimization Problem

Generally, an optimization problem has three components.

$$\text{minimize } f(x), \text{ w.r.t } x, \text{ subject to } a \leq x \leq b$$

1. **The objective function( $f(x)$ ):** The first component is an objective function  $f(x)$  which we are trying to either maximize or minimize. In general, we talk about minimization problems this is simply because if you have a maximization problem with  $f(x)$  we can convert it to a minimization problem with  $-f(x)$ . So, without loss of generality, we can look at minimization problems.
2. **Decision variables( $x$ ):** The second component is the decision variables which we can choose to minimize the function. So, we write this as **min  $f(x)$** .
3. **Constraints( $a \leq x \leq b$ ):** The third component is the constraint which basically constrains this  $x$  to some set.

So, whenever you look at an optimization problem you should look for these three components in an optimization problem.

## Types of Optimization Problems:

### Depending on the types of constraints only:

1. **Constrained optimization problems:** In cases where the constraint is given there and we have to have the solution satisfy these constraints we call them constrained optimization problems.
2. **Unconstrained optimization problems:** In cases where the constraint is missing we call them unconstrained optimization problems.

### Depending on the types of objective functions, decision variables and constraints:

1. **If the decision variable( $x$ ) is a continuous variable:** A variable  $x$  is said to be continuous if it takes an infinite number of values. In this case,  $x$  can take an infinite number of values between -2 to 2.

$$\text{min } f(x), \quad x \in (-2, 2)$$

2.
  - **Linear programming problem:** If the decision variable( $x$ ) is a continuous variable and if the objective function( $f$ ) is linear and all the constraints are also linear then this type of problem known as a linear programming problem.

So, in this case, the decision variables are continuous, the objective function is linear and the constraints are also linear.

- **Nonlinear programming problem:** If the decision variable(x) remains continuous; however, if either the objective function(f) or the constraints are non-linear then this type of problem known as a non-linear programming problem. So, a programming problem becomes non-linear if either the objective or the constraints become non-linear.

3. **If the decision variable(x) is an integer variable:** All numbers whose fractional part is 0 (zero) like -3, -2, 1, 0, 10, 100 are integers.

$$\min f(x), x \in [0, 1, 2, 3]$$

4.

- **Linear integer programming problem:** If the decision variable(x) is an integer variable and if the objective function(f) is linear and all the constraints are also linear then this type of problem known as a linear integer programming problem. So, in this case, the decision variables are integers, the objective function is linear and the constraints are also linear.

- **Nonlinear integer programming problem:** If the decision variable(x) remains integer; however, if either the objective function(f) or the constraints are non-linear then this type of problem known as a non-linear integer programming problem. So, a programming problem becomes non-linear if either the objective or the constraints become non-linear.

- **Binary integer programming problem:** If the decision variable(x) can take only binary values like 0 and 1 only then this type of problem known as a binary integer programming problem.

$$\min f(x), x \in [0, 1]$$

•

5. **If the decision variable(x) is a mixed variable:** If we combine both continuous variable and integer variable then this decision variable known as a mixed variable.

$$\min f(x_1, x_2), x_1 \in [0, 1, 2, 3] \text{ and } x_2 \in (-2, 2)$$

6.

- **Mixed-integer linear programming problem:** If the decision variable(x) is a mixed variable and if the objective function(f) is linear and all the constraints are also linear then this type of problem known as a mixed-integer linear programming problem. So, in this case, the decision variables are mixed, the objective function is linear and the constraints are also linear.

- **Mixed-integer non-linear programming problem:** If the decision variable(x) remains mixed; however, if either the objective function(f) or the constraints are non-linear then this type of problem known as a mixed-integer non-linear programming problem. So, a programming problem becomes non-linear if either the objective or the constraints become non-linear.

## What are Eigenvalues?

The eigenvalue is explained to be a scalar associated with a linear set of equations which, when multiplied by a nonzero vector, equals to the vector obtained by transformation operating on the vector.

Let us consider  $k \times k$  square matrix  $A$  and  $v$  be a vector, then  $\lambda$  is a scalar quantity represented in the following way:

$$AV = \lambda V$$

Here,  $\lambda$  is considered to be the eigenvalue of matrix  $A$ .

The above equation can also be written as:

$$(A - \lambda I) = 0$$

Where “ $I$ ” is the identity matrix of the same order as  $A$ .

This equation can be represented in the determinant of matrix form.

$$|A - \lambda I| = 0$$

The above relation enables us to calculate eigenvalues  $\lambda$  easily.

## Steps to Find Eigenvalues of a Matrix

In order to find the eigenvalues of a **matrix**, follow the steps below:

**Step 1:** Make sure the given matrix  $A$  is a square matrix. Also, determine the identity matrix  $I$  of the same order.

**Step 2:** Estimate the matrix  $A - \lambda I$ , where  $\lambda$  is a scalar quantity.

**Step 3:** Find the determinant of matrix  $A - \lambda I$  and equate it to zero.

**Step 4:** From the equation thus obtained, calculate all the possible values of  $\lambda$ , which are the required eigenvalues of matrix  $A$ .

## Eigenvector Definition

Eigenvector of a square matrix is defined as a non-vector in which when a given matrix is multiplied, it is equal to a scalar multiple of that vector. Let us suppose that  $A$  is an  $n \times n$  square matrix, and if  $v$  be a non-zero vector, then the product of matrix  $A$ , and vector  $v$  is defined as the product of a scalar quantity  $\lambda$  and the given vector, such that:

$$Av = \lambda v$$

Where

$v$  = Eigenvector and  $\lambda$  be the scalar quantity that is termed as eigenvalue associated with given matrix  $A$

-