

# **Topic 2 - Data Analysis and Data for Graphics**

# Introduction - Definition



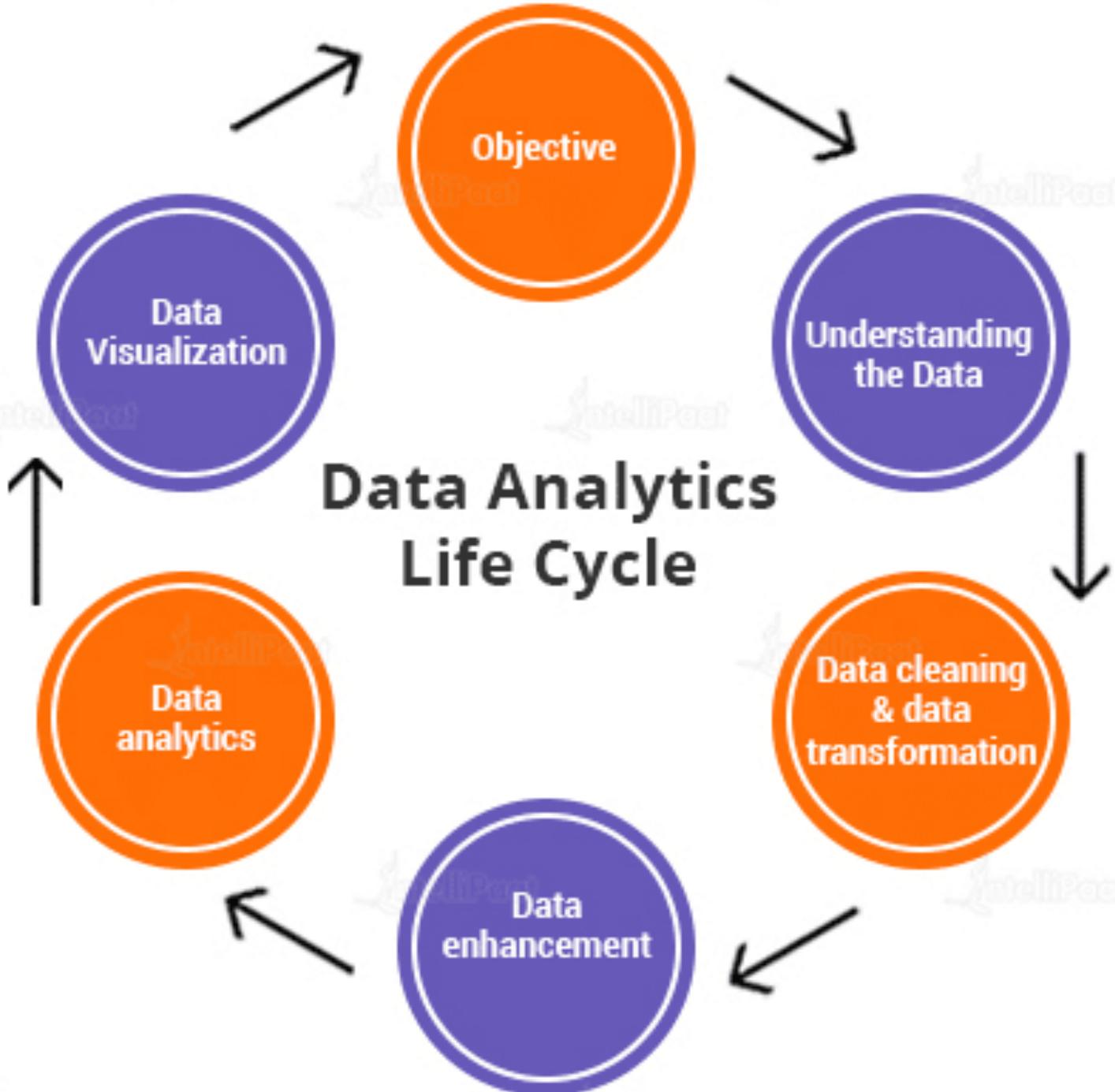
**Data analysis** is defined as a process of cleaning, transforming, and modeling data to **discover useful information for decision-making**.

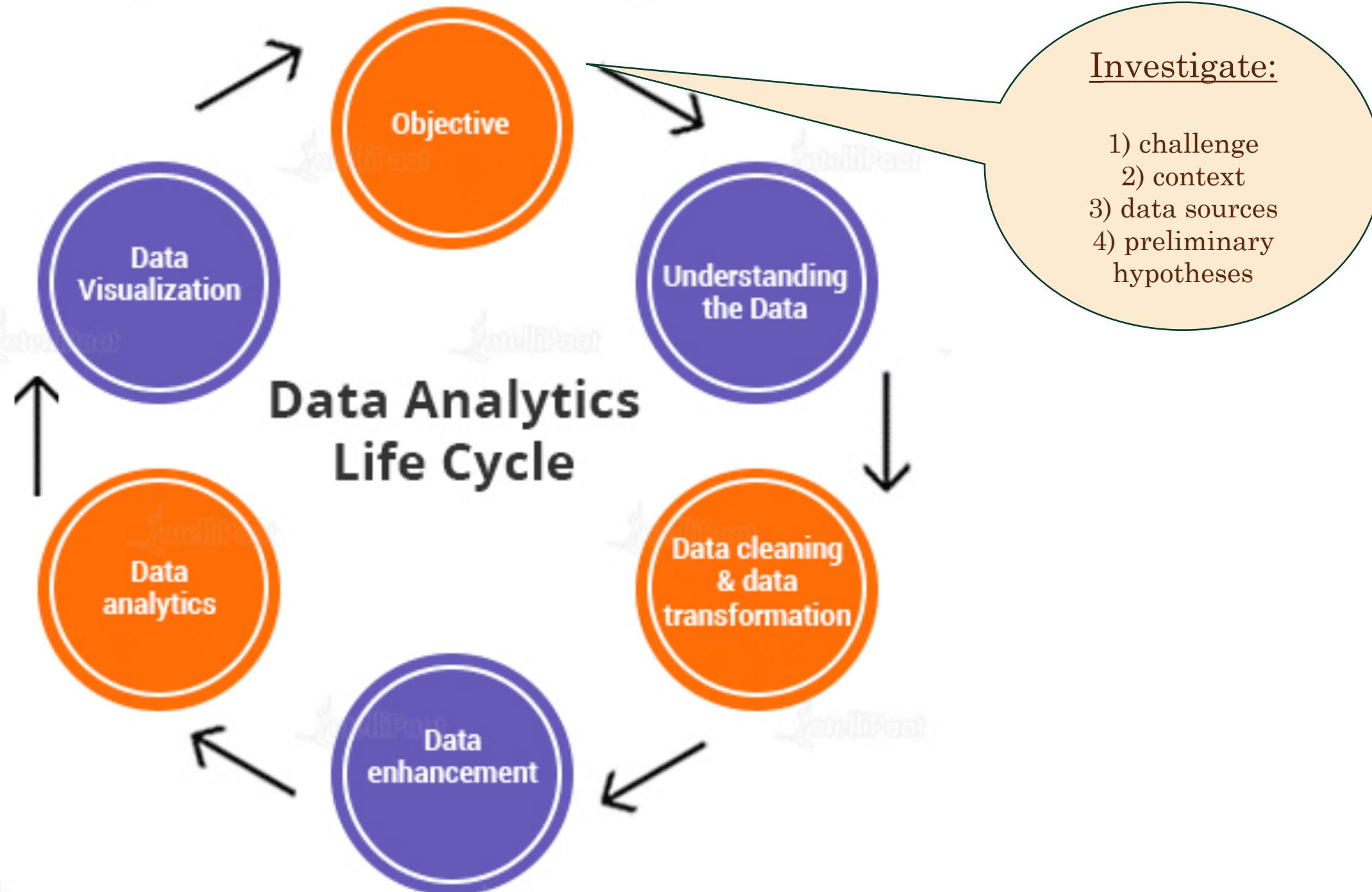
# Introduction

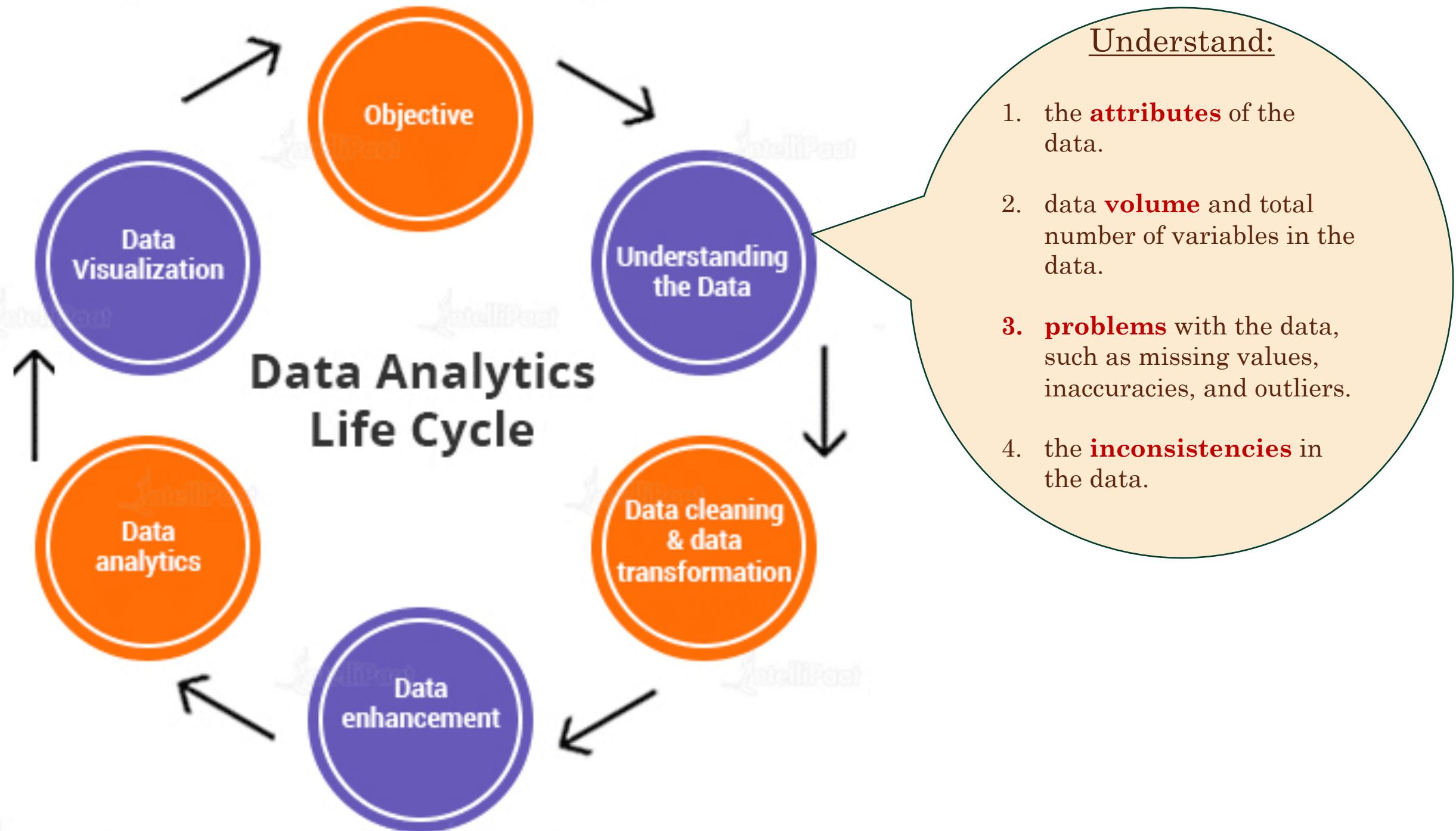
The purpose of data analysis are:

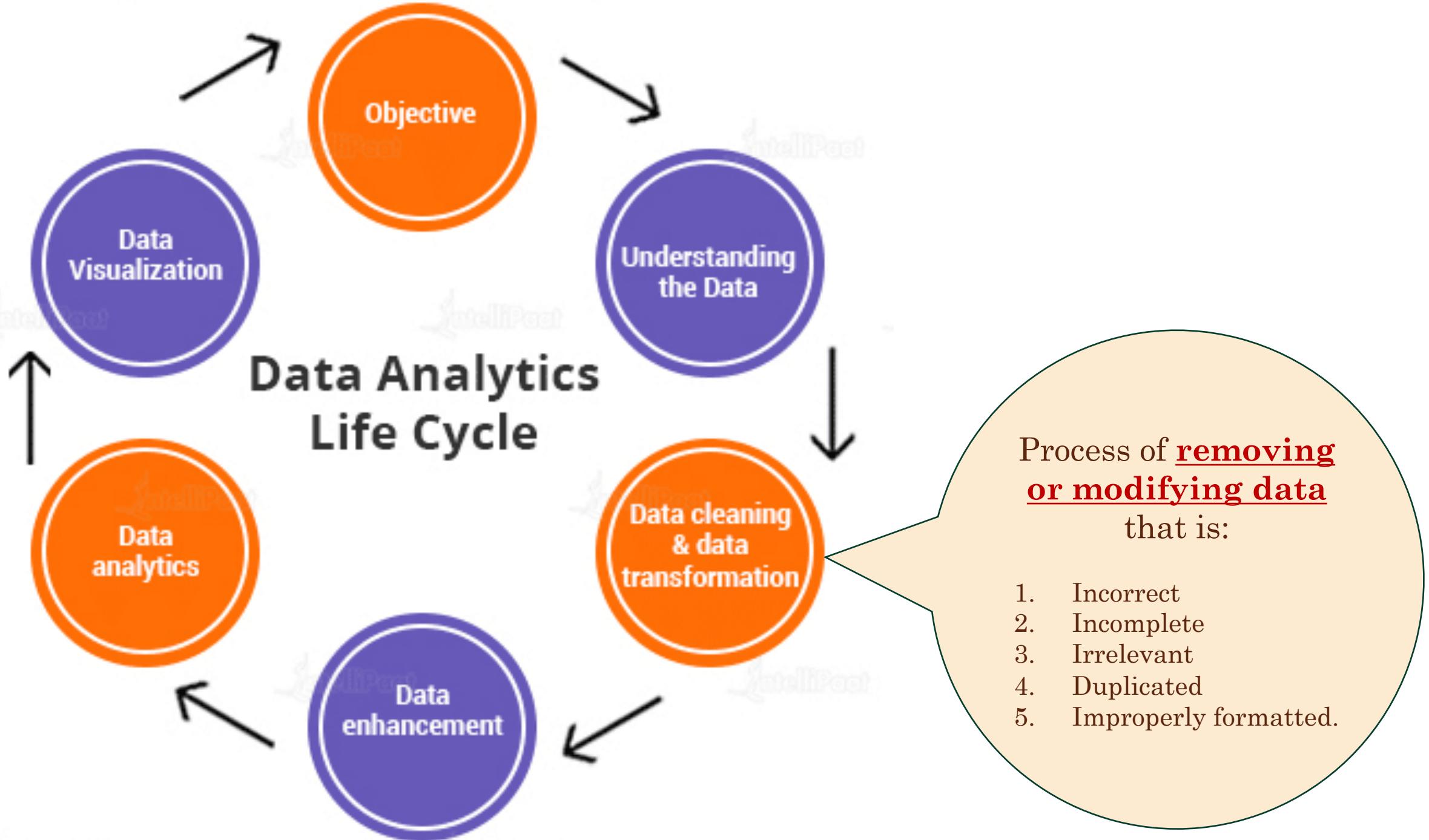
- To **extract useful information** from data and taking the decision based upon the data analysis.
- To **describe data** in meaningful terms.
- To **detect patterns**, relationships or trends.



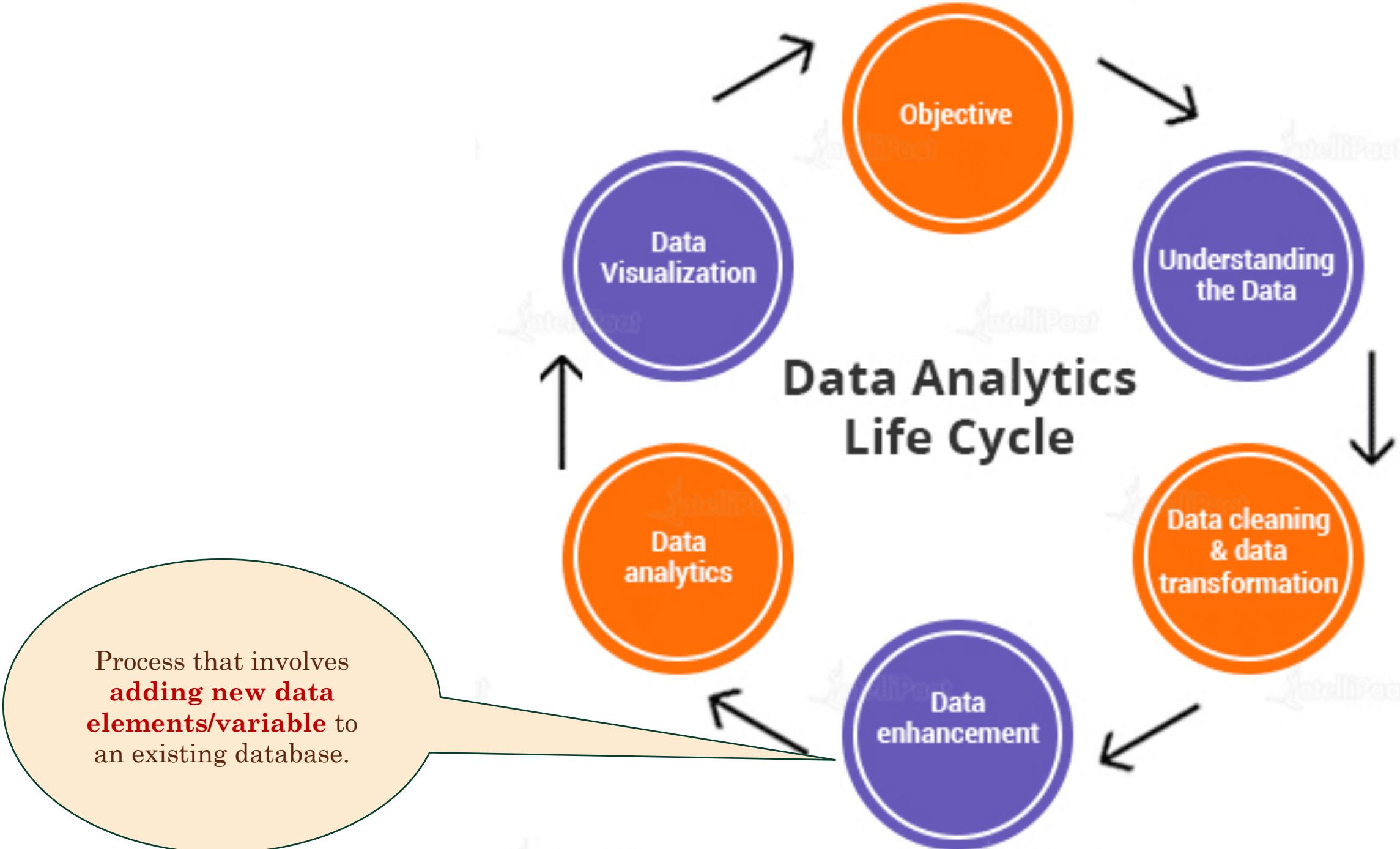








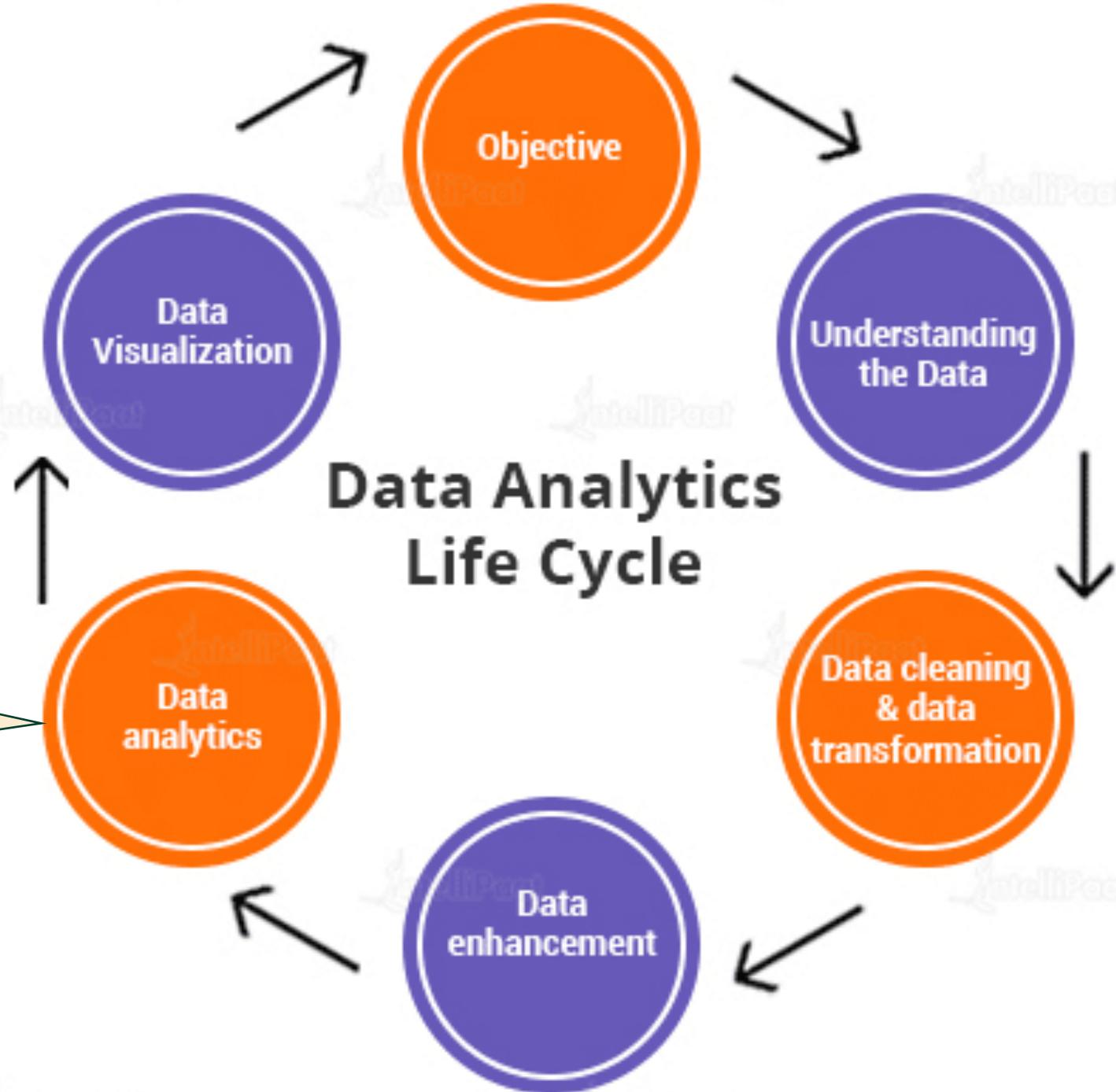
## Data Analytics Life Cycle



## Data Analytics Life Cycle

Use data analysis  
tool/software to:

1. **extract** raw data
  2. **organize** it
  3. **analyze** it
  4. **transform** it
- from incomprehensible numbers into coherent, intelligible information





## Data Analytics Life Cycle

Data visualization is **the graphical representation** of information and data.

By using visual elements like charts, graphs, and maps, data visualization tools provide an accessible way to see and understand trends, outliers, and patterns in data.

# Why is Data Analysis Important?

## 1) Better User Targeting

You don't want to waste your business's precious time, resources, and money putting together advertising campaigns targeted at demographic groups that have little to no interest in the goods and services you offer.

Data analysis helps you see where you should be focusing your advertising efforts.

## 2) Know Your Target User Better

Data analysis tracks how well your products and campaigns are performing within your target demographic.

Through data analysis, your business can get a better idea of your target audience's spending habits, disposable income, and most likely areas of interest

# Why is Data Analysis Important?

## 3) Reduce Operational Costs

Data analysis shows you which areas in your business need more resources and money, and which areas are not producing and thus should be scaled back or eliminated outright.

## 4) Better Problem-Solving Methods

Informed decisions are more likely to be successful decisions. Data provides businesses with information. You can see where this progression is leading.

Data analysis helps businesses make the right choices and avoid costly pitfalls.

# Why is Data Analysis Important?

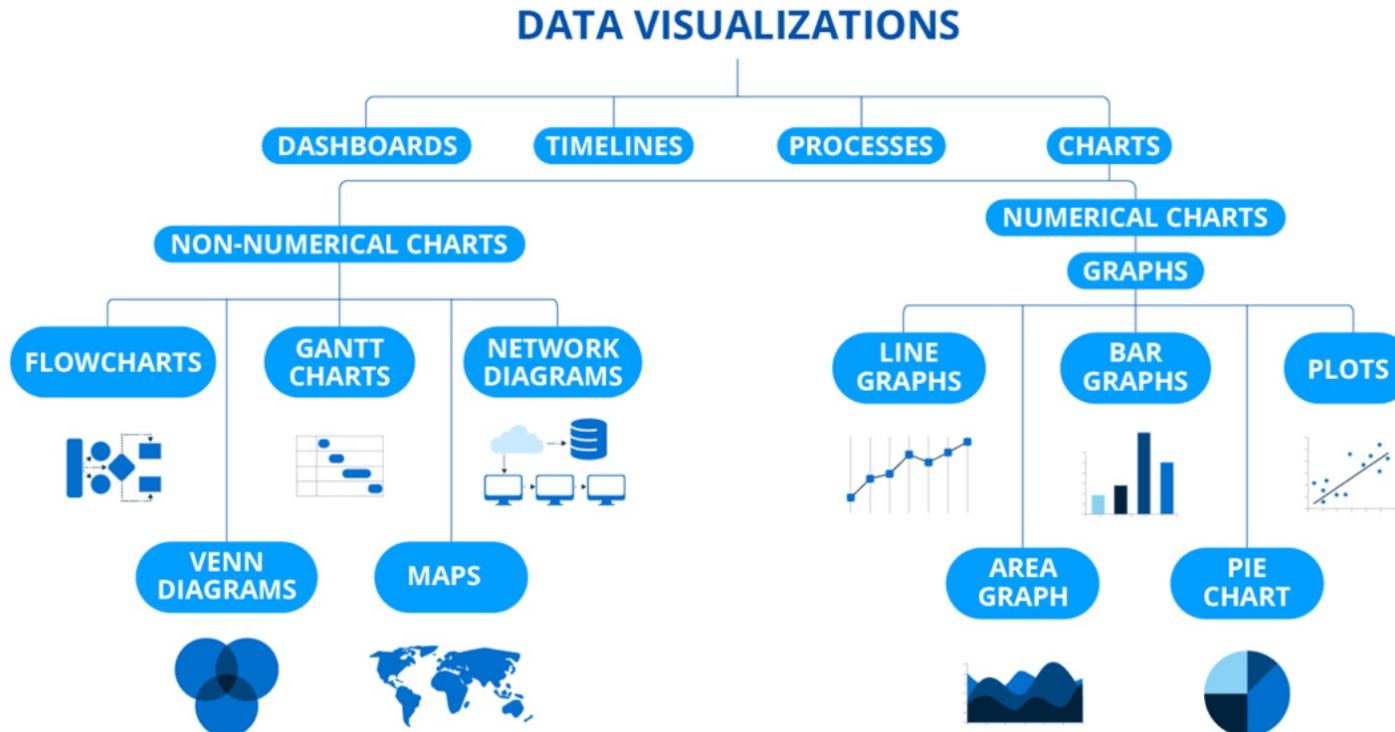
## 5) Get More Accurate Data

If you want to make informed decisions, you need data, but there's more to it. The data in question must be accurate.

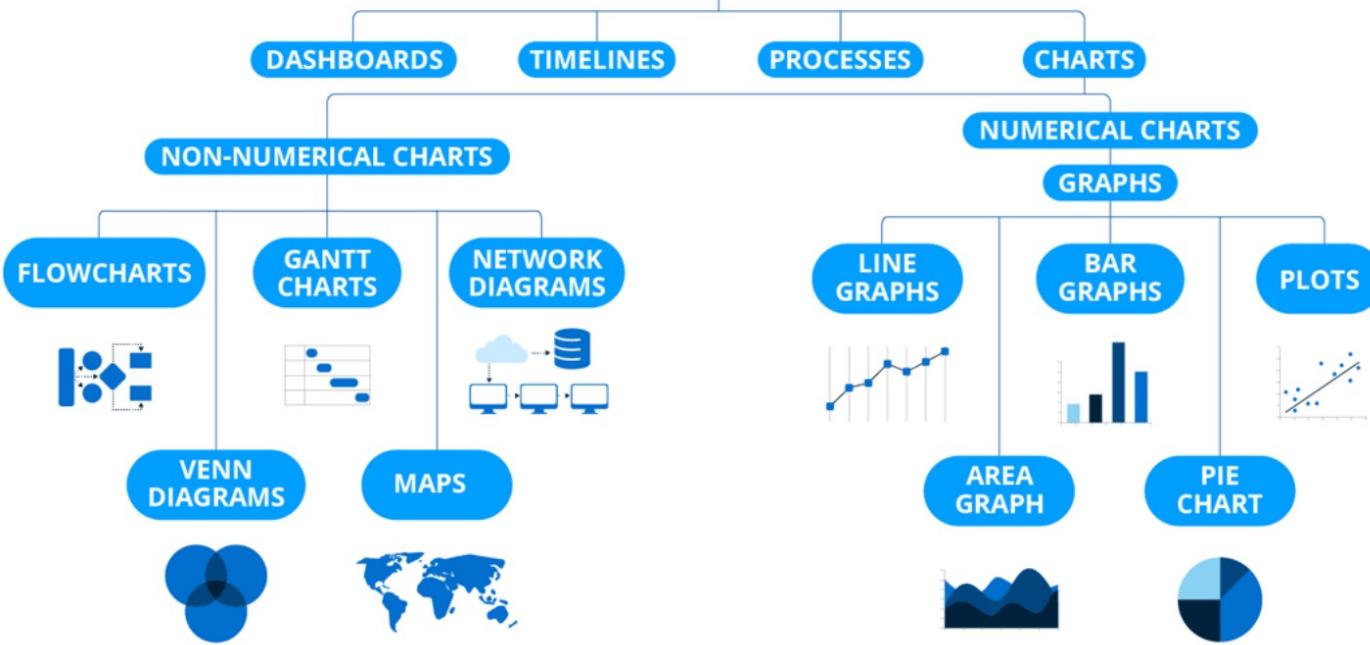
Data analysis helps businesses acquire relevant, accurate information, suitable for developing future marketing strategies, business plans, and realigning the company's vision or mission.

# Understanding the Graphical Representation of Data

An appropriate presentation of data involves organization of data in such a manner that meaningful conclusions and inferences can drawn to answer the research question.



# DATA VISUALIZATIONS



A **chart** is a visual representation of information or data.

The purpose of a chart is to help viewers understand and analyze information easily with the help of visuals.

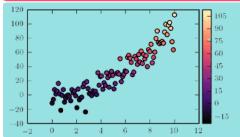
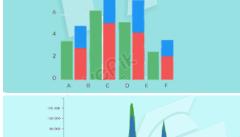
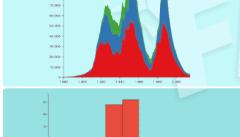
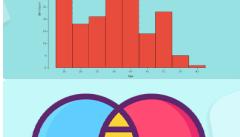
Charts can be stand-alone visuals or grouped to create infographics, dashboards, and other more complex data visualizations.

A **graph** is a chart that uses mathematical equations to visualize data and analyze relationships and trends.

For a chart to be a graph—and not another type of chart—it must involve a mathematical analysis, generally using the x and y axis to plot data points.

Common graphs use lines and bars to visualize data quantities, relationships, and trends.

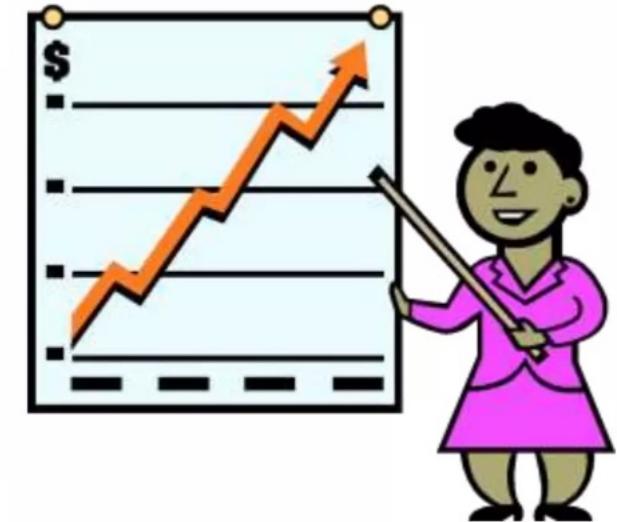
# How to Use Charts & Graphs

What if it Looks Like	What it is	When to Use it	What to Remember
	Scatter Plot	Reporting Relationships/Correlations Between Different Variables	When Using More than 10 DataPoints, Place in a Horizontal Axis.
	Column Chart	Showin Change Over Time Between Two Variables with the Same Unit of Measurement and Comparable Sizes and is Oriented Vertically.	Use When You Want to Show Minimum and Maximum Values.
	Bar Chart	Comparing Data Across Different Categories and is Oriented Horizontally.	Place Your Values on the x-axis, Categories on the y-axis, Use When the Number of Categories is Between 5 and 8.
	Line Chart	Showing Data Trends Over Time but Have too Big Dataset for a Column or Bar.	Great When the Number of Categories is 8+ and Works Well in Combination With Bar Charts.
	Combination Chart	Comparing Two Variables With Different units of Measurment or Sizes.	Great to Use When you Want to Measure Values that are Significantly Different in Range.
	Pie Chart	Showing Relative Proportions (or Even Percentages).	The Data Set Must Represent Part of a Whole, and Data Should be on Comparable Sizes so Nothing is Dwarfed.
	Stacked Column	Showing the Composition of 5-8 Categories.	This Follows the Same Rules of a Pie Graph, But With More Categories.
	Stacked Area	Emphasizing the Magnitude of Change Over Time With Fewer than 8 Data Points.	Use this to Show the Trend of Compositions.
	Histogram	Showing the Distribution of Data.	Use this When you Want to Show the Distribution of Data Across Groups.
	Venn Diagram	Showing Overlapping Data Sets.	Consider this for Attribution Modeling No Overlap = No Attribution Problems.

# How to Present Charts & Graphs

The most efficient way to present charts and graphs is to follow this formula:

- 1) **Introduce** the graph.
- 2) Identify the **variables**.
- 3) **Highlight** key info.
- 4) Share **conclusions**.



# The need of presenting data graphically

Graphics are used to represent large volume of data. They are necessary:

1

Graphical form makes it possible to **easily draw visual impressions of data.**

4

Create an **imprint on mind** for a longer time.

2

The graphic method of the representation of data **enhances our understanding.**

5

It represents characteristics in a **simplified way.**

3

It makes the **comparison easy.**

6

These makes it easy to understand the **patterns and trends**

# Graphical Representation of Data

**Tables:** This consists of rows and columns used to compare variables. Tables can show a great deal of information in a structured way, but they can also overwhelm users that are simply looking for high-level trends.

**Pie charts and stacked bar charts:** These graphs are divided into sections that represent parts of a whole. They provide a simple way to organize data and compare the size of each component to one other.

**Line charts and area charts:** These visuals show change in one or more quantities by plotting a series of data points over time and are frequently used within predictive analytics. Line graphs utilize lines to demonstrate these changes while area charts connect data points with line segments, stacking variables on top of one another and using color to distinguish between variables.

**Histograms:** This graph plots a distribution of numbers using a bar chart (with no spaces between the bars), representing the quantity of data that falls within a particular range. This visual makes it easy for an end user to identify outliers within a given dataset.

**Scatter plots:** These visuals are beneficial in revealing the relationship between two variables, and they are commonly used within regression data analysis. However, these can sometimes be confused with bubble charts, which are used to visualize three variables via the x-axis, the y-axis, and the size of the bubble.

**Heat maps:** These graphical representation displays are helpful in visualizing behavioral data by location. This can be a location on a map, or even a webpage.

**Tree maps:** which display hierarchical data as a set of nested shapes, typically rectangles. Treemaps are great for comparing the proportions between categories via their area size.

# Graphical Representation of Data

Resource:

- 1) <https://www.intellspot.com/types-graphs-charts/>
- 2) <https://piktochart.com/blog/types-of-graphs/>
- 3) <https://visme.co/blog/types-of-graphs/>