

The importance of data type identification

So far, you've learned that visualization planning and design are essential parts of the storytelling journey. In this reading, you'll learn more about identifying data types, and their importance in accurately communicating the story of your data.

Data types

A data type is an attribute that describes a piece of data based on its values, its programming language, or the operations it can perform.

For example, the number 1992 doesn't have enough context on its own. It could mean the year 1992, it could indicate a quantity of 1,992 things, it could indicate a dollar amount of \$1,992, or it could represent military time: 19:92. You can change the meaning and use of a number simply by changing the data type.

Understanding and accurately using data types is important because they:

- Affect how computers and systems use and process data
- Affect what types of operations can be performed on data
- Inform how analysts and programmers interpret data
- Impact which visuals you can use
- Help analysts process and present data correctly

Data type examples

You'll likely work with many different types of data throughout your cloud data analyst career. Here's a list of some of the data types you'll likely encounter:

Type	Expected data	Display examples
Number	A floating point number	<ul style="list-style-type: none">• 1,234.56
Percent	A floating point number representing a ratio	<ul style="list-style-type: none">• 61.73%
Duration	An integer representing a time duration in seconds	<ul style="list-style-type: none">• 01:43• 01:43:19
Currency	A floating point number representing a monetary value	<ul style="list-style-type: none">• \$1,234.56• 173.9 €• ¥30,093
Text	Zero or more letters, numbers, characters, or symbols	<ul style="list-style-type: none">• A1B2C3

Date & Time	An actual calendar date, a calendar date with time, or an abstract date or time index	<ul style="list-style-type: none"> • 2020-09-21 • 2020-09-21 12:35 PM • Q4 2020 • Week 1
Boolean	A logical value	<ul style="list-style-type: none"> • true • false
Geo	A value representing a geographic region, such as a country name, latitude and longitude, ISO code, or Google Ads criteria ID	<ul style="list-style-type: none"> • United Kingdom • 51.5074,-0.1278
URL	An address of a World Wide Web page	https://google.com/
Hyperlink	A clickable link to a URL	Hyperlink Example
Image	An image	
Image Link	A clickable image	

Note: The list of data types provided in this reading isn't comprehensive. Different tools and programming languages have different variations, but this chart provides a good foundation for understanding the basics.

Numerical vs. categorical data

Another important aspect of understanding data types is determining whether it's numerical or categorical. As a reminder, understanding these two data types will help you identify which types of visualizations are best suited for the type of data you're working with. These differences are also crucial for reporting purposes.

Numerical data

Numerical data is a specific and objective measure, like a number, quantity, or range. Numerical data is sometimes referred to as quantitative data because it describes things you can count, order, or measure. For example, numerical data could be the number of sales made in a day, transaction numbers, and the percent increase in sales from a previous year.

Data analysts use numerical data to answer questions like:

- How much?
- How many?
- How often?

In reports and visualizations, numerical data is used for measuring and ordering variables, like website traffic, operating expenses, or sales revenue.

Categorical data

Categorical data is a subjective and explanatory measure of a quality or characteristic. This means it describes the category or quality of the data being collected, which is why it's also sometimes referred to as qualitative data.

An easy way to determine if data is categorical, is if it's not numerical, but instead a descriptive name or label that can fit into a category, like countries, industries, colors, or products. Some examples of categorical business data are customer satisfaction ratings, like satisfied, neutral, dissatisfied, and vehicle type, like sedan, SUV, truck, or minivan.

Categorical data is subjective, often has no inherent order, and can't be objectively measured or calculated. Data analysts use categorical data to answer questions about data like:

- Where does it belong?
- What is it?

Categorical data is used in reports and visualizations for grouping data, like the sales region a store is located in, or the various products within a product line.

Key takeaways

Understanding the differences between data types is essential for creating reports and dashboards that are accurate, add value, and answer the questions your stakeholders have about their business. Accurately identifying numerical and categorical data types also allows you to effectively communicate what a value represents, how it can be interpreted, and how it can be processed. As a cloud data analyst, familiarizing yourself with these data types will allow you to present data more effectively, and help your stakeholders gain a clearer understanding of their business.