Tableau Lecture 1

Topics :
- Part A : Introduction to Data Visualization and Tableau

Part A : Introduction to Data Visualization and Tableau

Learning Objective:

- Discover what data visualization is, and how we can use it to better see and understand data.
- We acquaint ourselves with the history of data visualization and get introduced to Tableau.
- Then, we get an idea about the various Tableau product offerings.
- We also learn about the pros and cons of Tableau.
- Finally we see how to download and install Tableau Public.

Part A Topics :

1. A quick demonstration of Tableau's capability : *Netflix Viz & Dashboard*
2. The Basics
3. Brief History
4. Over the years
5. Why choose Tableau? And why not choose Tableau?
6. Tableau Products Suite
7. Installation
8. Quizzes

The Basics :

❖   What is Business Intelligence(BI)

BI combines business analytics, data mining, data visualization, data tools and infrastructure, and best practices to help organizations to make more data-driven decisions

❖   Why BI

It help companies make better decisions by showing present and historical data within their business context.

❖   What is Data Visualization

It 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.

Humans are good at deriving knowledge from visualizations.

Brief History

❖ Until the early 21st century, Database, Excel, Access etc were used to produce numbers and data.

❖ The main idea behind Tableau's creation was to make the database industry interactive and comprehensive.

❖ Tableau is a popular data visualization and business intelligence tool used for reporting and analyzing vast volumes of data.

❖ Tableau was founded by Pat Hanrahan, Christian Chabot, and Chris Stolte from Stanford University in 2003.

Over The Years..

- Tableau has been named a Leader in the Gartner Magic Quadrant for Analytics & Business Intelligence Platforms for the 10th consecutive year.

- Tableau Software has a market capitalization of $14.61 billion and generates $982.95 million in revenue each year.

Gartner Magic Quadrant for Analytics & Business Intelligence Platforms

- CHALLENGERS
- LEADERS
  - Microsoft
  - Salesforce (Tableau)
  - Qlik
  - ThoughtSpot
  - Sisense
- Google
- Domo
- MicroStrategy
- Amazon Web Services
- Alibaba Cloud
- Pyramid Analytics
- Incorta
- Zoho
- TIBCO Software
- SAP
- Oracle
- IBM
- SAS
- Yellowfin
- Tellius
- NICHE PLAYERS
- VISIONARIES

ABILITY TO EXECUTE

COMPLETENESS OF VISION

As of January 2022

© Gartner, Inc

Gartner

Why choose Tableau?

Pros :

- Quick and interactive visualization

- Easy to use for non programmers

- High Performance

- Mobile friendly

- Extensive customer resources(Tableau

  Community)

- Working with different Data sources

- Easy to upgrade

Tableau Limitations

However, Tableau still has several limitations:

- Tableau focuses primarily on visualization and cannot work with uncleaned data. In order to efficiently use Tableau, you need to do proper data cleaning in the underlying database first.

- Lacks data modeling and data dictionary capabilities for Data Analysts. This means that you've to separately maintain your metrics definitions elsewhere.

- Their support is very poor and some users said that they have to solve the issue by themselves.

- Lack of version control and collaboration when building data logic and dashboard.

Tableau Products Suite

- Tableau Prep
- Tableau Desktop
- Tableau Server
- Tableau Online
- Tableau Reader
- Tableau Public

References :

- Comparison of product suite
- Tableau Pricing
- Understanding License types of Tableau
- https://www.edureka.co/blog/tableau-desktop-vs-tableau-public-vs-tableau-reader/#:~:text=Tableau%20Desktop%20is%20meant%20for,%2C%20bloggers%2C%20students%20and%20more.
-

Installation

Tableau Public download link : https://public.tableau.com/en-us/s/

Tableau Desktop Installation steps:
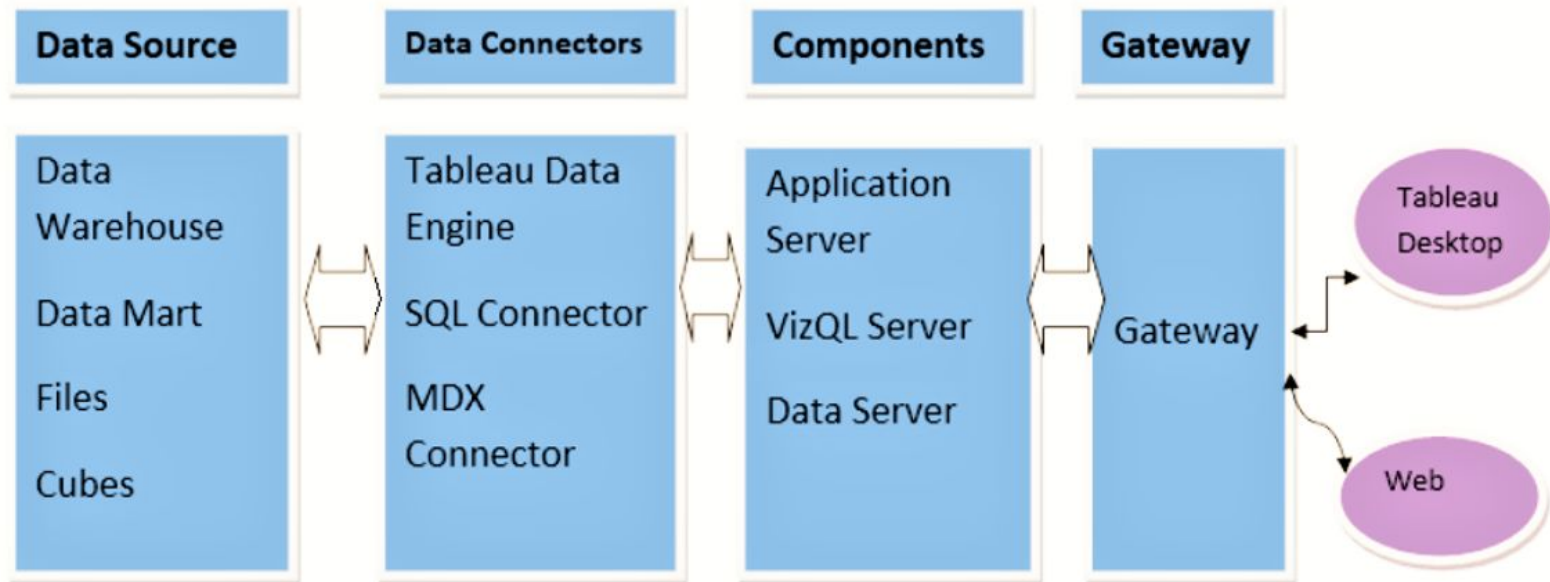https://help.tableau.com/current/desktopdeploy/en-us/desktop_deploy_download_and_install.htm#install-the-product

Part B : Tableau Architecture and Data Connection with Tableau Public

Learning Objective: In this module, you will get a brief idea about the Architecture, Tableau UI components, fields and various ways to establish data connection.

Part B Topics :

1. Tableau Architecture
2. Business Problem 1 covers :
   a. Data set loading
   b. Data interpreter
   c. Pivot
   d. Data source filtering
   e. Tableau Public Sign In
   f. Workbook Publish on Tableau Public Online
   g. Dashboard Publish on Tableau Public Online
3. Graphical User Interface and features of Tableau Public
4. Data Fields
   a. Data Types
      i. Change Data Type in Data Source page, Data Pane and View
   b. Dimension vs Measure
   c. Discrete vs Continuous
   d. Special Cases 1 & 2
5. Business Problem 2 (covers quick filter)
6. Alias
7. Problem 3 (Alias)
8. Simple Dashboard Creation and Upload it to Tableau Public Online
9. Tableau File Types
10. Tableau Data Connectors
11. OData
12. Business Problem 4
13. Web Data Connector
14. Business Problem 5

Tableau Architecture

Data Source	Data Connectors	Components	Gateway
Data Warehouse Data Mart Files Cubes	Tableau Data Engine SQL Connector MDX Connector	Application Server VizQL Server Data Server	Gateway

Tableau Desktop

Web

Reference:
- https://www.educba.com/tableau-architecture/
- https://www.tableau.com/learn/articles/business-intelligence
- https://www.tableau.com/drive/what-is-vizql
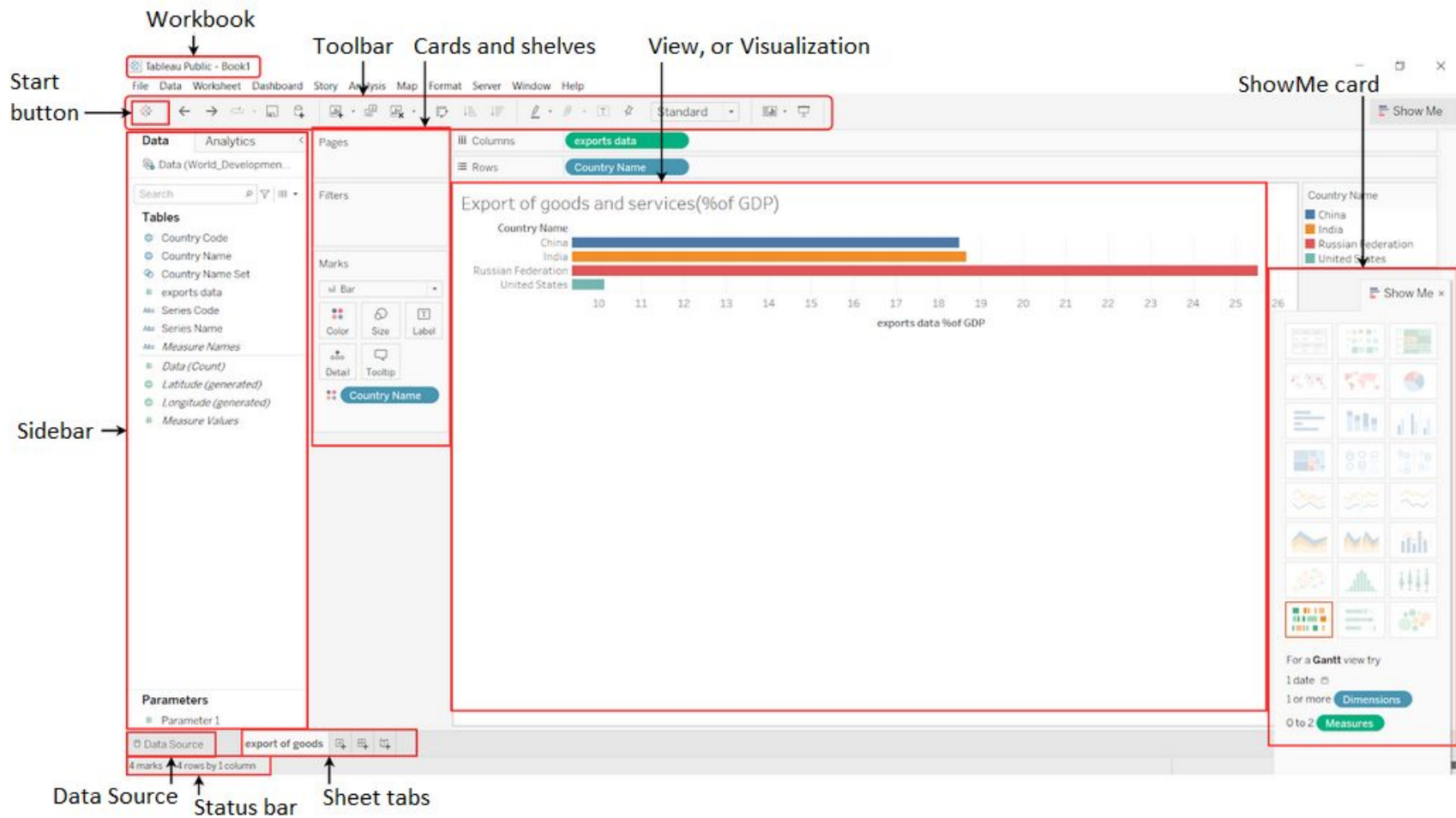- https://www.tableau.com/wp/dmr-20080215

Formatting of quick filter post implementation

- Filter card Options -After you show a filter, there are many different options that let you control how the filter works and appears. You can access these options by clicking the drop-down menu in the upper right corner of the filter card in the view.
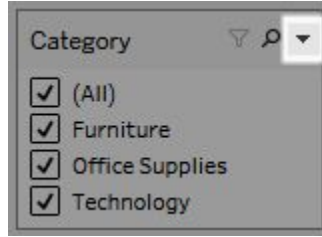
- Filter card modes -You can control the appearance and interaction of your filter card in the view by selecting a filter card mode. (Demo of it is added in the doc)

The Graphical User Interface (GUI)

Workbook

Toolbar   Cards and shelves

View, or Visualization

ShowMe card

Start button →

Sidebar →

Data Source

Status bar

Sheet tabs

Data Fields

- After you connect to your data and set up the data source with Tableau, the data source connections and fields appear on the left side of the workbook in the Data pane.
- The term "fields" refers to columns.
- When you connect to a new data source, Tableau assigns each field in the data source as dimension or measure in the Data pane, depending on the type of data the field contains.
- Each field is automatically assigned a :
  - Data type (such as integer, string, date),

  and each field is automatically assigned a role:
  - Discrete Dimension or Continuous Measure (more common), or
  - Continuous Dimension or Discrete Measure (less common)

Reference :
  - https://help.tableau.com/current/pro/desktop/en-us/datafields_typesandroles.htm
  - https://help.tableau.com/current/pro/desktop/en-us/datafields_understanddatawindow.htm

Data Types

- Tableau expresses fields and assigns data types automatically.
- If the data source appoints the data type, Tableau will use that data type.
- If the data source doesn't individually assign a data type, Tableau will assign one

Tableau Data Types

Data Type	Icon
String values (Text)	Abc
Integer values (Numbers)	#
Date values (DD/MM/YYYY or MM/DD/YYYY)	📅
Date & Time values	📅🕐
Boolean values (True or False; relational)	T|F
Geographic values (Region, Postal code etc.)	🌐
Cluster group or mixed values	

■ Changing Data Types in Tableau [*Hands-on*] :

1. Changing data type of a field in Data Source page

2. Changing data type of a field from Data pane.

3. Changing the data type of a field in the View
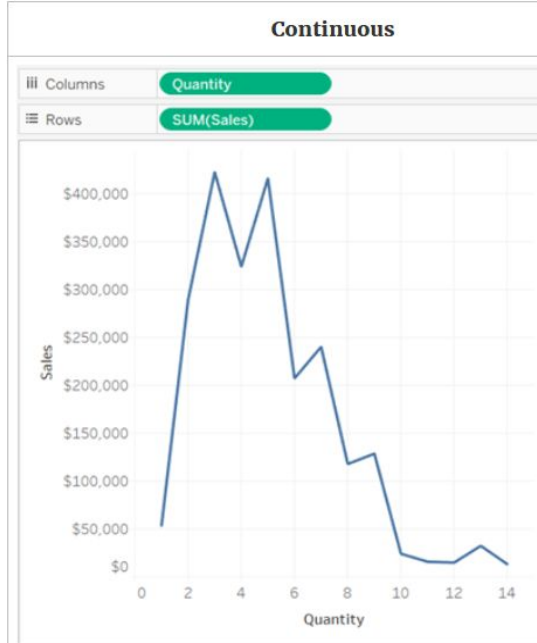
Dimension vs Measure

- A measure is a field that is a dependent variable; that is, its value is a function of one or more dimensions. Tableau treats any field containing numeric (quantitative) information as a measure.

- Dimension is a field that can be considered an independent variable. By default, Tableau treats any field containing qualitative, categorical information as a dimension.

- In case any field is misclassified by Tableau, it can easily be reclassified by right clicking the field from within the Dimensions or Measures area of the Data pane and choosing "Convert to dimension" or "Convert to measure" as appropriate.

Rule of Thumb : Generally, the measure is the number; the dimension is what you "slice and dice" the number by.

Discrete vs Continuous

Rule of thumb :

1. Discrete fields draw headers; continuous fields draw axes.
2. Discrete fields can be sorted; continuous fields cannot.
3. Blue colour Field indicated Discrete Field
4. Green colour field indicates Continuous Field

Continuous

Columns   Quantity

Rows   SUM(Sales)

Sales — $400,000 / $350,000 / $300,000 / $250,000 / $200,000 / $150,000 / $100,000 / $50,000 / $0

Quantity — 0 2 4 6 8 10 12 14

Discrete

Columns   Quantity

Rows   SUM(Sales)

Quantity

Sales — $400,000 / $350,000 / $300,000 / $250,000 / $200,000 / $150,000 / $100,000 / $50,000 / $0

2 4 6 8 10 12 14

In the example on the left (above), because the Quantity field is set to **Continuous**, it creates a horizontal axis along the bottom of the view. The green background and the axis help you to see that it's a continuous field.

In the example on the right(above), the **Quantity** field has been set to Discrete. It creates horizontal headers instead of an axis. The blue background and the horizontal headers help you to see that it's discrete.

In both examples, the **Sales** field is set to **Continuous**. It creates a vertical axis because it continuous and it's been added to the Rows shelf. If it was on the Columns shelf, it would create a horizontal axis. The green background and aggregation function (in this case, SUM) help to indicate that it's a measure.

The absence of an aggregation function in the **Quantity** field name help to indicate that it's a dimension

Special Case 1: Geographic data type :

Format geographic data in Tableau

- Depending on the type of map you want to create, you must assign certain data types, data roles, and geographic roles to your fields (or columns).
- For example, in most cases, your latitude and longitude fields should have a *data type* of **number (decimal)**, a *data role* of **measure**, and be assigned the **Latitude** and **Longitude** *geographic roles*.
- All other geographic fields should have a *data type* of **string**, a *data role* of **dimension**, and be assigned the appropriate geographic roles.

Change the data type of a column

- Tableau might incorrectly assign a Postal Code column a data type of Number (whole). To create map views, your Postal Code data must have a data type of String.

Assign geographic roles to your geographic data

- When you assign the correct geographic role to a field, Tableau assigns latitude and longitude values to each location in that field by finding a match that is already built in to the installed geocoding database. This is how Tableau knows where to plot your locations on the map.

- Geographic roles are sometimes automatically assigned to your data. You can tell a geographic role has been assigned to your data because the column includes a globe icon.

- If a geographic role is not automatically assigned, you can manually assign one to your field.

To assign or edit a geographic role:

1. On the Data Source page, click the globe icon.

2. Select **Geographic Role**, and then select a role that best matches your data.

   For example, in this case, the Country column does not have a geographic role assigned to it, so the Country/Region geographic role is assigned.
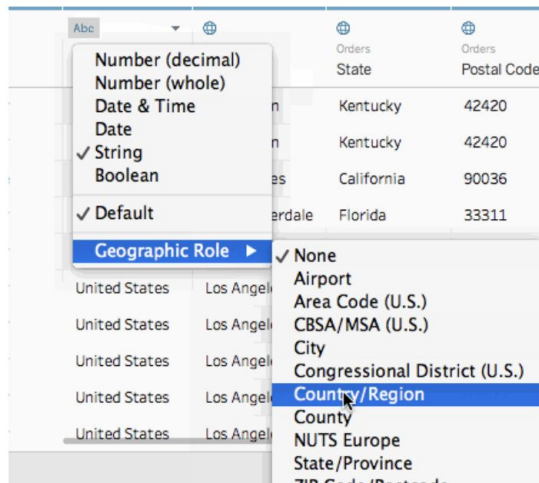
Special Case 2. **Fields that are treated as a dimension**

Field names that contain certain keywords are treated as dimensions, even if the values for those fields are numeric.

A. Keywords Code, Key, and ID

Field names that contain the following keywords and meet the conditions listed are treated as dimensions instead of measures.

Conditions:

- Contains keywords Code, ID, or Key.
- Keywords are either separated from other text in the field name by non-letter characters, all capitalized, or the first letter is capitalized in a field name that otherwise has mixed casing.
- Keywords are at the beginning or end of the field name, with leading or trailing non-letter characters. In Chinese, Japanese, and Korean, the key word must be located at the end of the field name.

Language	Keywords
English (UK and US)	Code, Id, Key
Chinese (Simplified and Traditional)	Id
French (Canada and France)	Id
German	Schlüssel, Schlussel
Italian	Chiave, Id, Codice

B. Keywords Number, Num, and Nbr

Field names that end with Number, Num, or Nbr are treated as a dimension instead of a measure. For example, "Record Number" is treated as a dimension but "Number of Records" is not. In addition, the Korean field name must be four or fewer characters long.

Language	Keywords
English (UK and US)	Number, Num, Nbr
Chinese (Simplified and Traditional)	–
French (Canada and France)	Num

C. Keywords related to dates

Field names containing keywords that are recognized as date parts are treated as dimensions. These field names can only contain only one additional word unrelated to dates to qualify. For example, "Fiscal Year" is treated as a dimension but "Fiscal Year Information" is not. In addition, Chinese field names must be four or fewer characters long and cannot contain digits. Japanese and Korean field names must be four or fewer characters long.

Language	Keywords
English (UK and US)	Year, Yr, Day, Day of Week, Week, Wk, Month, Quarter, Qtr, FY
Chinese (Simplified and Traditional)	年, 月

Tableau Connections

- Live and extracts are two ways you can make the data connection to the tableau. Live allows you real-time data while extracts are kind of batch which needs to be refreshed from time to time to get the updated data.

- In the case of live connection whatever changes will be done at the Datasource end that will be directly available to the Tableau Desktop (professional).

- While in case of extracting any changes made in the data source won't reflect in the report immediately. It will be reflected when the extract will be refreshed. Tableau Public only supports extract connections, while Tableau Desktop (professional) supports both live and extract connections.

Tableau Data Connectors

Tableau Public's native connectors can connect to the following types of data sources.

- **File Systems** such as CSV, Excel, JSon, Access, PDF, Spatial and Statistical file

- **OData :**
    - OData (Open Data Protocol) is a standard protocol for creating and consuming data, that is, it is an efficient way for business intelligence teams to get specific data from one system to another.
    - (https://data.seattle.gov/OData.svc/28ny-9ts8)

- **Web Data Connector** :

    - Web data connectors (WDCs) are web pages that **provide a data connection that is accessible over HTTP for data sources that don't already have a connector in Tableau**.
    - WDCs allow users to connect to almost any data that is accessible over the web and to create extracts for their workbooks. Data sources for a WDC can include internal web services, JSON data, REST APIs, and other sources that are available over HTTP or HTTPS.

- **Google Drive : Link**

OData : Open Data Protocol

URL : https://data.seattle.gov/OData.svc/28ny-9ts8

Business Problem 4 : Get total number of arrest flags raised by Male and Female Officers

<u>Web Data Connector</u> :

What is a ticker?

- A **ticker symbol** or **stock symbol** is an abbreviation used to uniquely identify publicly traded shares of a particular stock on a particular stock market.
- In short, ticker symbols are arrangements of symbols or characters (generally Latin letters or digits) representing specific assets or securities listed on a stock exchange or traded publicly
- Example Ticker : AAPL, DATA

Some examples of US Stock symbols include:

- A – Agilent Technologies
- AAPL – Apple
- BRK.(A/B) - Berkshire Hathaway (Class A or B shares marked by letter following period, BRK.A or BRK.B)
- C – Citigroup
- GOOG – Alphabet (parent company of Google)
- HOG – Harley-Davidson
- HPQ - Hewlett-Packard
- INTC – Intel

Note : Tableau Software trades on the New York Stock Exchange (NYSE) under the ticker symbol "DATA."

URL : https://alpha-vantage-wdc.glitch.me/

Business Problem 5 : Show the trend of stock volume

End