

# Visualization Questions

## Q1. What makes data visualisation good?

Though tons of research have been made for good data visualisation, it is mainly how people best perceive the data and information that is given or presented to them. The data visualisation should be light and must highlight essential aspects of the data; looking at important variables, what is relatively important, what are the trends and changes. Besides, data visualisation must be visually appealing but should not have unnecessary information in it.

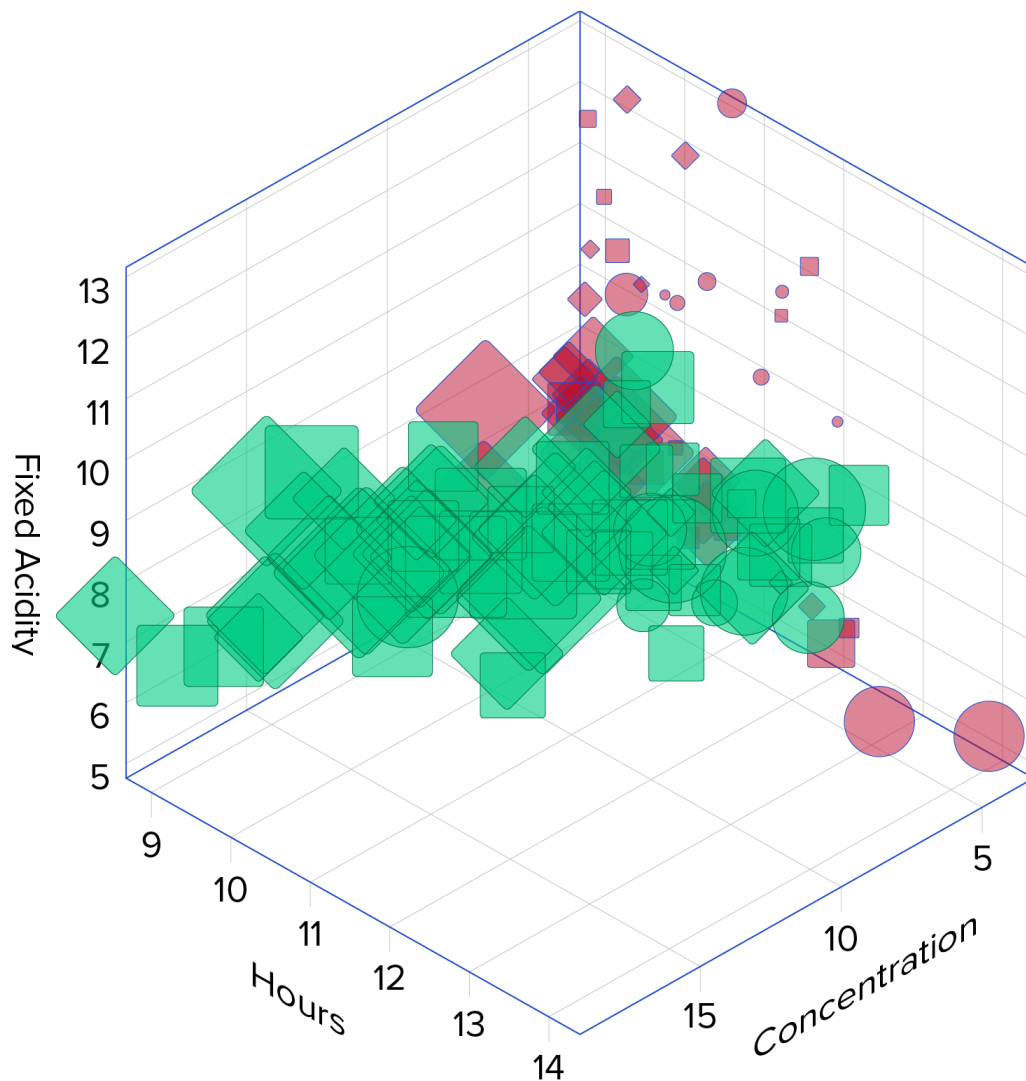
One can answer this question in multiple ways: from technical points to mentioning key aspects, but be sure to remember saying these points:

- Data positioning
- Bars over circle and squares
- Use of colour theory
- Reducing chart junk by avoiding 3D charts and eliminating the use of pie charts to show proportions
- And why sunburst visualisation is more effective for hierarchical plots.

## Q2. How can we visualize more than three dimensions of data in a single chart?

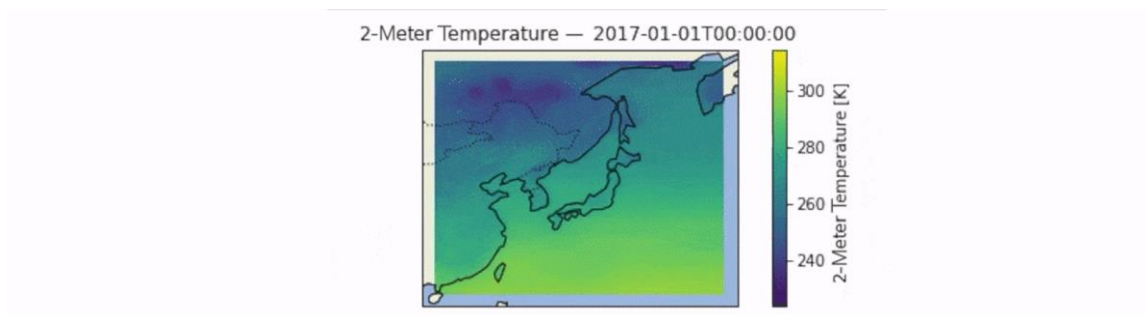
Usually, data is visually represented through a chart using locations in the image (height, width, and depth). Going beyond three dimensions, we need to make use of other visual cues to add more information. Some of the most common are:

- **Colour:** A visually appealing and intuitive way to depict both continuous and categorical data.
- **Size:** Marker size is also used to represent continuous data. Could be applied for categorical data as well, but since size differences are more difficult to detect than colour, it is not the most appropriate choice for this type of data.
- **Shape:** Lastly, we have shapes, which are an effective way to represent different classes.



Combining all of the above we can visualize up to six dimensions, though one could argue that cramming so much information in a single chart does not make for a very effective visualization.

Another possibility is to make an **animated** chart, which is quite useful to depict changes through time:



### Q3. What are the steps involved in 3D Transformation of data visualisation?

3D Transformation of data is essential because it gives a more comprehensive idea of the data plus one gets to visualise the data in depth.

The general steps are:

- Modelling Transformation
- Viewing Transformation
- Projection Transformation
- Workstation Transformation

### Q4. Why use Python Seaborn?

As mentioned earlier, the Python Seaborn library is used to ease the challenging task of data visualization and it's based on Matplotlib. Seaborn allows the creation of statistical graphics through the following functionalities:

- An API that is based on datasets allowing comparison between multiple variables
- Supports multi-plot grids that in turn ease building complex visualizations
- Univariate and bivariate visualizations available to compare between subsets of data
- Availability of different colour palettes to reveal various kinds of patterns
- Estimates and plots linear regression automatically

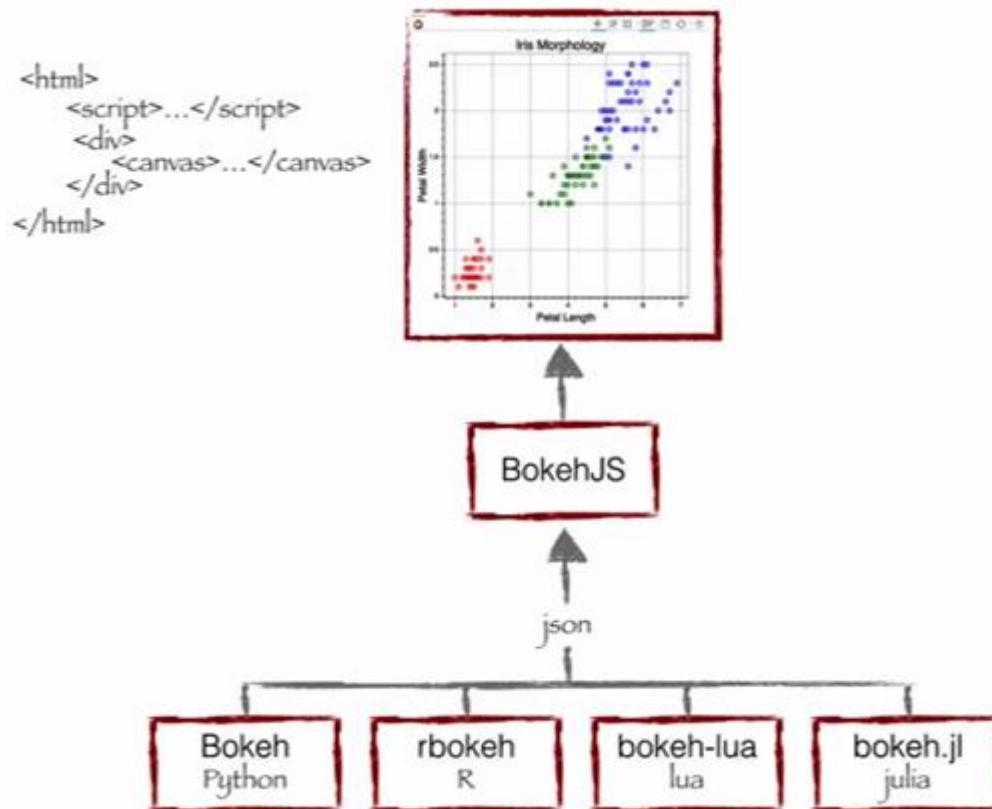
### Q5. Matplotlib vs. Seaborn

Factually, Matplotlib is good but Seaborn is better. There are basically two shortcomings of Matplotlib that Seaborn fixes:

1. Matplotlib can be personalized but it's difficult to figure out what settings are required to make plots more attractive. On the other hand, Seaborn comes with numerous customized themes and high-level interfaces to solve this issue.
2. When working with Pandas, Matplotlib doesn't serve well when it comes to dealing with DataFrames, while Seaborn functions actually work on DataFrames.

## Q6. What is Bokeh?

Bokeh is a Python library for interactive visualization that targets web browsers for representation. This is the core difference between Bokeh and other visualization libraries. Look at the snapshot below, which explains the process flow of how Bokeh helps to present data to a web browser.



## Q6. Benefits of Bokeh

Benefits of Bokeh:

- Bokeh allows you to build complex statistical plots quickly and through simple commands
- Bokeh provides you output in various medium like html, notebook and server
- We can also embed Bokeh visualization to flask and django app
- Bokeh can transform visualization written in other libraries like matplotlib, seaborn, ggplot
- Bokeh has flexibility for applying interaction, layouts and different styling option to visualization

## Q7. Methodology for creating a chart.

common methodology to create a chart:

1. Import the library and functions/ methods
2. Prepare the data
3. Set the output mode (Notebook, Web Browser or Server)
4. Create chart with styling option (if required)
5. Visualize the chart

## Q8. What is Tableau?

- Tableau is a business intelligence software.
- It allows anyone to connect to the respective data.
- Visualizes and creates interactive, shareable dashboards.

## Q9. What is the difference between .twb and .twbx extension?

- A .twb is an xml document which contains all the selections and layout made you have made in your Tableau workbook. It does not contain any data.
- A .twbx is a 'zipped' archive containing a .twb and any external files such as extracts and background images.

## Q10. How many maximum tables can you join in Tableau?

You can join a maximum of 32 tables in Tableau.

## Q11. What are the different connections you can make with your dataset?

We can either connect live to our data set or extract data onto Tableau.

- **Live:** Connecting live to a data set leverages its computational processing and storage. New queries will go to the database and will be reflected as new or updated within the data.
- **Extract:** An extract will make a static snapshot of the data to be used by Tableau's data engine. The snapshot of the data can be refreshed on a recurring schedule as a whole or incrementally append data. One way to set up these schedules is via the Tableau server.

The benefit of Tableau extract over live connection is that extract can be used anywhere without any connection and you can build your own visualization without connecting to database.

## Q12. What is Tableau Data Server?

Tableau server acts a middleman between Tableau users and the data. Tableau Data Server allows you to upload and share data extracts, preserve database connections, as well as reuse calculations and field metadata. This means any changes you make to the dataset, calculated fields, parameters, aliases, or definitions, can be saved and shared with others, allowing for a secure, centrally managed and standardized dataset. Additionally, you can leverage your server's resources to run queries on extracts without having to first transfer them to your local machine.

## Q13. What is a dual axis?

Dual Axis is an excellent phenomenon supported by Tableau that helps users view two scales of two measures in the same graph. Many websites like Indeed.com and other make use of dual axis to show the comparison between two measures and their growth rate in a septic set of years. Dual axes let you compare multiple measures at once, having two independent axes layered on top of one another.

## Q14. What is the difference between a tree map and heat map?

A heat map can be used for comparing categories with colour and size. With heat maps, you can compare two different measures together.

A tree map also does the same except it is considered a very powerful visualization as it can be used for illustrating hierarchical data and part-to-whole relationships.

## Q15. How to add Custom Colour to Tableau?

Adding a Custom Colour refers to a power tool in Tableau. Restart you Tableau desktop once you save .tps file. From the Measures pane, drag the one you want to add colour to **Colour**. From the colour legend menu arrow, select **Edit Colours**. When a dialog box opens, select the palette drop-down list and customize as per requirement.

### **Q16. What Is Depth Cueing in Visualization?**

Basic problem for visualization techniques is called depth cueing. Some 3D objects are without depth information visible line and surface identification to highlight the visible lines display visible lines as dashed lines removing the invisible lines.