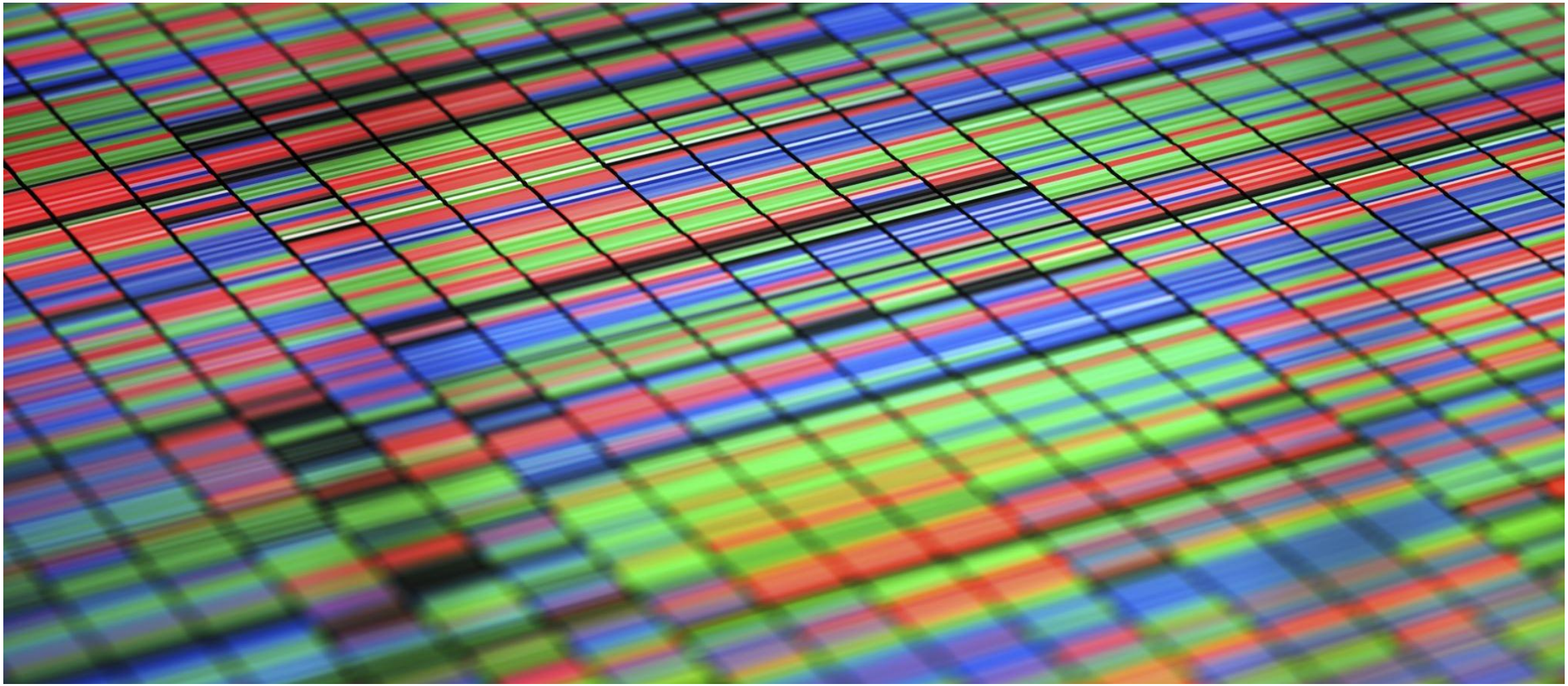
TYPES OF DATA.



TYPES OF DATA.

The foundational Stage - recognising that data exists and influences decisions.

Key Abilities:

Recognise that data is information we can capture, store and analyse.

Understand why we visualise data.

Notice when data is being used to persuade or inform.

Identify data types and simple charts.

Typical Learner Mindset:

"I know data is important, and I can spot when it is being used"

AGENDA.

- Types Of Data
- Categorical Data
- Ordinal Data
- Quantitative Data
- Encoding Data In Charts
- Use of Colour in Visualisations
- Summary

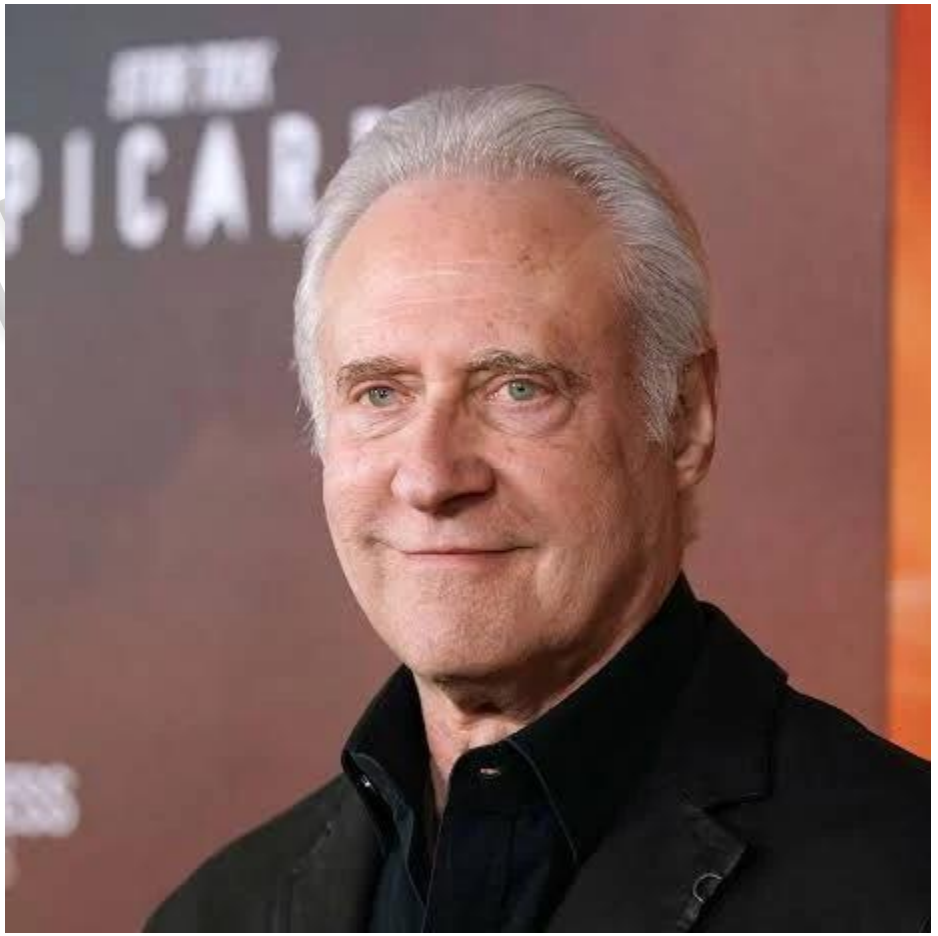
TYPES OF DATA.

- There are **three** types of data:

  - Categorical.
  - Ordinal.
  - Quantitative.

CATEGORICAL DATA.

- **Categorical** (or nominal) data **represents things**.
  - These things are **mutually exclusive** labels without any **numerical value**.
  - His **name** is Brent Spiner.
  - By **profession** he is an actor.
  - He **played** the character Data in the TV show Star Trek.



- Name, profession, character, and TV show are all **categorical data types**.
- Other types include - gender, product category and city.

ORDINAL DATA.

- **Ordinal data** is similar to categorial data, except it has a **clear order**.
  - Other types of ordinal data include education experience, satisfaction level, and salary bands in an organisation.
  - Although **ordinal values** often have **numbers** associated with them, the **interval between** those values is **arbitrary**.
  - For example, the **difference** in an organisation between pay scales 1 and 2 might be very different from that between pay scales 4 and 6.

- Brent Spiner's date of birth is Wednesday, February 2, 1949.
- He appeared in all seven seasons of Star Trek: The Next Generation.
- Data's rank was lieutenant commander.
- Data was the fifth of six androids made by Dr. Noonien Soong.

QUANTITATIVE DATA.

- **Quantitative data** are the **numbers**.
- **Quantitative** (or numerical) data is data that can be **measured** and **aggregated**.

  - Brent Spiner's date of birth is Wednesday, February 2, 1949.
  - His height is 5 ft 9 in (180 cm) tall.
  - He made 177 appearances in episodes of Star Trek.
  - Data's positronic brain is capable of 60 trillion operations per second.

- You'll have noticed that **date of birth** appears in **both ordinal** and **quantitative** data types.

  - **Time** is **unusual** in that it can be **both**.

QUANTITATIVE DATA.

- **Quantitative** data can be expressed in **two** ways:
  - **Discrete** or **Continuous.**

- **Discrete** data is presented at predefined, **exact points** there's **no "in-between"**.

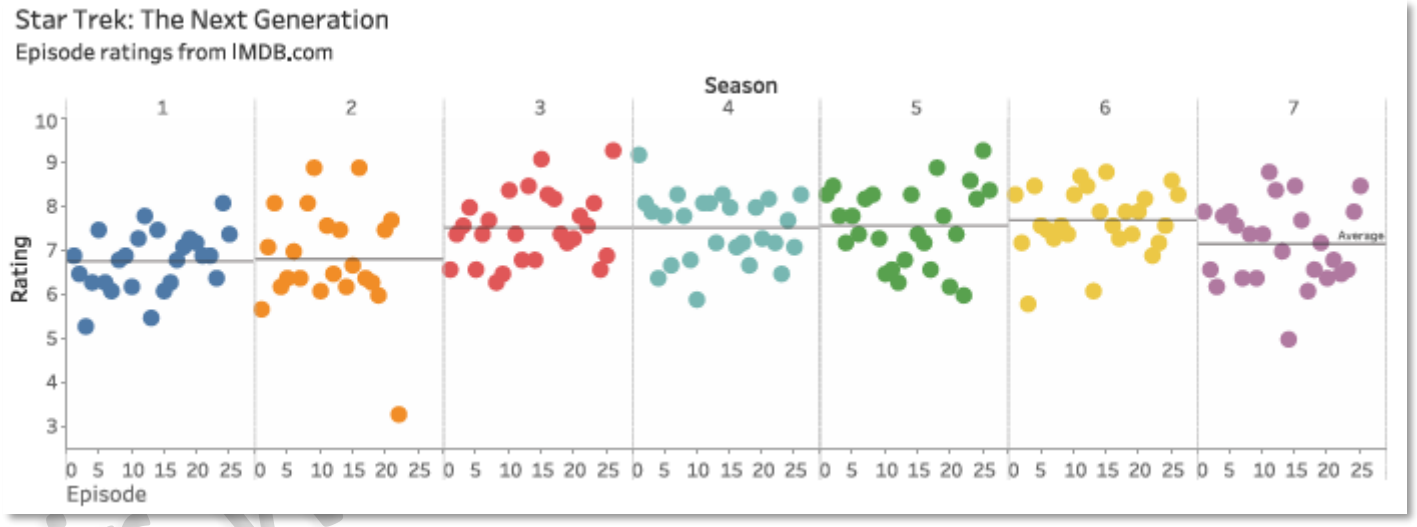
- For example:
  - Brent Spiner appeared in 177 episodes of Star Trek.
  - He couldn't have appeared in 177.5 episodes.

- **Continuous** data allows for the "**in-between**," as there is an infinite number of possible **intermediate values**.

- For example:
  - Brent Spiner grew to a height of 5 ft 9 in but at one point in his life he was 4 ft 7.5 in tall.

ENCODING DATA IN CHARTS.

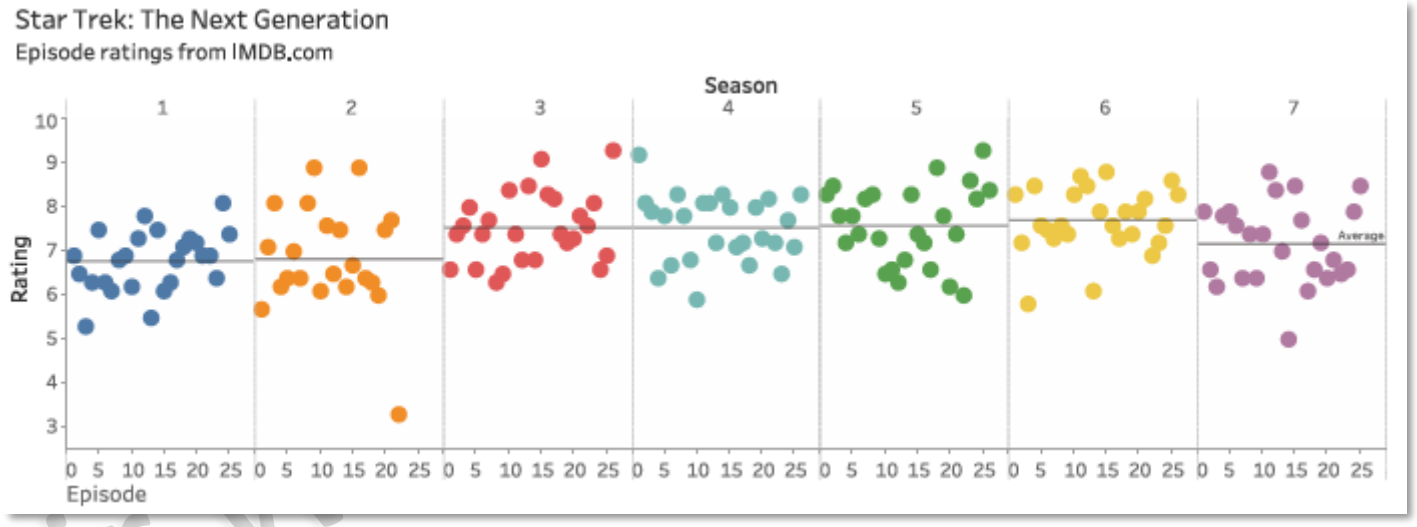
Have a try



Data	Data Type	Encoding	Note
Episode		Position	Each episode is represented by a dot. Each dot has its own position on the canvas.
Episode Number		Position	The x-axis shows the number of each episode in each season.
Season		Colour Position	Each season is represented by a different colour (hue). Each season also has its own section on the chart.
IMDB Rating		Position	The better the episode, the higher it is on the y-axis.
Average Season Rating		Position	The horizontal bar in each pane shows the average rating of the episodes in each season. There is some controversy over whether you should average ordinal ratings. We believe that the practice is so common with ratings it is acceptable.

ENCODING DATA IN CHARTS.

Have a try

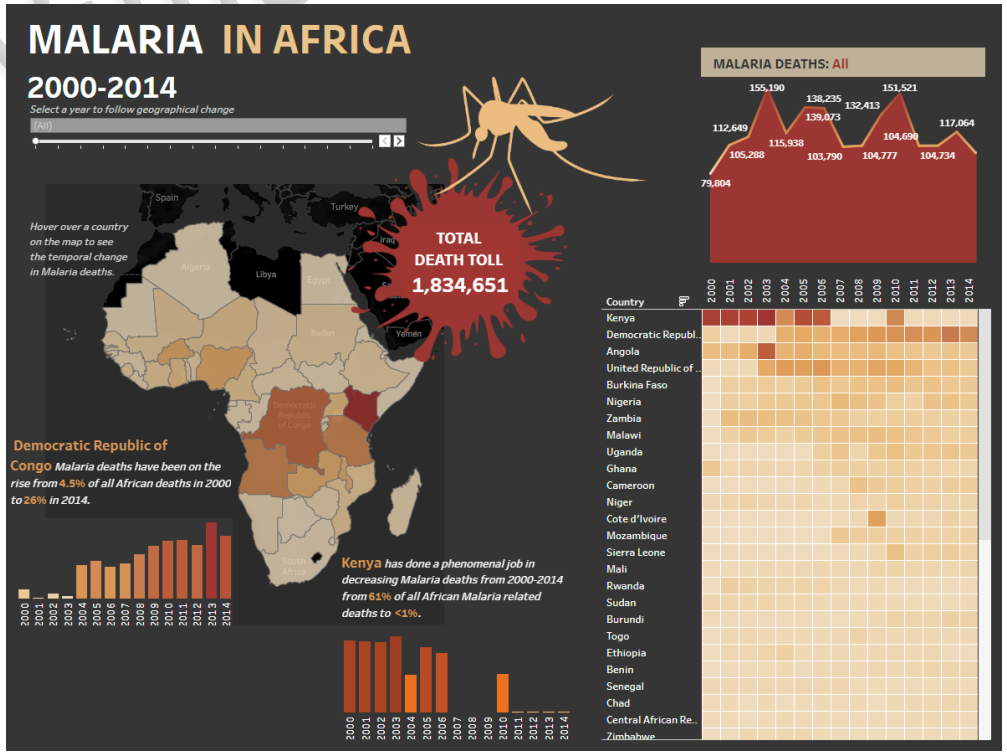


Data	Data Type	Encoding	Note
Episode	Categorical	Position	Each episode is represented by a dot. Each dot has its own position on the canvas.
Episode Number	Ordinal	Position	The x-axis shows the number of each episode in each season.
Season	Ordinal	Colour Position	Each season is represented by a different colour (hue). Each season also has its own section on the chart.
IMDB Rating	Ordinal	Position	The better the episode, the higher it is on the y-axis.
Average Season Rating	Quantitative	Position	The horizontal bar in each pane shows the average rating of the episodes in each season. There is some controversy over whether you should average ordinal ratings. We believe that the practice is so common with ratings it is acceptable.

ENCODING DATA IN CHARTS.

Data	Data Type	Encoding	Note
Country	Categorical	Position	The map shows the position of each country. In the highlight table, each country has its own row.
Deaths	Quantitative	Colour	The map and table use the same colour legend to show deaths.
Year	Ordinal	Position	Each year is a discrete column in the table.

University - Malaria In Africa | Tableau Public



ENCODING DATA IN CHARTS.

University - French Impressionist Painters | Tableau Public

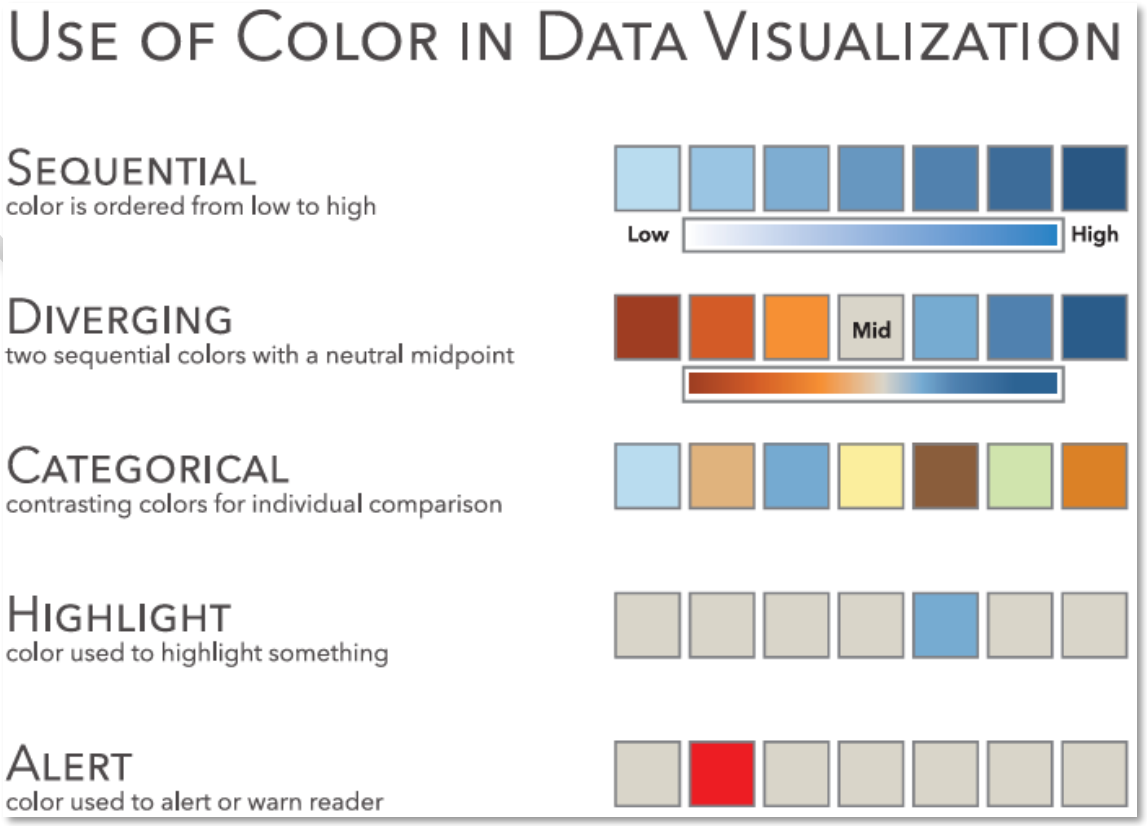
10 Steps to Winning the Tableau Iron Viz Championship

Tableau - YouTube

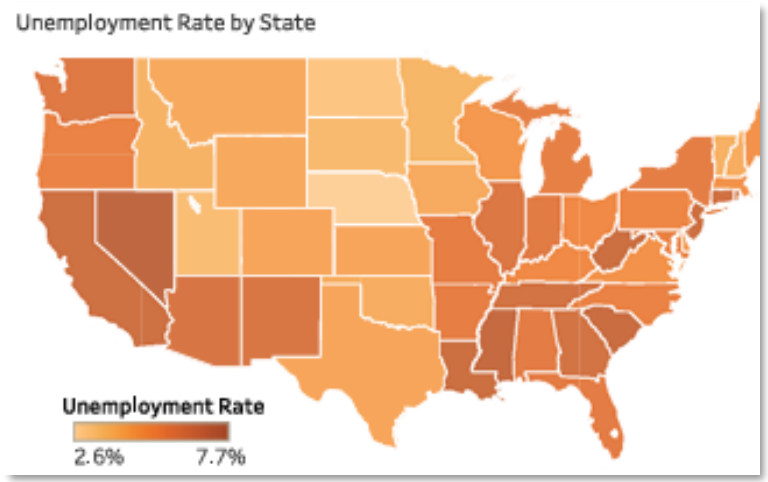
USE OF COLOUR IN DATA VISUALISATION.

- **Colour** is one of the most **important** things to understand in data visualisation.

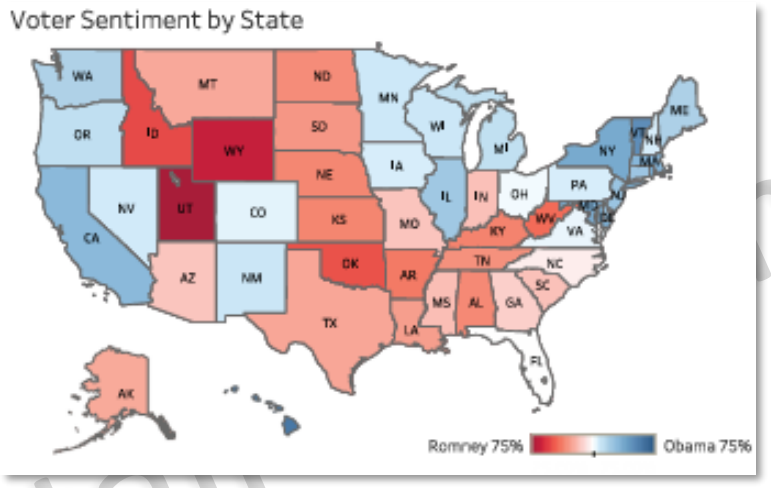
  - **Colour** should be used in data visualisation in **three** primary ways: **sequential**, **diverging**, and **categorical**.
  - In addition, there is often the need to **highlight data** or alert the reader of **something important**.



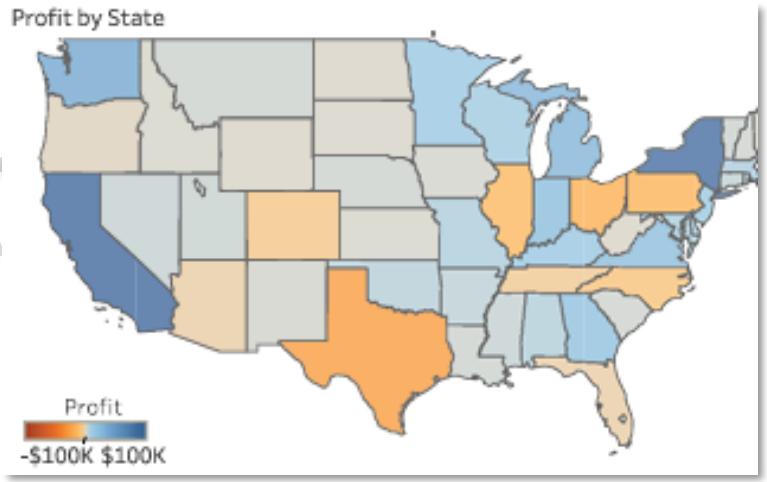
USE OF COLOUR IN DATA VISUALISATION.



Unemployment rate by state using a sequential colour scheme.



Degree of Democratic (blue) versus Republican (red) voter sentiment in each state.



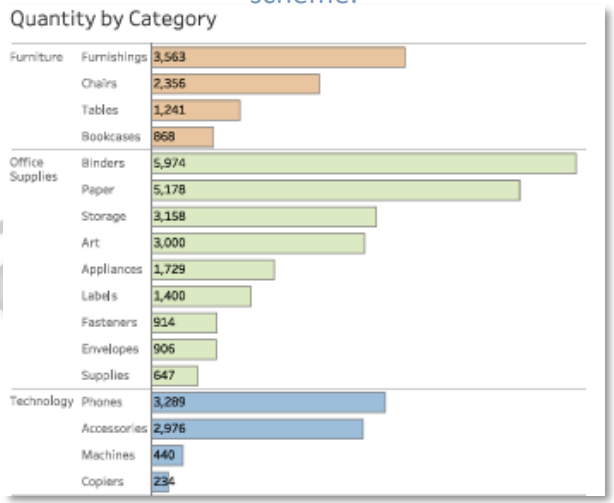
Profit by state using a diverging colour scheme.

- **Sequential colour** is the use of a **single colour** from light to dark.
  - The image shows the unemployment rate by state.
- **Diverging colour** is used to show a **range diverging** from a **midpoint**.
  - This colour can be used in the same manner as the sequential colour scheme but can **encode two different ranges of a measure**.
- **Diverging colour** can also be used to show the profit by state with the **midpoint being the average**.
  - The image shows an example where profit (positive) is shown in blue and (loss) in orange.

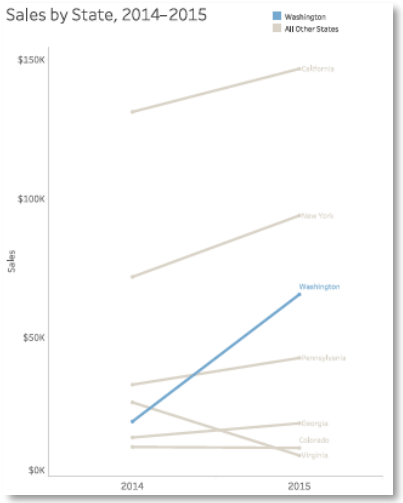
USE OF COLOUR IN DATA VISUALISATION.

- **Categorical** colour uses **different colour hues** to distinguish between different categories.
  - For example, we can establish categories involving apparel (e.g., shoes, socks, shirts, hats, and coats).
  - or vehicle types (e.g., cars, mini vans, sport utility vehicles, and motorcycles).

Quantity of office supplies in three categories using a categorical colour scheme.



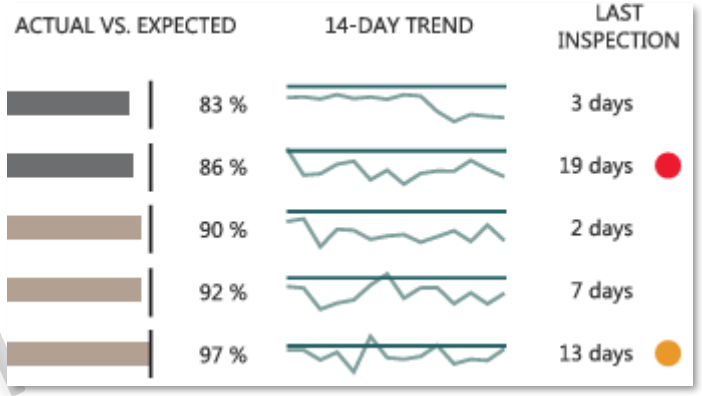
- **Highlight** colour is used when there is something that needs to **stand out** to the reader, but not alert or alarm them.
  - Highlights can be used in a number of ways, as in **highlighting a certain data point**, text in a table, a certain line on a line chart, or a specific bar in a bar chart.



Slopegraph showing sales by state, 2014–2015, using a single colour to highlight the state of Washington.

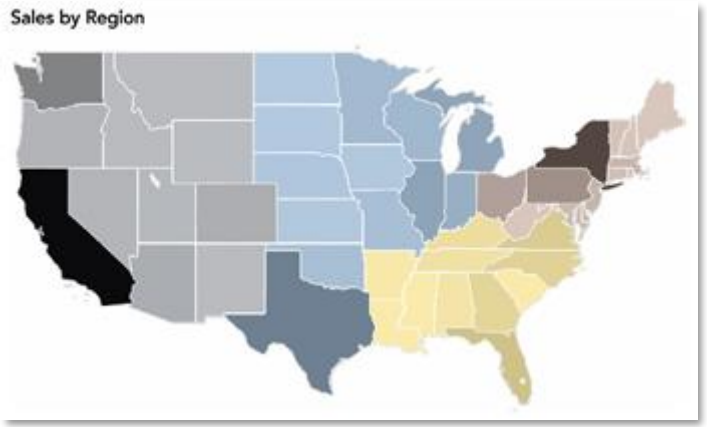
USE OF COLOUR IN DATA VISUALISATION.

- **Alerting colour** is used when there is a need to draw **attention** to something for the reader.
- In this case, it's often best to use **bright, alarming colours**, which will **quickly** draw the reader's **attention**.



Red and orange indicators to alert the reader that something on the dashboard needs attention.

- It is also possible to have a **categorical-sequential** colour scheme.
  - In this case, each category has a **distinct hue** that is darker or lighter depending on the **measurement** it is representing.
- An example is a four-region map using **categorical colours** (i.e., gray, blue, yellow, and brown) but at the same time **encoding** a measure in those regions using **sequential colour**; let's assume that sales are higher in states with darker shading.



Sales by region using four categorical colours and the total sales shown with sequential colour.

COLOUR VISION DEFICIENCY (CVD).

- **8% of males** have **colour vision blindness.** (Birch 1993).
  - **Protanopia** is the lack of long-wave cones (red weak).
  - **Deuteranopia** is the lack of medium-wave cones (weak green).
  - **Tritanopia** is the lack of short-wave cones (blue).

- Red, Green and Orange all appear brown for someone with strong CVD.
- Blue-Orange palette is often referred to as being "**colour-blind friendly**".
- Kazunori Asada's website allows users to upload images and simulate how they would appear to people with different forms of CVD.

Chromatic Vision Simulator

SUMMARY.

- To recap, we have discussed the different types of data any how they are used on dashboards, and we have also looked at the use of colour in data.

- Next, we will look at Common Chart Types.

SUMMARY.

- **Data Types:**
- Categorical, Ordinal, Quantitative (discrete or continuous).

- **Chart encoding:**
- Position, colour, and aggregation are used to represent different data types.

- **Colour Use:**
- Sequential, diverging, categorical, highlighted, and alerting schemes help convey meaning and emphasis.

- **Accessibility:**
- Colour-bind friendly palettes (e.g. blue-orange) ensure dashboards remain clear for all viewers.