



**Data Science  
Bootcamp**

Hyperiondev

# Data Visualisation

WELCOME TO THE EVENT HANDLING TASK

# Your Lecturer for This Session



**Christiaan Joubert**

# Lecture – Housekeeping

- ❑ The use of disrespectful language is prohibited in the questions, this is a supportive, learning environment for all - please engage accordingly.
- ❑ No question is daft or silly - **ask them!**
- ❑ There are Q/A sessions midway and at the end of the session, should you wish to ask any follow-up questions.
- ❑ You can also submit questions here:  
[hyperiondev.com/sbc4-ds-questions](https://hyperiondev.com/sbc4-ds-questions)
- ❑ For all non-academic questions, please submit a query:  
[hyperiondev.com/support](https://hyperiondev.com/support)
- ❑ Report a safeguarding incident:  
[hyperiondev.com/safeguardreporting](https://hyperiondev.com/safeguardreporting)
- ❑ We would love your feedback on lectures:  
<https://hyperiondev.wufoo.com/forms/zsgv4m40ui4i0g/>

# Lecture – Code Repo

Go to: [github.com/HyperionDevBootcamps](https://github.com/HyperionDevBootcamps)

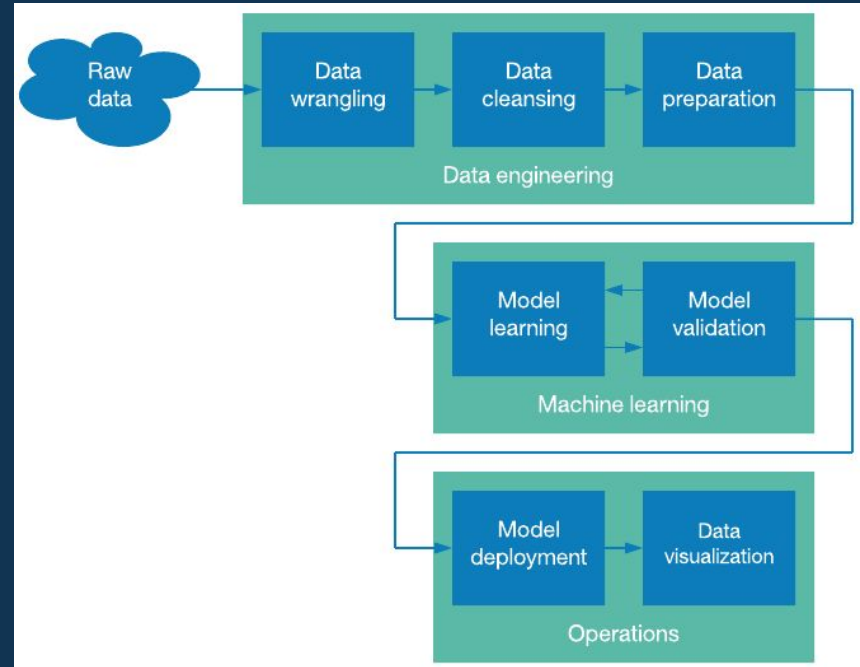
Then click on the “**C4\_DS\_lecture\_examples**” repository, do view or download the code.

# Objectives

1. Understand importance of data visualisation
2. Use Tableau to generate dynamic and interactive graphs

# The Data Science Pipeline

- Part of data wrangling is understanding the data.
- For large chunks of data, you need to visualise it.
- Therefore, it is crucial to Data Science.



# Approaching Data Visualisation

- **Start with a processed and clean dataset.**
  - Missing/unstructured data will likely not yield anything useful.
- **Know your dataset.**
  - Useful to understand how data was gathered
- **Determine what you want to find.**
  - Ask questions about the data, these can be answered using visualisation.

# Approaching Data Visualisation (Cont.)

- **Create data visualisations.**
  - Time to answer the questions that you asked previously. Depending on what you asked, you will need to find a visualisation that works.
- **Refine your visualisation.**
  - Graphs should be easy to read at a first glance.
- **Note down your findings.**
  - At this stage, you can start analysing your graphs and finding conclusions.

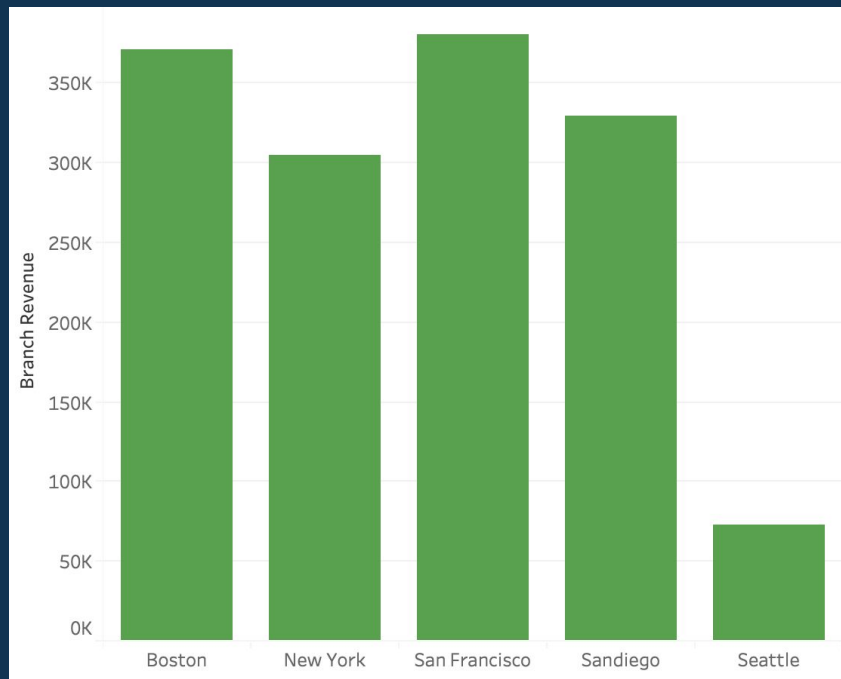


# A Note on Types of Data

- Discrete:
  - Can take on specific values, with an infinite range. E.g. [1, 2, 3, 4, 5, ...]
- Categorical:
  - Can take on specific values, with a limited range. E.g. [Dog, Cat, Hamster, Fish]
- Continuous:
  - Available values is a spectrum. This means that there is an infinite number of values. An example is temperature or distance.
- Time Series:
  - Data changes along with some time-related progression.

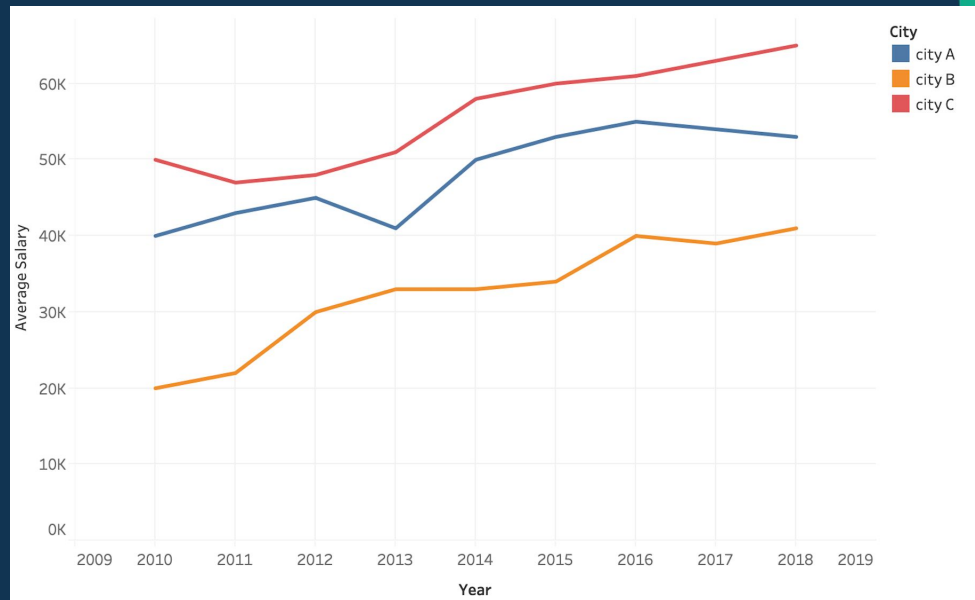
# Bar Chart

- Good for plotting data that is:
  - Categorical vs continuous/discrete.



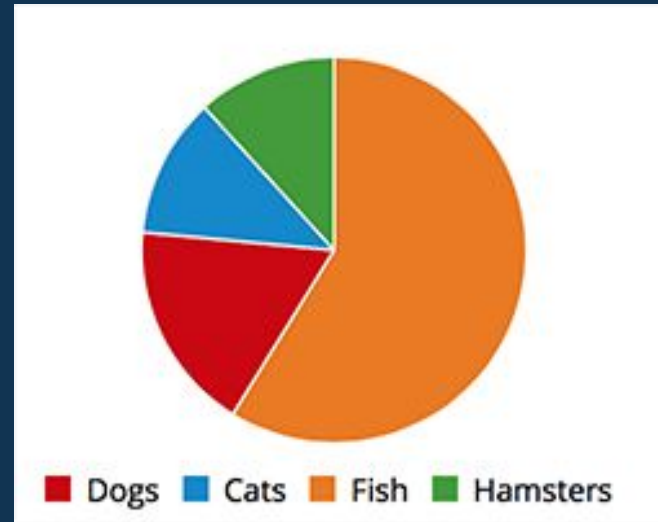
# Line Graphs

- Good for plotting data that is:
  - Discrete/continuous vs discrete/continuous
  - Time Series



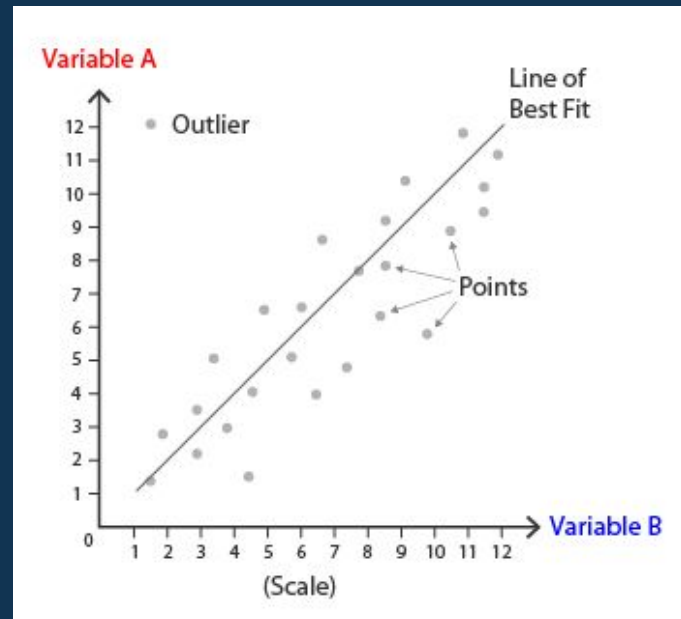
# Pie Chart

- Good for plotting data that is:
  - Categorical vs discrete
- Great for getting a sense of proportions.



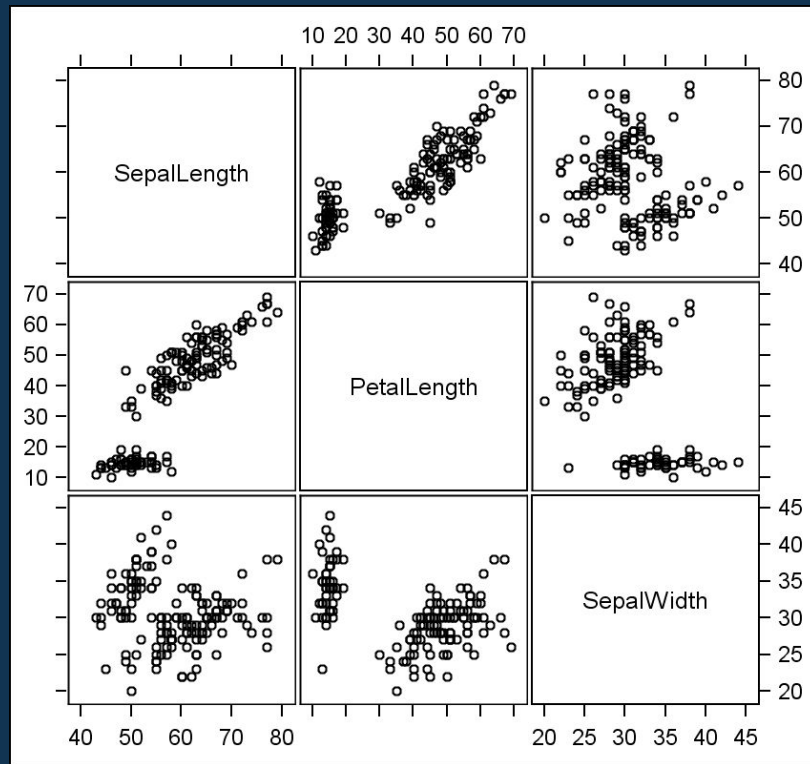
# Scatterplot

- Great for plotting:
  - Discrete vs Discrete
  - Continuous vs Continuous
- Can be useful for finding relationships between two variables.



# Scatterplot Matrix

- Shows relationships between multiple variables.
- Example: Iris dataset ->
- Examining relationships between Sepal Length, Petal Length and Sepal Width.
- What relationships can we see here? Are there any obvious separations?



# Double Axis Chart

- Orange line shows Sales.
- Blue bar graph shows profit.
- Useful graphic to understand costs.

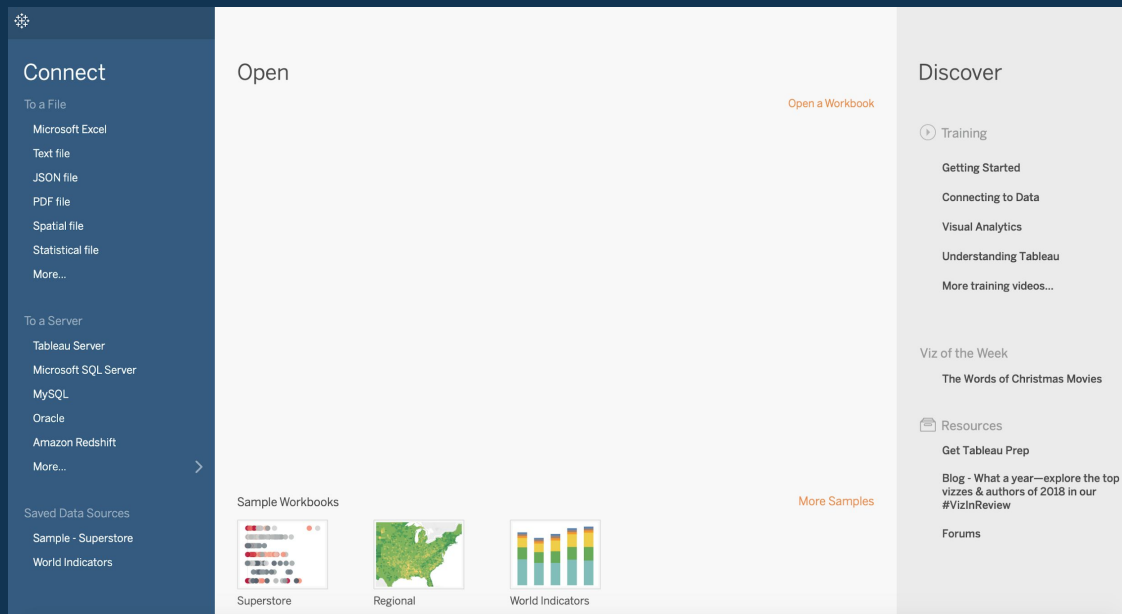


# Tableau

- Strong and interactive data visualisation tool.
- Allows you to easily create visualisations.
- Create many different and interactive visualisations and dashboards.



# Loading Data



You can connect to anything, from a CSV file to a server such as Google Analytics.

# Creating a Visualisation – Bar Graph

The screenshot displays a data visualization tool interface. On the left, the 'Data' pane shows a dataset named 'department\_sales'. Under 'Dimensions', the fields 'Date', 'Department', and 'Measure Names' are listed, with a red box highlighting them and the text 'The column names'. Under 'Measures', the fields 'Sales', 'Number of Records', and 'Measure Values' are listed, with a red box highlighting them and the text 'Values'. A red arrow points from the 'Measures' box to the 'Marks' area, with the text 'Drag and drop'. The 'Marks' area shows a dropdown set to 'Automatic' and buttons for 'Color', 'Size', 'Text', 'Detail', and 'Tooltip'. The main workspace is labeled 'Sheet 1' and contains two 'Drop field here' prompts. On the right, a 'Visualizations' gallery is shown, containing various chart types like bar graphs, line charts, and pie charts, with a red box highlighting it and the text 'Visualizations'. Below the gallery, text reads 'Select or drag data' and 'Use the Shift or Cmd key to select multiple fields'.

Drag and drop

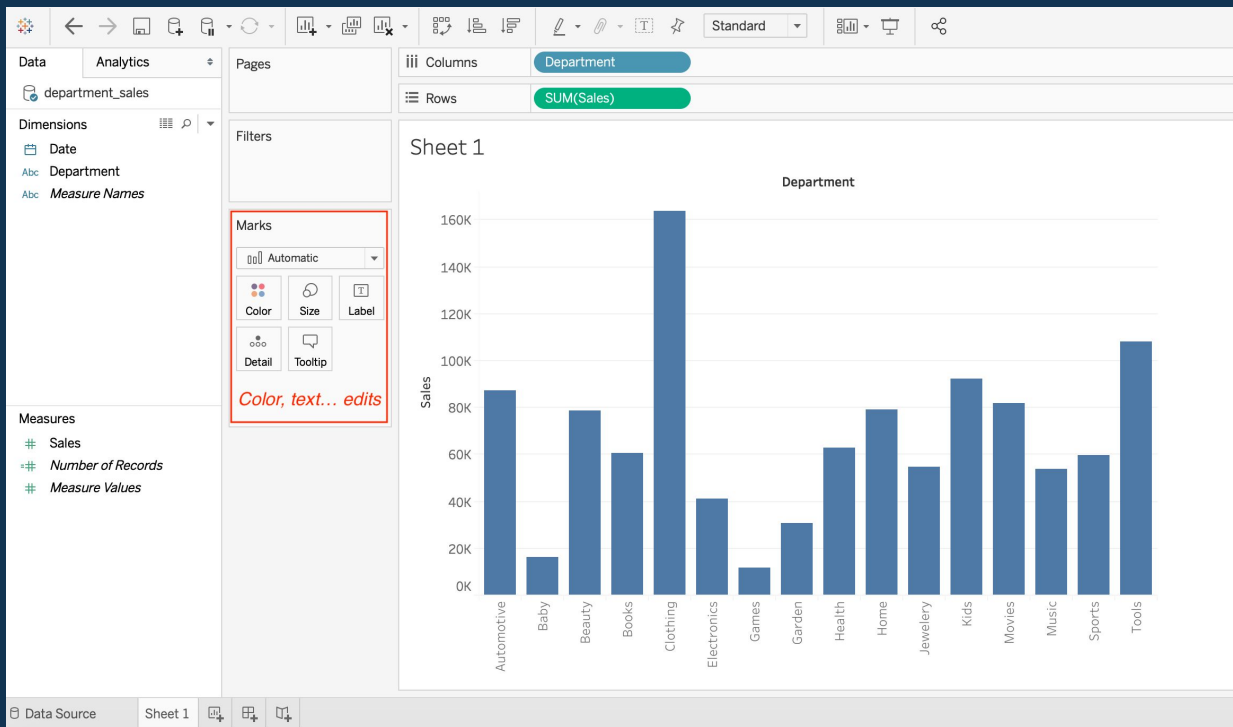
The column names

Values

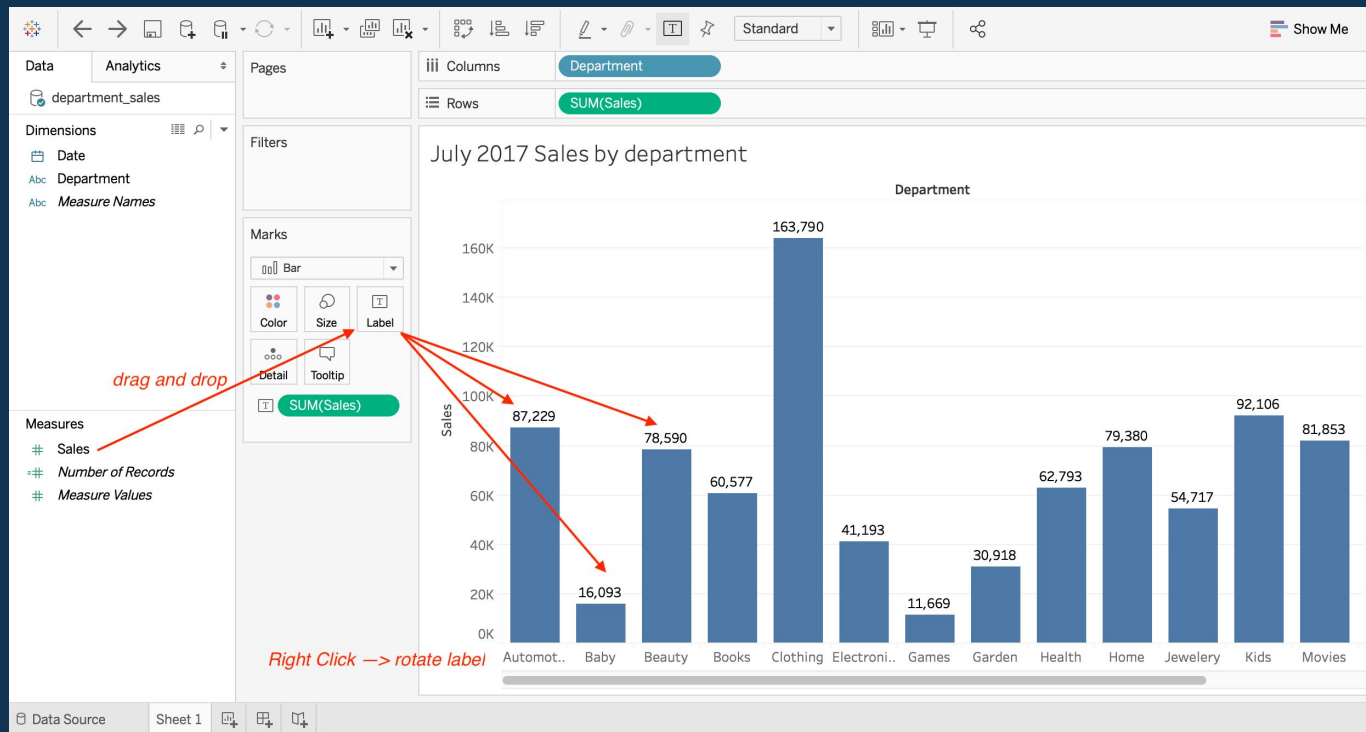
Visualizations

Simply select your columns, values and visualisations

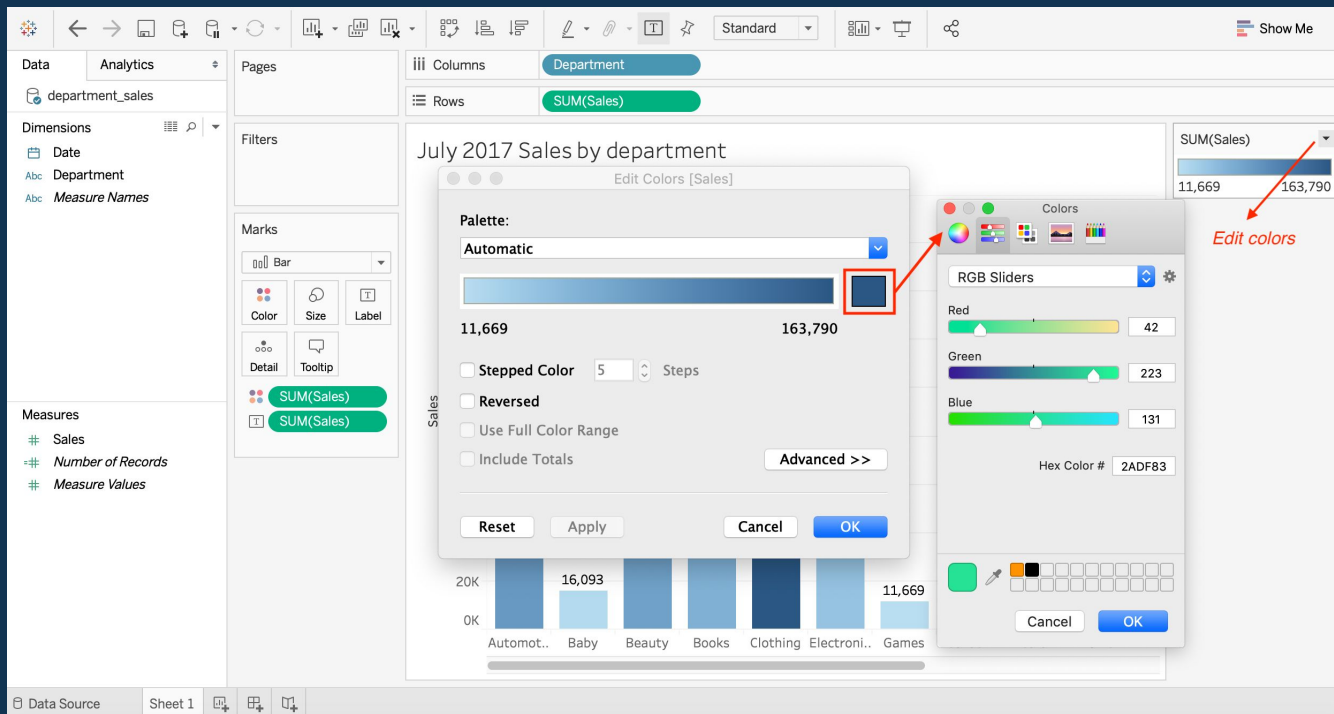
# Editing Your Graph



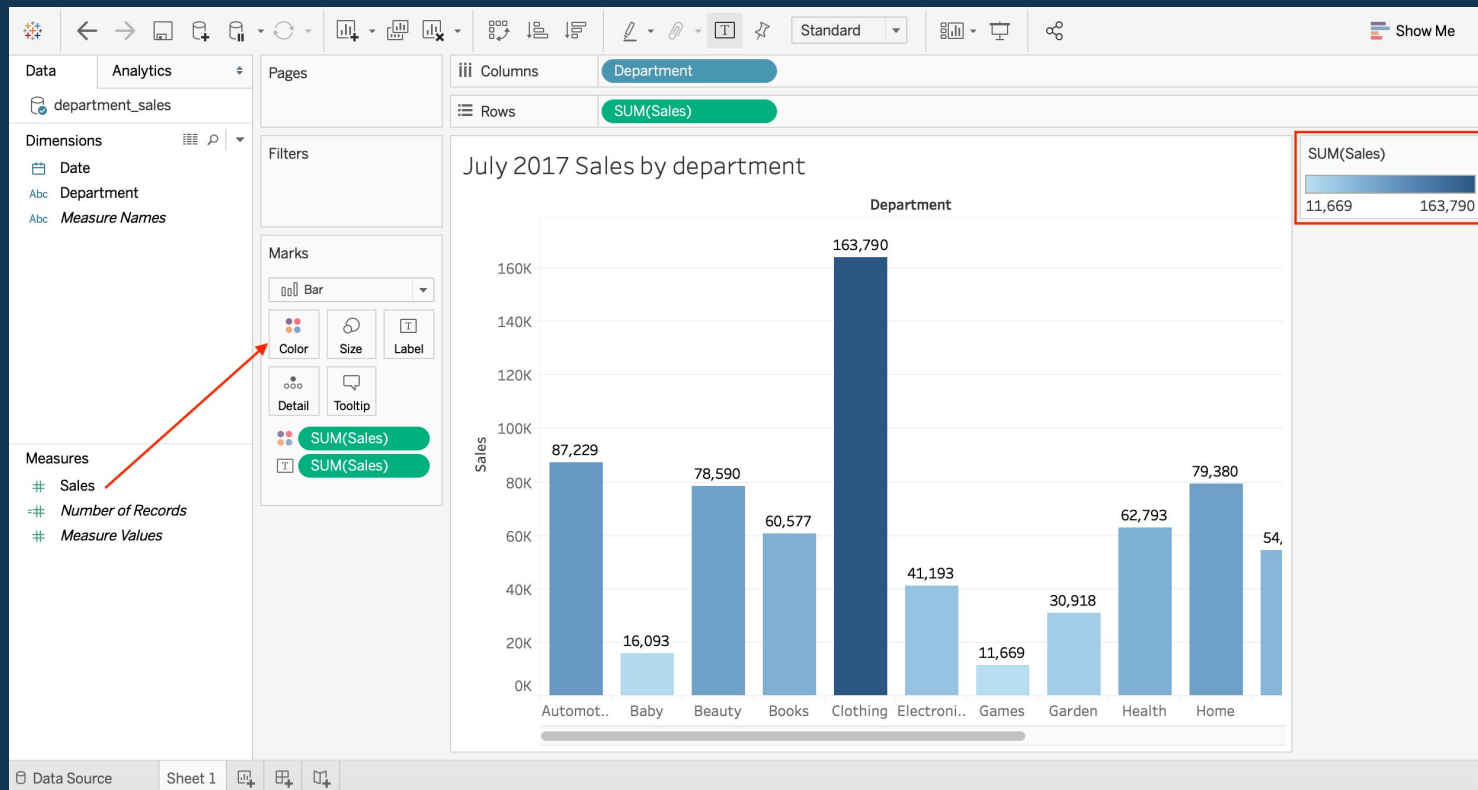
# Adding Labels to Bars



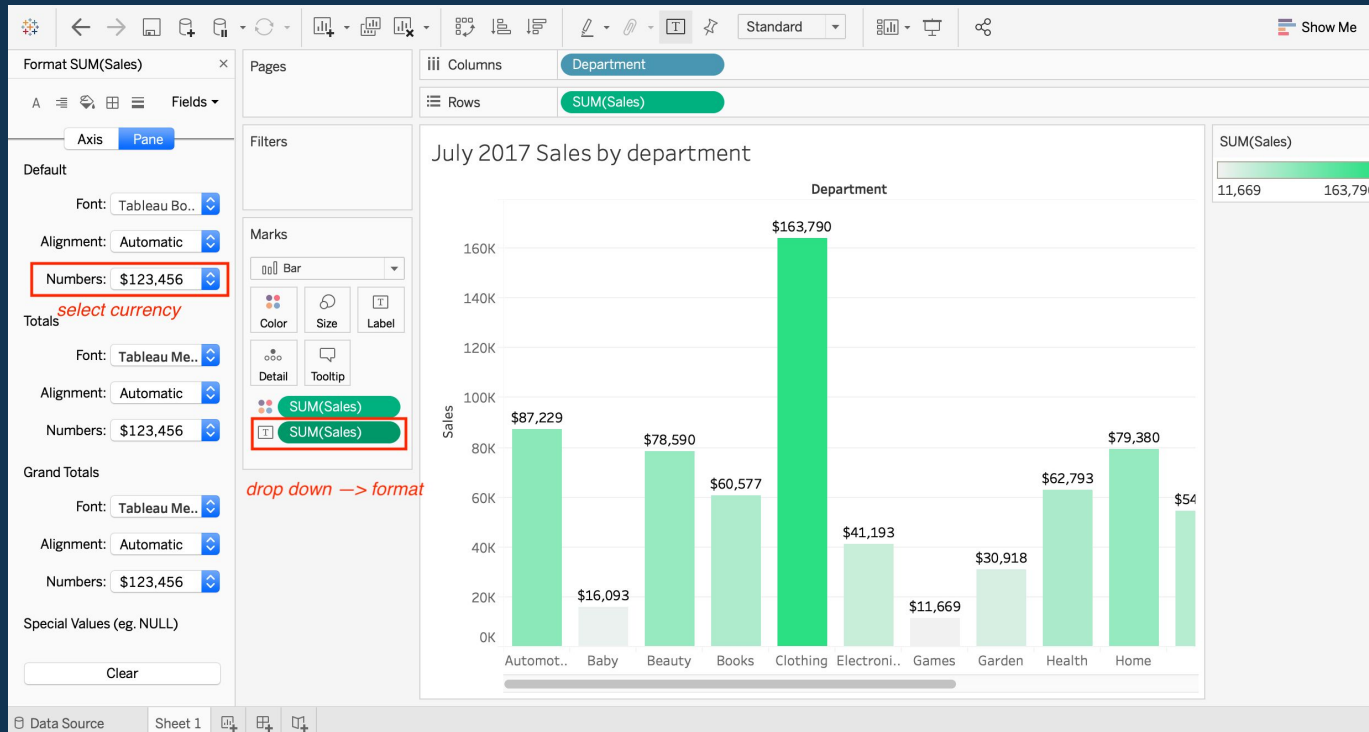
# Editing Your Colours



# Dynamic Bar Colouring



# Formatting for Currency

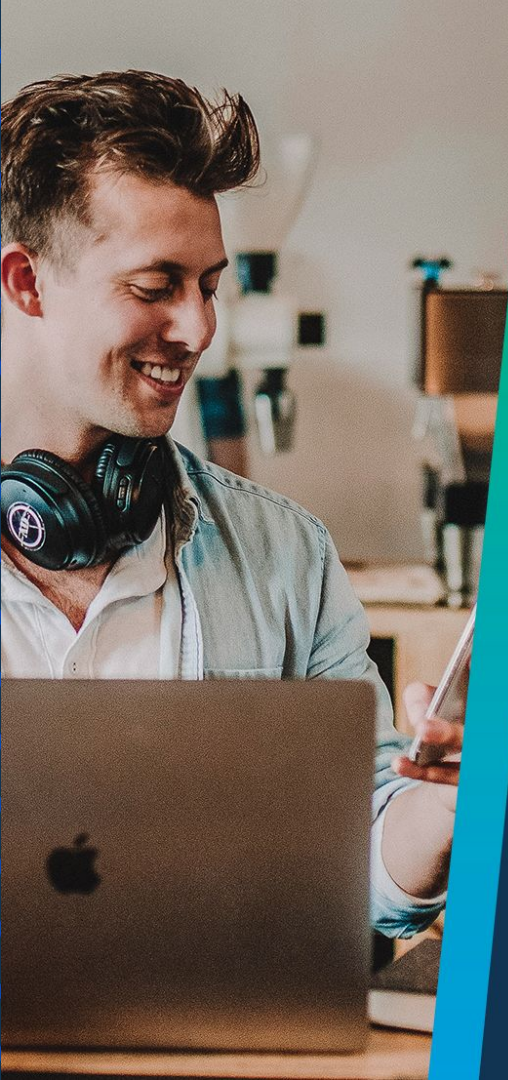


Hyperiondev

# Q & A Section

**Please use this time to ask any questions relating to the topic, should you have any.**





Hyperiondev

**Thank You for  
Joining Us**