FIT3179 DATA VISUALISATION

2nd Semester, 2018

Tutorial Week 5: Visualisation Analysis of Chart Chunk and Data-ink Ratio, Tableau: Dual Axis Visualisation

Overview

In this week's tutorial, you will do these following activities:

1. Chart Junk: Visualisation Analysis

2. Chart Junk: Cleaning Exercise

3. Tableau: Dual Axis Visualisation

4. Tableau: Dual Axis and Simple Label Exercise

1. Chart Junk: Visualisation Analysis

In groups of three, pick and examine one of the visualisations in the list and answer the questions below. Prepare a mini presentation of 3 minutes. Select one visualisation from the list below:

1. [Analyzing the Gender Representation of 34,476 Comic Book Characters](https://pudding.cool/2017/07/comics/)
   https://pudding.cool/2017/07/comics/

2. Gapminder [Human Development](#), 2005.

3. An example from the [Tableau Public Gallery](#).

Question
1. Is there any chart junk? Are there ornamental elements that could be removed? Is it better to retain them?
2. What is the data-ink ratio? How could you increase the data-ink ratio?
3. Is the visualisation lying in any way?
4. What is the macro reading of the visualisation? (What story is the designer probably trying to tell?)
5. What is the micro reading of the visualisation?
6. Which of the seven narrative visualisation genres is used?
7. Which of the Gestalt Principles and Graphical Elements guide the viewer's eye?

2. Chart Junk: Cleaning Exercise

2.1. Part 1

Download the Excel file *ChartJunk1.xlsx* from Moodle and open it in Excel. What elements of chart junk can you identify in this graph? Create a second graph using the same data without chart chunk that maximises the data-ink ratio.

2.2. Part 2

Download the Tableau file *ChartJunk2.twbx* from Moodle and open it in Tableau. Improve this Dashboard to reduce the cognitive load required to read this visualisation:

- remove chart chunk
- remove unnecessary elements (For example, could you remove some of the legends? Or could you replace some of the legends with annotations on the chart?)
- increase the data-ink ratio
- use consistent colours, prefer grey for less important marks, use colour to highlight important elements.
- Use typography consistently: Similar elements should have the same size and type weight (for example, labels on axes).

3. Tableau: Dual Axis Visualisation

In this tutorial, you will learn how to create a meaningful visualisation by combining two axes. We will use the Pedestrian Volume.csv dataset of the last tutorial for this activity. You can download again dataset on Week 5 Moodle page.

1. Create a new sheet

2. First, we want to create a **bar chart** showing the **average number of pedestrians each day**.

WHAT YOU CAN DO

Don't scroll don't to the next page, see if you can make it!

3. So, here is how to do it. Drag and drop the **Hourly Count** measure to Rows. Then, change the measure to **average**.
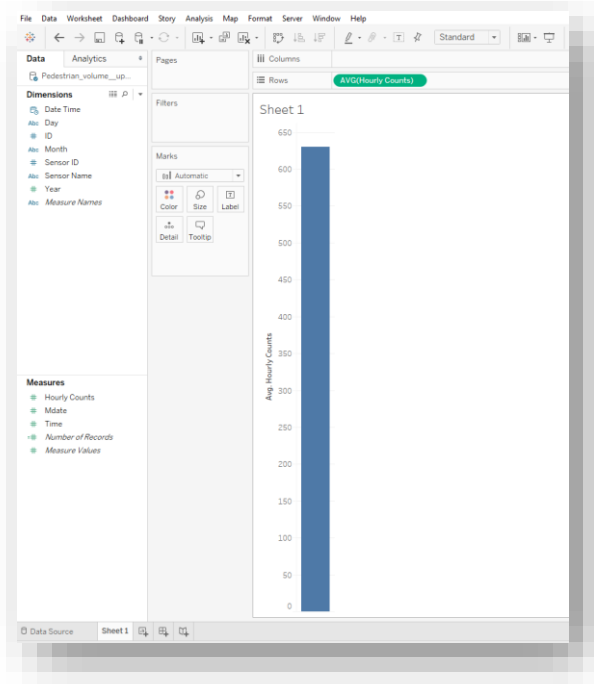

4. Drag and drop the **Day** dimension into Columns. You will immediately get the expected result: a **bar chart** showing the **daily average number of pedestrians**. Please note this is the average number of pedestrians in all sensors.
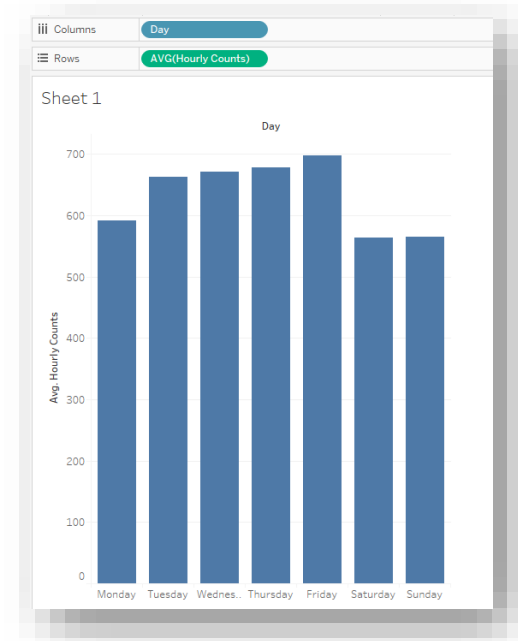
5. Now, plot the maximum number of pedestrians each day to see the difference between the **average** and **maximum** in the peak hours.

6. Drag and drop **Hourly Counts** to Rows and change the measure to **Maximum**.
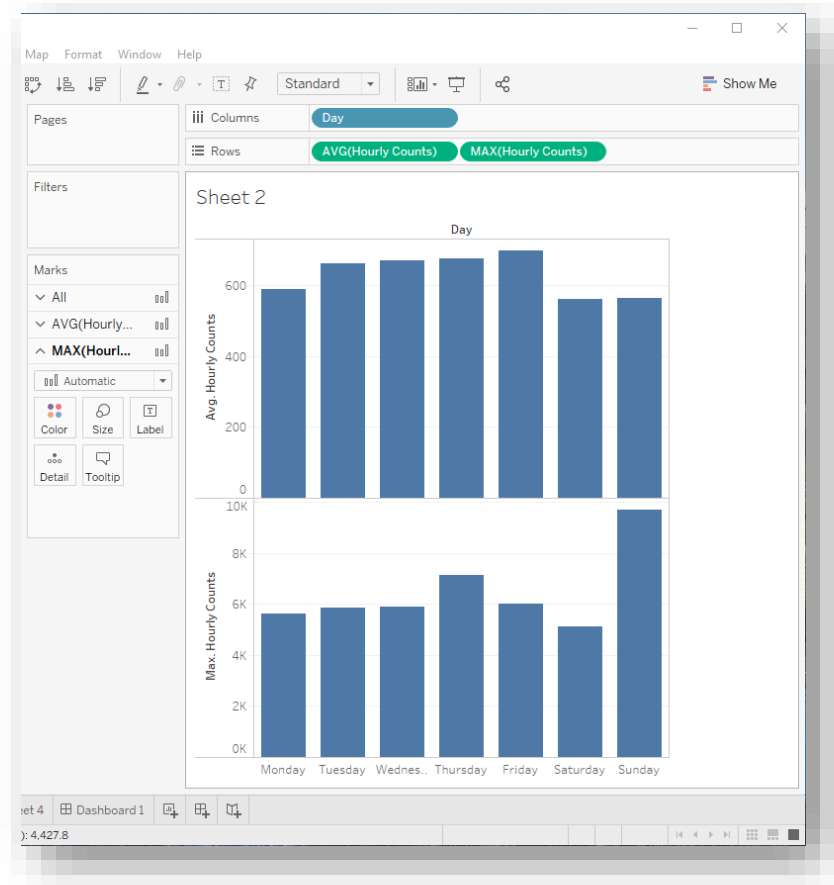
7. Now you have two bar charts with different measures. You can also see there are three sections in the **Marks** panel. This means you can change the **mark properties** of each bar chart separately. For instance, you can have a **line chart** for the maximum. Let's do it and change the **Max** mark to **line**.
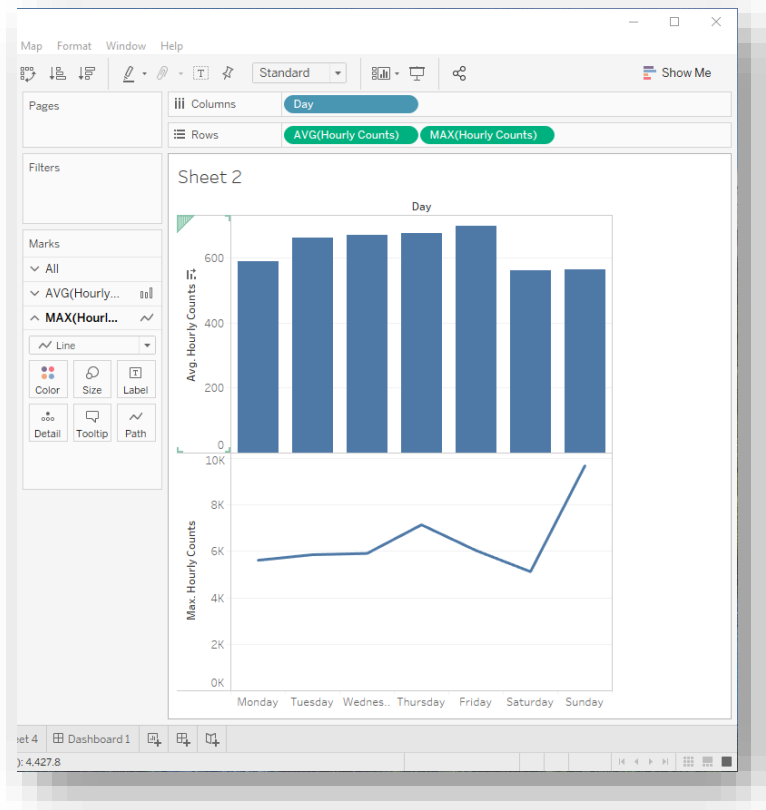
8. You could also change the **colour** of the line.
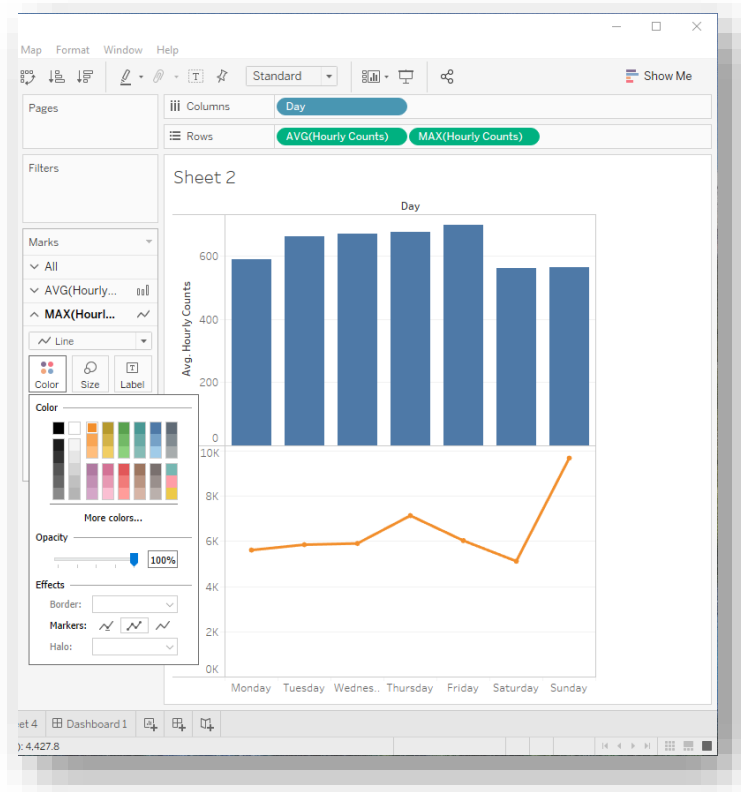
9. That's looking good. Now, recall the **"least ink" principle**. You would want to <u>minimise</u> the **amount of ink** used in your visualisation and <u>maximise</u> the **amount of information**. In this case, you can see that both charts share a common variable: **Hourly Counts**. It obviously makes sense to combine them into a single visualisation with a **shared axis**.

10. In Tableau, we can do it by utilising **dual axis**. First, right click on **Max. Hourly Counts** axis (the **line chart**). You will see an option called **Dual Axis**.
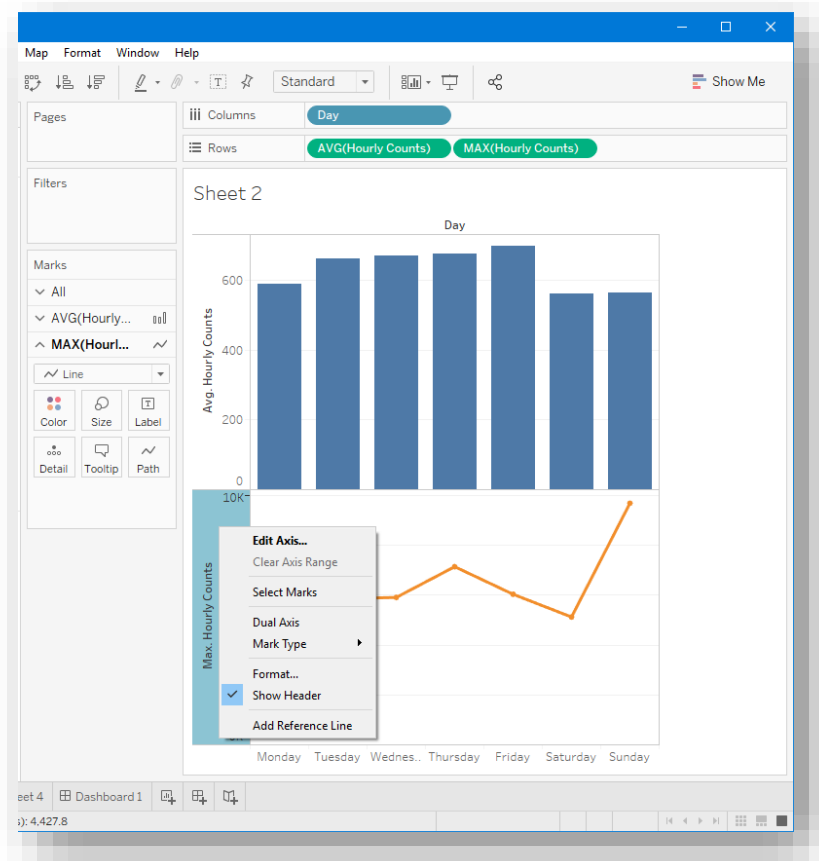
11. You can also see the dual axis option by clicking the **Max(Hourly Counts)** in the Rows.
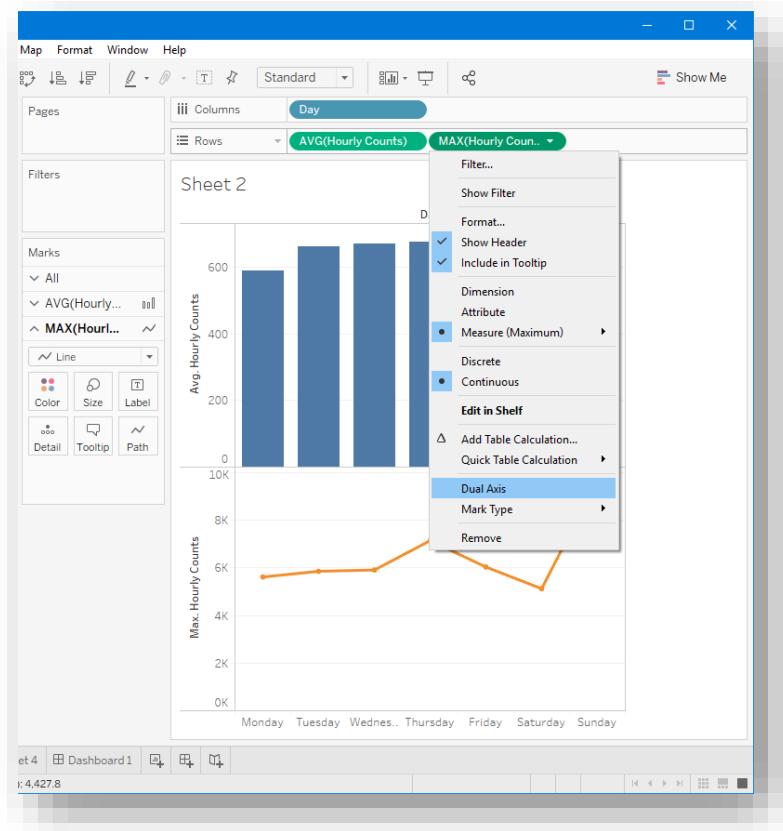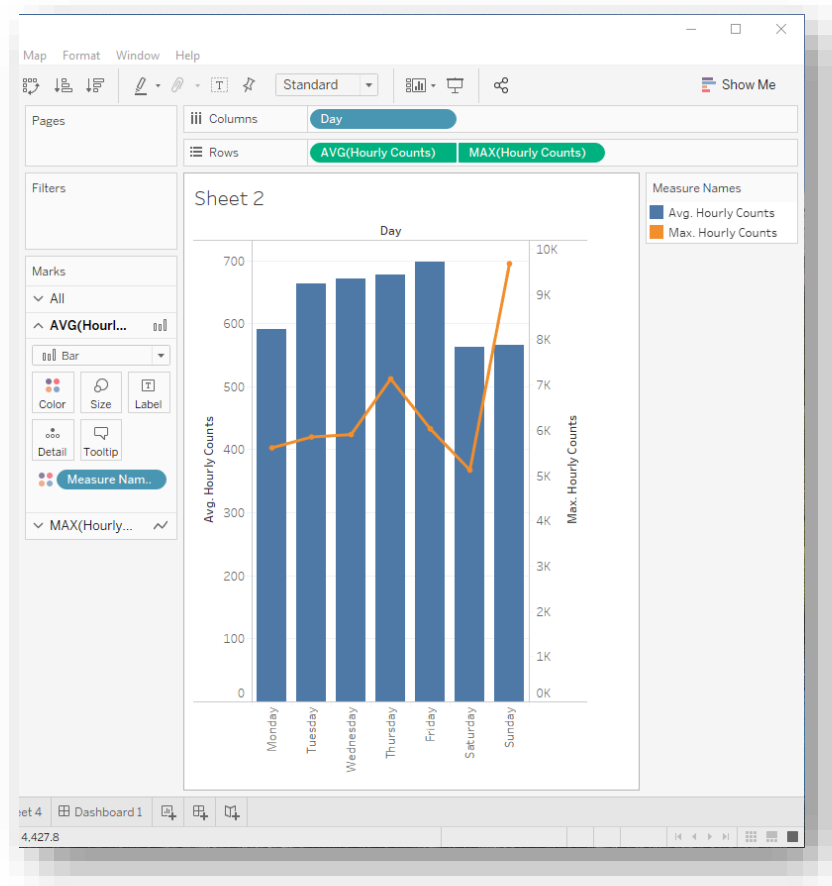

12. Once you've clicked on the **Dual Axis**, you will have a **composite chart** showing both the **Average** and **Maximum** number of pedestrians each day. You will also notice that the bar chart has turned into a **scatter plot**.

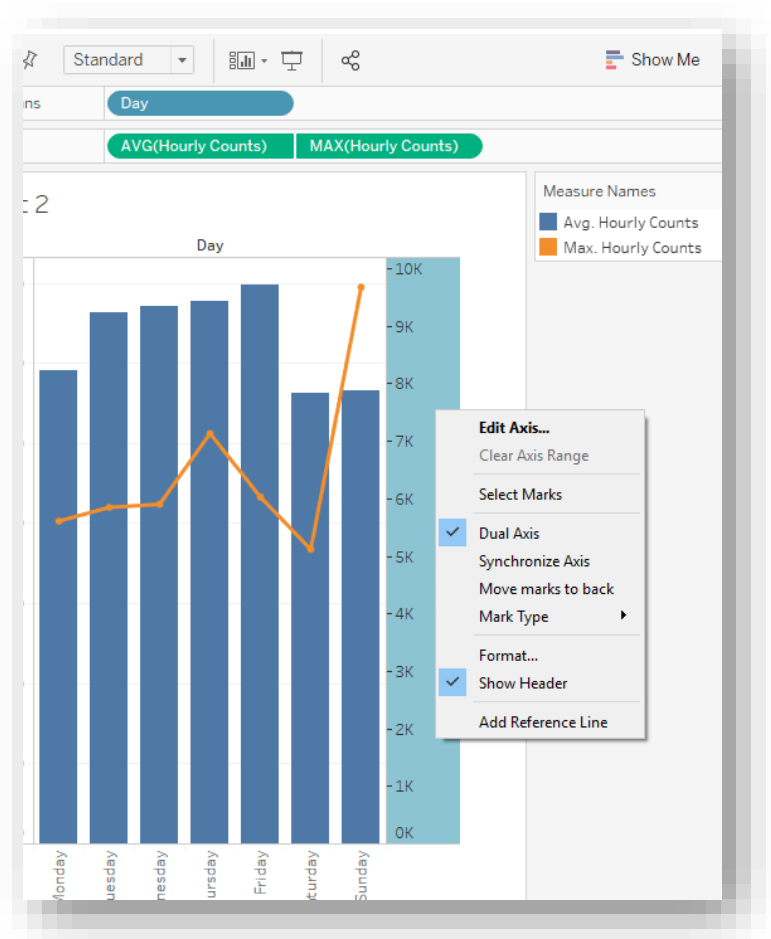| Question |
| --- |
| *Why do you think Tableau changed the bar chart into a scatter plot?* |

13. Let's revert to the bar chart. Change the **AVG(Hourly Counts)** mark to **bar**.


14. Nicely done! We now have a **composite chart**. But wait a minute, don't you see any problem in this chart?

Question
Before proceeding, discuss the problem with your peers and think of ideas on how to fix it!

15. If you find that the problem is <u>the maximum values being below the average values</u>, you are right. You can tell the difference from the marks used, but this is not a good way of visualising it. You might want to **synchronise** the axis and have a **single shared axis**. To do this, right click on the axis, and click **Synchronize Axis**.
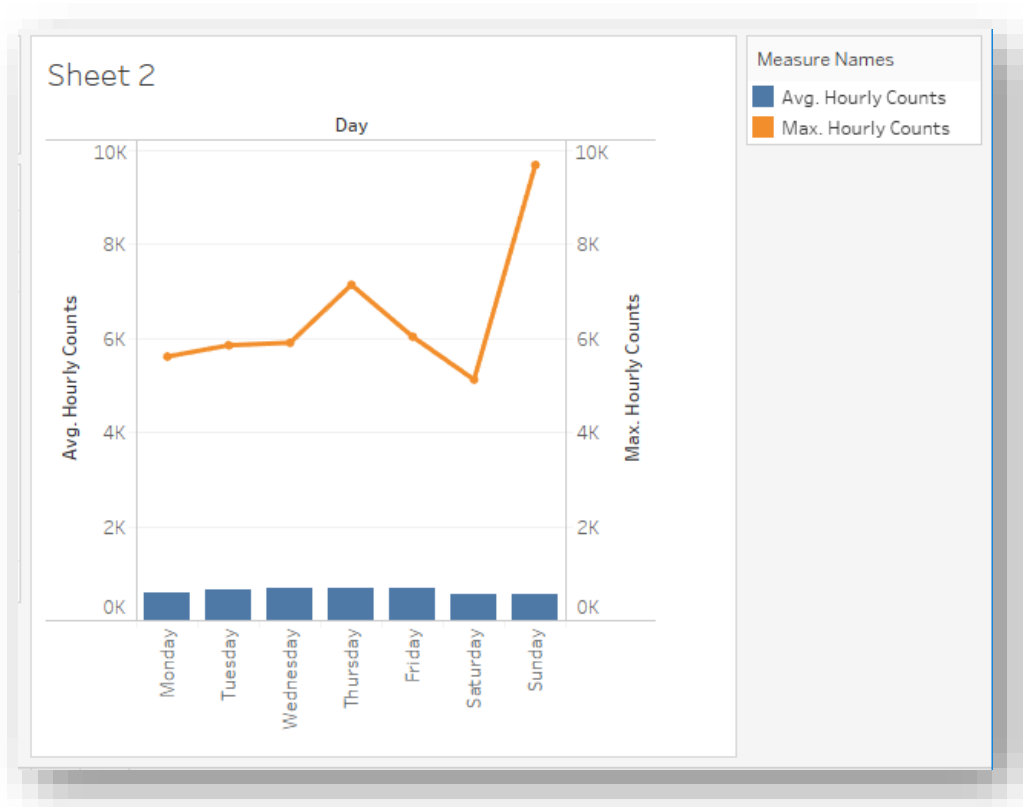
16. Now our visualisation is not lying. Remember, it is important to **synchronise** the axis when you have **multiple measures** sharing a **common axis** (in this case, they are sharing Day axis). This example maybe not the ideal one since it is very hard to see differences in Avg. Hourly Counts. Try to think about it if you decide to use dual axis in your visualisation project.

4. Tableau: Dual Axis and Simple Label Exercise

In this activity, you need to apply the skills you have learned from the previous activity to create a similar visualisation (with average and maximum) using the **Hour** dimension. It will be used later to see the distribution of the number of pedestrians among different hours within a day.

Create a new sheet and go ahead. Your visualisation should look like this:
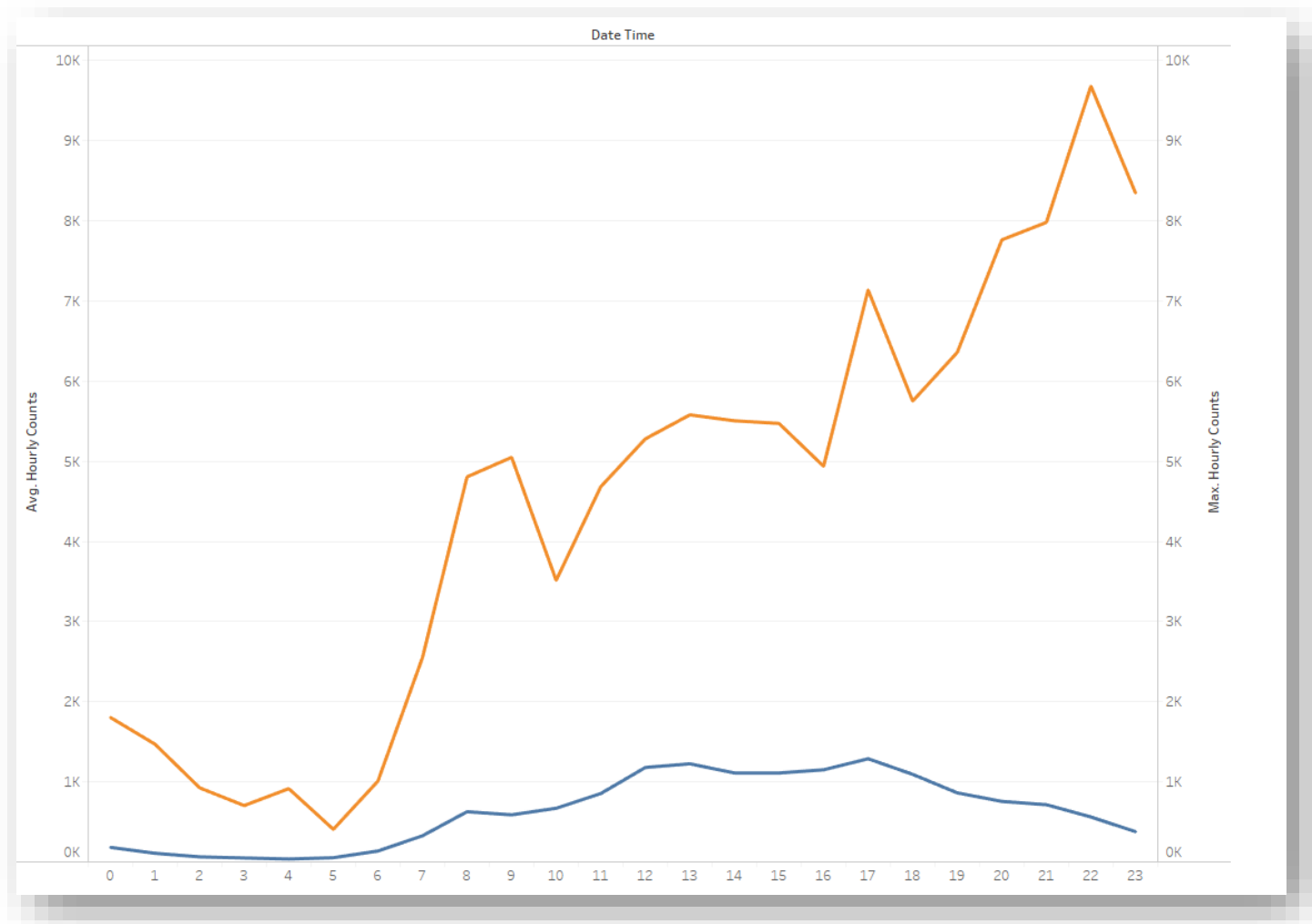
Note that the x-axis contains the **Hours** dimension while the y-axis illustrates both the **average** and the **maximum** number of pedestrians per hour. The hour dimension can be derived from **Date Time** dimension.

As a final touch, lets reduce the use of colour in this visualisation by using Label instead of Colour.

1. First, change **the Measure Names** of **All** marks from **Color** to **Label**.
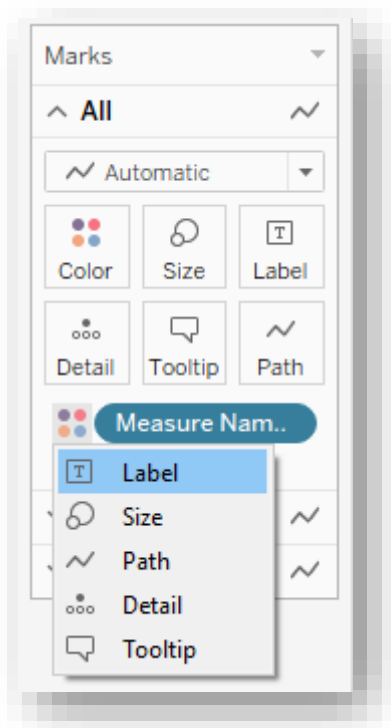
2. Then click the **Label** icon to show label options. Check **Show mark labels**. Set **Marks to Label** option to **Line Ends**. Most of us read from left to right, so it might be a good idea to only check **Label start of line** option to increase the data-ink ratio.
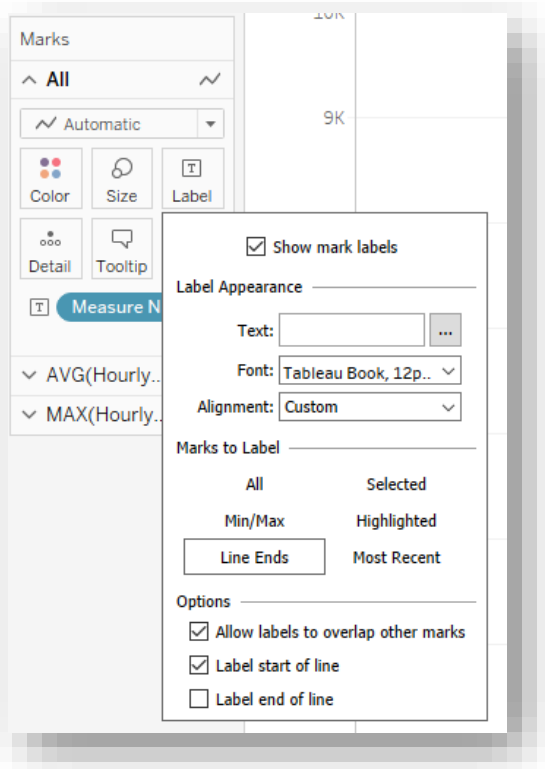
3. We could also increase the font size and weight as well as matching the font colour to the line colour. Click Font drop-down menu to adjust those options.
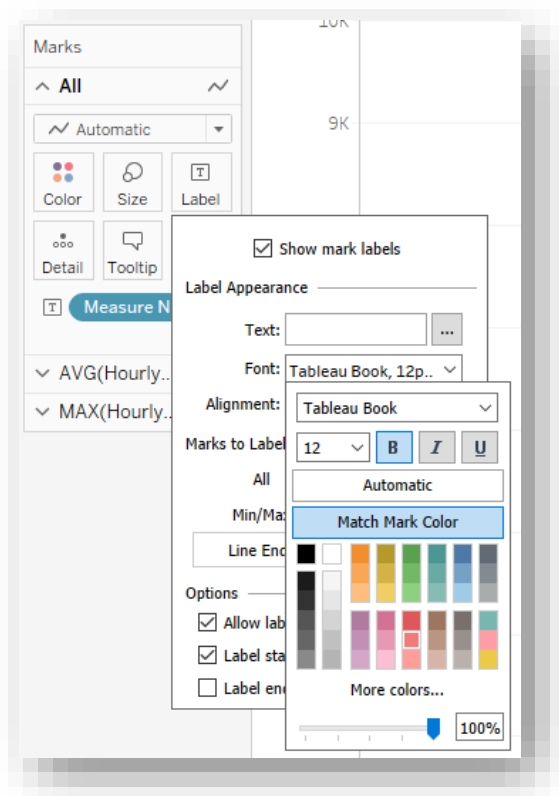
4. Your final line chart should look like this.