# 7PAM2000 Applied Data Science 1

## Assignment 1: Visualisation

Link to repository: [ads1\_assignment1\_repository](https://github.com/21062872/ads1_assignment1.git)

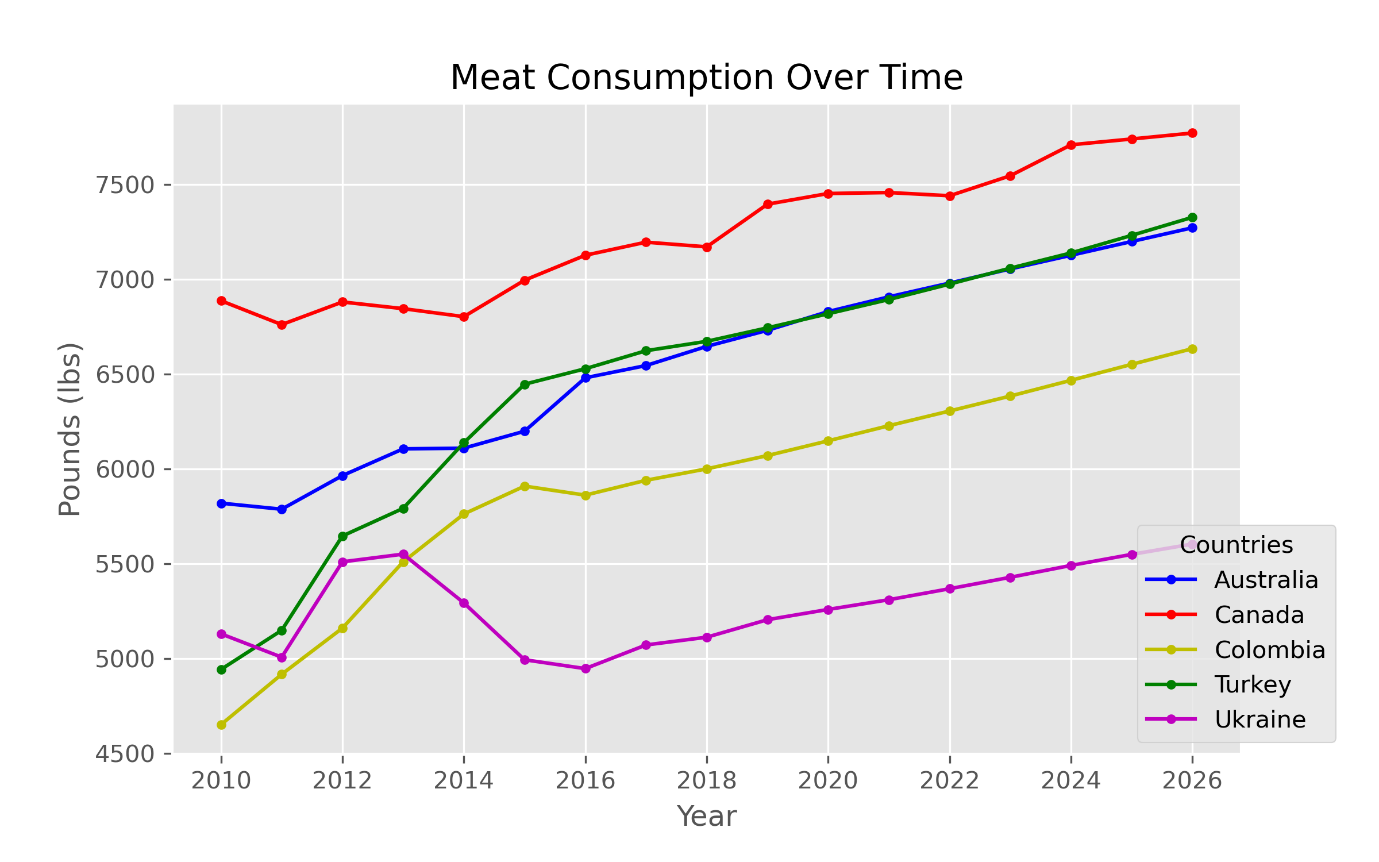
(Repository contains: data sources(folder: \data sources), plots(folder: \plots) and python files(folder: \python files, assignment answers (this word doc) and Turnitin report)

1. Q: **Produce a line plot showing multiple lines with proper labels and legend. Describe  
   what conclusions you can draw from this plot.**

A:

Line plot:

Data Source: <https://www.kaggle.com/datasets/vagifa/meatconsumption>



Conclusion:

This line plot depicts the meat consumption over a time period (from 2010) in Australia, Canada, Colombia, Turkey and Ukraine. Line graph is useful to track changes over a time period. Furthermore, line plots are used to determine a trend clearly and it can help to make predictions more accurately.

This chart shows, almost every country follows a similar pattern and across the years total meat consumption in these countries are steadily increased and expecting to reach the top values in year 2026. Also, by looking at this graph it is clear that Canada’s meat consumption is significantly higher than any other country. While lowest recorded meat consumption in Colombia is in year 2010, it has drastically increased over the years.

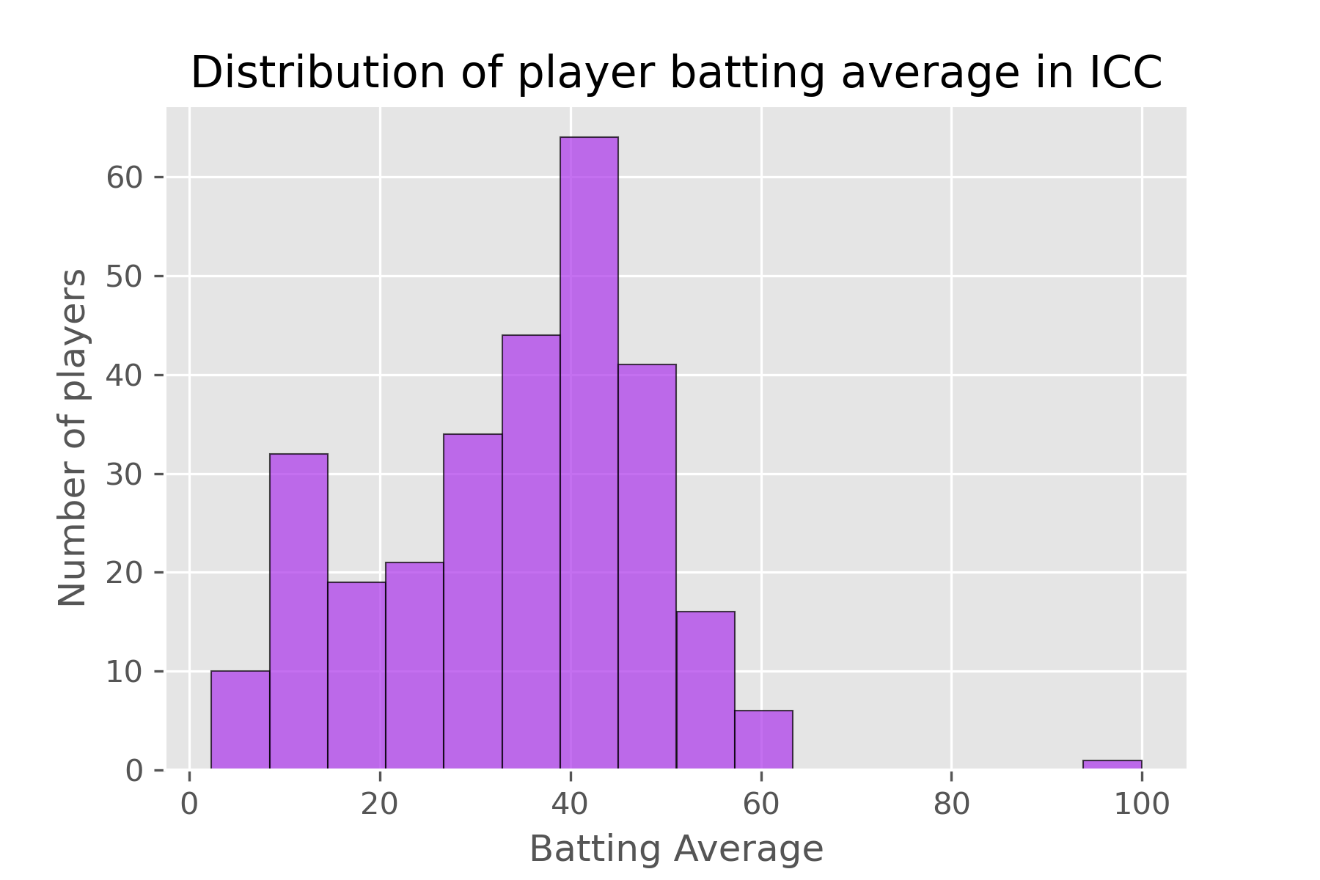
2. Q: **Produce graphs using two other visualisation methods. Explain why you picked this  
type of graph and describe what conclusions you can draw**

A:

1. Histogram:

Data source: <https://www.kaggle.com/datasets/veeralakrishna/icc-test-cricket-runs>

References used: <http://www.jtrive.com/determining-histogram-bin-width-using-the-freedman-diaconis-rule.html>



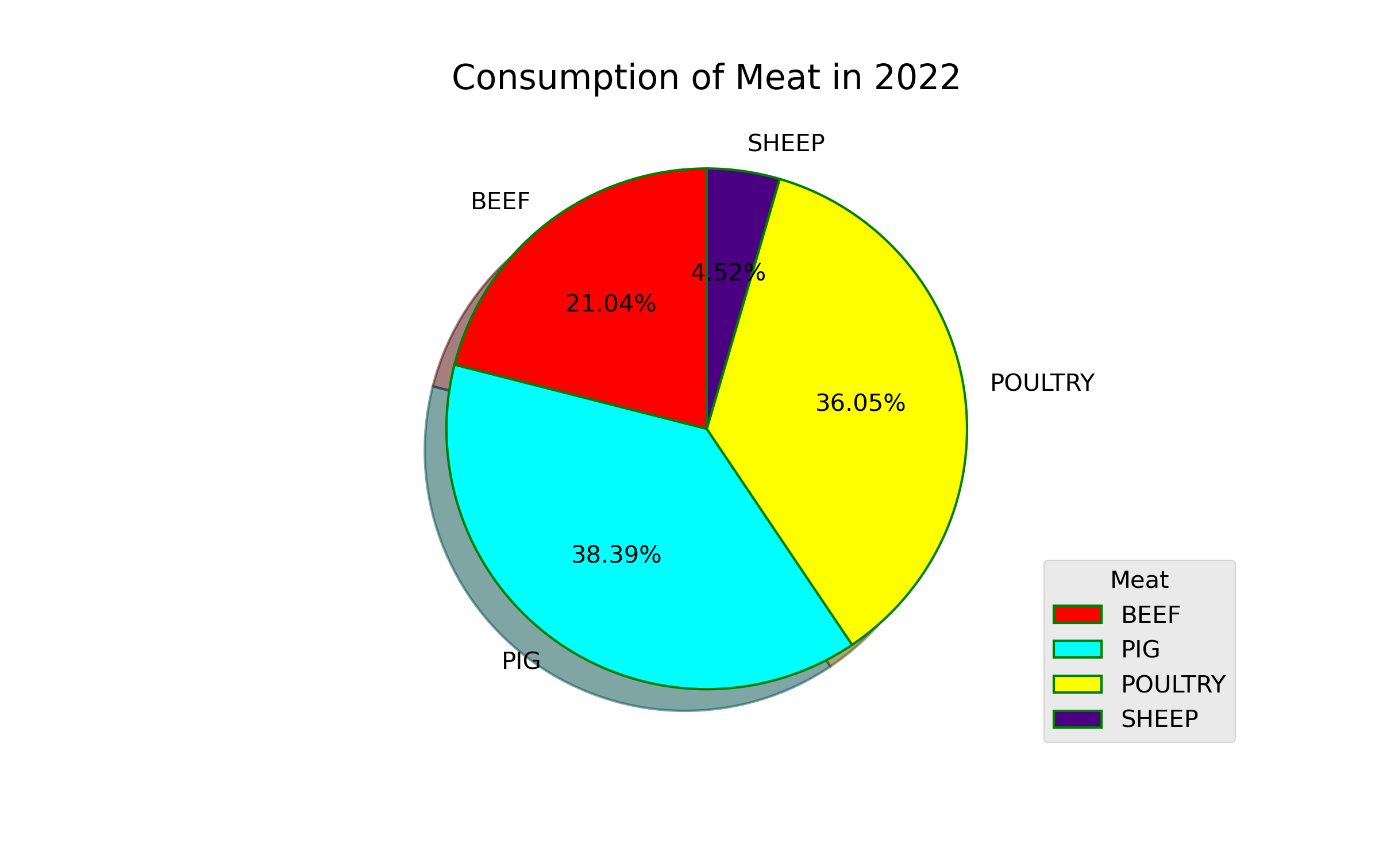
Description:

This histogram shows the distribution of player batting average for batters who has played at least 50 matches. I chose a histogram because it is a good and effective way to illustrate a distribution of the recorded data in a convenient form. This helps to understand the shape of the data distribution. This histogram is following a bell-shaped patten with one single peak in the middle of the distribution.

This histogram shows that, the majority of the player’s batting average is in 35-50 range and the highest number of players has an average within 40-45 range. While about 10 players have the lowest average, only one player has a recorded average within 95-100 range, which is exceptional.

1. Pie chart:

Data source: <https://www.kaggle.com/datasets/vagifa/meatconsumption>



Description:

This pie-chart illustrates shares of meat consumption across various meat types such as pig, beef, sheep and poultry in year 2022. I have chosen a pie chart, because it clearly differentiates each potion and helps to compare among various data segregations. Pie chart is also providing a way to illustrate data in very clear and easily understandable way.

According to the above pie chart, Pig is consumed the most and poultry has the second highest consumption level which is slightly lower than pig. The smallest consumed meat group is sheep in 2022. Beef consumption is almost 50% less than pig consumption and almost 5 times higher than sheep consumption.