Data Science

Lab Session 1

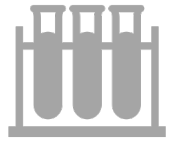
Lab Steps

Step 1 – Form groups and sit in your groups (2-3 members)

Step 2 – Send me the list of your members

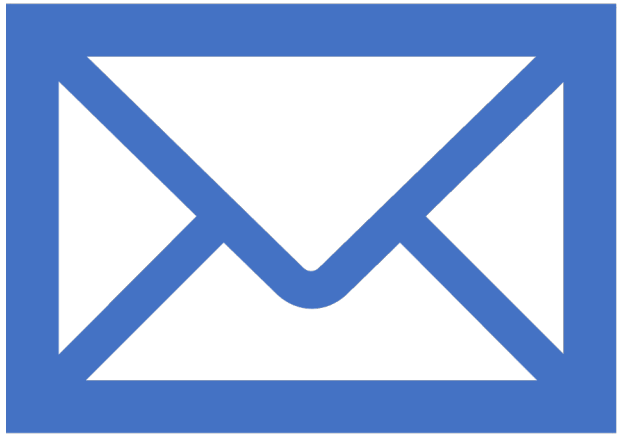
Step 3 – Have Google Colab up and running

Step 4 – Have you watched videos of this week?



Step 5 – Lab assignment for today

Form groups and sit in your groups (2-3 members)



Send me the list of your members!

Have Google Colab up and running

Have you watched videos of this week?

- Webscraping
  - Video: https://www.youtube.com/watch?v=_2kFa52pI6Q
  - https://github.com/sepinouda/Intro_to_Data_Science/blob/main/Webscraping/Webscraping.ipynb

- EDA and Data Visualization
- Video: https://www.youtube.com/watch?v=Zji-7tAfvEg&t=4s
  - https://github.com/sepinouda/Intro_to_Data_Science/tree/main/Lecture%203/EDA
  - https://github.com/sepinouda/Intro_to_Data_Science/tree/main/Lecture%202/Data%20Visualisation

Lab assignment

Step1 – Webscraping

- Scrape the data of the first 30 pages of the website https://www.skinnytaste.com

Step 2 – Filter interesting data

- Name of the food
- Image of the food
- Calories
- Personal Points
- Summary
- The recepie Key (Could be found on the website)

Step 3 – Visualise

- Use the appropriate visualisation method to provide information on
- Calories distribution
- Point distribution
- Recepie key distribution

Step 1 – Webscraping

Finding the solution – 20 min

Team presentation – 10 min

Step 2 – Filter interesting data

Finding the solution – 20 min

Team presentation – 10 min

Step 3 – Visualise

Finding the solution – 20 min

Team presentation – 10 min


What is next?

Step 4 – Interaction with the user

Your code should be able to perform the following task:

- Input a calorie range
- Input a point range
- Output the first 5 foods sorted based on calories **and points**, with their image and their summary.