Machine Medicine Technologies Ltd
Analytics Data Challenge

Instructions
- o Aim to spend a <u>maximum</u> of 3 hours working on the challenge, this is enough time for us to get a good gauge of your skills.

- o Provide answers in a Python Jupyter notebook, using markdown cells to add any necessary comments. We will run the notebook ourselves, so it should not reference any files other than the dataset provided.

- o Email answers to gareth@machinemedicine.com and cc joanna@machinemedicine.com

1. **Body Part Visualisation**
   The file df.csv contains the length in pixels of body parts of patients from several videos, and several categorical variables.
   - Write a function that, given two body parts (and any other arguments you feel appropriate), returns a figure plotting the distribution of both body parts on a single axis.
   - This function should also computer relevant summary statistics, perform an appropriate statistical test to compare these two distributions, and present these values.

2. **Body Part Tabulation**
   The file df.csv contains the length in pixels of body parts of patients from several videos, and several categorical variables.
   - Write a function that, given either one or two categorical variables, and one body part, (and any other arguments you feel appropriate), creates a tabular summary of that body parts.

3. **Activity Detection Function**
   The file ts.csv contains the distance between two body parts of a patient during a video.
   For less than half of the video the patient is performing an action.
   - Write a function that, given this CSV, (and any other arguments you feel appropriate), returns when the activity occurred (i.e. two numbers – the frame when it starts and the frame when it ends)