*Prior to the course: Participants are requested to install R (*[*https://cran.r-project.org/*](https://cran.r-project.org/)*) and RStudio (*[*https://posit.co/download/rstudio-desktop/*](https://posit.co/download/rstudio-desktop/) *)* ***before 5th July****. The instructors will be available from 9:00 on Wednesday 5th July to sort out any installation issues.*

**Thursday 4th July, AM**

9:00-9:30 Software installation issues troubleshooting

**Session I: 9:30-10:15**

* Introductions
* Course structure
* Goals
* Motivation – NYC datasets: research questions and hypotheses

**Session II: 10:30-11:15**

* Reading and summarising dataframes (NYC Squirrel)
* Read/import datasets
* Exploratory data analysis (NYC Squirrel)

**Session III: 11:30-12:30**

* Research question 1 (basic R)
* Subsetting/filtering

**Lunch break and library tour: 12:30-14:30**

**Thursday 4th July, PM**

**Session IV: 14:30-15:30**

* Installing packages
* Tidyverse introduction

**Session V: 15:45-16:30**

* Research question 2 exercise

**Session VI: 16:45-17:30**

* Data structures (atomic vectors, data frames, lists, matrices)
* Research questions

Day 2

**Monday 8th July, AM**

**Session I: 9:30-10:15**

* Introduction to new dataset: superheroes
* Work on your own: data exploration
* Motivation

**Session II: 10:30-11:30**

* Handling missing values
* Basic statistics

**Lunch break & supervisor meetings: 11:30 – 14:00**

**Monday 8th July, PM**

**Session IV: 14:00-15:00**

* Plotting with base R
* Histograms, boxplots
* Export graphs
* Exercise
* Plotting with ggplot2
* Introduction/installation
* Export graphs
* Exercise

**Session V: 15:15-17:30**

* Exercise & discussion

Day 3

**Monday 15th July, AM**

**Session I: 9:30-10:15**

* Conditionals
* Loops

**Session II: 10:30-11:30**

* Exercise & discussion

**Session III: 11:45-12:45**

* Functions

**Lunch break: 13:00-14:00**

**Monday 15th July, PM**

Day 4

**Monday 22nd July, AM**

**Session IV: 14:00-15:15**

* Reproducible data analysis (commenting, Rmarkdown)
* Capstone project, work in groups
* Introduction to datasets
* FitBit (few files)
* Diabetes
* Best songs on Spotify 2000-2023
* US births
* Glass door gender gap
* Life expectancy (WHO)
* Student mental health
* Bladder Cancer Recurrence Dataset
* UFO sightings

**Session V: 15:30-17:30**

* Presentation of results in groups
* Work through solutions
* Additional resources