**TU Dublin City Campus**

**MSc in Computing (Data Analytics) – Probability and Statistical Inference**

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**Week 3 Practical**

**Correlation**

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

1. Webinar on R notebooks with markdown <https://rstudio.com/resources/webinars/getting-started-with-r-markdown/>
2. R markdown cheatsheet - <https://www.rstudio.com/wp-content/uploads/2015/02/rmarkdown-cheatsheet.pdf>

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Preliminaries**

Download the zip file from Brightspace that contains the datafiles we are using this week (survey.dat and the codebook for the survey that was used to generate the data – this includes a description of the variables - Note all the Total variable are calculated to be scale variables).

Unzip to a location that makes sense for you.

Make sure you can load the dataset – either by saving the dataset to your working directory or setting your working directory to where the dataset has been downloaded.

**NOTE**: this zip file includes a code book which outlines the variable names given for each concept. For this lab you are interested in the variables Total scale scores (e.g. total optimism, total self-esteem etc)

Step 1:

Launch R Studio and open the .rmd file for the lecture.

Review the commands.

Step 2:

Try to write the code to investigate the following relationships using a Pearson correlation test (create scatterplots first)

* Relationship between Total Positive Affect and Total Perceived Stress
* Relationship between Total Life Satisfaction and Total Perceived Control of Internal States.
* Relationship between Total Mastery and Total Optimism.

Step 3:

* Investigate Spearman and Kendall for the above pairs of variables