Data Science with R - Course Description and pre-requisites

This course gives the essential knowledge to get started with a data science project in R using the Tidyverse. We learn to tell stories with data using the scientific data analysis workflow and analytic tools based on the R system.
In this two-day course we learn about the data science workflow

Reflection - Collection - Preparation - Analysis - Reporting

and how the Tidyverse, a collection of R packages specifically developed for contemporary data science, assists at each stage of the flow in a standardised, coherent and reproducible way.
The focus of the course is on the reflection, collection and preparation stages of the data science process. We learn about importing data presented in almost any format into R, a standard concept of tidy data and how to transform messy data sets into tidy ones, and how to explore a data set, using techniques such as visualisation and other tools, so that it is ready for analysis. The approach is hands-on using case-studies.

At the end of this course participants will be able to:

- Understand the scientific approach to the data analysis workflow and why and how R contributes to the process.
- Be able to import data into R in different text formats, flat files, excel, SAS, STATA and SPSS files as well as data from the web.
- Know the concept of tidy data, identify messy features in a data set and tidy it ready for analysis.
- Use simple exploratory analysis, including visualisation, to understand the data structure and some information it contains and also to detect, be aware of, and possibly correct, data anomalies.
- Be able to create a basic report of data analysis using a R notebook.
- Work with the Tidyverse packages readr, tidyr, dplyr, stringr, ggplot2, forcats, lubridate, etc.

Intended audience
Data analysis practitioners working in Government who use other systems for their work with an interest in migrating to R. If you have never used R before or have very limited experience, please follow the recommendations below in pre-requisites.

Pre-requisites
Participants must be familiar with the basic R syntax, object types and be able to use and be familiar with RStudio. Please, print this base R cheat sheet. You should be fairly comfortable with the functions in this cheatsheet except perhaps what is listed in the third column of the sheets.
RStudio has some free online learning resources and links that give an introduction to R and RStudio.
https://education.rstudio.com/learn/beginner/