



Swansea University  
Prifysgol Abertawe

# Introduction to



**By: Fatemeh Torabi**



Improving our health  
through data science



Medical School  
Ysgol Feddygaeth

Population Data Science  
at Swansea University

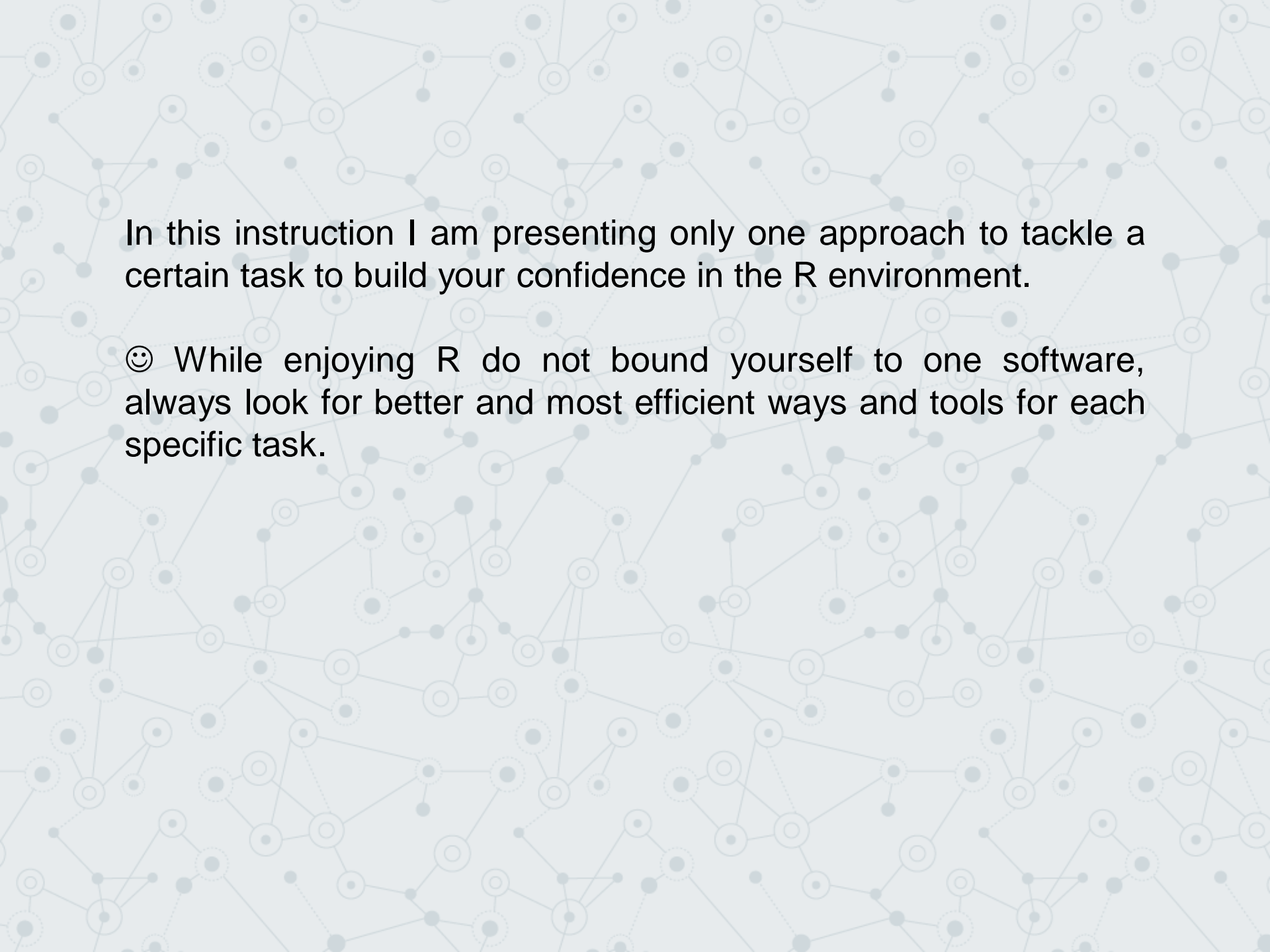


**Powerful statistical software**

**Free & open source**

**With over 9000+ packages**

**Great community**

The background of the slide features a complex, light blue network pattern. It consists of numerous small circles, some of which are solid and others are hollow, connected by thin, light blue lines. These lines form a web-like structure that fills the entire background, creating a sense of interconnectedness and data flow.

In this instruction I am presenting only one approach to tackle a certain task to build your confidence in the R environment.

☺ While enjoying R do not bound yourself to one software, always look for better and most efficient ways and tools for each specific task.



“

## Learning Outcomes

At the end of the session you will be able to:

1. Use R as a tool
2. Import your required datasets into R
3. Conduct initial investigation on a dataset
4. Interpret outputs of descriptive statistics
5. Visualise variables in the dataset.
6. Conduct and interpret a simple statistical tests

# Getting started

## SOFTWARES IN USE

### ◎ Install R:

◎ Go to <https://cran.r-project.org/> and follow the link for your operating system.



CRAN  
[Mirrors](#)  
[What's new?](#)  
[Task Views](#)  
[Search](#)

[About R](#)  
[R Homepage](#)  
[The R Journal](#)

[Software](#)

## The Comprehensive R Archive Network

### Download and Install R

Precompiled binary distributions of the base system and contributed packages, **Windows** and **Mac** users most likely want one of these versions of R:

- [Download R for Linux](#)
- [Download R for MacOS X](#)
- [Download R for Windows](#)

R is part of many Linux distributions, you should check with your Linux package management system in addition to the link above.

### Source Code for all Platforms

Windows and Mac users most likely want to download the precompiled binaries listed in the upper box, not the source code. The sources have to be compiled before you can use

## Install RStudio:

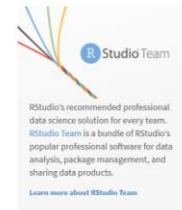
Go to <https://www.rstudio.com/products/rstudio/download/> and click on the installer link for your operating system.



### Choose Your Version

The RStudio IDE is a set of integrated tools designed to help you be more productive with R and Python. It includes a console, syntax-highlighting editor that supports direct code execution, and a variety of robust tools for plotting, viewing history, debugging and managing your workspace.

[LEARN MORE ABOUT RSTUDIO FEATURES](#)





## Example:

### Effect of mindfulness on task completion time


#### Aim:

To explore whether brief mindfulness meditation can improve attentional skills and reduce distraction on tasks of attention.

#### Design:

Experiment – time for completion of a distraction task (the Stroop) before and after a brief period of mindfulness meditation ( 3 min)

The **Stroop test** can be used to **measure** a person's selective attention capacity and skills to evaluate overall executive processing abilities.



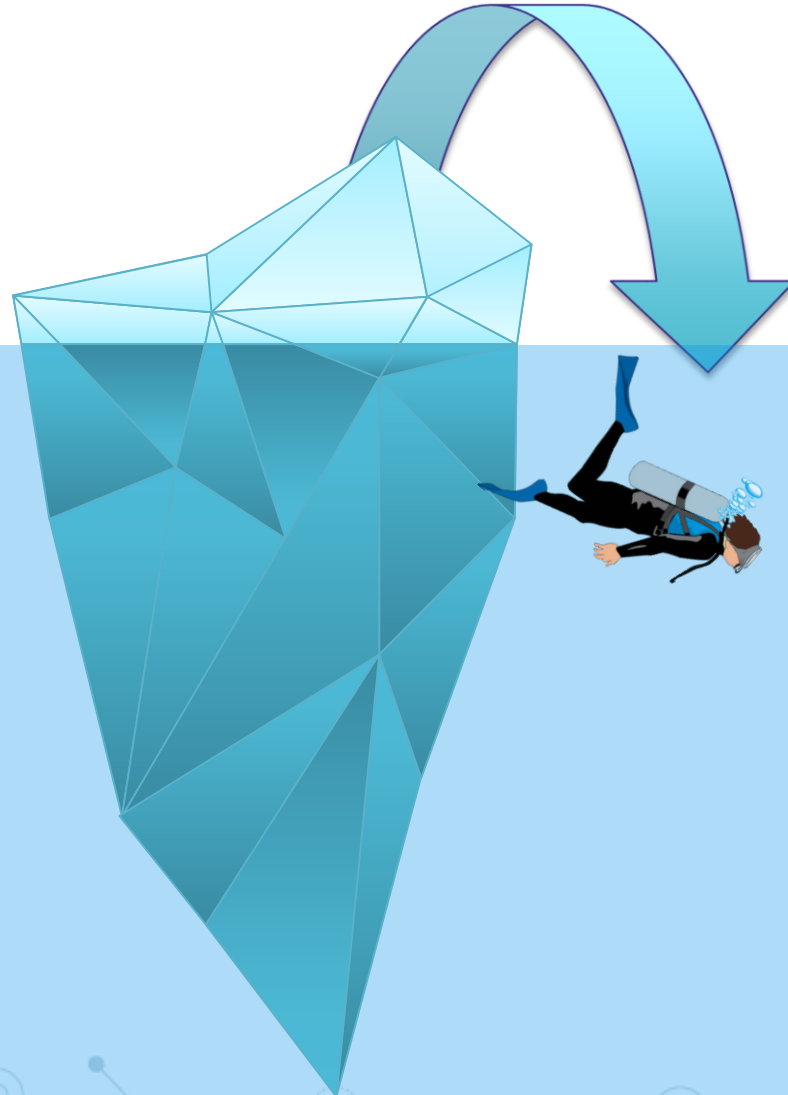


Pt no.	DOB	Age	Sex	Education	Occ	ST1	S1 incorrect	ST2	S2 incorrect
1	09/03/2000	19	male	Alevels	Student	20	0	15	0
2	15/10/1963	56	male	degree	retired	27.49	0	28	0
3	25/05/1962	57	female	Degree	GP	39.53	2	25.96	1
4	22/11/1997	22	male	Degree	Web developer	21.16	0	18.9	1
5	15/02/1999	20	female	Alevels	student	19	0	17.16	0
6	22/04/1965	54	Female	Graduate level - Qualified RGN & Midwife - (but without a degree)	Health Services Manager	42	2	37	1
7	31/12/1967	51	Male	Secondary school	Enforcement Officer	40	1	37	1
8	21/05/1994	25	Male	College	Head Chef	41	4	40	2
9	04/03/1999	20	Female	Secondary School	Deputy Manager (Hospitality)	48	3	45	1
10	08/06/1964	55	Female	College	Care Manager	49	3	46	2
11	10/01/1978	41	female	Degree	Teacher	33	0	32	0

# Exploratory Analysis on dataset

## Initial exploration:

- 1-Descriptive Statistics
- 2-Visualizing variables
- 3- Simple statistical tests



## Data Preparation

Data preparation  
Handling missing data  
And ...

## Data Analysis

Linear Models:  
Regression Analysis  
GLM  
More complex models  
Statistical Tests  
Interpreting results  
And  
...





# Let's move on to R

