University of Exeter

Introduction to Business Analytics BEM2031

Term2: 2023/24

University of Exeter

Dr Alison Harper

a.l.harper@exeter.ac.uk

Office hours: Tuesdays 3-5pm

Building:One, 2nd Floor, CSAM

https://bit.ly/Book_Alison_Harper Home_BEM2031_Teams

University of Exeter

Teaching team:

Dr Alison Harper
a.l.harper@exeter.ac.uk

Dr Han Lin
H.Lin2@exeter.ac.uk

University of Exeter

There will be
no workshops
on
Monday,
22 January

University of Exeter

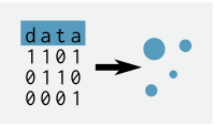
Module Overview

Peer-reviewed homework due:
16 February (formative)

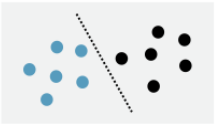
Week 01 - Data Analytic Thinking

Week 02 - Managing and Cleaning Data

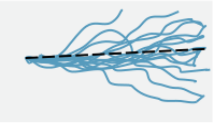
Week 03 - Data Visualization



Week 04 - Clusters and Similarity



Week 05 - Predictive Modeling



Week 06 - Reading Week

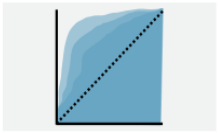
University of Exeter

Module Overview

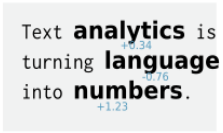
Assignment due:
15 March 3pm
(summative – 30%)

Final project due:
28 March 3pm
(summative – 70%)

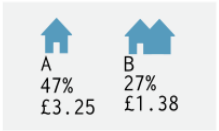
Week 07 - What is a good model?



Week 08 - Text Analytics



Week 09 - Data Driven Decisions



Week 10 - AI and Organizations

Week 11 - Project Week

University of Exeter

Module Structure

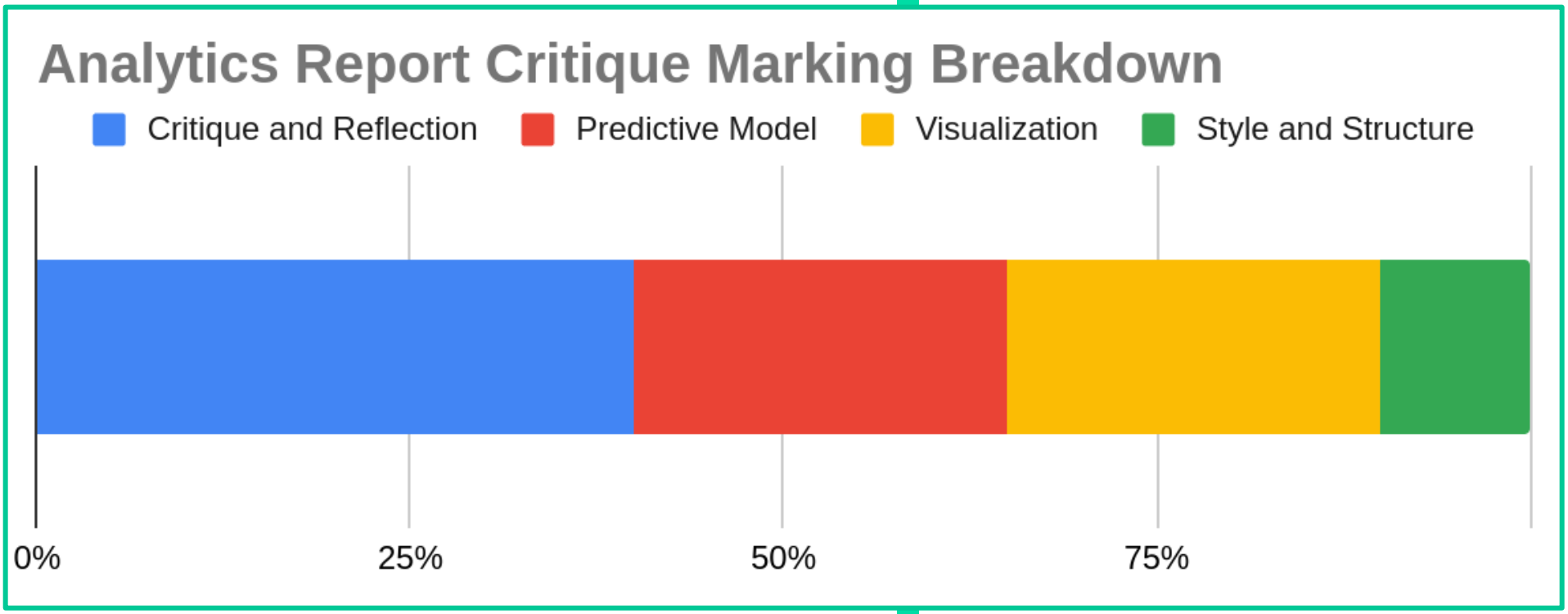
Week	Date	Topics	Deadlines
T2:01		Data analytic thinking	
T2:02		Managing and cleaning data	
T2:03		Data visualisation	
T2:04		Clustering and similarity	HOMEWORK
T2:05		Predictive modelling	
T2:06		READING WEEK	
T2:07		What is a good model?	
T2:08		Text analytics	
T2:09		Data-driven decisions	ASSIGNMENT
T2:10		Artificial intelligence and ethics	
T2:11			FINAL PROJECT

ASSIGNMENT CRITERIA (15 March):

Fully correct answers that complete the task in the expected manner will be given a high distinction of 8/10. For a full 10/10 I have left some room for innovation and personal exploration. Students who go above the expected, integrate a new package, attempt a new plot, try a new analysis, can be rewarded here.

Score	Description
0	The problem was not attempted.
2	The problem was attempted but largely incomplete or incorrect.
4	Concepts are understood, but not well explained in the context of the problem. Calculations yield the wrong answer due to minor or major errors. Plots are incorrectly generated.
6	The approach is generally correct. Calculations yield the wrong answer due to minor errors. Plots are roughly correct.
8	The solution is correct, well-documented, and the writing is clear. Reproducible code provides a correct step-by-step solution and is easy to follow. Plots are correct, detailed, and clearly explained.
10	The solutions are exceptional, clear, and creative. The solutions provided innovate and expand on existing knowledge.

FINAL REPORT CRITERIA
(28 March)



University of Exeter

Getting started

1. Read the handbook and ask questions
2. Download and install R and RStudio
3. Go to ELE and review/prep for Week 1:
   - Watch the videos
   - Access the textbook

This is not a programming class

University of Exeter

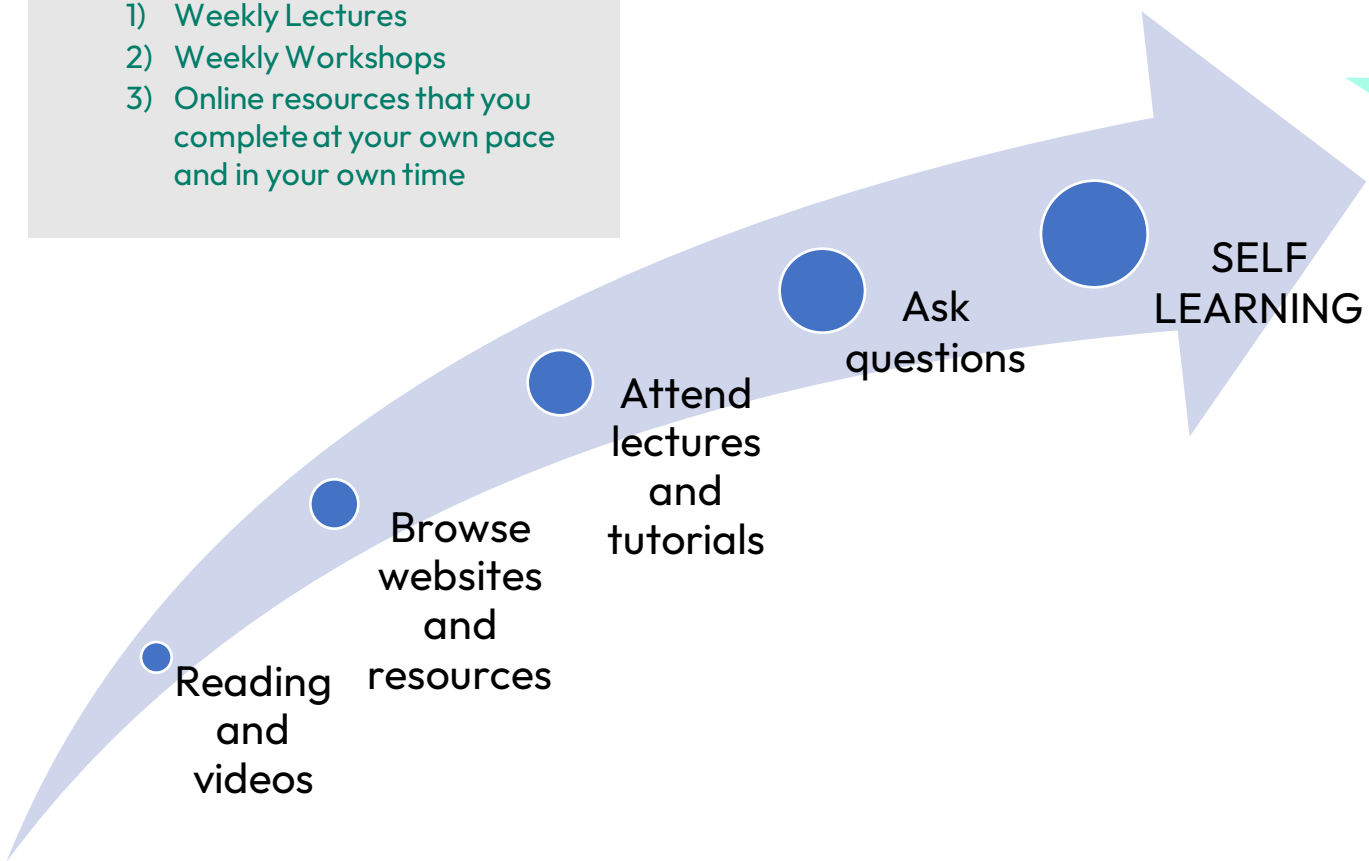
LEARNING OBJECTIVES

After this course you should be able to...

- Critically evaluate current approaches used for collection, management, communication and analysis of commercial, operational and sustainability data, and how this data is used to support decision-making.
- Apply *Design Thinking* techniques to the analysis of a specific business challenge and use these to identify required information flows.
- Use data visualisation techniques to share original content and insight with a general management audience
- Demonstrate familiarity with analytical tools available for the analysis of numerical and textual data and use these to find, derive and evaluate information.
- Discuss current developments and thinking in the information management industry, specifically around big data management, analytics, cloud, and visualisation techniques.

The module will be taught through a combination of:

1) Weekly Lectures
2) Weekly Workshops
3) Online resources that you complete at your own pace and in your own time



How to achieve the LEARNING OBJECTIVES

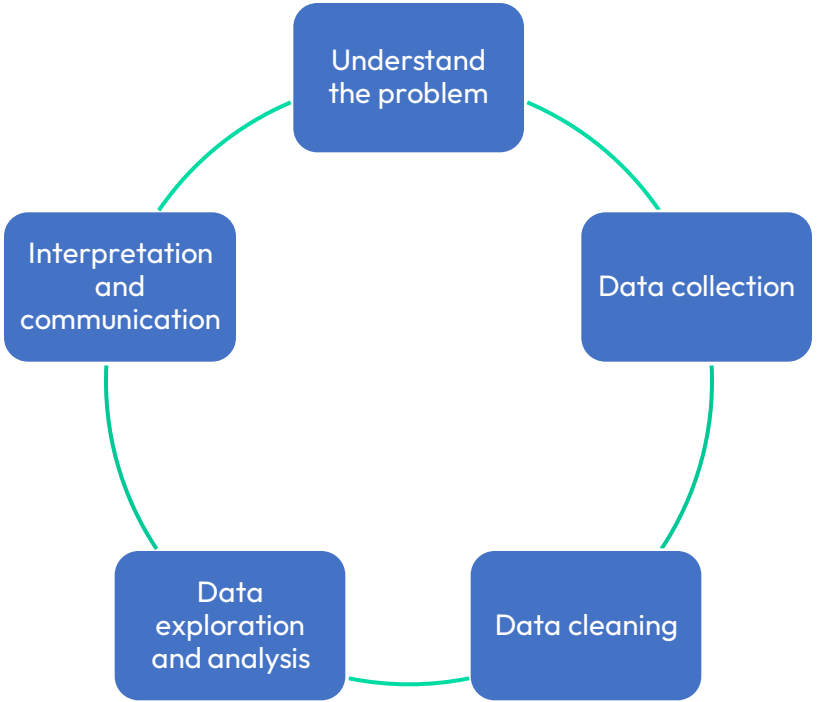
University of Exeter

Introduction to Business Analytics

University of Exeter

Data Analytics is the process of collecting, transforming, and analysing large volumes of data to gain information, knowledge and insights that inform decision-making and drive improvements.

It involves using various tools, techniques, and algorithms to examine data and extract valuable information. This, in turn, enables businesses and organisations to make data-driven decisions and thus gain a competitive edge.



University of Exeter

Types of data analytics

✓ **Descriptive analytics: What happened / is happening?**
Uses historical and real time data to describe the current or past situation, aggregate data into summaries, separate out patterns
e.g. visualisations, dashboards, data mining, clustering

✓ **Diagnostic analytics: Why did it happen?**
Predicts potential future outcomes based on past data
e.g. statistical forecasting, machine learning

✓ **Predictive analytics: What will happen?**
Predicts potential future outcomes based on past data
e.g. statistical forecasting, machine learning

✓ **Prescriptive analytics: What can we do about it?**
Offers potential or optimal solutions
e.g. simulation, optimisation, game theoretical approaches

Data-driven decision making

Using data to gather insights, identify patterns, and draw conclusions to guide decisions.

E.g., customer behaviour, market trends, operational efficiency, strategic choices for growth or success.

Competitive advantage

Data and effective data analysis capability is a competitive asset toward meeting customer needs.

E.g., pricing strategies, marketing and branding campaigns, safety and quality.

Personalisation and customer experience

Understanding individual preferences and customer behaviour to increase customer loyalty and satisfaction.

E.g., customer segmentation, predictive analysis, sentiment analysis, churn prediction.

Operational efficiency

Data analysis can be used to optimise internal processes, identify inefficiencies and bottlenecks, reduce costs.

E.g., pathway modelling, simulation, optimisation, real-time monitoring, predictive maintenance.

Predictive analytics

Predictive modelling and forecasting can be used to anticipate future trends to stay ahead of challenges and opportunities.

E.g., demand prediction, drivers of customer choice, features predicting churn.

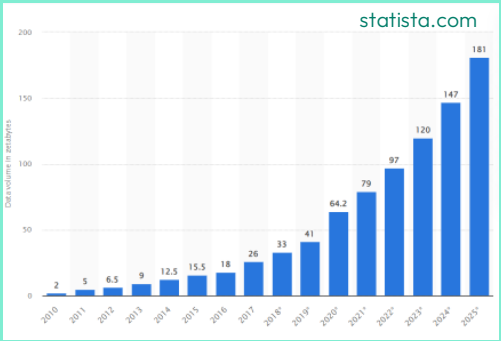
Internet-of-things (IoT) and Big Data

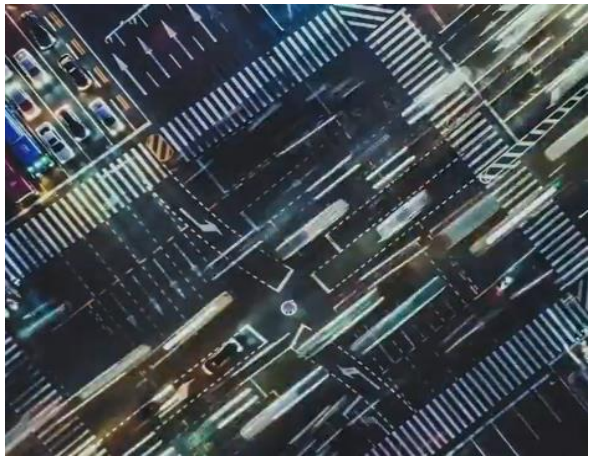
Using sensors and IoT devices connecting everyday objects and machines to the internet creates vast amount s of data that can be used to monitor, predict and optimise systems and processes.

University of Exeter

Future and Scope in Data Analytics

Volume of data/information created, captured, copied, and consumed worldwide from 2010 to 2020, with forecasts from 2021 to 2025





"Now that we're paying such close attention, we can see how data can also be inconclusive, misunderstood and even abused.

We sense now that data has a Big Data problem, opening the door to opportunists who manipulate and misrepresent data to promote their own agenda, undermining both public health as well as civil liberties. From the politicization of data, to the growing realization of data biases and lack of appropriate investment in data analysis..."

Forbes.com

University of Exeter

Week 1 tasks:

Read Chapters 1 and 2 of textbook and watch videos

1. Download and install R and RStudio
2. Work on introductory workshop for Week 1

Get in touch through office hours for any questions

University of Exeter

Week 2 tasks:

Read the materials

1. Prepare for workshop 2

Get in touch through office hours for any questions

DOWNLOAD

RStudio Desktop

Used by millions of people weekly, the RStudio integrated development environment (IDE) is a set of tools built to help you be more productive with R and Python.

Don't want to download or install anything? Get started with RStudio on Posit Cloud for free. If you're a professional data scientist looking to download RStudio and also need common enterprise features, don't hesitate to book a call with us.

1: Install R

RStudio requires R 3.3.0+. Choose a version of R that matches your computer's operating system.

DOWNLOAD AND INSTALL R

2: Install RStudio

DOWNLOAD RSTUDIO DESKTOP FOR WINDOWS

Size: 214.34 MB | SHA-256: FE62B784 | Version: 2023.09.1+494 | Released: 2023-10-17

wnload/rstudio-desktop/

posit PRODUCTS SOLUTIONS LEARN & SUPPORT EXPLORE MORE

DOWNLOAD

RStudio Desktop

Used by millions of people weekly, the RStudio integrated development environment (IDE) is a set of tools built to help you be more productive with R and Python.

Don't want to download or install anything? Get started with RStudio on Posit Cloud for free. If you're a professional data scientist looking to download RStudio and also need common enterprise features, don't hesitate to book a call with us.

1: Install R

RStudio requires R 3.3.0+. Choose a version of R that matches your computer's operating system.

DOWNLOAD AND INSTALL R

1

Download and Install R

Precompiled binary distributions of the base system and contributed p
of R:

- Download R for Linux (Debian, Fedora/Redhat, Ubuntu)
- Download R for macOS
- Download R for Windows

R is part of many Linux distributions, you should check with your Li

2

Subdirectories:

base	Binaries for base distribution. This is what you want to **install R for the first time**.
contrib	Binaries of contributed CRAN packages (for R >= 3.4.x).
old contrib	Binaries of contributed CRAN packages for outdated versions of R (for R < 3.4.x).
Rtools	Tools to build R and R packages. This is what you want to build your own packages on Windows, or to build R itself.

3

Download R-4.3.1 for Windows (79 megabytes, 64 bit)

README on the Windows binary distribution
New features in this version

DOWNLOAD

RStudio Desktop

Used by millions of people weekly, the RStudio integrated development environment (IDE) is a set of tools built to help you be more productive with R and Python.

Don't want to download or install anything? Get started with RStudio on Posit Cloud for free. If you're a professional data scientist looking to download RStudio and also need common enterprise features, don't hesitate to book a call with us.

1: Install R

RStudio requires R 3.3.0+. Choose a version of R that matches your computer's operating system.

DOWNLOAD AND INSTALL R

2: Install RStudio

DOWNLOAD RSTUDIO DESKTOP FOR WINDOWS

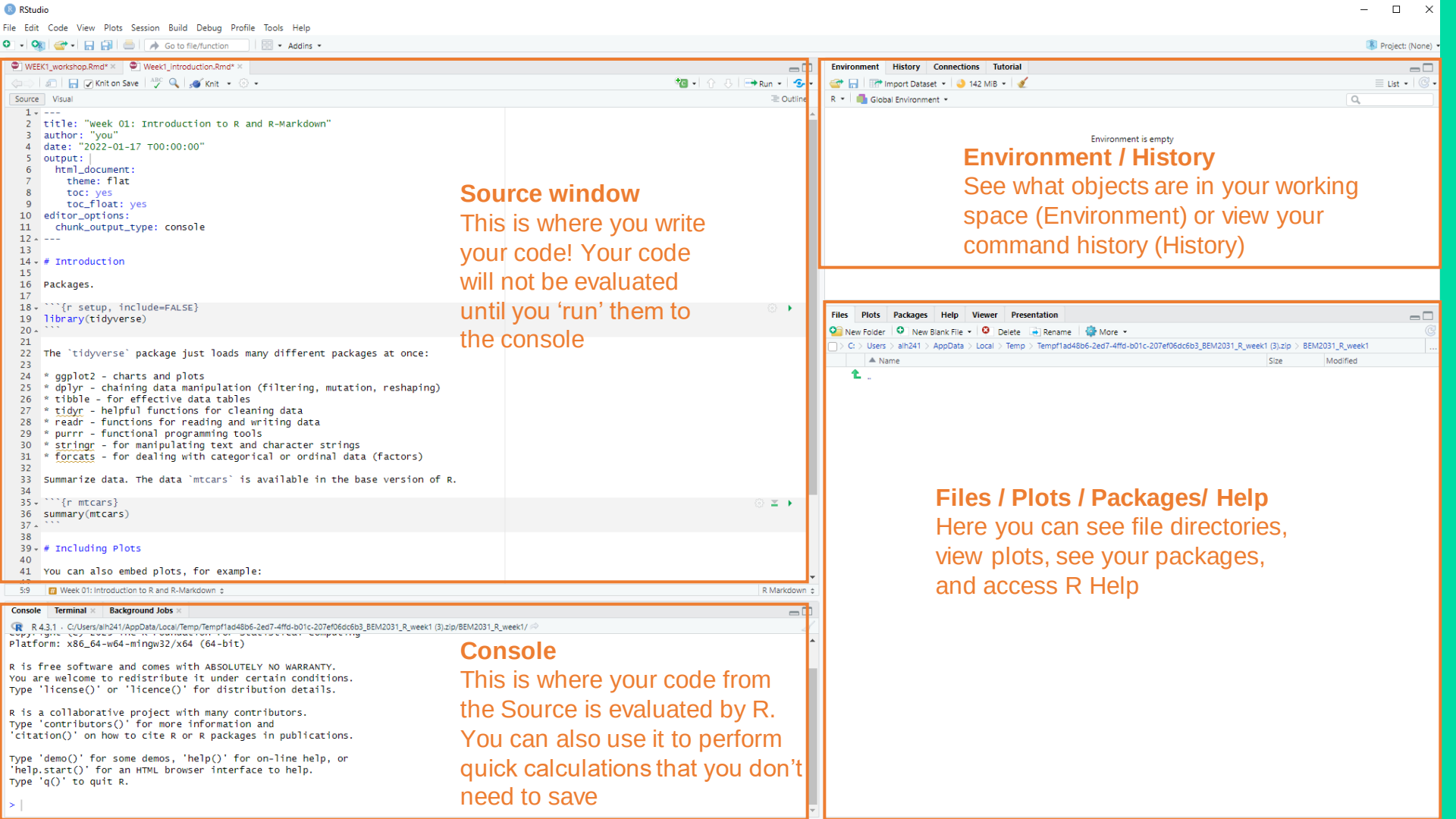
Size: 214.34 MB | SHA-256: FE62B784 | Version: 2023.09.1+494 | Released: 2023-10-17

University of Exeter

Download files for Workshop 1

Extract and save them somewhere sensible, for example:

Desktop > Y2_T2_Modules > BEM2031 > Workshops > Week 1


RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
WEEK1_workshop.Rmd*
Week1_introduction.Rmd*
Knit on Save
Knit
Run
Source
Visual
1 ---
2 title: "Week 01: Introduction to R and R-Markdown"
3 author: "you"
4 date: "2022-01-17 T00:00:00"
5 output:
6 html_document:
7 theme: flat
8 toc: yes
9 toc_float: yes
10 editor_options:
11 chunk_output_type: console
12 ---
14 # Introduction
16 Packages.
18 ```{r setup, include=FALSE}
19 library(tidyverse)
20 ```
22 The `tidyverse` package just loads many different packages at once:
24 * ggplot2 - charts and plots
25 * dplyr - chaining data manipulation (filtering, mutation, reshaping)
26 * tibble - for effective data tables
27 * tidyr - helpful functions for cleaning data
28 * readr - functions for reading and writing data
29 * purrr - functional programming tools
30 * stringr - for manipulating text and character strings
31 * forcats - for dealing with categorical or ordinal data (factors)
33 Summarize data. The data `mtcars` is available in the base version of R.
35 ```{r mtcars}
36 summary(mtcars)
37 ```
39 # Including Plots
41 You can also embed plots, for example:
5:9 Week 01: Introduction to R and R-Markdown
R Markdown
Source window
This is where you write your code! Your code will not be evaluated until you 'run' them to the console
Environment History Connections Tutorial
Import Dataset
142 MiB
List
R
Global Environment
Environment is empty
Environment / History
See what objects are in your working space (Environment) or view your command history (History)
Files Plots Packages Help Viewer Presentation
New Folder
New Blank File
Delete
Rename
More
C: Users alh241 AppData Local Temp Tempf1ad48b6-2ed7-4ffd-b01c-207ef06dc6b3_BEM2031_R_week1 (3).zip BEM2031_R_week1
Name
Size
Modified
Files / Plots / Packages/ Help
Here you can see file directories, view plots, see your packages, and access R Help
Console Terminal Background Jobs
R 4.3.1 · C:/Users/alh241/AppData/Local/Temp/Tempf1ad48b6-2ed7-4ffd-b01c-207ef06dc6b3_BEM2031_R_week1 (3).zip/BEM2031_R_week1/
Platform: x86_64-w64-mingw32/x64 (64-bit)
R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.
R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.
Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.
>
Console
This is where your code from the Source is evaluated by R. You can also use it to perform quick calculations that you don't need to save

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function | Addins

WEEK1_workshop.Rmd* | Week1_introduction.Rmd*

Knit on Save | Knit

Source | Visual

```
10  editor_options:
11    chunk_output_type: console
12  ---
13
14  # Introduction
15
16  Packages.
17
18  ```{r setup, include=FALSE}
19  library(tidyverse)
20  ```
21
22  The `tidyverse` package just loads many different packages at once
23
24  * ggplot2 - charts and plots
25  * dplyr - chaining data manipulation (filtering, mutation, reshapi
26  * tibble - for effective data tables
27  * tidyr - helpful functions for cleaning data
28  * readr - functions for reading and writing data
29  * purrr - functional programming tools
30  * stringr - for manipulating text and character strings
31  * forcats - for dealing with categorical or ordinal data (factors)
32
33  Summarize data. The data `mtcars` is available in the base version
34
35  ```{r mtcars}
36  summary(mtcars)
37  ```
38
39  # Including Plots
40
41  You can also embed plots, for example:
42
43  ```{r displacement_mpg, echo=FALSE}
44  ggplot(mtcars, aes(x = disp, y = mpg)) +
45    geom_point()
46  ```
47
48  ```{r header, echo=TRUE}
49  head(mtcars)
50  ```
```

5:9 | Week 01: Introduction to R and R-Markdown

Console | Terminal | Background Jobs

R 4.3.1 · C:/Users/alh241/AppData/Local/Temp/Temp1ad48b6-2ed7-4ffd-b01c-207ef06dc6b3_BEM2031_R_wee

```
>
> # Chunk 3: displacement_mpg
> ggplot(mtcars, aes(x = disp, y = mpg)) +
+   geom_point()
>
> # Chunk 4: header
> head(mtcars)
                   mpg cyl disp  hp drat    wt  qsec vs am gear carb
Mazda RX4         21.0   6  160 110 3.90 2.620 16.46  0  1    4    4
Mazda RX4 Wag     21.0   6  160 110 3.90 2.875 17.02  0  1    4    4
Datsun 710        22.8   4  108  93 3.85 2.320 18.61  1  1    4    1
Hornet 4 Drive    21.4   6  258 110 3.08 3.215 19.44  1  0    3    1
Hornet Sportabout 18.7   8  360 175 3.15 3.440 17.02  0  0    3    2
Valiant           18.1   6  225 105 2.76 3.460 20.22  1  0    3    1
>
```

Week 01: Introduction to R and R-Markdown

you

2022-01-17 T00:00:00

Contents

Introduction

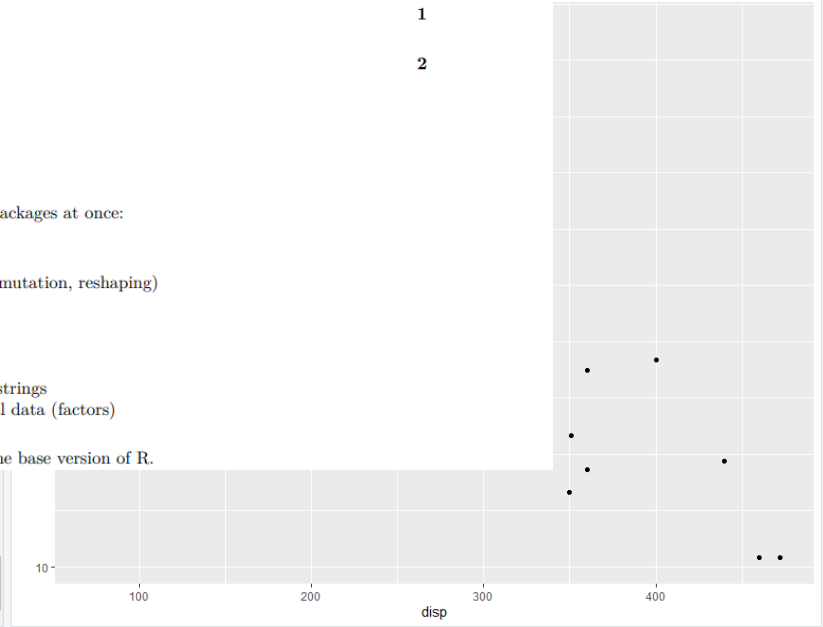
Packages.

The **tidyverse** package just loads many different packages at once:

- ggplot2 - charts and plots :)
- dplyr - chaining data manipulation (filtering, mutation, reshaping)
- tibble - for effective data tables
- tidyr - helpful functions for cleaning data
- readr - functions for reading and writing data
- purrr - functional programming tools
- stringr - for manipulating text and character strings
- forcats - for dealing with categorical or ordinal data (factors)

Summarize data. The data **mtcars** is available in the base version of R.

Publish



University of Exeter

Any questions?

?