

R Markdown with Bump Charts, Heatmaps and Networks

EdinbR User Group
18 Apr 2018

Nevil Hopley

This talk will contain:

- the experience of moving to using R Markdown
- UCAS application data for one Scottish secondary school
- stacked bar charts, bump charts, heat maps and interactive networks

Still an R Newbie

- R user since January 2017
- Teaching in Secondary Schools since 1993
- Teaching S6 AH Statistics since 1999
- Previous EdinbR Talks:
 - Breakdown Plots (June 2017)
 - Postcodes on Maps (Sep 2017)
 - Sankey Plots (Nov 2017)

Nevil Hopley

Sources of Code, Help & Inspiration

<https://rmarkdown.rstudio.com/lesson-1.html>

http://data-slinky.com/2016/07/31/bump_charts.html



Social Network Analysis Workshop

Friday, March 23, 2018
9:00 AM – 5:00 PM

Scotland's Rural College (SRUC)



stackoverflow

R Packages Used

tidyverse

ggmap

mapproj

RColorBrewer

zoo for `na.locf`

tools for `toTitleCase`

broom

visNetwork

igraph

ggrepel

scales

knitr

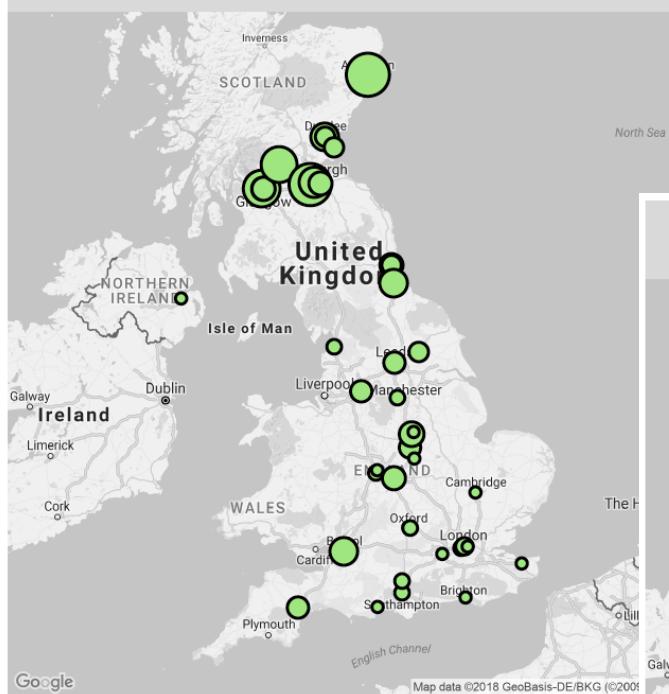
UCAS Applications Status Reports

1	G28	GLASG	P390	REJ	Film and Television Studies
2	E59	ENAP	P303	W	Film
3	S75	STIRL	P300	U	Film and Media
4	N77	NORTH	P310	C	Film and TV Production
5	A20	ABRDN	W690	C	Film and Visual Culture
1	S01	SRUC	4J25	REJ	Animal Care
1	A20	ABRDN	M114	UF	Law
2	G28	GLASG	M114	REJ	Law
3	D65	DUND	M114	UD	Law (Scots)
4	E56	EDINB	M114	REJ	Law
1	D65	DUND	V140	UD	History
2	A20	ABRDN	VL13	UD	History and Sociology
3	S36	STA	V140	REJ	Modern History
4	G28	GLASG	LV31	UD	History/Sociology (SocSci)
5	E56	EDINB	VL13	UF	History and Sociology
1	U80	UCL	R990	C	European Social and Political Studies (4 years)
2	C05	CAM	L000	C	Human, Social and Political Sciences
3	K60	KCL	L240	U	Politics
4	E56	EDINB	8M9D	U	Politics with Quantitative Methods

Application Status Codes

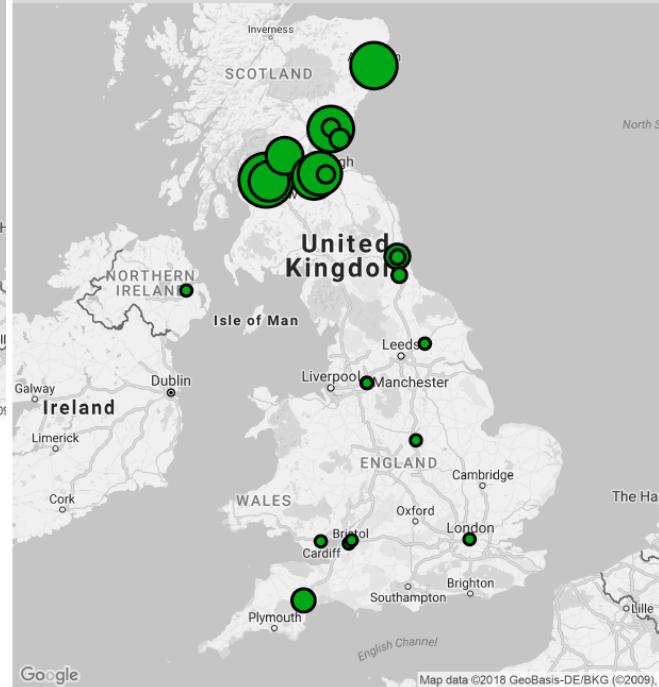
Awaiting Response	Unconditional Offer Accepted
Conditional Offer Accepted	Unconditional Offer as Insurance
Conditional Offer as Insurance	Unconditional Offer Made
Conditional Offer Made	Unsuccessful
Interview	Withdrawn
Offer Declined	

Conditional Offer Made

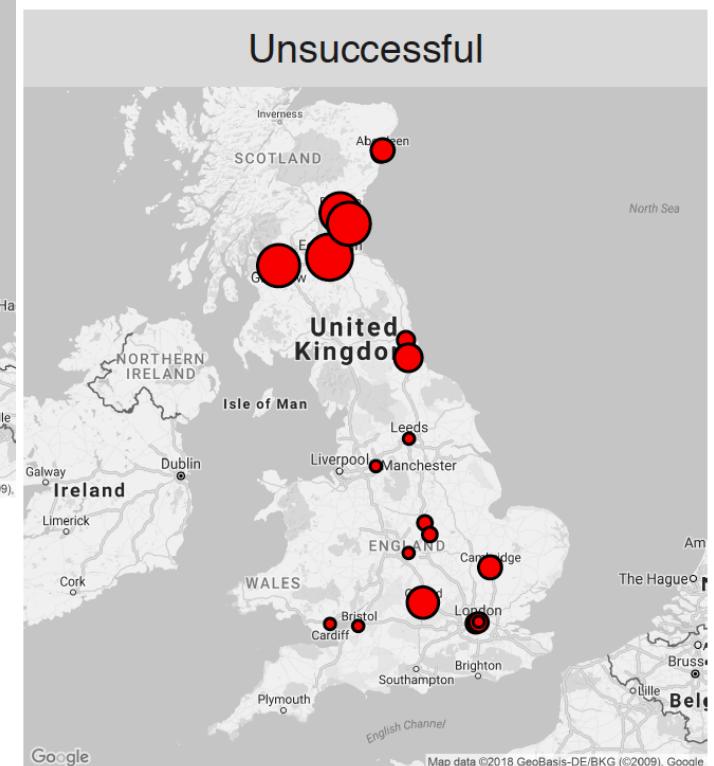


Merge University codes with Postcodes, with Latitude and Longitude and with ggmap imagery!

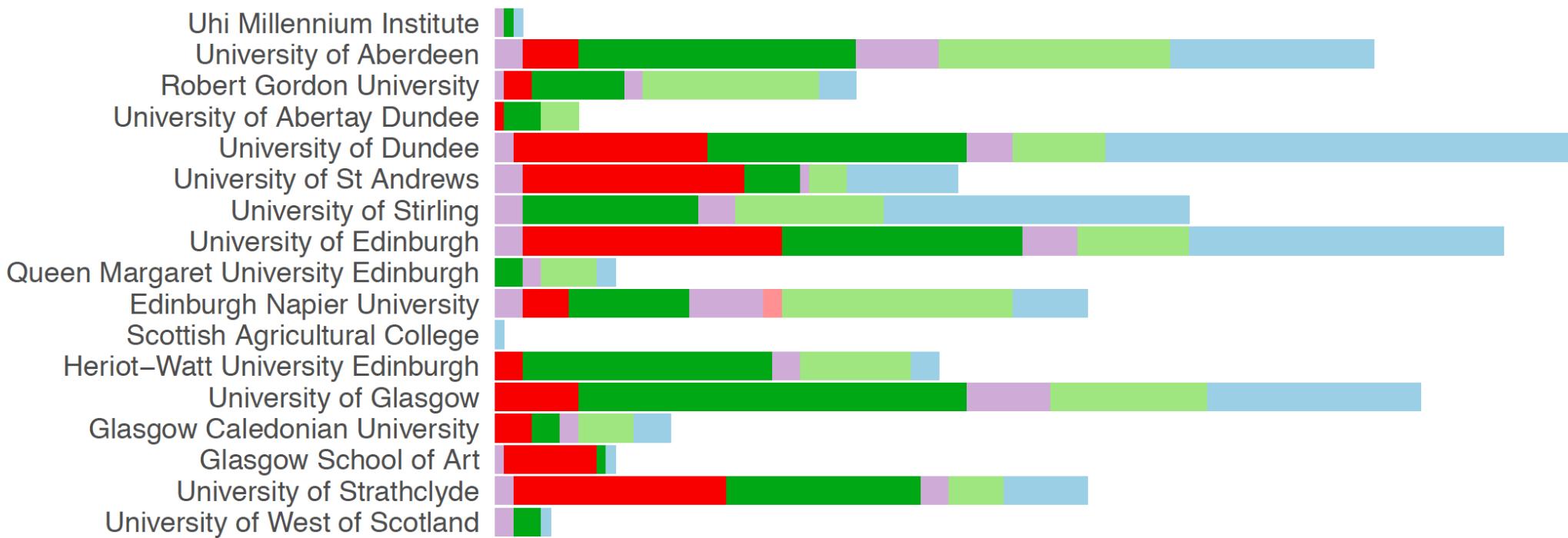
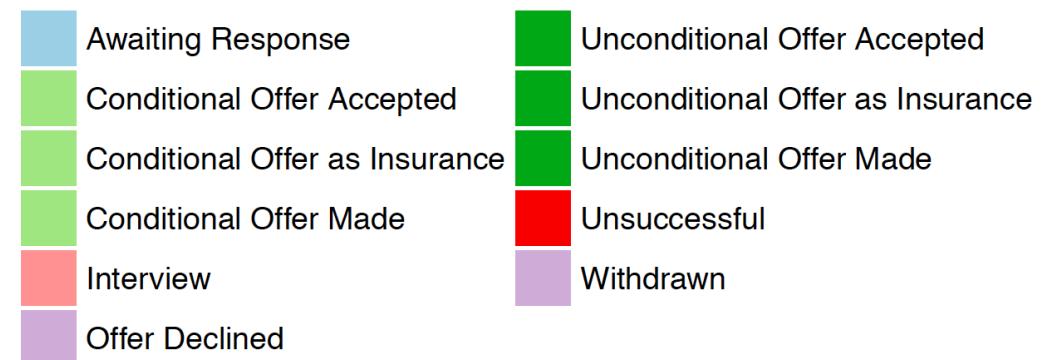
Unconditional Offer Made



Unsuccessful



Scottish Universities Applications



Non-Scottish Universities Applications



A4 pdf size page is no longer long enough, so hurray for R markdown and html web-pages that can be any length!

UCAS Course Codes

1 letter codes

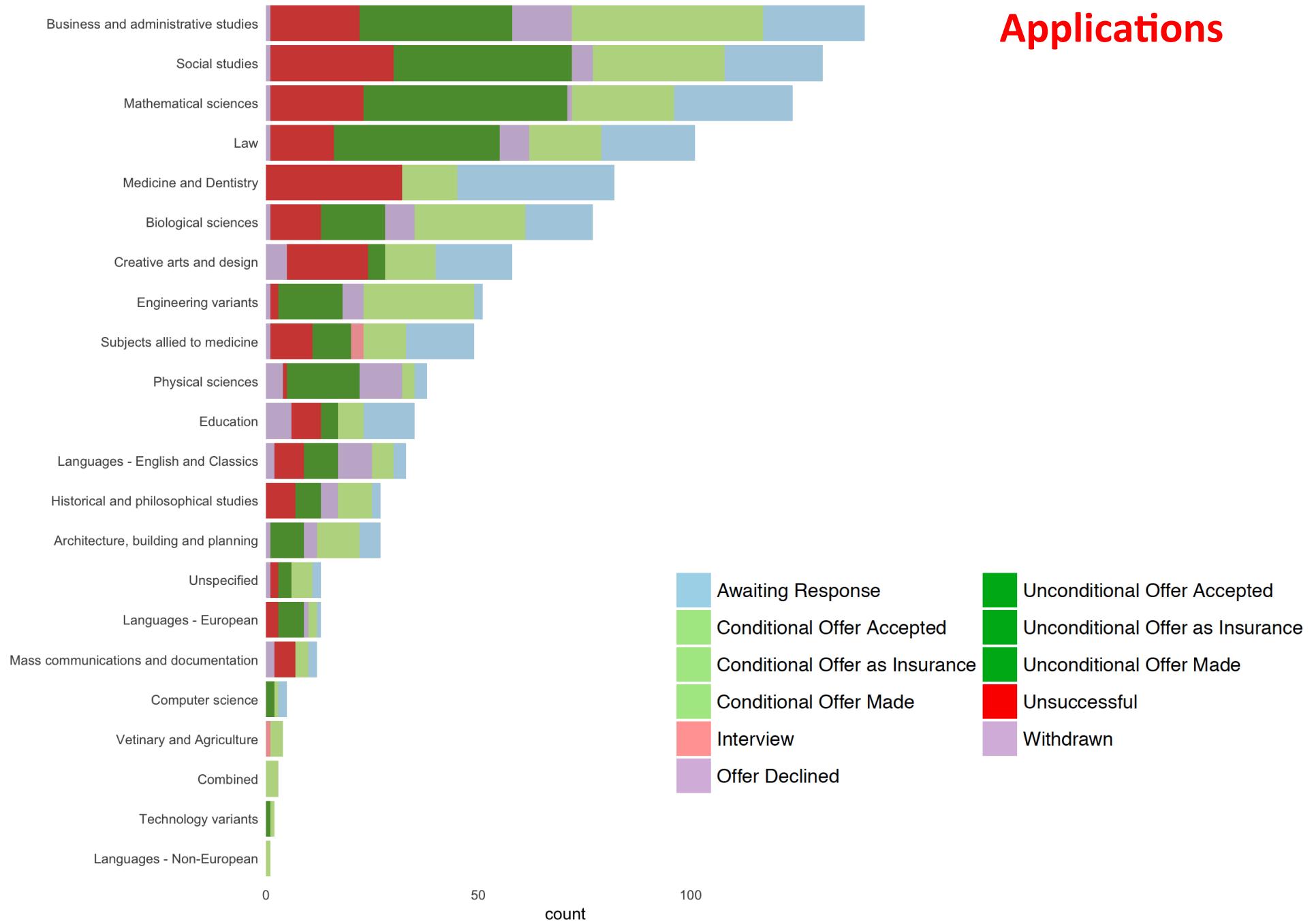
- A Medicine and Dentistry
- B Subjects allied to medicine
- C Biological sciences
- D Veterinary and Agriculture
- F Physical sciences
- G Mathematical sciences
- H Engineering variants
- J Technology variants
- I Computer science
- K Architecture, building and planning
- L Social studies
- M Law
- N Business and administrative studies
- P Mass communications and documentation
- Q Languages - English and Classics
- R Languages - European
- T Languages - Non-European
- V Historical and philosophical studies
- W Creative arts and design
- X Education
- Y Combined

2 letter codes

- A0 Broadly-based programmes within medicine & dentistry
- A1 Pre-clinical medicine
- A2 Pre-clinical dentistry
- A3 Clinical medicine
- A4 Clinical dentistry
- A9 Others in medicine & dentistry
- B0 Broadly-based programmes within subjects allied to medicine
- B1 Anatomy, physiology & pathology
- B2 Pharmacology, toxicology & pharmacy
- B3 Complementary medicines, therapies & well-being
- B4 Nutrition
- B5 Ophthalmics
- B6 Aural & oral sciences
- B7 Nursing
- B8 Medical technology
- B9 Others in subjects allied to medicine
- C0 Broadly-based programmes within biological sciences
- C1 Biology
- C2 Botany
- C3 Zoology

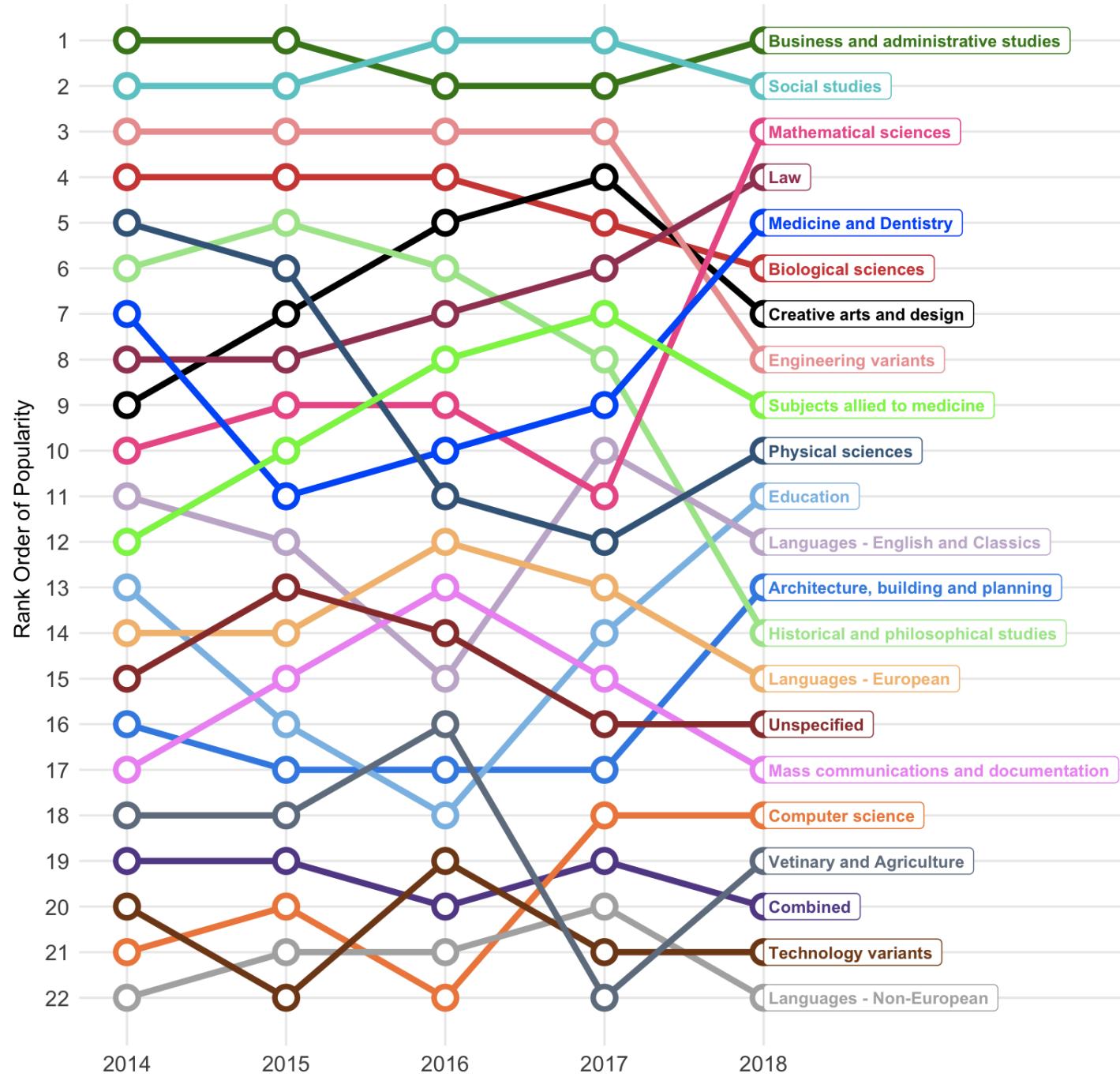
... a total of 188 of them!

Applications



This summary graphic was produced by Nevil Hopley, using R.

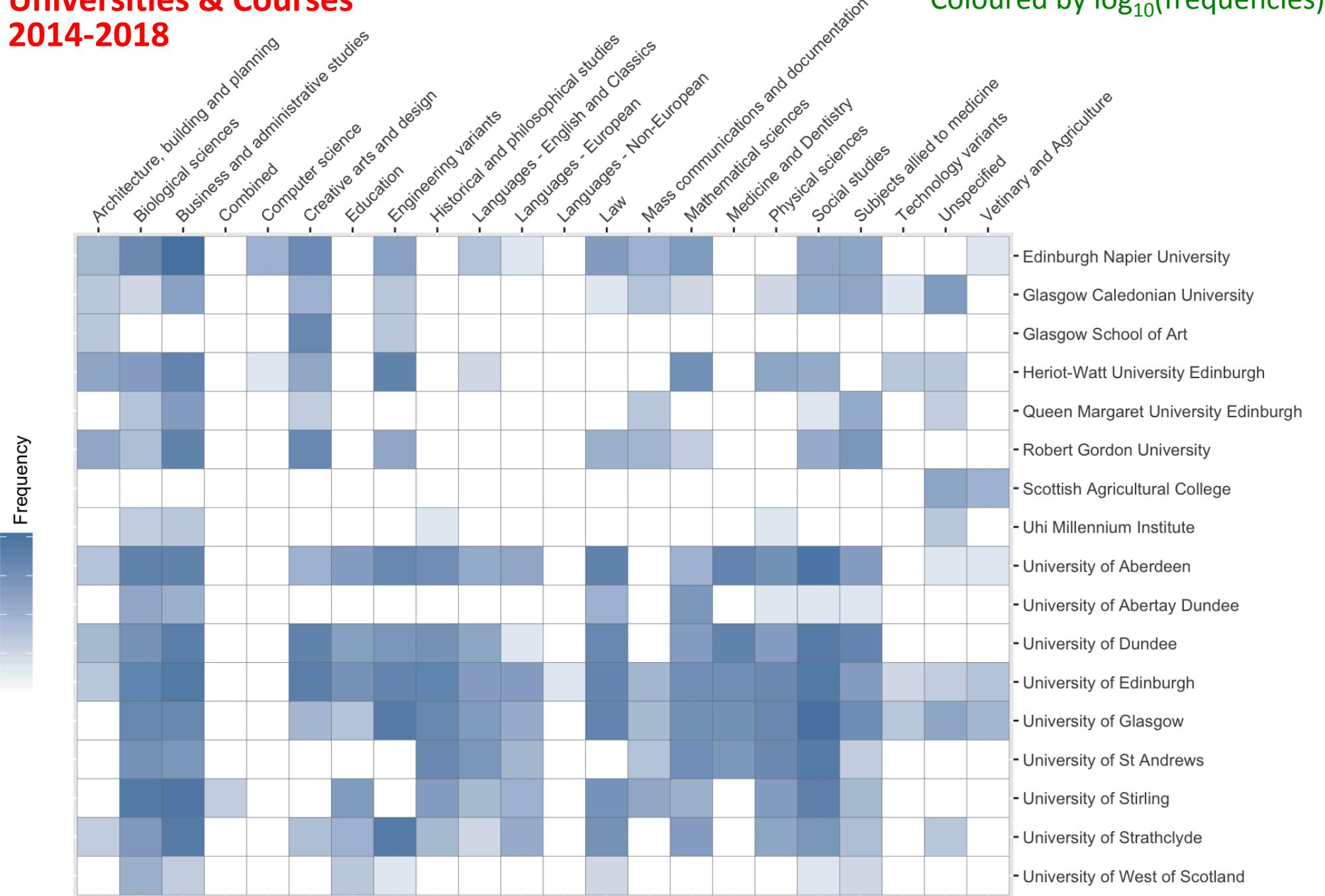
Courses Applied to 2014-2018



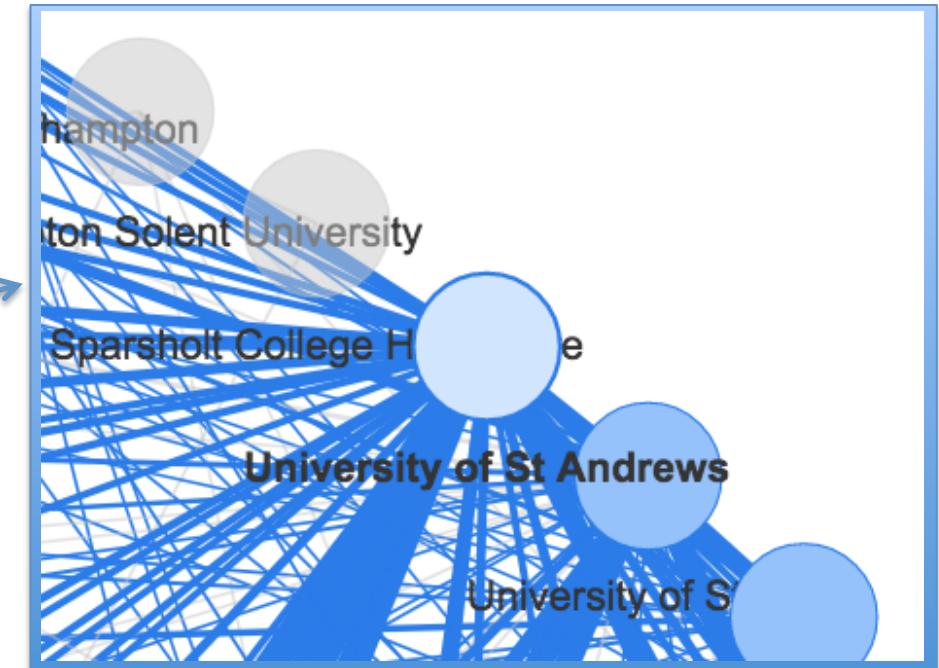
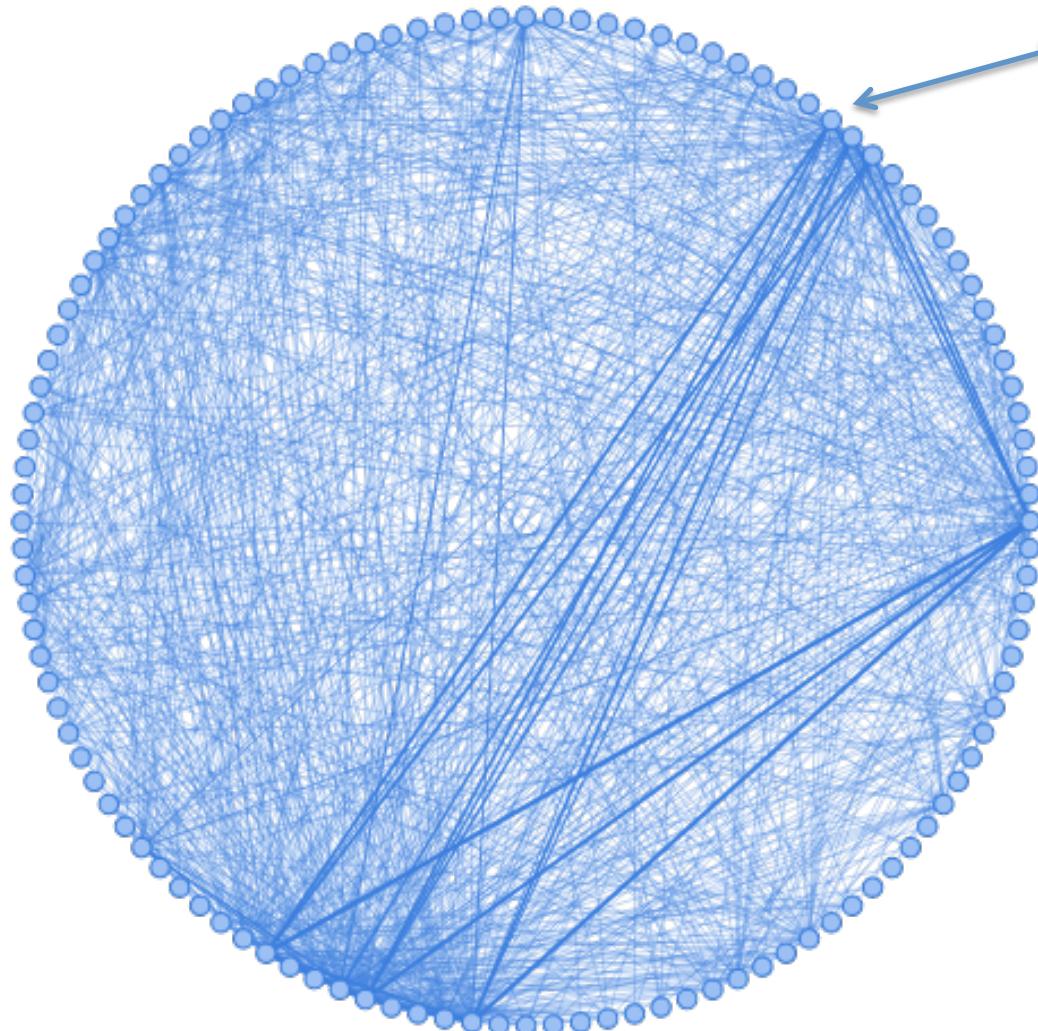
Will encounter colour palette problems if 2 letter course codes are used

Universities & Courses 2014-2018

Coloured by $\log_{10}(\text{frequencies})$



UCAS Application Form Clusters of Applications 2014-2018



Before Datafest at SRUC...

Default panels

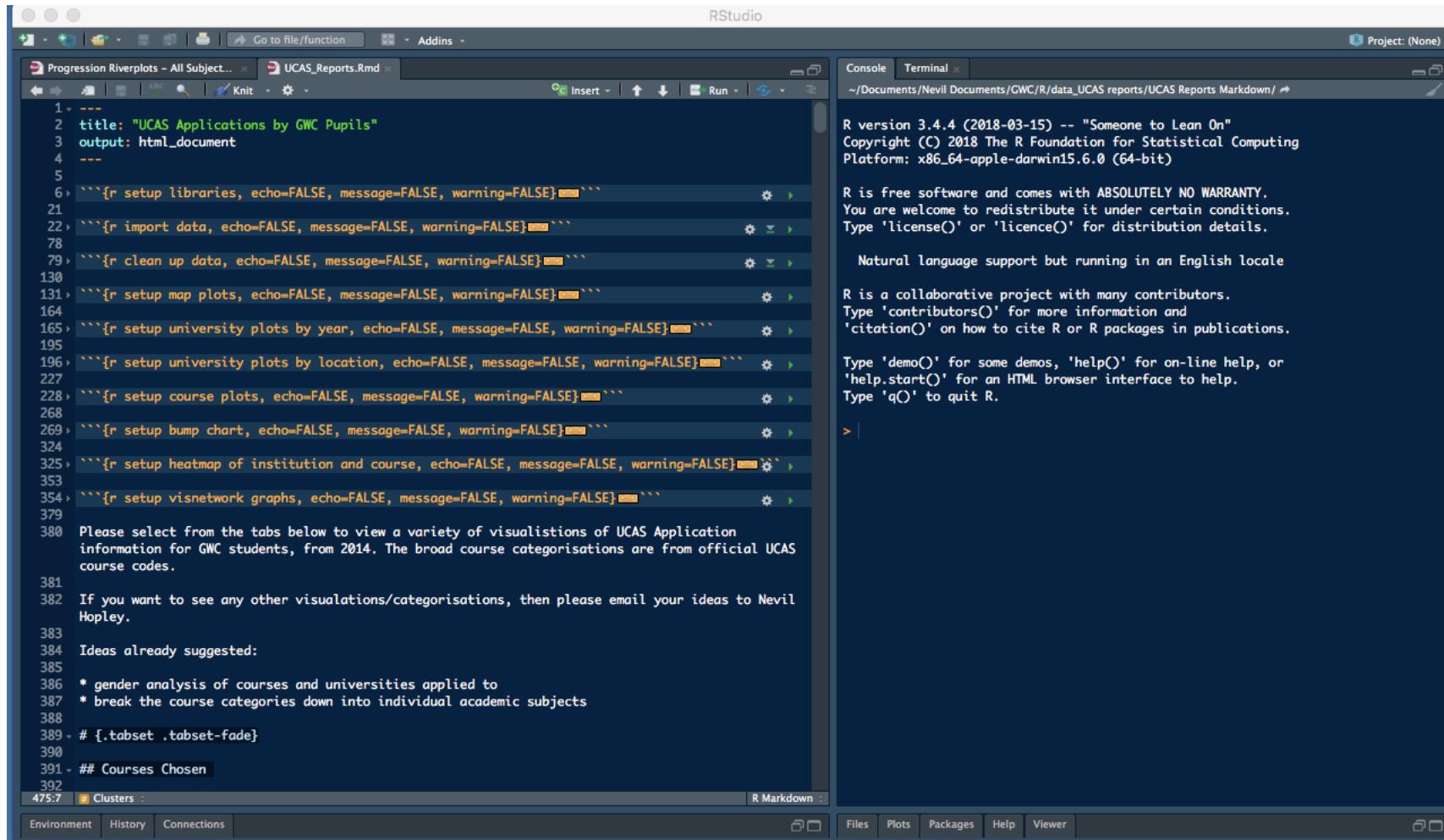
The screenshot shows the RStudio interface with several panels open:

- Data View:** Displays a table titled "labike" with 38 observations of 6 variables. The variables are name, longitude, latitude, type, bike_count_pm, and ped_cc.
- History:** Shows the R session history with commands like `load`, `weather`, `library`, and `plot`.
- Console:** Displays the R startup message and the command history.
- Plots:** Shows a box plot comparing bike count (pm) across four categories: bike lane, bike path, bike route, and none.

	name	longitude	latitude	type	bike_count_pm	ped_cc
1	1st & Alameda	-118.2381	34.04917	none	62	241
2	4th & Wilton	-118.3134	34.06713	bike route	48	87
3	7th & Figueroa	-118.2599	34.04939	none	216	1979
4	8th & La Brea	-118.3446	34.06045	none	72	272
5	9th & Pacific	-118.2873	33.73512	none	58	160

After Datafest at SRUC...

Maximise size of both code and console panels



The screenshot shows the RStudio interface with the following details:

- Code Editor (Left Panel):** Displays the R Markdown file "UCAS_Reports.Rmd". The code includes various R chunks (e.g., `r setup libraries`, `r import data`) and a large block of text explaining course categorizations and visualization ideas.
- Console (Right Panel):** Shows the R environment information and a welcome message from R version 3.4.4.
- Bottom Navigation Bar:** Includes tabs for Environment, History, Connections, Files, Plots, Packages, Help, and Viewer.

Tools > Global Options > Code > Soft-wrap R source files

VisNetwork interactive graphs

Before R Markdown...

```
wd <- getwd()
setwd(paste0(wd, "/csv ucas"))
core_info_filename <- list.files(getwd())
ucas_report <- read.csv(core_info_filename,
                         header = TRUE,
                         sep = ",",
                         fileEncoding = "latin1",
                         na.strings = c("NA", ""))
```



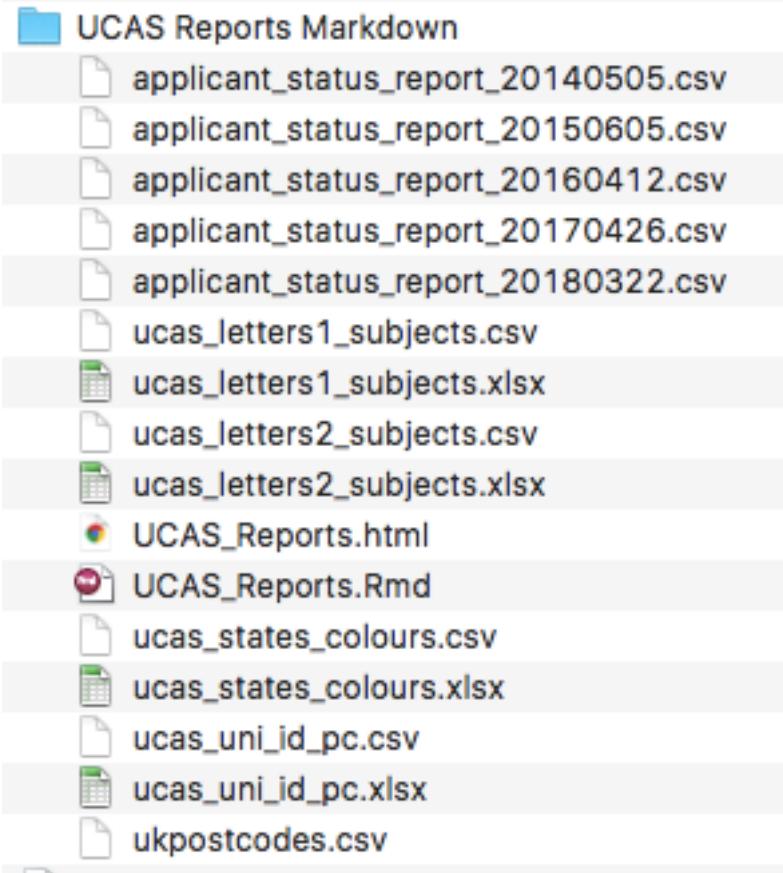
▼	csv data
	JACS3.csv
	ucas_letters1_subjects.csv
	ucas_letters1_subjects.xlsx
	ucas_letters2_subjects.csv
	ucas_letters2_subjects.xlsx
	ucas_states_colours.csv
	ucas_states_colours.xlsx
	ucas_uni_id_pc.csv
	ucas_uni_id_pc.xlsx
	ukpostcodes.csv
▼	csv ucas
	applicant_status_report_20180322.csv
►	pdf
	UCAS Reports v1.R

```
pdf_filename <- paste0(core_info_filename, ".pdf")
pdf(file= pdf_filename, paper="a4", width=8.2, height=11.6, onefile=TRUE)
grid.newpage()
rows <- 1
cols <- 1
pushViewport(viewport(layout = grid.layout(rows, cols)))
```



After R Markdown...

setwd(..)



```
register_google(key = "AIzaSyAaDei50mj_j-ECOicQx8BF_64Knf0")  
map <- get_map(location = 'UK BB7',  
                 zoom = 6,  
                 maptype = "roadmap",  
                 color = "bw")
```

HTML web pages with tabs!

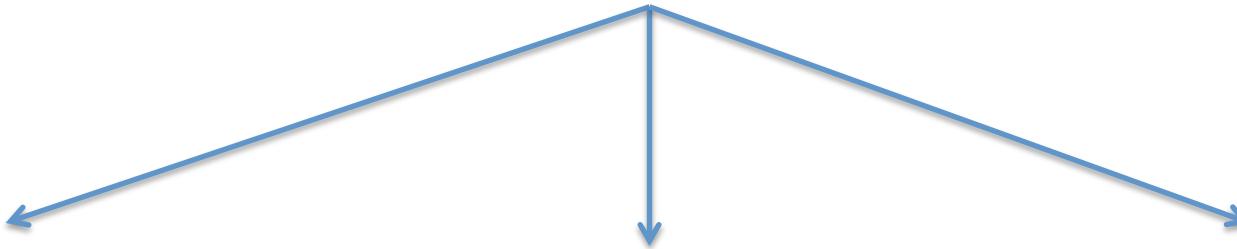
The screenshot shows a tabbed interface from an R Markdown generated HTML file. The main tabs are "Courses Chosen", "Universities Chosen", and "Study What and Where". Below these, there is a sub-tab labeled "Popularity" and a row of year tabs: "2017-18", "2016-17", "2015-16", "2014-15", and "2013-14".

Collapsible code chunks that can be independently executed, with animated progress bars. Bliss!

The screenshot shows a code editor in RStudio with several code chunks collapsed. Each chunk has a small icon indicating it can be expanded. The numbers on the left likely represent line numbers.

```
6: `r setup libraries, echo=FALSE, message=FALSE, warning=FALSE`  
21: `r import data, echo=FALSE, message=FALSE, warning=FALSE`  
78: `r clean up data, echo=FALSE, message=FALSE, warning=FALSE`  
130: `r setup map plots, echo=FALSE, message=FALSE, warning=FALSE`  
164: `r setup university plots by year, echo=FALSE, message=FALSE, warning=FALSE`  
195:
```

Secure distribution of R Markdown HTML files...



Easy password protection for any site!

(other webserver companies are available)

Free for educational institutions
that teach courses that cite 'R' in
their official syllabus.

**Do we have time
for a quick
demonstration?**



Next Steps Help Needed!

- 1. Colour palette for large numbers of categories**

- 2. What package to give html-widget-based
interactive plots where data categories appear
and disappear, rather than just dimmed?**

- 3. Any advice for optimising using R Markdown.**