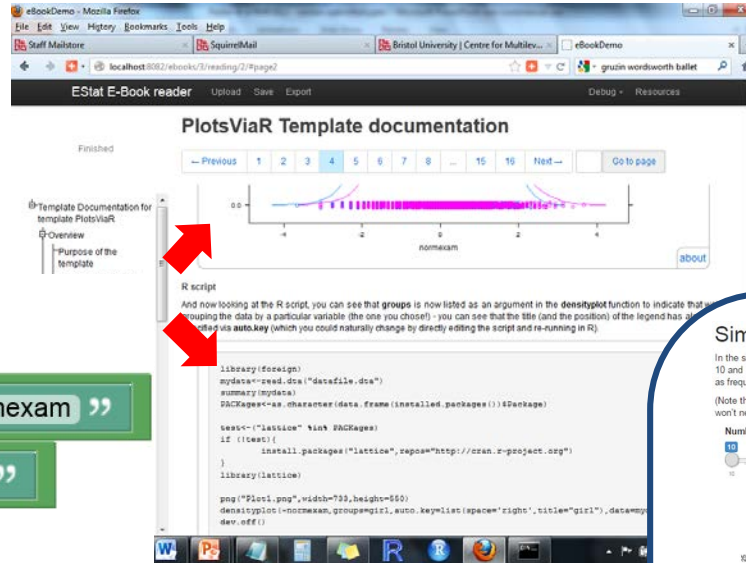


Developing
new tools to
promote stats
education...

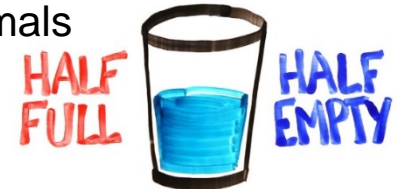


PlotsViaR Template documentation

R script

```
library(fooreign)
mydata<-read_dsv("datafile.dsv")
summary(mydata)
DACKages<-as.character(data.frame(installed.packages()))$Package
sear<-"lattice" %>% link DACKages
if (isect){
  install.packages("lattice", repos="http://cran.z-project.org")
}
library(lattice)
png("Plot1.png",width=750,height=550)
densityplot(~normexam,group=gl(2),auto.key=list(space="right",title="gl1"),deseq=
dens.ndf())
```

PhD: cognition & emotion in
animals



Start

Select dataset “ tutorial ”

Set Input “ vals ” = “ normexam ”

Set Input “ bins ” = “ 15 ”

Template “ Histogram ”

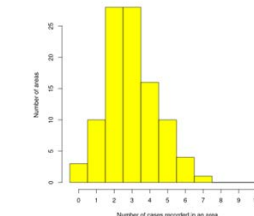
Show “ histogram.svg ”

Simulating for different population sizes and risks

In the simulation above the probability of being reported as a case was fixed at 0.2, as was the number of people residing in each area (fixed at 10 and 100). Below you can vary these parameters and explore the effect they have on the chart below: plotting the number of cases reported as frequencies, and the schematic 'maps' showing 'hot' and 'cool' spots.



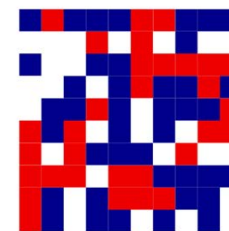
Simulation of 100 areas, each with population of 10, with each individual having a 0.3 probability of being recorded as a case



Histogram

- The histogram to the left plots the number of cases recorded across all 100 areas.
- As you can see, the horizontal axis is scaled to cover all possible outcomes (from no-one being reported as a case in an area, to everyone being reported as a case in an area) given the number of people in each area you choose above.

Map 1



- This plot, to the left, is a schematic representation of red 'hotspots' (areas which have reported more cases than the most probable outcome) and blue 'coolspots' (areas which have reported less cases than the most probable outcome). The white areas are those which happen to match the most probable outcome.
- Given the probability (0.3) and the number of people in each area (10) you specified above, the most probable outcome for each area would be 3, but by chance it is likely that some will record more cases than this, and some will record less cases (and of course health authorities don't really know the underlying probability from which samples are drawn).
- Whilst the allocation of areas to each square of this 'map' is random, sometimes clusters will appear by chance. NOTE: could scale this differently, e.g. differences from median or possibly mode, or some other way of representing extent of spread.



Breaking down big numbers

ROYAL STATISTICAL SOCIETY

Statistics for journalists

RoyalStatSoc • 10 videos • 108 views • Last u

Statistics for Journalists workshop
Filmed at Science and statistics for journalists
Bloomberg in London

▶ Play all < Share + Save