

Using org-mode for scientific posters

Philipp Homan¹

¹Department of Psychiatry, The Donald and Barbara Zucker School of Medicine at Northwell/Hofstra, Hempstead, NY



DONALD AND BARBARA
ZUCKER SCHOOL of MEDICINE
AT HOFSTRA/NORTHWELL

Background

- Here we show how org-mode (version 9.1.7) and emacs (version 25.2.2) can be used to make decent looking scientific posters
- With org-mode we can populate the poster with code, graphs and numbers from inline code in languages such as R, python, Matlab and even shell scripting
- For example, this poster was created on 2018-05-01 00:19 on Ubuntu 17.10.
- Inline code could look like this (which will produce a graph; Fig. 1):

```
1 set.seed(20180402)
2 x1 <- rnorm(100, 0, 1)
3 x2 <- rnorm(100, 0.5, 1)
4 hist(x1, col="red")
5 hist(x2, col="blue", add=TRUE)
```

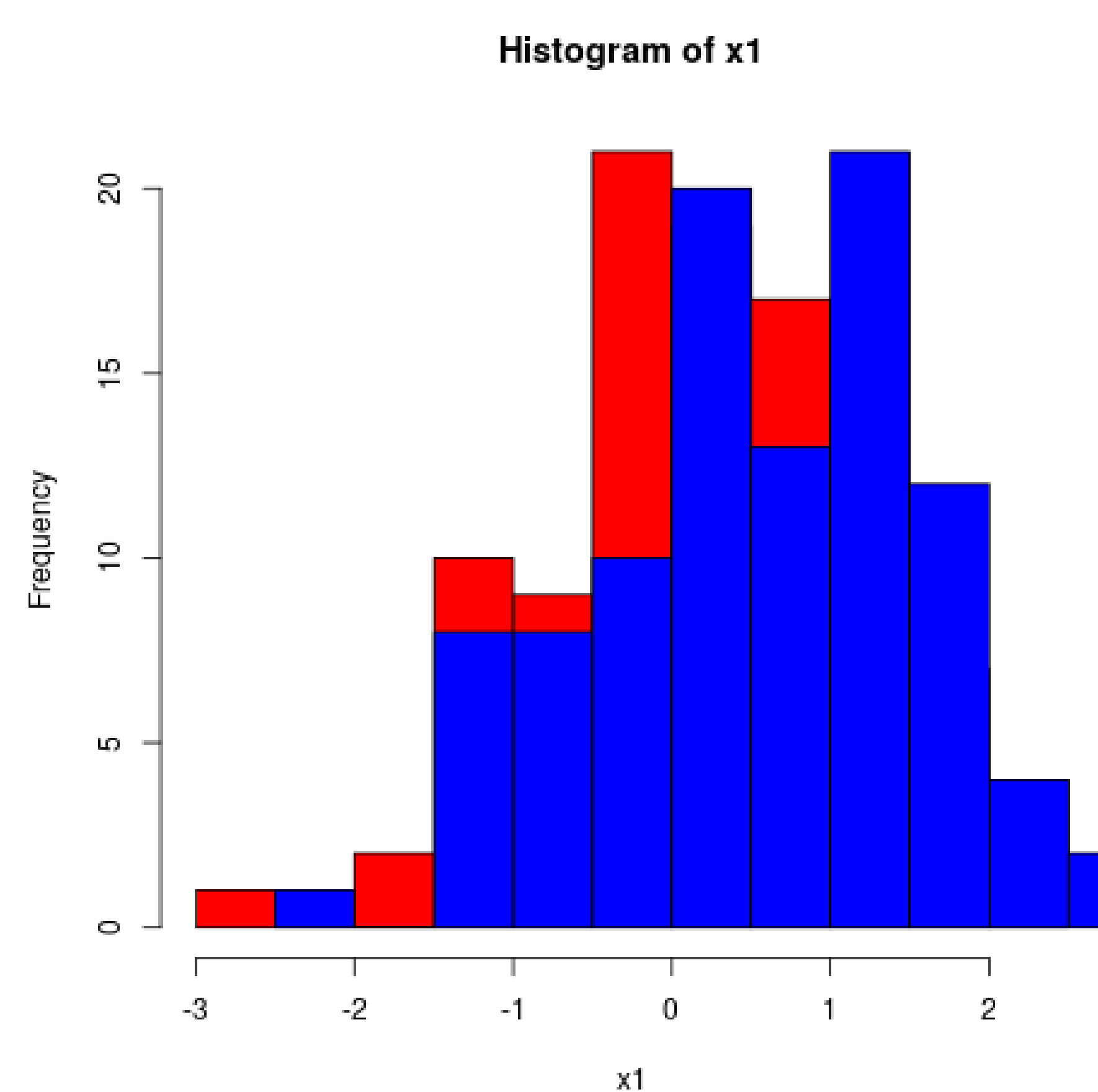


Figure 1: This is the output.

Inline code and tables

- Tables are powerful in org-mode and even include spreadsheet capabilities
- Some code to process the first vector from above to make a table out of its summary could look like this, which would result in a little table (Table 1):

```
1 library(broom)
2 library(dplyr)
3 t1 <- tidy(round(summary(x1), 2))
4 t2 <- tidy(round(summary(x2), 2))
5
6 # This will export as a table
7 rbind(t1, t2) %>%
8 mutate(name=c("x1", "x2"))
```

minimum	q1	median	mean	q3	maximum	name
-2.29	-0.49	0.11	0.14	0.8	2.47	x1
-2.17	-0.45	0.07	0.13	0.85	2.23	x2

Table 1: A table summarizing the two distributions.

Graphics

- We can use shell scripting to grab an image with curl from the internet (Fig. 2):

```
1 # Download emacs icon from gnu.org
2 curl -O https://www.gnu.org/software/emacs/images/emacs.png
```



Figure 2: This is the downloaded image.

Math

- Let's describe how to compute the distance between the two simulated distributions x_1 and x_2 from before:

The Kullback-Leibler (KL) divergence measures the difference between two probability distributions (i.e., the loss of information when one distribution is used to approximate another). The KL divergence is thus defined as

$$D_{KL}(P||Q) = \sum_{i=1}^n P(i) \log \frac{P(i)}{Q(i)} \quad (1)$$

with P and Q being two probability distribution functions and n the number of sample points. Since $D_{KL}(P||Q)$ is not equal to $D_{KL}(Q||P)$, a symmetric variation of the KL divergence can be derived as follows:

$$D_{KL}(P, Q) = \sum_{i=1}^n \left(P(i) \log \frac{P(i)}{Q(i)} + Q(i) \log \frac{Q(i)}{P(i)} \right). \quad (2)$$

Columns

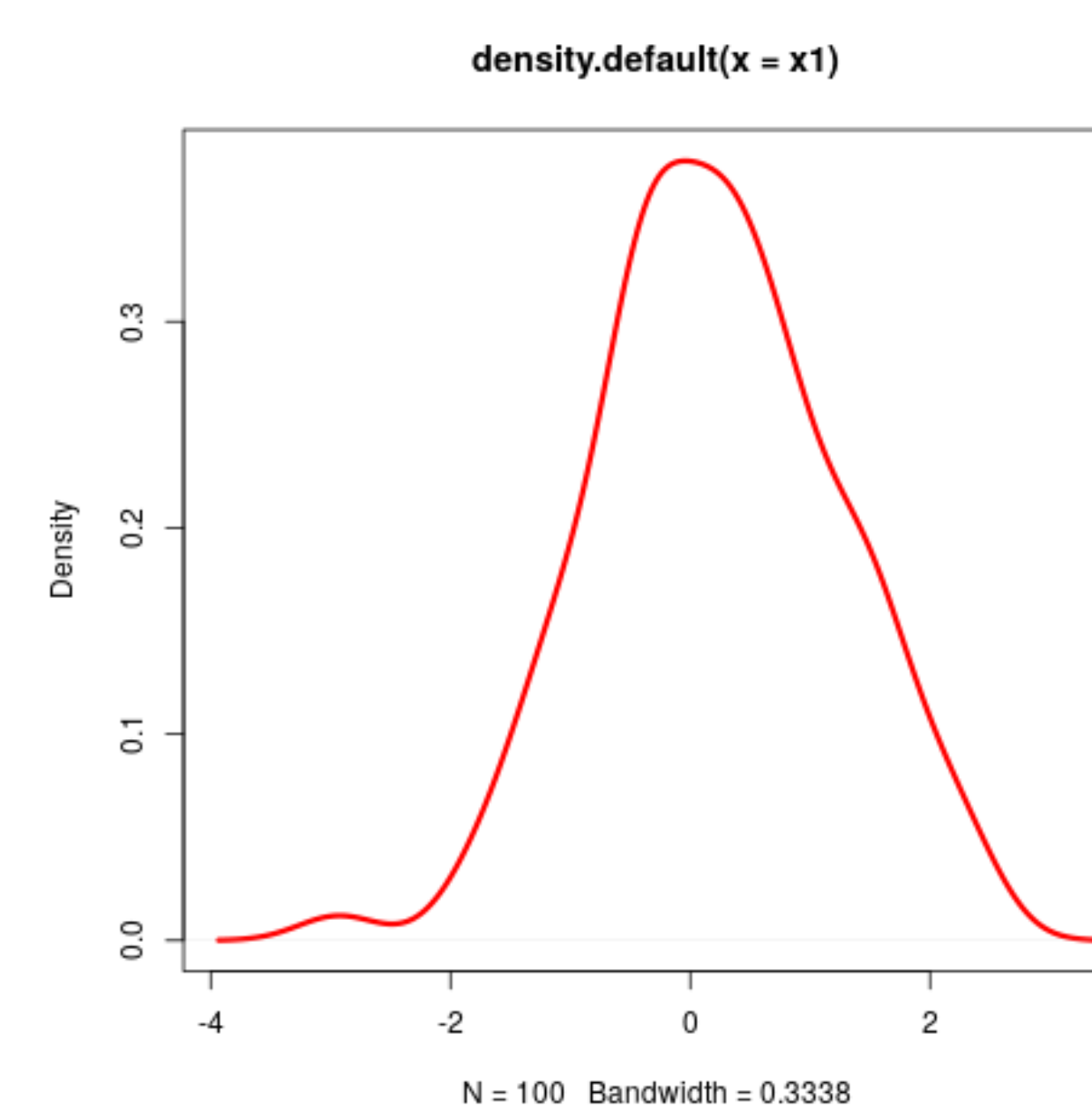


Figure 3: This is the left figure of a two-column block, showing the density of x_1 .

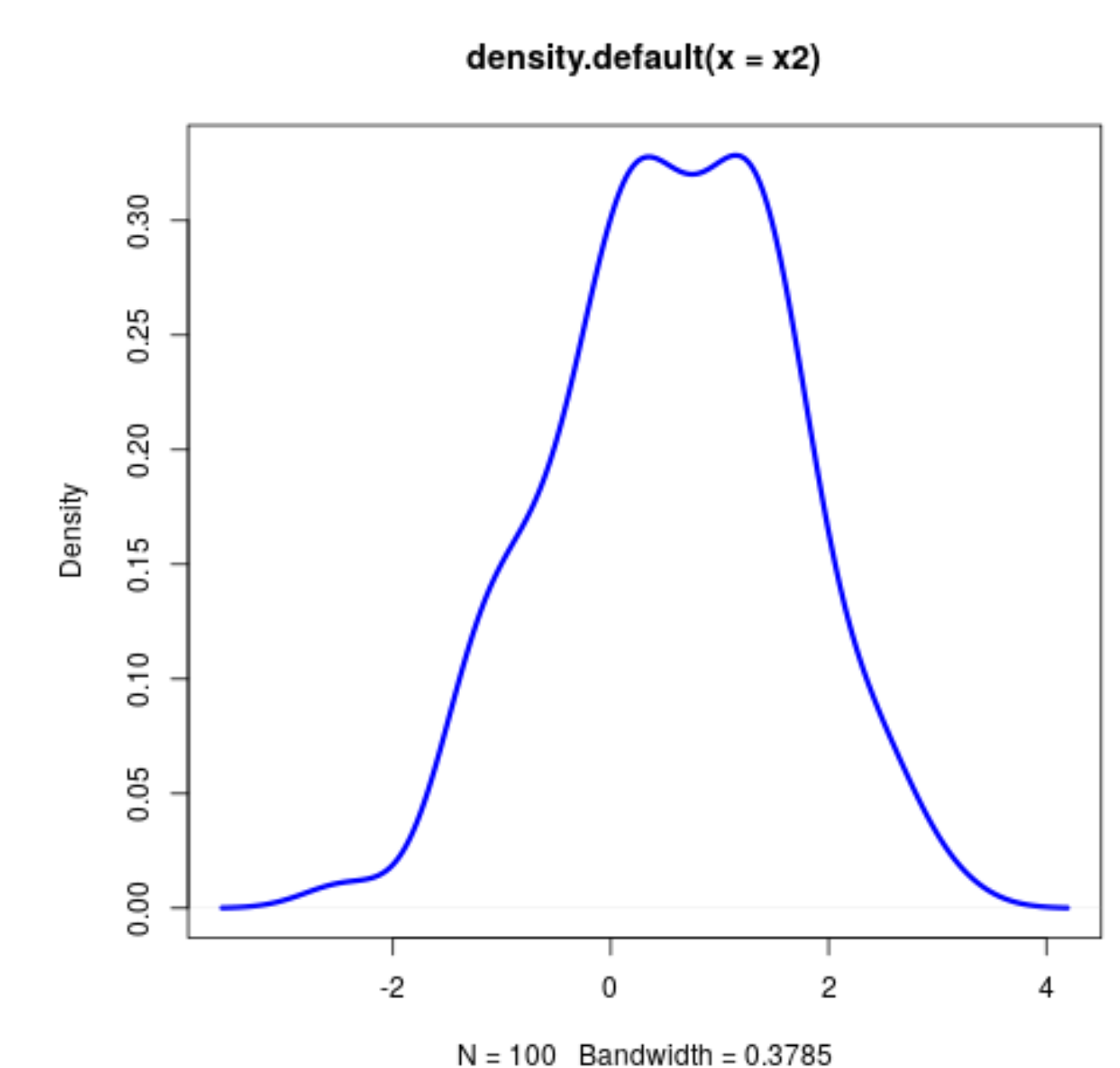


Figure 4: This is the right figure. It shows the density of x_2 .

Conclusions

- This example is meant to show how versatile org-mode is
- Scientific posters can be produced with a simple text editor