Reporting data results #1

Plot guidelines

Guidelines for good plots

There are a number of very thoughtful books and articles about creating graphics that effectively communicate information.

Some of the authors I highly recommend (and from whose work I've pulled the guidelines for good graphics we'll talk about this week) are:

- Edward Tufte
- Howard Wainer
- Stephen Few
- Nathan Yau

You should plan, in particular, to read *The Visual Display of Quantitative Information* by Edward Tufte before you graduate.

Guidelines for good plots

This week, we'll focus on six guidelines for good graphics, based on the writings of these and other specialists in data display.

The guidelines are:

1. Aim for high data density.
2. Use clear, meaningful labels.
3. Provide useful references.
4. Highlight interesting aspects of the data.
5. Make order meaningful.
6. When possible, use small multiples.

Packages for examples

For the examples, I'll use `dplyr` for data cleaning and, for plotting, the packages `ggplot2`, `gridExtra`, and `ggthemes`.

```
library("dplyr")

library("ggplot2")
library("gridExtra")
library("ggthemes")
```

Example data

You can load the data for today's examples with the following code:

```
library("faraway")
data("worldcup")
```

Example data

```
library("dlnm")
data("chicagoNMMAPS")
chic <- chicagoNMMAPS
chic_july <- chic %>%
  filter(month == 7 & year == 1995)
```

```
slice(chic_july, 1:3)
```

```
##         date time year month doy      dow death cvd resp     temp   dptp   r
## 1 1995-07-01 3104 1995     7 182 Saturday   109  52    4 18.61111 47.750 55.
## 2 1995-07-02 3105 1995     7 183   Sunday   103  49    4 17.50000 49.375 47.
## 3 1995-07-03 3106 1995     7 184   Monday   112  51   10 21.66667 57.500 51.
##       pm10       o3
## 1 15.64046 27.47199
## 2 27.12486 24.81275
## 3 49.11362 43.63302
```