GettingStarted

Statistics 4868/6610 Data Visualization

Prof. Eric A. Suess

January 6, 2015

Introduction

Be sure to read the Introduction to the book. The discussion about unemployment is very interested.

The figures in FIGURE 1-1 shows the change over time very nicely.

And FIGURE 1-2 show the last Map for 2009.

page XVII

So what is visualization?

Well, it depends on who you talk to. Some people say it is strictly traditional graphs and charts.

Others have a more liberal view where anything that displays data is visualization.

page XXI

Statistics isn't just about hypothesis testing

Statistics is about **storytelling with data**.

page XXIII

Chapter 1

Journalism is one of the main places where visualization is being used extensively. Certainly in print, but also online.

New York Times

Life in New York City, Where Arrests Are Down and Tickets Are Rarities

FiveThiryEight

DataDrivenJournalism

Gapminder

Public Health.

This is Hans Roslings website.

[Gapminder](#)

Watch some of the videos. They are all very interesting.

Download the software and give it a try. Note that you can watch the dynamic display of the data both on a bubble graph but also on a map.

page 7

What to Look For

- Time Series plot, look for Patterns
- [yahoo finance](#)

Questionable Data

Data-checking and verification is one of the most important – if not the most important – part of graph design.

Design

Cleveland and McGill [Graphical Preception and Graphical Methods for Analyzing Data](#)

page 13

Design

- Label Axes
- Keep Your Geometry in Check - this is very important!
- Include Your Sources - give credit where credit is due
- Consider Your Audience

Simple R example

```r
summary(cars)
```

```
     speed           dist
 Min.   : 4.0   Min.   :  2.00
 1st Qu.:12.0   1st Qu.: 26.00
 Median :15.0   Median : 36.00
 Mean   :15.4   Mean   : 42.98
 3rd Qu.:19.0   3rd Qu.: 56.00
 Max.   :25.0   Max.   :120.00
```

Simple R example

```r
boxplot(cars)
```