Class Notes

Jan 26, 2015

**Data Visualization**

* **Three categories:**

1. **Statistical Graphics**

The first was not around until 1750-1800.

Some important mile-stones: Playfair was the first one to create charts tables and graphics around 1800.

Tukey created boxplot 1969

1. **Time Series**

Economic time (1786, Playfair)

Geophysical (Lambert 1779)

Span-time (Ibbry 1885)

1. **Data Maps**

First data maps didn’t emerge until 17th century.

Trade Winds (Haley, 1686)

EPT – Cholera (Snow, 1854)

* **Statistical Graphics**

1. **Univariate:**

Histogram/Probability Density Function (PDF)/Pie Charts/ Boxplot

1. **Bivariate**

Scatterplot (show the relationship between two variables)

1. **Multivariate**

**R studio**

**require(stats)/require(graphics)**

Assumptions of a linear regression:

Independent, identical normally distributed errors

1. The error term (residual/difference between the estimated Y and the real Y) should not be correlated with the Y variable
2. Each observation should be independent
3. The error term should be normally distributed (constant variance)

Submission: gitub

1. Open up Terminal

git add README.md

git commit –m “office hoursW 2:30-3:30 in IAB270F”

git push origin gh-pages/master (folder name)

git fetch upstream

git merge upstream/master