How social science can prepare you for a career in data science

Drew Linzer
Chief Data Scientist, Daily Kos

Emory University
December 3, 2015

✉️ drew@votamatic.org
🌐 votamatic.org
🐦 @DrewLinzer
What social scientists do

• Understand human behavior, history, and institutions
• Think critically about hard and complex problems
• Find new questions that people care about
• Research using statistics and quantitative data
Data science combines diverse skills

Source: Drew Conway
Data science combines diverse skills

Josh Wills
@josh_wills

Data Scientist (n.): Person who is better at statistics than any software engineer and better at software engineering than any statistician.
9:55 AM - 3 May 2012

1,255 Retweets
714 Likes
Data scientist salaries

Source: O'Reilly 2015 Data Science Salary Survey
Raw coding skills for working with data

Languages
- R
- Python
- SQL
- Excel

Environments and packages
- RStudio
- Jupyter
- Pandas
- scikit-learn

Version control and collaboration
- Git
- GitHub
Typical data workflow

1. **Acquisition**
   Internal databases, APIs, Web scraping

2. **Manipulation**
   Data munging/Data wrangling

3. **Exploration**
   Descriptive statistics, “Analytics,” Data visualization

4. **Inference**
   Model building, Prediction, Machine learning,
   Pattern recognition, Experimentation (A/B Testing)

5. **Collaboration**
   Communication with managers and engineers,
   Writing maintainable, encapsulated code
Typical data workflow

In Data Science, 80% of time spent prepare data, 20% of time spent complain about need for prepare data.

6:47 PM - 26 Feb 2013

425 169
Getting a job

Talk to people
Build a GitHub portfolio
Publish your work online
Getting a job

- Talk to people
- Build a GitHub portfolio
- Publish your work online

Apply!
Contact me

✉️ drew@votamatic.org
🌐 votamatic.org
🐦 @DrewLinzer