

POLS/CS&SS 503:  
Advanced Quantitative Political Methodology  
**INTRODUCTION**

March 31, 2015

Jeffrey B. Arnold



# Introductions

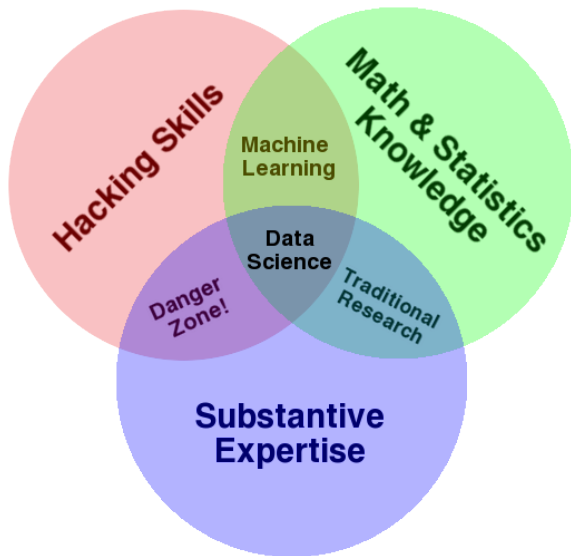
What are we doing here?

# What are the course objectives?

At the end of this course, you will be able to ...

1. estimate and interpret linear models
2. identify and explain the assumptions of the linear model
3. diagnose problems linear models and use appropriate solutions
4. represent statistical models in matrix algebra and compute basic matrix operations
5. use R to implement the statistical methods introduced here
6. take POLS/CSSS 510: MLE and other advanced CSSS courses

# What are we covering in this course?



# Why Linear Models?

# The Linear Model

$$y = X\beta + \epsilon$$

*Essentially, all models are wrong, but some are useful.*

*- Box, G. E. P. and Draper, N. P. (1987) Empirical Model Building and Response Surface*



# Deep Thoughts

# Recent Trends in (Quantitative Social) Science

- Reproducibility and Open Science
- NHST (Null Hypothesis Significance Test) doubts
- Causal inference
- Prediction
- Computation
- Data munging

Why the Tools we are Using?



# Why R?

- R is free (as in beer)
- R is free (as in freedom)
- R is a language
- Graphics and data viz capabilities
- Widely used

# R is popular and becoming more so

- r4stats.com, The Popularity of Data Analysis Software
- KDnuggets Languages for Analytics/Data Mining/Data Science
- Kagglers' Favorite Tools
- TIOBE Index for March 2015. R is #20 *all programming languages*
- The RedMonk Programming Language Rankings: January 2015. R is #13 of *all programming languages*

# R Markdown

- R + markup language
- Easier to combine code with results
- Increases reproducibility

# How to Learn R

And How to Succeed in this Course!

<https://youtu.be/JxwxefRAu70?t=30m1s>



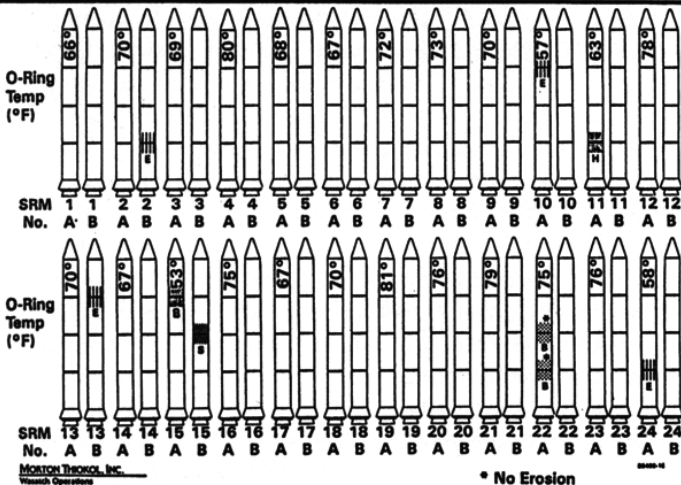
## Challenger Example

# Space Shuttle Challenger Exploded in 1986



# What Pattern Do You See?

## History of O-Ring Damage in Field Joints (Cont)



INFORMATION ON THIS PAGE WAS PREPARED TO SUPPORT AN ORAL PRESENTATION  
AND CANNOT BE CONSIDERED COMPLETE WITHOUT THE ORAL DISCUSSION

# Outline of this Course



# Review Syllabus on Course Website

Go to [http://pols503.github.io/pols\\_503\\_sp15](http://pols503.github.io/pols_503_sp15)

# References

- Drew Conway, “The Data Science Venn Diagram”,  
<http://drewconway.com/zia/2013/3/26/the-data-science-venn-diagram>,  
CC-BY-NC
- R Logo. Tobias Wolf.  
<http://developer.r-project.org/Logo/Rlogo.pdf> CC-SA
- [http://commons.wikimedia.org/wiki/File:Under\\_Construction.jpeg](http://commons.wikimedia.org/wiki/File:Under_Construction.jpeg)CC-BY-SA
- Challenger example inspired by Edward Tufte, \*The Visual Display of Quantitative Information\*
- Idea for using the Challenger example in this course from Christopher Adolph, “Introduction to the Course and R”, \*POLS/CSSS 221: Advanced Quantitative Political Methodology\*, Spring 2014.  
<<http://faculty.washington.edu/cadolph/503/topic1.pw.pdf>>