



Welcome to Data Science

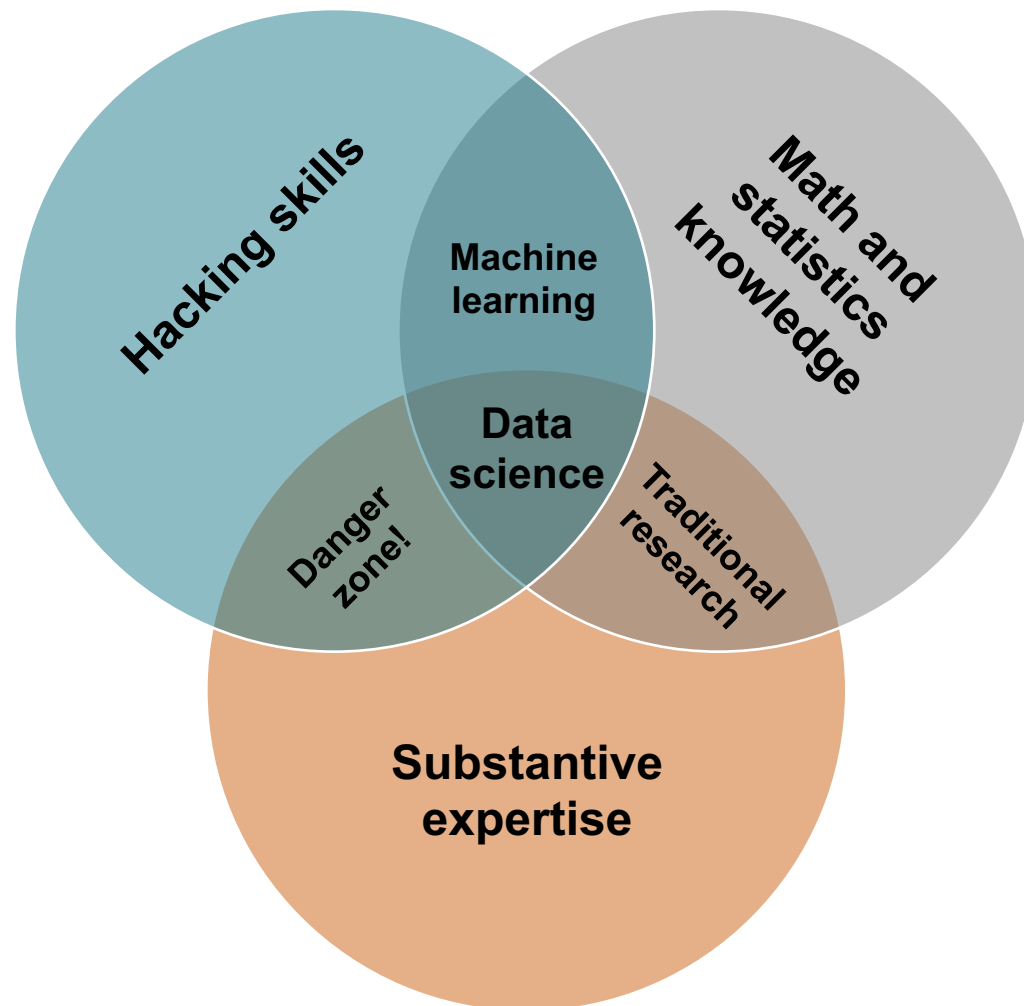
Will Doyle

What Is Data Science?

Elements of data science

1. Programming (“hacking”) skills
2. Statistics
3. Content knowledge

The Data Science Venn Diagram



This Class

- Introduction to *practice* of data science
- No previous expertise expected
- Teaches concepts through hands-on applications
- Uses real-world data and examples

Topics for the Class

- Getting data
- Analyzing data
- Presenting results

Topics: Getting Data

- Tools of the trade: R and Rstudio
- Flat data
- Data from the Web
- Databases

Topics: Analyzing Data

- Conditional means
- Linear regression
- Classification
- Cross validation
- Machine learning: clustering

Topics: Presenting Results

- Descriptive plots
- Scatterplots
- Plots and tables for classification
- Interactive graphics

Putting It All Together

- Final project: predict an outcome using a dataset
- This will challenge you to combine all of the skills you will master, including:
 - Getting data
 - Conducting an analysis
 - Communicating your results



VANDERBILT
PEABODY COLLEGE