Math Camp 2017

Christopher D. Grady

August 11, 2017

Resources:

- Basic R tutorial
  - Going to recommend Swirl: http://swirlstats.com/students.html
  - Could use codeschool: https://www.codeschool.com/courses/try-r
  - Could use Datacamp: https://www.datacamp.com/
  - Others?
- Github with R
  - http://happygitwithr.com/

Instructors:

- Faculty: Alicia.
- Graduate Students: Jae, Chris, Gustavo, Wei, Ekrem, Nuole

Topics:

- Algebra and Calculus
- Probability & Statistics
- General R
- Workflow (Organized & replicable data, Rmarkdown/LaTex, GitHub)
- Interpreting tables/figures

Style

- Interactive: ask the students to do things, rather than tell them how to do things.

Schedule

- Five days, two 2-hour sessions per day, start with 30 minute review of previous day.
- Review: 9:30am-10:00am
- Session1: 10:00am-12:00pm
- Lunch: 12:00pm-1:30pm
- Session2: 1:30pm-3:30pm

1. Day 1 (Aug 14): Algebra & calculus.
   - Professor Uribe-McGuire
2. Day 2 (Aug 15): Probability & statistics.
   - Ekrem and Lula
3. Day 3 (Aug 16) Introduction to R.

    - Gustavo and ??
4. Day 4 (Aug 17): Workflow – Rmarkdown, LaTex, GitHub.
    - Wei (Rmarkdown/Latex) and Chris (GitHub)
5. Day 5 (Aug 18) : Reading tables/plots & conclusion.
    - Ekrem (conclusion) and Jae (reading tables/plots)

Details

Day 1 (Aug 14): Algebra + calc

- Introduction (9:30am)
    - What we use math for:
    - Identifying an effect of X on Y
    - Conditions: when X –> Y?
    - Mechanism: how X –> Y?
    - Need mathematics to determine if Y increases as X increases (simple version).
- Module One (10:00am-12:00pm)
    - Algebra & calculus
- Module Two (1:30pm-3:30pm)
    - Algebra & calculus

Day 2 (Aug 15): Probability & Statistics

- Review (9:30am)
    - Review & questions from algebra & calculus
- Module One (10:00am-12:00pm)
    - Probability & statistics
    - Statistics as describing variation
    - Descriptive statistics (means, modes, medians, etc. . .)
    - Logic
    - More advanced topics
- Module Two (1:30pm-3:30pm)
    - Probability & statistics
    - Statistics as describing variation
    - Descriptive statistics (means, modes, medians, etc. . .)
    - Logic
    - More advanced topics

Day 3 (Aug 16): Introduction to R

- Review (9:30am)
    - Review & questions from probability & statistics
- Module One (10:00am-12:00pm)
    - Types of data
        * Characters, factors, numerics
        * Types of objects
        * Scalars, vectors, matrices, dataframes.
    - R programming as a language
        * Functions are verbs
        * Objects as direct objects (thing verb acts on)

- Module Two (1:30pm-3:30pm)
  - Basic functions
  - Structure of data
    - rows (units of analysis) and columns (variables/characteristics that describe units of analysis)
    - Indexing [row,column]
  - Troubleshooting with Google (stackoverflow)

Day 4 (Aug 17): Workflow – Rmarkdown, LaTex, GitHub

- Review (9:30am)
  - Review & questions from probability & statistics
- Module One (10:00am-12:00pm)
  - Workflow: what is workflow?
  - Data management & replicability
  - Tools for Workflow: Rmarkdown
    - code chunks & their options
    - plots in rmarkdown
    - inline code
    - knitting
  - Tools for Workflow: LaTex
- Module Two (1:30pm-3:30pm)
  - Tools for Workflow: GitHub
    - Repositories, initializing, cloning, status, pulling, committing, pushing.

Day 5 (Aug 18) : Reading tables/plots & conclusion

- Review
  - Workflow
- Module One (10:00am-12:00pm)
  - Regression Tables
    - coefficients
    - standard errors
    - statistical significance
  - Typical plots:
    - barplots
    - histograms
    - line graphs
    - others
- Module Two (1:30pm-3:30pm)
  - Goal of methods:
    - detect effect of X on Y
    - conditions under which X effects Y
    - mechanism through which X effects Y
  - How do algebra, calculus, and statistics help us do social science?
  - What is a statistical model?
  - Why do we want to use statistical models?

Modified/Shortened Schedule

- Three days, short sessions.
- Review: 9:30am-10:00am

- Session1: 10:00am-12:00pm
- Lunch: 12:00pm-1:30pm
- Session2: 1:30pm-3:30pm

1. Day 1 (Aug 24): Algebra & calculus.
   - Professor Uribe-McGuire
2. Day 2 (Aug 25): Probability & statistics + reading tables/plots
   - Ekrem and Lula, Jae
3. Day 3 (Aug 26) R, workflow (Rmarkdown, LaTex, GitHub basics), conclusion.
   - Gustavo, Wei, Ekrem

Details

Day 1 (Aug 24): Algebra + calc

- Introduction (1:30pm): math camp, why we did it, how it will prepare them. who is teaching it, any questions?
- **Module Zero: Point of it all**
  - What we use math for:
  - Identifying an effect of X on Y
  - Conditions: when X –> Y?
  - Mechanism: how X –> Y?
  - Need mathematics to determine if Y increases as X increases (simple version).
- **Module One (1:40pm-3:30pm)**
  - Algebra & calculus
- **Module Two (3:45pm-5:30pm)**
  - Algebra & calculus

Day 2 (Aug 15): Probability & Statistics, Reading tables & plots

- Review (9:30am)
  - Review & questions from algebra & calculus
- **Module One (9:45am-12:00pm)**
  - Probability & statistics
  - Statistics as describing variation
  - Descriptive statistics (means, modes, medians, etc. . .)
  - Logic
  - More advanced topics
- **Module Two (1:30pm-3:30pm)**
  - Probability & statistics
  - Statistics as describing variation
  - Descriptive statistics (means, modes, medians, etc. . .)
  - Logic
  - More advanced topics
- **Module Three (3:45pm-5:30pm)**
  - Reading tables & plots
  - Regression Tables
    - coefficients
    - standard errors
    - statistical significance
  - Typical plots:
    - barplots

        * histograms
        * line graphs
        * others

Day 3 (Aug 16): R, Workflow, Conclusion

- Review (9:30am)
  - Review & questions from probability & statistics + reading tables/plots
- **Module One (9:45am-12:00pm)**
  - Types of data
    * Characters, factors, numerics
    * Types of objects
    * Scalars, vectors, matrices, dataframes.
  - R programming as a language
    * Functions are verbs
    * Objects as direct objects (thing verb acts on)
  - Basic functions
  - Structure of data
    * rows (units of analysis) and columns (variables/characteristics that describe units of analysis)
    * Indexing [row,column]
  - Troubleshooting with Google (stackoverflow)
- **Module Two (1:30pm-3:30pm)**
  - Workflow: what is workflow?
  - Data management & replicability
  - Tools for Workflow: Rmarkdown
    * code chunks & their options
    * plots in rmarkdown
    * inline code
    * knitting
  - Tools for Workflow: LaTex
  - Tools for Workflow: GitHub
    * Repositories, initializing, cloning, status, pulling, committing, pushing.
- **Module Three (3:45pm-5:30pm)**
  - Conclusion
    * Goal of methods:
    * detect effect of X on Y
    * conditions under which X effects Y
    * mechanism through which X effects Y
  - How do algebra, calculus, and statistics help us do social science?
  - What is a statistical model?
  - Why do we want to use statistical models?
  - Asking political science questions