

MPP-E1180 Lecture 9: Automatic Tables and Static Visualisation

Christopher Gandrud

10 November 2014

Objectives for the week

- ▶ Assignment 3
- ▶ Review
- ▶ Static results presentation
 - ▶ Automatic table creation
 - ▶ Zelig simulations for showing results
 - ▶ ggplot2 for general graphing
 - ▶ ggmap for mapping

Assignment 3

Purpose: Gather, clean, and analyse data

Deadline: 14 November 2015

You will submit a GitHub repo that:

- ▶ Gathers web-based data from at least **two sources**. Cleans and merges the data so that it is ready for statistical analyses.
- ▶ Conducts basic descriptive and inferential statistics with the data to address a relevant research question.
- ▶ Briefly describes the results including with dynamically generated tables and figures.
- ▶ Has a write up of 1,500 words maximum that describes the data gathering and analysis and uses literate programming.

Assignment 3

Note: I will be traveling/at a conference/not able to check my email much on the **13th** and **14th**.

So try to ask all of your **questions by the 12th (Wednesday)**.

I will have **normal office hours** on Wednesday.

Review

- ▶ What is the basic R syntax for a regression model?
- ▶ What is a model function?
- ▶ How do you find a confidence interval for a parameter point estimate (both mathematically and in R)?
- ▶ What is a good way to interpret and present results from a logistic regression?

Motivation

Today we will learn how to **communicate your research findings** with automatically generated tables and static plots.

Why automatically generate?

- ▶ **Saves time:** don't have to re-enter numbers by hand into a table or restyle a graph.
- ▶ Easier to **find and correct errors:** all source code that created all tables and figures is linked and output updated when corrections are made.
- ▶ **More reproducible:** everything is clearly linked together.

Automatic table generation

There are a number of tools for automatically generating tables in R/R Markdown.

- ▶ `kable` in the `knitr` package
- ▶ `xtable` package
- ▶ `texreg` package
- ▶ `stargazer` package

Today

We will focus on `kable` and `texreg`.

- ▶ `kable` is a good, **simple** tool for creating tables from **data frames** (or matrices).
- ▶ `texreg` is useful for creating more complex tables with **regression model output**.

kable example: predicted probabilities

```
# Load data
URL <- 'http://www.ats.ucla.edu/stat/data/binary.csv'
Admission <- read.csv(URL)

# Estimate model
Logit1 <- glm(admit ~ gre + gpa + as.factor(rank),
              data = Admission, family = 'binomial')

# Create fitted data
fitted <- with(Admission,
              data.frame(gre = mean(gre),
                          gpa = mean(gpa),
                          rank = factor(1:4)))
```


kable example: predicted probabilities