# LING 410X: Language as Data

## Semester: Spring '18

Instructor: Sowmya Vajjala

Iowa State University, USA

11 Jan 2018

# Class outline

- Familiarizing yourself with R
- Doing R tutorial by installing "Swirl" library.
- Installing R and R Studio on your computers
- Assignment 1 description

# R and RStudio

- R is the actual programming language, which we use to write code and which does all the processing.
- R studio is a "development environment", which is like an interface to R, making it more convenient to use.
- Installing both is easy - you can find this online and choose according to your operating system (windows, macos, linux variants etc)
- Links are in the next slide.

# Installations

- On lab computers, R and R studio are already installed. You can open Rstudio by going to Applications and clicking Rstudio.
- For personal laptops: You have to first download R and then Rstudio. Choose R version 3.4 and above.
  - download R: `https://cran.rstudio.com/`
  - download R Studio: `https://www.rstudio.com/products/rstudio/download/`

Note: R Studio uses and needs R. Install R first.

# Swirl - R package

- an R package is a collection of R programs which you can use while working with your own data in R.
- a collection of such packages is called a library. I will use package and library interchangeably.

# Swirl - R package

- an R package is a collection of R programs which you can use while working with your own data in R.
- a collection of such packages is called a library. I will use package and library interchangeably.
- Swirl is a R package that contains various interactive learning tutorials on topics related to R.
- We will be working with this for the first 2 weeks to learn the basic vocabulary and grammar of R language.

# Installing Swirl

- Open R-studio, Go to Tools $->$ Install packages, and search for swirl in the prompt. Click install.
- Alternatively, you can type install.packages("swirl") in the console of Rstudio (usually seen in bottom left panel).
- Doing one of these will install swirl library into your computer's R.
- Note: if you are using lab computers, your installations and changes will be lost once you logout.

# using Swirl

- Once you finished the installation, enter library("swirl") in the console of R studio.
- If you see an error, it means you did not install swirl successfully. Let me know.
- If you successfully installed swirl, entering the library(swirl) should show you a prompt like this:
"Hi! Type swirl() when you are ready to begin."
- If you see that, you are ready to begin.
- Let me walk you through for the first and second lesson, and you can continue doing other lessons after that.

# Swirl Tutorial

There are several small lessons inside the basic R programming course in Swirl. I want you to do atleast 2–3 lessons today. Try to finish up to Lesson 5 or 6, whenever you get time, and the rest, we can learn as we go. You are welcome to do as many exercises as you want, ofcourse!

# Assignment 1 Description

- 2 questions, 10 marks.
- Deadline: 27 January, midnight
- first question: literature review - nothing about R.
- second question: using R to process strings - you should finish the swirl() tutorial first and also follow an additional tutorial document I prepared. We will have time today, and on next thursday to familiarize ourselves with R.

# Resources to learn or get help with R

- `https://www.r-project.org/help.html`
- `https://www.r-bloggers.com`
- R Programming for Data Science - free ebook by Roger Peng (available on leanpub.com, uploaded to Canvas)
- Coursera R courses by Johns Hopkins University (focus is on statistical analysis, not text)
- Discuss with classmates, talk to me during office hours.
- Google search with the error messages you get or with your general questions
- for Linguistics students: Look for a R book by Haraald Baayen (again, focus is more on statistics for linguists)
- Other, once you are comfortable with R: `https://www.tidytextmining.com/tidytext.html`

Please fill up the questionnaire