

Stat 491CI / 697P: Project Course

MWF 12:20-1:10, LGRT 1114

www.math.umass.edu/~jstauden/projcourse.htm

Professor John Staudenmayer, jstauden@math.umass.edu

413 648 7248

Office hours: Thurs 10-11, Fri 3-4 or by appointment

You will need to know how to use R and regression models in order to take this course.

Course goals:

Overall: We are going to use statistical models to try to understand the rental property market in the Amherst area. This will involve 3 (+3) phases:

- 1) Write program to get data on properties and prices from rentals.com.
- 2) Learn to use ggplot2 to graphically explore how property characteristics influence price.
- 3) Fit regression models to predict price as a function of property characteristics.

4) Develop spatial models for price as a function of property characteristics and location.

5) Make a map of property prices (controlling for property characteristics).

6) Explore machine learning models to predict property price.

(All do 1-3, split into groups for 4-6.)

Specific Goals:

- Develop web data scrubbing skills (rvest library).
- Learn to make informative graphics (ggplot2 library).
- Reinforce regression modeling skills.
- Learn new methods: spatial, statistical learning, GIS, etc.
- Become a better R programmer.
- Practice speaking to a group about statistics.
- Create a technical writing sample.

Tentative Schedule:

Jan 23-Feb 3: Make dataset using rvest

Feb 6-Feb 17: Graphics: how property characteristics influence price.

Feb 20-Mar 24: Linear regression: how property characteristics influence price.

Rest of semester: proj 4-6 above.

Individual presentation dates will be determined.

First draft of project write up is due March 3rd. Suggested outline given out Feb 10.

Grading:

50% participation (presentations) / 50% project write up