

# TYP QR: Data Analysis and Visualization

## Fall 2014

**Time :** Mondays and Wednesdays 2:00PM - 3:20PM

**Instructor :** Te Rutherford (tet@brandeis.edu)

**Teaching Assistant :** TBA

**Course Website :** [www.typ-stats.com](http://www.typ-stats.com)

### Course Description

A survey of some of the most popular tools and techniques in statistics. The emphasis is on understanding the importance of using data to make decisions and convincing arguments. Students will gain hands-on experience in R statistical programming language and spreadsheet software for basic data analysis and manipulation. Through examples and practical exercises, students will learn how to summarize data in a visually appealing way, use graphs as a supporting argument, conduct a statistical test, and fit a linear regression model.

### Class format

Each class is a mix of discussion, lecture, and lab components. During lecture, we will discuss some examples and concepts at a higher level. During the lab component, students will conduct an analysis using statistical software during class to get practical experience with the tools. They will also be assigned a dataset to re-apply the technique that they just learn in class.

There is no textbook for this course. The print-out of class notes will be distributed, and the updated versions are available on the course website.

### Topics to cover

1. Introduction to Data, R, and Google Spreadsheet.
2. Summarize single variable real-valued data
3. Histograms with R and G Spreadsheet
4. Summarize single variable real-valued data with multiple groups
5. Summary tables, barplots, and overlaying histograms
6. Percentiles, quartiles, and boxplots
7. Statistical hypothesis testing
8. Bootstrap hypothesis testing
9. Correlation and linear regression

## **Academic Integrity**

Most of the individual work done in this class is on the computer, so it is very easy and possibly tempting to copy and paste other people's work and submit it. It is never acceptable to do that. If you have problems with any part of the assignment, talk to me or the TA as soon as possible.

You are expected to be familiar with and to follow the University's policies on academic integrity (see <http://www.brandeis.edu/studentlife/sdc/ai> ). Faculty may refer any suspected instances of alleged dishonesty to the Office of Student Development and Conduct, which actually happened before in this class in the previous years.

## **Grading**

Final grades are based on the following components:

1. Attendance and participation (40%)  
We will do a lot of hands-on exercise in class. It is extremely important that you show up to class and complete the exercise during class. It will be difficult to learn the material on your own time as the class is not based on a textbook.
2. Quizzes (15%)  
There will be a few short quizzes, which focus on the more theoretical concepts. The dates for the quizzes will be announced in class.
3. Labs (30%)  
Students will repeat some of the analysis done in class on the new datasets.
4. Final project (15%)  
Students will apply all of the techniques in data analysis learned in class to datasets of their choice and present their results in a written report and in an oral presentation.