due in 2day 19h

Submission Phase

1. Do assignment ☐ (/dataanalysis-001/human_grading/view/courses/294/assessments/5/submissions)

Evaluation Phase

2. Evaluate peers 🔒 (/dataanalysis-001/human_grading/view/courses/294/assessments/5/peerGradingSets)

3. Self-evaluate 🔒 (/dataanalysis-001/human_grading/view/courses/294/assessments/5/selfGradingSets)

Results Phase

4. See results 🔒 (/dataanalysis-001/human_grading/view/courses/294/assessments/5/results/mine)

☐ In accordance with the Honor Code, I certify that my answers here are my own work, and that I have appropriately acknowledged all external sources (if any) that were used in this work.

Save draft   Submit for grading

Data

The data for this assignment are the Samsung activity data available from the course website:

https://spark-public.s3.amazonaws.com/dataanalysis/samsungData.rda (https://spark-public.s3.amazonaws.com/dataanalysis/samsungData.rda%22%3ELink)

These data are slightly processed to make them easier to load into R. You can also find the raw data here:

http://archive.ics.uci.edu/ml/datasets/Human+Activity+Recognition+Using+Smartphones (http://archive.ics.uci.edu/ml/datasets/Human+Activity+Recognition+Using+Smartphones)

All of the columns of the data set (except the last two) represents one measurement from the Samsung phone. The variable subject indicates which subject was performing the tasks when the measurements were taken. The variable activity tells what activity they were performing.

Prompt

Your task is to build a function that predicts what activity a subject is performing based on the quantitative measurements from the Samsung phone. For this analysis your training set must include the data from subjects 1, 3, 5, and 6. But you may use more subjects data to train if you wish. Your test set is the data from subjects 27, 28, 29, and 30, but you may use more data to test. Be careful that your training/test sets do not overlap.

You should perform all of the steps in building a predictive model and describe your analysis in a report as explained below.

What you should submit

Your data analysis submission will consist of the following components:
1. The main text of your document including a numbered list of references. This can be uploaded either as a pdf document or typed into the text box (not both!). The limit for the text and references is 2000 words. Your main text should be written in the form of an essay with an introduction, methods, results, and conclusions section.
2. One figure for your data analysis uploaded as a .png, .jpg, or .pdf file, along with a figure caption of up to 500 words.

Reproducibility

Due to security concerns with the exchange of R code, you will no longer be asked to submit code to reproduce your analyses. I still believe reproducibility is a key component of data analysis and I encourage you to create reproducible code for your data analysis.

Submission Deadline

You must submit your data analysis by March 11, 2013 at 11:00AM UTC-5:00 (Baltimore time). No late days may be applied to the data analysis. Note that this is an extension of the original date posted on the class website.

Please either enter the body of your data analysis in the text box or upload a pdf file with your analysis. This file should both contain the main text of your analysis and the numbered list of references. It may be no more than 2000 words.

	Format			

Words: 0 / 2000

Upload file : [Choose File] No file chosen

Please upload the figure for your data analysis here. It must be in .png, .jpg, or .pdf format. Type in a corresponding figure caption (no more than 500 words).

Format

Words: 0 / 500

Upload file : Choose File No file chosen

In accordance with the Honor Code, I certify that my answers here are my own work, and that I have appropriately acknowledged all external sources (if any) that were used in this work.

Save draft   Submit for grading