due in 2wk 1d

Submission Phase

1. Do assignment ☐ (/dataanalysis-001/human_grading/view/courses/294/assessments/4/submissions)

Evaluation Phase

2. Evaluate peers 🔒 (/dataanalysis-001/human_grading/view/courses/294/assessments/4/peerGradingSets)

3. Self-evaluate 🔒 (/dataanalysis-001/human_grading/view/courses/294/assessments/4/selfGradingSets)

Results Phase

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

For this analysis you will use the loans data available from here:

https://spark-public.s3.amazonaws.com/dataanalysis/loansData.csv (https://spark-public.s3.amazonaws.com/dataanalysis/loansData.csv)
https://spark-public.s3.amazonaws.com/dataanalysis/loansData.rda (https://spark-public.s3.amazonaws.com/dataanalysis/loansData.rda)

There is a code book for the variables in the data set available here:

https://spark-public.s3.amazonaws.com/dataanalysis/loansCodebook.pdf (https://spark-public.s3.amazonaws.com/dataanalysis/loansCodebook.pdf)

Prompt

The data above consist of a sample of 2,500 peer-to-peer loans issued through the Lending Club (https://www.lendingclub.com/home.action (https://www.lendingclub.com/home.action)). The interest

rate of these loans is determined by the Lending Club on the basis of characteristics of the person asking for the loan such as their employment history, credit history, and creditworthiness scores.

The purpose of your analysis is to identify and quantify associations between the interest rate of the loan and the other variables in the data set. In particular, you should consider whether any of these variables have an important association with interest rate after taking into account the applicant's FICO score. For example, if two people have the same FICO score, can the other variables explain a difference in interest rate between them?

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 February 18th, 2013 at 7:00AM 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

Evaluation/feedback on the above work

Note: this section can only be filled out during the evaluation phase.

Does the analysis have an introduction, methods, results, and conclusions section?

[select field]

Are figures labeled and referred to by number in the text?

[select field]

Is the analysis written in clear and understandable English?

[select field]

Are the names of variables reported in plain language, rather than in coded names?

[select field]

Does the analysis report the number of observations/samples?

[select field]

Does the analysis report any missing data or other unusual features?

[select field]

Does the analysis include description and justification for data transformations?

[dropdown]

Does the analysis include a discussion of potential confounders?

[dropdown]

Are the statistical models correctly applied?

[dropdown]

Are estimates reported with appropriate units and measures of uncertainty?

[dropdown]

Are estimators/predictions appropriately interpreted?

[dropdown]

Does the analysis make concrete conclusions?

[dropdown]

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