

QMB 6315: Python for Business Analytics
College of Business
University of Central Florida
Spring 2025

Assignment 5

Due Friday, April 25, 2025 at 11:59 PM
in digital form in your GitHub repository.

Instructions:

Complete this assignment within the space on your *private* GitHub repo (not a fork of the course repo QMB6315S25) in a folder called assignment.05. In this folder, save your answers to Question 1 in a file called `credit_logit.py`, following the sample script demonstrated in class `python_logistic_regression.py` in the folder `demo.13.Classification` in the course repository. When you are finished, submit it by uploading your files to your GitHub repo using the approach outlined in Question 2. You are free to discuss your approach to each question with your classmates but you must push your own work.

Question 1:

Start with the script `python_logistic_regression.py` and modify it to produce a model using the data from `credit_data.csv`, both of which can be found in the folder `demo.13.Classification` in the course repository. The dataset includes the following variables.

D_i	=	an indicator of whether borrower i defaulted on a loan
$bmaxrate_i$	=	the maximum rate of interest on any part of loan i
$amount_i$	=	the amount funded on loan i
$bankcardutil_i$	=	the utilization rate on the borrower's other credit card accounts (the percentage of the balance divided by the credit limit)
$close_i$	=	an indicator that the borrower takes the option of closing the listing before it is fully funded (that is, they might settle for a partial loan)
$rating_i$	=	a categorical variable based on the credit rating of borrower i which is sorted into the following categories:
AA		indicates that the borrower's FICO score is greater than 760
A		indicates that the borrower's FICO score is between 720 and 759
B		indicates that the borrower's FICO score is between 680 and 719
C		indicates that the borrower's FICO score is between 640 and 679
D		indicates that the borrower's FICO score is between 600 and 639
$Missing$		indicates that the borrower's FICO score was unobserved

This dataset consists of 2377 loans provided on a peer-to-peer lending platform. The variables that you include in the model will depend on the particular decision that you are analyzing.

- a) Estimate a logistic regression model to predict default D_i as a function of the other variables.
- b) Copy and paste any printed results from the logistic regression model in (a) in the allocated space on the file `README.md`.
- c) Create a plot of the ROC curve for the logistic regression model in (a) and save the image it in a file `Logit_ROC_full.png` so that it displays in the file `README.md`. *[Note that the image has extension PNG, so that it will display on GitHub.]*
- d) So far, you have estimated a model by considering all available variables, including those that might be unknown at the time of the loan. Now assume that you are an investor considering the loan and must make a decision using only the information available to a potential lender. That is, estimate a logistic regression model without the variables $bmaxrate_i$ and $amount_i$.
- e) Copy and paste any printed results from the logistic regression model in (d) in the allocated space on the file `README.md`.
- f) Create a plot of the ROC curve for the logistic regression model in (d) and save the image it in a file `Logit_ROC_decision.png` so that it displays in the file `README.md`. *[You should find that the model is less accurate; however, it is more useful, since the prediction can be used to make a decision.]*

Question 2:

Upload your code to your GitHub repo using the interface in GitHub Desktop.

1. Save your file within the folder in your repository in GitHub Desktop.
2. When you see the changes in GitHub Desktop, add a description of the changes you are making in the bottom left panel.
3. Press the button “Commit to main” to commit those changes.
4. Press the button “Push origin” to push the changes to the online repository. After this step, the changes should be visible on a browser, after refreshing the page.