Columbia Business School
AT THE VERY CENTER OF BUSINESS™

Business Analytics

Homework 2

UP TO AND INCLUDING MODULE 4

This homework will be graded on effort, like every other homework in the class. You must submit your homework as **one single PDF file**, containing a clear description of everything you did, and including as many screenshots from Jupyter Notebook as required to illustrate the work you did. You *may not* submit any other files to support your solutions – anything we should know should be included in your report. This is as much a test of technical communication as technical skill.

In this homework we return to the Nomis data we looked at in class, which can be found in `Nomis_data.xlsx`. Additionally, we will focus only on the loans with the following characteristics:

- Used cars
- Borrowers with FICO scores between 684 and 712
- Term of 60 months
- Loan amounts between $17,800 and $25,000

Our goal is to use the data from this segment to *predict* the likelihood of acceptance for two new loans.

1. Use `statsmodels` to fit a logistic regression on the Nomis data. Use the borrower's FICO score, the loan's APR, as well as the monthly payment as parameters. *Note:* the monthly payment is not found in the data and will need to be calculated as shown in class.
2. For each of the three parameters, discuss if the coefficients found by the model are statistically significant, and whether the relationship of their parameter to the outcome is positive or negative.
3. Use the logistic model you fit in Part 1 to predict if the new loans found in `predict.xlsx` will be accepted.