CST 383: Intro to Data Science

Dr. Glenn Bruns

Missing data in campaign contribution data

Purpose. One of the most important and first steps in data preprocessing is to identify and treat missing data. The purpose of this lab is to help you develop skill in this step. In particular, I want you to become skilled in identifying missing data in its various forms, and to understand how to treat the missing data.

Instructions. Download <u>campaign-cleaning.py</u> and insert your solution code after each problem number. Please read the instructions at the top of the file.

File <u>campaign-cleaning-out.txt</u> is the output from my solutions to the problems. You can use this file to compare your output to my output. The data set for this problem is a sample from a larger data set. Please note that I will use a different sample when testing your code.

Submission. Submit your edited campaign-cleaning.py on iLearn.

Grading. Each problem is worth the same number of points.