Product Science Homework

Indeed Product Science, Data Science

April 2, 2019

1 Instructions

Given the data included in this archive, please answer the following questions. Include the code you used to answer them as well as an explanation of your thought process and any issues you encountered when working through the data.

You should include the file with your code and any other files with your answers compressed in a .zip file. If attaching a write-up, it should be in a .pdf file.

All answers should be anonymous. Please exclude all personally identifiable information information from folder or file names.

Any questions should be directed in an email to your recruiter.

*This data is not real and was invented for the purposes of this exercise.

2 Background

Indeed routes leads to the sales team based on the potential quality of each lead. The system predicts the potential value of each lead and assigns out the leads from highest to lowest probability until each sales representative has a full sales book.

The attached data set shows leads that were assigned and never assigned to a sales rep. The purpose of this exercise is to estimate the incremental impact the sales representatives had on revenue. In other words, how much more did these leads spend because there was sales intervention?

3 Questions

Question 1

How many leads are represented in this dataset? Describe both the assigned and unassigned populations. What is the average revenue of each group?

Question 2

What are the most important metrics to considering when answering the problem statement? Why?

Question 3

Analyze any existing relationship between account age and revenue.

Question 4

What is the incremental value of assigning a lead to the sales team?

Question 5 (Bonus Question)

Investigate the data however you wish and discuss any interesting insights you can find in the data. Don't feel pressured to spend hours on this.