

Research Topics

Statistics - Will Lieber, Wania Iftikhar Khan, AJ Tenser, Kami Akala

Research Topic 1

Physical therapy outcomes depending on different approaches to doing rehabilitation exercises.

Research Topic 2

Another possible topic is looking into housing loans in North Carolina. There is a large federal database which contains much of the information pertaining to loan transactions - from the loan details (type, amount, purchaser) to the demographic information of the borrower (location, race, gender, income). With over 180,000 observations per year, and data from 2007-2017, the Consumer Financial Protection Bureau database is a rich source to generate a regression model to gain insight and generate predictions.

We decided to look into this topic because banks are faced with the challenging task of either approving or denying a loan amount, and if banks were able to better compare a given loan to what would be expected (as suggested by a model) they would have the tools to properly address risk. Housing is a public good, so the data is easily available, and given the fact that everyone needs housing, there is a vast amount of data to analyze.

This model would be most useful for banks who would benefit from the risk analytics, but this model could also be used by the borrowers themselves to reason what would be the expected amount for them to ask to borrow. In short, this model would help both the lender and the borrower to make more data-informed decisions.

Research Questions

1. What is the total amount a borrower will request for a loan in North Carolina? Using location (county and/or city), property type, type of loan, race, gender, and median income, as well as possibly interaction effects, we envision creating a model to predict the expected amount a borrower will request in North Carolina, likely using the 2017 data.
2. How have demographic factors influenced the amount borrowers request for loans? This would use the aforementioned demographic variables but would use data from likely 2007 and 2017 to find what shifts have occurred.
3. How do loan characteristics vary by region? This would rely on the metropolitan area variable to create groups. It would also be interesting to look at the observations outside of these areas to compare the urban/rural loan requests (especially regarding FHA/VA types which signals information about the borrower).

[CFMD Data](#): this is a link to the data that we have referenced throughout this section - it contains the information about demographics and loan type. The wide range of variables and combination of both categorical and continuous variables will hopefully allow us to construct a strong model to create meaning.

Research Topic 3

Lead levels in water as related to incarceration rates in the area.