

Student Loan Analysis

Kemal Fidan, Logan Courtney, Chandler Lampe
December 4, 2020

Outline

- Motivation
- Approach
- Data
- Findings
- Issues
- Future Work

Motivation

- Students have loans (we're all students)
- Not many analysis on trends in the data
- Help upcoming students
- Uncover any inequalities in loans
- Still prevalent today



ELLE

How Joe Biden's Plans
to Tackle the Student
Debt Crisis Could Help
You

1 day ago



yahoo!finance

Student loan forgiveness
'doesn't solve the crisis
long term,' expert
explains

1 day ago



CGTN

The U.S. student-debt
crisis – and why it is
strange to China

10 hours ago

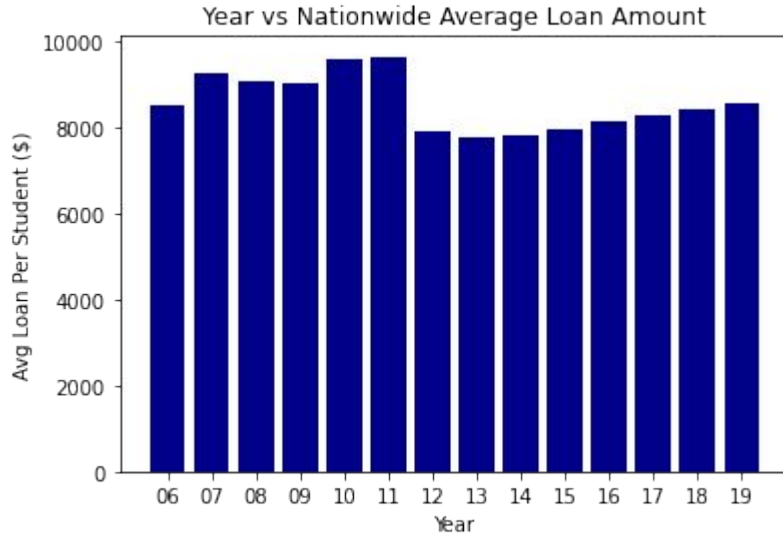
Questions

- Average Loan Per State
- Average Loan For Different Types of Institutions (Public, Private, etc)
- Average Loan Vs Year – look at change over year
- Average Loan Vs Cost of School
- Average Loan Vs Size of School
- Loan Differences for Rank of Institution
- Average Loan Vs ACT/SAT
- Average Loan Vs Region
- Average Loan Vs School's Gender/Race %
- Average Loan Vs Size of City

Approach & Data

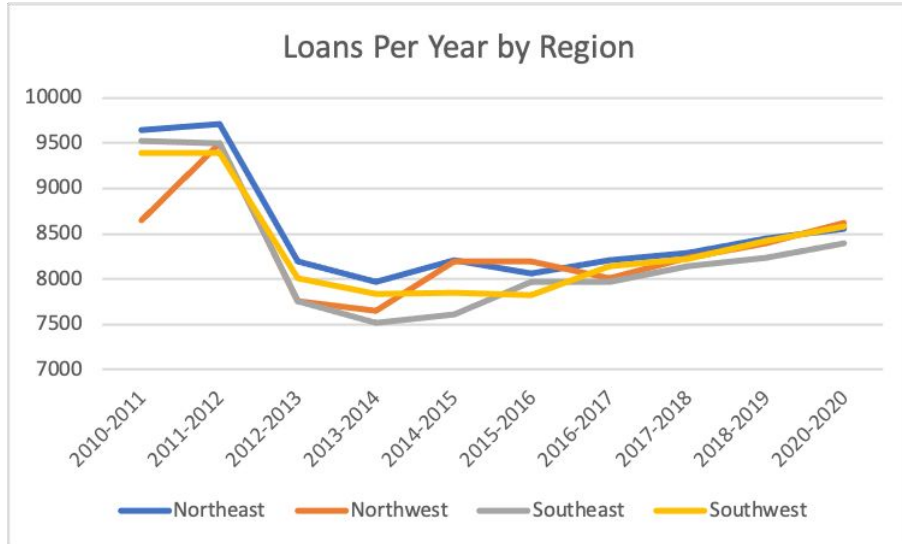
- Drafted questions
- Used data from the government's student loan office
- Aggregate data for analysis
- Preprocessing done at runtime:
 - Ignoring empty ('-') fields
 - Merging datasets

Findings



- Steady growth over time
- New loans vs outstanding loans
- Taking out loans vs paying them back
- What happened in 2012?
 - Student Loan Forgiveness Act, Election year

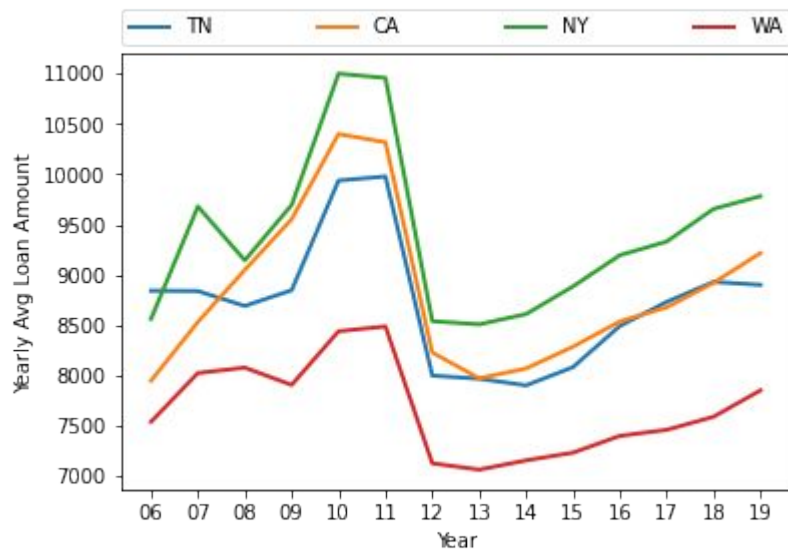
Findings



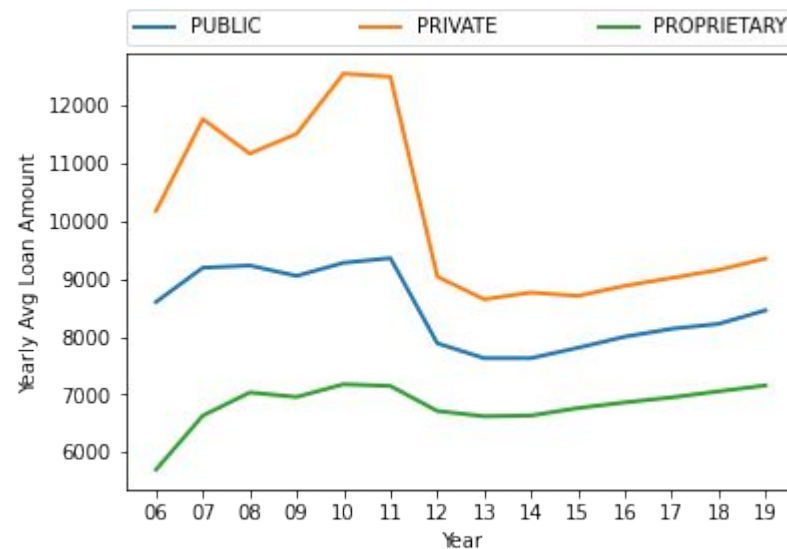
Does the region affect the amount of loans are taken out?

- Northeast remained to be the region with the most loans
- Southeast was on average the region with the fewest amount of loans

Findings

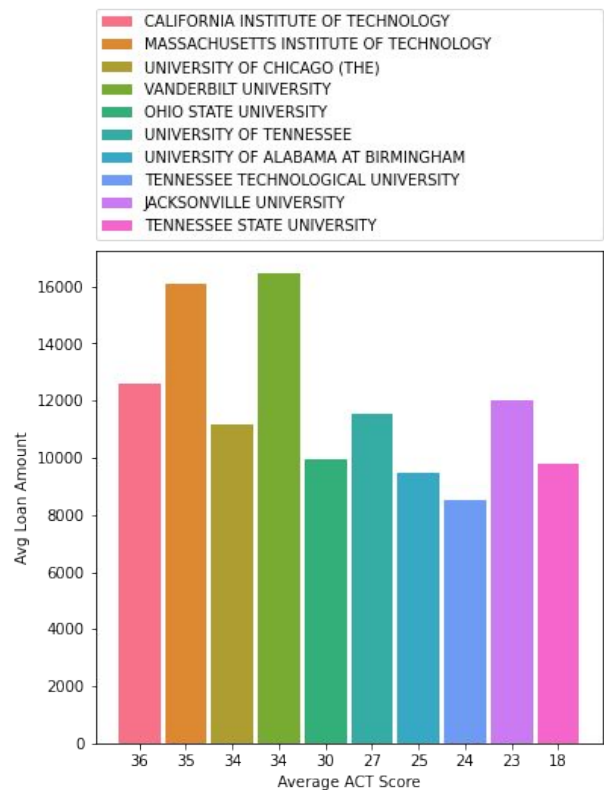


- CA & NY cost of living
- Generally better schools → more aid?



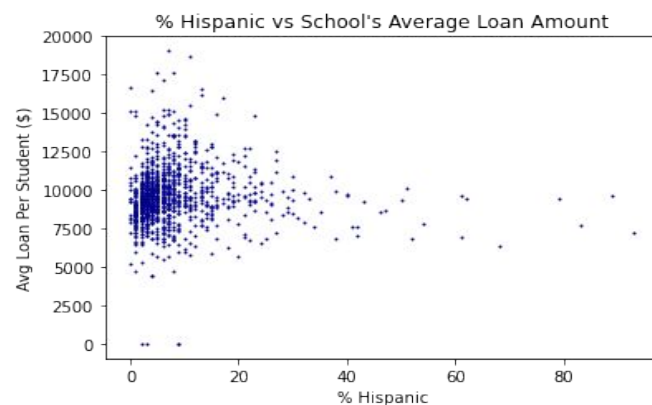
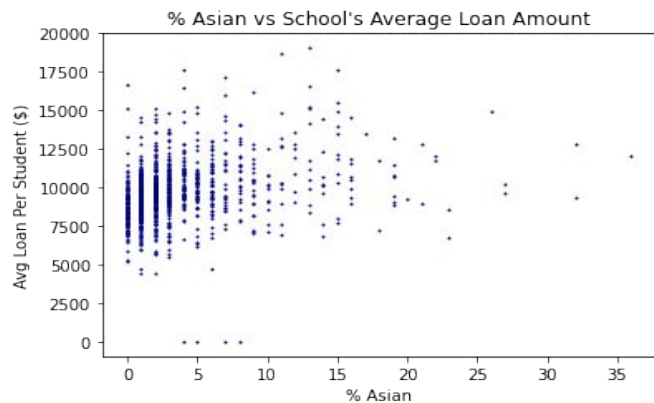
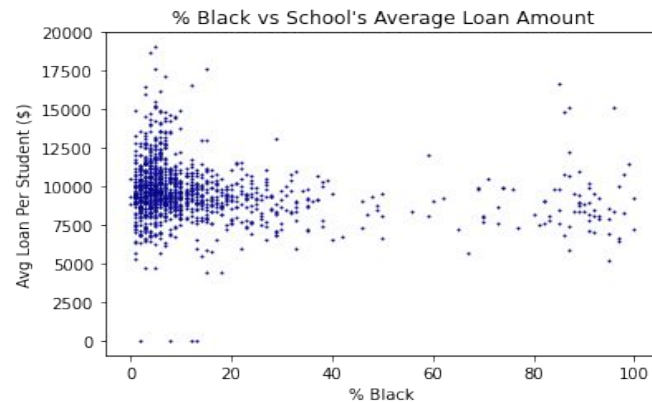
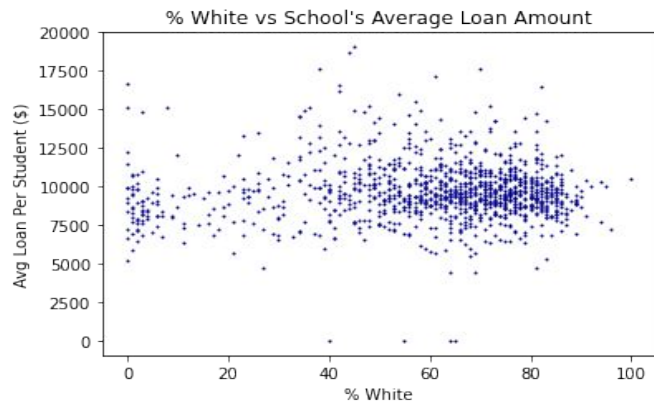
- Expected. Private = for-profit
- Quality of school vs scholarships offered

Findings

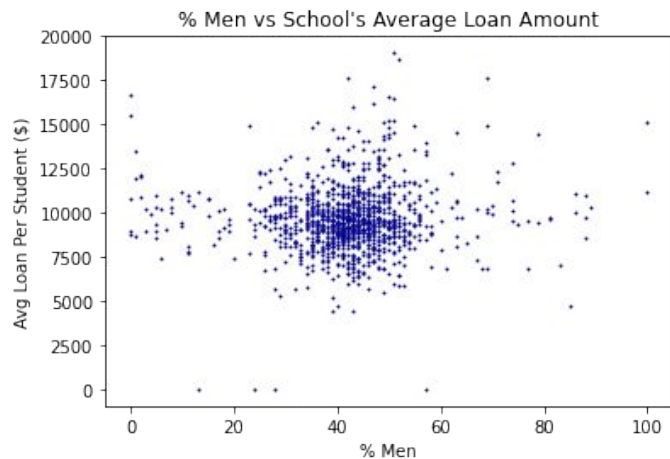


- This was tough to hypothesize:
 - smarter students probably go to better and more expensive school
 - might receive more aid in the process
- Lower ACT scores correspond to lower loan rates
- There are some outliers like MIT and Vanderbilt

Findings



Findings

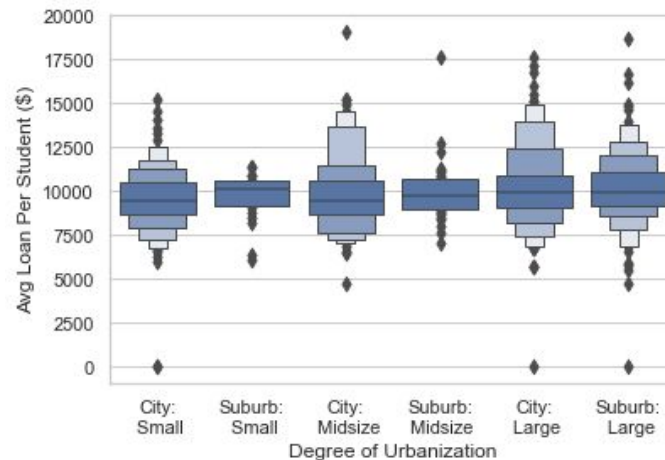


- No apparent trend

Small: 0-100,000 population

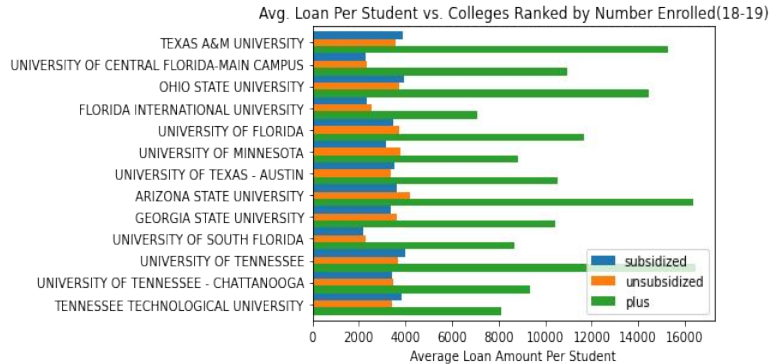
Medium: 100,000-250,000 population

Large: 250,000+ population



- Schools in suburban areas tend to have less of a spread
- All urban types have a median ~10k

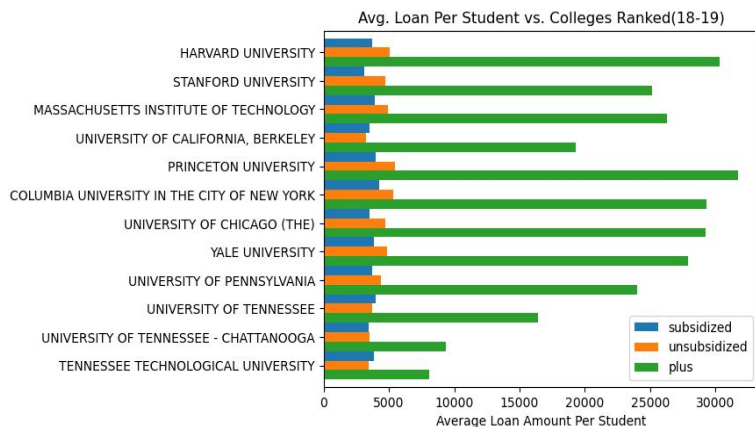
Findings



Does the amount of students enrolled in a college affect the average loan amount?

- More students = less loans?
- Subsidized/Unsubsidized loans of the high costing schools were relatively similar.
- Average plus loans are very sporadic, showing no correlation between enrollment and average loan amount.

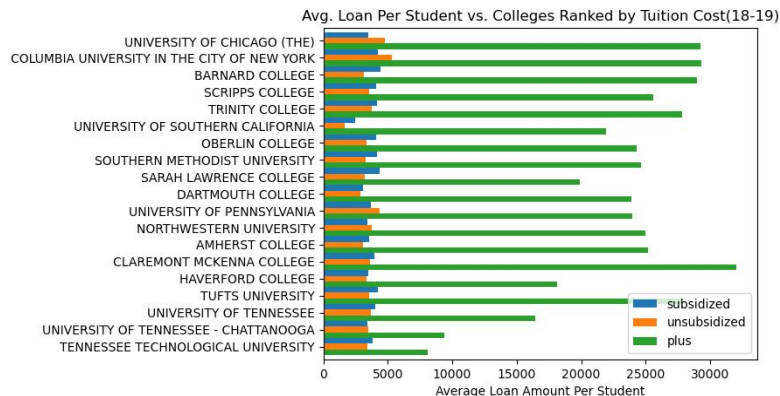
Findings



Does the college's rank correlate to the average loan amount?

- Higher rank = More loans or Higher rank = More scholarships
- Ranked list wasn't very extensive.
- A correlation could be present, but we couldn't fully confirm it without a larger ranked college list.

Findings



Does tuition cost reflect the amount students take out in loans?

- Higher cost = more loans
- Again we see subsidized and unsubsidized somewhat similar for every college.
- Most of the plus loans are significantly higher than the three reference colleges.
- Higher costing colleges seem to also have higher average student loan amounts.

Issues

- Finding datasets (demographics, ACT scores, etc)
- Combining datasets together
- Processing many excel files, split quarterly
- Difference in private and public university loan amounts added confusing outliers

Future Work

1. Keep data in a database
2. More info on student demographics, like race, gender, ethnicity, etc
3. Collect per-person info
4. Given information about the borrower, ML to predict student loan info
5. Site or article to hand info to the public