



MEASURING SUCCESS

**THROUGH THE
PERFORMANCE OF**

**AMERICA'S
POSTSECONDARY EDUCATION**



Bining Lu
Clarence Li
Sumiera Taqui
William Sit

QUESTION:

What college factors lead to successful students?



DEFINITIONS:

- **College**

- A four-year institution
- Status = operating
- Predominantly bachelor-granting

- **Factors**

- Admission rate, school ownership
- School selectivity (based on average SAT score)
- Retention rate, completion rate
- Full-time faculty rate, average faculty earnings

- **Success**

- Students' mean earnings (6 & 10 years after entry)

SOURCES:



1. **College Scorecard Data API** (from U.S. Department of Education)

- a. data.gov link = <https://collegescorecard.ed.gov/data/>
- b. Internal sources:
 - i. **Integrated Postsecondary Education Data System (IPEDS)**
 1. Admission rate, school ownership
 2. School selectiveness
 3. Retention rate, completion rate
 4. Full-time faculty rate, average faculty salary
 - ii. **U.S Department of Treasury**
 1. Students' mean earnings (6 & 10 years after entry)



DATA COLLECTION & CLEANSING



FINDING THE METRICS:

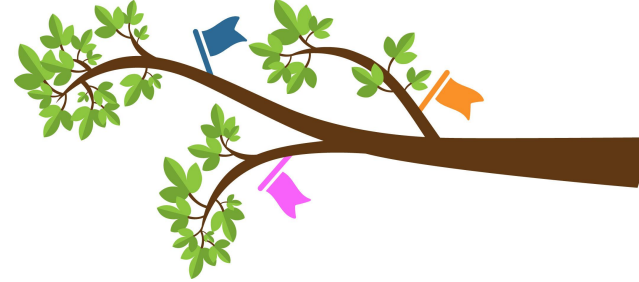
- **School**
 - School name, school ownership
 - Faculty rate, faculty salary
- **Student**
 - Retention rate
- **Completion**
 - Completion rate
- **Admissions**
 - Admission rate
 - SAT score
- **Earnings**
 - Student mean earnings 6 years after entry
 - Student mean earnings 10 years after entry



DATA COLLECTION & CLEANSING

- Created a dictionary to hold key column names and the corresponding request metric values
- Sent an API request to get the total number of response records to determine the number of iterations needed for obtaining the complete resultset since each request from the API source returns 100 records max
- Created a master DataFrame from the json response
- Updated "School Ownership" column with descriptive text value instead of the original numeric code values
- Formatted all columns containing numeric and percentage values

DATA BRANCHING



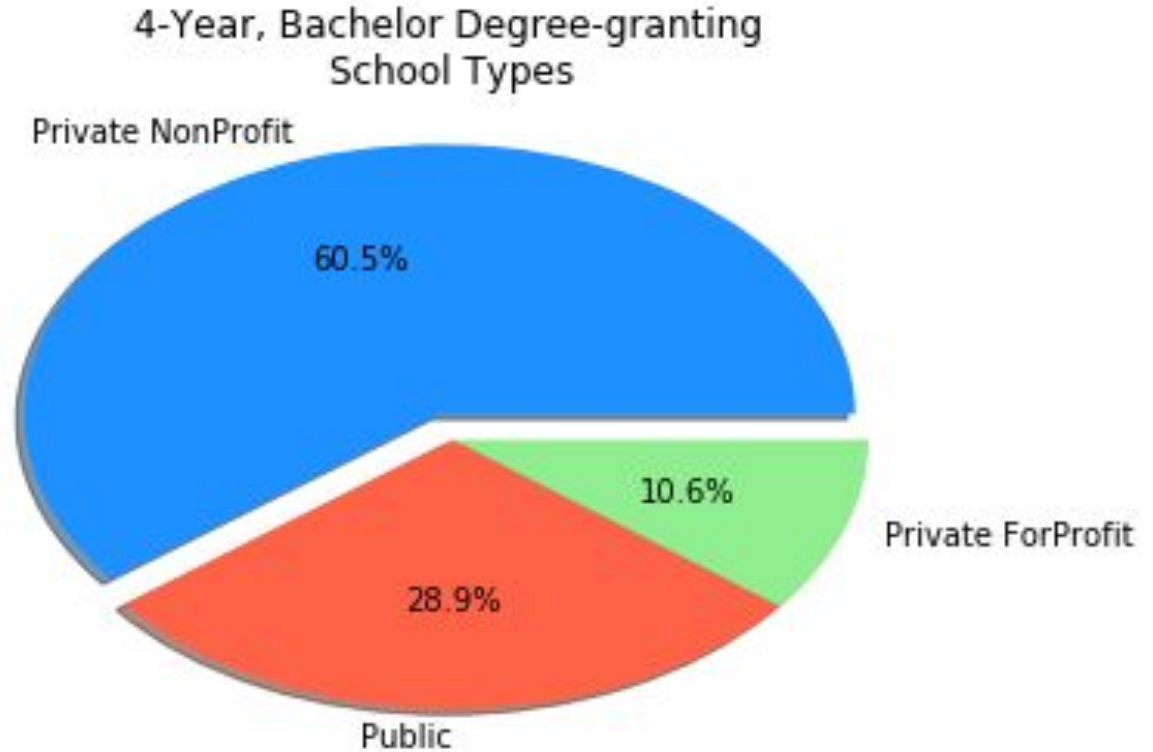
- Created branching DataFrames for each metric to compare against Earnings
 - **Metrics:**
 - School ownership, admission rate
 - School selectiveness
 - Retention rate, completion rate
 - Full-time faculty rate, average faculty salary
 - **Earnings:** Students' mean earnings (6 & 10 years after entry)
- Further cleansing of the sub DataFrames
 - Removed rows with null values
 - Binned data on certain metrics to study trends based on specific range of values
 - Grouped data by specific metrics and performed aggregated calculations
 - Data cleansing codes: **dropna()**, **loc()**, **groupby()**, **sort_values**, **style.format()**, etc.

DATA VISUALIZATIONS & ANALYSIS

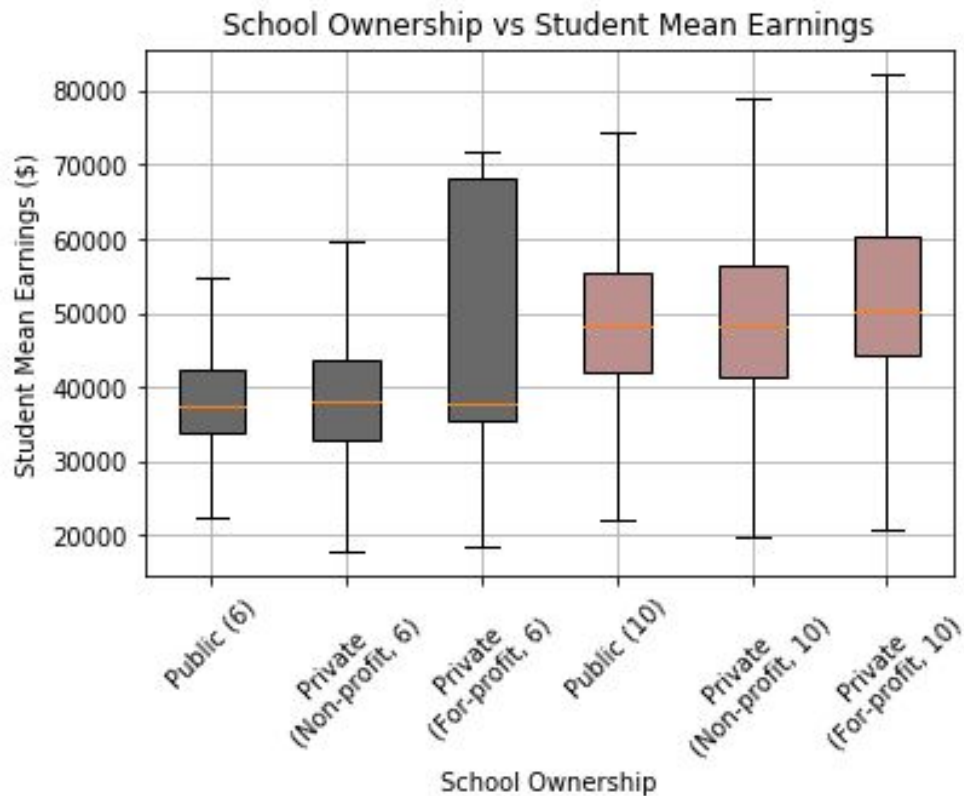


SCHOOL OWNERSHIP

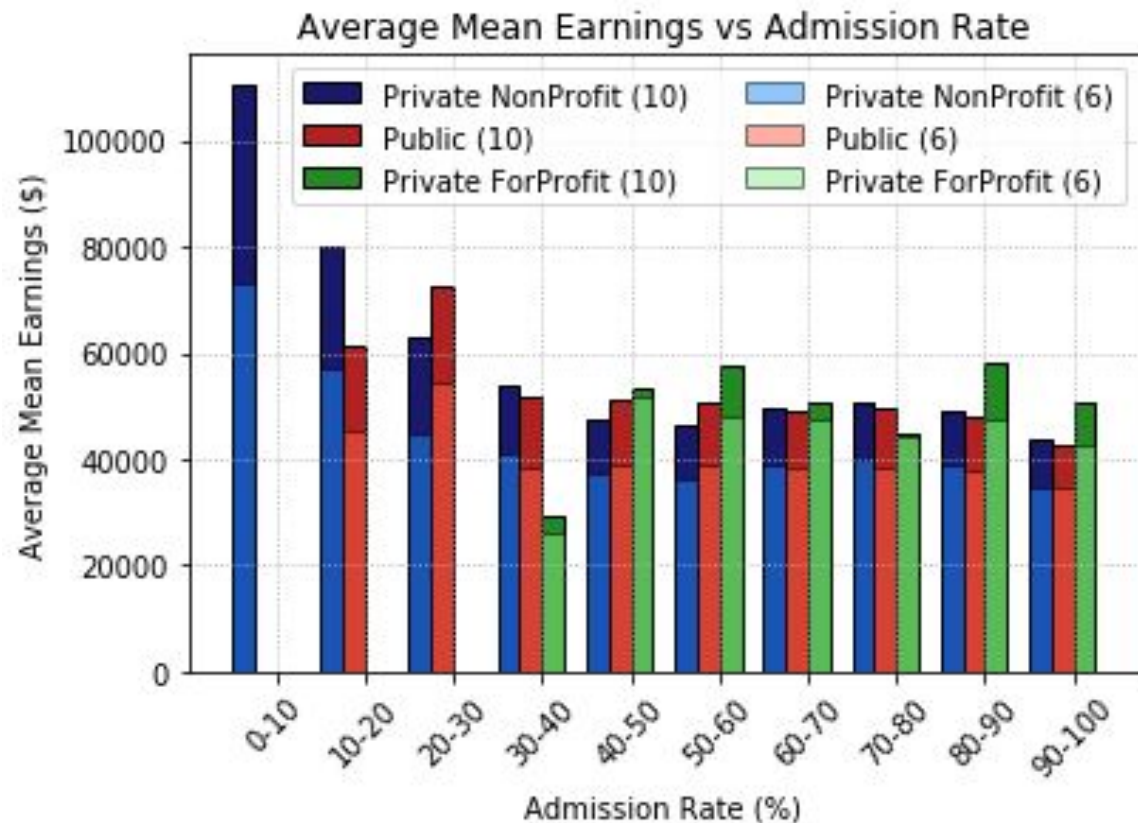
- **Private Non-Profit**
 - 1239 colleges
- **Public**
 - 593 colleges
- **Private For-Profit**
 - 217 colleges



EARNINGS DISTRIBUTION BY SCHOOL OWNERSHIP



ADMISSION RATES



COMPARING 6 AND 10-YEAR DATA

Private Non-profit

Admission Rate Range (%)	School Count	Mean Earnings Difference	% Difference
0-10	13	\$37,553.85	51.39%
10-20	26	\$22,992.31	40.10%
20-30	37	\$18,283.78	40.87%
30-40	50	\$13,018.00	31.87%
40-50	98	\$10,311.22	27.67%
50-60	157	\$9,971.97	27.40%
60-70	199	\$10,886.93	28.15%
70-80	181	\$10,649.17	26.62%
80-90	102	\$10,115.69	25.87%
90-100	77	\$8,764.94	25.21%

Public

School Count	Mean Earnings Difference	% Difference
0	\$0.00	0.00%
3	\$15,966.67	35.12%
8	\$17,837.50	32.69%
17	\$13,564.71	35.28%
42	\$12,269.05	31.57%
77	\$11,580.52	29.66%
98	\$10,758.16	27.92%
118	\$11,128.81	28.78%
103	\$10,100.00	26.62%
56	\$8,341.07	24.24%

Private For-profit

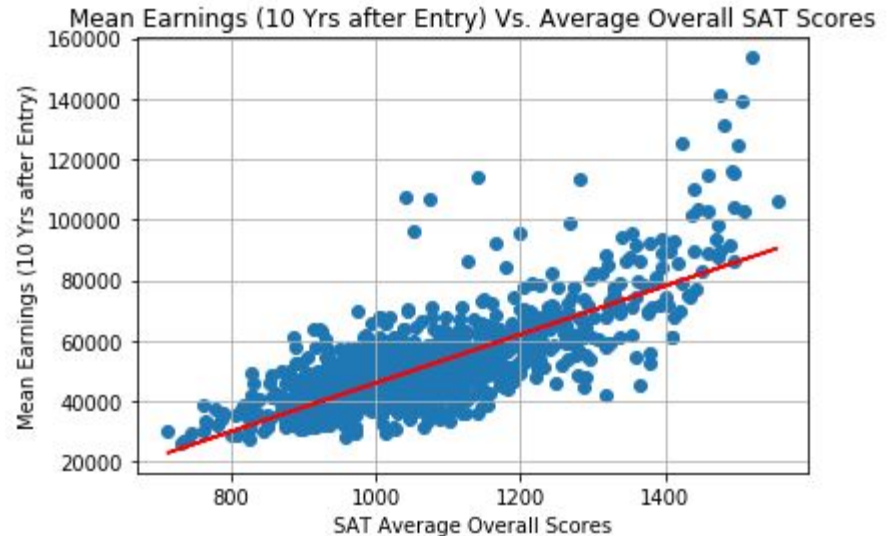
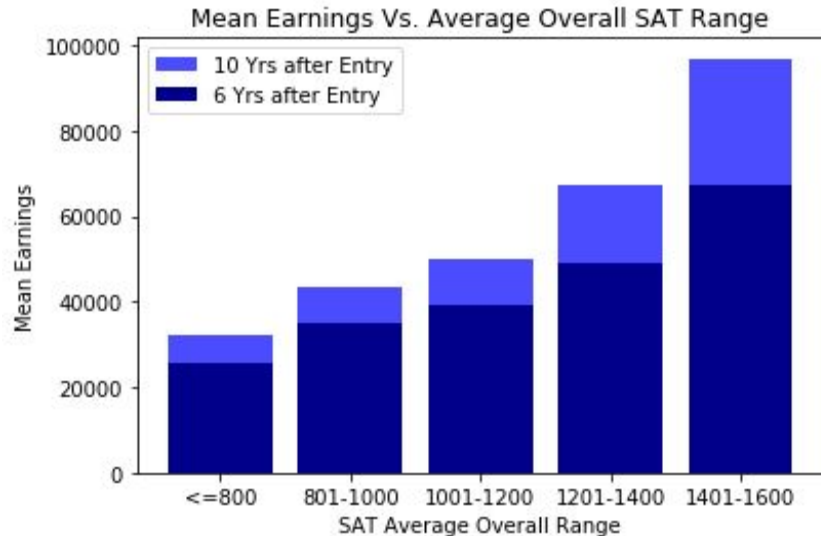
School Count	Mean Earnings Difference	% Difference
0	\$0.00	0.00%
0	\$0.00	0.00%
0	\$0.00	0.00%
1	\$3,300.00	12.64%
4	\$1,750.00	3.38%
4	\$9,975.00	20.88%
5	\$3,440.00	7.28%
10	\$360.00	0.81%
10	\$10,380.00	21.79%
20	\$7,900.00	18.54%

COLLEGE SELECTIVITY

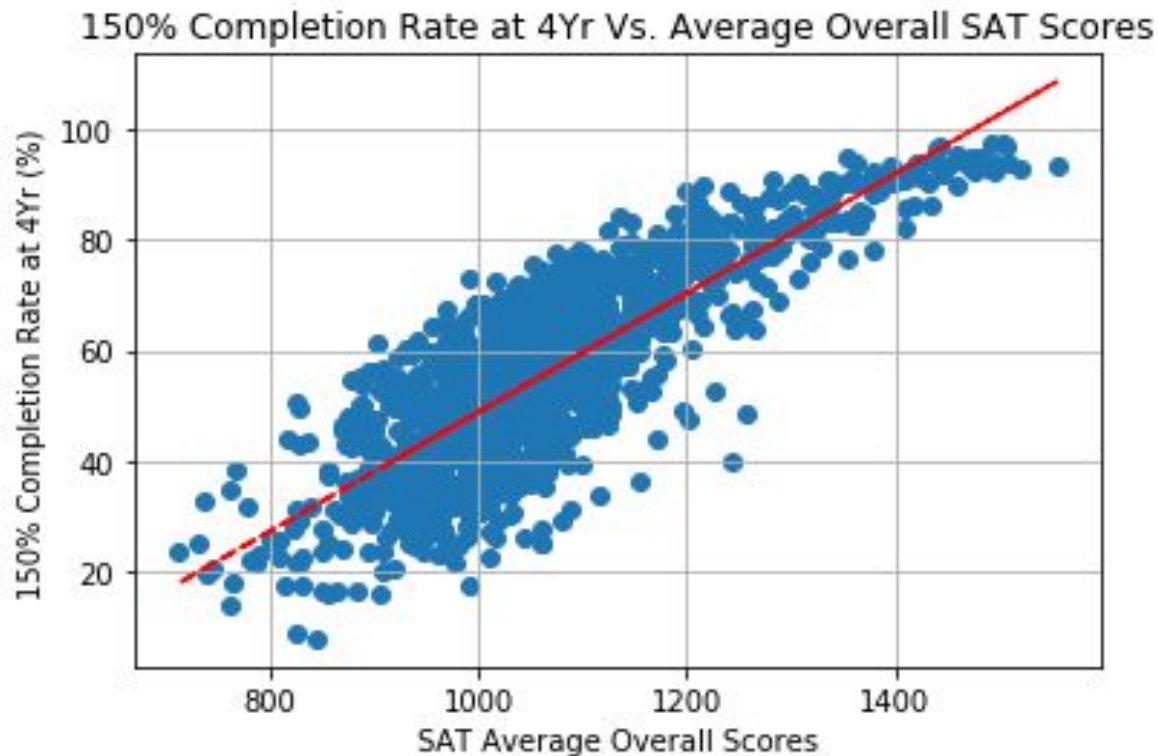
(Based on Average SAT Score)

College Selectivity and Earnings:

College selectivity has a positive correlation with the mean earnings of graduates.

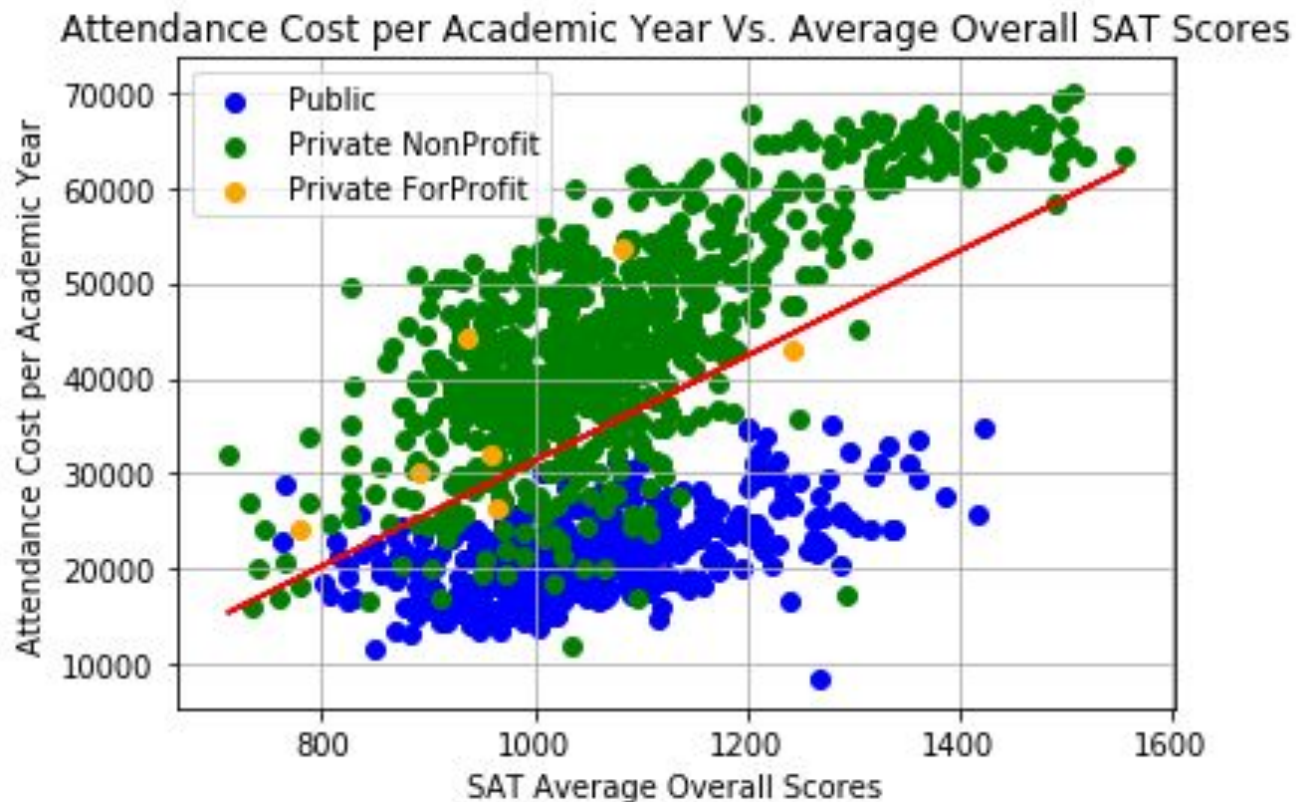


College Selectivity and Completion Rate:

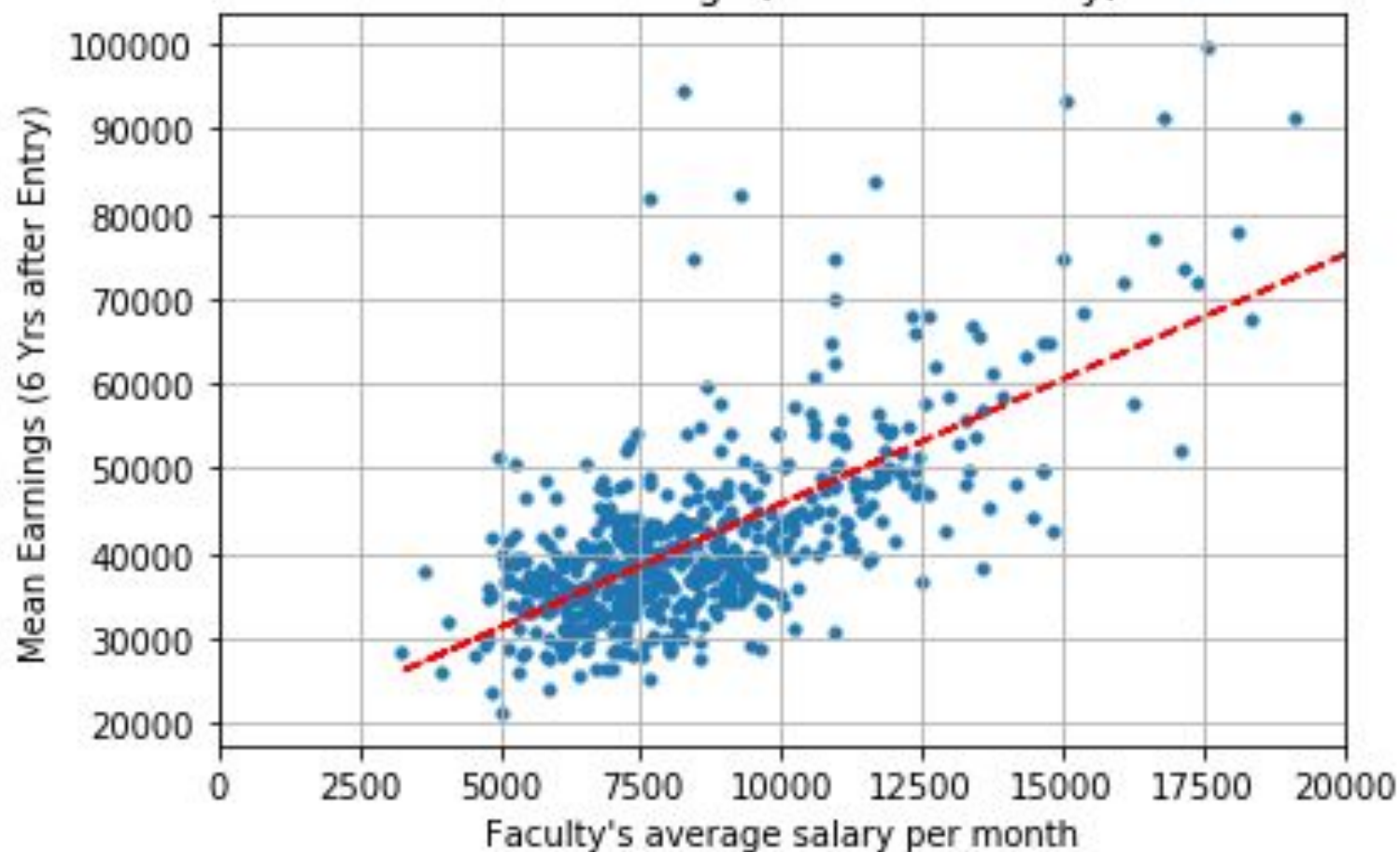


SAT Range	Count of Schools (%)
<=800	1.07
801-1000	31.90
1001-1200	52.76
1201-1400	11.13
1401-1600	3.13

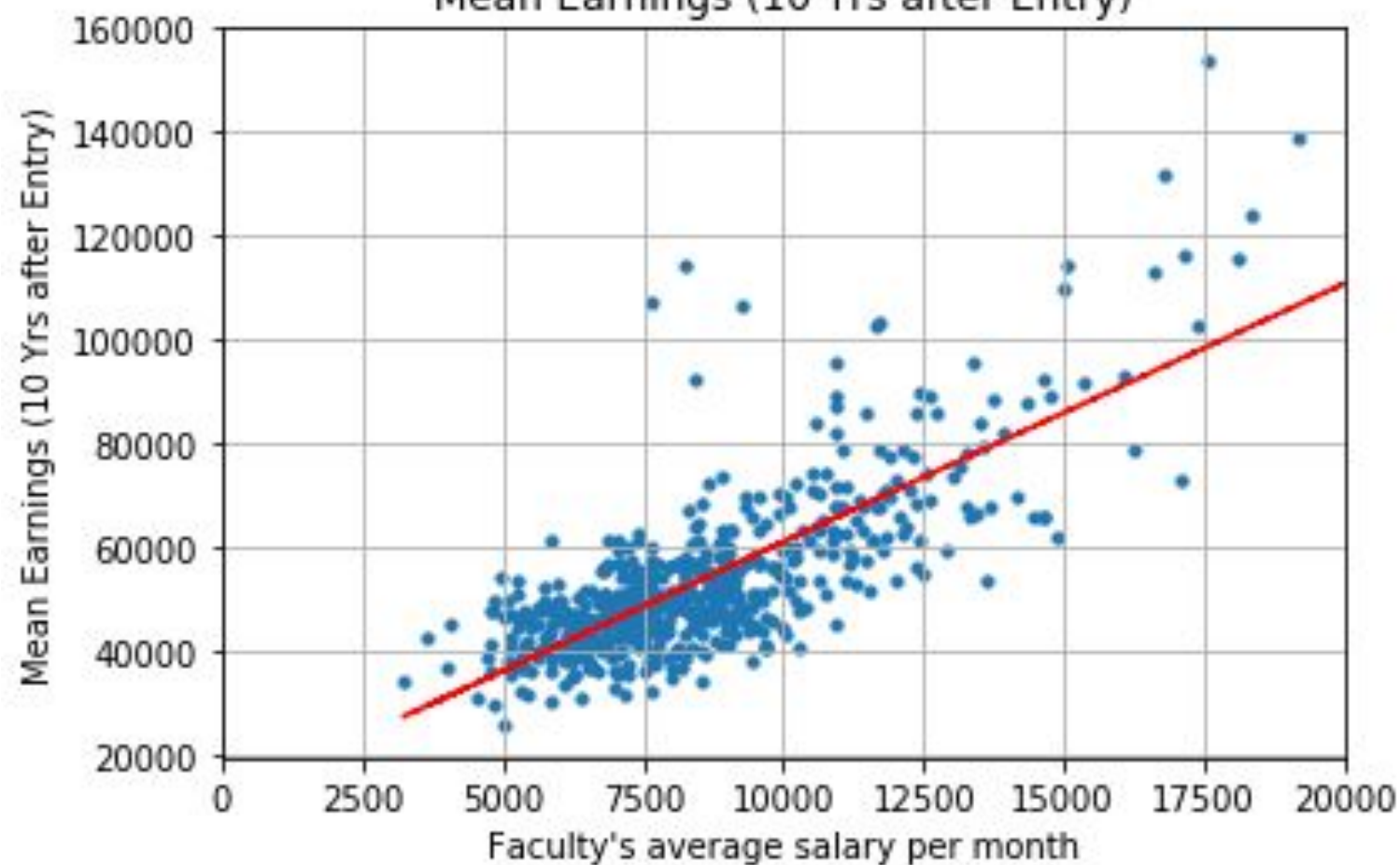
College Selectivity and Cost per Academic Year:



Faculty's average salary per month vs
Mean Earnings (6 Yrs after Entry)

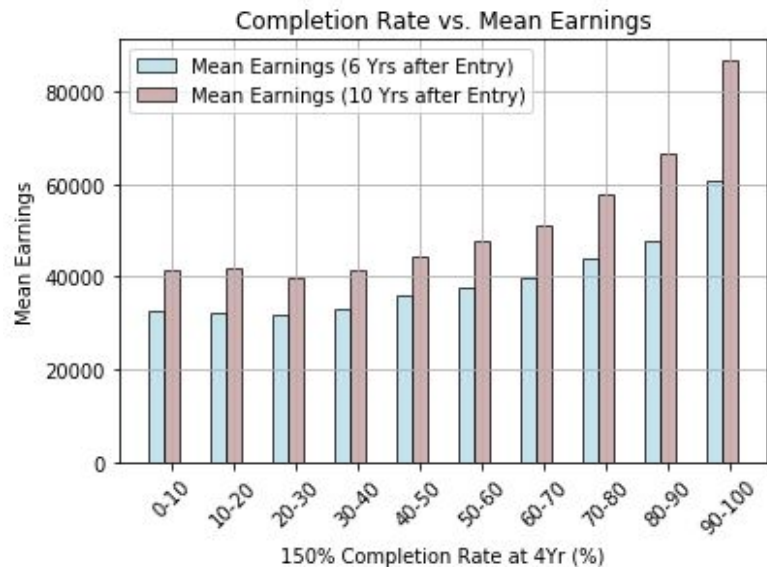
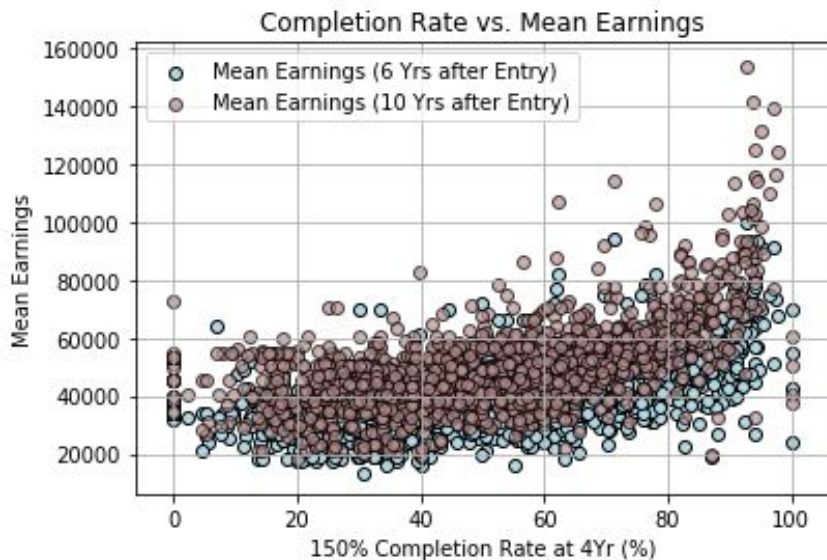


Faculty's average salary per month vs
Mean Earnings (10 Yrs after Entry)

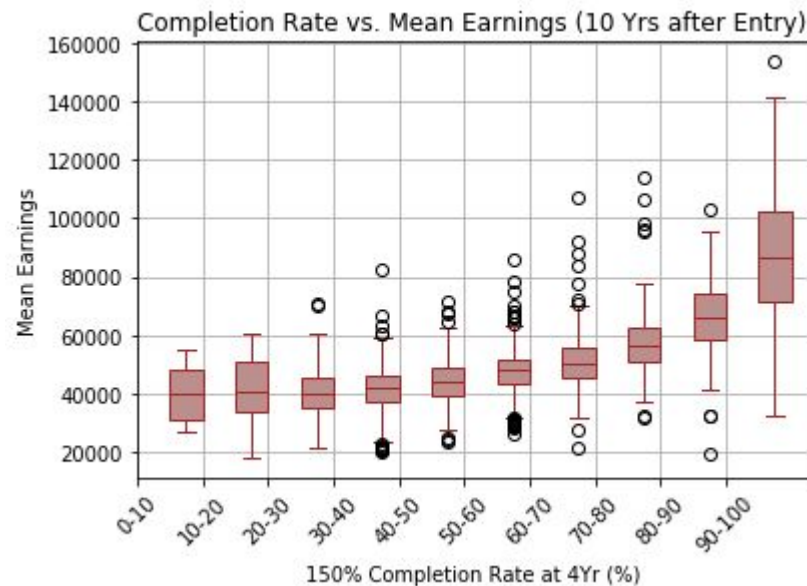
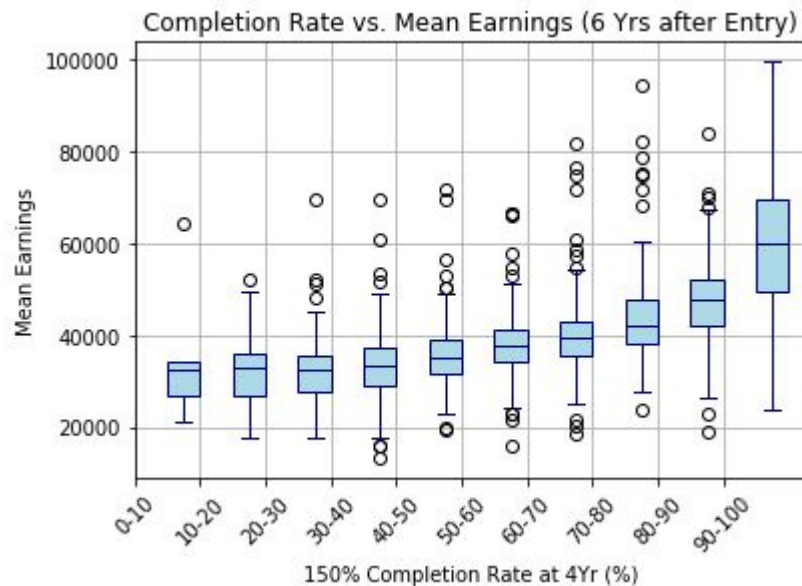


COMPLETION RATE VS MEAN EARNINGS

Completion rate: 150% Completion Rate at 4Yr (%)

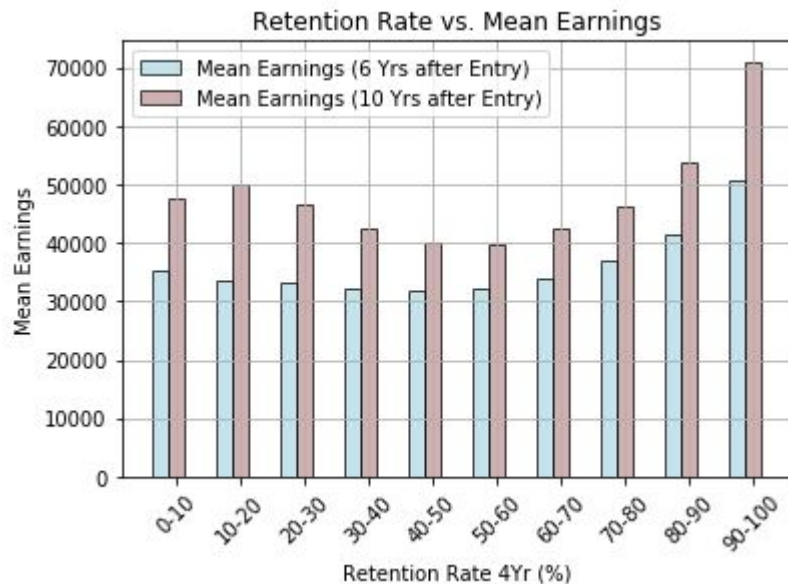
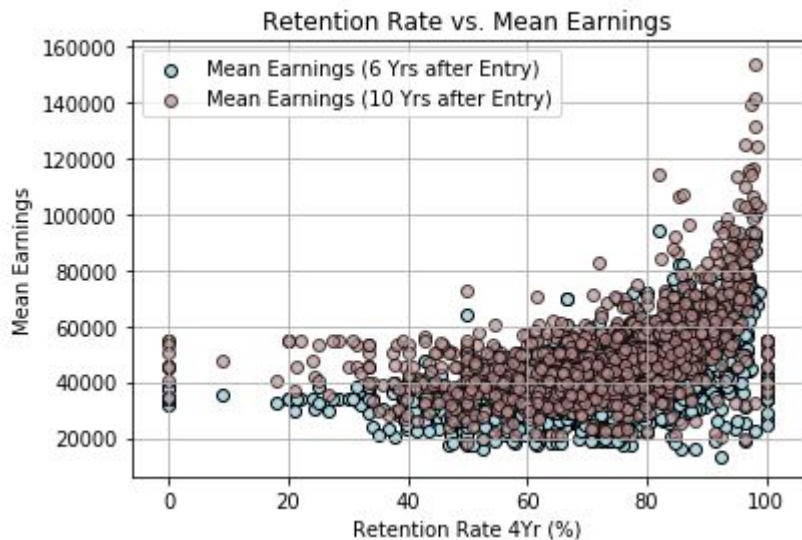


COMPLETION RATE

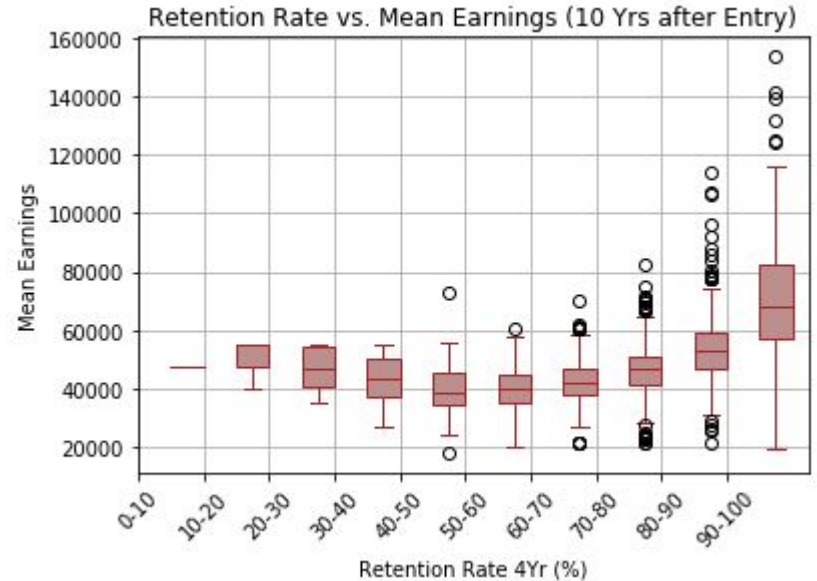
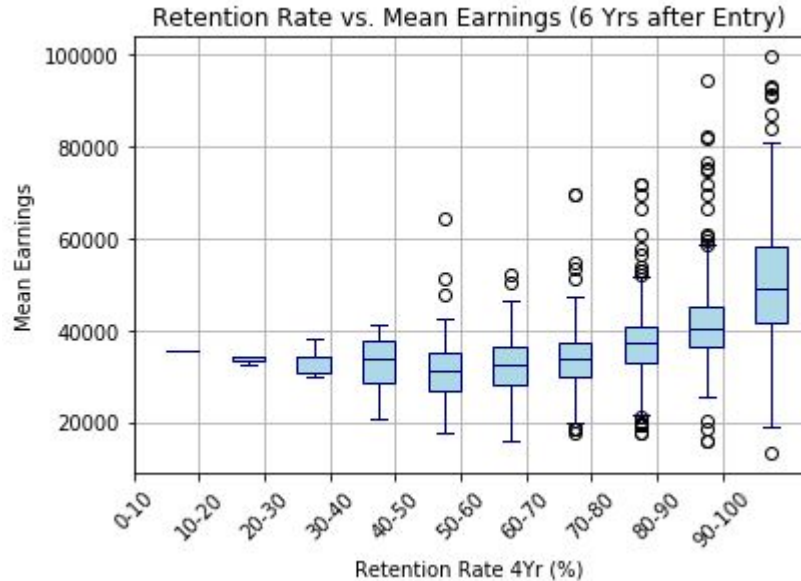


RETENTION RATE VS MEAN EARNINGS

Retention rate: Retention Rate 4Yr (%)



RETENTION RATE





CONCLUSIONS



TOP 5

School ID	School Name	School State	School Ownership	Admission Rate (%)	Mean Earnings (10 Yrs after Entry)	Mean Earnings (6 Yrs after Entry)
166683	Massachusetts Institute of Technology	MA	Private NonProfit	7.94	\$153,600.00	\$99,600.00
243744	Stanford University	CA	Private NonProfit	4.81	\$141,300.00	\$92,800.00
166027	Harvard University	MA	Private NonProfit	5.4	\$139,100.00	\$91,300.00
215062	University of Pennsylvania	PA	Private NonProfit	9.44	\$131,600.00	\$91,200.00
131496	Georgetown University	DC	Private NonProfit	16.85	\$125,200.00	\$81,100.00

BOTTOM 5

School ID	School Name	School State	School Ownership	Admission Rate (%)	Mean Earnings (10 Yrs after Entry)	Mean Earnings (6 Yrs after Entry)
458469	National University College-Ponce	PR	Private ForProfit	72.78	\$20,900.00	\$18,600.00
468723	National University College-Caguas	PR	Private ForProfit	76.81	\$20,900.00	\$18,600.00
241951	Escuela de Artes Plasticas y Diseno de Puerto Rico	PR	Public	83.52	\$20,600.00	\$13,300.00
197735	Yeshivath Viznitz	NY	Private NonProfit	91.47	\$19,600.00	\$19,000.00
187745	Institute of American Indian and Alaska Native Culture and Arts Development	NM	Public	100	\$18,000.00	\$17,800.00

Admission Rate(%) Mean Earnings (10 Yrs after Entry)

Max	100.00	\$153,600.00
Min	0.00	\$18,000.00
Mean	65.50	\$50,388.47

Private Non-profit TOP 5

School ID	School Name	School State	School Ownership	Admission Rate (%)	Mean Earnings (10 Yrs after Entry)	Mean Earnings (6 Yrs after Entry)
166683	Massachusetts Institute of Technology	MA	Private NonProfit	7.94	\$153,600.00	\$99,600.00
243744	Stanford University	CA	Private NonProfit	4.81	\$141,300.00	\$92,800.00
166027	Harvard University	MA	Private NonProfit	5.4	\$139,100.00	\$91,300.00
215062	University of Pennsylvania	PA	Private NonProfit	9.44	\$131,600.00	\$91,200.00
131496	Georgetown University	DC	Private NonProfit	16.85	\$125,200.00	\$81,100.00

Public TOP 5

161299	Maine Maritime Academy	ME	Public	75.51	\$107,200.00	\$82,000.00
197027	United States Merchant Marine Academy	NY	Public	20.3	\$98,600.00	\$78,600.00
166692	Massachusetts Maritime Academy	MA	Public	80.48	\$96,000.00	\$75,200.00
126775	Colorado School of Mines	CO	Public	40.35	\$95,600.00	\$74,700.00
111188	California State University Maritime Academy	CA	Public	68.2	\$87,800.00	\$76,700.00

Private For-profit TOP 5

443410	DigiPen Institute of Technology	WA	Private ForProfit	53.32	\$82,400.00	\$60,900.00
445692	Neumont College of Computer Science	UT	Private ForProfit	82.39	\$75,100.00	\$66,400.00
485272	West Coast University-Miami	FL	Private ForProfit	100	\$67,700.00	\$71,900.00
443331	West Coast University-Los Angeles	CA	Private ForProfit	100	\$67,700.00	\$71,900.00
458210	West Coast University-Orange County	CA	Private ForProfit	88.24	\$67,700.00	\$71,900.00

TOP 5

School Name	Completion Rate Group	150% Completion Rate at 4Yr (%)	Retention Rate Group	Retention Rate 4Yr (%)	Mean Earnings (10 Yrs after Entry)
Massachusetts Institute of Technology	90-100	92.76	90-100	98.28	153600.00
Stanford University	90-100	93.90	90-100	98.02	141300.00
Harvard University	90-100	97.23	90-100	97.35	139100.00
University of Pennsylvania	90-100	95.26	90-100	98.23	131600.00
Georgetown University	90-100	94.26	90-100	96.49	125200.00

BOTTOM 5

School Name	Completion Rate Group	150% Completion Rate at 4Yr (%)	Retention Rate Group	Retention Rate 4Yr (%)	Mean Earnings (10 Yrs after Entry)
Institute of American Indian and Alaska Native...	10-20	20.00	40-50	50.00	18000.00
Yeshivath Viznitz	80-90	86.96	90-100	96.30	19600.00
Rabbinical College of Ohr Shimon Yisroel	30-40	40.00	50-60	52.63	20000.00
Escuela de Artes Plasticas y Diseno de Puerto ...	30-40	30.77	90-100	92.31	20600.00
National University College-Ponce	30-40	34.82	50-60	58.00	20900.00

TOP 5

School ID	School Name	School State	Faculty's average salary per month	Mean Earnings (10 Yrs after Entry)
166683	Massachusetts Institute of Technology	MA	\$17,545.00	\$153,600.00
243744	Stanford University	CA	\$22,146.00	\$141,300.00
166027	Harvard University	MA	\$19,152.00	\$139,100.00
215062	University of Pennsylvania	PA	\$16,796.00	\$131,600.00
130794	Yale University	CT	\$18,333.00	\$124,400.00

BOTTOM 5

School ID	School Name	School State	Faculty's average salary per month	Mean Earnings (10 Yrs after Entry)
187745	Institute of American Indian and Alaska Native Culture and Arts Development	NM	\$6,269.00	\$18,000.00
197735	Yeshivath Viznitz	NY	\$3,048.00	\$19,600.00
405854	Rabbinical College of Ohr Shimon Yisroel	NY	\$1,742.00	\$20,000.00
241951	Escuela de Artes Plasticas y Diseno de Puerto Rico	PR	\$3,322.00	\$20,600.00
468723	National University College-Caguas	PR	\$3,543.00	\$20,900.00

	Faculty's average salary per month	Mean Earnings (10 Yrs after Entry)
Max	\$27,570.00	\$153,600.00
Min	\$588.00	\$26,100.00
Mean	\$7,621.11	\$51,170.65

TOP 5

School ID	School Name	School State	SAT Average Overall	Mean Earnings (10 Yrs after Entry)
166683	Massachusetts Institute of Technology	MA	1519	\$153,600.00
243744	Stanford University	CA	1476	\$141,300.00
166027	Harvard University	MA	1506	\$139,100.00
215062	University of Pennsylvania	PA	1481	\$131,600.00
131496	Georgetown University	DC	1422	\$125,200.00

BOTTOM 5

School ID	School Name	School State	SAT Average Overall	Mean Earnings (10 Yrs after Entry)
198862	Livingstone College	NC	730	\$26,100.00
176318	Rust College	MS	735	\$26,600.00
176044	Mississippi Valley State University	MS	825	\$27,200.00
220598	Lane College	TN	740	\$27,500.00
220808	Memphis College of Art	TN	957	\$28,300.00

	SAT Average Overall	Mean Earnings (10 Yrs after Entry)
Max	1555	\$153,600.00
Min	712	\$26,100.00
Mean	1064.76	\$51,170.65



ISSUES

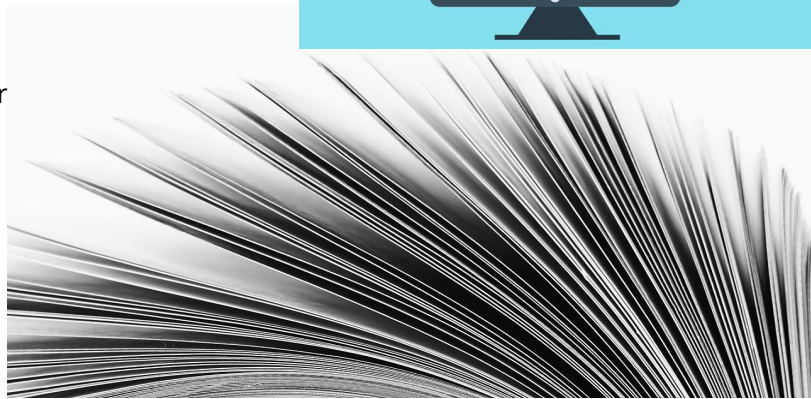
ITERATING THROUGH PAGES OF THE SCORECARD DATA

```
{'metadata': {'page': 0, 'per_page': 100, 'total': 2113},  
  'results': [{'2010': {'academics': {'program': {'assoc': {'agriculture': 0,  
                                                         'architecture': 0,  
                                                         'biological': 0,  
                                                         'business_marketing': 0,
```

- College Scorecard API provides data into sections

- **'metadata'**
 - Master JSON object containing summary of returned results
- **'page'**
 - Page number of returned results
 - Default at *'page'=0*
- **'per_page'**
 - Number of results sets (colleges) returned per page number
 - Default at *'per_page'=20* and maxed at *'per_page'=100*
- **'total'**
 - Maximum number of results sets in returned results
- **'results'**
 - Returned dataset results

- So how to collect all the desired datasets?



```
# Getting number of records returned to set the max page number
query_url = f"{base_url}{fields_url}{parameters_url}&page=0"
response = requests.get(query_url).json()
max_page_num = response["metadata"]["total"]//100 + 1
max_page_num
```

21

```
# Constructing the dataframe from the API request response
```

```
#Initializing variables
```

```
school_df = []
```

```
per_page = 100
```

```
for page_num in range(0,max_page_num):
```

```
    query_url = f"{base_url}{fields_url}{parameters_url}&page={page_num}&per_page={per_page}"
```

```
    response = requests.get(query_url).json()
```

```
    for x in range(len(response["results"])):
```

```
        result_row = {}
```

```
        for key, val in fields.items():
```

```
            try:
```

```
                result_row[key] = response["results"][x][val]
```

```
            except KeyError:
```

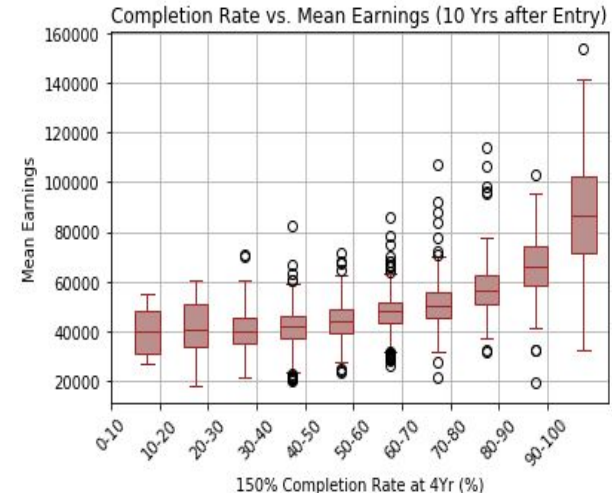
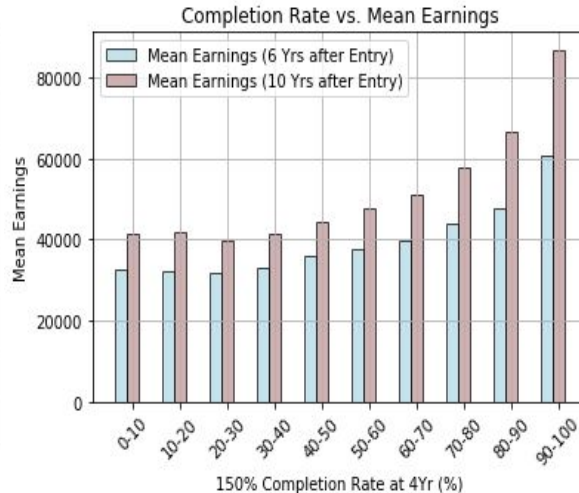
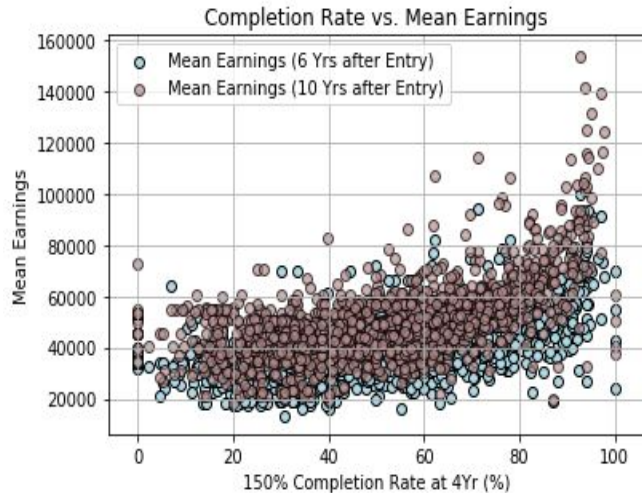
```
                print(f"{key} key not found")
```

```
        school_df.append(result_row)
```

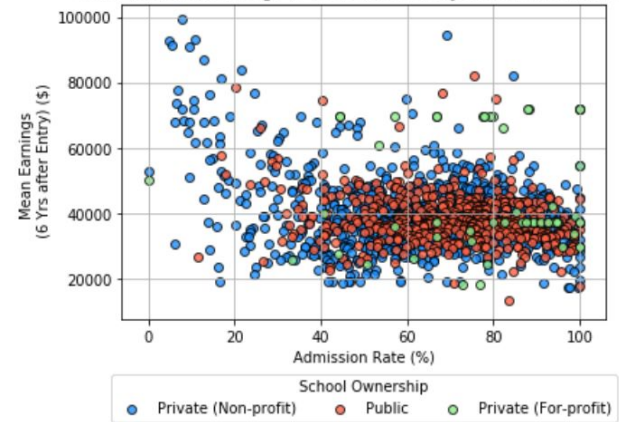
```
school_df = pd.DataFrame(school_df)
```

FINDING THE BEST VISUALIZATION TO TELL THE CLEAREST STORY

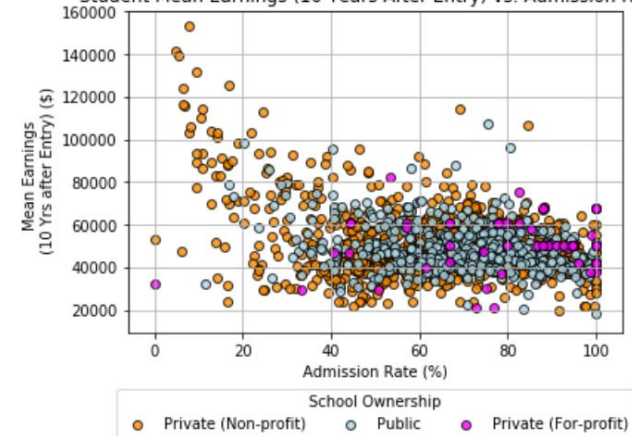
- Which type of plot to use?
 - Scatter plot? Bar plot? Box plot?
- What story does it tell?



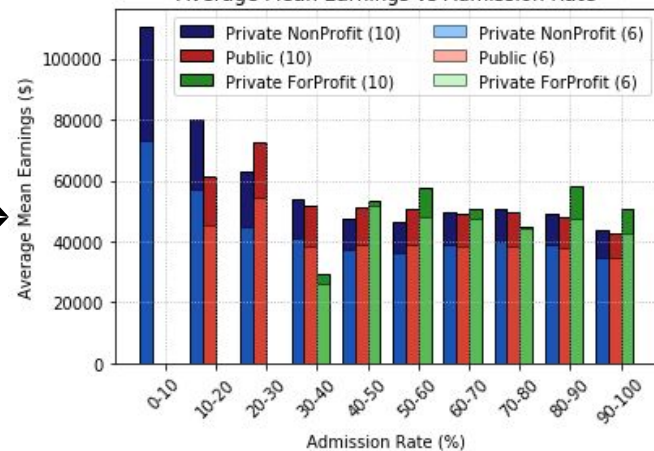
Student Mean Earnings (6 Years After Entry) vs. Admission Rates



Student Mean Earnings (10 Years After Entry) vs. Admission Rates



Average Mean Earnings vs Admission Rate





LIMITATIONS OF CURRENT DATA



COLLEGE SCORECARD



- **Does not** include every student at each college
 - Data collected only from students who participate in financial aid programs with federal tax records
 - Title IV of the Higher Education Act of 1965
- **Does not** further evaluate schools with many different majors
 - High earnings skewed towards specialized schools
 - Especially favors schools in high-paying fields: healthcare, engineering, technology, business, etc.
- **Does not** report separate earnings of branching school campuses.
 - Only system level (overall college) is reported
 - EX: Rutgers-New Brunswick, Rutgers-Camden, Rutgers-Newark share similar datasets (not distincted)
 - EX: West Coast University campuses: Miami, Los Angeles, Orange County



END

