

SS22 CMSE 201 Final Project

Student

Section XXX

Exploration of College Major Salary Outcomes

Background and Motivation

As I embark on the journey throughout college and towards adulthood, the consequences of my decisions will become more impactful. In particular, which major is the most financially rewarding? Does the region of the college have an affect on income-levels? Despite being set spending the rest of my college career to be at MSU and being a data science major, this project would have been very useful when I was first applying to universities. This would also be really helpful to future students who don't have any specific goals or dreams. With this in mind, I set out to research the forementioned questions.

The first step was to find a usable dataset. Searching on Kaggle, a website with a plethora of datasets, I was able to find one titled "Where it Pays to Attend College." With the data being gathered by The Wall Street Journal and Payscale, Inc, it seemed to be a reliable dataset for this project.

There were three different dataframes included in this set: degrees-that-pay-back.csv, salaries-by-college-type.csv, and salaries-by-region.csv. However, only the first and last ones were relevant to my research aim.

Looking at the degrees-that-pay-back.csv, there are a total of eight columns and 50 unique values, aka 50 different listed majors. The other seven columns described the following for each major: Starting Median Salary, Mid-Career Median Salary, Percent change from Starting to Mid-Career Salary, Mid-Career 10th Percentile Salary, Mid-Career 25th Percentile Salary, Mid-Career 75th Percentile Salary, and Mid-Career 90th Percentile Salary.

Since that's quite a few of variables, I decided to focus on the Starting Median Salary, Mid-Career Median Salary, and Mid-Career 90th Percentile Salary.

For the salaries-by-region.csv dataframe, there were a total of 320 different observations that accounted for: the Starting Median Salary, Mid-Career Median Salary, Mid-Career 10th Percentile Salary, Mid-Career 25th Percentile Salary, Mid-Career 75th Percentile Salary, and Mid-Career 90th Percentile Salary.

To stay consistent with the first dataframe, I also focused on the Starting Median Salary, Mid-Career Median Salary, and Mid-Career 90th Percentile Salary.

Methodology

```
In [1]: #loading in necessary libraries
import pandas as pd
import matplotlib.pyplot as plt
import matplotlib
import seaborn as sns
from matplotlib.lines import Line2D
```

```
In [15]: #load in first dataset
major_salary_df = pd.read_csv("degrees-that-pay-back.csv")
major_salary_df
```

	Undergraduate Major	Starting Median Salary	Mid-Career Median Salary	Percent change from Starting to Mid-Career Salary	Mid-Career 10th Percentile Salary	Mid-Career 25th Percentile Salary	Mid-Career 75th Percentile Salary	Mid-Career 90th Percentile Salary
0	Accounting	\$46,000.00	\$77,100.00		\$67.6	\$42,200.00	\$56,100.00	\$108,000.00
1	Aerospace Engineering	\$57,700.00	\$101,000.00		75.0	\$64,300.00	\$82,100.00	\$127,000.00
2	Agriculture	\$42,600.00	\$71,900.00		68.8	\$36,300.00	\$52,100.00	\$96,300.00
3	Anthropology	\$36,800.00	\$61,500.00		67.1	\$33,800.00	\$45,500.00	\$89,300.00
4	Architecture	\$41,600.00	\$76,800.00		84.6	\$50,600.00	\$62,200.00	\$97,000.00
5	Art History	\$35,800.00	\$64,900.00		81.3	\$28,800.00	\$42,200.00	\$87,400.00
6	Biology	\$38,800.00	\$64,800.00		67.0	\$36,900.00	\$47,400.00	\$94,500.00
7	Business Management	\$43,000.00	\$72,100.00		67.7	\$38,800.00	\$51,500.00	\$102,000.00
8	Chemical Engineering	\$63,200.00	\$107,000.00		69.3	\$71,900.00	\$87,300.00	\$143,000.00
9	Chemistry	\$42,600.00	\$79,900.00		87.6	\$45,300.00	\$60,700.00	\$108,000.00
10	Civil Engineering	\$53,900.00	\$90,500.00		67.9	\$43,400.00	\$57,100.00	\$115,000.00
11	Communications	\$38,100.00	\$70,000.00		83.7	\$37,000.00	\$49,700.00	\$98,800.00
12	Computer Engineering	\$61,400.00	\$105,000.00		71.0	\$66,100.00	\$84,100.00	\$135,000.00
13	Computer Science	\$55,900.00	\$95,500.00		70.8	\$56,000.00	\$74,900.00	\$122,000.00
14	Construction	\$53,700.00	\$88,900.00		65.5	\$56,300.00	\$68,100.00	\$118,000.00
15	Criminal Justice	\$35,000.00	\$56,300.00		60.9	\$32,200.00	\$41,600.00	\$80,700.00
16	Drama	\$35,900.00	\$56,900.00		58.5	\$36,700.00	\$41,300.00	\$79,100.00
17	Economics	\$50,100.00	\$98,600.00		96.8	\$50,600.00	\$70,600.00	\$145,000.00
18	Education	\$34,900.00	\$52,000.00		49.0	\$29,300.00	\$37,900.00	\$73,400.00
19	Electrical Engineering	\$60,900.00	\$103,000.00		69.1	\$69,300.00	\$83,800.00	\$130,000.00
20	English	\$38,000.00	\$64,700.00		70.3	\$33,400.00	\$44,800.00	\$93,200.00
21	Finance	\$37,900.00	\$68,500.00		80.7	\$37,000.00	\$46,500.00	\$100,000.00
22	Film	\$47,900.00	\$88,300.00		84.3	\$47,200.00	\$62,100.00	\$128,000.00
23	Forestry	\$39,100.00	\$62,600.00		60.1	\$41,000.00	\$49,300.00	\$97,200.00
24	Geography	\$41,200.00	\$65,500.00		59.0	\$40,000.00	\$50,000.00	\$90,800.00
25	Geology	\$43,500.00	\$79,500.00		82.8	\$45,000.00	\$59,600.00	\$101,000.00
26	Graphic Design	\$35,700.00	\$59,800.00		67.5	\$36,000.00	\$45,500.00	\$80,800.00
27	Health Care Administration	\$38,800.00	\$60,600.00		56.2	\$34,600.00	\$45,600.00	\$78,800.00
28	History	\$39,200.00	\$71,000.00		81.1	\$37,000.00	\$49,200.00	\$103,000.00
29	Hospitality & Tourism	\$37,800.00	\$57,500.00		52.1	\$35,500.00	\$43,600.00	\$81,900.00
30	Industrial Engineering	\$57,700.00	\$94,700.00		64.1	\$57,100.00	\$72,300.00	\$132,000.00
31	Information Technology (IT)	\$49,100.00	\$74,800.00		52.3	\$44,500.00	\$56,700.00	\$96,700.00
32	Interior Design	\$36,100.00	\$53,200.00		47.4	\$35,700.00	\$42,600.00	\$72,500.00
33	International Relations	\$40,900.00	\$80,900.00		97.8	\$38,200.00	\$56,000.00	\$111,000.00
34	Journalism	\$35,600.00	\$66,700.00		87.4	\$38,400.00	\$48,300.00	\$97,700.00
35	Management Information Systems (MIS)	\$49,200.00	\$82,300.00		67.3	\$45,300.00	\$60,500.00	\$108,000.00
36	Marketing	\$40,800.00	\$79,600.00		95.1	\$42,100.00	\$55,600.00	\$119,000.00
37	Math	\$45,400.00	\$92,400.00		103.5	\$45,200.00	\$64,200.00	\$128,000.00
38	Mechanical Engineering	\$57,900.00	\$93,600.00		61.7	\$63,700.00	\$76,200.00	\$120,000.00
39	Music	\$35,900.00	\$55,000.00		53.2	\$26,700.00	\$40,200.00	\$88,000.00
40	Nursing	\$54,200.00	\$67,000.00		23.6	\$47,600.00	\$56,400.00	\$80,900.00
41	Nutrition	\$39,900.00	\$55,300.00		38.6	\$33,900.00	\$44,500.00	\$70,500.00
42	Philosophy	\$39,900.00	\$81,200.00		103.5	\$35,500.00	\$52,800.00	\$127,000.00
43	Physician Assistant	\$74,300.00	\$91,700.00		23.4	\$66,400.00	\$75,200.00	\$108,000.00
44	Physics	\$50,300.00	\$97,300.00		93.4	\$56,000.00	\$74,200.00	\$132,000.00
45	Political Science	\$40,800.00	\$78,200.00		91.7	\$41,200.00	\$55,300.00	\$114,000.00
46	Psychology	\$35,900.00	\$60,400.00		68.2	\$31,600.00	\$42,100.00	\$87,500.00
47	Religion	\$34,100.00	\$52,000.00		52.5	\$29,700.00	\$36,500.00	\$70,900.00
48	Sociology	\$36,500.00	\$58,200.00		59.5	\$30,700.00	\$40,400.00	\$81,200.00
49	Spanish	\$34,000.00	\$53,100.00		56.2	\$31,000.00	\$40,000.00	\$76,800.00

```
In [3]: #dropping unnecessary columns for better focus
major_salary_df.drop(["Percent change from Starting to Mid-Career Salary", "Mid-Career 75th Percentile Salary"], axis=1, inplace=True)
```

	Undergraduate Major	Starting Median Salary	Mid-Career Median Salary	Mid-Career 90th Percentile Salary
0	Accounting	\$46,000.00	\$77,100.00	\$152,000.00
1	Aerospace Engineering	\$57,700.00	\$101,000.00	\$161,000.00
2	Agriculture	\$42,600.00	\$71,900.00	\$150,000.00
3	Anthropology	\$36,800.00	\$61,500.00	\$138,000.00
4	Architecture	\$41,600.00	\$76,800.00	\$136,000.00
5	Art History	\$35,800.00	\$64,900.00	\$125,000.00
6	Biology	\$38,800.00	\$64,800.00	\$135,000.00
7	Business Management	\$43,000.00	\$72,100.00	\$147,000.00
8	Chemical Engineering	\$63,200.00	\$107,000.00	\$194,000.00
9	Chemistry	\$42,600.00	\$79,900.00	\$148,000.00
10	Civil Engineering	\$53,900.00	\$90,500.00	\$148,000.00
11	Communications	\$38,100.00	\$70,000.00	\$143,000.00
12	Computer Engineering	\$61,400.00	\$105,000.00	\$162,000.00
13	Computer Science	\$55,900.00	\$95,500.00	\$154,000.00
14	Construction	\$53,700.00	\$88,900.00	\$171,000.00
15	Criminal Justice	\$35,000.00	\$56,300.00	\$107,000.00
16	Drama	\$35,900.00	\$56,900.00	\$153,000.00
17	Economics	\$50,100.00	\$98,600.00	\$210,000.00
18	Education	\$34,900.00	\$52,000.00	\$102,000.00
19	Electrical Engineering	\$60,900.00	\$103,000.00	\$168,000.00
20	English	\$38,000.00	\$64,700.00	\$133,000.00
21	Film	\$37,900.00	\$68,500.00	\$136,000.00
22	Finance	\$47,900.00	\$88,300.00	\$195,000.00
23	Forestry	\$39,100.00	\$62,600.00	\$111,000.00
24	Geography	\$41,200.00	\$65,500.00	\$122,000.00
25	Geology	\$43,500.00	\$79,500.00	\$156,000.00
26	Graphic Design	\$35,700.00	\$59,800.00	\$112,000.00
27	Health Care Administration	\$38,800.00	\$60,600.00	\$101,000.00
28	History	\$39,200.00	\$71,000.00	\$149,000.00
29	Hospitality & Tourism	\$37,800.00	\$57,500.00	\$124,000.00
30	Industrial Engineering	\$57,700.00	\$94,700.00	\$173,000.00
31	Information Technology (IT)	\$49,100.00	\$74,800.00	\$129,000.00
32	Interior Design	\$36,100.00	\$53,200.00	\$107,000.00
33	International Relations	\$40,900.00	\$80,900.00	\$157,000.00
34	Journalism	\$35,600.00	\$66,700.00	\$145,000.00
35	Management Information Systems (MIS)	\$49,200.00	\$82,300.00	\$146,000.00
36	Marketing	\$40,800.00	\$79,600.00	\$175,000.00
37	Math	\$45,400.00	\$92,400.00	\$183,000.00
38	Mechanical Engineering	\$57,900.00	\$93,600.00	\$163,000.00
39	Music	\$35,900.00	\$55,000.00	\$134,000.00
40	Nursing	\$54,200.00	\$67,000.00	\$98,300.00
41	Nutrition	\$39,900.00	\$55,300.00	\$99,200.00
42	Philosophy	\$39,900.00	\$81,200.00	\$168,000.00
43	Physician Assistant	\$74,300.00	\$91,700.00	\$124,000.00
44	Physics	\$50,300.00	\$97,300.00	\$178,000.00
45	Political Science	\$40,800.00	\$78,200.00	\$168,000.00
46	Psychology	\$35,900.00	\$60,400.00	\$127,000.00
47	Religion	\$34,100.00	\$52,000.00	\$96,400.00
48	Sociology	\$36,500.00	\$58,200.00	\$118,000.00
49	Spanish	\$34,000.00	\$53,100.00	\$96,400.00

```
In [4]: #creating a list of the column names I want to loop through
cols = ["Starting Median Salary", "Mid-Career Median Salary", "Mid-Career 90th Percentile Salary"]
```

```
In [5]: #cleaning the columns with the $ and , to make them usable
for col in cols:
    major_salary_df[col] = major_salary_df[col].astype(str).str.replace("$", '', regex=True).replace(",", '', regex=True)
```

```
In [6]: #renaming the columns
major_salary_df.rename(columns={'Undergraduate Major': 'major', 'Starting Median Salary': 'start_med', 'Mid-Career Median Salary': 'mid_med', 'Mid-Career 90th Percentile Salary': 'mid90'})
```

```
In [7]: #changing columns from strings to floats
major_salary_cleaned = major_salary_df.astype({"start_med": float, "mid_med": float, "mid90": float})
#sorting the values by the 90th percentile for mid-career to plot later
major_salary_sorted = major_salary_cleaned.sort_values(by = 'mid90', ascending = False)
```

```
In [14]: #looking at dataframe sorted by starting salary
major_start_salary_sorted = major_salary_cleaned.sort_values(by = 'start_med', ascending = False)
major_start_salary_sorted
```

		(WIS)							
31	Information Technology (IT)	49100.0	74800.0	52.3	\$44,500.00	\$56,700.00	\$96,700.00	129000.0	
22	Finance	47900.0	88300.0	84.3	\$47,200.00	\$62,100.00	\$128,000.00	195000.0	
0	Accounting	46000.0	77100.0	67.6	\$42,200.00	\$56,100.00	\$108,000.00	152000.0	
37	Math	45400.0	92400.0	103.5	\$45,200.00	\$64,200.00	\$128,000.00	183000.0	
25	Geology	43500.0	79500.0	82.8	\$45,000.00	\$59,600.00	\$101,000.00	156000.0	
7	Business Management	43000.0	72100.0	67.7	\$38,800.00	\$51,500.00	\$102,000.00	147000.0	
2	Agriculture	42600.0	71900.0	68.8	\$36,300.00	\$52,100.00	\$96,300.00	150000.0	
9	Chemistry	42600.0	79900.0	87.6	\$45,300.00	\$60,700.00	\$108,000.00	148000.0	
4	Architecture	41600.0	76800.0	84.6	\$50,600.00	\$62,200.00	\$97,000.00	136000.0	
24	Geography	41200.0	65500.0	59.0	\$40,000.00	\$50,000.00	\$90,800.00	132000.0	
33	International Relations	40900.0	80900.0	97.8	\$38,200.00	\$56,000.00	\$111,000.00	157000.0	
45	Political Science	40800.0	78200.0	91.7	\$41,200.00	\$55,300.00	\$114,000.00	168000.0	
36	Marketing	40800.0	79600.0	95.1	\$42,100.00	\$55,600.00	\$119,000.00	175000.0	
42	Philosophy	39900.0	81200.0	103.5	\$35,500.00	\$52,800.00	\$127,000.00	168000.0	
41	Nutrition	39900.0	55300.0	38.6	\$33,900.00	\$44,500.00	\$70,500.00	99200.0	
28	History	39200.0	71000.0	81.1	\$37,000.00	\$49,200.00	\$103,000.00	149000.0	
23	Forestry	39100.0	62600.0	60.1	\$41,000.00	\$49,300.00	\$78,200.00	111000.0	
27	Health Care Administration	38800.0	60600.0	56.2	\$34,600.00	\$45,600.00	\$78,800.00	101000.0	
6	Biology	38800.0	64800.0	67.0	\$36,900.00	\$47,400.00	\$94,500.00	135000.0	
11	Communications	38100.0	70000.0	83.7	\$37,500.00	\$49,700.00	\$98,800.00	143000.0	
20	English	38000.0	64700.0	70.3	\$33,400.00	\$44,800.00	\$93,200.00	133000.0	
21	Film	37900.0	68500.0	80.7	\$33,900.00	\$45,500.00	\$100,000.00	136000.0	
29	Hospitality & Tourism	37800.0	57500.0	52.1	\$35,500.00	\$43,600.00	\$81,900.00	124000.0	
3	Anthropology	36800.0	61500.0	67.1	\$33,800.00	\$45,500.00	\$89,300.00	138000.0	
38	Sociology	36500.0	58200.0	59.5	\$30,700.00	\$40,400.00	\$81,200.00	118000.0	
42	Interior Design	36100.0	53200.0	47.4	\$35,700.00	\$42,600.00	\$72,500.00	107000.0	
46	Psychology	35900.0	60400.0	68.2	\$31,600.00	\$42,100.00	\$87,500.00	127000.0	
39	Music	35900.0	55000.0	53.2	\$26,700.00	\$40,200.00	\$88,000.00	134000.0	
16	Drama	35900.0	56900.0	58.5	\$36,700.00	\$41,300.00	\$79,100.00	153000.0	
5	Art History	35800.0	64900.0	81.3	\$28,800.00	\$42,200.00	\$87,400.00	125000.0	
26	Graphic Design	35700.0	59800.0	67.5	\$36,000.00	\$45,500.00	\$80,800.00	112000.0	
34	Journalism	35600.0	66700.0	87.4	\$38,400.00	\$48,300.00	\$97,700.00	145000.0	
15	Criminal Justice	35000.0	56300.0	60.9	\$32,200.00	\$41,600.00	\$80,700.00	107000.0	
18	Education	34900.0	52000.0	49.0	\$29,300.00	\$37,900.00	\$73,400.00	96200.0	
47	Religion	34100.0	52000.0	52.5	\$29,700.00	\$36,500.00	\$70,900.00	104000.0	
49	Spanish	34000.0	53100.0	56.2	\$31,000.00	\$40,000.00	\$76,800.00	96400.0	
<div> <div>#looking at dataframe sorted by mid-career salary</div> <div>major_mid_salary_sorted = major_salary_cleaned.sort_values(by = 'mid_med', ascending = False)</div> <div>major_mid_salary_sorted</div> </div>									
	major	start_med	mid_med	Percent change from Starting to Mid-Career Salary	Mid-Career 10th Percentile Salary	Mid-Career 25th Percentile Salary	Mid-Career 75th Percentile Salary	mid90	
8	Chemical Engineering	63200.0	107000.0	69.3	\$71,900.00	\$87,300.00	\$143,000.00	194000.0	
12	Computer Engineering	61400.0	105000.0	71.0	\$66,100.00	\$84,100.00	\$135,000.00	162000.0	
19	Electrical Engineering	60900.0	103000.0	69.1	\$69,300.00	\$83,800.00	\$130,000.00	168000.0	
38	Mechanical Engineering	57900.0	101000.0	75.0	\$64,300.00	\$82,100.00	\$127,000.00	161000.0	
1	Aerospace Engineering	57700.0	101000.0	75.0	\$64,300.00	\$82,100.00	\$127,000.00	161000.0	
17	Economics	50100.0	98600.0	96.8	\$50,600.00	\$70,600.00	\$145,000.00	210000.0	
44	Physics	50300.0	97300.0	93.4	\$56,000.00	\$74,200.00	\$132,000.00	178000.0	
13	Computer Science	55900.0	95500.0	70.8	\$56,000.00	\$74,900.00	\$122,000.00	154000.0	
40	Nursing	54200.0	67000.0	23.6	\$47,600.00	\$56,400.00	\$80,900.00	98300.0	
10	Civil Engineering	53900.0	90500.0	67.9	\$63,400.00	\$75,100.00	\$115,000.00	148000.0	
14	Construction	53700.0	88900.0	65.5	\$56,300.00	\$68,100.00	\$118,000.00	171000.0	