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-byregion.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 %matplotlib inline 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 Out[15]: Starting Mid-Career **Mid-Career Mid-Career Mid-Career** Mid-Career Percent change from Undergraduate 75th Percentile 90th Percentile Median Median **Starting to Mid-Career** 10th Percentile 25th Percentile Major Salary Salary Salary Salary Salary Salary Salary \$42,200.00 0 \$46,000.00 \$77,100.00 67.6 \$56,100.00 \$108,000.00 \$152,000.00 Accounting Aerospace \$64,300.00 \$57,700.00 \$101,000.00 75.0 \$82,100.00 \$127,000.00 \$161,000.00 Engineering 2 \$42,600.00 \$71,900.00 68.8 \$36,300.00 \$52,100.00 \$96,300.00 \$150,000.00 Agriculture \$36,800.00 3 \$61,500.00 67.1 \$33,800.00 \$45,500.00 \$89,300.00 \$138,000.00 Anthropology Architecture 84.6 \$50,600.00 \$97,000.00 4 \$41,600.00 \$76,800.00 \$62,200.00 \$136,000.00 \$35,800.00 \$28,800.00 5 Art History \$64,900.00 81.3 \$42,200.00 \$87,400.00 \$125,000.00 6 Biology \$94,500.00 \$38,800.00 \$64,800.00 67.0 \$36,900.00 \$47,400.00 \$135,000.00 **Business** \$38,800.00 \$51,500.00 \$102,000.00 7 \$43,000.00 \$72,100.00 67.7 \$147,000.00 Management Chemical \$87,300.00 \$143,000.00 8 \$63,200.00 \$107,000.00 69.3 \$71,900.00 \$194,000.00 Engineering 9 Chemistry \$42,600.00 \$79,900.00 87.6 \$45,300.00 \$60,700.00 \$108,000.00 \$148,000.00 10 Civil Engineering \$90,500.00 \$115,000.00 \$53,900.00 67.9 \$63,400.00 \$75,100.00 \$148,000.00 \$98,800.00 \$143,000.00 11 Communications \$38,100.00 \$70,000.00 83.7 \$37,500.00 \$49,700.00 Computer 12 \$61,400.00 \$105,000.00 71.0 \$66,100.00 \$84,100.00 \$135,000.00 \$162,000.00 Engineering Computer Science 13 \$55,900.00 \$95,500.00 70.8 \$56,000.00 \$74,900.00 \$122,000.00 \$154,000.00 14 Construction \$53,700.00 \$88,900.00 65.5 \$56,300.00 \$68,100.00 \$118,000.00 \$171,000.00 15 Criminal Justice \$35,000.00 \$56,300.00 60.9 \$32,200.00 \$41,600.00 \$80,700.00 \$107,000.00 58.5 16 \$35,900.00 \$56,900.00 \$36,700.00 \$41,300.00 \$79,100.00 \$153,000.00 Drama 17 **Economics** \$50,100.00 \$98,600.00 96.8 \$50,600.00 \$70,600.00 \$145,000.00 \$210,000.00 18 Education \$34,900.00 \$52,000.00 49.0 \$29,300.00 \$37,900.00 \$73,400.00 \$102,000.00 Electrical 19 \$60,900.00 \$103,000.00 69.1 \$69,300.00 \$83,800.00 \$130,000.00 \$168,000.00 Engineering 70.3 20 English \$38,000.00 \$64,700.00 \$33,400.00 \$44,800.00 \$93,200.00 \$133,000.00 21 \$37,900.00 \$68,500.00 80.7 \$33,900.00 \$45,500.00 \$100,000.00 \$136,000.00 Film 22 Finance \$47,900.00 \$88,300.00 84.3 \$47,200.00 \$62,100.00 \$128,000.00 \$195,000.00 23 \$39,100.00 \$62,600.00 60.1 \$41,000.00 \$49,300.00 \$78,200.00 \$111,000.00 Forestry 24 Geography \$41,200.00 \$65,500.00 59.0 \$40,000.00 \$50,000.00 \$90,800.00 \$132,000.00 25 \$43,500.00 \$79,500.00 82.8 \$45,000.00 \$59,600.00 \$101,000.00 \$156,000.00 Geology 26 Graphic Design \$35,700.00 \$59,800.00 67.5 \$36,000.00 \$45,500.00 \$80,800.00 \$112,000.00 Health Care 27 \$38,800.00 \$60,600.00 56.2 \$34,600.00 \$45,600.00 \$78,800.00 \$101,000.00 Administration History 28 \$39,200.00 \$71,000.00 81.1 \$37,000.00 \$49,200.00 \$103,000.00 \$149,000.00 Hospitality & 29 \$37,800.00 \$57,500.00 52.1 \$35,500.00 \$43,600.00 \$81,900.00 \$124,000.00 **Tourism** Industrial \$57,700.00 \$94,700.00 \$72,300.00 30 64.1 \$57,100.00 \$132,000.00 \$173,000.00 Engineering Information 31 \$49,100.00 \$74,800.00 52.3 \$44,500.00 \$56,700.00 \$96,700.00 \$129,000.00 Technology (IT) Interior Design \$72,500.00 32 \$36,100.00 \$53,200.00 47.4 \$35,700.00 \$42,600.00 \$107,000.00 International 33 \$40,900.00 \$80,900.00 97.8 \$38,200.00 \$56,000.00 \$111,000.00 \$157,000.00 Relations 34 Journalism \$35,600.00 \$66,700.00 87.4 \$38,400.00 \$48,300.00 \$97,700.00 \$145,000.00 Management \$146,000.00 35 Information \$49,200.00 \$82,300.00 67.3 \$45,300.00 \$60,500.00 \$108,000.00 Systems (MIS) 36 Marketing \$40,800.00 \$79,600.00 95.1 \$42,100.00 \$55,600.00 \$119,000.00 \$175,000.00 \$45,200.00 \$183,000.00 37 \$45,400.00 \$92,400.00 103.5 \$64,200.00 \$128,000.00 Math Mechanical \$57,900.00 \$93,600.00 \$63,700.00 \$76,200.00 \$120,000.00 \$163,000.00 Engineering Music \$35,900.00 39 \$55,000.00 53.2 \$26,700.00 \$40,200.00 \$88,000.00 \$134,000.00 23.6 \$98,300.00 40 Nursing \$54,200.00 \$67,000.00 \$47,600.00 \$56,400.00 \$80,900.00 \$70,500.00 41 Nutrition \$39,900.00 \$55,300.00 38.6 \$33,900.00 \$44,500.00 \$99,200.00 Philosophy 42 \$39,900.00 \$81,200.00 103.5 \$35,500.00 \$52,800.00 \$127,000.00 \$168,000.00 43 Physician Assistant \$74,300.00 \$91,700.00 23.4 \$66,400.00 \$75,200.00 \$108,000.00 \$124,000.00 \$56,000.00 \$178,000.00 44 **Physics** \$50,300.00 \$97,300.00 93.4 \$74,200.00 \$132,000.00 **Political Science** \$78,200.00 \$114,000.00 45 \$40,800.00 91.7 \$41,200.00 \$55,300.00 \$168,000.00 Psychology \$42,100.00 46 \$35,900.00 \$60,400.00 68.2 \$31,600.00 \$87,500.00 \$127,000.00 52.5 \$70,900.00 47 Religion \$34,100.00 \$52,000.00 \$29,700.00 \$36,500.00 \$96,400.00 48 Sociology \$36,500.00 \$58,200.00 59.5 \$30,700.00 \$40,400.00 \$81,200.00 \$118,000.00 \$53,100.00 49 Spanish \$34,000.00 56.2 \$31,000.00 \$40,000.00 \$76,800.00 \$96,400.00 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', Out[3]: Undergraduate Major Starting Median Salary Mid-Career Median Salary Mid-Career 90th Percentile Salary \$46,000.00 \$152,000.00 0 Accounting \$77,100.00 1 Aerospace Engineering \$57,700.00 \$101,000.00 \$161,000.00 2 Agriculture \$42,600.00 \$71,900.00 \$150,000.00 Anthropology \$36,800.00 3 \$61,500.00 \$138,000.00 4 Architecture \$41,600.00 \$76,800.00 \$136,000.00 \$125,000.00 5 Art History \$35,800.00 \$64,900.00 Biology 6 \$38,800.00 \$64,800.00 \$135,000.00 **Business Management** 7 \$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 \$55,900.00 \$95,500.00 \$154,000.00 **Computer Science** 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 \$50,100.00 \$98,600.00 \$210,000.00 **Economics** 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 \$132,000.00 25 \$43,500.00 \$79,500.00 \$156,000.00 Geology 26 Graphic Design \$35,700.00 \$59,800.00 \$112,000.00 Health Care Administration 27 \$38,800.00 \$60,600.00 \$101,000.00 28 \$39,200.00 \$71,000.00 \$149,000.00 History 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 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 \$45,400.00 \$92,400.00 \$183,000.00 Math 38 Mechanical Engineering \$57,900.00 \$93,600.00 \$163,000.00 39 Music \$35,900.00 \$55,000.00 \$134,000.00 40 \$54,200.00 \$67,000.00 \$98,300.00 Nursing 41 \$39,900.00 \$55,300.00 \$99,200.00 Nutrition 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 \$52,000.00 47 \$34,100.00 \$96,400.00 Religion 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-Car 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 Percent change from Out[14]: Mid-Career 25th Mid-Career 10th Mid-Career 75th mid90 major start_med mid_med **Starting to Mid-Career Percentile Salary Percentile Salary Percentile Salary** Salary Physician Assistant 43 74300.0 91700.0 23.4 \$66,400.00 \$75,200.00 \$108,000.00 124000.0 8 63200.0 107000.0 69.3 \$71,900.00 \$87,300.00 \$143,000.00 194000.0 Chemical Engineering 12 **Computer Engineering** 61400.0 105000.0 71.0 \$66,100.00 \$84,100.00 \$135,000.00 162000.0 **Electrical Engineering** 19 60900.0 103000.0 69.1 \$69,300.00 \$83,800.00 \$130,000.00 168000.0 Mechanical 93600.0 38 57900.0 61.7 \$63,700.00 \$76,200.00 \$120,000.00 163000.0 Engineering Aerospace Engineering 57700.0 101000.0 75.0 \$64,300.00 \$82,100.00 \$127,000.00 161000.0 Industrial Engineering 30 57700.0 94700.0 64.1 \$57,100.00 \$72,300.00 \$132,000.00 173000.0 \$122,000.00 13 55900.0 95500.0 70.8 \$56,000.00 \$74,900.00 154000.0 **Computer Science** \$56,400.00 40 Nursing 54200.0 67000.0 23.6 \$47,600.00 \$80,900.00 98300.0 \$115,000.00 10 Civil Engineering 53900.0 90500.0 67.9 \$63,400.00 \$75,100.00 148000.0 \$68,100.00 \$118,000.00 14 53700.0 88900.0 65.5 \$56,300.00 171000.0 Construction **Physics** 50300.0 97300.0 93.4 \$56,000.00 \$74,200.00 \$132,000.00 178000.0 44 50100.0 17 **Economics** 98600.0 96.8 \$50,600.00 \$70,600.00 \$145,000.00 210000.0 Management Information Systems \$45,300.00 35 49200.0 82300.0 67.3 \$60,500.00 \$108,000.00 146000.0 (MIS) Information 31 49100.0 74800.0 52.3 \$44,500.00 \$56,700.00 \$96,700.00 129000.0 Technology (IT) 22 47900.0 88300.0 84.3 \$47,200.00 \$62,100.00 \$128,000.00 195000.0 Finance 0 Accounting 46000.0 77100.0 67.6 \$42,200.00 \$56,100.00 \$108,000.00 152000.0 \$64,200.00 37 45400.0 92400.0 103.5 \$45,200.00 \$128,000.00 183000.0 Math 43500.0 25 Geology 79500.0 82.8 \$45,000.00 \$59,600.00 \$101,000.00 156000.0 **Business Management** 7 43000.0 72100.0 67.7 \$38,800.00 \$51,500.00 \$102,000.00 147000.0 2 42600.0 71900.0 68.8 \$36,300.00 \$52,100.00 \$96,300.00 150000.0 Agriculture \$108,000.00 9 42600.0 79900.0 87.6 \$45,300.00 \$60,700.00 148000.0 Chemistry 76800.0 \$50,600.00 4 Architecture 41600.0 84.6 \$62,200.00 \$97,000.00 136000.0 \$40,000.00 24 Geography 41200.0 65500.0 59.0 \$50,000.00 \$90,800.00 132000.0 International Relations \$38,200.00 33 40900.0 80900.0 97.8 \$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 \$55,600.00 \$119,000.00 36 40800.0 79600.0 95.1 \$42,100.00 175000.0 Marketing \$127,000.00 42 39900.0 81200.0 103.5 \$35,500.00 \$52,800.00 168000.0 Philosophy Nutrition 39900.0 55300.0 41 38.6 \$33,900.00 \$44,500.00 \$70,500.00 99200.0 28 39200.0 71000.0 81.1 \$37,000.00 \$49,200.00 \$103,000.00 149000.0 History 23 39100.0 62600.0 60.1 \$41,000.00 \$49,300.00 \$78,200.00 111000.0 Forestry Health Care \$45,600.00 101000.0 27 38800.0 60600.0 56.2 \$34,600.00 \$78,800.00 Administration 6 38800.0 64800.0 67.0 \$36,900.00 \$47,400.00 \$94,500.00 135000.0 Biology 11 Communications 38100.0 70000.0 83.7 \$37,500.00 \$49,700.00 \$98,800.00 143000.0 20 38000.0 64700.0 70.3 \$33,400.00 \$44,800.00 \$93,200.00 133000.0 English 21 Film 37900.0 68500.0 80.7 \$33,900.00 \$45,500.00 \$100,000.00 136000.0 37800.0 57500.0 29 Hospitality & Tourism 52.1 \$35,500.00 \$43,600.00 \$81,900.00 124000.0 Anthropology 3 36800.0 61500.0 67.1 \$33,800.00 \$45,500.00 \$89,300.00 138000.0 48 Sociology 36500.0 58200.0 59.5 \$30,700.00 \$40,400.00 \$81,200.00 118000.0 Interior Design 32 36100.0 53200.0 47.4 \$35,700.00 \$42,600.00 \$72,500.00 107000.0 35900.0 46 60400.0 68.2 \$31,600.00 \$42,100.00 \$87,500.00 127000.0 Psychology \$26,700.00 \$40,200.00 39 35900.0 55000.0 53.2 \$88,000.00 134000.0 Music 16 35900.0 56900.0 58.5 \$36,700.00 \$41,300.00 \$79,100.00 153000.0 Drama 5 35800.0 64900.0 81.3 \$28,800.00 \$42,200.00 \$87,400.00 125000.0 **Art History** Graphic Design 35700.0 26 59800.0 67.5 \$36,000.00 \$45,500.00 \$80,800.00 112000.0 34 35600.0 66700.0 87.4 \$38,400.00 \$48,300.00 \$97,700.00 145000.0 Journalism 15 Criminal Justice 35000.0 56300.0 60.9 \$32,200.00 \$41,600.00 \$80,700.00 107000.0 18 34900.0 52000.0 49.0 \$29,300.00 \$37,900.00 \$73,400.00 102000.0 Education 47 34100.0 52000.0 52.5 \$29,700.00 \$36,500.00 \$70,900.00 96400.0 Religion 49 34000.0 53100.0 56.2 \$31,000.00 \$40,000.00 \$76,800.00 96400.0 Spanish In [16]: #looking at dataframe sorted by mid-careert salary major_mid_salary_sorted = major_salary_cleaned.sort_values(by = 'mid_med', ascending = False) major_mid_salary_sorted Out[16]: Percent change from Mid-Career 10th Mid-Career 25th Mid-Career 75th **Starting to Mid-Career** mid90 major start_med mid_med **Percentile Salary Percentile Salary Percentile Salary** Salary 63200.0 107000.0 69.3 \$71,900.00 \$87,300.00 \$143,000.00 194000.0 8 Chemical Engineering \$66,100.00 61400.0 105000.0 71.0 \$84,100.00 \$135,000.00 162000.0 12 Computer Engineering 19 60900.0 103000.0 69.1 \$69,300.00 \$83,800.00 \$130,000.00 168000.0 **Electrical Engineering** Aerospace Engineering 57700.0 101000.0 75.0 \$64,300.00 \$82,100.00 \$127,000.00 161000.0 17 50100.0 98600.0 96.8 \$50,600.00 \$70,600.00 \$145,000.00 210000.0 **Economics** 50300.0 93.4 \$56,000.00 \$132,000.00 178000.0 44 **Physics** 97300.0 \$74,200.00 13 55900.0 95500.0 70.8 \$56,000.00 \$74,900.00 \$122,000.00 154000.0 **Computer Science** 30 57700.0 94700.0 64.1 \$72,300.00 \$132,000.00 173000.0 **Industrial Engineering** \$57,100.00 Mechanical 38 57900.0 93600.0 61.7 \$63,700.00 \$76,200.00 \$120,000.00 163000.0 Engineering 37 45400.0 92400.0 103.5 \$45,200.00 \$64,200.00 \$128,000.00 183000.0 Math 43 74300.0 91700.0 23.4 \$66,400.00 \$75,200.00 \$108,000.00 124000.0 Physician Assistant 10 Civil Engineering 53900.0 90500.0 67.9 \$63,400.00 \$75,100.00 \$115,000.00 148000.0 14 53700.0 88900.0 65.5 \$56,300.00 \$68,100.00 \$118,000.00 171000.0 Construction 22 47900.0 88300.0 84.3 \$47,200.00 \$128,000.00 195000.0 Finance \$62,100.00 Management \$45,300.00 35 Information Systems 49200.0 82300.0 67.3 \$60,500.00 \$108,000.00 146000.0 (MIS) 42 39900.0 81200.0 103.5 \$35,500.00 \$52,800.00 \$127,000.00 168000.0 Philosophy International Relations 40900.0 80900.0 97.8 \$38,200.00 \$56,000.00 \$111,000.00 157000.0 33 42600.0 9 Chemistry 79900.0 87.6 \$45,300.00 \$60,700.00 \$108,000.00 148000.0 36 40800.0 79600.0 95.1 \$42,100.00 \$55,600.00 \$119,000.00 175000.0 Marketing Geology 43500.0 79500.0 82.8 \$45,000.00 \$59,600.00 \$101,000.00 156000.0 25 45 Political Science 40800.0 78200.0 91.7 \$41,200.00 \$55,300.00 \$114,000.00 168000.0 0 46000.0 77100.0 67.6 \$42,200.00 \$56,100.00 \$108,000.00 152000.0 Accounting 4 Architecture 41600.0 76800.0 84.6 \$50,600.00 \$62,200.00 \$97,000.00 136000.0 Information 31 49100.0 74800.0 \$44,500.00 \$56,700.00 \$96,700.00 129000.0 52.3 Technology (IT) 43000.0 72100.0 67.7 \$38,800.00 \$51,500.00 \$102,000.00 147000.0 7 **Business Management** 2 Agriculture 42600.0 71900.0 68.8 \$36,300.00 \$52,100.00 \$96,300.00 150000.0 28 39200.0 71000.0 81.1 \$37,000.00 \$49,200.00 \$103,000.00 149000.0 History 11 Communications 38100.0 70000.0 83.7 \$37,500.00 \$49,700.00 \$98,800.00 143000.0 21 37900.0 68500.0 80.7 \$33,900.00 \$45,500.00 \$100,000.00 136000.0 Film 40 54200.0 67000.0 23.6 \$47,600.00 \$56,400.00 \$80,900.00 98300.0 Nursing 34 Journalism 35600.0 66700.0 87.4 \$38,400.00 \$48,300.00 \$97,700.00 145000.0 41200.0 65500.0 59.0 \$40,000.00 \$50,000.00 \$90,800.00 132000.0 24 Geography 5 Art History 35800.0 64900.0 81.3 \$28,800.00 \$42,200.00 \$87,400.00 125000.0 6 38800.0 64800.0 67.0 \$36,900.00 \$94,500.00 135000.0 Biology \$47,400.00 20 38000.0 64700.0 70.3 \$33,400.00 \$44,800.00 \$93,200.00 133000.0 English 23 Forestry 39100.0 62600.0 60.1 \$41,000.00 \$49,300.00 \$78,200.00 111000.0 3 Anthropology 36800.0 61500.0 67.1 \$33,800.00 \$45,500.00 \$89,300.00 138000.0 Health Care \$34,600.00 27 38800.0 60600.0 56.2 \$45,600.00 \$78,800.00 101000.0 Administration 46 Psychology 35900.0 60400.0 68.2 \$31,600.00 \$42,100.00 \$87,500.00 127000.0 Graphic Design 35700.0 59800.0 67.5 \$36,000.00 \$45,500.00 \$80,800.00 112000.0 26 48 36500.0 58200.0 59.5 \$30,700.00 \$40,400.00 \$81,200.00 118000.0 Sociology Hospitality & Tourism 29 37800.0 57500.0 52.1 \$35,500.00 \$43,600.00 \$81,900.00 124000.0 16 35900.0 56900.0 58.5 \$36,700.00 \$41,300.00 \$79,100.00 153000.0 Drama 15 Criminal Justice 35000.0 56300.0 60.9 \$32,200.00 \$41,600.00 \$80,700.00 107000.0 41 Nutrition 39900.0 55300.0 38.6 \$33,900.00 \$44,500.00 \$70,500.00 99200.0 39 35900.0 55000.0 \$26,700.00 \$40,200.00 \$88,000.00 134000.0 Music 53.2 32 Interior Design 36100.0 53200.0 47.4 \$35,700.00 \$42,600.00 \$72,500.00 107000.0 49 34000.0 53100.0 56.2 \$31,000.00 \$40,000.00 \$76,800.00 96400.0 Spanish 18 Education 34900.0 52000.0 49.0 \$29,300.00 \$37,900.00 \$73,400.00 102000.0 47 34100.0 52000.0 52.5 \$29,700.00 \$70,900.00 96400.0 Religion \$36,500.00 In [11]: #setting up the figure configurations sns.set_theme(style="whitegrid") f, ax = plt.subplots(figsize=(6, 10)) sns.set_color_codes("muted") #the plotting sns.barplot(x="mid90", y="major", data=major_salary_sorted, label="90th Percentile Mid-Career Salary for Each Ma sns.barplot(x="mid_med", y="major", data=major_salary_sorted, label="Median Mid-Career Salary for Each Major", co sns.barplot(x="start_med", y="major", data=major_salary_sorted, label="Median Starting Salary for Each Major", data=major_salary_s #adding in extra features/parameters ax.legend(ncol=1, loc="lower right", frameon=True, fontsize=10) ax.set(xlim=(0, 200000), ylabel="", xlabel="Salary (USD)", title="Salary for Each Major") sns.despine(left=True, bottom=True) Salary for Each Major Economics Finance Chemical Engineering Math Physics Marketing Industrial Engineering Construction Political Science Philosophy Electrical Engineering Mechanical Engineering Computer Engineering Aerospace Engineering International Relations Geology Computer Science Drama Accounting Agriculture History Civil Engineering Chemistry Business Management Management Information Systems (MIS) Journalism Communications Anthropology Architecture Biology Music English Geography Information Technology (IT) Psychology Art History Hospitality & Tourism Physician Assistant Sociology Graphic Design Forestry Interior Design Criminal Justice Education Health Care Administration 90th Percentile Mid-Career Salary for Each Major Nutrition Median Mid-Career Salary for Each Major Nursing Religion Median Starting Salary for Each Major Spanish 75000 100000 125000 150000 175000 200000 Salary (USD) In [18]: #load in second dataset region_salary_df = pd.read_csv("salaries-by-region.csv") In [19]: #dropping columns region salary df.drop(['Mid-Career 75th Percentile Salary', 'Mid-Career 25th Percentile Salary', 'Mid-Career 10 In [20]: #cleaning \$ and , punctuation symbolsfor col in cols: region_salary_df[col] = region_salary_df[col].astype(str).str.replace("\$", '', regex=True).replace(",", '', regex=True). region_salary_df.rename(columns={'School Name': 'name', 'Region': 'region', 'Starting Median Salary': 'start me

In [21]: #changing columns from strings to floats region salary cleaned = region salary df.astype({"start med": float, "mid med": float, "mid90": float}) #sorting the values by the 90th percentile for mid-career to plot region salary sorted = region salary cleaned.sort values(by = 'mid90', ascending = False) In [34]: #setting up the figure size plt.figure(figsize=[15,6]) #plotting them separately plt.subplot(1,3,1)sns.boxplot(x="region", y="start med", data=region salary sorted, color="b") plt.subplot(1,3,2)sns.boxplot(x="region", y="mid_med", data=region_salary_sorted, color="g") plt.subplot(1,3,3)sns.boxplot(x="region", y="mid90", data=region_salary_sorted, color="y") <AxesSubplot:xlabel='region', ylabel='mid90'> Out[34]: 300000 70000 120000 250000 60000 100000 start_med 200000 50000 150000 60000 40000 100000 40000 Northeaster@aliforniaMidwesternSouthern Western NortheastenCaliforniaMidwesternSouthern Western Northeaster@aliforniaMidwesternSouthern Western region region region In [33]: #plotting salaries #setting up the figure size plt.figure(figsize=[10,10]) #the plotting sns.boxplot(x="region", y="start med", data=region salary sorted, color="b") sns.boxplot(x="region", y="mid_med", data=region_salary_sorted, color="g") sns.boxplot(x="region", y="mid90", data=region_salary_sorted, color="y") #adding in extra features/parameters plt.xlabel("Regions", size=16) plt.ylabel("Salary (USD)", size=16) plt.title("Salary Distribution for Each Region", size=19) colors = ['#dfb31f','#75bf71','#597dbf'] lines = [Line2D([0], [0], color=c, linewidth=7) for c in colors] labels = ["90th Percentile Mid-Career Salary", 'Mid-Career Median Salary', 'Starting Median Salary'] plt.legend(lines, labels) <matplotlib.legend.Legend at 0x7fd0013bc190> Out[33]: Salary Distribution for Each Region 90th Percentile Mid-Career Salary Mid-Career Median Salary Starting Median Salary 300000 250000 Salary (USD) 200000 150000 100000

50000

Northeastern

actually make so much more.

highest median salary.

References

Discussion and Conclusion

most beneficial if the university was from the northeast region.

California

Midwestern

Regions

Southern

At the start of our careers, the most promising major would be a physician's assistant. That major starts off with a median salary 74,300, beating all the engineering majors, computer science, and nursing. However, it should be noted that a physician's assistant usually have to spend six years in school, compared to the average four undergraduate years. More importantly, salaries don't stay the same as our careers

progress. Unfortunately for physician's assistants, they only experience a 23.4% increase in salaries from starting to mid-career stretch. When looking at the long run, many different engineering majors top the chart with numbers breaking the 100,000 bracket. The top ten also sees an addition in economics, physics, and math majors. Moreover, looking at the 90th percentile for mid-career salary outcomes, economic and finance majors actually come out in the top two spots. It does make sense for prestigious jobs dealing with money to

Next are the results for the region of colleges. This dataset listed a total of five different regions: northeast, midwest, south, west, and California with a category of its own. For the starting salary, Californian colleges has the highest median starting salaries but Northeastern schools have highest maximum salary. As we transition to the mid-career stage, it is still the same trend. However, in the 90th percentile for mid-career, the Northeast's median and maximum dramatically passes the other four regions. This is most likely due to the economic and finance majors working near New York City. The region with the lowest salary for the 90th percentile is the Midwest but it has the 3rd

In summary, I learned that economic and finance majors can make the most at the 90th percentile during the long run. It would be the

I would also learn how to make the straight lines in the boxplots for regional salary be in different colors for easier differentiation.

Journal, The Wall Street. "Where It Pays to Attend College." Kaggle, 29 Apr. 2017, https://www.kaggle.com/wsj/college-salaries.

"Seaborn.boxplot"." Seaborn.boxplot - Seaborn 0.11.2 Documentation, https://seaborn.pydata.org/generated/seaborn.boxplot.html.

If I were do to anything different, I would like to find a dataset that can show salary data up until the senior stages to complete the full run.

Western