A Deep Dive Into Higher Education

Katherine Potter

Abstract

Using datasets from Statista.com, DataHub.io, Data.World, & NSF.gov, the following analysis aims to determine how higher education trends have changed since the early 1900’s. Exploratory methods of analysis as well as mean and other Python and Pandas methods led to numerous findings including that the Business major has yielded the highest rates of employment for recent graduates. While, Engineering majors i.e. Aerospace, Chemical, Civil Engineering, etcetera, are among the highest paying career paths, the employment rates are somewhat low, particularly for women, & as a result, the graduation rates for Engineering majors fall to the middle of the pack for men and are at the bottom for women. Finally, as expected, the number of people participating in the agricultural industry has declined as the number of people in non labor fields has steadily increased.

Motivation

The workforce has changed dramatically over the last 50 years. More women are working than ever before and more men are diversifying their field of choice. The number of people pursuing graduate and doctoral degrees is also increasing as certain fields are becoming more and more competitive.

Throughout this analysis, I plan to put numbers to these facts. I plan to learn which college majors were male dominated in previous years, which were female dominated, and how those numbers have changed in recent years.

I plan to learn which college majors provide the highest earning potential, how many people are earning graduate and doctoral degrees, and which degrees yield the highest rates of employment. Are there any patterns that may provide insight as to why people are choosing certain majors?

The answers to the questions above could be beneficial to prospective students and college administrators as knowing which degrees are the most popular may help to determine where money should be spent and may advise of the level of competition in a certain course of study.

Dataset(s)

Throughout this analysis, I utilized the following datasets:

- Dataset 1: **Percentage of the U.S. population with a college degree 1940-2018, by gender**
  - This dataset includes the percentage of the U.S. population that earned a college degree from 1940 to 2018 broken down by gender
    - Source: Statista > Statistics > Education & Science | https://www.statista.com/statistics/185157/number-of-bachelor-degrees-by-gender-since-1950/
- Dataset 2: **Bachelor's degrees earned in the United States by gender 1950-2028**
  - This dataset includes the number of individuals that earned a bachelor's degree from 1950 to a projected 2028 broken down by gender
    - Source: Statista > Statistics > Education & Science | https://www.statista.com/statistics/185157/number-of-bachelor-degrees-by-gender-since-1950/
- Dataset 3: **Master's degrees earned in the United States by gender 1949/50-2027/28**
  - This dataset includes the number of individuals that earned a master's degree from 1949/50 to a projected 2027/2028 broken down by gender
    - Source: Statista > Statistics > Education & Science | https://www.statista.com/statistics/185160/number-of-masters-degrees-by-gender-since-1950/
- Dataset 4: **Doctoral degrees earned in the United States by gender 1949/50-2027/28**
  - This dataset includes the number of individuals that earned a doctoral degree from 1949/50 to a projected 2027/2028 broken down by gender
    - Source: Statista > Statistics > Education & Science | https://www.statista.com/statistics/185167/number-of-doctoral-degrees-by-gender-since-1950/
- Dataset 5: **US Employment and Unemployment rates since 1940**
  - This dataset includes the employment & unemployment rates of people in agriculture, not in agriculture, and not in labor from 1940 to 2010
    - Source: Data Hub.io > Core > Employment > US | https://datahub.io/core/employment-us#data
- Dataset 6: **Employment by College Major**
  - This dataset includes the employment & unemployment rates of recent graduates with a variety of college majors
    - Source: Data.world > Five Thirty Eight > College-Majors | https://data.world/fivethirtyeight/college-majors/file/recent-grads.csv
- Dataset 7: **Salary by College Major**
  - This dataset includes the starting median & mid-career median salary as well as the percentage change from starting to mid-career median salary by college major
    - Source: Kaggle.com > WSJ > College-Salaries | https://www.kaggle.com/wsj/college-salaries
- Dataset 8 & 8.1: **Bachelor's degrees by College Major (two formats)**
  - This dataset includes the number of bachelor's degrees earned from 2004 – 2014 by gender and college major
    - Source: NSF.gov > Data | https://www.nsf.gov/statistics/2017/nsf17310/data.cfm

** I utilized the entirety of each dataset throughout this analysis. The shape (amount of data) of each dataset can be found in the top right corner of this slide.*

Size of Datasets 1 - 8

```
Dataset 1:
(64, 4)
Dataset 2:
(63, 3)
Dataset 3:
(63, 3)
Dataset 4:
(63, 3)
Dataset 5:
(71, 12)
Dataset 6:
(173, 21)
Dataset 7:
(115, 12)
Dataset 8:
(55, 14)
```

Data Preparation and Cleaning

Each dataset that I utilized in this analysis required different cleaning methods.

- Dataset 1: **Percentage of the U.S. population with a college degree 1940-2018, by gender**
    - For Dataset 1, I began by removing the top two header rows and renamed the columns. I then converted the year and gender columns to floats and saved them to their own DataFrame. This left me with a clean, easy to graph DataFrame
- Dataset 2: **Bachelor's degrees earned in the United States by gender 1950-2028**
    - Dataset 2 required more cleaning than the first. I started as with the first table by removing the top two header rows and renaming the columns. After that step, I broke the newly named "Year Range" column into "Lower Year" & "Upper Year" for graphing purposes
        - Dataset 2 provided a few challenges primarily due to the "Year Range" column. The upper year in each range did not includes a decade as each range was a string. To resolve this issue, I split the columns as described above
        - The other issue I encountered with Dataset 2 were the commas in the gender columns. I resolved this issue by replacing all commas with blanks
- Dataset 3: **Master's degrees earned in the United States by gender 1949/50-2027/28**
    - Dataset 3 required the exact same cleaning and issue resolution as Dataset 2
- Dataset 4: **Doctoral degrees earned in the United States by gender 1949/50-2027/28**
    - Dataset 4 required the exact same cleaning and issue resolution as Dataset 2
        - ✓ **After cleaning Datasets 1-4, I merged them into one, comprehensive DataFrame for ease of use and visualizations**
- Dataset 5: **US Employment and Unemployment rates since 1940**
    - Dataset 5 required almost no cleaning. The only element that needed to be changed was one misspelling of the word agriculture in the original dataset
- ❑ Dataset 6: **Employment by College Major**
    - Dataset 6 required very little cleaning. I removed the "Rank", "Major_code", & "Share_women" columns as I would not be utilizing that information and sorted the DataFrame by the highest to lowest value in the "Total" column
- ❑ Dataset 7: **Salary by College Major**
    - Dataset 7 required very little cleaning. I replaced all punctuation with blanks and converted the salary columns to floats and was ready to explore the data
- ❑ Dataset 8 & 8.1: **Bachelor's degrees awarded, by sex and field: 2004–14 (two formats)**
    - Datasets 8 & 8.1 required similar cleaning to Dataset 2 in that the top 3 rows had to be removed and renamed
    - I also sliced the DataFrame into 3 DataFrames: male, female, & both sexes as the genders were listed horizontally with the years listed beneath each

Research Question(s)

- **How have trends in higher education changed since the early 1900's?**
    - How many people have pursued higher education as technology has developed and agriculture is no longer the most important industry?
    - How has the ratio of men to women in engineering fields changed?
    - Which college majors were the most popular over the years?
    - Which college majors yield the highest rates of employment?
    - Which college majors yield the most earning potential?

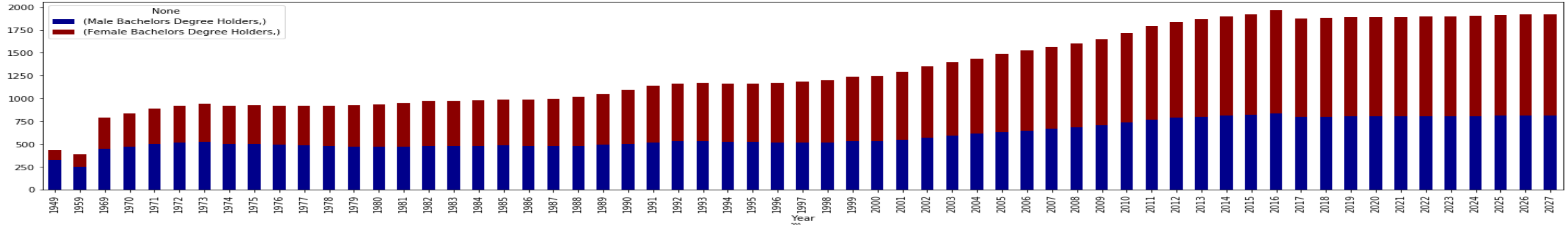
Methods

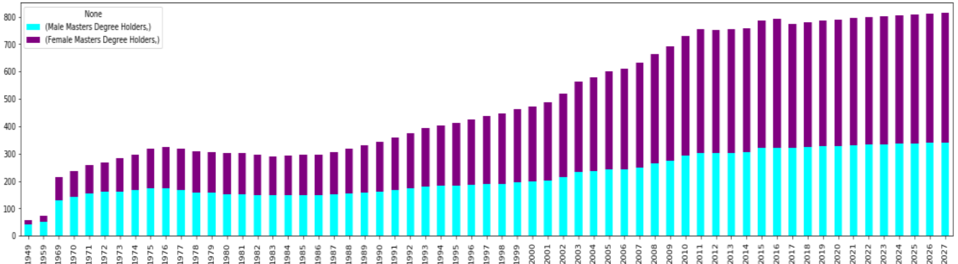
- For this analysis, I used exploratory methods to analyze the data
    - These methods were appropriate because:
        - I set out to discover what factors have contributed to the change in higher education trends as well as employment rates
            - In order to answer this question, a deep dive into multiple datasets and an analysis of the connections between each dataset was required.
        - I utilized exploratory methods by examining and cleaning each dataset to discover what information each dataset could provide
            - I then merged, graphed, sorted, and otherwise manipulated the data in order to create a final product of clean, succinct, valuable data that could be used to answer my research questions
            - I also utilized a number of Pandas methods such as mean, idxmax, drop_duplicates, & more throughout my analysis
                - All Python & Pandas methods were used to clean, manipulate, & transform the data to allow for more exploratory analysis

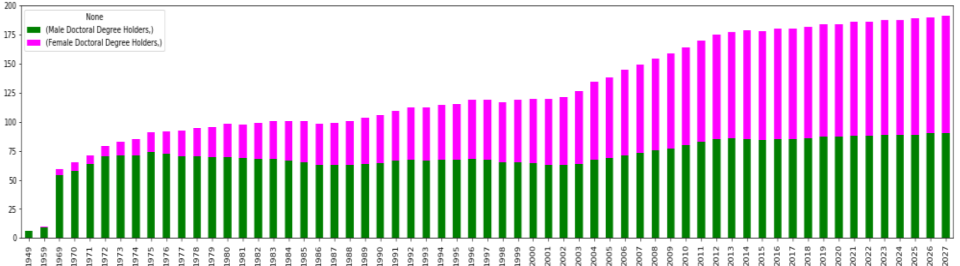
Findings: Higher Education Trends - Men vs Women

- My findings for this analysis fell into buckets
  - First, I compared the number of men and women who have earned bachelors, masters, & graduate degrees from 1940 to 2018
  - The charts below supported my hypothesis that more people tend to pursue bachelor's degrees than master's degrees and more pursue master's degrees than doctoral degrees
    - Similarly, as I expected, the number of men pursuing each degree level is higher than the number of women at the beginning of the timespan while the distribution has become more even over time
      - The number of women even surpassed the number of men earning bachelors and masters degrees in the early to mid 1980's and surpassed the number of men earning doctoral degrees in 2006

Data Source: Datasets 1 – 4

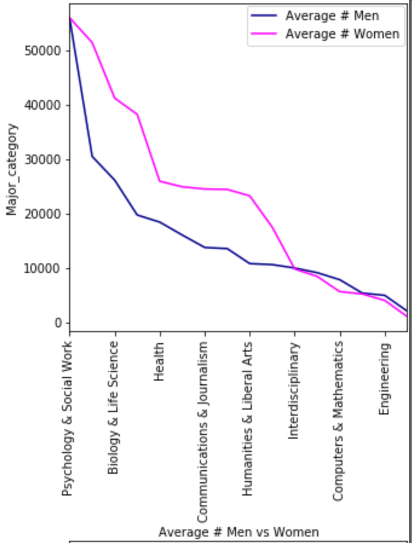


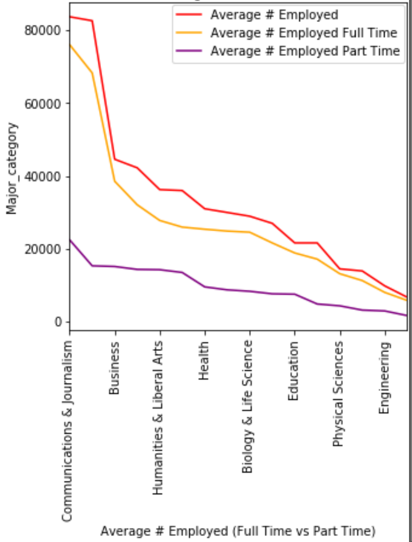




Findings: Employment Rates by Major & Gender

- For the second part of my analysis, I dug into the employment rates datasets
  - Datasets 5 & 6 provided a wealth of information
  - For instance, my analysis revealed that Psychology & Social Work, Social Science, and Biology & Life Science were the top three most employable majors amongst men & women
    - Interestingly yet unsurprisingly, Engineering was the 6th most employable major for men, with over 13,815 men finding work in the field, while serving as the 14th (second to last) most employable major for women, with only 4,070 women reported as having found work in the field
    - Interestingly, Computer & Mathematics also ranked low for both men and women, with only 5,420 men and 5,690 women finding work in those fields
      - My analysis of the specified datasets also revealed that the college major with the best employment rates overall was Business
        - Over 76,066 people found full time employment in the Business field, 13,609 of whom were reported as men and 8,489 of whom were reported as women (I believe the discrepancy in the numbers can be attributed to the available choice not to report a gender)
  - Business ranked third when it came to part time employment as well
  - Communications & Journalism ranked #1 for part time employment, followed by Psychology & Social Work
  - Computers & Mathematics and Engineering came in low again in the overall employment rate analysis, with only 21,626 and 14,495 individuals finding work in these fields, respectively





Data Source : Dataset 6

Findings: Salary Potential & Popularity by College Major

- In the third phase of my analysis, I compared the mean graduation rate by degree to the percent change from starting to mid-career salary
- Unsurprisingly, there seems to be a strong correlation between the graduation rate and salary potential of certain college majors
  - For instance, engineering has the 7th highest graduation rate over the years while Aerospace, Computer, Chemical, Electrical, & Civil Engineering all fall within the top 25 majors with the highest salary growth potential
  - Economics, Physics, Chemistry, Psychology, Anthropology, & Sociology also appear in the top 25 majors in both datasets
  - ** One limitation of the datasets I worked with is the diversity of naming conventions
    - Dataset 7 included a more detailed breakdown of the Engineering field in particular, while Dataset 8 used a more high level grouping such as "Mathematics & Statistics" as opposed to listing the two majors out separately
    - Another limitation that I faced was the lack of available public salary records by year and college major. Without such data, no statistical correlation could calculated

	Major	Mean
0	All fields	1.643906e+06
1	Non-S&E	1.115218e+06
2	S&E	5.286876e+05
3	Science	4.536545e+05
4	Social sciences	1.604335e+05
5	Psychology	9.802055e+04
6	Biological sciences	8.640055e+04
7	Engineering	7.503309e+04
8	Political science and public administration	5.600991e+04
9	Computer sciences	4.744900e+04
10	Sociology	2.925709e+04
11	Other	2.922391e+04
12	Economics	2.760973e+04
13	Agricultural sciences	2.103918e+04
14	Mechanical engineering	1.852909e+04
15	Electrical engineering	1.849200e+04
16	Physical sciences	1.815764e+04
17	Mathematics and statistics	1.726600e+04
18	Civil engineering	1.283100e+04
19	Chemistry	1.208927e+04
20	Other	1.022955e+04
21	Anthropology	9.374455e+03
22	Area and ethnic studies	7.236545e+03
23	Chemical engineering	6.600364e+03
24	Physics	5.060091e+03

Data Source : Dataset 8

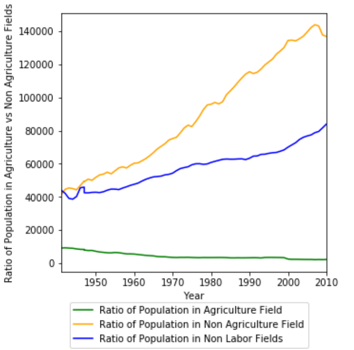
	Undergraduate Major	Percent change from Starting to Mid-Career Salary
0	Math	103.5
1	Philosophy	103.5
2	International Relations	97.8
3	Economics	96.8
4	Marketing	95.1
5	Physics	93.4
6	Political Science	91.7
7	Chemistry	87.6
8	Journalism	87.4
9	Architecture	84.6
10	Finance	84.3
11	Communications	83.7
12	Geology	82.8
13	Art History	81.3
14	History	81.1
15	Film	80.7
16	Aerospace Engineering	75.0
17	Computer Engineering	71.0
18	Computer Science	70.8
19	English	70.3
20	Chemical Engineering	69.3
21	Electrical Engineering	69.1
22	Agriculture	68.8
23	Psychology	68.2
24	Civil Engineering	67.9

Data Source : Dataset 7

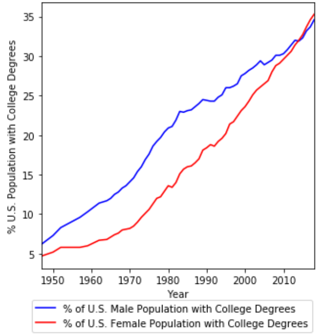
Limitations

- There were limitations to each of the datasets that I explored as well as to the amount of data that is publicly available regarding salaries
  - Data diversity:
    - I used datasets from numerous sources, which resulted in the occasional mismatch in naming conventions or connections between the data
    - I resolved this by joining certain datasets as well as slicing others to provide workable datasets that could be compared
  - Availability of salary data:
    - Initially I set out to compare the salaries of certain career paths over the years to the trend in the number of people pursuing corresponding college majors
    - However, I was unable to find a publicly available dataset that explored the earning potential of college majors over a timespan of any kind
    - I found datasets like dataset 7, that explored the starting & mid-career median salaries by college major, but none that provided a trend of changes in the earning potential of each major
      - As a result of this limitation, I added the employment rates datasets, providing a third dimension to the analysis and more factors to compare to the college major trend

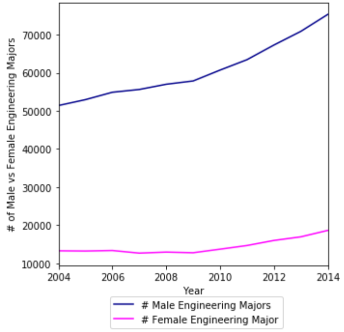
Conclusions : Higher Education Trends & College Major Popularity



- **How have trends in higher education changed since the early 1900's?**
  - *How many people have pursued higher education as technology has developed and agriculture is no longer the most important industry?*
    - As the chart in the upper right corner as well as those on slide 8 show, the number of people participating in the agriculture industry has decreased since 1950, while the number of people working in non agriculture fields has drastically increased
    - The % of the male & female populations earning college degrees has faced a similarly drastic and consistent increase



  - *How has the ratio of men to women majoring in engineering changed?*
    - Though the ratio of men to women majoring in engineering has changed, the changes are not extreme
      - There are still many more men obtaining engineering degrees than there are women and the increase of male degree earners dwarfs that of the females
  - *Which college majors were the most popular over the years?*
    - While coming to a conclusion for this part of the question, I realized that cleaning and splicing data set 8 would take unnecessary time because NSF.gov offers the exact same data in a transposed format
      - Reading in what I will call Dataset 8.1 allowed me to learn very quickly that Science & Engineering was the most popular major category in every year between 2004 & 2014
        - Next most popular major was Science followed by Social Science, Psychology, Biological Science, & Political Science



Conclusions : Rates of Employment & Earning Potential

- **How have trends in higher education changed since the early 1900’s?**
  - *Which college majors yield the highest rates of employment?*
    - As discussed on slide 9, the Business major yields the highest overall rate of employment
      - Communications & Journalism, Social Sciences, Psychology & Social Work, and Humanities & Liberal Arts round out the top 5 most employable majors in Dataset 6
      - Biology & Life Sciences, Engineering, and Computers & Mathematics are also in the top 20 most employable majors while also making appearances in the most popular majors dataset as well
  - *Which college majors yield the most earning potential?*
    - Interestingly, Physician Assistant took the number one top spot when it came to starting median salary
    - The rest of the top 10 top earners right out of college were Chemical, Computer, Electrical, Mechanical, Aerospace, & Industrial Engineering, Computer Science, Nursing, & Civil Engineering

Data Source: Dataset 7

Undergraduate Major	Starting Median Salary
Physician Assistant	74300.00
Chemical Engineering	63200.00
Computer Engineering	61400.00
Electrical Engineering	60900.00
Mechanical Engineering	57900.00
Aerospace Engineering	57700.00
Industrial Engineering	57700.00
Computer Science	55900.00
Nursing	54200.00
Civil Engineering	53900.00

Undergraduate Major	Mid-Career Median Salary
Economics	98600.00
Physics	97300.00
Computer Science	95500.00
Industrial Engineering	94700.00
Mechanical Engineering	93600.00
Math	92400.00
Physician Assistant	91700.00
Civil Engineering	90500.00
Construction	88900.00
Finance	88300.00

Major	% Change from Starting to Mid-Career Salary
Math	103.5
Philosophy	103.5
International Relations	97.8
Economics	96.8
Marketing	95.1
Physics	93.4
Political Science	91.7
Chemistry	87.6
Journalism	87.4
Architecture	84.6

Data Source: Dataset 6

Major	Men	Women	Employed
Business	13609.000000	8489.769231	83749.384615
Communications & Journalism	9173.000000	24569.500000	82665.000000
Social Science	55928.555556	51512.888889	44610.333333
Psychology & Social Work	26205.222222	58073.555556	42260.444444
Humanities & Liberal Arts	10663.000000	23309.066667	36274.533333
Arts	10847.625000	17558.625000	36014.250000
Health	7885.833333	26002.166667	31012.250000

Conclusions : Final Thoughts

- This analysis was incredibly enlightening in that many illusions that I held were shattered
- The popularity & earning potential of the Business & Psychology majors were not surprising to me, but the relatively low number of people that pursue the Engineering major was unexpected
- The discrepancy between men & women in the Engineering & Computer Science fields was predictable, but after this analysis, I would attribute the low number of people pursuing the major overall to the competitiveness of the industry
- Pair the cost of an Engineering education with the extensive skillsets that acquiring and maintaining a job in the industry requires and it makes sense that, despite the high salaries that jobs in the Engineering field can offer, relatively few people choose that path
- The other surprise, in my opinion, was the Math major
- While not all of the datasets I explored listed a general “Math” major, it did hold the top spot in the percentage change from starting to mid-career median salaries
- Philosophy was similar in that the major did not stand out in any dataset other than the percentage change/earning growth potential dataset
- Overall, I believe that the largest contributing factor to college major selection seems to be employment rates
- A large number of people seem to forgo the jobs that yield the most earning potential in favor of those that may provide more stability and a better chance at securing a job in the first place
- There would be even more information to be gleaned from the datasets I utilized if a dataset providing salary or employment by major by year becomes available

Acknowledgements

My datasets came from a variety of sources:

- Datasets 1 – 4 can be found on Statista.com (direct links can be found on slide 4)
- Dataset 5 can be found on DataHub.io (direct link can be found on slide 4)
- Dataset 6 can be found on Data.world (direct link can be found on slide 4)
- Dataset 7 can be found on Kaggle.com (direct link can be found on slide 4)
- Dataset 8 & 8.1 can be found on NSF.gov (direct link can be found on slide 4)

Special Thanks to the data scientists on my team who were gracious enough to listen to my project idea and hash it out with me as well as to lend a pair of fresh eyes and a bit of input whenever I asked.

References

The only reference materials used throughout this analysis were the datasets listed on slide 4 and acknowledged on slide 15.

All data cleaning, exploring, manipulation, and visuals were completed on my own

Import Libraries

In [33]:
```
import pandas as pd
import numpy as np
import os.path
import matplotlib.pyplot as plt
import matplotlib
import sys
import locale
from functools import reduce
```

Import College Degree Datasets

In [34]:
```
DEGREEANALYSIS = "C:/Users/kacco/OneDrive/Documents/Python_For_Data_Science/We
ek_9_Final_Project/"

# Percentage of the U.S. population with a college degree 1940-2018, by gender
percentbygender = pd.read_csv(os.path.join(DEGREEANALYSIS, 'statistic_id184272
_percentage-of-the-us-population-with-a-college-degree-1940-2018-by-gender.cs
v'), sep=',')

# Bachelor's degrees earned in the United States by gender 1950-2028
bachelorsbygender= pd.read_csv(os.path.join(DEGREEANALYSIS, 'statistic_id18515
7_bachelors-degrees-earned-in-the-united-states-by-gender-1950-2028.csv'), sep
=',')

# Master's degrees earned in the United States by gender 1950-2028
mastersbygender = pd.read_csv(os.path.join(DEGREEANALYSIS, 'statistic_id185160
_masters-degrees-earned-in-the-united-states-by-gender-1949-50-2027-28.csv'),
sep=',')

# Doctoral degrees earned in the United States by gender 1950-2028
doctoralbygender = pd.read_csv(os.path.join(DEGREEANALYSIS, 'statistic_id18516
7_doctoral-degrees-earned-in-the-united-states-by-gender-1950-2028.csv'), sep=
',')
```

Import Employment Rates & Salary Datasets

In [35]:
```
EMPLOYMENTANALYSIS = "C:/Users/kacco/OneDrive/Documents/Python_For_Data_Scienc
e/Week_9_Final_Project/"

# US Employment and Unemployment rates since 1940
employmentbyyear = pd.read_csv(os.path.join(EMPLOYMENTANALYSIS, 'employment_by
_year.csv'), sep=',')

# Employment rates by college major
employmentbymajor= pd.read_csv(os.path.join(EMPLOYMENTANALYSIS, 'recent-grads.
csv'), sep=',')

# Salaries by college major
salarybymajor = pd.read_csv(os.path.join(EMPLOYMENTANALYSIS, 'degrees-that-pay
-back.csv'), sep=',')

# Bachelor's degrees awarded, by sex and field: 2004-14
bachelorsdegreesbymajor = pd.read_csv(os.path.join(EMPLOYMENTANALYSIS, 'Bachel
ors_degrees_awarded_by_field_and_sex_2004-14.csv'), sep=',')

# Bachelor's degrees awarded, by sex and field: 2004-14
bachelorsdegreesbymajor_v2 = pd.read_csv(os.path.join(EMPLOYMENTANALYSIS, 'Bac
helors_degrees_awarded_by_sex_field_2004-14.csv'), sep=',')
```

Dataset size calculation

In [36]:
```
print('Dataset 1:')
print(percentbygender.shape)
print('Dataset 2:')
print(bachelorsbygender.shape)
print('Dataset 3:')
print(mastersbygender.shape)
print('Dataset 4:')
print(doctoralbygender.shape)
print('Dataset 5:')
print(employmentbyyear.shape)
print('Dataset 6:')
print(employmentbymajor.shape)
print('Dataset 7:')
print(bachelorsdegreesbymajor.shape)
print('Dataset 8:')
print(bachelorsdegreesbymajor_v2.shape)
```

```
Dataset 1:
(64, 4)
Dataset 2:
(63, 3)
Dataset 3:
(63, 3)
Dataset 4:
(63, 3)
Dataset 5:
(71, 12)
Dataset 6:
(173, 21)
Dataset 7:
(115, 12)
Dataset 8:
(55, 14)
```

Degree Analysis : Degrees Earned by Level, Gender, & Year

The Percentage of the U.S. Male & Female Population with Bachelor's Degrees by Year

In [37]:

```
percentbygender = percentbygender.iloc[2:]
percentbygender.columns = ['Year','Male','Female','%']
percentbygender[['Male','Female']] = percentbygender[['Male','Female']].astype
(float)
percentbygender
```

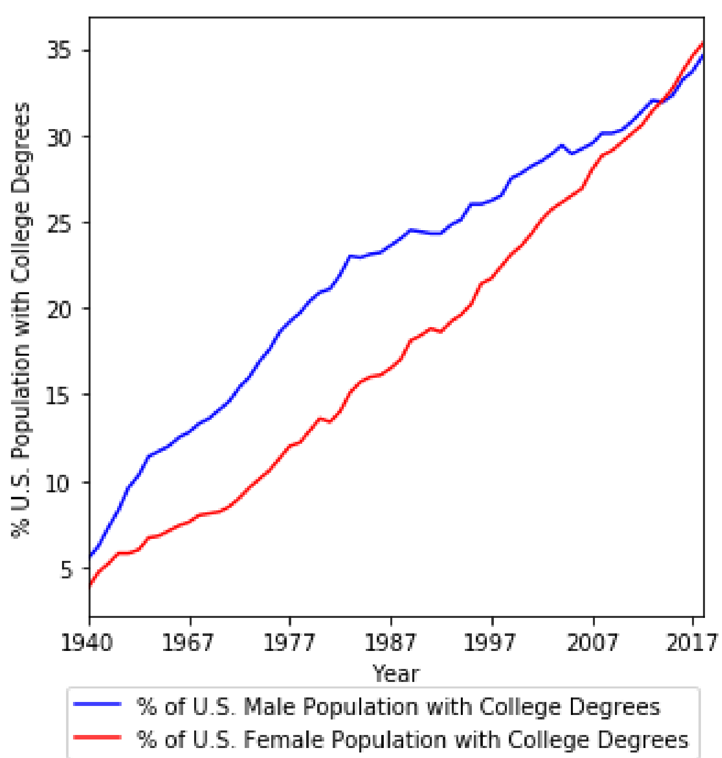
Out[37]:

	Year	Male	Female	%
2	1940	5.5	3.8	in %
3	1947	6.2	4.7	in %
4	1950	7.3	5.2	in %
5	1952	8.3	5.8	in %
6	1957	9.6	5.8	in %
7	1959	10.3	6.0	in %
8	1962	11.4	6.7	in %
9	1964	11.7	6.8	in %
10	1965	12.0	7.1	in %
11	1966	12.5	7.4	in %
12	1967	12.8	7.6	in %
13	1968	13.3	8.0	in %
14	1969	13.6	8.1	in %
15	1970	14.1	8.2	in %
16	1971	14.6	8.5	in %
17	1972	15.4	9.0	in %
18	1973	16.0	9.6	in %
19	1974	16.9	10.1	in %
20	1975	17.6	10.6	in %
21	1976	18.6	11.3	in %
22	1977	19.2	12.0	in %
23	1978	19.7	12.2	in %
24	1979	20.4	12.9	in %
25	1980	20.9	13.6	in %
26	1981	21.1	13.4	in %
27	1982	21.9	14.0	in %
28	1983	23.0	15.1	in %
29	1984	22.9	15.7	in %
30	1985	23.1	16.0	in %
31	1986	23.2	16.1	in %
...	...	...	...	...
34	1989	24.5	18.1	in %
35	1990	24.4	18.4	in %
36	1991	24.3	18.8	in %
37	1992	24.3	18.6	in %
38	1993	24.8	19.2	in %
39	1994	25.1	19.6	in %
40	1995	26.0	20.2	in %
41	1996	26.0	21.4	in %
42	1997	26.2	21.7	in %
43	1998	26.5	22.4	in %
44	1999	27.5	23.1	in %
45	2000	27.8	23.6	in %
46	2001	28.2	24.3	in %
47	2002	28.5	25.1	in %

	Year	Male	Female	%
48	2003	28.9	25.7	in %
49	2004	29.4	26.1	in %
50	2005	28.9	26.5	in %
51	2006	29.2	26.9	in %
52	2007	29.5	28.0	in %
53	2008	30.1	28.8	in %
54	2009	30.1	29.1	in %
55	2010	30.3	29.6	in %
56	2011	30.8	30.1	in %
57	2012	31.4	30.6	in %
58	2013	32.0	31.4	in %
59	2014	31.9	32.0	in %
60	2015	32.3	32.7	in %
61	2016	33.2	33.7	in %
62	2017	33.7	34.6	in %
63	2018	34.6	35.3	in %

62 rows × 4 columns

In [38]:
```
ax = percentbygender.plot(kind='line',x='Year',y='Male',color='blue',label='% of U.S. Male Population with College Degrees')
percentbygender.plot(kind='line',x='Year',y='Female',color='red',label='% of U.S. Female Population with College Degrees',ax=ax)

plt.xlabel('Year')
plt.ylabel('% U.S. Population with College Degrees')

ax.legend(bbox_to_anchor=(1.01, -.1))
plt.gcf().set_size_inches((5, 5))
```



The # of Bachelors Degree Holders by Year by Gender (in 1,000)

In [39]:

```
bachelorsbygender = bachelorsbygender.iloc[2:]
bachelorsbygender.columns = ['Year Range','Male','Female']

bachelorsbygender['Lower Year'] = bachelorsbygender['Year Range'].astype(str).
str[:4]
bachelorsbygender['Upper Year'] = bachelorsbygender['Year Range'].astype(str).
str[:2] + bachelorsbygender['Year Range'].astype(str).str[5:7]

bachelorsbygender = bachelorsbygender[['Lower Year','Upper Year','Male','Femal
e']].apply(lambda x: x.str.replace(',','')).astype(float)
bachelorsbygender['Upper Year'] = bachelorsbygender['Upper Year'].replace(190
0.0,2000.0)

bachelorsbygender['Lower Year'] = bachelorsbygender['Lower Year'].astype(int)
bachelorsbygender['Upper Year'] = bachelorsbygender['Upper Year'].astype(int)

bachelorsbygender
```

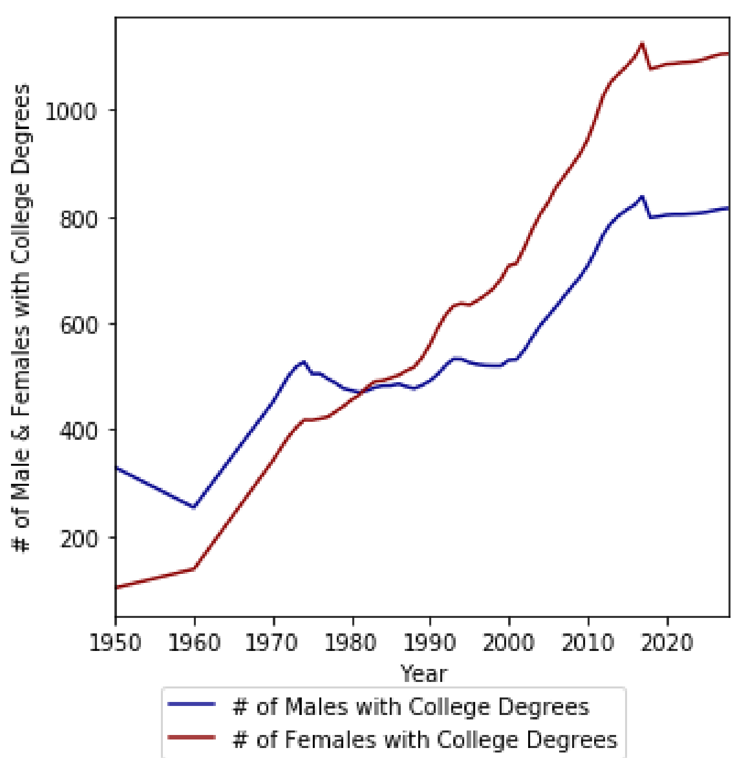
Out[39]:

	Lower Year	Upper Year	Male	Female
2	1949	1950	328.84	103.22
3	1959	1960	254.06	138.38
4	1969	1970	451.10	341.22
5	1970	1971	475.59	364.14
6	1971	1972	500.59	386.68
7	1972	1973	518.19	404.17
8	1973	1974	527.31	418.46
9	1974	1975	504.84	418.09
10	1975	1976	504.93	420.82
11	1976	1977	495.55	424.00
12	1977	1978	487.35	433.86
13	1978	1979	477.34	444.05
14	1979	1980	473.61	455.81
15	1980	1981	469.88	465.26
16	1981	1982	473.36	479.63
17	1982	1983	479.14	490.37
18	1983	1984	482.32	491.99
19	1984	1985	482.53	496.95
20	1985	1986	485.92	501.90
21	1986	1987	480.78	510.48
22	1987	1988	477.20	517.63
23	1988	1989	483.35	535.41
24	1989	1990	491.70	559.65
25	1990	1991	504.05	590.49
26	1991	1992	520.81	615.74
27	1992	1993	532.88	632.30
28	1993	1994	532.42	636.85
29	1994	1995	526.13	634.00
30	1995	1996	522.45	642.34
31	1996	1997	520.52	652.36
...	...	...	...	...
33	1998	1999	519.96	682.28
34	1999	2000	530.37	707.51
35	2000	2001	531.84	712.33
36	2001	2002	549.82	742.08
37	2002	2003	573.26	775.55
38	2003	2004	595.43	804.12
39	2004	2005	613.00	826.26
40	2005	2006	630.60	854.64
41	2006	2007	649.57	874.52
42	2007	2008	667.93	895.14
43	2008	2009	685.42	915.98
44	2009	2010	706.66	943.26
45	2010	2011	734.16	981.89
46	2011	2012	765.77	1026.39

	Lower Year	Upper Year	Male	Female
47	2012	2013	787.41	1052.97
48	2013	2014	801.91	1068.25
49	2014	2015	812.69	1082.28
50	2015	2016	821.78	1098.94
51	2016	2017	838.00	1125.00
52	2017	2018	798.00	1077.00
53	2018	2019	800.00	1081.00
54	2019	2020	803.00	1086.00
55	2020	2021	804.00	1087.00
56	2021	2022	804.00	1089.00
57	2022	2023	805.00	1090.00
58	2023	2024	806.00	1092.00
59	2024	2025	808.00	1096.00
60	2025	2026	811.00	1101.00
61	2026	2027	814.00	1105.00
62	2027	2028	816.00	1106.00

61 rows × 4 columns

In [40]:
```
ax = bachelorsbygender.plot(kind='line',x='Upper Year',y='Male',color='darkblue',label='# of Males with College Degrees')
bachelorsbygender.plot(kind='line',x='Upper Year',y='Female',color='darkred',label='# of Females with College Degrees',ax=ax)

plt.xlabel('Year')
plt.ylabel('# of Male & Females with College Degrees')

ax.legend(bbox_to_anchor=(.85, -.1))
plt.gcf().set_size_inches((5, 5))
```



The # of Masters Degree Holders by Year by Gender (in 1,000)

In [41]:
```
mastersbygender = mastersbygender.iloc[2:]
mastersbygender.columns = ['Year Range','Male','Female']

mastersbygender['Lower Year'] = mastersbygender['Year Range'].astype(str).str
[:4]
mastersbygender['Upper Year'] = mastersbygender['Year Range'].astype(str).str
[:2] + mastersbygender['Year Range'].astype(str).str[5:7]

mastersbygender = mastersbygender[['Lower Year','Upper Year','Male','Female']]
.apply(lambda x: x.str.replace(',','')).astype(float)
mastersbygender['Upper Year'] = mastersbygender['Upper Year'].replace(1900.0,2
000.0)

mastersbygender['Lower Year'] = mastersbygender['Lower Year'].astype(int)
mastersbygender['Upper Year'] = mastersbygender['Upper Year'].astype(int)

mastersbygender
```

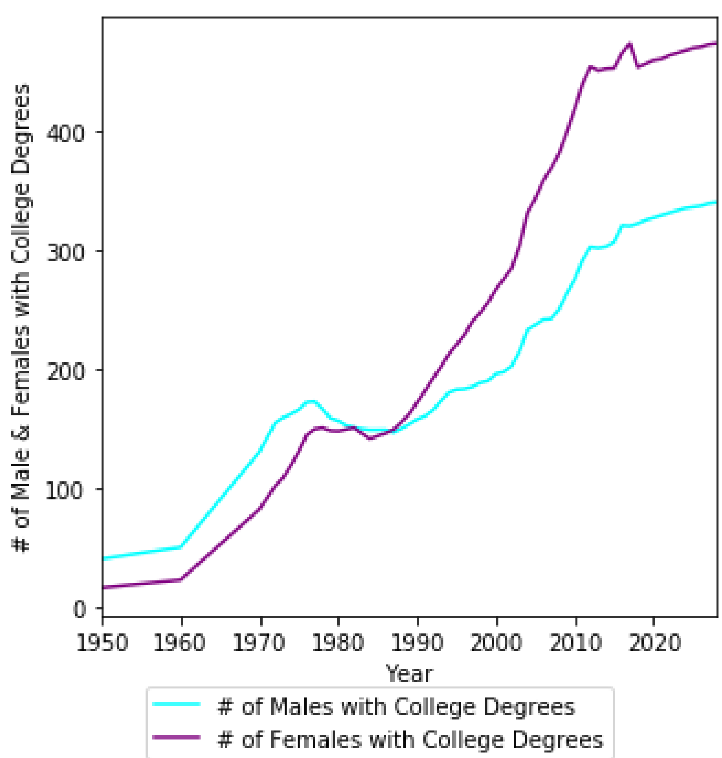
Out[41]:

	Lower Year	Upper Year	Male	Female
2	1949	1950	41.22	16.96
3	1959	1960	50.90	23.54
4	1969	1970	130.80	82.79
5	1970	1971	143.08	92.48
6	1971	1972	155.01	102.19
7	1972	1973	159.57	109.09
8	1973	1974	162.61	119.47
9	1974	1975	166.32	131.23
10	1975	1976	172.52	144.96
11	1976	1977	173.09	149.94
12	1977	1978	166.86	151.13
13	1978	1979	159.11	148.58
14	1979	1980	156.88	148.31
15	1980	1981	152.98	149.66
16	1981	1982	151.35	151.10
17	1982	1983	150.09	146.32
18	1983	1984	149.27	141.87
19	1984	1985	149.28	144.20
20	1985	1986	149.37	146.48
21	1986	1987	147.06	149.47
22	1987	1988	150.24	155.54
23	1988	1989	153.99	162.63
24	1989	1990	158.05	172.10
25	1990	1991	160.84	182.02
26	1991	1992	165.87	192.22
27	1992	1993	173.35	201.68
28	1993	1994	180.57	212.47
29	1994	1995	183.04	220.57
30	1995	1996	183.48	228.70
31	1996	1997	185.27	239.99
...	...	...	...	...
33	1998	1999	190.23	255.81
34	1999	2000	196.13	267.06
35	2000	2001	197.77	275.73
36	2001	2002	202.60	284.71
37	2002	2003	215.17	303.53
38	2003	2004	233.06	331.22
39	2004	2005	237.16	343.00
40	2005	2006	241.66	358.08
41	2006	2007	242.19	368.41
42	2007	2008	250.17	380.50
43	2008	2009	263.52	398.57
44	2009	2010	275.32	418.00
45	2010	2011	291.68	439.24
46	2011	2012	302.48	453.48

	Lower Year	Upper Year	Male	Female
47	2012	2013	301.55	450.17
48	2013	2014	302.85	451.74
49	2014	2015	306.62	452.19
50	2015	2016	320.44	465.15
51	2016	2017	320.00	473.00
52	2017	2018	322.00	453.00
53	2018	2019	325.00	456.00
54	2019	2020	327.00	459.00
55	2020	2021	329.00	460.00
56	2021	2022	331.00	463.00
57	2022	2023	333.00	465.00
58	2023	2024	335.00	467.00
59	2024	2025	336.00	469.00
60	2025	2026	337.00	470.00
61	2026	2027	339.00	472.00
62	2027	2028	340.00	473.00

61 rows × 4 columns

In [42]:
```
ax = mastersbygender.plot(kind='line',x='Upper Year',y='Male',color='aqua',label='# of Males with College Degrees')
mastersbygender.plot(kind='line',x='Upper Year',y='Female',color='purple',label='# of Females with College Degrees',ax=ax)

plt.xlabel('Year')
plt.ylabel('# of Male & Females with College Degrees')

ax.legend(bbox_to_anchor=(.85, -.1))
plt.gcf().set_size_inches((5, 5))
```



The # of Doctoral Degree Holders by Year by Gender (in 1,000)

In [43]:

```
doctoralbygender = doctoralbygender.iloc[2:]
doctoralbygender.columns = ['Year Range','Male','Female']

doctoralbygender['Lower Year'] = doctoralbygender['Year Range'].astype(str).str[:4]
doctoralbygender['Upper Year'] = doctoralbygender['Year Range'].astype(str).str[:2] + doctoralbygender['Year Range'].astype(str).str[5:7]

doctoralbygender = doctoralbygender[['Lower Year','Upper Year','Male','Female']].apply(lambda x: x.str.replace(',','')).astype(float)
doctoralbygender['Upper Year'] = doctoralbygender['Upper Year'].replace(1900.0,2000.0)

doctoralbygender['Lower Year'] = doctoralbygender['Lower Year'].astype(int)
doctoralbygender['Upper Year'] = doctoralbygender['Upper Year'].astype(int)

doctoralbygender
```

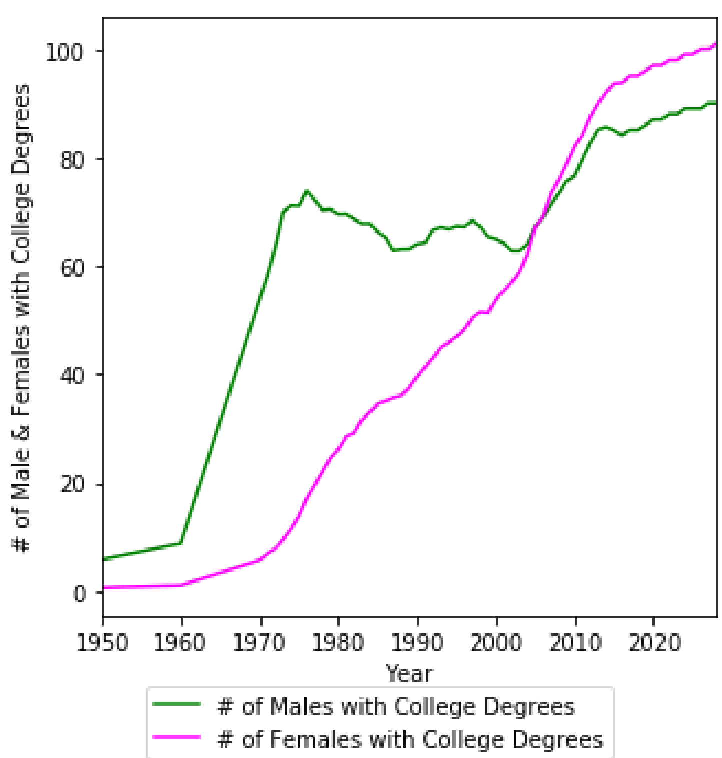
Out[43]:

	Lower Year	Upper Year	Male	Female
2	1949	1950	5.80	0.62
3	1959	1960	8.80	1.03
4	1969	1970	53.79	5.69
5	1970	1971	58.14	6.86
6	1971	1972	63.35	7.85
7	1972	1973	69.96	9.55
8	1973	1974	71.13	11.46
9	1974	1975	71.03	13.88
10	1975	1976	73.89	17.12
11	1976	1977	72.21	19.52
12	1977	1978	70.28	22.06
13	1978	1979	70.45	24.52
14	1979	1980	69.53	26.11
15	1980	1981	69.57	28.45
16	1981	1982	68.63	29.21
17	1982	1983	67.76	31.58
18	1983	1984	67.77	33.03
19	1984	1985	66.27	34.52
20	1985	1986	65.22	35.07
21	1986	1987	62.79	35.69
22	1987	1988	63.02	36.12
23	1988	1989	63.06	37.52
24	1989	1990	63.96	39.55
25	1990	1991	64.24	41.31
26	1991	1992	66.60	42.95
27	1992	1993	67.13	44.94
28	1993	1994	66.77	45.86
29	1994	1995	67.32	46.94
30	1995	1996	67.19	48.32
31	1996	1997	68.39	50.36
...	...	...	...	...
33	1998	1999	65.34	51.36
34	1999	2000	64.93	53.81
35	2000	2001	64.17	55.41
36	2001	2002	62.73	56.93
37	2002	2003	62.73	58.85
38	2003	2004	63.98	62.11
39	2004	2005	67.26	67.13
40	2005	2006	68.91	69.14
41	2006	2007	71.31	73.38
42	2007	2008	73.45	75.93
43	2008	2009	75.67	78.89
44	2009	2010	76.61	81.98
45	2010	2011	79.67	84.16
46	2011	2012	82.67	87.55

	Lower Year	Upper Year	Male	Female
47	2012	2013	85.08	89.95
48	2013	2014	85.59	92.00
49	2014	2015	84.92	93.63
50	2015	2016	84.09	93.78
51	2016	2017	85.00	95.00
52	2017	2018	85.00	95.00
53	2018	2019	86.00	96.00
54	2019	2020	87.00	97.00
55	2020	2021	87.00	97.00
56	2021	2022	88.00	98.00
57	2022	2023	88.00	98.00
58	2023	2024	89.00	99.00
59	2024	2025	89.00	99.00
60	2025	2026	89.00	100.00
61	2026	2027	90.00	100.00
62	2027	2028	90.00	101.00

61 rows × 4 columns

In [44]:
```
ax = doctoralbygender.plot(kind='line',x='Upper Year',y='Male',color='green',label='# of Males with College Degrees')
doctoralbygender.plot(kind='line',x='Upper Year',y='Female',color='magenta',label='# of Females with College Degrees',ax=ax)

plt.xlabel('Year')
plt.ylabel('# of Male & Females with College Degrees')

ax.legend(bbox_to_anchor=(.85, -.1))
plt.gcf().set_size_inches((5, 5))
```



The total # of Bachelors, Masters, & Doctoral Degree Holders by Year by Gender (in 1,000)

In [45]:

```
degreesbygender_base = [bachelorsbygender, mastersbygender, doctoralbygender]
degreesbygender = reduce(lambda left,right: pd.merge(left,right,on='Upper Yea
r'), degreesbygender_base)
degreesbygender = degreesbygender[['Lower Year_x','Male_x','Female_x','Male_y'
,'Female_y','Male','Female']]
degreesbygender.columns = [['Year','Male Bachelors Degree Holders','Female Bac
helors Degree Holders','Male Masters Degree Holders','Female Masters Degree Ho
lders','Male Doctoral Degree Holders','Female Doctoral Degree Holders']]
degreesbygender
```

Out[45]:

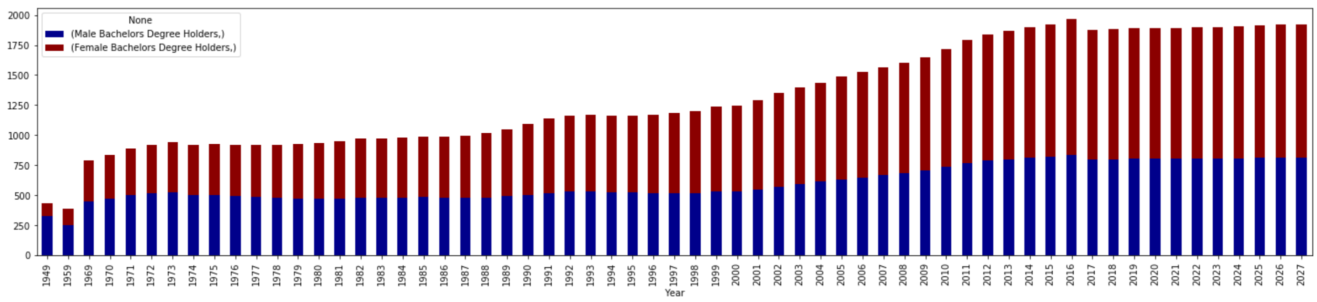
	Year	Male Bachelors Degree Holders	Female Bachelors Degree Holders	Male Masters Degree Holders	Female Masters Degree Holders	Male Doctoral Degree Holders	Female Doctoral Degree Holders
0	1949	328.84	103.22	41.22	16.96	5.80	0.62
1	1959	254.06	138.38	50.90	23.54	8.80	1.03
2	1969	451.10	341.22	130.80	82.79	53.79	5.69
3	1970	475.59	364.14	143.08	92.48	58.14	6.86
4	1971	500.59	386.68	155.01	102.19	63.35	7.85
5	1972	518.19	404.17	159.57	109.09	69.96	9.55
6	1973	527.31	418.46	162.61	119.47	71.13	11.46
7	1974	504.84	418.09	166.32	131.23	71.03	13.88
8	1975	504.93	420.82	172.52	144.96	73.89	17.12
9	1976	495.55	424.00	173.09	149.94	72.21	19.52
10	1977	487.35	433.86	166.86	151.13	70.28	22.06
11	1978	477.34	444.05	159.11	148.58	70.45	24.52
12	1979	473.61	455.81	156.88	148.31	69.53	26.11
13	1980	469.88	465.26	152.98	149.66	69.57	28.45
14	1981	473.36	479.63	151.35	151.10	68.63	29.21
15	1982	479.14	490.37	150.09	146.32	67.76	31.58
16	1983	482.32	491.99	149.27	141.87	67.77	33.03
17	1984	482.53	496.95	149.28	144.20	66.27	34.52
18	1985	485.92	501.90	149.37	146.48	65.22	35.07
19	1986	480.78	510.48	147.06	149.47	62.79	35.69
20	1987	477.20	517.63	150.24	155.54	63.02	36.12
21	1988	483.35	535.41	153.99	162.63	63.06	37.52
22	1989	491.70	559.65	158.05	172.10	63.96	39.55
23	1990	504.05	590.49	160.84	182.02	64.24	41.31
24	1991	520.81	615.74	165.87	192.22	66.60	42.95
25	1992	532.88	632.30	173.35	201.68	67.13	44.94
26	1993	532.42	636.85	180.57	212.47	66.77	45.86
27	1994	526.13	634.00	183.04	220.57	67.32	46.94
28	1995	522.45	642.34	183.48	228.70	67.19	48.32
29	1996	520.52	652.36	185.27	239.99	68.39	50.36
...	...	...	...	...	...	...	...
31	1998	519.96	682.28	190.23	255.81	65.34	51.36
32	1999	530.37	707.51	196.13	267.06	64.93	53.81
33	2000	531.84	712.33	197.77	275.73	64.17	55.41
34	2001	549.82	742.08	202.60	284.71	62.73	56.93
35	2002	573.26	775.55	215.17	303.53	62.73	58.85
36	2003	595.43	804.12	233.06	331.22	63.98	62.11
37	2004	613.00	826.26	237.16	343.00	67.26	67.13
38	2005	630.60	854.64	241.66	358.08	68.91	69.14
39	2006	649.57	874.52	242.19	368.41	71.31	73.38
40	2007	667.93	895.14	250.17	380.50	73.45	75.93
41	2008	685.42	915.98	263.52	398.57	75.67	78.89
42	2009	706.66	943.26	275.32	418.00	76.61	81.98

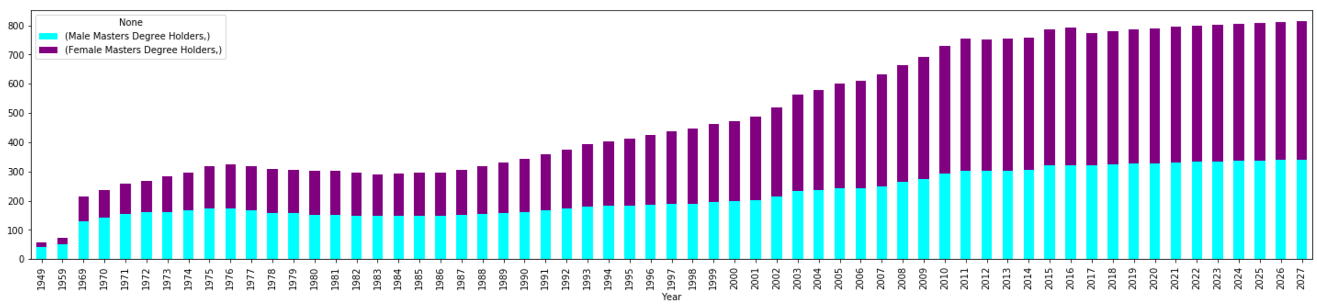
	Year	Male Bachelors Degree Holders	Female Bachelors Degree Holders	Male Masters Degree Holders	Female Masters Degree Holders	Male Doctoral Degree Holders	Female Doctoral Degree Holders
43	2010	734.16	981.89	291.68	439.24	79.67	84.16
44	2011	765.77	1026.39	302.48	453.48	82.67	87.55
45	2012	787.41	1052.97	301.55	450.17	85.08	89.95
46	2013	801.91	1068.25	302.85	451.74	85.59	92.00
47	2014	812.69	1082.28	306.62	452.19	84.92	93.63
48	2015	821.78	1098.94	320.44	465.15	84.09	93.78
49	2016	838.00	1125.00	320.00	473.00	85.00	95.00
50	2017	798.00	1077.00	322.00	453.00	85.00	95.00
51	2018	800.00	1081.00	325.00	456.00	86.00	96.00
52	2019	803.00	1086.00	327.00	459.00	87.00	97.00
53	2020	804.00	1087.00	329.00	460.00	87.00	97.00
54	2021	804.00	1089.00	331.00	463.00	88.00	98.00
55	2022	805.00	1090.00	333.00	465.00	88.00	98.00
56	2023	806.00	1092.00	335.00	467.00	89.00	99.00
57	2024	808.00	1096.00	336.00	469.00	89.00	99.00
58	2025	811.00	1101.00	337.00	470.00	89.00	100.00
59	2026	814.00	1105.00	339.00	472.00	90.00	100.00
60	2027	816.00	1106.00	340.00	473.00	90.00	101.00

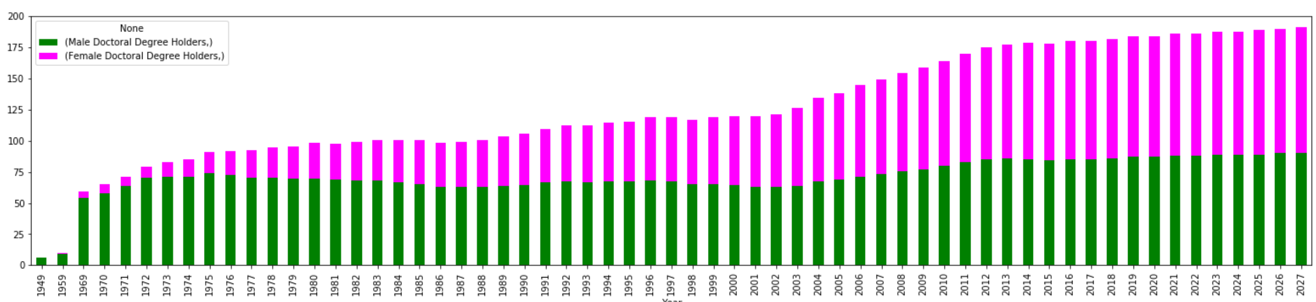
61 rows × 7 columns

In [46]:

```
degreesbygender.set_index('Year')[['Male Bachelors Degree Holders','Female Bachelors Degree Holders']].plot.bar(color = ['darkblue','darkred'], stacked=True)
plt.gcf().set_size_inches((25, 5))

degreesbygender.set_index('Year')[['Male Masters Degree Holders','Female Masters Degree Holders']].plot.bar(color = ['aqua','purple'], stacked=True)
plt.gcf().set_size_inches((25, 5))

degreesbygender.set_index('Year')[['Male Doctoral Degree Holders','Female Doctoral Degree Holders']].plot.bar(color = ['green','magenta'], stacked=True)
plt.gcf().set_size_inches((25, 5))
```







Degree Analysis : Bachelor's Degrees Earned by Major, Gender, & Year

The Bachelor's Degrees by Sex & Year table is segmented by gender (both sexes, male, & female). Segmentation is vertical so I had split the table into subsets of Both Sexes, Male, & Female below.

In [47]:
```
bachelorsdegreesbymajor.columns = ['Major','2004','2005','2006','2007','2008',
'2009','2010','2011','2012','2013','2014']
bachelorsdegreesbymajor
```

Out[47]:

	Major	2004	2005	2006	2007	2008	2009	2010	20
0	Field and sex	2004	2005	2006	2007	2008	2009	2010	
1	Both sexes	NaN	NaN	NaN	NaN	NaN	NaN	NaN	
2	All fields	1,417,421	1,456,401	1,502,922	1,541,704	1,580,413	1,619,028	1,668,227	1,
3	S&E	458,658	470,214	478,858	485,772	496,168	505,435	525,374	
4	Science	393,978	404,062	410,631	417,498	426,260	434,835	450,975	
5	Agricultural sciences	17,058	17,120	17,307	17,696	18,474	19,152	20,400	
6	Biological sciences	64,750	67,972	72,972	79,348	82,398	85,574	89,615	
7	Computer sciences	59,968	54,588	48,000	42,596	38,922	38,496	40,107	
8	Earth, atmospheric, and ocean sciences	3,903	3,959	3,987	4,077	4,314	4,542	4,802	
9	Atmospheric sciences	572	677	651	745	736	721	691	
10	Earth sciences	3,170	3,146	3,231	3,193	3,453	3,679	3,931	
11	Ocean sciences	161	136	105	139	125	142	180	
12	Mathematics and statistics	13,735	14,816	15,310	15,551	15,841	16,208	16,832	
13	Physical sciences	14,219	15,005	16,390	17,007	17,652	17,942	18,402	
14	Astronomy	292	331	366	332	346	335	388	
15	Chemistry	9,300	9,923	10,887	11,250	11,832	12,144	12,338	
16	Physics	4,140	4,199	4,566	4,870	4,876	4,842	5,000	
17	Other	487	552	571	555	598	621	676	
18	Psychology	82,606	86,031	88,551	90,498	92,989	94,743	97,746	
19	Social sciences	137,739	144,571	148,114	150,725	155,670	158,178	163,071	
20	Anthropology	7,631	7,742	8,100	8,086	8,853	9,153	9,566	
21	Area and ethnic studies	6,244	6,612	6,975	7,120	7,323	7,598	7,364	
22	Economics	24,927	25,160	24,707	24,825	26,281	27,628	29,090	
23	History of science	140	117	113	114	103	100	110	
24	Linguistics	1,047	1,053	1,201	1,340	1,505	1,478	1,682	
25	Political science and public administration	47,509	51,540	53,589	54,598	55,889	55,319	57,649	
26	Sociology	27,020	28,556	28,541	29,050	28,908	28,836	29,000	
27	Other	23,221	23,791	24,888	25,592	26,808	28,066	28,610	
28	Engineering	64,680	66,152	68,227	68,274	69,908	70,600	74,399	
29	Aerospace engineering	2,318	2,384	2,753	2,828	2,934	3,037	3,207	
...	...	...	...	...	...	...	...	...	
85	Ocean sciences	75	68	36	59	63	79	93	
86	Mathematics and statistics	7,428	8,207	8,435	8,724	8,884	9,237	9,584	

	Major	2004	2005	2006	2007	2008	2009	2010	20
87	Physical sciences	8,221	8,590	9,437	10,021	10,368	10,491	10,804	
88	Astronomy	169	188	235	200	226	189	248	
89	Chemistry	4,548	4,779	5,253	5,636	5,923	6,038	6,176	
90	Physics	3,233	3,299	3,621	3,846	3,888	3,917	3,985	
91	Other	271	324	328	339	331	347	395	
92	Psychology	18,322	19,099	19,975	20,446	21,307	21,579	22,410	
93	Social sciences	62,660	66,270	68,576	69,706	72,423	73,398	75,445	
94	Anthropology	2,350	2,362	2,494	2,456	2,733	2,723	2,863	
95	Area and ethnic studies	2,163	2,345	2,463	2,561	2,626	2,732	2,663	
96	Economics	16,650	17,021	17,043	17,195	18,237	19,282	20,195	
97	History of science	67	53	46	44	51	48	34	
98	Linguistics	330	324	447	461	502	503	558	
99	Political science and public administration	23,101	25,130	26,462	27,053	27,734	27,184	27,850	
100	Sociology	7,718	8,430	8,548	8,698	8,787	8,707	8,858	
101	Other	10,281	10,605	11,073	11,238	11,753	12,219	12,424	
102	Engineering	51,417	52,949	54,889	55,621	56,987	57,850	60,706	
103	Aerospace engineering	1,905	1,980	2,252	2,376	2,493	2,600	2,744	
104	Chemical engineering	3,352	3,113	3,179	3,338	3,670	3,977	4,543	
105	Civil engineering	7,121	7,451	8,241	8,819	9,491	10,032	10,589	
106	Electrical engineering	18,331	18,614	17,346	16,438	15,450	14,578	14,550	
107	Industrial engineering	2,543	2,666	2,716	2,473	2,557	2,823	3,048	
108	Materials engineering	595	624	737	753	878	807	844	
109	Mechanical engineering	12,395	12,933	14,050	14,894	15,592	15,671	16,695	
110	Other	5,175	5,568	6,368	6,530	6,856	7,362	7,693	
111	Non-S&E	373,727	384,834	398,530	413,696	427,069	440,686	452,245	
112	S&E = science and engineering.	NaN	NaN	NaN	NaN	NaN	NaN	NaN	
113	NOTE: Data are based on degree-granting instit...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	
114	SOURCE: National Science Foundation, National ...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	

115 rows × 12 columns

The # of Bachelor's Degrees Earned by Both Sexes by Field by Year

In [48]:
```
bachelorsdegreesbymajor_bothsexes = bachelorsdegreesbymajor[2:38].apply(lambda
x: x.str.replace(',','')).astype(int, errors = 'ignore')
bachelorsdegreesbymajor_bothsexes.columns = ['Major','2004','2005','2006','200
7','2008','2009','2010','2011','2012','2013','2014']

bachelorsdegreesbymajor_bothsexes['2004'] = bachelorsdegreesbymajor_bothsexes[
'2004'].astype(int)
bachelorsdegreesbymajor_bothsexes['2005'] = bachelorsdegreesbymajor_bothsexes[
'2005'].astype(int)
bachelorsdegreesbymajor_bothsexes['2006'] = bachelorsdegreesbymajor_bothsexes[
'2006'].astype(int)
bachelorsdegreesbymajor_bothsexes['2007'] = bachelorsdegreesbymajor_bothsexes[
'2007'].astype(int)
bachelorsdegreesbymajor_bothsexes['2008'] = bachelorsdegreesbymajor_bothsexes[
'2008'].astype(int)
bachelorsdegreesbymajor_bothsexes['2009'] = bachelorsdegreesbymajor_bothsexes[
'2009'].astype(int)
bachelorsdegreesbymajor_bothsexes['2010'] = bachelorsdegreesbymajor_bothsexes[
'2010'].astype(int)
bachelorsdegreesbymajor_bothsexes['2011'] = bachelorsdegreesbymajor_bothsexes[
'2011'].astype(int)
bachelorsdegreesbymajor_bothsexes['2012'] = bachelorsdegreesbymajor_bothsexes[
'2012'].astype(int)
bachelorsdegreesbymajor_bothsexes['2013'] = bachelorsdegreesbymajor_bothsexes[
'2013'].astype(int)
bachelorsdegreesbymajor_bothsexes['2014'] = bachelorsdegreesbymajor_bothsexes[
'2014'].astype(int)

bachelorsdegreesbymajor_bothsexes['Mean'] = bachelorsdegreesbymajor_bothsexes
[['2004','2005','2006','2007','2008','2009','2010','2011','2012','2013','2014'
]].mean(axis=1)
bachelorsdegreesbymajor_bothsexes
```

Out[48]:

	Major	2004	2005	2006	2007	2008	2009	2010	2011	2
2	All fields	1417421	1456401	1502922	1541704	1580413	1619028	1668227	1734229	1
3	S&E	458658	470214	478858	485772	496168	505435	525374	554365	
4	Science	393978	404062	410631	417498	426260	434835	450975	476266	
5	Agricultural sciences	17058	17120	17307	17696	18474	19152	20400	22759	
6	Biological sciences	64750	67972	72972	79348	82398	85574	89615	93654	
7	Computer sciences	59968	54588	48000	42596	38922	38496	40107	43586	
8	Earth atmospheric and ocean sciences	3903	3959	3987	4077	4314	4542	4802	5299	
9	Atmospheric sciences	572	677	651	745	736	721	691	678	
10	Earth sciences	3170	3146	3231	3193	3453	3679	3931	4410	
11	Ocean sciences	161	136	105	139	125	142	180	211	
12	Mathematics and statistics	13735	14816	15310	15551	15841	16208	16832	18021	
13	Physical sciences	14219	15005	16390	17007	17652	17942	18402	19198	
14	Astronomy	292	331	366	332	346	335	388	364	
15	Chemistry	9300	9923	10887	11250	11832	12144	12338	12888	
16	Physics	4140	4199	4566	4870	4876	4842	5000	5221	
17	Other	487	552	571	555	598	621	676	725	
18	Psychology	82606	86031	88551	90498	92989	94743	97746	101568	
19	Social sciences	137739	144571	148114	150725	155670	158178	163071	172181	
20	Anthropology	7631	7742	8100	8086	8853	9153	9566	10022	
21	Area and ethnic studies	6244	6612	6975	7120	7323	7598	7364	7817	
22	Economics	24927	25160	24707	24825	26281	27628	29090	30013	
23	History of science	140	117	113	114	103	100	110	138	
24	Linguistics	1047	1053	1201	1340	1505	1478	1682	1808	
25	Political science and public administration	47509	51540	53589	54598	55889	55319	57649	59752	
26	Sociology	27020	28556	28541	29050	28908	28836	29000	29614	
27	Other	23221	23791	24888	25592	26808	28066	28610	33017	
28	Engineering	64680	66152	68227	68274	69908	70600	74399	78099	
29	Aerospace engineering	2318	2384	2753	2828	2934	3037	3207	3342	
30	Chemical engineering	5185	4889	4857	5081	5504	5894	6692	7535	
31	Civil engineering	9400	9800	10663	11318	12243	12787	13482	14840	
32	Electrical engineering	21374	21459	19933	18547	17357	16184	16290	16485	
33	Industrial engineering	3808	3936	3890	3529	3566	3879	4184	4293	

	Major	2004	2005	2006	2007	2008	2009	2010	2011	2
34	Materials engineering	865	906	1006	1025	1150	1070	1138	1140	
35	Mechanical engineering	14342	14880	16163	16911	17685	17643	18858	19470	
36	Other	7388	7898	8962	9035	9469	10106	10548	10994	
37	Non-S&E	958763	986187	1024064	1055932	1084245	1113593	1142853	1179864	1

The Mean # of Bachelor's Degrees Earned by Both Sexes by Field Across all Years

In [49]:
```
bachelorsdegreesbymajor_bothsexes_filtered = bachelorsdegreesbymajor_bothsexes
[['Major','Mean']]
bachelorsdegreesbymajor_bothsexes_filtered = bachelorsdegreesbymajor_bothsexes
_filtered.sort_values(['Mean'], ascending=False).reset_index()
bachelorsdegreesbymajor_bothsexes_filtered = bachelorsdegreesbymajor_bothsexes
_filtered[['Major','Mean']]
bachelorsdegreesbymajor_bothsexes_filtered
```

Out[49]:

	Major	Mean
0	All fields	1.643906e+06
1	Non-S&E	1.115218e+06
2	S&E	5.286876e+05
3	Science	4.536545e+05
4	Social sciences	1.604335e+05
5	Psychology	9.802055e+04
6	Biological sciences	8.640055e+04
7	Engineering	7.503309e+04
8	Political science and public administration	5.600991e+04
9	Computer sciences	4.744900e+04
10	Sociology	2.925709e+04
11	Other	2.922391e+04
12	Economics	2.760973e+04
13	Agricultural sciences	2.103918e+04
14	Mechanical engineering	1.852909e+04
15	Electrical engineering	1.849200e+04
16	Physical sciences	1.815764e+04
17	Mathematics and statistics	1.726600e+04
18	Civil engineering	1.283100e+04
19	Chemistry	1.208927e+04
20	Other	1.022955e+04
21	Anthropology	9.374455e+03
22	Area and ethnic studies	7.236545e+03
23	Chemical engineering	6.600364e+03
24	Physics	5.060091e+03
25	Earth atmospheric and ocean sciences	4.888091e+03
26	Industrial engineering	4.170545e+03
27	Earth sciences	4.013000e+03
28	Aerospace engineering	3.036636e+03
29	Linguistics	1.602909e+03
30	Materials engineering	1.143909e+03
31	Atmospheric sciences	7.005455e+02
32	Other	6.508182e+02
33	Astronomy	3.574545e+02
34	Ocean sciences	1.745455e+02
35	History of science	1.190000e+02

The # of Bachelor's Degrees Earned by Males by Field by Year

In [50]:

```
bachelorsdegreesbymajor_male = bachelorsdegreesbymajor.iloc[76:112]
bachelorsdegreesbymajor_male = bachelorsdegreesbymajor[76:112].apply(lambda x: x.str.replace(',','')).astype(int, errors = 'ignore')
bachelorsdegreesbymajor_male.columns = ['Major','2004','2005','2006','2007','2008','2009','2010','2011','2012','2013','2014']

bachelorsdegreesbymajor_male['2004'] = bachelorsdegreesbymajor_male['2004'].astype(int)
bachelorsdegreesbymajor_male['2005'] = bachelorsdegreesbymajor_male['2005'].astype(int)
bachelorsdegreesbymajor_male['2006'] = bachelorsdegreesbymajor_male['2006'].astype(int)
bachelorsdegreesbymajor_male['2007'] = bachelorsdegreesbymajor_male['2007'].astype(int)
bachelorsdegreesbymajor_male['2008'] = bachelorsdegreesbymajor_male['2008'].astype(int)
bachelorsdegreesbymajor_male['2009'] = bachelorsdegreesbymajor_male['2009'].astype(int)
bachelorsdegreesbymajor_male['2010'] = bachelorsdegreesbymajor_male['2010'].astype(int)
bachelorsdegreesbymajor_male['2011'] = bachelorsdegreesbymajor_male['2011'].astype(int)
bachelorsdegreesbymajor_male['2012'] = bachelorsdegreesbymajor_male['2012'].astype(int)
bachelorsdegreesbymajor_male['2013'] = bachelorsdegreesbymajor_male['2013'].astype(int)
bachelorsdegreesbymajor_male['2014'] = bachelorsdegreesbymajor_male['2014'].astype(int)

bachelorsdegreesbymajor_male['Mean'] = bachelorsdegreesbymajor_male[['2004','2005','2006','2007','2008','2009','2010','2011','2012','2013','2014']].mean(axis=1)
bachelorsdegreesbymajor_male
```

Out[50]:

	Major	2004	2005	2006	2007	2008	2009	2010	2011	2012
76	All fields	601588	618758	636559	655393	673788	691428	713336	740357	772175
77	S&E	227861	233924	238029	241697	246719	250742	261091	275258	291791
78	Science	176444	180975	183140	186076	189732	192892	200385	211817	224509
79	Agricultural sciences	8336	8383	8398	8781	9017	9334	9691	10904	11616
80	Biological sciences	24319	25698	27917	31347	33136	34477	36737	37860	40683
81	Computer sciences	44902	42429	38061	34652	32038	31602	32801	35886	39230
82	Earth atmospheric and ocean sciences	2256	2299	2341	2399	2559	2774	2913	3219	3570
83	Atmospheric sciences	369	436	432	483	473	486	455	461	472
84	Earth sciences	1812	1795	1873	1857	2023	2209	2365	2650	2978
85	Ocean sciences	75	68	36	59	63	79	93	108	120
86	Mathematics and statistics	7428	8207	8435	8724	8884	9237	9584	10276	11283
87	Physical sciences	8221	8590	9437	10021	10368	10491	10804	11404	12137
88	Astronomy	169	188	235	200	226	189	248	228	239
89	Chemistry	4548	4779	5253	5636	5923	6038	6176	6560	6984
90	Physics	3233	3299	3621	3846	3888	3917	3985	4220	4495
91	Other	271	324	328	339	331	347	395	396	419
92	Psychology	18322	19099	19975	20446	21307	21579	22410	23401	25590
93	Social sciences	62660	66270	68576	69706	72423	73398	75445	78867	80400
94	Anthropology	2350	2362	2494	2456	2733	2723	2863	2998	3201
95	Area and ethnic studies	2163	2345	2463	2561	2626	2732	2663	2763	2715
96	Economics	16650	17021	17043	17195	18237	19282	20195	20826	20677
97	History of science	67	53	46	44	51	48	34	60	49
98	Linguistics	330	324	447	461	502	503	558	573	689
99	Political science and public administration	23101	25130	26462	27053	27734	27184	27850	28975	29267
100	Sociology	7718	8430	8548	8698	8787	8707	8858	8988	9433
101	Other	10281	10605	11073	11238	11753	12219	12424	13684	14369
102	Engineering	51417	52949	54889	55621	56987	57850	60706	63441	67282
103	Aerospace engineering	1905	1980	2252	2376	2493	2600	2744	2903	3082
104	Chemical engineering	3352	3113	3179	3338	3670	3977	4543	5228	5809
105	Civil engineering	7121	7451	8241	8819	9491	10032	10589	11562	11747
106	Electrical engineering	18331	18614	17346	16438	15450	14578	14550	14744	15457
107	Industrial engineering	2543	2666	2716	2473	2557	2823	3048	3129	3413

	Major	2004	2005	2006	2007	2008	2009	2010	2011	2012
108	Materials engineering	595	624	737	753	878	807	844	819	954
109	Mechanical engineering	12395	12933	14050	14894	15592	15671	16695	17262	18351
110	Other	5175	5568	6368	6530	6856	7362	7693	7794	8469
111	Non-S&E	373727	384834	398530	413696	427069	440686	452245	465099	480384

The Mean # of Bachelor's Degrees Earned by Males by Field Across All Years

In [51]:
```
bachelorsdegreesbymajor_male_filtered = bachelorsdegreesbymajor_male[['Major',
'Mean']]
bachelorsdegreesbymajor_male_filtered = bachelorsdegreesbymajor_male_filtered.
sort_values('Mean', ascending = False).reset_index()
bachelorsdegreesbymajor_male_filtered
```

Out[51]:

	index	Major	Mean
0	76	All fields	700698.818182
1	111	Non-S&E	437889.363636
2	77	S&E	262809.454545
3	78	Science	202141.818182
4	93	Social sciences	73584.636364
5	102	Engineering	60667.636364
6	81	Computer sciences	38178.545455
7	80	Biological sciences	34519.727273
8	99	Political science and public administration	27245.909091
9	92	Psychology	22429.727273
10	96	Economics	19072.545455
11	106	Electrical engineering	16319.363636
12	109	Mechanical engineering	16275.272727
13	101	Other	12501.181818
14	87	Physical sciences	10709.909091
15	79	Agricultural sciences	10046.727273
16	105	Civil engineering	9920.818182
17	86	Mathematics and statistics	9731.181818
18	100	Sociology	8842.909091
19	110	Other	7289.909091
20	89	Chemistry	6071.636364
21	104	Chemical engineering	4472.545455
22	90	Physics	4049.636364
23	107	Industrial engineering	2975.818182
24	82	Earth atmospheric and ocean sciences	2941.363636
25	94	Anthropology	2796.363636
26	103	Aerospace engineering	2583.909091
27	95	Area and ethnic studies	2550.000000
28	84	Earth sciences	2394.545455
29	108	Materials engineering	830.000000
30	98	Linguistics	526.090909
31	83	Atmospheric sciences	460.000000
32	91	Other	369.545455
33	88	Astronomy	219.090909
34	85	Ocean sciences	86.818182
35	97	History of science	49.636364

The # of Bachelor's Degrees Earned by Females by Field by Year

In [52]:

```
bachelorsdegreesbymajor_female = bachelorsdegreesbymajor.iloc[39:75]
bachelorsdegreesbymajor_female = bachelorsdegreesbymajor[39:75].apply(lambda x
: x.str.replace(',','')).astype(int, errors = 'ignore')
bachelorsdegreesbymajor_female.columns = ['Major','2004','2005','2006','2007',
'2008','2009','2010','2011','2012','2013','2014']

bachelorsdegreesbymajor_female['2004'] = bachelorsdegreesbymajor_female['2004'
].astype(int)
bachelorsdegreesbymajor_female['2005'] = bachelorsdegreesbymajor_female['2005'
].astype(int)
bachelorsdegreesbymajor_female['2006'] = bachelorsdegreesbymajor_female['2006'
].astype(int)
bachelorsdegreesbymajor_female['2007'] = bachelorsdegreesbymajor_female['2007'
].astype(int)
bachelorsdegreesbymajor_female['2008'] = bachelorsdegreesbymajor_female['2008'
].astype(int)
bachelorsdegreesbymajor_female['2009'] = bachelorsdegreesbymajor_female['2009'
].astype(int)
bachelorsdegreesbymajor_female['2010'] = bachelorsdegreesbymajor_female['2010'
].astype(int)
bachelorsdegreesbymajor_female['2011'] = bachelorsdegreesbymajor_female['2011'
].astype(int)
bachelorsdegreesbymajor_female['2012'] = bachelorsdegreesbymajor_female['2012'
].astype(int)
bachelorsdegreesbymajor_female['2013'] = bachelorsdegreesbymajor_female['2013'
].astype(int)
bachelorsdegreesbymajor_female['2014'] = bachelorsdegreesbymajor_female['2014'
].astype(int)

bachelorsdegreesbymajor_female['Mean'] = bachelorsdegreesbymajor_female[['200
4','2005','2006','2007','2008','2009','2010','2011','2012','2013','2014']].mea
n(axis=1)
bachelorsdegreesbymajor_female
```

Out[52]:

	Major	2004	2005	2006	2007	2008	2009	2010	2011	2012
39	All fields	815833	837643	866363	886311	906625	927600	954891	993872	1038472
40	S&E	230797	236290	240829	244075	249449	254693	264283	279107	297539
41	Science	217534	223087	227491	231422	236528	241943	250590	264449	281558
42	Agricultural sciences	8722	8737	8909	8915	9457	9818	10709	11855	13444
43	Biological sciences	40431	42274	45055	48001	49262	51097	52878	55794	59217
44	Computer sciences	15066	12159	9939	7944	6884	6894	7306	7700	8730
45	Earth atmospheric and ocean sciences	1647	1660	1646	1678	1755	1768	1889	2080	2295
46	Atmospheric sciences	203	241	219	262	263	235	236	217	263
47	Earth sciences	1358	1351	1358	1336	1430	1470	1566	1760	1926
48	Ocean sciences	86	68	69	80	62	63	87	103	106
49	Mathematics and statistics	6307	6609	6875	6827	6957	6971	7248	7745	8536
50	Physical sciences	5998	6415	6953	6986	7284	7451	7598	7794	8284
51	Astronomy	123	143	131	132	120	146	140	136	153
52	Chemistry	4752	5144	5634	5614	5909	6106	6162	6328	6730
53	Physics	907	900	945	1024	988	925	1015	1001	1062
54	Other	216	228	243	216	267	274	281	329	339
55	Psychology	64284	66932	68576	70052	71682	73164	75336	78167	84126
56	Social sciences	75079	78301	79538	81019	83247	84780	87626	93314	96926
57	Anthropology	5281	5380	5606	5630	6120	6430	6703	7024	7927
58	Area and ethnic studies	4081	4267	4512	4559	4697	4866	4701	5054	5145
59	Economics	8277	8139	7664	7630	8044	8346	8895	9187	8780
60	History of science	73	64	67	70	52	52	76	78	73
61	Linguistics	717	729	754	879	1003	975	1124	1235	1420
62	Political science and public administration	24408	26410	27127	27545	28155	28135	29799	30777	31493
63	Sociology	19302	20126	19993	20352	20121	20129	20142	20626	21169
64	Other	12940	13186	13815	14354	15055	15847	16186	19333	20919
65	Engineering	13263	13203	13338	12653	12921	12750	13693	14658	15981
66	Aerospace engineering	413	404	501	452	441	437	463	439	463
67	Chemical engineering	1833	1776	1678	1743	1834	1917	2149	2307	2535
68	Civil engineering	2279	2349	2422	2499	2752	2755	2893	3278	3428
69	Electrical engineering	3043	2845	2587	2109	1907	1606	1740	1741	1923
70	Industrial engineering	1265	1270	1174	1056	1009	1056	1136	1164	1276

	Major	2004	2005	2006	2007	2008	2009	2010	2011	2012
71	Materials engineering	270	282	269	272	272	263	294	321	368
72	Mechanical engineering	1947	1947	2113	2017	2093	1972	2163	2208	2538
73	Other	2213	2330	2594	2505	2613	2744	2855	3200	3450
74	Non-S&E	585036	601353	625534	642236	657176	672907	690608	714765	740933

The Mean # of Bachelor's Degrees Earned by Females by Field Across All Years

In [53]:
```
bachelorsdegreesbymajor_female_filtered = bachelorsdegreesbymajor_female[['Major','Mean']]
bachelorsdegreesbymajor_female_filtered = bachelorsdegreesbymajor_female_filtered.sort_values('Mean', ascending = False).reset_index()
bachelorsdegreesbymajor_female_filtered
```

Out[53]:

	index	Major	Mean
0	39	All fields	943207.272727
1	74	Non-S&E	677329.090909
2	40	S&E	265878.181818
3	41	Science	251512.727273
4	56	Social sciences	86848.909091
5	55	Psychology	75590.818182
6	43	Biological sciences	51880.818182
7	62	Political science and public administration	28764.000000
8	63	Sociology	20414.181818
9	64	Other	16722.727273
10	65	Engineering	14365.454545
11	42	Agricultural sciences	10992.454545
12	44	Computer sciences	9270.454545
13	59	Economics	8537.181818
14	49	Mathematics and statistics	7534.818182
15	50	Physical sciences	7447.727273
16	57	Anthropology	6578.090909
17	52	Chemistry	6017.636364
18	58	Area and ethnic studies	4686.545455
19	73	Other	2939.636364
20	68	Civil engineering	2910.181818
21	72	Mechanical engineering	2253.818182
22	69	Electrical engineering	2172.636364
23	67	Chemical engineering	2127.818182
24	45	Earth atmospheric and ocean sciences	1946.727273
25	47	Earth sciences	1618.454545
26	70	Industrial engineering	1194.727273
27	61	Linguistics	1076.818182
28	53	Physics	1010.454545
29	66	Aerospace engineering	452.727273
30	71	Materials engineering	313.909091
31	54	Other	281.272727
32	46	Atmospheric sciences	240.545455
33	51	Astronomy	138.363636
34	48	Ocean sciences	87.727273
35	60	History of science	69.363636

The # of Bachelor's Degrees by Sex & Field Version 2

In [54]:
```
# bachelorsdegreesbymajor_v2 = bachelorsdegreesbymajor_v2.iloc[3:]
bachelorsdegreesbymajor_v2.columns =['Year','All Fields','Science & Eng','All Sciences','Agricultural Sciences','Biological Sciences','Computer Sciences','Earth, Atmospheric, & Ocean Sciences','Mathematics & Statistics','Physical Sciences','Psychology','Social Sciences','Engineering','Non Science & Eng']
bachelorsdegreesbymajor_v2
```

Out[54]:

	Year	All Fields	Science & Eng	All Sciences	Agricultural Sciences	Biological Sciences	Computer Sciences	Earth, Atmospheric & Ocean Sciences
0	Sex and year	All fields	S&E	Science	NaN	NaN	NaN	Na
1	NaN	NaN	NaN	All \nsciences	Agricultural sciences	Biological sciences	Computer sciences	Earth atmospheric and ocea scienc
2	NaN	Number	NaN	NaN	NaN	NaN	NaN	Na
3	Both sexes	NaN	NaN	NaN	NaN	NaN	NaN	Na
4	2004	1,417,421	458,658	393,978	17,058	64,750	59,968	3,90
5	2005	1,456,401	470,214	404,062	17,120	67,972	54,588	3,95
6	2006	1,502,922	478,858	410,631	17,307	72,972	48,000	3,98
7	2007	1,541,704	485,772	417,498	17,696	79,348	42,596	4,07
8	2008	1,580,413	496,168	426,260	18,474	82,398	38,922	4,31
9	2009	1,619,028	505,435	434,835	19,152	85,574	38,496	4,54
10	2010	1,668,227	525,374	450,975	20,400	89,615	40,107	4,80
11	2011	1,734,229	554,365	476,266	22,759	93,654	43,586	5,29
12	2012	1,810,647	589,330	506,067	25,060	99,900	47,960	5,86
13	2013	1,861,034	615,475	527,663	27,609	104,703	51,586	6,29
14	2014	1,890,941	635,915	541,965	28,796	109,520	56,130	6,73
15	Female	NaN	NaN	NaN	NaN	NaN	NaN	Na
16	2004	815,833	230,797	217,534	8,722	40,431	15,066	1,64
17	2005	837,643	236,290	223,087	8,737	42,274	12,159	1,66
18	2006	866,363	240,829	227,491	8,909	45,055	9,939	1,64
19	2007	886,311	244,075	231,422	8,915	48,001	7,944	1,67
20	2008	906,625	249,449	236,528	9,457	49,262	6,884	1,75
21	2009	927,600	254,693	241,943	9,818	51,097	6,894	1,76
22	2010	954,891	264,283	250,590	10,709	52,878	7,306	1,88
23	2011	993,872	279,107	264,449	11,855	55,794	7,700	2,08
24	2012	1,038,472	297,539	281,558	13,444	59,217	8,730	2,29
25	2013	1,066,182	309,698	292,764	14,826	61,971	9,209	2,40
26	2014	1,081,488	317,900	299,274	15,525	64,709	10,144	2,59
27	Male	NaN	NaN	NaN	NaN	NaN	NaN	Na
28	2004	601,588	227,861	176,444	8,336	24,319	44,902	2,25
29	2005	618,758	233,924	180,975	8,383	25,698	42,429	2,29
30	2006	636,559	238,029	183,140	8,398	27,917	38,061	2,34
31	2007	655,393	241,697	186,076	8,781	31,347	34,652	2,39
32	2008	673,788	246,719	189,732	9,017	33,136	32,038	2,55
33	2009	691,428	250,742	192,892	9,334	34,477	31,602	2,77
34	2010	713,336	261,091	200,385	9,691	36,737	32,801	2,91
35	2011	740,357	275,258	211,817	10,904	37,860	35,886	3,21
36	2012	772,175	291,791	224,509	11,616	40,683	39,230	3,57
37	2013	794,852	305,777	234,899	12,783	42,732	42,377	3,89
38	2014	809,453	318,015	242,691	13,271	44,811	45,986	4,13
39	NaN	Percent	NaN	NaN	NaN	NaN	NaN	Na
40	Female	NaN	NaN	NaN	NaN	NaN	NaN	Na

	Year	All Fields	Science & Eng	All Sciences	Agricultural Sciences	Biological Sciences	Computer Sciences	Earth, Atmospheric & Ocean Sciences
41	2004	57.6	50.3	55.2	51.1	62.4	25.1	42.
42	2005	57.5	50.3	55.2	51.0	62.2	22.3	41.
43	2006	57.6	50.3	55.4	51.5	61.7	20.7	41.
44	2007	57.5	50.2	55.4	50.4	60.5	18.6	41.
45	2008	57.4	50.3	55.5	51.2	59.8	17.7	40.
46	2009	57.3	50.4	55.6	51.3	59.7	17.9	38.
47	2010	57.2	50.3	55.6	52.5	59.0	18.2	39.
48	2011	57.3	50.4	55.5	52.1	59.6	17.7	39.
49	2012	57.4	50.5	55.6	53.7	59.3	18.2	39.
50	2013	57.3	50.3	55.5	53.7	59.2	17.9	38.
51	2014	57.2	50.0	55.2	53.9	59.1	18.1	38.
52	S&E = science and engineering.	NaN	NaN	NaN	NaN	NaN	NaN	Na
53	NOTE: Data are based on degree-granting instit...	NaN	NaN	NaN	NaN	NaN	NaN	Na
54	SOURCE: National Science Foundation, National ...	NaN	NaN	NaN	NaN	NaN	NaN	Na

The # of Bachelor's Degrees Earned by Both Sexes by Field Version 2

In [55]:

```
bachelorsdegreesbymajor_v2_bothsexes = bachelorsdegreesbymajor_v2[4:15][['Year','All Fields','Science & Eng','All Sciences','Agricultural Sciences','Biological Sciences','Computer Sciences','Earth, Atmospheric, & Ocean Sciences','Mathematics & Statistics','Physical Sciences','Psychology','Social Sciences','Engineering','Non Science & Eng']].apply(lambda x: x.str.replace(',','')).astype(int)
bachelorsdegreesbymajor_v2_bothsexes
```

Out[55]:

	Year	All Fields	Science & Eng	All Sciences	Agricultural Sciences	Biological Sciences	Computer Sciences	Earth, Atmospheric, & Ocean Sciences	Mathem & Statis
4	2004	1417421	458658	393978	17058	64750	59968	3903	1
5	2005	1456401	470214	404062	17120	67972	54588	3959	1
6	2006	1502922	478858	410631	17307	72972	48000	3987	1
7	2007	1541704	485772	417498	17696	79348	42596	4077	1
8	2008	1580413	496168	426260	18474	82398	38922	4314	1
9	2009	1619028	505435	434835	19152	85574	38496	4542	1
10	2010	1668227	525374	450975	20400	89615	40107	4802	1
11	2011	1734229	554365	476266	22759	93654	43586	5299	1
12	2012	1810647	589330	506067	25060	99900	47960	5865	1
13	2013	1861034	615475	527663	27609	104703	51586	6291	2
14	2014	1890941	635915	541965	28796	109520	56130	6730	2

The # of Bachelor's Degrees Earned by Males by Field Version

In [56]:
```
bachelorsdegreesbymajor_v2_male = bachelorsdegreesbymajor_v2.iloc[28:39].apply
(lambda x: x.str.replace(',','')).astype(int)
bachelorsdegreesbymajor_v2_male
```

Out[56]:

	Year	All Fields	Science & Eng	All Sciences	Agricultural Sciences	Biological Sciences	Computer Sciences	Earth, Atmospheric, & Ocean Sciences	Mathema & Statisti
28	2004	601588	227861	176444	8336	24319	44902	2256	7
29	2005	618758	233924	180975	8383	25698	42429	2299	8
30	2006	636559	238029	183140	8398	27917	38061	2341	8
31	2007	655393	241697	186076	8781	31347	34652	2399	8
32	2008	673788	246719	189732	9017	33136	32038	2559	8
33	2009	691428	250742	192892	9334	34477	31602	2774	9
34	2010	713336	261091	200385	9691	36737	32801	2913	9
35	2011	740357	275258	211817	10904	37860	35886	3219	10
36	2012	772175	291791	224509	11616	40683	39230	3570	11
37	2013	794852	305777	234899	12783	42732	42377	3891	12
38	2014	809453	318015	242691	13271	44811	45986	4134	12

The # of Bachelor's Degrees Earned by Female by Field Version

In [57]:
```
bachelorsdegreesbymajor_v2_female = bachelorsdegreesbymajor_v2.iloc[16:27].app
ly(lambda x: x.str.replace(',','')).astype(int)
bachelorsdegreesbymajor_v2_female
```

Out[57]:

	Year	All Fields	Science & Eng	All Sciences	Agricultural Sciences	Biological Sciences	Computer Sciences	Earth, Atmospheric, & Ocean Sciences	Mathem & Statis
16	2004	815833	230797	217534	8722	40431	15066	1647	
17	2005	837643	236290	223087	8737	42274	12159	1660	
18	2006	866363	240829	227491	8909	45055	9939	1646	
19	2007	886311	244075	231422	8915	48001	7944	1678	
20	2008	906625	249449	236528	9457	49262	6884	1755	
21	2009	927600	254693	241943	9818	51097	6894	1768	
22	2010	954891	264283	250590	10709	52878	7306	1889	
23	2011	993872	279107	264449	11855	55794	7700	2080	
24	2012	1038472	297539	281558	13444	59217	8730	2295	
25	2013	1066182	309698	292764	14826	61971	9209	2400	
26	2014	1081488	317900	299274	15525	64709	10144	2596	

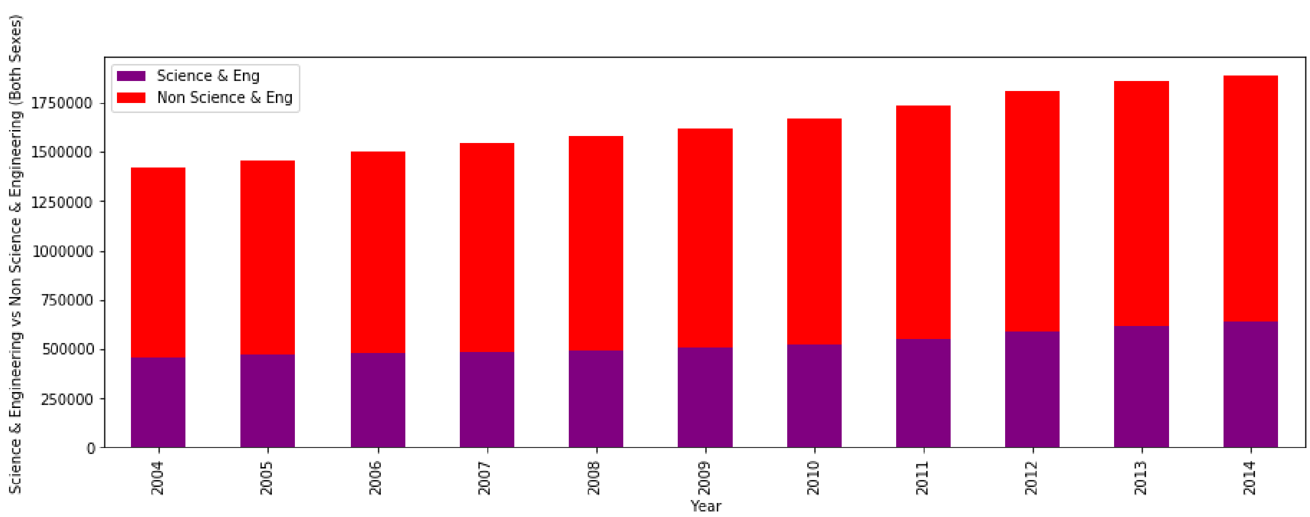
The # of Science & Engineering Versus Non Science & Engineering by Gender & Year

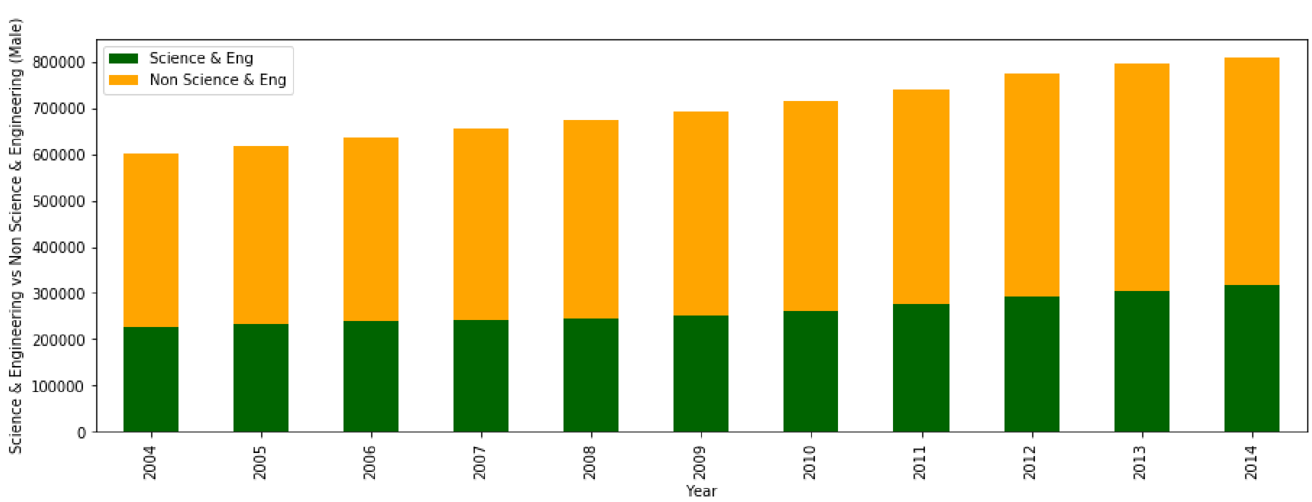
In [58]:

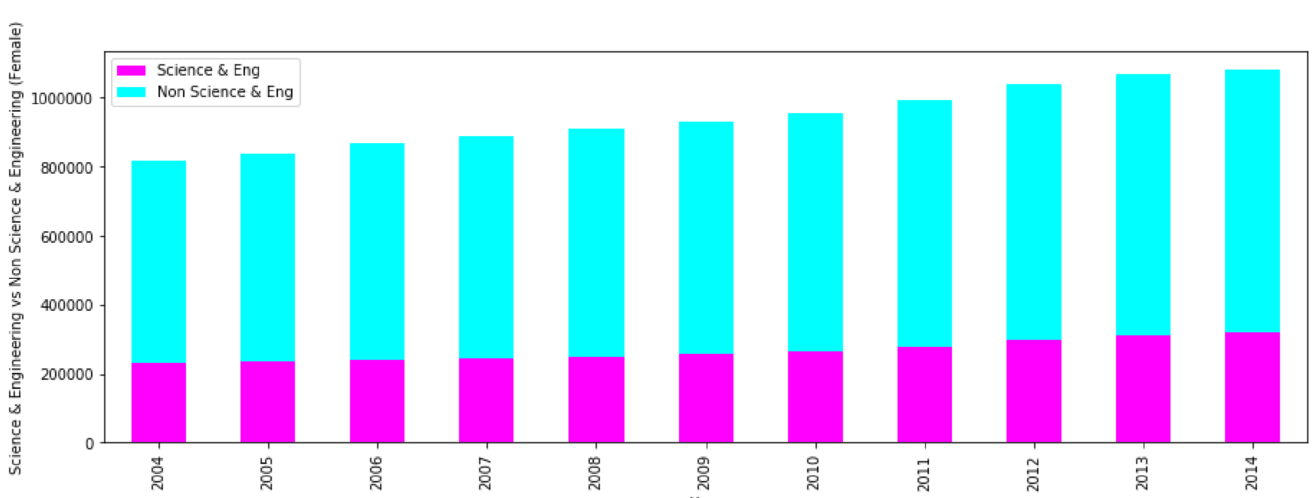
```python
bachelorsdegreesbymajor_v2_bothsexes.set_index('Year')[['Science & Eng','Non Science & Eng']].plot.bar(color = ['purple','red'], stacked=True)
plt.gcf().set_size_inches((15, 5))
plt.xlabel('Year')
plt.ylabel('Science & Engineering vs Non Science & Engineering (Both Sexes)')
plt.legend()

bachelorsdegreesbymajor_v2_male.set_index('Year')[['Science & Eng','Non Science & Eng']].plot.bar(color = ['darkgreen','orange'], stacked=True)
plt.gcf().set_size_inches((15, 5))
plt.xlabel('Year')
plt.ylabel('Science & Engineering vs Non Science & Engineering (Male)')
plt.legend()

bachelorsdegreesbymajor_v2_female.set_index('Year')[['Science & Eng','Non Science & Eng']].plot.bar(color = ['magenta','aqua'], stacked=True)
plt.gcf().set_size_inches((15, 5))
plt.xlabel('Year')
plt.ylabel('Science & Engineering vs Non Science & Engineering (Female)')
plt.legend()
```

Out[58]:

```
<matplotlib.legend.Legend at 0x1b0fcb702e8>
```







The # of Males that Earned Bachelor's Degrees in the Science & Engineering, All Sciences, & Engineering Fields by Year

In [59]:
```
bachelorsdegreesbymajor_v2_filtered = bachelorsdegreesbymajor_v2[['Year','All Fields','Science & Eng','Engineering']]
bachelorsdegreesbymajor_v2_filtered
```

Out[59]:

	Year	All Fields	Science & Eng	Engineering
0	Sex and year	All fields	S&E	Engineering
1	NaN	NaN	NaN	NaN
2	NaN	Number	NaN	NaN
3	Both sexes	NaN	NaN	NaN
4	2004	1,417,421	458,658	64,680
5	2005	1,456,401	470,214	66,152
6	2006	1,502,922	478,858	68,227
7	2007	1,541,704	485,772	68,274
8	2008	1,580,413	496,168	69,908
9	2009	1,619,028	505,435	70,600
10	2010	1,668,227	525,374	74,399
11	2011	1,734,229	554,365	78,099
12	2012	1,810,647	589,330	83,263
13	2013	1,861,034	615,475	87,812
14	2014	1,890,941	635,915	93,950
15	Female	NaN	NaN	NaN
16	2004	815,833	230,797	13,263
17	2005	837,643	236,290	13,203
18	2006	866,363	240,829	13,338
19	2007	886,311	244,075	12,653
20	2008	906,625	249,449	12,921
21	2009	927,600	254,693	12,750
22	2010	954,891	264,283	13,693
23	2011	993,872	279,107	14,658
24	2012	1,038,472	297,539	15,981
25	2013	1,066,182	309,698	16,934
26	2014	1,081,488	317,900	18,626
27	Male	NaN	NaN	NaN
28	2004	601,588	227,861	51,417
29	2005	618,758	233,924	52,949
30	2006	636,559	238,029	54,889
31	2007	655,393	241,697	55,621
32	2008	673,788	246,719	56,987
33	2009	691,428	250,742	57,850
34	2010	713,336	261,091	60,706
35	2011	740,357	275,258	63,441
36	2012	772,175	291,791	67,282
37	2013	794,852	305,777	70,878
38	2014	809,453	318,015	75,324
39	NaN	Percent	NaN	NaN
40	Female	NaN	NaN	NaN
41	2004	57.6	50.3	20.5
42	2005	57.5	50.3	20.0
43	2006	57.6	50.3	19.6
44	2007	57.5	50.2	18.5

	Year	All Fields	Science & Eng	Engineering
45	2008	57.4	50.3	18.5
46	2009	57.3	50.4	18.1
47	2010	57.2	50.3	18.4
48	2011	57.3	50.4	18.8
49	2012	57.4	50.5	19.2
50	2013	57.3	50.3	19.3
51	2014	57.2	50.0	19.8
52	S&E = science and engineering.	NaN	NaN	NaN
53	NOTE: Data are based on degree-granting instit...	NaN	NaN	NaN
54	SOURCE: National Science Foundation, National ...	NaN	NaN	NaN

In [60]:

```
bachelorsdegreesbymajor_v2_filtered_male = bachelorsdegreesbymajor_v2_filtered
.iloc[28:39].apply(lambda x: x.str.replace(',','')).astype(int)
bachelorsdegreesbymajor_v2_filtered_male
```

Out[60]:

	Year	All Fields	Science & Eng	Engineering
28	2004	601588	227861	51417
29	2005	618758	233924	52949
30	2006	636559	238029	54889
31	2007	655393	241697	55621
32	2008	673788	246719	56987
33	2009	691428	250742	57850
34	2010	713336	261091	60706
35	2011	740357	275258	63441
36	2012	772175	291791	67282
37	2013	794852	305777	70878
38	2014	809453	318015	75324

The # of Females that Earned Bachelor's Degrees in the Science & Engineering, All Sciences, & Engineering Fields by Year

In [61]:

```
bachelorsdegreesbymajor_v2_filtered_female = bachelorsdegreesbymajor_v2_filter
ed.iloc[16:27].apply(lambda x: x.str.replace(',','')).astype(int)
bachelorsdegreesbymajor_v2_filtered_female
```

Out[61]:

	Year	All Fields	Science & Eng	Engineering
16	2004	815833	230797	13263
17	2005	837643	236290	13203
18	2006	866363	240829	13338
19	2007	886311	244075	12653
20	2008	906625	249449	12921
21	2009	927600	254693	12750
22	2010	954891	264283	13693
23	2011	993872	279107	14658
24	2012	1038472	297539	15981
25	2013	1066182	309698	16934
26	2014	1081488	317900	18626

The # of Bachelor's Degrees Earned by Males & Females in the Science & Engineering, All Sciences, & Engineering Fields by Year

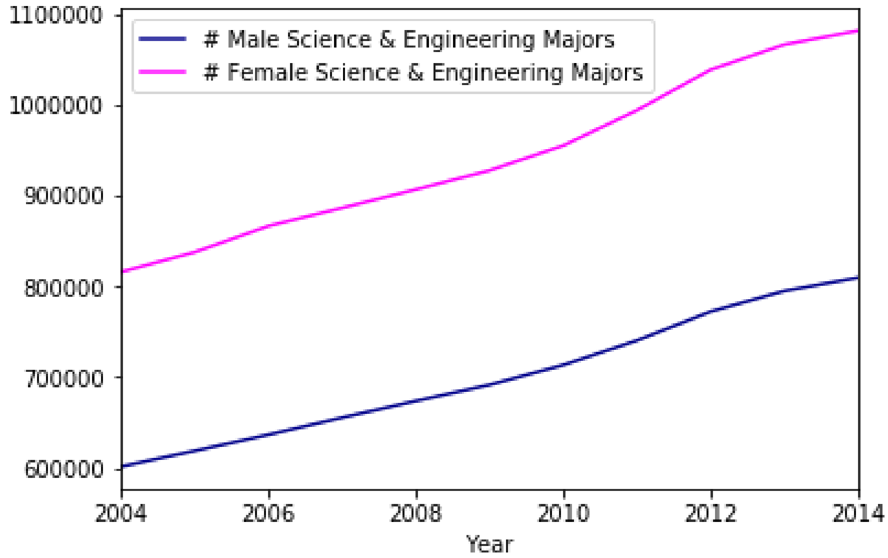
In [62]:
```python
bachelorsdegreesbymajor_malevsfemale = pd.merge(bachelorsdegreesbymajor_v2_filtered_male, bachelorsdegreesbymajor_v2_filtered_female, on='Year', how='outer')
bachelorsdegreesbymajor_malevsfemale.columns = ['Year','Males in Science & Eng','Males in All Sciences','Males in Engineering','Females in Science & Eng','Females in All Sciences','Females in Engineering']
bachelorsdegreesbymajor_malevsfemale[['Year','Males in Science & Eng','Males in All Sciences','Males in Engineering','Females in Science & Eng','Females in All Sciences','Females in Engineering']] = bachelorsdegreesbymajor_malevsfemale[['Year','Males in Science & Eng','Males in All Sciences','Males in Engineering','Females in Science & Eng','Females in All Sciences','Females in Engineering']].astype(int)
bachelorsdegreesbymajor_malevsfemale
```

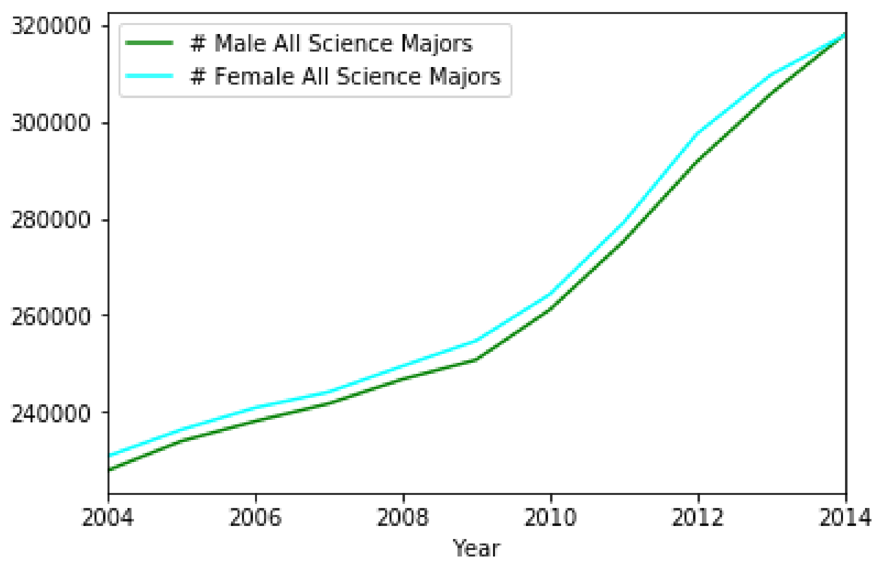
Out[62]:

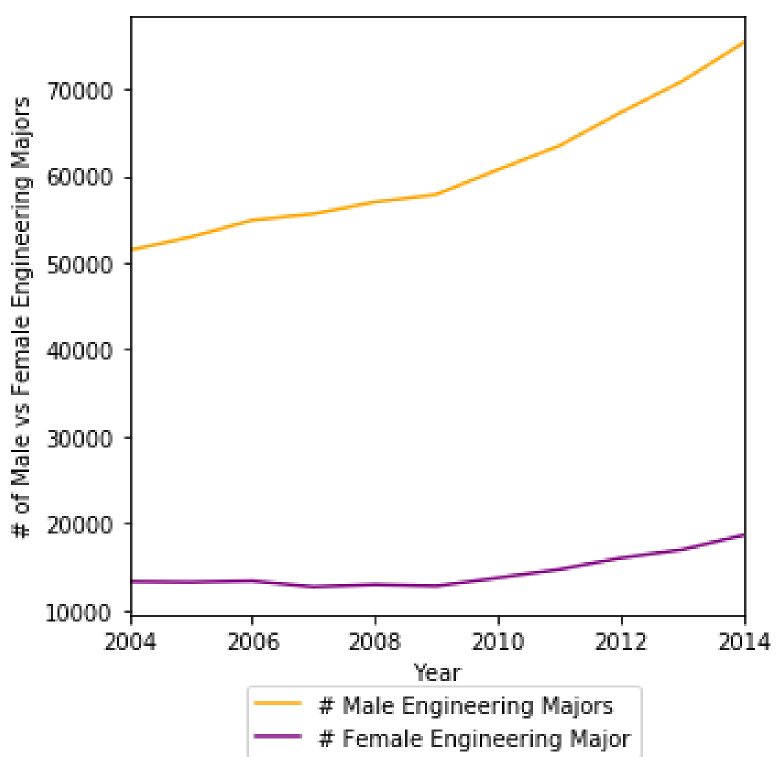
	Year	Males in Science & Eng	Males in All Sciences	Males in Engineering	Females in Science & Eng	Females in All Sciences	Females in Engineering
0	2004	601588	227861	51417	815833	230797	13263
1	2005	618758	233924	52949	837643	236290	13203
2	2006	636559	238029	54889	866363	240829	13338
3	2007	655393	241697	55621	886311	244075	12653
4	2008	673788	246719	56987	906625	249449	12921
5	2009	691428	250742	57850	927600	254693	12750
6	2010	713336	261091	60706	954891	264283	13693
7	2011	740357	275258	63441	993872	279107	14658
8	2012	772175	291791	67282	1038472	297539	15981
9	2013	794852	305777	70878	1066182	309698	16934
10	2014	809453	318015	75324	1081488	317900	18626

In [63]:

```
ax = bachelorsdegreesbymajor_malevsfemale.plot(kind='line',x='Year',y='Males in Science & Eng',color='darkblue',label='# Male Science & Engineering Majors')
bachelorsdegreesbymajor_malevsfemale.plot(kind='line',x='Year',y='Females in Science & Eng',color='magenta',label='# Female Science & Engineering Majors',ax=ax)
ax = bachelorsdegreesbymajor_malevsfemale.plot(kind='line',x='Year',y='Males in All Sciences',color='green',label='# Male All Science Majors')
bachelorsdegreesbymajor_malevsfemale.plot(kind='line',x='Year',y='Females in All Sciences',color='aqua',label='# Female All Science Majors',ax=ax)
ax = bachelorsdegreesbymajor_malevsfemale.plot(kind='line',x='Year',y='Males in Engineering',color='orange',label='# Male Engineering Majors')
bachelorsdegreesbymajor_malevsfemale.plot(kind='line',x='Year',y='Females in Engineering',color='purple',label='# Female Engineering Major',ax=ax)

plt.xlabel('Year')
plt.ylabel('# of Male vs Female Engineering Majors')

ax.legend(bbox_to_anchor=(.85, -.1))
plt.gcf().set_size_inches((5, 5))
```







Employment Analysis : Employment Rates by Industry, Major, Year, & Gender

In [64]:
```
employmentbyyear = employmentbyyear.rename(columns = {'year': 'Year','agricultture_ratio': 'agriculture_ratio'})
employmentbyyear['Year'] = employmentbyyear['Year'].astype(int)
employmentbyyear
```

Out[64]:

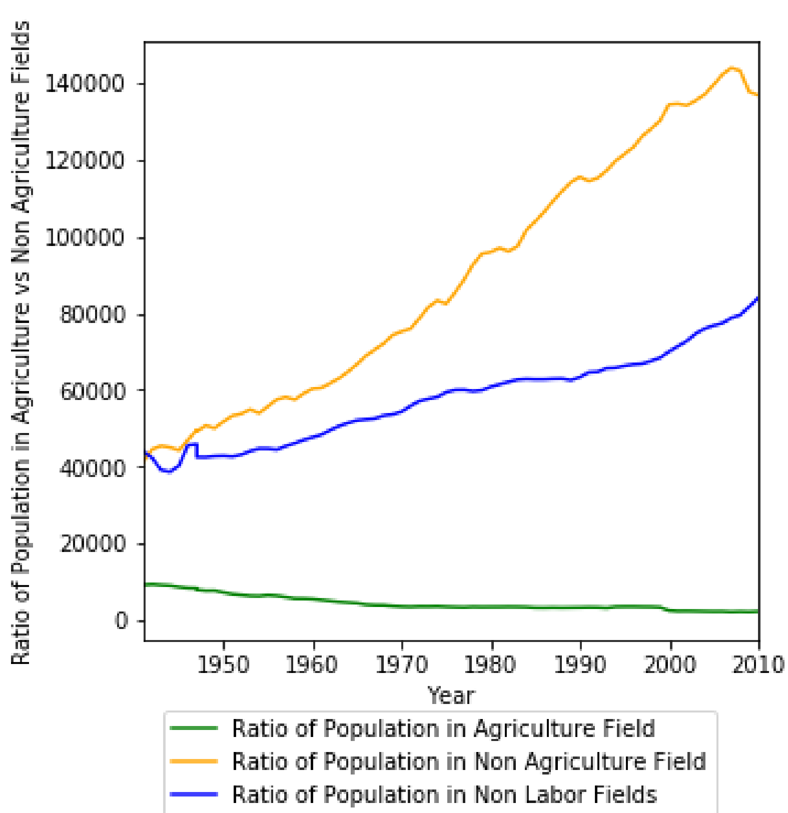
	Year	population	labor_force	population_percent	employed_total	employed_percent	agricul
0	1941	99900	55910	56.0	50350	50.4	
1	1942	98640	56410	57.2	53750	54.5	
2	1943	94640	55540	58.7	54470	57.6	
3	1944	93220	54630	58.6	53960	57.9	
4	1945	94090	53860	57.2	52820	56.1	
5	1946	103070	57520	55.8	55250	53.6	
6	1947	106018	60168	56.8	57812	54.5	
7	1947	101827	59350	58.3	57038	56.0	
8	1948	103068	60621	58.8	58343	56.6	
9	1949	103994	61286	58.9	57651	55.4	
10	1950	104995	62208	59.2	58918	56.1	
11	1951	104621	62017	59.2	59961	57.3	
12	1952	105231	62138	59.0	60250	57.3	
13	1953	107056	63015	58.9	61179	57.1	
14	1954	108321	63643	58.8	60109	55.5	
15	1955	109683	65023	59.3	62170	56.7	
16	1956	110954	66552	60.0	63799	57.5	
17	1957	112265	66929	59.6	64071	57.1	
18	1958	113727	67639	59.5	63036	55.4	
19	1959	115329	68369	59.3	64630	56.0	
20	1960	117245	69628	59.4	65778	56.1	
21	1961	118771	70459	59.3	65746	55.4	
22	1962	120153	70614	58.8	66702	55.5	
23	1963	122416	71833	58.7	67762	55.4	
24	1964	124485	73091	58.7	69305	55.7	
25	1965	126513	74455	58.9	71088	56.2	
26	1966	128058	75770	59.2	72895	56.9	
27	1967	129874	77347	59.6	74372	57.3	
28	1968	132028	78737	59.6	75920	57.5	
29	1969	134335	80734	60.1	77902	58.0	
...	...	...	...	...	...	...	
41	1981	170130	108670	63.9	100397	59.0	
42	1982	172271	110204	64.0	99526	57.8	
43	1983	174215	111550	64.0	100834	57.9	
44	1984	176383	113544	64.4	105005	59.5	
45	1985	178206	115461	64.8	107150	60.1	
46	1986	180587	117834	65.3	109597	60.7	
47	1987	182753	119865	65.6	112440	61.5	
48	1988	184613	121669	65.9	114968	62.3	
49	1989	186393	123869	66.5	117342	63.0	
50	1990	189164	125840	66.5	118793	62.8	
51	1991	190925	126346	66.2	117718	61.7	
52	1992	192805	128105	66.4	118492	61.5	
53	1993	194838	129200	66.3	120259	61.7	
54	1994	196814	131056	66.6	123060	62.5	

	Year	population	labor_force	population_percent	employed_total	employed_percent	agricul
55	1995	198584	132304	66.6	124900	62.9	
56	1996	200591	133943	66.8	126708	63.2	
57	1997	203133	136297	67.1	129558	63.8	
58	1998	205220	137673	67.1	131463	64.1	
59	1999	207753	139368	67.1	133488	64.3	
60	2000	212577	142583	67.1	136891	64.4	
61	2001	215092	143734	66.8	136933	63.7	
62	2002	217570	144863	66.6	136485	62.7	
63	2003	221168	146510	66.2	137736	62.3	
64	2004	223357	147401	66.0	139252	62.3	
65	2005	226082	149320	66.0	141730	62.7	
66	2006	228815	151428	66.2	144427	63.1	
67	2007	231867	153124	66.0	146047	63.0	
68	2008	233788	154287	66.0	145362	62.2	
69	2009	235801	154142	65.4	139877	59.3	
70	2010	237830	153889	64.7	139064	58.5	

71 rows × 12 columns

Ratio of population in Agricultural Field vs Non Agricultural Field vs Non Labor Fields

In [65]:
```
ax = employmentbyyear.plot(kind='line',x='Year',y='agriculture_ratio',color='green',label='Ratio of Population in Agriculture Field')
employmentbyyear.plot(kind='line',x='Year',y='nonagriculture_ratio',color='orange',label='Ratio of Population in Non Agriculture Field',ax=ax)
employmentbyyear.plot(kind='line',x='Year',y='not_in_labor',color='blue',label='Ratio of Population in Non Labor Fields',ax=ax)

plt.xlabel('Year')
plt.ylabel('Ratio of Population in Agriculture vs Non Agriculture Fields')

ax.legend(bbox_to_anchor=(.95,-.1))
plt.gcf().set_size_inches((5, 5))
```



Employment Rates by College Major

In [66]:

```
# employmentbymajor = employmentbymajor.drop(['Rank', 'Major_code','ShareWome
n'], axis = 1)
employmentbymajor = employmentbymajor.sort_values(['Total'], ascending = False
)
employmentbymajor = employmentbymajor.rename(index=str, columns={'Men' : 'Mal
e','Women' : 'Female'})
employmentbymajor
```

Out[66]:

	Rank	Major_code	Major	Major_category	Total	Sample_size	Male	Fem
145	146	5200	PSYCHOLOGY	Psychology & Social Work	393735	2584	86648	307
76	77	6203	BUSINESS MANAGEMENT AND ADMINISTRATION	Business	329927	4212	6383	8
123	124	3600	BIOLOGY	Biology & Life Science	280709	1370	111762	168
57	58	6200	GENERAL BUSINESS	Business	234590	2380	6053	4
93	94	1901	COMMUNICATIONS	Communications & Journalism	213996	2394	476	
34	35	6107	NURSING	Health	209394	2554	21773	187
77	78	6206	MARKETING AND MARKETING RESEARCH	Business	205211	2684	11404	7
40	41	6201	ACCOUNTING	Business	198633	2042	27392	9
137	138	3301	ENGLISH LANGUAGE AND LITERATURE	Humanities & Liberal Arts	194673	1436	4897	2
78	79	5506	POLITICAL SCIENCE AND GOVERNMENT	Social Science	182621	1387	93880	88
35	36	6207	FINANCE	Business	174506	2189	89749	49
138	139	2304	ELEMENTARY EDUCATION	Education	170862	1629	13029	157
94	95	5301	CRIMINAL JUSTICE AND FIRE PROTECTION	Law & Public Policy	152824	1728	3156	
113	114	2300	GENERAL EDUCATION	Education	143718	919	26893	116
114	115	6402	HISTORY	Humanities & Liberal Arts	141951	1058	280	
36	37	5501	ECONOMICS	Social Science	139247	1322	115030	59
20	21	2102	COMPUTER SCIENCE	Computers & Mathematics	128319	1196	1837	2
139	140	4101	PHYSICAL FITNESS PARKS RECREATION AND LEISURE	Industrial Arts & Consumer Services	125074	1014	2049	4
124	125	5507	SOCIOLOGY	Social Science	115433	1024	24704	28
95	96	6004	COMMERCIAL ART AND GRAPHIC DESIGN	Arts	103480	1186	8617	5
8	9	2414	MECHANICAL ENGINEERING	Engineering	91227	1029	12953	2
9	10	2408	ELECTRICAL ENGINEERING	Engineering	81527	631	8407	6
149	150	6000	FINE ARTS	Arts	74440	623	24786	49
96	97	1902	JOURNALISM	Communications & Journalism	72619	843	8739	22
41	42	3700	MATHEMATICS	Computers & Mathematics	72397	541	9005	2
140	141	3401	LIBERAL ARTS	Humanities & Liberal Arts	71369	569	58227	136
74	75	5003	CHEMISTRY	Physical Sciences	66530	353	32923	33

	Rank	Major_code	Major	Major_category	Total	Sample_size	Male	Fem
97	98	5098	MULTI-DISCIPLINARY OR GENERAL SCIENCE	Physical Sciences	62052	427	70619	143
17	18	2400	GENERAL ENGINEERING	Engineering	61152	425	45683	15
146	147	6002	MUSIC	Arts	60633	419	15670	12
...	...	...	...	...	...	...	...	
70	71	5205	INDUSTRIAL AND ORGANIZATIONAL PSYCHOLOGY	Psychology & Social Work	3014	24	132238	102
23	24	2413	MATERIALS ENGINEERING AND MATERIALS SCIENCE	Engineering	2993	22	8181	3
50	51	2501	ENGINEERING AND INDUSTRIAL MANAGEMENT	Engineering	2906	29	2400	
169	170	5201	EDUCATIONAL PSYCHOLOGY	Psychology & Social Work	2854	7	522	2
170	171	5202	CLINICAL PSYCHOLOGY	Psychology & Social Work	2838	13	568	2
18	19	2403	ARCHITECTURAL ENGINEERING	Engineering	2825	26	1835	
5	6	2418	NUCLEAR ENGINEERING	Engineering	2573	17	2200	
71	72	1102	AGRICULTURAL ECONOMICS	Agriculture & Natural Resources	2439	44	10624	15
75	76	5701	ELECTRICAL, MECHANICAL, AND PRECISION TECHNOLO...	Industrial Arts & Consumer Services	2435	37	1869	
49	50	5006	OCEANOGRAPHY	Physical Sciences	2418	36	752	1
0	1	2419	PETROLEUM ENGINEERING	Engineering	2339	36	2057	
39	40	5102	NUCLEAR, INDUSTRIAL RADIOLOGY, AND BIOLOGICAL ...	Physical Sciences	2116	31	7575	5
90	91	5005	GEOSCIENCES	Physical Sciences	1978	18	1589	11
7	8	5001	ASTRONOMY AND ASTROPHYSICS	Physical Sciences	1792	10	2110	1
48	49	3607	PHARMACOLOGY	Biology & Life Science	1762	3	94519	104
161	162	1199	MISCELLANEOUS AGRICULTURE	Agriculture & Natural Resources	1488	24	4754	12
72	73	5000	PHYSICAL SCIENCES	Physical Sciences	1436	10	10285	7
91	92	5206	SOCIAL PSYCHOLOGY	Psychology & Social Work	1386	8	6184	18
83	84	3602	BOTANY	Biology & Life Science	1329	9	626	
3	4	2417	NAVAL ARCHITECTURE AND MARINE ENGINEERING	Engineering	1258	16	1123	
19	20	3201	COURT REPORTING	Law & Public Policy	1148	14	877	

	Rank	Major_code	Major	Major_category	Total	Sample_size	Male	Fem
172	173	3501	LIBRARY SCIENCE	Education	1098	2	134	
2	3	2415	METALLURGICAL ENGINEERING	Engineering	856	3	725	
55	56	2303	SCHOOL STUDENT COUNSELING	Education	818	4	1667	2
120	121	2301	EDUCATIONAL ADMINISTRATION AND SUPERVISION	Education	804	5	78253	63
1	2	2416	MINING AND MINERAL ENGINEERING	Engineering	756	7	679	
33	34	2411	GEOLOGICAL AND GEOPHYSICAL ENGINEERING	Engineering	720	5	7921	3
112	113	1106	SOIL SCIENCE	Agriculture & Natural Resources	685	4	4266	13
52	53	4005	MATHEMATICS AND COMPUTER SCIENCE	Computers & Mathematics	609	7	803	10
73	74	3801	MILITARY TECHNOLOGIES	Industrial Arts & Consumer Services	124	4	1756	1

173 rows × 21 columns

Most Employable College Majors for Males Ranked Highest to Lowest

In [67]:

```
employmentbymajor_men = employmentbymajor.groupby('Major_category')['Male'].me
an().reset_index()
employmentbymajor_men = employmentbymajor_men.sort_values(['Male'], ascending
= False).reset_index()
employmentbymajor_men = employmentbymajor_men[['Major_category','Male']]
employmentbymajor_men
```

Out[67]:

	Major_category	Male
0	Social Science	55928.555556
1	Biology & Life Science	30557.357143
2	Psychology & Social Work	26205.222222
3	Agriculture & Natural Resources	19787.500000
4	Physical Sciences	18487.900000
5	Education	16107.000000
6	Engineering	13815.310345
7	Business	13609.000000
8	Arts	10847.625000
9	Humanities & Liberal Arts	10663.000000
10	Interdisciplinary	10031.000000
11	Communications & Journalism	9173.000000
12	Health	7885.833333
13	Computers & Mathematics	5420.272727
14	Industrial Arts & Consumer Services	5006.142857
15	Law & Public Policy	2096.800000

Most Employable College Majors for Females Ranked Highest to Lowest

In [68]:
```
employmentbymajor_women = employmentbymajor.groupby('Major_category')['Female'
].mean().reset_index()
employmentbymajor_women = employmentbymajor_women.sort_values(['Female'], asce
nding = False).reset_index()
employmentbymajor_women = employmentbymajor_women[['Major_category','Female']]
employmentbymajor_women
```

Out[68]:

	Major_category	Female
0	Psychology & Social Work	56073.555556
1	Social Science	51512.888889
2	Biology & Life Science	41295.142857
3	Education	38309.875000
4	Health	26002.166667
5	Agriculture & Natural Resources	24981.200000
6	Communications & Journalism	24569.500000
7	Physical Sciences	24499.600000
8	Humanities & Liberal Arts	23309.066667
9	Arts	17558.625000
10	Interdisciplinary	9848.000000
11	Business	8489.769231
12	Computers & Mathematics	5690.818182
13	Industrial Arts & Consumer Services	5260.571429
14	Engineering	4070.724138
15	Law & Public Policy	1095.600000

Most Employable Colleges Majors for Males & Females Ranked Highest to Lowest

In [69]:
```
employmentbymajor_employed = employmentbymajor.groupby('Major_category')['Male','Female','Employed'].mean().reset_index()
employmentbymajor_employed = employmentbymajor_employed.sort_values(['Employed'], ascending = False).reset_index()
employmentbymajor_employed = employmentbymajor_employed[['Major_category','Male','Female','Employed']]
employmentbymajor_employed
```

Out[69]:

	Major_category	Male	Female	Employed
0	Business	13609.000000	8489.769231	83749.384615
1	Communications & Journalism	9173.000000	24569.500000	82665.000000
2	Social Science	55928.555556	51512.888889	44610.333333
3	Psychology & Social Work	26205.222222	56073.555556	42260.444444
4	Humanities & Liberal Arts	10663.000000	23309.066667	36274.533333
5	Arts	10847.625000	17558.625000	36014.250000
6	Health	7885.833333	26002.166667	31012.250000
7	Education	16107.000000	38309.875000	29989.937500
8	Law & Public Policy	2096.800000	1095.600000	28958.000000
9	Industrial Arts & Consumer Services	5006.142857	5260.571429	27006.142857
10	Biology & Life Science	30557.357143	41295.142857	21628.357143
11	Computers & Mathematics	5420.272727	5690.818182	21626.727273
12	Engineering	13815.310345	4070.724138	14495.586207
13	Physical Sciences	18487.900000	24499.600000	13923.100000
14	Interdisciplinary	10031.000000	9848.000000	9821.000000
15	Agriculture & Natural Resources	19787.500000	24981.200000	6694.300000

Most Employable College Majors for Full Time Positions for Males & Females Ranked Highest to Lowest

In [70]:
```
employmentbymajor_fulltime = employmentbymajor.groupby('Major_category')['Male','Female','Full_time'].mean().reset_index()
employmentbymajor_fulltime = employmentbymajor_fulltime.sort_values(['Full_time'], ascending = False).reset_index()
employmentbymajor_fulltime = employmentbymajor_fulltime[['Major_category','Male','Female','Full_time']]
employmentbymajor_fulltime
```

Out[70]:

	Major_category	Male	Female	Full_time
0	Business	13609.000000	8489.769231	76066.923077
1	Communications & Journalism	9173.000000	24569.500000	68332.500000
2	Social Science	55928.555556	51512.888889	38571.222222
3	Psychology & Social Work	26205.222222	56073.555556	32111.111111
4	Humanities & Liberal Arts	10663.000000	23309.066667	27795.933333
5	Arts	10847.625000	17558.625000	25971.625000
6	Law & Public Policy	2096.800000	1095.600000	25388.000000
7	Education	16107.000000	38309.875000	24878.687500
8	Health	7885.833333	26002.166667	24568.250000
9	Industrial Arts & Consumer Services	5006.142857	5260.571429	21626.142857
10	Computers & Mathematics	5420.272727	5690.818182	18867.727273
11	Biology & Life Science	30557.357143	41295.142857	17169.785714
12	Engineering	13815.310345	4070.724138	13167.827586
13	Physical Sciences	18487.900000	24499.600000	11285.200000
14	Interdisciplinary	10031.000000	9848.000000	8032.000000
15	Agriculture & Natural Resources	19787.500000	24981.200000	5814.300000

Most Employable College Majors for Part Time Positions for Males & Females Ranked Highest to Lowest

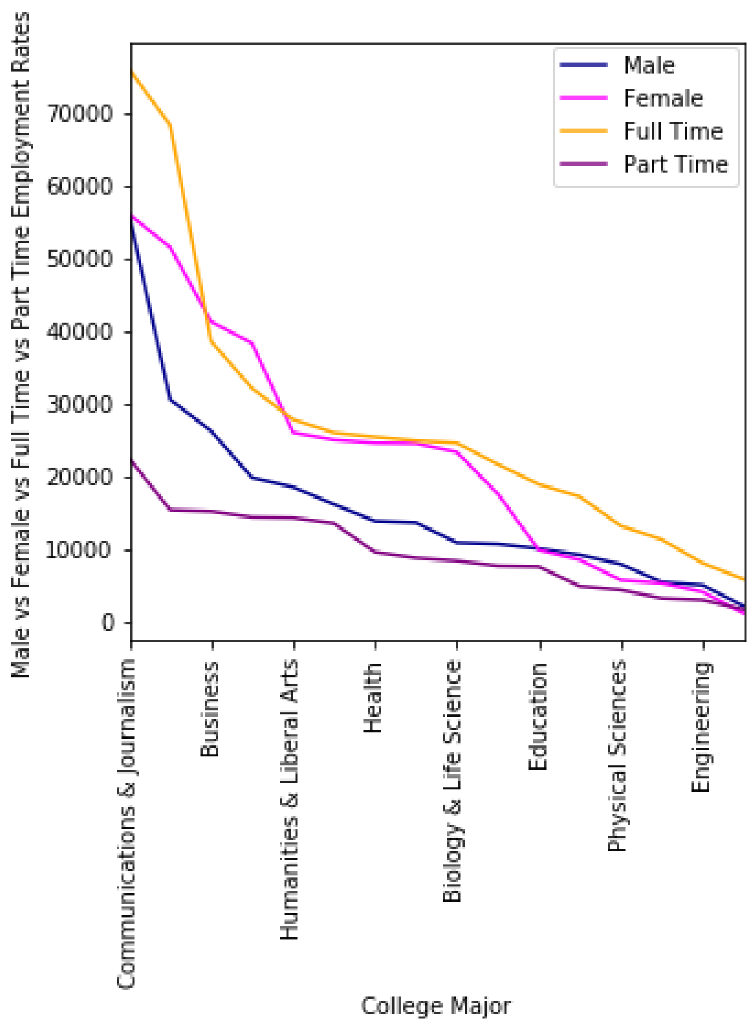
In [71]:
```
employmentbymajor_parttime = employmentbymajor.groupby('Major_category')['Male','Female','Part_time'].mean().reset_index()
employmentbymajor_parttime = employmentbymajor_parttime.sort_values(['Part_time'], ascending = False).reset_index()
employmentbymajor_parttime = employmentbymajor_parttime[['Major_category','Male','Female','Part_time']]
employmentbymajor_parttime
```

Out[71]:

	Major_category	Male	Female	Part_time
0	Communications & Journalism	9173.000000	24569.500000	22454.250000
1	Psychology & Social Work	26205.222222	56073.555556	15332.444444
2	Business	13609.000000	8489.769231	15148.923077
3	Arts	10847.625000	17558.625000	14348.875000
4	Humanities & Liberal Arts	10663.000000	23309.066667	14268.666667
5	Social Science	55928.555556	51512.888889	13507.666667
6	Health	7885.833333	26002.166667	9549.333333
7	Industrial Arts & Consumer Services	5006.142857	5260.571429	8731.714286
8	Biology & Life Science	30557.357143	41295.142857	8338.285714
9	Law & Public Policy	2096.800000	1095.600000	7642.600000
10	Education	16107.000000	38309.875000	7537.062500
11	Computers & Mathematics	5420.272727	5690.818182	4842.727273
12	Physical Sciences	18487.900000	24499.600000	4344.400000
13	Interdisciplinary	10031.000000	9848.000000	3173.000000
14	Engineering	13815.310345	4070.724138	2935.724138
15	Agriculture & Natural Resources	19787.500000	24981.200000	1659.100000

Most Employable Majors (Male vs Female vs Full Time vs Part Time)

In [72]:
```
ax = employmentbymajor_men.plot(kind='line',x='Major_category',y='Male',color=
'darkblue',label='Male')
employmentbymajor_women.plot(kind='line',x='Major_category',y='Female',color=
'magenta',label='Female',ax=ax)
employmentbymajor_fulltime.plot(kind='line',x='Major_category',y='Full_time',c
olor='orange',label='Full Time',ax=ax)
employmentbymajor_parttime.plot(kind='line',x='Major_category',y='Part_time',c
olor='purple',label='Part Time',ax=ax)

plt.xlabel('College Major')
plt.ylabel('Male vs Female vs Full Time vs Part Time Employment Rates')

plt.xticks(rotation=90)

ax.legend(bbox_to_anchor=(1.01, 1.01))
plt.gcf().set_size_inches((5, 5))
```



Salary Analysis : Starting Median, Mid-Career Median, & Percent Change from Starting to Mid-Career Salary by College Major

In [73]:

```
salarybymajor =  salarybymajor.rename(index=str, columns = ({'Undergraduate Major' : 'Major'}))
salarybymajor
```

Out[73]:

	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-C 90th Perce Salar
0	Accounting	$46,000.00	$77,100.00	67.6	$42,200.00	$56,100.00	$108,000.00	$152,
1	Aerospace Engineering	$57,700.00	$101,000.00	75.0	$64,300.00	$82,100.00	$127,000.00	$161,
2	Agriculture	$42,600.00	$71,900.00	68.8	$36,300.00	$52,100.00	$96,300.00	$150,
3	Anthropology	$36,800.00	$61,500.00	67.1	$33,800.00	$45,500.00	$89,300.00	$138,
4	Architecture	$41,600.00	$76,800.00	84.6	$50,600.00	$62,200.00	$97,000.00	$136,
5	Art History	$35,800.00	$64,900.00	81.3	$28,800.00	$42,200.00	$87,400.00	$125,
6	Biology	$38,800.00	$64,800.00	67.0	$36,900.00	$47,400.00	$94,500.00	$135,
7	Business Management	$43,000.00	$72,100.00	67.7	$38,800.00	$51,500.00	$102,000.00	$147,
8	Chemical Engineering	$63,200.00	$107,000.00	69.3	$71,900.00	$87,300.00	$143,000.00	$194,
9	Chemistry	$42,600.00	$79,900.00	87.6	$45,300.00	$60,700.00	$108,000.00	$148,
10	Civil Engineering	$53,900.00	$90,500.00	67.9	$63,400.00	$75,100.00	$115,000.00	$148,
11	Communications	$38,100.00	$70,000.00	83.7	$37,500.00	$49,700.00	$98,800.00	$143,
12	Computer Engineering	$61,400.00	$105,000.00	71.0	$66,100.00	$84,100.00	$135,000.00	$162,
13	Computer Science	$55,900.00	$95,500.00	70.8	$56,000.00	$74,900.00	$122,000.00	$154,
14	Construction	$53,700.00	$88,900.00	65.5	$56,300.00	$68,100.00	$118,000.00	$171,
15	Criminal Justice	$35,000.00	$56,300.00	60.9	$32,200.00	$41,600.00	$80,700.00	$107,
16	Drama	$35,900.00	$56,900.00	58.5	$36,700.00	$41,300.00	$79,100.00	$153,
17	Economics	$50,100.00	$98,600.00	96.8	$50,600.00	$70,600.00	$145,000.00	$210,
18	Education	$34,900.00	$52,000.00	49.0	$29,300.00	$37,900.00	$73,400.00	$102,
19	Electrical Engineering	$60,900.00	$103,000.00	69.1	$69,300.00	$83,800.00	$130,000.00	$168,
20	English	$38,000.00	$64,700.00	70.3	$33,400.00	$44,800.00	$93,200.00	$133,
21	Film	$37,900.00	$68,500.00	80.7	$33,900.00	$45,500.00	$100,000.00	$136,
22	Finance	$47,900.00	$88,300.00	84.3	$47,200.00	$62,100.00	$128,000.00	$195,
23	Forestry	$39,100.00	$62,600.00	60.1	$41,000.00	$49,300.00	$78,200.00	$111,
24	Geography	$41,200.00	$65,500.00	59.0	$40,000.00	$50,000.00	$90,800.00	$132,
25	Geology	$43,500.00	$79,500.00	82.8	$45,000.00	$59,600.00	$101,000.00	$156,
26	Graphic Design	$35,700.00	$59,800.00	67.5	$36,000.00	$45,500.00	$80,800.00	$112,
27	Health Care Administration	$38,800.00	$60,600.00	56.2	$34,600.00	$45,600.00	$78,800.00	$101,
28	History	$39,200.00	$71,000.00	81.1	$37,000.00	$49,200.00	$103,000.00	$149,
29	Hospitality & Tourism	$37,800.00	$57,500.00	52.1	$35,500.00	$43,600.00	$81,900.00	$124,
30	Industrial Engineering	$57,700.00	$94,700.00	64.1	$57,100.00	$72,300.00	$132,000.00	$173,
31	Information Technology (IT)	$49,100.00	$74,800.00	52.3	$44,500.00	$56,700.00	$96,700.00	$129,
32	Interior Design	$36,100.00	$53,200.00	47.4	$35,700.00	$42,600.00	$72,500.00	$107,
33	International Relations	$40,900.00	$80,900.00	97.8	$38,200.00	$56,000.00	$111,000.00	$157,
34	Journalism	$35,600.00	$66,700.00	87.4	$38,400.00	$48,300.00	$97,700.00	$145,

	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-C 90th Perce Salar
35	Management Information Systems (MIS)	$49,200.00	$82,300.00	67.3	$45,300.00	$60,500.00	$108,000.00	$146,
36	Marketing	$40,800.00	$79,600.00	95.1	$42,100.00	$55,600.00	$119,000.00	$175,
37	Math	$45,400.00	$92,400.00	103.5	$45,200.00	$64,200.00	$128,000.00	$183,
38	Mechanical Engineering	$57,900.00	$93,600.00	61.7	$63,700.00	$76,200.00	$120,000.00	$163,
39	Music	$35,900.00	$55,000.00	53.2	$26,700.00	$40,200.00	$88,000.00	$134,
40	Nursing	$54,200.00	$67,000.00	23.6	$47,600.00	$56,400.00	$80,900.00	$98,
41	Nutrition	$39,900.00	$55,300.00	38.6	$33,900.00	$44,500.00	$70,500.00	$99,
42	Philosophy	$39,900.00	$81,200.00	103.5	$35,500.00	$52,800.00	$127,000.00	$168,
43	Physician Assistant	$74,300.00	$91,700.00	23.4	$66,400.00	$75,200.00	$108,000.00	$124,
44	Physics	$50,300.00	$97,300.00	93.4	$56,000.00	$74,200.00	$132,000.00	$178,
45	Political Science	$40,800.00	$78,200.00	91.7	$41,200.00	$55,300.00	$114,000.00	$168,
46	Psychology	$35,900.00	$60,400.00	68.2	$31,600.00	$42,100.00	$87,500.00	$127,
47	Religion	$34,100.00	$52,000.00	52.5	$29,700.00	$36,500.00	$70,900.00	$96,
48	Sociology	$36,500.00	$58,200.00	59.5	$30,700.00	$40,400.00	$81,200.00	$118,
49	Spanish	$34,000.00	$53,100.00	56.2	$31,000.00	$40,000.00	$76,800.00	$96,

Starting Median Salary by Major Ranked Highest to Lowest

In [74]:
```
salarybymajor_starting = salarybymajor[['Major','Starting Median Salary']].apply(lambda x: x.str.replace(',','')).apply(lambda x: x.str.replace('$',''))
salarybymajor_starting = salarybymajor_starting.sort_values('Starting Median Salary', ascending = False).reset_index()
salarybymajor_starting['Starting Median Salary'] = salarybymajor_starting['Starting Median Salary'].astype(float)
salarybymajor_starting
```

Out[74]:

	index	Major	Starting Median Salary
0	43	Physician Assistant	74300.0
1	8	Chemical Engineering	63200.0
2	12	Computer Engineering	61400.0
3	19	Electrical Engineering	60900.0
4	38	Mechanical Engineering	57900.0
5	1	Aerospace Engineering	57700.0
6	30	Industrial Engineering	57700.0
7	13	Computer Science	55900.0
8	40	Nursing	54200.0
9	10	Civil Engineering	53900.0
10	14	Construction	53700.0
11	44	Physics	50300.0
12	17	Economics	50100.0
13	35	Management Information Systems (MIS)	49200.0
14	31	Information Technology (IT)	49100.0
15	22	Finance	47900.0
16	0	Accounting	46000.0
17	37	Math	45400.0
18	25	Geology	43500.0
19	7	Business Management	43000.0
20	2	Agriculture	42600.0
21	9	Chemistry	42600.0
22	4	Architecture	41600.0
23	24	Geography	41200.0
24	33	International Relations	40900.0
25	45	Political Science	40800.0
26	36	Marketing	40800.0
27	42	Philosophy	39900.0
28	41	Nutrition	39900.0
29	28	History	39200.0
30	23	Forestry	39100.0
31	27	Health Care Administration	38800.0
32	6	Biology	38800.0
33	11	Communications	38100.0
34	20	English	38000.0
35	21	Film	37900.0
36	29	Hospitality & Tourism	37800.0
37	3	Anthropology	36800.0
38	48	Sociology	36500.0
39	32	Interior Design	36100.0
40	46	Psychology	35900.0
41	39	Music	35900.0
42	16	Drama	35900.0
43	5	Art History	35800.0
44	26	Graphic Design	35700.0

	index	Major	Starting Median Salary
45	34	Journalism	35600.0
46	15	Criminal Justice	35000.0
47	18	Education	34900.0
48	47	Religion	34100.0
49	49	Spanish	34000.0

Mid-Career Median Salary by Major Ranked Highest to Lowest

In [75]:
```
salarybymajor_midcareer = salarybymajor[['Major','Mid-Career Median Salary']].
apply(lambda x: x.str.replace(',','')).apply(lambda x: x.str.replace('$',''))
salarybymajor_midcareer = salarybymajor_midcareer.sort_values('Mid-Career Medi
an Salary', ascending = False).reset_index()
salarybymajor_midcareer['Mid-Career Median Salary'] = salarybymajor_midcareer[
'Mid-Career Median Salary'].astype(float)
salarybymajor_midcareer
```

Out[75]:

	index	Major	Mid-Career Median Salary
0	17	Economics	98600.0
1	44	Physics	97300.0
2	13	Computer Science	95500.0
3	30	Industrial Engineering	94700.0
4	38	Mechanical Engineering	93600.0
5	37	Math	92400.0
6	43	Physician Assistant	91700.0
7	10	Civil Engineering	90500.0
8	14	Construction	88900.0
9	22	Finance	88300.0
10	35	Management Information Systems (MIS)	82300.0
11	42	Philosophy	81200.0
12	33	International Relations	80900.0
13	9	Chemistry	79900.0
14	36	Marketing	79600.0
15	25	Geology	79500.0
16	45	Political Science	78200.0
17	0	Accounting	77100.0
18	4	Architecture	76800.0
19	31	Information Technology (IT)	74800.0
20	7	Business Management	72100.0
21	2	Agriculture	71900.0
22	28	History	71000.0
23	11	Communications	70000.0
24	21	Film	68500.0
25	40	Nursing	67000.0
26	34	Journalism	66700.0
27	24	Geography	65500.0
28	5	Art History	64900.0
29	6	Biology	64800.0
30	20	English	64700.0
31	23	Forestry	62600.0
32	3	Anthropology	61500.0
33	27	Health Care Administration	60600.0
34	46	Psychology	60400.0
35	26	Graphic Design	59800.0
36	48	Sociology	58200.0
37	29	Hospitality & Tourism	57500.0
38	16	Drama	56900.0
39	15	Criminal Justice	56300.0
40	41	Nutrition	55300.0
41	39	Music	55000.0
42	32	Interior Design	53200.0
43	49	Spanish	53100.0
44	47	Religion	52000.0

	index	Major	Mid-Career Median Salary
45	18	Education	52000.0
46	8	Chemical Engineering	107000.0
47	12	Computer Engineering	105000.0
48	19	Electrical Engineering	103000.0
49	1	Aerospace Engineering	101000.0

Percent Change from Starting to Mid-Career Median Salary by Major Ranked Highest to Lowest

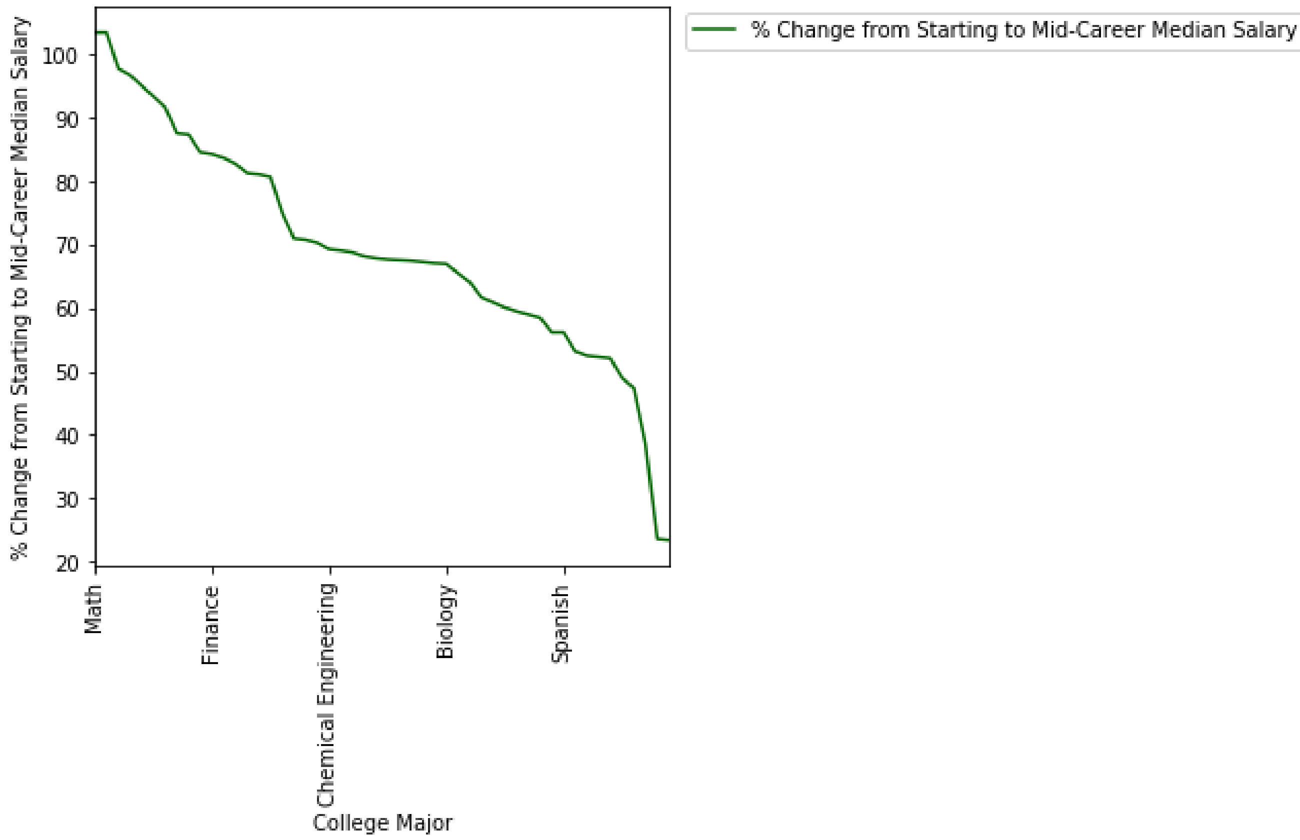
In [76]:
```
salarybymajor_percentchange = salarybymajor.groupby('Major')['Percent change from Starting to Mid-Career Salary'].mean().reset_index()
salarybymajor_percentchange = salarybymajor_percentchange.sort_values(['Percent change from Starting to Mid-Career Salary'], ascending = False).reset_index()
salarybymajor_percentchange = salarybymajor_percentchange[['Major','Percent change from Starting to Mid-Career Salary']]
salarybymajor_percentchange = pd.DataFrame(salarybymajor_percentchange[['Major','Percent change from Starting to Mid-Career Salary']])
salarybymajor_percentchange.columns = ['Major','% Change from Starting to Mid-Career Salary']
salarybymajor_percentchange
```

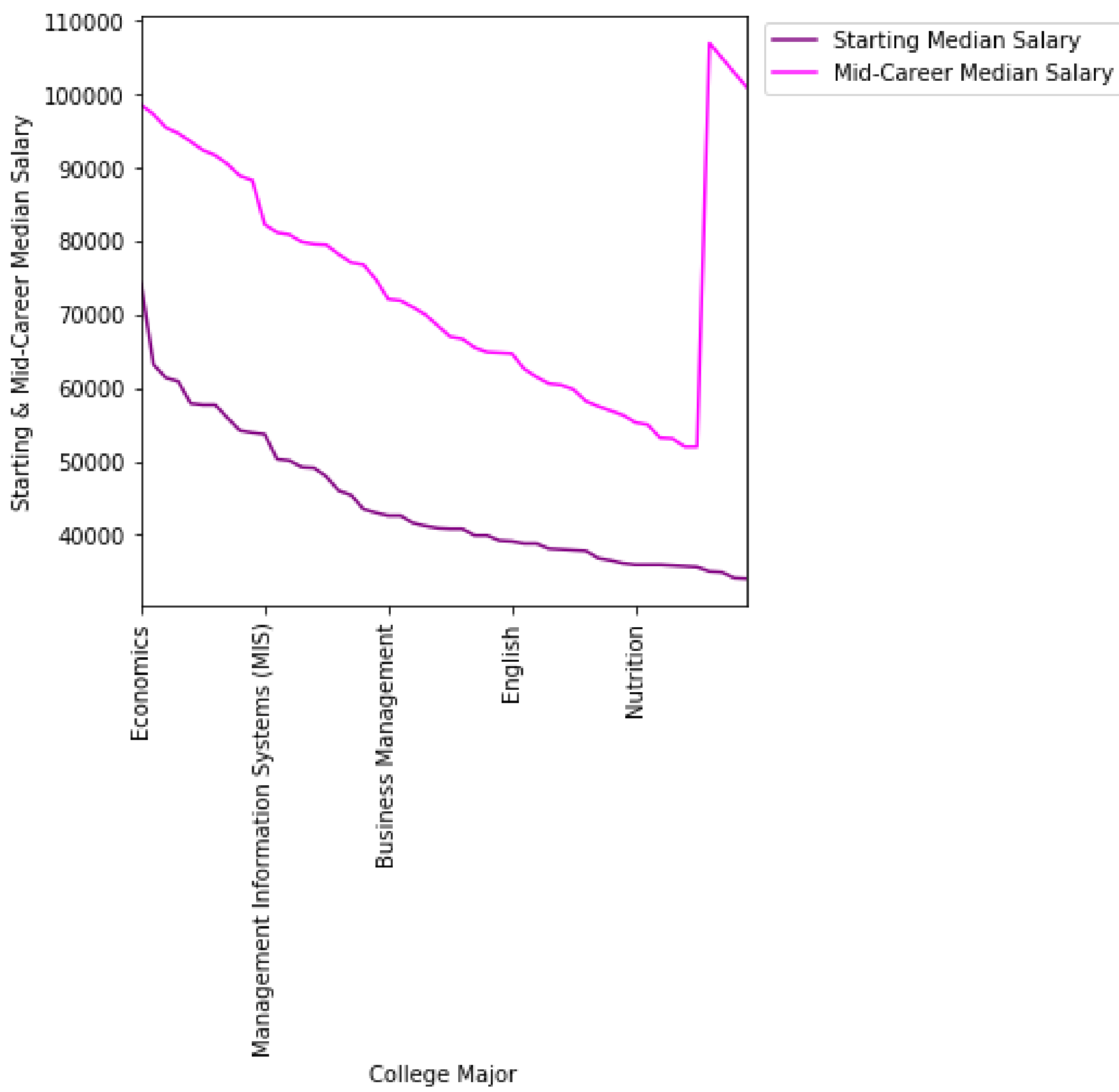
Out[76]:

	Major	% Change from Starting to Mid-Career Salary
0	Math	103.5
1	Philosophy	103.5
2	International Relations	97.8
3	Economics	96.8
4	Marketing	95.1
5	Physics	93.4
6	Political Science	91.7
7	Chemistry	87.6
8	Journalism	87.4
9	Architecture	84.6
10	Finance	84.3
11	Communications	83.7
12	Geology	82.8
13	Art History	81.3
14	History	81.1
15	Film	80.7
16	Aerospace Engineering	75.0
17	Computer Engineering	71.0
18	Computer Science	70.8
19	English	70.3
20	Chemical Engineering	69.3
21	Electrical Engineering	69.1
22	Agriculture	68.8
23	Psychology	68.2
24	Civil Engineering	67.9
25	Business Management	67.7
26	Accounting	67.6
27	Graphic Design	67.5
28	Management Information Systems (MIS)	67.3
29	Anthropology	67.1
30	Biology	67.0
31	Construction	65.5
32	Industrial Engineering	64.1
33	Mechanical Engineering	61.7
34	Criminal Justice	60.9
35	Forestry	60.1
36	Sociology	59.5
37	Geography	59.0
38	Drama	58.5
39	Health Care Administration	56.2
40	Spanish	56.2
41	Music	53.2
42	Religion	52.5
43	Information Technology (IT)	52.3
44	Hospitality & Tourism	52.1

	Major	% Change from Starting to Mid-Career Salary
45	Education	49.0
46	Interior Design	47.4
47	Nutrition	38.6
48	Nursing	23.6
49	Physician Assistant	23.4

In [77]:
```
ax = salarybymajor_percentchange.plot(kind='line',x='Major',y='% Change from Starting to Mid-Career Salary',color='darkgreen',label='% Change from Starting to Mid-Career Median Salary')

plt.xlabel('College Major')
plt.ylabel('% Change from Starting to Mid-Career Median Salary')

plt.xticks(rotation=90)

ax.legend(bbox_to_anchor=(1.01, 1.01))
plt.gcf().set_size_inches((5, 5))

ax = salarybymajor_starting.plot(kind='line',x='Major',y='Starting Median Salary',color='purple',label='Starting Median Salary')
salarybymajor_midcareer.plot(kind='line',x='Major',y='Mid-Career Median Salary',color='magenta',label='Mid-Career Median Salary',ax=ax)

plt.xlabel('College Major')
plt.ylabel('Starting & Mid-Career Median Salary')

plt.xticks(rotation=90)

ax.legend(bbox_to_anchor=(1.01, 1.01))
plt.gcf().set_size_inches((5, 5))

# ax = salarybymajor_midcareer.plot(kind='line',x='Major',y='Mid-Career Median Salary',color='magenta',label='Mid-Career Median Salary')

# plt.xlabel('Average # Men vs Women')
# plt.ylabel('Major_category')

# plt.xticks(rotation=90)

# ax.legend(bbox_to_anchor=(1.01, 1.01))
# plt.gcf().set_size_inches((5, 5))
```


% Change from Starting to Mid-Career Median Salary
% Change from Starting to Mid-Career Median Salary
100
90
80
70
60
50
40
30
20
Math
Finance
Chemical Engineering
Biology
Spanish
College Major


Starting Median Salary
Mid-Career Median Salary
Starting & Mid-Career Median Salary
110000
100000
90000
80000
70000
60000
50000
40000
Economics
Management Information Systems (MIS)
Business Management
English
Nutrition
College Major