

# College Major With MySQL

YoungJin Seo

# Description

- My goal is to analyze the most popular major in Georgetown University.
- All data is from American Community Survey 2010-2012 Public Use Microdata Series.
- Kaggle : <https://www.kaggle.com/tunguz/college-majors>

## **Majors.csv**

- List of majors with their FOD1P codes and major categories.
- Major categories are from Carnevale et al, "What's It Worth?: The Economic Value of College Majors." Georgetown University Center on Education and the Workforce, 2011. <http://cew.georgetown.edu/whatsitworth>

## **Grads.csv**

- It contains a more detailed breakdown, including by sex and by the type of job they got.

## **Recentgrads.csv**

- It contains a more detailed breakdown, including by sex and by the type of job they got.

## **Stem.csv**

- It contains data for scatter plot in associated DataLab post on women in science/technology jobs.

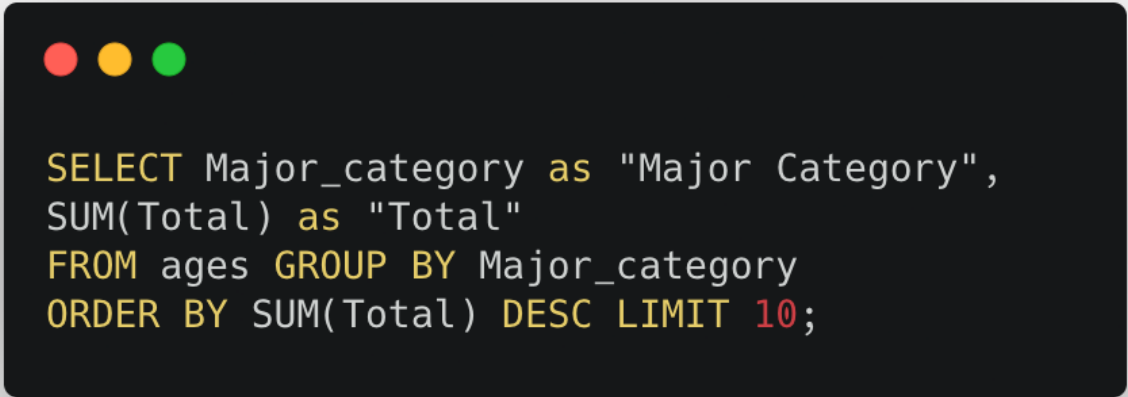
## **Ages.csv**

- It contains majors, employed , unemployment and so on .

Headers for recentgrads.csv are shown below:

Header	Description		
Rank	Rank by median earnings		
Major_code	Major code, FO1DP in ACS PUMS		
Major	Major description		
Major_category	Category of major from Carnevale et al		
Total	Total number of people with major	Unemployment_rate	Unemployed / (Unemployed + Employed)
Sample_size	Sample size (unweighted) of full-time, year-round O (used for earnings)	Median	Median earnings of full-time, year-round workers
Men	Male graduates	P25th	25th percentile of earnings
Women	Female graduates	P75th	75th percentile of earnings
ShareWomen	Women as share of total	College_jobs	Number with job requiring a college degree
Employed	Number employed (ESR == 1 or 2)	Non_college_jobs	Number with job not requiring a college degree
Full_time	Employed 35 hours or more	Low_wage_jobs	Number in low-wage service jobs
Part_time	Employed less than 35 hours		
Full_time_year_round	Employed at least 50 weeks (WKW == 1) and at least 35 hours (WKHP >= 35)		
Unemployed	Number unemployed (ESR == 3)		

# Make top 10 list of the major category



```
SELECT Major_category as "Major Category",  
SUM(Total) as "Total"  
FROM ages GROUP BY Major_category  
ORDER BY SUM(Total) DESC LIMIT 10;
```

Major Category	Total
Business	9858741
Education	4700118
Humanities & Liberal Arts	3738335
Engineering	3576013
Health	2950859
Social Science	2654125
Psychology & Social Work	1987278
Arts	1805865
Communications & Journalism	1803822
Computers & Mathematics	1781378

What was the famous major among graduate students? (Over 20000)

```
SELECT FOD1P, Major
FROM majors
WHERE Major = ANY (SELECT Major
FROM grad WHERE Grad_total > 20000)
LIMIT 5;
```

FOD1P	Major
1100	GENERAL AGRICULTURE
1103	ANIMAL SCIENCES
1105	PLANT SCIENCE AND AGRONOMY
1302	FORESTRY
1303	NATURAL RESOURCES MANAGEMENT

## Make the column “Percentage”

```
SELECT Major, Total * 100/  
sub.sum_total 'Percentage'  
FROM recentgrads  
CROSS JOIN  
(SELECT SUM(Total) sum_Total FROM  
recentgrads) sub;
```

Major	Percentage
PETROLEUM ENGINEERING	0.0345
MINING AND MINERAL ENGINEERING	0.0112
METALLURGICAL ENGINEERING	0.0126
NAVAL ARCHITECTURE AND MARINE ENGIN...	0.0186
CHEMICAL ENGINEERING	0.4764
NUCLEAR ENGINEERING	0.0380
ACTUARIAL SCIENCE	0.0558
ASTRONOMY AND ASTROPHYSICS	0.0265
MECHANICAL ENGINEERING	1.3472
ELECTRICAL ENGINEERING	1.2039
COMPUTER ENGINEERING	0.6135
AEROSPACE ENGINEERING	0.2224
BIOMEDICAL ENGINEERING	0.2208
MATERIALS SCIENCE	0.0632
ENGINEERING MECHANICS PHYSICS AND S...	0.0638
BIOLOGICAL ENGINEERING	0.1318

## Create view as result

```
CREATE VIEW Result
AS SELECT stem.Major_code,
          majors.FOD1P,
          stem.Major,
          stem.Total
FROM stem, majors
WHERE stem.Major = majors.Major
ORDER BY stem.Total DESC;

SELECT * FROM Result;
```

Major_code	FOD1P	Major	Total
3600	3600	BIOLOGY	280709
6107	6107	NURSING	209394
2102	2102	COMPUTER SCIENCE	128319
2414	2414	MECHANICAL ENGINEERING	91227
2408	2408	ELECTRICAL ENGINEERING	81527
3700	3700	MATHEMATICS	72397
5003	5003	CHEMISTRY	66530
5098	5098	MULTI-DISCIPLINARY OR GENERAL SCIENCE	62052
2400	2400	GENERAL ENGINEERING	61152
2406	2406	CIVIL ENGINEERING	53153
6109	6109	TREATMENT THERAPY PROFESSIONS	48491
1401	1401	ARCHITECTURE	46420
2407	2407	COMPUTER ENGINEERING	41542
3601	3601	BIOCHEMICAL SCIENCES	39107
6102	6102	COMMUNICATION DISORDERS SCIENCES A...	38279
2100	2100	COMPUTER AND INFORMATION SYSTEMS	36698



Find a major as “CO%”

```
SELECT MAJOR,MEN FROM recentgrads
WHERE Major LIKE 'CO%'
ORDER BY MEN DESC ;
```

MAJOR	MEN
COMPUTER SCIENCE	99743
COMMUNICATIONS	70619
COMPUTER ENGINEERING	33258
COMMERCIAL ART AND GRAPHIC DESIGN	32041
COMPUTER AND INFORMATION SYSTEMS	27392
CONSTRUCTION SERVICES	16820
COMMUNICATION TECHNOLOGIES	11431
COMPOSITION AND RHETORIC	7022
COMPUTER ADMINISTRATION MANAGEMEN...	6607
COMPUTER NETWORKING AND TELECOMM...	5291
COSMETOLOGY SERVICES AND CULINARY...	4364
COMMUNITY AND PUBLIC HEALTH	4103
COMPUTER PROGRAMMING AND DATA PRO...	3046
COGNITIVE SCIENCE AND BIOPSYCHOLOGY	1667

Find a major as “A%”

```
SELECT Major, WOMEN FROM recentgrads
WHERE Major LIKE 'A%'
ORDER BY WOMEN DESC;
```

Major	WOMEN
ACCOUNTING	104114
ADVERTISING AND PUBLIC RELATIONS	40300
ANTHROPOLOGY AND ARCHEOLOGY	27468
ART AND MUSIC EDUCATION	23449
AREA ETHNIC AND CIVILIZATION STUDIES	22456
ARCHITECTURE	20957
ART HISTORY AND CRITICISM	17790
ANIMAL SCIENCES	16226
AGRICULTURE PRODUCTION AND MANAGE...	4582
APPLIED MATHEMATICS	2145
AEROSPACE ENGINEERING	2105
ACTUARIAL SCIENCE	1667
ATMOSPHERIC SCIENCES AND METEOROL...	1299
ARCHITECTURAL ENGINEERING	990
ASTRONOMY AND ASTROPHYSICS	960

```

SELECT Major_code, Major, Major_category,
CASE
    WHEN MEN > WOMEN THEN 'This major is famous for men'
    WHEN MEN = WOMEN THEN 'This major is famous for women'
    ELSE "This major is famous for women"
END AS MajorText
FROM stem
ORDER BY Total;

```

Create a column “MajorText”

Major_code	Major	Major_category	MajorText
4005	MATHEMATICS AND COMPUTER SCIENCE	Computers & Mathematics	This major is famous for men
2411	GEOLOGICAL AND GEOPHYSICAL ENGINEERING	Engineering	This major is famous for men
2416	MINING AND MINERAL ENGINEERING	Engineering	This major is famous for men
2415	METALLURGICAL ENGINEERING	Engineering	This major is famous for men
2417	NAVAL ARCHITECTURE AND MARINE ENGINEERING	Engineering	This major is famous for men
3602	BOTANY	Biology & Life Science	This major is famous for women
5000	PHYSICAL SCIENCES	Physical Sciences	This major is famous for men
3607	PHARMACOLOGY	Biology & Life Science	This major is famous for women
5001	ASTRONOMY AND ASTROPHYSICS	Physical Sciences	This major is famous for women
5005	GEOSCIENCES	Physical Sciences	This major is famous for women
5102	NUCLEAR, INDUSTRIAL RADIOLOGY, AND BIOLOGICAL SCIENCES	Physical Sciences	This major is famous for women
2419	PETROLEUM ENGINEERING	Engineering	This major is famous for men
5006	OCEANOGRAPHY	Physical Sciences	This major is famous for women
2418	NUCLEAR ENGINEERING	Engineering	This major is famous for men
2403	ARCHITECTURAL ENGINEERING	Engineering	This major is famous for men
2501	ENGINEERING AND INDUSTRIAL MANAGEMENT	Engineering	This major is famous for men

# Result

- Business was the most popular college major among college students.
- General Agriculture was the most popular college major among graduate students.
- STEM students took Biology a lot.
- There were 99, 743 male students who recently graduated took Computer Science.
- There were above 100,000 female students who recently graduated took Accounting.

Please click on those links below.

Medium : <https://medium.com/@annsyj94>

Github : <https://github.com/annsyj94>

LinkedIn : <https://www.linkedin.com/in/kelly94/>

Thank You