

♥ How College Major and Occupation Impact LOVE ♥

Dating MAJJic: Milla (mshin7), Alan (atang5), Jenny (jhe19), Joseph (jchen88)

Hypothesis

Dating is an important part of one's adulthood. We were specifically interested in compatibility and what allows two individuals to find each other and make good partners, depending on their stage in adulthood. As such, we investigated the link between one's academic and professional background and their tendency to date or marry those of the same college major or career.

Our main hypothesis is that an individual's initial attraction to someone, while in higher education, is positively influenced by whether they have the same college major, but not whether they have the same intended career.

Conversely, marriage is more strongly influenced by whether partners have the same current occupation than whether they had the same college major. This is because we believe that dating, relationships, and marriage are products of proximity to other people and who someone is more likely to meet and interact with on a day-to-day basis.

Data

Kaggle speed-dating dataset (publically available from Kaggle): collected by a research group led by Professors Ray Fisman and Sheena Iyengar from Columbia University Business School. The research team performed a speed dating experiment in 2002-2004 to determine what factors influence an individual's degree of attraction for another individual during a date and the likelihood of the two individuals going on a second date.

US Census dataset: collected from the ACS (American community survey) PUMS (public use microdata sample) files. We retrieved the data from IPUMS, which allows you to extract custom rows and columns from the census data. The data was chosen to match the Kaggle dataset; i.e. filtered for individuals from the state of New York who have both obtained a degree beyond a bachelor's. The data comes from 2009 and 2010 (earliest years with degree info available).

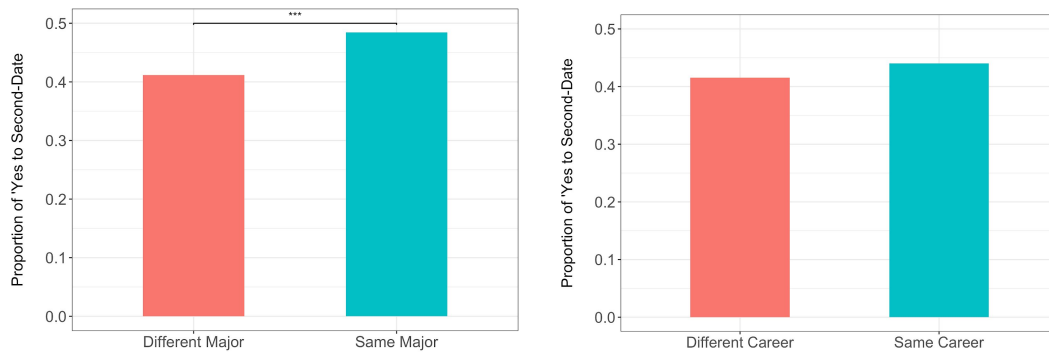
Findings

We divided our study into two parts: an analysis using the Kaggle data on how college major and occupation impact initial attraction, and an analysis using the census data on how these two factors impact marriage.

Claim #1: Having the same (or different) college major influences the likelihood that two students will go out on a second date, but having the same (or different) intended career does not.

Support for Claim #1: We first compared the proportion of individuals who had the same major and said yes to a second date with the proportion of individuals who had different majors and said yes to a second date. Using a z-test for differences in proportions, we found that there is a significant difference in the proportion of individuals who say yes to a second date between the

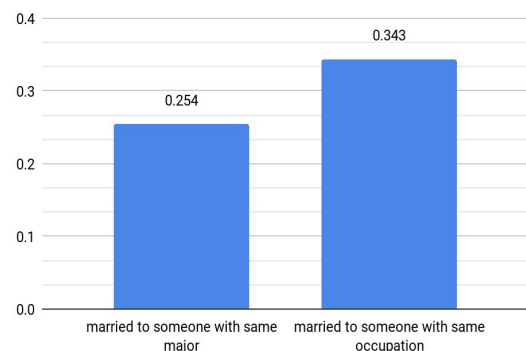
two populations ($p\text{-value} = 1.9\text{e-}5$). When performing the same test between the same career oriented and different career oriented populations, there was no significant difference in proportions ($p\text{-value} = 0.08$). Then, using logistic regression analysis to construct a model for determining the likelihood of an individual agreeing to a second date, having the same major was a significant and independent predictor ($OR = 1.33$, $p\text{-value} = 0.0001$) but having the same intended career was not ($OR = 1.04$, $p\text{-value} = 0.53$).



Claim #2: More married couples have the same occupation compared to the same college major

Support for Claim #2: We compared the proportion of married people with the same occupation as their spouse (p_o) and the proportion of married people with the same college major as their spouse (p_m). Using a z-test, we found that we can reject the null hypothesis ($p_o = p_m$), and accept the alternate hypothesis ($p_o \neq p_m$). The test statistic was 15.59 and the p-value was less than $1\text{e-}16$.

Marriage based on major and occupation in US census dataset



Claim #3 (Subclaim of Claim 2): When focusing on specific fields and occupations, not all of them have significant differences between the distribution of people who marry the same occupation and the distribution of people who marry the same college major.

Support for Claim #3:

We compared the proportion of people in specific majors and their corresponding occupations who married the same field or occupation to see if there was a significant difference. After doing chi-squared tests for independence to compare these proportions, there was a significant difference in the distribution of people who got married to the same field compared to the same occupation for lawyers ($p\text{-value} = 0.005$), doctors ($p\text{-value} = 2.3\text{e-}13$), and people who worked in business ($p\text{-value} = 6.3\text{e-}07$). There was no significant difference for engineers ($p\text{-value} = 0.59$), journalists and creative artists ($p\text{-value} = 0.27$), and psychologists ($p\text{-value} = 0.48$).

As an example, we have the proportion of students in the medical sciences who marry each other compared to the proportion of doctors who marry each other. Using the chi-squared test for independence we get a p-value of $2.3e-13$, meaning that we reject the null hypothesis and conclude that there is a significant difference in the distribution of doctors who marry a doctor and medical sciences students who end up marrying other medical sciences students. Looking at the graphs, we can see that the proportion of people who marry that same occupation is higher than the proportion of people that marry that same major.

