Understanding the effect of sex on income

Hlaseter, jx24

Hypothesis:

Gender pay gap has increasingly entered the national conversation, and is an important topic of discussion for gender equality. We want to explore the numbers behind the pay gap and analyze how the US census data reflect the national discussion. We hypothesize not only that a worker's sex relates with what they are paid, but this relationship is affected by factors like college degree, years of experience, and age.

Data:

We obtained the data from the U.S. census bearau, from the Public Use Microdata Series (PUMS). We combined the data from the 2018 PUMS surveys, crawling the census archives to combine data from all fifty states. From this dataset, we selected the relevant fields (sex, age, experience, wage, degree type, hours worked) and cleaned the data to construct our male and female datasets.

Findings:

Claim 1: Sex is correlated with wage, where males tend to earn more than females .

Support: The relationship between sex and income was a baseline for our project, and the data clearly supports it. We performed a t-test, representing sex as a dummy variable. We let H_o: B = 0; H_a: B < 0 and $\alpha = 0.5$. We easily rejected the null, confirming the existence of this relationship.

Variable:	B_i	t-stat	p-value
constant	60.754	261.113	p<.001
Sex (1 if M, 2 if F)	-14.9712	-102.667	p<.001

Claim 2: Sex continues to relate with wage when combined with factors like education, age, experience.

Support: We performed a hypothesis on a multiple regression of age, years experience, education level, and a dummy variable for sex. Again, We let H_o: $B_i = 0$ for all i; Ha: $B_i \neq 0$ and $\alpha = 0.5$. In the event of a failure to reject, we would perform a t-test to check for multicollinearity. However, we again clearly rejected the null, confirming the persistence of the effect of sex.

Variable:	B_i	t-stat	p-value
constant	-140.669	-72.931	p<.001
Experience	0.0956	13.767	p<.001
Age	0.3984	74.855	p<.001
Schooling	8.429	94.125	p<.001
Sex (1 if M, 2 if F)	-13.7775	-95.653	p<.001

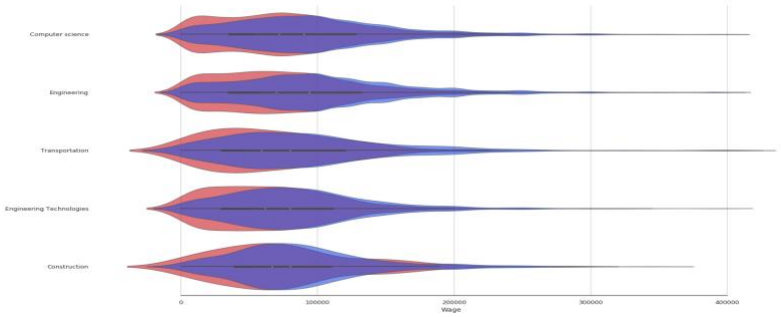
Claim 3: Age impacts the gender wage gap.

Support: We first verified via similar hypothesis test on a regression of sex and age on wages that neither sex and age both affect wage. Next, we performed a series of tests analyzing the difference between mean and median wages of men and women at various ages. We found that in general, the wage gap increases with age. However, we found that this increase is only significant for younger workers. We performed the following t-tests, with let H_o:

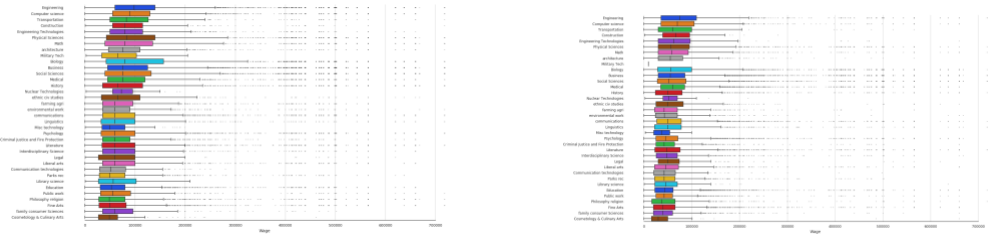
B = 0; H_a: B ≠0 and α = 0.5. As we can see, the trend is very significant for younger ages, but the relative growth stabilizes when men get older.

Age range:	Age Coefficient	t-stat	p-value
20-70	.4646	14.042	p<.001
20-45	.7893	27.548	p<.001
Sex (1 if M, 2 if F)	-13.7775	-95.653	P=0.88

Display: We completed a number of visualizations to better showcase the effect of gender, especially considering common degree choices.



Here we depict the difference in earnings by gender across five extremely common degree choices.



These figures (men on the left, women on the right) show the distribution of earnings for many common majors, and they show the contrast between the earnings of each sex.

Conclusion: We found that, even when incorporating a variety of additional variables, there clearly exists a significant disparity between the earnings of men and women.