0. Say you have data about race and income: how would you go about testing whether Hispanics make less than Asians? What do you need to argue causality?

1. A researcher is interested in explaining variation in religiosity (the frequency of church attendance) among Mexicans. She believes that education, age, and whether the survey respondent is catholic all affect the frequency of church attendance. She thinks that the lower educated, older, non-catholic respondents will have higher levels of religiosity. She is concerned that there is a difference in the structure of the regression between men and women. She runs three regressions: (1) only for women, (2) only for men, and (3) men and women combined

	Women Coef.	Men Coef.	Combined Coef.
Educ	-0.00	-0.01	-0.03
Age	0.01*	0.01*	0.01*
Cath	-0.24	0.39*	0.20*
_constant	3.13	2.36*	2.75*
N	777	769	1546
Residual Sum of Squares	787.266	1109.88	1982.13

* p<0.05

Coding:

Religiosity	1=Never Attends Church 2=Attends on holidays 3=Attends once a month 4=Attends once a week 5=Attends more than once a week
Educ	1=No Education 2=Completed 1-3 years of education 3=Completed 3-7 years of education 4=Completed high school 5=Attended University
Age	Age in 2000 (years)
Cath	1=Catholic 0=Non-Catholic

A. Are her hypotheses supported by the data? (25%)

B. What is the predicted value of religiosity (how frequently would we expect church attendance) for a catholic woman who attended university and was 25 years old in 2000? What is the predicted value for a man with the same characteristics? (25%)

C. ~~Perform a Chow Test to determine if the regressions for men and women are different.~~

D. What alternative(s) to running three regressions (as above) is/are available to a researcher who wants to compare the effects of education, age, and catholic among men to those effects among women? (write out the equation for the new model)(25%)

S05B2

Table II on the following page is drawn from Alan B. Krueger, "How Computers Have Changed the Wage Structure: Evidence from Microdata, 1984-1989," Quarterly Journal of Economics (1993). Kruger's argument is that the use of a computer at work became a more important determinant of wages during the late 1980s.

a) Do the results support Prof. Krueger's hypothesis? Why or why not?

b) Based on Model 1, what is the predicted hourly wage for someone who used a computer at work?

c) Experience is modeled in quadratic form. Based on Model 6, at what level of experience does the effect of experience become negative?

d) Discuss the interpretation of the "female" coefficient in Model 6.

e) Discuss an appropriate method and/or statistical test to determine whether there has been a change in the wage structure between 1984 and 1989.

OLS Regression Estimates of the Effect of Computer Use on Pay
(Dependent variable: ln (hourly wage))

Independent variable	October 1984			October 1989		
	(1)	(2)	(3)	(4)	(5)	(6)
Intercept	1.937	0.750	0.928	2.086	0.905	1.094
	(0.005)	(0.023)	(0.026)	(0.006)	(0.024)	(0.026)
Uses computer at work (1 = yes)	0.276	0.170	0.140	0.325	0.188	0.162
	(0.010)	(0.008)	(0.008)	(0.009)	(0.008)	(0.008)
Years of education	—	0.069	0.048	—	0.075	0.055
		(0.001)	(0.002)		(0.002)	(0.002)
Experience	—	0.027	0.025	—	0.027	0.025
		(0.001)	(0.001)		(0.001)	(0.001)
Experience-squared ÷ 100	—	−0.041	−0.040	—	−0.041	−0.040
		(0.002)	(0.002)		(0.002)	(0.002)
Black (1 = yes)	—	−0.098	−0.066	—	−0.121	−0.092
		(0.013)	(0.012)		(0.013)	(0.012)
Other race (1 = yes)	—	−0.105	−0.079	—	−0.029	−0.015
		(0.020)	(0.019)		(0.020)	(0.020)
Part-time (1 = yes)	—	−0.256	−0.216	—	−0.221	−0.183
		(0.010)	(0.010)		(0.010)	(0.010)
Lives in SMSA (1 = yes)	—	0.111	0.105	—	0.138	0.130
		(0.007)	(0.007)		(0.007)	(0.007)
Veteran (1 = yes)	—	0.038	0.041	—	0.025	0.031
		(0.011)	(0.011)		(0.012)	(0.011)
Female (1 = yes)	—	−0.162	−0.135	—	−0.172	−0.151
		(0.012)	(0.012)		(0.012)	(0.012)
Married (1 = yes)	—	0.156	0.129	—	0.159	0.143
		(0.011)	(0.011)		(0.011)	(0.011)
Married*Female	—	−0.168	−0.151	—	−0.141	−0.131
		(0.015)	(0.015)		(0.015)	(0.015)
Union member (1 = yes)	—	0.181	0.194	—	0.182	0.189
		(0.009)	(0.009)		(0.010)	(0.010)
8 Occupation dummies	No	No	Yes	No	No	Yes
R^2	0.051	0.446	0.491	0.082	0.451	0.486

Notes. Standard errors are shown in parentheses. Sample size is 13,335 for 1984 and 13,379 for 1989. Columns (2), (3), (5), and (6) also include three region dummy variables.

During the decennial Census, the U.S. Census Bureau hires many thousands of temporary workers who go door to door collecting Census surveys. They receive a variable amount of training. In an effort to learn more about the effectiveness of training, the Census Bureau sent follow-up workers to check the work of the temporary employees, and rated their performance on a scale of 1 to 100. Table 1 shows the results of several OLS regression models, including the hours in training, the years of prior work experience, and dummy variables (female = 1 if female, 0 otherwise; college graduate=1 if the worker had a college degree, 0 otherwise; and flextime=1 if the workers could choose their own hours of work, 0 otherwise). Table 2 shows descriptive statistics for the variables.

a) What does each model say about the impact of training on performance?

b) What is the predicted performance of a female worker, who is not a college graduate and not on flextime? Assume she has average values of the other variables.

c) According to Model 2, how much experience produces the optimal performance level?

d) Are the interaction terms in Model 3 jointly significant at the 0.05 level?

e) It may not be realistic to assume that training has the same effect no matter how much training a person receives. After a while, you have learned everything you need to know and further training may be pointless. Write down a regression model that will take this issue into account.

Table 1: OLS Regressions on Job Performance, Census Takers, 2000

Variable	*Model 1*	*Model 2*	*Model 3*
Constant	45.45	32.089	35.348
	(1.265)	(2.876)	(3.432)
Training (hours)	0.817	0.943	0.902
	(.061)	(.051)	(.056)
Experience (years)		2.291	1.938
		(.444)	(.525)
Experience Squared		-0.115	-0.102
		(.018)	(.021)
College Graduate (0/1)		0.924	0.629
		(.508)	(.548)
Flextime (0/1)		-3.515	-3.538
		(.767)	(.863)
Female (0/1)		7.181	-2.31
		(.774)	(4.194)
Female*Training			0.239
			(.136)
Female*Experience			0.417
			(.325)
Female*College Grad.			1.815
			(1.446)
Female*Flextime			-0.98
			(1.93)
R^2	0.281	0.529	0.536
N	454	454	454

Note: Standard errors in parentheses.

Table 2: Descriptive Statistics

Variable	*Obs*	*Mean*	*Std. Dev.*
Performance	454	62	8
Training	454	20	5
Experience	454	12	3
College Graduate	454	0	1
Female	454	0	0
Flextime	454	0	0

There are 3 presidential candidates, with follwong incomes and education: does education predict income?

candidate	educ	income
bernie	27	.1m
hilary	25	.7m
donald	20	50m