Lily Klau and Charlotte Taylor
Replication of Gender Differences in Major Choice: The Importance of Female Role Models
December 12, 2024

Task 1

Table A – Year 2015 Sample Data

Descriptive Statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
Took intermediate the semester following Principles	712	0.1	0.3	0	1
Took intermediate the year after Principles	712	0.195	0.397	0	1
Took another econ class after Principles	712	0.324	0.468	0	1
Number of econ classes taken after Principles (total)	712	1.817	3.614	0	15
Economics Major	712	0.17	0.376	0	1
Female Student	712	0.449	0.498	0	1
Student belongs to fraternity/sorority	533	0.585	0.493	0	1
Took econ in high school	530	0.609	0.488	0	1
Student is an athlete	532	0.056	0.231	0	1
Year 2016	712	0	0	0	0
Treatment class	712	0.499	0.5	0	1
In-state student	712	0.222	0.416	0	1
Freshman	712	0.86	0.348	0	1
American student	712	0.914	0.28	0	1
Cumulative GPA	712	3.284	0.524	1.2	4
Grade in Principles	712	2.921	0.917	0	4
High earning majors	527	2.863	1.301	1	5
Lower earning majors	170	2.218	1.298	1	4

Table A summarizes the sample data for the year 2015 describing the number of observations, mean, standard deviation, min and max for each variable. In the data, the most surprising variable was the proportion of students that took an economics class in high school. It is interesting that 60.9% of the students sampled took a prior class during high school and chose to take principles. This data shows that the correlation between those two variables might be statistically significant, and we would be interested in exploring this data. Furthermore, other interesting data in the table includes the average number of econ classes taken after principles, which was 1.817 classes. Also, the number of observations for the year 2015 is at most 712, which is greater than the year 2016 number of observations.

TASK 2:

Table B – Year 2016 Sample Data
Descriptive Statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
Took intermediate the semester following Principles	685	0.089	0.285	0	1
Took intermediate the year after Principles	685	0.194	0.396	0	1
Took another econ class after Principles	685	0.298	0.458	0	1
Number of econ classes taken after Principles (total)	685	1.807	3.584	0	14
Economics Major	685	0.18	0.384	0	1
Female Student	685	0.448	0.498	0	1
Student belongs to fraternity/sorority	480	0.531	0.5	0	1
Took econ in high school	481	0.615	0.487	0	1
Student is an athlete	479	0.071	0.257	0	1
Year 2015	685	1	0	1	1
Treatment class	685	0.495	0.5	0	1

In-state student	685	0.206	0.405	0	1
Freshman	685	0.885	0.32	0	1
American student	685	0.889	0.314	0	1
Cumulative GPA	685	3.331	.466	1.18	4
Grade in Principles	685	2.892	0.947	0	4
High earning majors	534	2.854	1.297	1	5
Lower earning majors	141	2.546	1.312	1	4

Table B summarizes the sample data for the year 2016 describing the number of observations, mean, standard deviation, min and max for each variable. In the data, it is unexpected that the proportion of students choosing to major in high earning rather than low earning. Additionally, the change in the number of females in the sample data was higher in 2015 than 2016. Since the highest number of observations in the sample of year 2016 was 685, the proportion of female students in Principles was 44.8% of 685; this total number of female students is much lower than the 44.9% of the 712 observed female students in 2016. Moreover, this is a significant decrease in the total number of students that were sampled between the two years.

TASK 3:

$\underline{BASE\ FORMULA:}$	$p: [\,\overline{p} \pm z_{\propto/2} * \sigma_{\overline{p}}\,] \rightarrow \sigma_{\overline{p}} = \sqrt{\frac{\overline{p}\,(1-\overline{p})}{n}}$

Confidence Interval for Population Proportion of Economics Majors in 2015 and 2016

	2015	**2016**
90% Confidence Level	$p: [\,0.1699438 \pm z_{0.10/2} * \sqrt{\frac{0.1699438 * (1-0.1699438)}{712}}\,]$ $p: [\,0.1699438 \pm 0.0231543352]$	$p: [\,0.179562 \pm z_{0.10/2} * \sqrt{\frac{0.179562 * (1-0.179562)}{685}}\,]$ $p: [\,0.179562 \pm 0.0241240777]$

	$p: [\ 0.1467894648, 0.1930981352]$	$p: [\ 0.1554379223, 0.2036860777]$
95% Confidence Level	$p: [\ 0.1699438 \pm z_{0.05/2} * \sqrt{\frac{0.1699438 * (1-0.1699438)}{712}}\ ]$ $p: [\ 0.1699438 \pm 0.0275881441]$ $p: [0.1423556559, 0.1975319441\ ]$	$p: [\ 0.179562 \pm z_{0.05/2} * \sqrt{\frac{0.179562 * (1-0.179562)}{685}}]$ $p: [\ 0.179562 \pm 0.0287435819\ ]$ $p: [\ 0.1508184181, 0.2083055819]$
99% Confidence Level	$p: [\ 0.1699438 \pm z_{0.01/2} * \sqrt{\frac{0.1699438 * (1-0.1699438)}{712}}\ ]$ $p: [\ 0.1699438 \pm 0.036244628\ ]$ $p: [\ 0.133699172, 0.206188428]$	$p: [\ 0.179562 \pm z_{0.01/2} * \sqrt{\frac{0.179562 * (1-0.179562)}{685}}]$ $p: [\ 0.179562 \pm 0.037762614]$ $p: [\ 0.141799386, 0.217324614]$

```
Variable |        Obs  Proportion    Std. err.       [90% conf. interval]
-------------+----------------------------------------------------------------
   econmajor |        712    .1699438    .0140756        .1467915    .1930961

Variable |        Obs  Proportion    Std. err.       [95% conf. interval]
-------------+----------------------------------------------------------------
   econmajor |        712    .1699438    .0140756        .1423562    .1975315

Variable |        Obs  Proportion    Std. err.       [99% conf. interval]
-------------+----------------------------------------------------------------
   econmajor |        712    .1699438    .0140756        .1336875    .2062001

Variable |        Obs  Proportion    Std. err.       [90% conf. interval]
-------------+----------------------------------------------------------------
   econmajor |        685     .179562    .0146651        .1554401     .203684

Variable |        Obs  Proportion    Std. err.       [95% conf. interval]
-------------+----------------------------------------------------------------
   econmajor |        685     .179562    .0146651         .150819    .2083051

Variable |        Obs  Proportion    Std. err.       [99% conf. interval]
-------------+----------------------------------------------------------------
   econmajor |        685     .179562    .0146651        .1417873    .2173368
```

Based on the results of the confidence intervals for population proportion of economics majors in the year 2015 and 2016, we can say that exactly 21% of the student population was an economics major in 2016 at 99% confidence level (CL). At both 90% and 95% confidence levels we cannot make that claim; we must reject because 21% falls beyond the max interval of 20.37% at 90% CL and 20.83% at 95% CL.

Task 4:

Step 1:
$H_o: \mu \geq 2.95$
$H_a: \mu < 2.95$ ® lower tailed test

Step 2:
$\propto = 0.1; 0.05; 0.01$

Step 3:
$Degrees\ of\ freedom = n - 1 = 1397 - 1 = 1396$
$t = \frac{\bar{x} - \mu_0}{s_{\bar{x}}} = \frac{2.906743 - 2.95}{0.9316177/\sqrt{1397}} = -1.74$

Step 4a:
$P - Value = P\ (t \leq\ -1.74) =\ \ 0.025\ \leq P\ (t \geq\ 1.74)\ \leq 0.05$
$0.1 \geq 0.05 \geq P - Value \geq 0.01$
Can reject null at 90% and 95% confidence levels, cannot reject null at 99%.

Step 4b:
$Critical\ Value =\ t_{\propto}$
$t_{0.1}$= 1.282
$t_{0.05}$ = 1.645
$t_{0.01}$ = 2.326
$t_{0.1} \leq t_{0.05}\ \leq t \leq t_{0.01}$
Can reject null at 90% and 95% confidence levels, cannot reject null at 99%.

```
One-sample t test
------------------------------------------------------------------------------
Variable |     Obs        Mean    Std. err.   Std. dev.   [95% conf. interval]
---------+--------------------------------------------------------------------
gradeP~s |   1,397    2.906743    .0249253    .9316177    2.857848    2.955638
------------------------------------------------------------------------------
    mean = mean(gradePrinciples)                                  t =  -1.7355
H0: mean = 2.95                                  Degrees of freedom =     1396

   Ha: mean < 2.95             Ha: mean != 2.95              Ha: mean > 2.95
 Pr(T < t) = 0.0414         Pr(|T| > |t|) = 0.0829          Pr(T > t) = 0.9586
```

The hypothesis that the average grade in Economics principles classes (intros) is at least 2.95 (between B- and B) can be rejected at 90% and 95% confidence levels, however, cannot be rejected at 99% confidence level. When we estimated this in STATA the result of the lower test gave a p-value of 0.0414, which is less than 0.1 and 0.05. These results mean we can reject at both of those confidence levels (90% and 95% respectively), but not at 0.01 (99%).

Task 5:

Step 1:
$H_o: p \leq 0.6$
$H_a: p > 0.6$ ® upper tailed test

Step 2:
$\propto = 0.1; 0.05; 0.01$

Step 3:
n = 1011
x = 619
$\overline{p} = \frac{x}{n} = \frac{619}{1011} = 0.6122650841$
$\sigma_{\overline{p}} = \sqrt{\frac{p_0 * (1-p_0)}{n}} = \sqrt{\frac{0.6 * (1-0.6)}{1011}} = 0.0154074243$
$z = \frac{\overline{p}-p_0}{\sigma_{\overline{p}}} = \frac{0.6122650841-0.6}{0.0154074243} = 0.7961 \approx 0.80$
*Standard error given when we run this is not correct it calculates it using p bar and not p null, however, it uses the correct standard error when calculating the z-value

Step 4a:
$P - Value = P\ (z \geq\ 0.80) = \ P\ (z \leq -\ 0.80) = P\ (z \geq\ -0.80) = 0.2119$
$P - Value \geq 0.1 \geq 0.05 \geq 0.01$
Cannot reject null at any confidence level.

Step 4b:
$Critical\ Value = z_{\propto}$
$z_{0.1}$= 1.285
$z_{0.05}$ = 1.645
$z_{0.01}$ = 2.325
$z\ \leq z_{0.1} \leq z_{0.05}\ \leq z_{0.01}$
Cannot reject null at any confidence level.

```
One-sample test of proportion                  Number of obs      =      1011
```

```
------------------------------------------------------------------------------
    Variable |       Mean   Std. err.                     [95% conf. interval]
-------------+----------------------------------------------------------------
     econ_hs |   .6122651   .0153236                      .5822313    .6422988
------------------------------------------------------------------------------
    p = proportion(econ_hs)                                       z =   0.7961
H0: p = 0.6

     Ha: p < 0.6                 Ha: p != 0.6                 Ha: p > 0.6
 Pr(Z < z) = 0.7870         Pr(|Z| > |z|) = 0.4260          Pr(Z > z) = 0.2130
```

The hypothesis that at most 60% of students in Principles took Economics course in high school cannot be rejected at any confidence level. When we estimated this in STATA the result of the upper test gave a p-value of 0.2130, which is greater than the p-values of 0.1, 0.05, and 0.01 (confidence levels 90%, 95%, and 99% respectively). Thus, in summary, using the results of STATA the null hypothesis cannot be rejected at any confidence level.

Task 6:

Quantitative variable: Cumulative GPA in 2016

Step 1:

$H_o: \mu_1 - \mu_2 = 0$

$H_a: \mu_1 - \mu_2 \neq 0$ → two tailed test

Step 2:

$\propto = 0.1; 0.05; 0.01$

Step 3:

$n_1 = 177$

$n_2 = 130$

$\overline{x}_1 = 3.479136$

$\overline{x}_2 = 3.389015$

$S_{\overline{x_1}-\overline{x_2}} = \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}} = \sqrt{\frac{(0.3586944)^2}{177} + \frac{(0.4621972)^2}{130}} = 0.0486845044$

$t = \frac{(\overline{x}_1 - \overline{x}_2) - 0}{S_{\overline{x_1}-\overline{x_2}}} = \frac{(3.479136 - 3.389015) - 0}{0.0486845044} = 1.85112288 \approx 1.85$

Step 4a:

$P - Value = 2 * P\ (t \geq\ 1.85) = 2 * [0.025 \leq P\ (t \geq\ 1.85) \leq 0.05] =\ 0.05 \leq 2 * P\ (t \geq\ 1.85) \leq 0.10$

$0.1 \geq P - Value \geq 0.05 \geq 0.01$

$$df = \frac{\left(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}\right)^2}{\frac{1}{n_1 - 1}\left(\frac{S_1^2}{n_1}\right)^2 + \frac{1}{n_2 - 1}\left(\frac{S_2^2}{n_2}\right)^2} = \frac{\left(\frac{(0.3586944)^2}{177} + \frac{(0.4621972)^2}{130}\right)^2}{\frac{1}{177 - 1}\left(\frac{(0.3586944)^2}{177}\right)^2 + \frac{1}{130 - 1}\left(\frac{(0.4621972)^2}{130}\right)^2} = 122$$

Reject null at 90% confidence level; cannot reject null at 95% or 99% confidence level.

Step 4b:

$Critical\ Value = t_{\propto/2}$

$t_{0.1/2}$= 1.645

$t_{0.05/2}$ = 1.96

$t_{0.01/2}$ = 2.575

$t_{0.1/2} \leq t \leq t_{0.05/2} \leq t_{0.01/2}$

Reject null at 90% confidence level; cannot reject null at 95% or 99% confidence level.

In testing one quantitative variable, the hypothesis that the difference between mean cumulative GPA of females in 2016 is equal to 0 can be rejected at 90% confidence level but cannot be rejected at either 95% or 99% confidence levels. This is supported by both p-value and critical value tests.

Table C -- All Variables by Treatment Class and Year

	Control classes 2015 (untreated)	**Treatment classes 2015 (untreated)**	**p-value diff.**	**Control classes 2016 (untreated)**	**Treatment classes 2016 (treated)**	**p-value diff.**
American student	92.95	94.51	(0.56)	92.09	73.23	(0.00)
In-state student	22.44	24.39	(0.68)	20.34	21.54	(0.80)
Freshman	83.97	85.37	(0.73)	83.05	95.38	(0.00)
Cumulative GPA	3.44	3.41	(0.52)	3.48	3.39	(0.06)
Belongs to sorority	65.08	65.00	(0.99)	67.46	58.72	(0.17)

Took econ in high school	53.97	61.87	(0.19)	59.06	55.96	(0.63)
Athlete	7.14	5.71	(0.63)	6.40	10.09	(0.30)
Took Intermediate Micro within year	14.74	10.98	(0.31)	10.73	18.46	(0.05)
Took another econ class	23.72	18.90	(0.29)	17.51	28.46	(0.02)
Number of further econ classes taken	1.03	0.90	(0.66)	0.85	1.42	(0.08)
Majored in economics	10.26	7.93	(0.46)	7.91	15.38	(0.04)

Task 7:

$$weighted\ mean = \bar{x} = \frac{2.708975\ (244) + 2.834581\ (155) + 3.06295\ (261) + 3.263366\ (309) + 3.181087\ (92)}{1061} = 3.016793731$$

n_T: 1061
k: 5

Step 1:
$H_0: \mu_{Econ} = \mu_{STEM} = \mu_{Business} = \mu_{Finance} = \mu_{Marketing}$
$H_a: not\ all\ population\ means\ are\ equal$

Step 2:
$\propto = 0.1; 0.05; 0.01$

Step 3:

$$F = \frac{MSTR}{MSE} = \frac{12.52291957}{0.7338165976} = 17.07$$

$$MSTR = \frac{\sum_{j=1}^{k} nj * \left(\bar{x}_j - \bar{x}\right)^2}{k-1}$$

$$= \frac{244(2.708975 - 3.016793731)^2 + 155(2.834581 - 3.016793731)^2 + 261(3.06295 - 3.016793731)^2 + 309(3.263366 - 3.016793731)^2 + 92(3.181087 - 3.016793731)^2}{5-1}$$

$= 12.52291957$

$$MSE = \frac{\sum_{j=1}^{k}(nj-1)*S_j^2}{n_T - k}$$

$$= \frac{(244-1).9384784^2 + (155-1)\,1.033806^2 + (261-1)\,.8703188^2 + (309-1)\,.7012255^2 + (92-1)\,.7256274^2}{1061-5}$$

$$= \frac{774.9103271}{1056} = 0.7338165976$$

Step 4a:

$P - value\ = 0.1 \geq 0.05 \geq 0.01 \geq P(F \geq 17.07)$

$Numerator\ df: k - 1 = 5 - 1 = 4$

$Denominator\ df: n_T - k =$ 1061 – 5 = 1056

Reject null at all confidence levels.

Step 4b:

$CV = F_{\propto}$

$F_{0.1} = 1.95$

$F_{0.05} = 2.38$

$F_{0.01} = 3.34$

$1.95 \leq 2.38 \leq 3.34 \leq 17.07$

Reject null at all confidence levels.

Using the sample data provided, the hypothesis being tested is whether there are any differences in grades earned in Principles by different major categories within "high-earning" majors. Based on the data the null hypothesis can be rejected at all confidence levels. This conclusion is supported by the p-value and critical value tests. The p-value of the test statistic F when F=17.07 is less than 0.1, 0.05, and 0.01 (confidence levels 90%, 95%, and 99% respectively). Thus, the rejection of the null hypothesis is supported at all confidence levels.

Table D – Analysis of Variance (ANOVA) Table

Source of Variation	**Sum of Squares**	**Degrees of Freedom**	**Mean of squares**	**F**	**Prob > F**
Between groups	50.0915948	4	12.5228987		

Within Groups	774.910316	1056	0.733816587	17.07	0.0000
Total	825.001911	1060	0.778303689		

Task 8:

Step 1:
$H_o: p_1 - p_2 = 0$
$H_a: p_1 - p_2 \neq 0$ → two tailed test

Step 2:
$\propto = 0.1; 0.05; 0.01$

Step 3:
$n_1 = 320$
$n_2 = 307$
$x_1 = 29$
$x_2 = 34$
$\overline{p}_1 = \frac{x}{n} = \frac{29}{320} = 0.090625$
$\overline{p}_2 = \frac{x}{n} = \frac{34}{307} = 0.1107491857$
$\bar{p} = \frac{n_1\bar{p}_1 + n_2\bar{p}_2}{n_1 + n_2} = \frac{320(0.090625) + 307(0.1107491857)}{320 + 307} = 0.1004784689$
$\overline{p}_1 - \overline{p}_2 = 0.090625 - 0.1107491857 = -0.0201241857$
$\sigma_{\overline{p_1} - \overline{p_2}} = \sqrt{\overline{p} * (1 - \overline{p}) * (\frac{1}{n_1} + \frac{1}{n_2})} =$
$\sqrt{0.1004784689 * (1 - 0.1004784689) * (\frac{1}{320} + \frac{1}{307})} = 0.0240177256$
$z = \frac{(\overline{p}_1 - \overline{p}_2) - 0}{\sigma_{\overline{p_1} - \overline{p_2}}} = \frac{(0.090625 - 0.1107491857) - 0}{0.0240177256} = -0.8378889007 \approx -0.84$

Step 4a:
$P - Value = 2 * P\ (z \leq -0.84) = 2 * 0.2005 = 0.4010 \leq 0.1 \leq 0.05 \leq 0.10$
$P - Value \geq 0.1 \geq 0.05 \geq 0.01$

Cannot reject null at any confidence level.

Step 4b:
$Critical\ Value = z_{\propto/2}$
$z_{0.1/2} = 1.645$
$z_{0.05/2} = 1.96$
$z_{0.01/2} = 2.575$
$z \leq z_{0.1/2} \leq z_{0.05/2} \leq z_{0.01/2}$

Cannot reject null at any confidence level.

The hypothesis tests whether proportion of female students majoring in Economics increased from year 2015 to year 2016. Based on the p-value and critical value tests, we cannot reject this null hypothesis at all confidence levels. The p-value of the test statistic z when z≈ -0.84 is greater than 0.1, 0.05, and 0.01 (confidence levels 90%, 95%, and 99% respectively). Thus, the rejection of the null hypothesis cannot be supported at all confidence levels.

```
Two-sample test of proportions                     0: Number of obs =      320
                                                   1: Number of obs =      307
------------------------------------------------------------------------------
       Group |       Mean   Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
           0 |    .090625    .016048                      .0591715    .1220785
           1 |   .1107492   .0179107                      .0756448    .1458536
-------------+----------------------------------------------------------------
        diff |  -.0201242   .0240485                     -.0672584    .0270101
             |  under H0:   .0240177    -0.84   0.402
------------------------------------------------------------------------------
        diff = prop(0) - prop(1)                                  z =  -0.8379
    H0: diff = 0

Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(Z < z) = 0.2010         Pr(|Z| > |z|) = 0.4021          Pr(Z > z) = 0.7990
```

Task 9:

Step 1:
H_o: *Choice of major is independent of being an American*
H_a: *Choice of major is not independent of being an American*

Step 2:
$\propto = 0.1; 0.05; 0.01$

Step 3:
$P(yes) = \frac{total\ yes}{total} = \frac{1242}{1372} = 0.9052478134$
$P(no) = \frac{total\ no}{total} = \frac{130}{1372} = 0.0947521866$

Table E - Chi-square Test (χ^2) for Independence of Choice of Major and Being an American

Yes/No (i)	**Populations (j)**	**Observed frequency (fij)**	**Expected frequency (eij)**	$(fij - eij)^2$	$\frac{(fij - eij)^2}{eij}$
Yes	Econ major	204	$244 * 0.9052478134 = 220.8804665$	$(204 - 220.8804665)^2 = 284.9501493$	$\frac{284.9501493}{220.8804665} = 1.290064956$

Yes	STEM major	133	$155 * 0.9052478134 = 140.3134111$	$(133 - 140.3134111)^2 = 53.48598192$	$\frac{53.48598192}{140.3134111} = 0.3811893781$
Yes	Business major	232	$261 * 0.9052478134 = 236.2696793$	$(232 - 236.269673)^2 = 18.23016132$	$\frac{18.23016132}{236.2696793} = 0.0771582768$
Yes	Finance major	287	$309 * 0.9052478134 = 279.7215743$	$(287 - 279.7215743)^2 = 52.97548067$	$\frac{52.97548067}{279.7215743} = 0.1893864669$
Yes	Marketing major	90	$92 * 0.9052478134 = 83.28279883$	$(90 - 83.28279883)^2 = 45.12079156$	$\frac{45.12079156}{83.28279883} = 0.5417780405$
Yes	Social Science major	127	$132 * 0.9052478134 = 119.4927114$	$(127 - 119.4927114)^2 = 56.35938212$	$\frac{56.35938212}{119.4927114} = 0.471655396$
Yes	Arts major	31	$32 * 0.9052478134 = 28.96793003$	$(31 - 28.96793003)^2 = 4.129308363$	$\frac{4.129308363}{28.96793003} = 0.1425475814$
Yes	Communications major	47	$48 * 0.9052478134 = 43.45189504$	$(47 - 43.45189504)^2 = 12.58904881$	$\frac{12.58904881}{43.45189504} = 0.2897238152$
Yes	Humanities major	91	$99 * 0.9052478134 = 89.61953353$	$(91 - 89.61953353)^2 = 1.905687675$	$\frac{1.905687675}{89.61953353} = 0.0212641999$
No	Econ major	40	$244 * 0.0947521866 = 23.11953353$	$(40 - 23.11953353)^2 = 284.9051482$	$\frac{284.9051482}{23.11953353} = 12.32508207$
No	STEM major	22	$155 * 0.0947521866 = 14.68658892$	$(22 - 14.68658892)^2 = 53.488598163$	$\frac{53.488598163}{14.68658892} = 3.641824655$
No	Business major	29	$261 * 0.0947521866 = 24.7303207$	$(29 - 24.7303207)^2 = 18.23016132 =$	$\frac{18.23016132}{24.7303207} = 0.7371583062$
No	Finance major	22	$309 * 0.0947521866 = 29.27842566$	$(22 - 29.27842566)^2 = 52.97548009$	$\frac{52.97548009}{29.27842566} = 1.809369148$
No	Marketing major	2	$92 * 0.0947521866 = 8.717201167$	$(2 - 8.717201167)^2 = 45.126079152$	$\frac{45.126079152}{8.717201167} = 5.176064043$
No	Social Science major	5	$132 * 0.0947521866 = 12.50728863$	$(5 - 12.50728863)^2 = 56.35938257$	$\frac{56.35938257}{12.50728863} = 4.506123128$
No	Arts major	1	$32 * 0.0947521866 = 3.032069971$	$(1 - 3.032069971)^2 = 4.12930837$	$\frac{4.12930837}{3.032069971} = 1.361877663$
No	Communications major	1	$48 * 0.0947521866 = 4.548104957$	$(1 - 4.548104957)^2 = 12.58904879$	$\frac{12.58904879}{4.548104957} = 2.767976753$

No	Humanities major	8	$99 * 0.0947521866 = 9.380466473$	$(8 - 9.380466473)^2 = 1.905687683$	$\frac{1.905687683}{9.380466473}$=0.2031548952

$$\chi^2 = \sum i \sum j \frac{(fij - eij)^2}{eij}$$

$$\chi^2 = 35.9334$$

Step 4a:
$df = (i-1) * (j-1) = (2-1) * (9-1) = 1 * 8 = 8$
$P - Value = P\ (\chi^2 \geq\ 35.9334) = \ P(\chi^2 \geq\ 1.74)\ \leq 0.005$
Reject null at all confidence levels.

Step 4b:
$\chi^2{}_{0.1} = 13.362$
$\chi^2{}_{0.05} = 15.507$
$\chi^2{}_{0.01} = 20.090$
$\chi^2 \geq \chi^2{}_{0.01} \geq \chi^2{}_{0.05} \geq \chi^2{}_{0.1}$
Reject null at all confidence levels.

Tabulation of task9_major american

1 econmajor 2 STEM 3 Business 4 Finance 5 Marketing 6 social science 7 Arts 8 Co	American student 0	1	Total
1	40	204	244
2	22	133	155
3	29	232	261
4	22	287	309
5	2	90	92
6	5	127	132
7	1	31	32
8	1	47	48
9	8	91	99
Total	130	1242	1372

Pearson Chi2 = 35.93   Prob = 0.0000

Based on the sample provided, the tested hypothesis was whether the choice of major (Econ, STEM, Business, Finance, Marketing, Social Science, Arts, Communication and Humanities) is independent of being an American. This null hypothesis cannot be rejected at any confidence level. This is supported by the chi-square tests (χ^2 tests), particularly in the context independence.

Task 10:

Table F - Treatment Effects on Intermediate Outcomes

	Took Micro within year	Took Micro within year	Took another econ class	Took another econ class
Treatment class × 2016	0.1149506 (0.036)	0.0781419 (0.221)	0.1576292 (0.017)	0.0528822 (0.475)
Year 2016	-0.0400913 (0.284)	-0.0246214 (0.569)	-0.0620382 (0.170)	-0.0085011 (0.868)
Treatment class (in 2015)	-0.0376798 (0.323)	-0.0378552 (0.369)	-0.0481551 (0.296)	-0.0234745 (0.638)
Constant	0.1474359 (0.000)	0.1920393 (0.230)	0.2371795 (0.000)	-0.0234745 (0.002)
Controls	No	Yes	No	Yes
Observations	627	499	627	499

Table G - Treatment Effects on Final Outcomes

	Number of econ classes taken	Number of econ classes taken	Major in economics	Major in economics
Treatment class × 2016	0.6915768 (0.109)	0.0472285 (0.917)	0.0980459 (0.043)	0.0177759 (0.735)
Year 2016	-0.1725337 (0.558)	0.0091182 (0.977)	-0.0234681 (0.477)	-0.0061458 (0.866)
Treatment class (in 2015)	-0.1292996 (0.666)	-0.1740662 (0.567)	-0.0232958 (0.488)	-0.0245735 (0.487)
Constant	1.025641 (0.000)	4.750939 (0.000)	0.1025641 (0.000)	0.396417 (0.003)
Controls	No	Yes	No	Yes
Observations	627	499	627	499

Task 11:

Table H - Treatment Effects on Other High-Earning Majors

	Major STEM	Major finance	Major business	Major marketing
Treatment class × 2016	0.0035659 (0.949)	0.0526226 (0.434)	0.0568617 (0.441)	-0.0091784 (0.883)
Year 2016	0.0280636 (0.469)	-0.048784 (0.294)	-0.0266869 (0.601)	0.0059678 (0.890)
Treatment class (in 2015)	0.0002356 (0.995)	-0.0395528 (0.383)	-0.0267902 (0.590)	-0.0285957 (0.498)
Constant	0.0847695 (0.555)	-0.3642913 (0.035)	-0.3944023 (0.037)	-0.2627721 (0.101)
Controls	Yes	Yes	Yes	Yes
Observations	499	499	499	499

Table I - Treatment Effects on Low-Earning Majors

	Major social sciences	Major arts	Major communication	Major humanities
Treatment class × 2016	-0.042313 (0.430)	0.0459001 (0.143)	-0.002004 (0.964)	-0.0881199 (0.080)
Year 2016	-0.0023723 (0.949)	0.0073853 (0.733)	-0.0190966 (0.535)	0.0526361 (0.130)
Treatment class (in 2015)	-0.0018961 (0.958)	-0.0114572 (0.587)	-0.0053421 (0.859)	0.0458014 (0.176)
Constant	0.8147332 (0.000)	-0.0299544 (0.708)	0.3692348 (0.001)	0.3234909 (0.012)
Controls	Yes	Yes	Yes	Yes
Observations	499	499	499	499