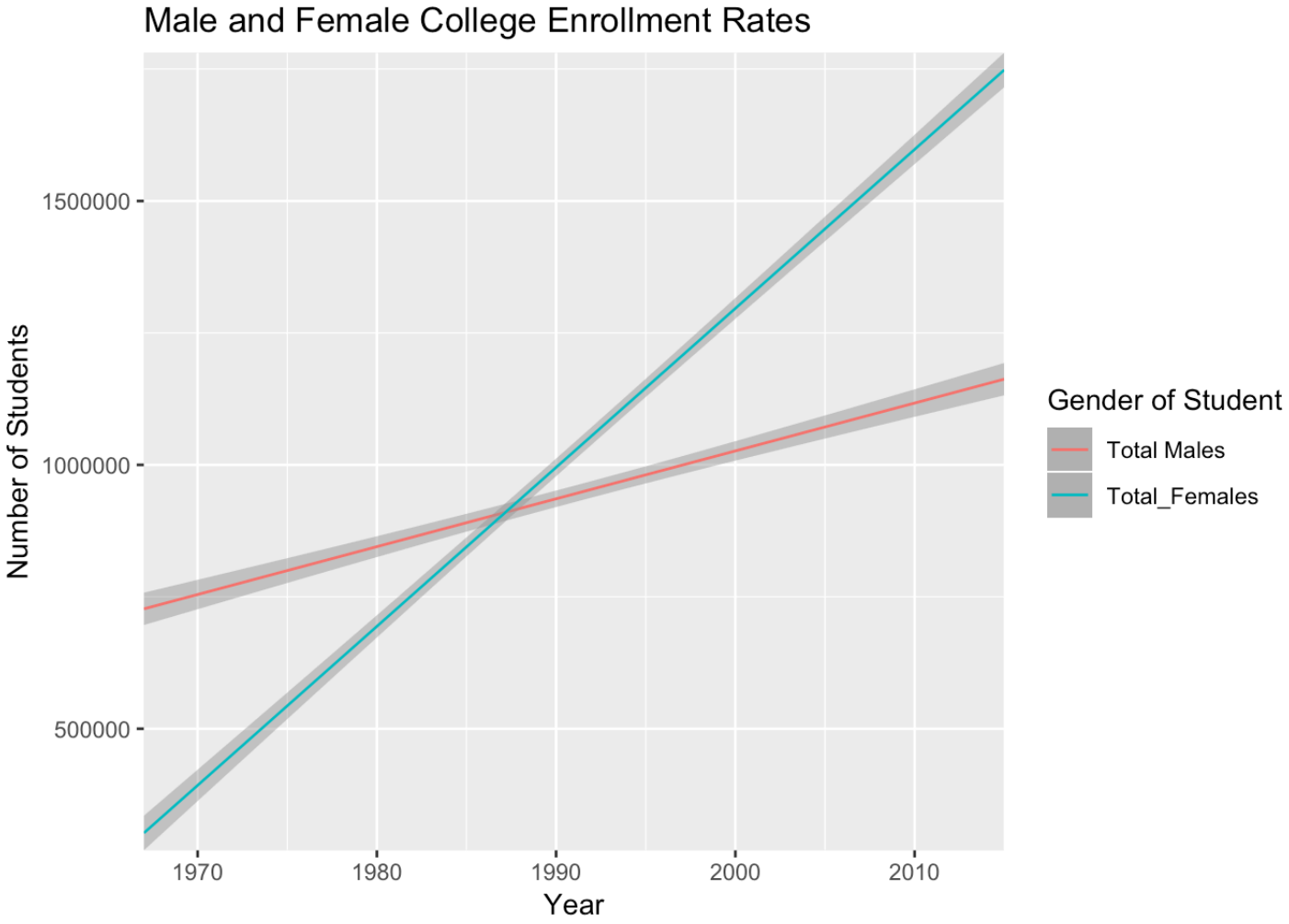
Assignment 5

Alex Ivina

November 26, 2018

1. Male and female graduate enrollment (1967 - 2015). Compare trends in total graduate enrollment for males and females (including full-time/part-time and private/public universities) in the United States from 1967 - 2015. Describe your results statistically, graphically and in text.



Written Results: Total_Males- Year significantly predicts college enrollment rates for males (b = 9069, t(47) = 16.61, p < 0.001) with a strong positive correlation between the two (Pearson's r = 0.92). The overall model (Total Male enrollment = -17,112,153 + 9069 (Year) explains a significant amount of variance in enrollment rate (F(1,47) = 476, p < 0.001, R2 = 0.8545)."

male_enrollment = -17112153 + 9069(Year)

Total_Females- Year significantly predicts college enrollment rates for females (b = 30126, t(47) = 51.66, p < 0.001) with a strong positive correlation between the two (Pearson's r = 0.99). The overall model (Total Female enrollment = -58955502 + 30126 (Year) explains a significant amount of variance in enrollment rate (F(1,47) = 2669, p < 0.001, R2 = 0.9827).

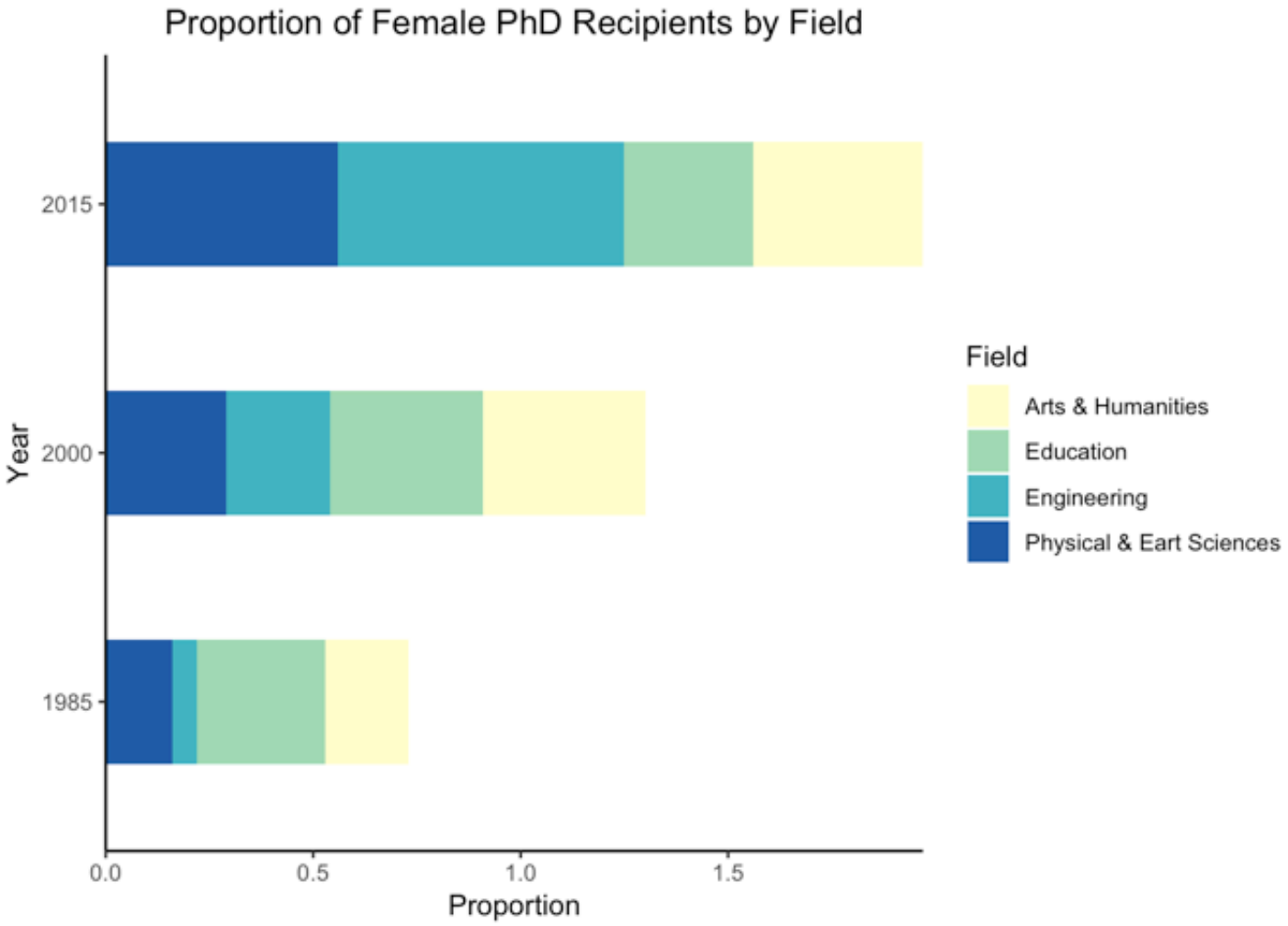
female_enrollment = -0.0000005896 + 30130(Year)

2. Shifts in female PhD recipients by field (1985, 2000, and 2015). Describe if and how there was a shift in PhDs awarded to females in four fields (Physical and Earth Sciences, Engineering, Education, and Humanities & Arts) in 1985, 2000, and 2015. Describe your results statistically, in a graph or table, and in text. Note: There are several ways that you can interpret this question. You are invited to decide which you think is/are most interesting. Just be really clear about what you are asking/answering in your report.

```
##
##  Pearson's Chi-squared test
##
## data:  Phds_new
## X-squared = 2898, df = 14, p-value < 2.2e-16
```

Changes in field of study

	1985	2000	2015
Female_Physical_science	0.16	0.29	0.56
Female_Engineering	0.06	0.25	0.69
Female_Education	0.31	0.37	0.31
Female_Humanities	0.20	0.39	0.41



The proportion of female phD recipients increased in three fields (Physical & Earth Sciences, Engineering, and Arts & Humanities) in 1985, 2000, 2015.

3. Male and female salaries for starting postdoctoral and other employment positions (2015). Compare median salaries for male and female doctorate recipients in 2015. Answer these two questions: Does median salary differ significantly between male and female starting postdoc positions? Does median salary differ significantly between male and female PhD recipients in non-postdoc employment positions?

=======

Communicating Statistical Results: Non-Parametric Wilcomxon Signed Rank anaylsis for data collected from 15 fields of study in the year 2015 revealed that the median salaries (in dollars) differed significantly for male and female PhD recipients in non-postdoc employment positions (V= 101, p=0.003). In all fields except Physics & Astonomy, men recieved either equal or moderately higher salaries(cliffs delta=0.213).

However, a separate non-parametric wimcoxon signed rank test for PhD recipients starting posdoc positions revealed that there was *not* a signficant difference between salaries for males and females (V=19.5, p=.888). In several fields of study, males and females recieved the exact same salary. Females recieved a slightly higher median salary in some fields, while males were paid slightly more in others (cliffs delta=.04).

4. Exploring academic salaries for professors in U.S. colleges. Explore relationships between variables in the 'Faculty salary data (2008 - 2009 survey)' dataset. Develop a model describing faculty salary based on data for faculty sex, rank, years in current position, field, and number of years since doctoral degree was earned. You should make decisions regarding which variables should remain in your final model. Describe the results qualitatively and quantitatively (i.e., don't just report the statistical results of the model – make sure you describe interesting findings in text). You can also discuss any concerns that you have with the model(s) you present, if any

	Dependent variable:
	salary
faculty_rankAssocProf	-32,928.400***
	(3,544.403)
faculty_rankAsstProf	-46,032.550***
	(4,240.120)
years_since_phd	61.011
	(127.010)
sexMale	4,349.366
	(3,875.393)
disciplineTheoretical	-13,937.470***
	(2,346.534)
Constant	127,854.300***
	(5,084.988)
Observations	397
R^2	0.447
Adjusted R^2	0.440
Residual Std. Error	22,663.240 (df = 391)
F Statistic	63.266*** (df = 5; 391)
Note:	*p<0.1;* ***p<0.05;*** p<0.01

Write Up:

Formula: salary = 127854.34 -32928.4(Associate_Professor) - 46032.55(Assitant_Professor) + 61.01(years_since_phd) + 4349.37(sexMale) - 13937.47(disciplineTheoretical)

The formula for faculty salary will predict salary as a function of a professors rank (here we are using professor as the refrence level and comparing against associate professor and assitant professor), years since someone has gained their PHD, their sex (here female is the refrence level and male is included in the graph) and lastly discipline (we only look at if the diswcipline is theoretical or applied and here we use applied as the refrence level). What we found was that professors were paid the most and assitant profressors the least by rank and that years every year after your PHD you earned slightly more. It was also apparent that males earned higher wages than femlaes and that those in theoretical disciplines earned less than those in technical disciplines. Over all these variables accounted for 44% (R^2 = 0.4401) of changes in salary. Looking at the variables observed the rank of a professor and the deispline showed significant results (p-value<0.001) where as sex and years since one earned their PHD did not yeild significant results (p-value>0.001).

```
##
## Call:
## lm(formula = salary ~ years_since_phd + years_employed + sex +
##     discipline, data = faculty_salary1)
##
## Coefficients:
##           (Intercept)        years_since_phd         years_employed
##               87651.2                 1804.1                 -770.1
##               sexMale  disciplineTheoretical
##                7545.3               -16325.4
```

```
##
## Call:
## lm(formula = salary ~ years_since_phd + years_employed + sex +
##     discipline, data = faculty_salary1)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -75974 -17094  -3799  16073  97055
##
## Coefficients:
##                      Estimate Std. Error t value Pr(>|t|)
## (Intercept)           87651.2     4666.2  18.784  < 2e-16 ***
## years_since_phd        1804.1      248.9   7.249 2.25e-12 ***
## years_employed         -770.1      244.1  -3.155  0.00173 **
## sexMale                7545.3     4462.6   1.691  0.09167 .
## disciplineTheoretical -16325.4     2708.9  -6.027 3.87e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 26130 on 392 degrees of freedom
## Multiple R-squared:  0.2634, Adjusted R-squared:  0.2558
## F-statistic: 35.04 on 4 and 392 DF,  p-value: < 2.2e-16
```

```
##
## Call:
## lm(formula = salary ~ years_since_phd + years_employed + sex +
##     discipline + years_since_phd * years_employed, data = faculty_salary1)
##
## Coefficients:
##                     (Intercept)                  years_since_phd
##                        67764.10                          2391.08
##                 years_employed                          sexMale
##                         1473.92                          8188.37
##          disciplineTheoretical  years_since_phd:years_employed
##                       -14999.33                           -62.23
```

```
##
## Call:
## lm(formula = salary ~ years_since_phd + years_employed + sex +
##     discipline + years_since_phd * years_employed, data = faculty_salary1)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -55902 -14672  -1759  12687  98576
##
## Coefficients:
##                                  Estimate Std. Error t value Pr(>|t|)
## (Intercept)                     67764.105   4849.353  13.974  < 2e-16 ***
## years_since_phd                  2391.078    237.860  10.052  < 2e-16 ***
## years_employed                   1473.923    341.469   4.316 2.01e-05 ***
## sexMale                          8188.367   4090.693   2.002    0.046 *
## disciplineTheoretical          -14999.332   2487.418  -6.030 3.80e-09 ***
## years_since_phd:years_employed    -62.226      7.153  -8.699  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 23950 on 391 degrees of freedom
## Multiple R-squared:  0.3828, Adjusted R-squared:  0.3749
## F-statistic:  48.5 on 5 and 391 DF,  p-value: < 2.2e-16
```

```
##                 years_since_phd                 years_employed
##                        6.488255                      13.619794
##                             sex                     discipline
##                        1.026208                       1.062488
## years_since_phd:years_employed
##                       13.062127
```

```
## [1] 9140.98
```

	Dependent variable:
	salary
years_since_phd	2,391.078***
	(237.860)
years_employed	1,473.923***
	(341.469)
sexMale	8,188.367**
	(4,090.693)
disciplineTheoretical	-14,999.330***
	(2,487.418)
years_since_phd:years_employed	-62.226***
	(7.153)
Constant	67,764.100***
	(4,849.353)
Observations	397
R^2	0.383
Adjusted R^2	0.375
Residual Std. Error	23,947.360 (df = 391)
F Statistic	48.501*** (df = 5; 391)
Note:	*p<0.1;* ***p<0.05;*** p<0.01

Model: $67,764.10+ $2391.08(years since phd)+ $1473.92(years_employed)+ $8188.37(sexMale)- $14999.33(disciplineTheoretical) -62.23(years_since_phd:years_employed)

According to the model, professors expect: -base pay of $67,764.10 -Additional $2,391 for each year since phd -Additional $1,473 for each successive year employed -Males to earn $8,188 more per year than females -Professors in the theoretical disciplines to earn about $15,000 less per year than professors in the applied disiplines.

Each of these factors is statistically significant (years_since_phd, p=<.001, years_employed, p=<.001, sexMale P=0.046, disciplineTheoretical p=<.001). Overall model significantly predicts outcomes (p<.001).