DATA MANAGEMENT CULMINATING PROJECT

STAGE 2: Final Proposal of Project

Anais Chen, Cheryl Lee, Vanesse Yuen
C. Arabian
MDM 4U1c
May 15, 2017

TABLE OF CONTENTS

THESIS

What is the correlation between cost of education and employment rates?

HYPOTHESIS

Expensive post-secondary education, i.e. university, is said to be the most effective route to employment. However, we predict that the cost of a bachelor's degree in Canada has only a weak or moderate correlation with employment, due to the many other effective paths one can take after high school.

INTRODUCTION

The topic of getting a job lingers in every high school student's mind. Many parents and students are concerned with whether it is worth spending money on post-secondary education if employment is not guaranteed. In this analysis, we would like to see the true correlation between these two topics, which can be important in providing a definite answer for future generations of students.

CHARACTERISTICS OF DATA

Population & Sample

The population is all university graduates in Canada. The sample is Canadian university graduates of the year 1999-2000, of several different degrees, surveyed a year after graduation.

Type of Study

A cross sectional study would be the most appropriate. We would use data collected from different groups: university graduates of several degrees, from one point in time: 2000.

Method of Data Collection

We would collect data from different groups of graduates using cluster random sampling (e.g. degree and field of study), and sample everyone from each group. We would first put together our sample after requesting information from all public post-secondary institutions in Canada. Using this information, we would divide our sample according to each graduate's respective field(s) of study (which universities can provide to us). We would then survey everyone in each group using information questions (e.g. "How much did you tuition cost?", "Are you employed?"). This would be

the most effective sampling technique because it is organized and accurate. Despite sampling everyone from every group, this would not consume too many resources because the data can be easily collected from school records.

VARIABLES

Two Main Variables

1. Cost of education— Independent
2. Employment rate— Dependent

Our research is focused on the correlation between the cost of education and the employment rate. Cost of education is our independent variable because it will affect employment rate.

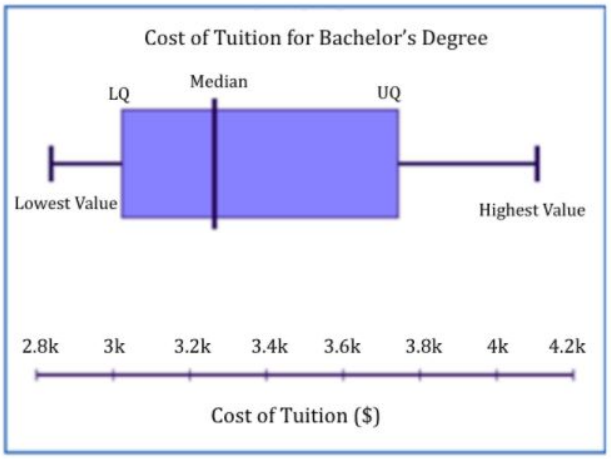
SINGLE VARIABLE ANALYSIS

Measure of Central Tendency (See Appendix 1, 2)

	Mean	Median	Mode
Total Tuition for all subjects	$3359.2	$3264	-
Total Employed (%)	89.125	88	87, 88
Total Number Employed	7558.375	7047	-
Unemployed/Out of Labour Force (%)	12.25	12.5	13

Measures of Spread (See Appendix 1)

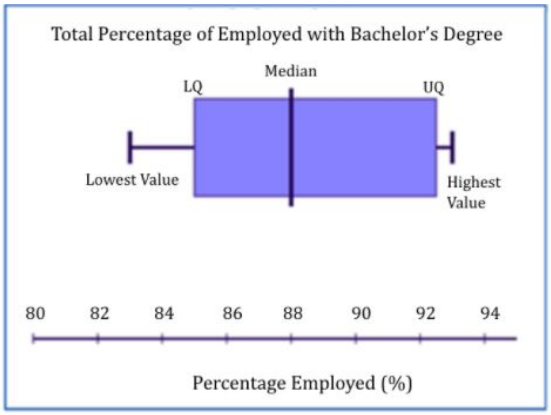
Total Tuition of Graduates of Each Degree (See Appendix 2)



Range: 1268
IQR: 720
Standard Deviation: 464.1435123

The price of tuition varies by little, the greatest difference being $1268. The box-plot is skewed right, meaning that the costs below the median are distributed more closely together than the costs above the median.

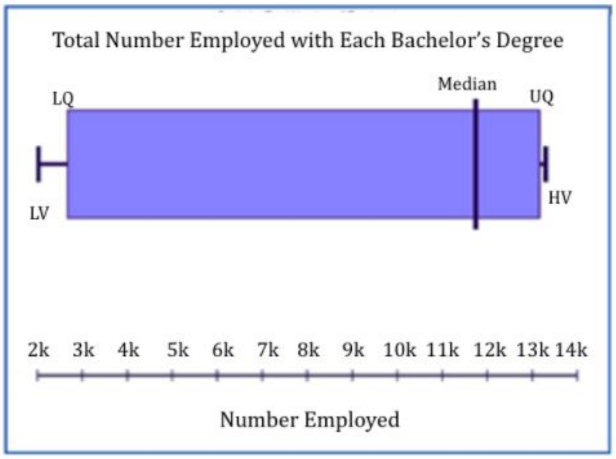
Percent of Employment of Graduates of Each Degree



Range: 10
IQR: 7.5
Standard Deviation: 4.037325848

The percent of employment rates among degrees varies by up to 10%. The box-plot is skewed right, meaning that the percents below the median are distributed more closely together than the percents above the median.

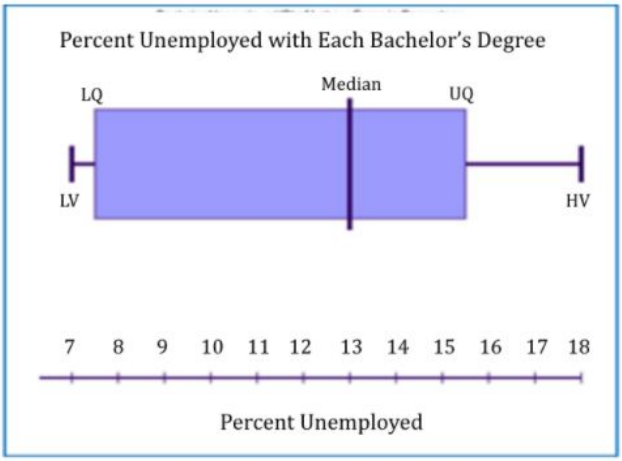
Total Number of Employed of Graduates of Each Degree



Range: 11 272
IQR: 10488
Standard Deviation: 5536.056494

The employment amounts vary by 11 272. The box-plot is skewed far left, meaning that the amounts below the median are distributed much farther together than the amounts above the median.

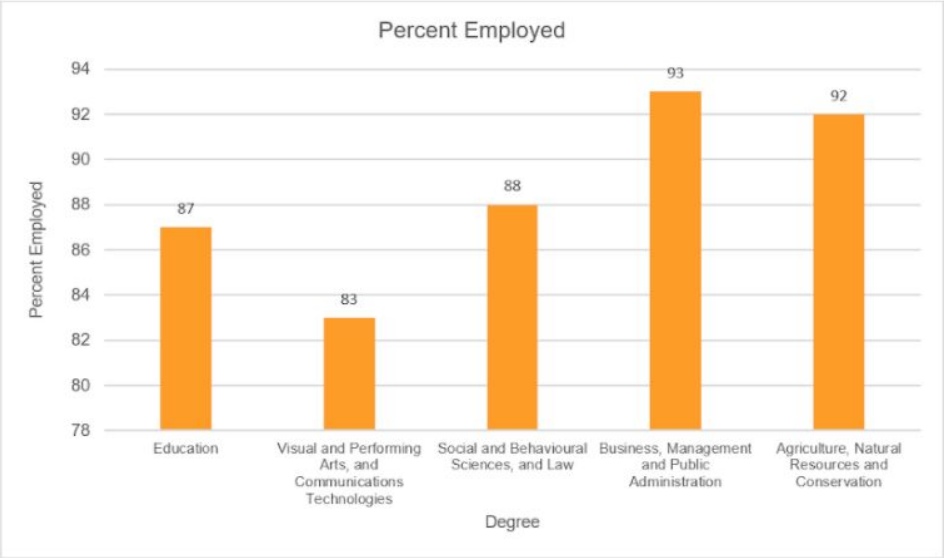
Percent of Unemployment of Graduates of Each Bachelor Degree



Range: 11
IQR: 8
Standard Deviation: 4.438468204

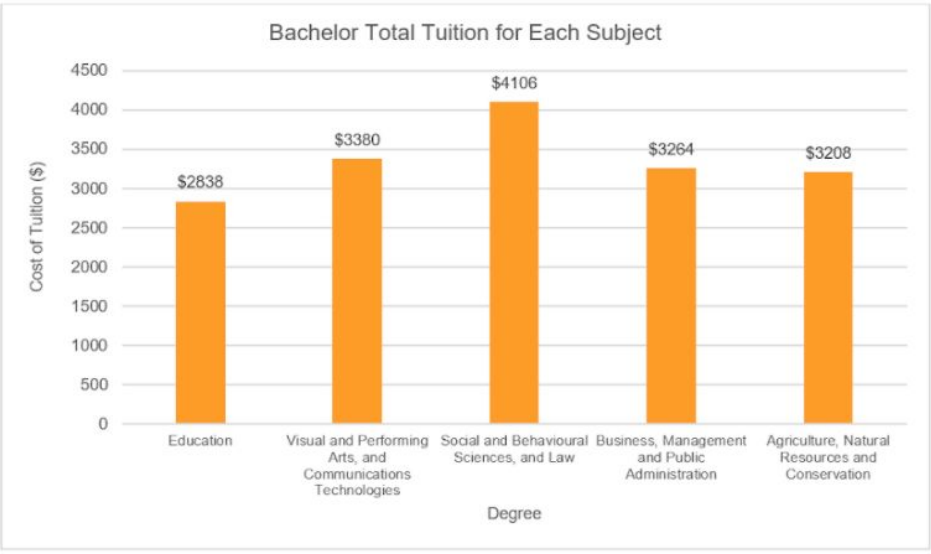
The unemployment rate varies by 11%. The box-plot is skewed left, meaning that the amounts below the median are distributed more closely together than the amounts above the median.

Graphical Displays of Data (See Appendix 1)



Graduates who invested in the degree of "Visual and Performing Arts, and Communications Technologies" became the least likely to be employed, while those with the degree of "Business, Management and Public Administration" became the most employed; Business graduates were 10% more employed than Arts graduates.

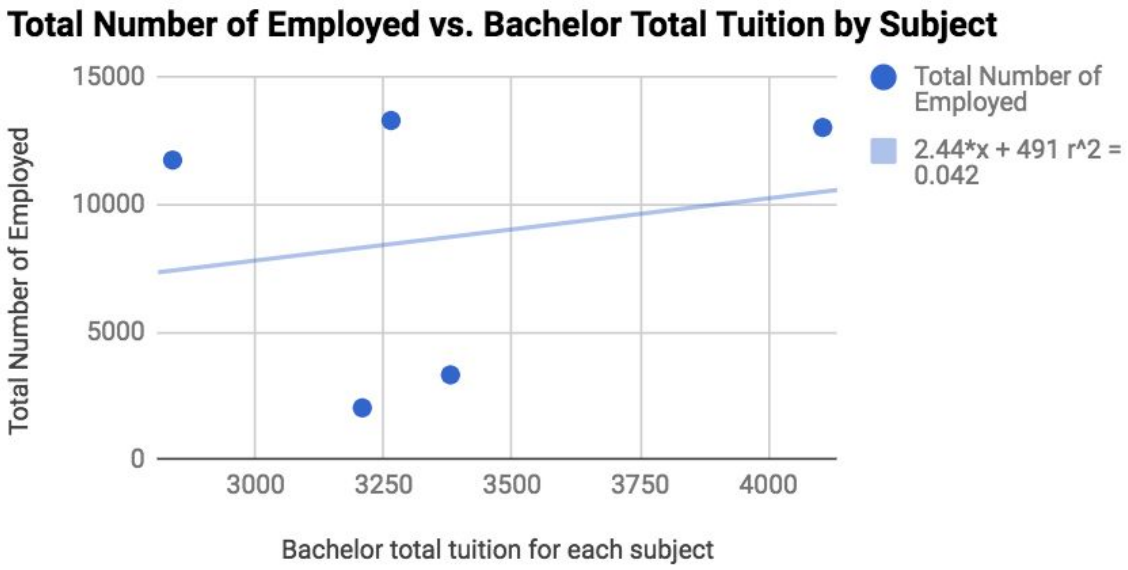
(See Appendix 2)



Students pay the most for a "Social and Behavioural Sciences, and Law" degree and the least for an "Education" degree. However, these two fields are not the most or least employable (according to the previous graph), and therefore does not show strong correlation. The remaining three degrees are fairly similar in tuition cost.

TWO VARIABLE STATISTICS

Regression Analysis (See Appendix 1, 2)



Outliers

There are no outliers because there is a very weak correlation. Therefore they are all outliers.

Explain Which Model is More Suitable and What the Model Explains About Your Thesis

We chose the linear model for this regression because we are looking at the linear correlation between the x and y variables, which are the "Total Number of Employed" and the "Bachelor Total Tuition for Each Subject". The linear model indicates the strength and the direction of this relationship and shows the value of r, which gives us a definite answer on whether it is a strong relation between "Total Number of Employed" and "Bachelor Total Tuition for Each Subject".

Correlation Coefficient and Coefficient of Determination/ Weak or Positive Correlation

Correlation r, r^2. According to the linear regression, the coefficient of determination indicates that about 4.2% of the variation in the total number of employed is due to the variation in the bachelor

total tuition for each subject. Also, the correlation coefficient is about 0.2, which means it has a weak positive linear correlation. This is because the closer the number of correlation coefficient is to 0, the weaker of a linear correlation is.

Cause and Effect

Because there is a weak correlation, there is no cause and effect because neither variable is significantly dependent on the other.

Analysis of Regression Model

The regression model proves that the cost of education does not have a significant influence on employment. According to the graph, we examine two variables: total number of employed versus total tuition for each subject. The lack of a strong correlation proves that the cost of tuition is not as big of a factor in employment as some may think.

Article

https://www.forbes.com/sites/johnebersole/2012/08/08/why-a-college-degree/#7eeb78535ed8
According to John Ebersole's article, there is approximately a 4.5% difference in unemployment rate between those with high school diplomas and bachelor's degrees. This may seem like a large percentage difference, but considering the amount of debt students must go into to obtain the bachelor's degree, the 4.5% is not very significant. The article states that the average cost to obtain the degree is $64,000, in comparison to $0 for a high school diploma. This article supports our hypothesis which states that employment rate is not strongly correlated with tuition costs. Our thesis questions the relationship between the two variables: "cost of tuition" and "employment rates". The answer this article provides is that there is not a strong correlation. However, one factor that the article brought up that we could look farther into is the monthly salary. Ebersole states that there is an average of $400 in difference between high school diploma and bachelor degree holders. Although our hypothesis has been proven and our thesis has been answered, we could look farther into other factors such as monthly salary.

CONCLUSION

Through our analysis, our hypothesis has been proven. Looking at all the graphs we have made, there is not a strong correlation between the cost of tuition and employment rates. We gathered that there is no clear relationship between the amount paid for a degree in a certain subject and the employment rate. The regression model, which gave us r and r^2, proves that a strong correlation between our two variables does not exist, and neither is very dependent on the other. This knowledge is useful for those who are choosing post-secondary pathways as it puts university and college degrees in a new light. This can help students to decide what futures are worth it and the best for them.

APPENDIX

Appendix 1

Bachelor						
Total	**78,900**	**81**	**8**	**90**	**7**	**4**
Education	13,500	74	12	87	9	4
Visual and Performing Arts, and Communications Technologies	4,000	65	18*	83	9*	9**
Humanities	6,300	73	14	88	7*	5*
Social and Behavioural Sciences, and Law	14,800	80	7	88	8*	5*
Social Sciences and Related Interdisciplinary Fields	6,800	79	8**	87	8**	F
Psychology and Related Interdisciplinary Fields	3,100	80	F	88	6**	F
Legal Professions and Studies	1,700	x	x	95	F	x
Business, Management and Public Administration	14,300	89	F	93	5*	2**
Physical and Life Sciences, and Technologies	4,300	81	4**	86	7*	7*
Mathematics, Computer and Information Sciences	3,400	84	F	87	11*	F
Mathematics and Statistics and Related Interdisciplinary Fields	800	x	x	83	F	x
Computer and Information Sciences and Support Services and Related Interdisciplinary Fields	2,500	86	F	88	12*	x
Architecture, Engineering and Related Technologies	6,700	88	F	90	8*	F
Architecture and Related Services and Related Interdisciplinary Studies	500*	x	x	98	x	x
Engineering	6,000	89	F	90	8*	F
Agriculture, Natural Resources and Conservation	2,200	90	F	92	F	x
Agriculture, Agricultural Operations, and Related Sciences	1,200	x	x	94	F	x
Natural Resources and Conservation	1,100	x	x	89	F	F
Health, Parks, Recreation and Fitness	9,000	84	12	95	2*	F
Medicine	600	x	x	99	F	F
Nursing	3,400	84	13	97	F	F
Other Health Professions and Related Clinical Sciences	3,200	85	12*	98	F	F
Parks, Recreation, Leisure and Fitness Studies	1,600	74	12**	86	F	F
Personal, Protective and Transportation Services	300**	x	x	F	F	F
Other	100**	x	x	F	F	F

Appendix 2

Average tuition fees[1]

	1998/1999	1999/2000	1998/1999 to 1999/2000
	$		% change
Agriculture	3,183	3,364	5.7
Architecture	3,132	3,372	7.7
Arts	3,156	3,379	7.1
Commerce	3,131	3,391	8.3
Dentistry	5,576	7,377	32.3
Education	3,032	3,245	7.0
Engineering	3,292	3,606	9.5
Household sciences	3,031	3,260	7.6
Law	3,274	3,639	11.2
Medicine	4,930	5,699	15.6
Music	3,096	3,299	6.6
Science	3,150	3,357	6.6
Undergraduate	3,155	3,433	8.8
Graduate	3,400	3,681	8.3

BIBLIOGRAPHY

Allen, Mary, and Chantal Vaillancourt. "Class of 2000: Profile of Postsecondary Graduates and Student Debt." (n.d.): n. pag. Apr. 2004. Web. 12 Apr. 2017. <http://www.statcan.gc.ca/pub/81-595-m/81-595-m2004016-eng.pdf>

Ebersole, John. "Why a College Degree?" *Forbes*. Forbes Magazine, 08 Aug. 2012. Web. 15 May 2017. <https://www.forbes.com/sites/johnebersole/2012/08/08/why-a-college-degree/#7eeb78535ed8>

Frenette, Marc, and René Morissette. "Wages and Full-time Employment Rates of Young High School Graduates and Bachelor's Degree Holders, 1997 to 2012." Statistics Canada. Government of Canada, 27 Nov. 2015. Web. 12 Apr. 2017. <http://www.statcan.gc.ca/pub/11f0019m/11f0019m2014360-eng.htm>.

National Graduates Survey, 2013. "Description for Chart 4.5." Statistics Canada. Government of Canada, 27 Nov. 2015. Web. 12 Apr. 2017. <http://www.statcan.gc.ca/pub/81-595-m/2014101/c-g/desc/desc45-eng.htm>.

"Study: Labour Market Outcomes of Graduates from Universities in the Maritime Provinces, 2006 to 2011." Statcan.gc.ca. N.p., 4 Nov. 2017. Web. 12 Apr. 2017. <http://www.statcan.gc.ca/daily-quotidien/170411/dq170411a-eng.htm?HPA=1>.

"University Tuition Fees." The Daily, Statistics Canada. Statistic Canada, 25 July 1999. Web. 15 May 2017. <http://www.statcan.gc.ca/daily-quotidien/990825/dq990825a-eng.htm>.