EDUC 641 Assignment 05 Key

## Research question 1.

Whether having a school-based mentor relates to whether participants' high school overall GPA above 3.0?

In this question, the predictor variable is “mentor”, a binary variable, which is coded one for individuals reporting having a mentor after 14 years old. The outcome variable is “gpa\_3”, a binary variable, which is coded 1 for individuals whose high school final GPA was above 3.0. Summary statistics in Table 1 shows that there are 1,395 individual-level observations in the sample, in which 657 (47%) reported to have a school-based mentor while 738 (53%) not. There were 525 (38%) individuals whose high school overall GPA was above 3.0.

The plot in Figure 1 visualizes the differences in counts of individuals whose high school GPA was above/below 3.0 between the individuals who had a mentor and those who had not. Among the 525 individuals having a high school GPA above 3.0, the majority (336, 64%) had a mentor. Among the 870 individuals whose GPA was below 3.0, the majority (549, 63%) didn’t have a mentor. These differences indicate that there may be a relationship between mentor and high school GPA but we don’t know whether this relationship is statistically significant.

My null hypothesis is that there was no relationship between mentor and high school GPA above 3.0. I conduct a chi-squared test in R to test the hypothesis. Results show that there was a significant, positive relationship between having a mentor and high school overall GPA above 3.0 (*X2* = 95.449, *p* < 0.01).

Table 1. Summary Statistics

Graphical user interface, text

Description automatically generated

Figure 1.

Chart, bar chart

Description automatically generated

## Research question 2.

Whether and to what extent having a school-based mentor relates to high school overall GPA?

In this question, the predictor variable is “mentor”, a binary variable, which is coded one for individuals reporting having a mentor after 14 years old. The outcome variable is “gpa”, a continuous variable, which represents the individual’s high school overall GPA. Summary statistics in Table 1 shows that there are 1,395 individual-level observations in the sample, in which 657 (47%) reported to have a school-based mentor while 738 (53%) not. The average GPA is 2.64 with a standard deviation of 0.83.

The boxplot in Figure 1 visualizes the difference in average outcome (high school GPA) between the individuals who had a mentor and those who had not. The average GPA of the former group is higher than that of the latter but we do not know whether this is significant. Moreover, the distribution of the outcome variable seems to be positively skewed for both groups and both have a few outliers.

My null hypothesis is that the two means are equal and I conduct a two-sample *t*-test to test the hypothesis. Results show that the average high school GPA of individuals who had a mentor is about 0.51 higher than that of the individuals who had not (*t*(1392.4) = -12.34, *p* < 0.05). I reject the null hypothesis and conclude that having a school-based mentor is significantly related to an increase of 0.51 in overall high school GPA.

Table 1. Summary Statistics

Graphical user interface, text, application

Description automatically generated

Figure 1.

Chart, box and whisker chart

Description automatically generated

## Research question 3.

Whether and to what extent mentee’s age relates to high school overall GPA?

In this question, the predictor variable is “mentee\_age”, a continuous variable, which represents the individual’s age when they first felt being influenced by the mentor. The outcome variable is “gpa”, a continuous variable, which represents the individual’s high school overall GPA. Summary statistics in Table 1 shows that there are 657 individual-level observations in the sample and the average GPA is 2.92 with a standard deviation of 0.73.

Figure 1 visualizes the relationship between the predictor and outcome variables. The trend-line estimated by Ordinary Least Squares (OLS) regression suggests a positive relationship between these two variables but I do not know whether it is statistically significant.

My null hypothesis is there is no relationship between the mentee’s age and high school overall GPA. I fit an ordinary-least-squares regression of high school GPA on mentee’s age to see whether the slope of the estimated trend-line is significantly different from zero. The *p* value of the slope is 0.262, which is above the alpha-threshold of 0.05. I fail to reject the null hypothesis and conclude that mentee’s age is not significantly related to high school overall GPA.

Table 1. Summary Statistics

Graphical user interface, text, table

Description automatically generated with medium confidence

Figure 1.

Chart, scatter chart

Description automatically generated