**Assignment 4: Categorical IVs**

**Due on March 26**

The data file is the Student Survey Data file. Five variables include:

**ID:** case id.

**Gender:** males = 1 and females = 2.

**Age:** students’ age.

**Party :** an categorical variable which contains three political parties: Republican = 1, Democrat = 2, and Independent = 3.

**Marijuan**: attitudes toward legalization of marijuana. High scores on the MARIJUAN variable reflect more positive attitudes toward legalization.

1. Suppose you wanted to conduct a regression analysis to determine whether males and females (GENDER) differ in their attitudes toward legalization of marijuana (MARIJUAN). Create a dummy variable for gender and conduct the regression analysis. Provide the table of dummy codes and syntax you used for the analysis. Provide an interpretation of the slope and intercept values on the printout.

The dummy code for gender in this analysis is:

|  |  |
| --- | --- |
| Gender | Female |
| Gender=Female | 1 |
| Gender=Male | 0 |

The syntax is:

MarijuanaByGender<-glm(Marijuana~Female, data=ds)

The results of the regression are:

=

=

The average of the measure of opinion on the legalization of marijuana for the reference group (Female=0) is =2.47. The effect of the variable “Female” on the average “Marijuana” measure is to increase the average by =0.12, a difference that fails to reach statistical significance, t(248)=0.07, p>.05.

1. Suppose you wanted to conduct a regression analysis to determine whether there were differences among the three political parties (PARTY) with respect to attitudes toward legalization of marijuana (MARIJUAN). Create dummy variables for PARTY and conduct the regression analysis. Provide the table of dummy codes and syntax you used for the analysis. For the purpose of this analysis, assume that Republicans are the reference group. Answer the following questions based on the printout:

The dummy code for Party in this analysis is:

|  |  |  |
| --- | --- | --- |
| Political Party | DEM | IND |
| Republican | 0 | 0 |
| Democrat | 1 | 0 |
| Independent | 0 | 1 |

* 1. Did Democrats have significantly different attitudes than Republicans? What quantities on the printout did you use to determine this?

The Democrats were significantly more positive towards Marijuana, the regression coefficient associated with the DEM variable is equal to =0.72, suggesting that their average is 0.72 higher than the average for Republicans, t(1)=4.122, p<0.01.

* 1. Who scored higher (i.e., more favorably)? What quantity on the printout did you use to determine this?

The Democrats tended to be most positive, with independents falling in the middle. The regression coefficient for the Democrats, =0.72, suggests that their average is 0.72 higher than the average for Republicans, giving them an average of 2.875. The regression coefficient for Independents, , suggests that their average 0.38 higher than the average for Republicans, giving them an average of 2.5429, a statistically significant difference (t(1)=2.192, p<0.05).

* 1. What is the proportion of variation in attitudes toward legalization of marijuana explained by political party?

The proportion of variation in attitudes toward legalization of marijuana explained by political party is approximately 6.6%, , F(2,246)=8.674, p<0.001.

1. Create unweighted effect code for PARTY to contrast Republican and Democrat with the unweighted overall mean of the three parties. Run the regression analysis with the effect code variables. Provide the table of effect codes and syntax you used for the analysis. Answer the following questions based on the printout:
   1. What is the overall mean attitude of the three parties?

The mean of the Independents = 2.53 – (.35) – (-.37) = 2.55

The mean of the Democrats = 2.53 + (.35) = 2.88

The mean of the Republicans = 2.53 + (-.37) = 2.16

The Grand Mean is

* 1. Did Republican have significantly different attitudes than the average attitude of the three parties? What quantities on the printout did you use to determine this?

Republicans had significantly less positive attitudes towards Marijuana than the grand mean, , t(1)=-3.77, p<0.001.

* 1. Did Democrats have significantly different attitudes than overall mean? What quantities on the printout did you use to determine this?

Democrats had an average attitude significantly higher than the grand mean, , t(1)= 3.268, p<0.01.

* 1. What is the proportion of variation in attitudes toward legalization of marijuana explained by political party? Is the proportion the same as the result in 2(c)?

Political party explained approximately 6.6%, F(2,246)=8.674, p<0.001. This is exactly what we saw in question 2c.

1. Suppose a researcher hypothesized that Democrats and Independents have very similar attitude toward legalization than republican. Both of them have more positive attitudes than republican. Create contrast code for PARTY to test the hypothesis. Provide the table of contrast codes and syntax you used for the analysis. Answer the following questions based on the printout:

|  |  |  |
| --- | --- | --- |
| Party | C1 | C2 |
| Republican | -2/3 | 0 |
| Democrat | 1/3 | 1/2 |
| Independent | 1/3 | -1/2 |

* 1. Did Democrats and Independents have similar attitude toward legalization? What quantities on the printout did you use to determine this?

Democrats and Independents did NOT have significantly different average attitudes towards marijuana legalization, t(1)=1.736, p=0.084.

* 1. Did Democrats and Independents have more positive attitudes than republican? What quantities on the printout did you use to determine this?

Democrats and Independents, as a group, did have a significantly more positive average opinion of marijuana legalization, t(1)=3.77, p<0.001.

* 1. What is the proportion of variation in attitudes toward legalization of marijuana explained by political party? Is the proportion the same as the result in 2(c)?

According to this model, political party explained approximately 6.6% of the variation in attitude toward marijuana, a statistically significant portion, F(2/246)=8.674, p<0.001. This is the same as in question 2(c).

1. Suppose the dummy codes created in question 2 are used. You then included age in the model to test the joint effect of age and party on MARIJUAN. Write down the estimated regression equation and interpret the intercept and slopes in the equation. What is the proportion of variation in attitudes toward legalization of marijuana explained by both age and political party?

The estimated regression equation for this model is

This suggests that the average attitude toward marijuana legalization for Republicans of zero years of age is 1.88. This average is expected to increase by 0.66 for Democrats, holding age constant. The average for Independents is expected to be larger than the average for Republicans (of Age=0) by 0.35, holding age constant. The average for all groups is expected to increase by 0.01 for each year of life.

This model explains approximately 7.2% of the variation in attitude toward marijuana legalization, a statistically significant proportion, F(3,245)=6.316, p<0.001.