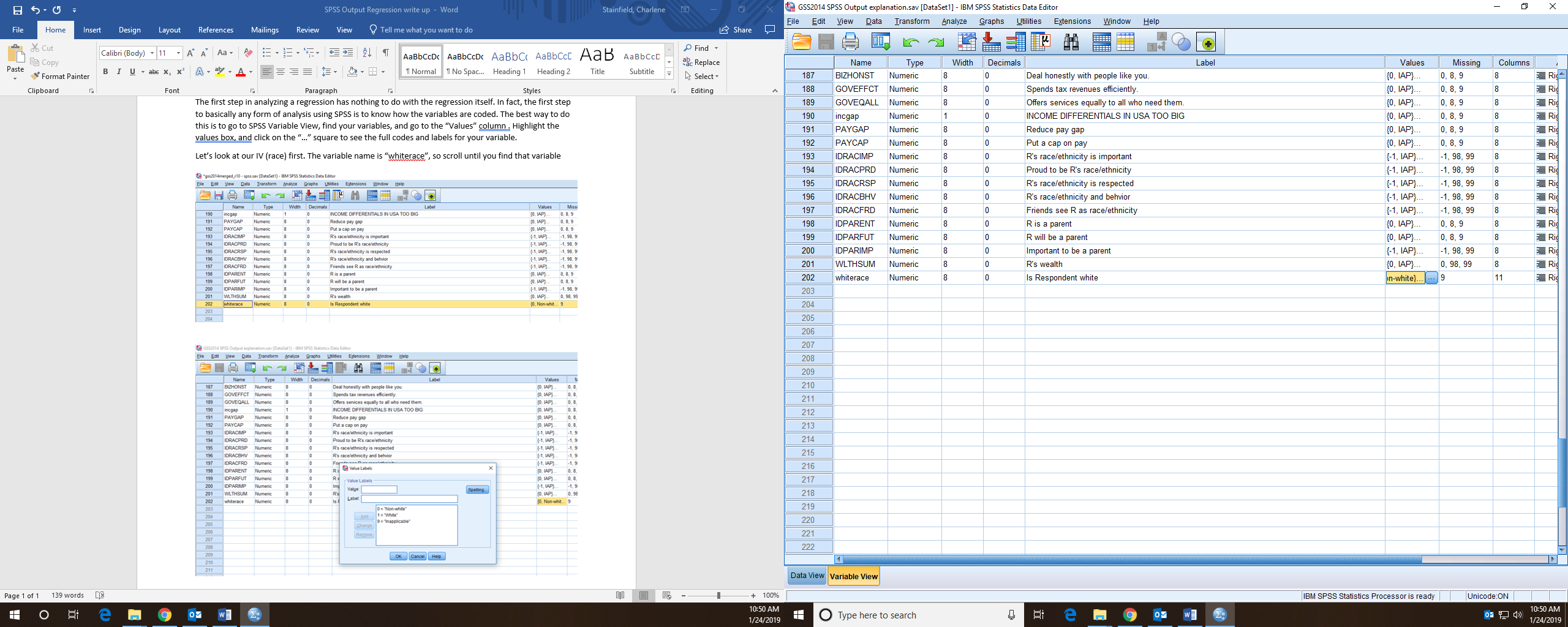
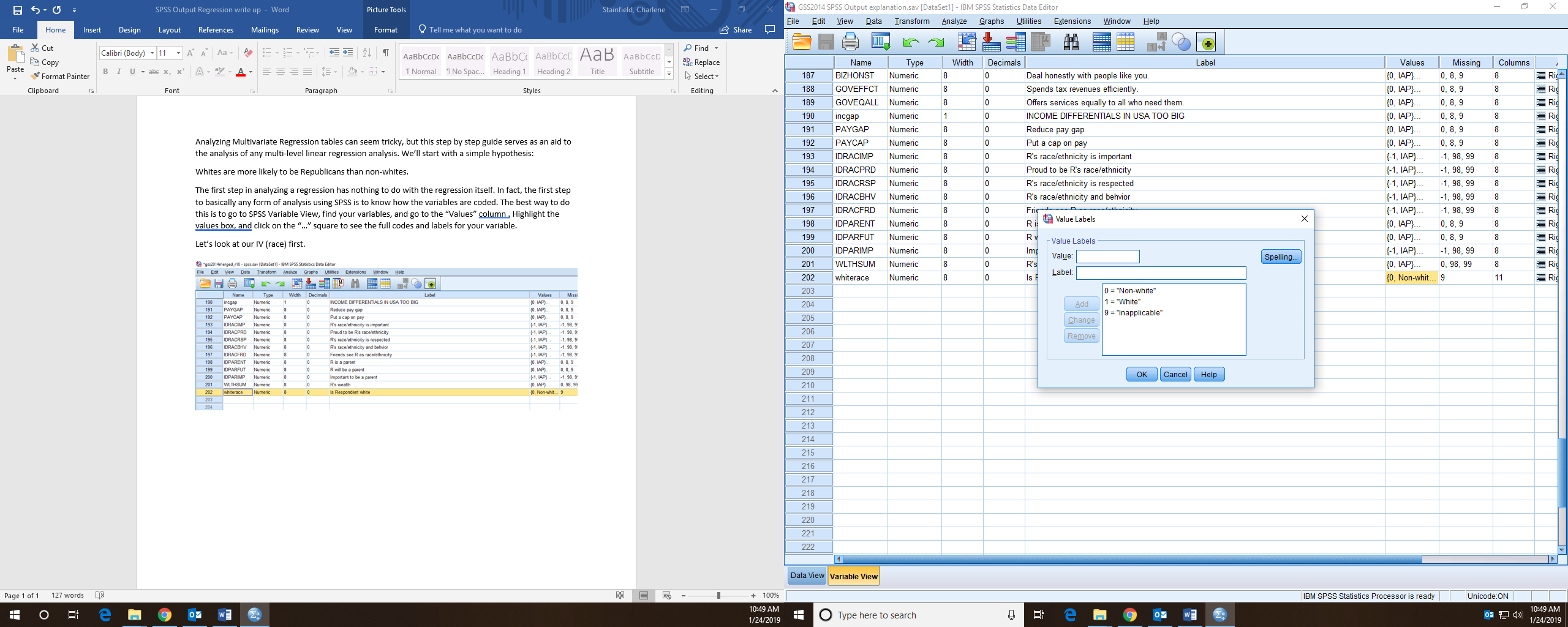
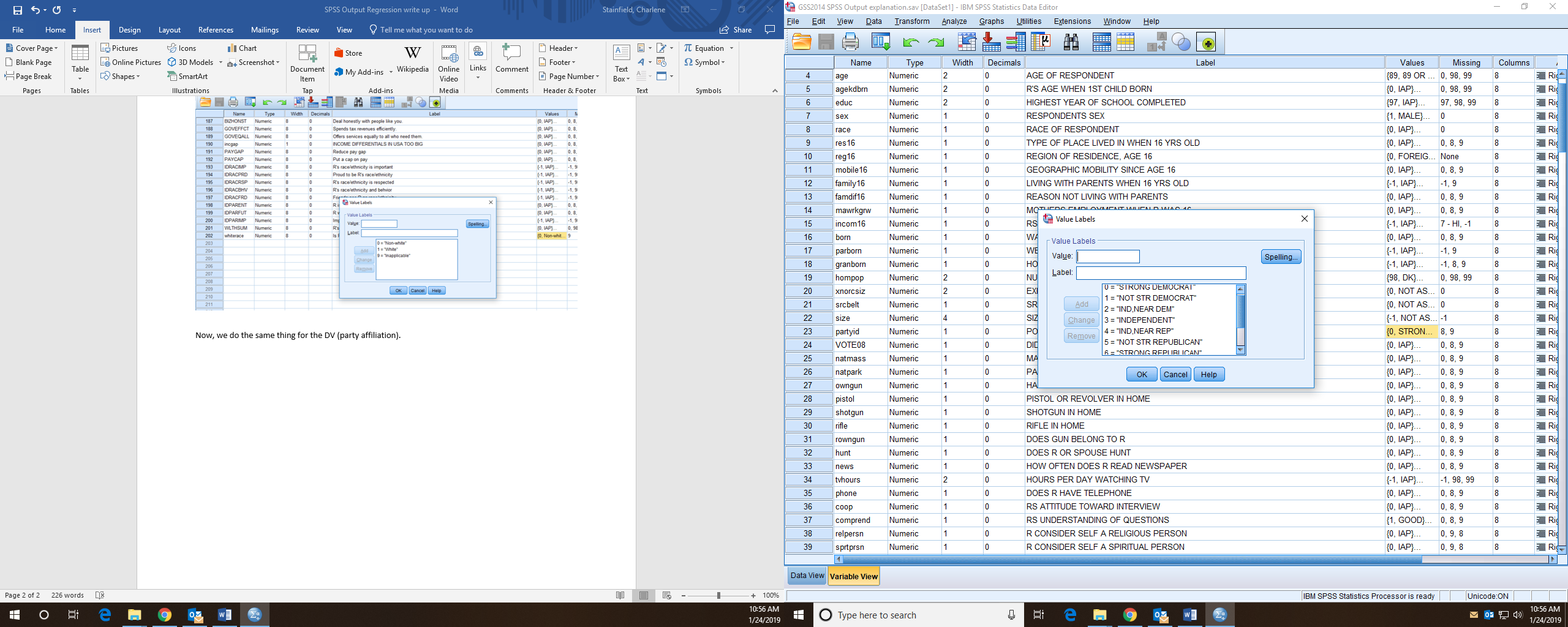
Analyzing Multivariate Regression tables can seem tricky, but this step by step guide serves as an aid to the analysis of any multi-level linear regression analysis. We’ll start with a simple hypothesis: Whites are more likely to be Republicans than non-whites.

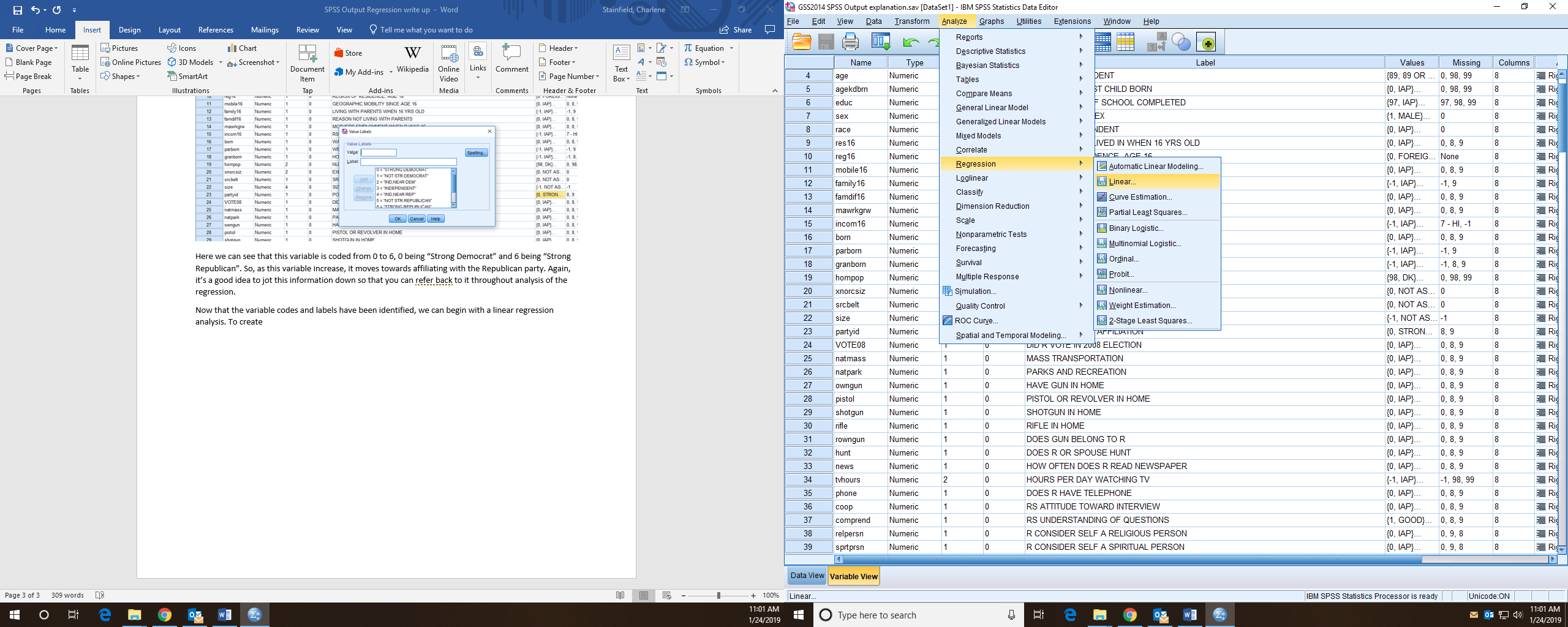
The first step in analyzing a regression has nothing to do with the regression itself. In fact, the first step to basically any form of analysis using SPSS is to know how the variables are coded. The best way to go about this is to go to SPSS Variable View, find your variable, and go to the “Values” column. Highlight the values box and click on the “…” square to see the full codes and labels for your variable.

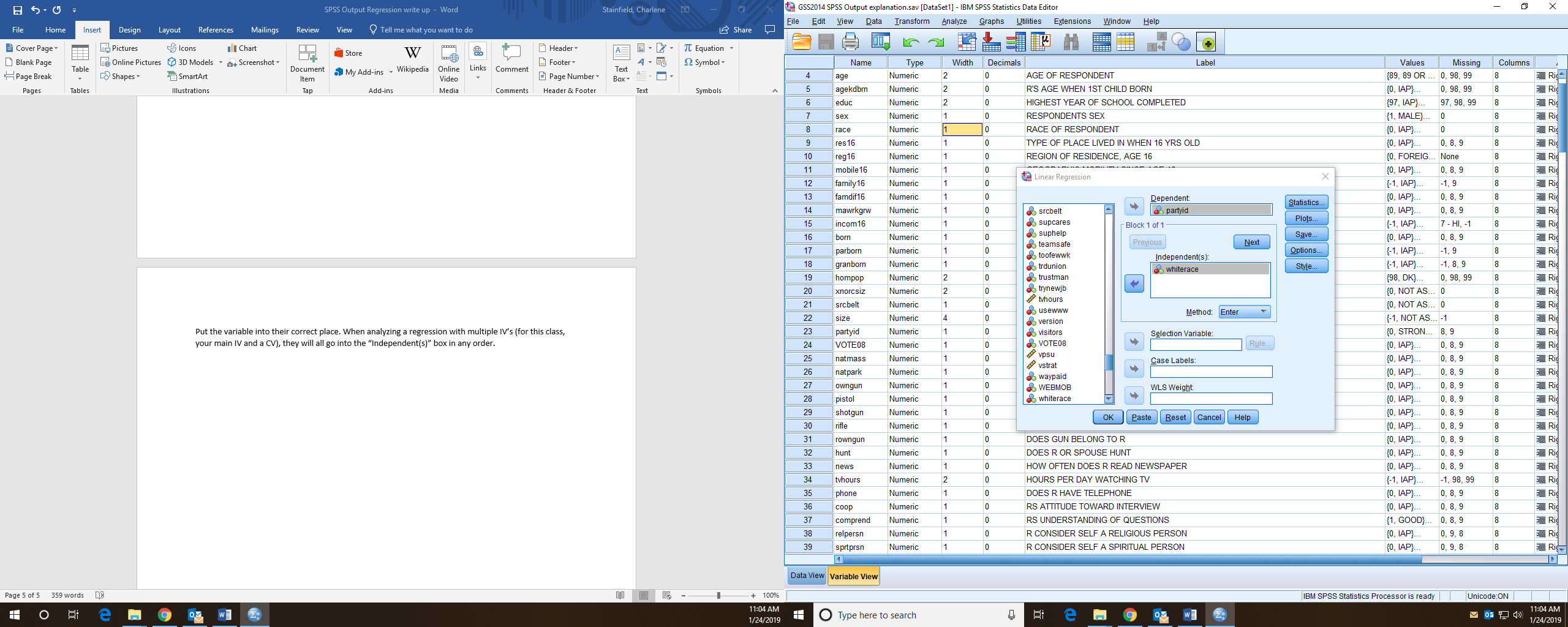
Let’s look at the IV (race) first. The variable name is “whiterace”, so scroll until you find that variable and highlight the “Values” column.

Click on the “…” square to display the codes and labels. For this variable, non-whites are coded 0 and whites are coded 1. So, as this variable increases from 0 to 1, it moves from the respondent not being white to being white. You can either make a mental note of this, or write the coding out somewhere, so you don’t need to keep coming back to this screen to double-check.

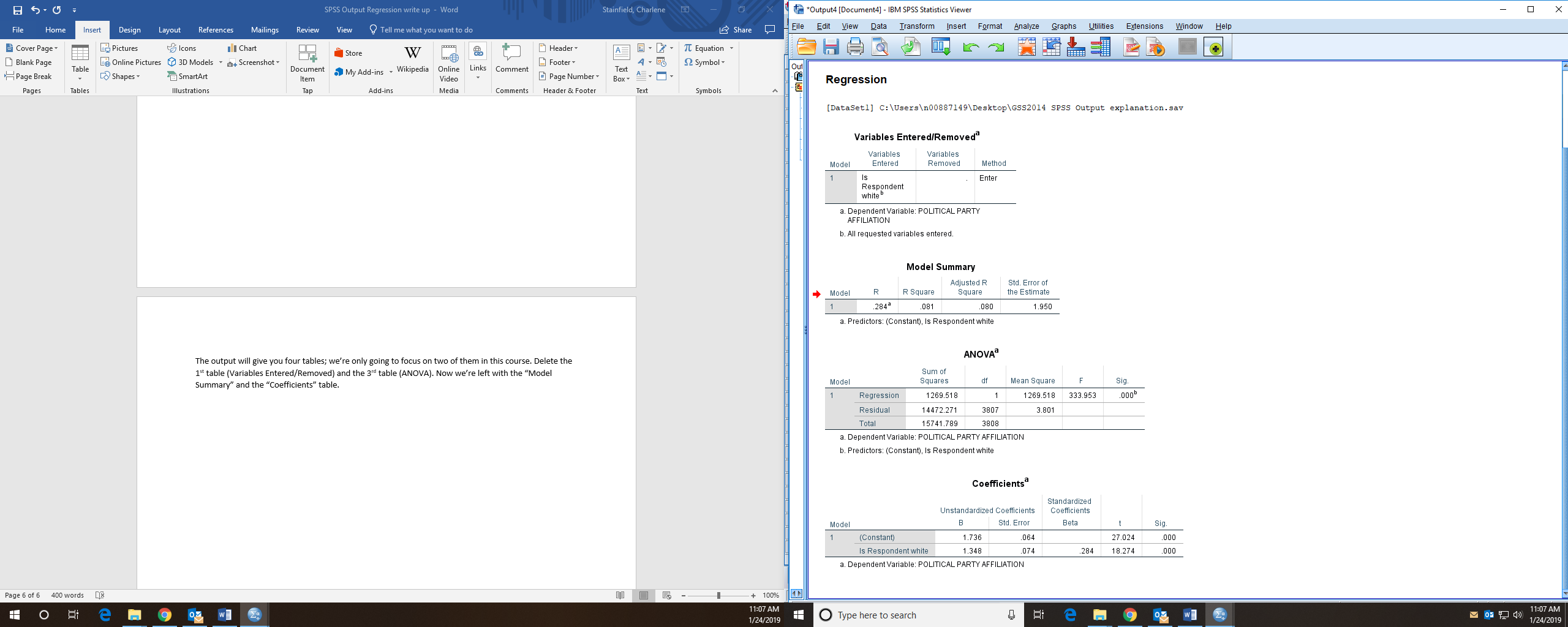
Now, we do the same thing for the DV (party affiliation), named “partyid” in the dataset.

Here we can see that this variable is coded from 0 to 6, 0 being “Strong Democrat” and 6 being “Strong Republican”. So, as this variable increase, it moves towards affiliating with the Republican party. Again, it’s a good idea to jot this information down so that you can refer back to it throughout analysis of the regression.

Now that the variable codes and labels have been identified, we can begin with a linear regression analysis. To create regression tables, go to “Analyze” > “Regression” > “Linear”, and a new screen with pop up.

Put the variable into their correct place. When analyzing a regression with multiple IV’s (for this class, your main IV and a control variable), they will all go into the “Independent(s)” box in any order.

The output will give you four tables; we’re only going to focus on two of them in this course. Delete the 1st table (Variables Entered/Removed) and the 3rd table (ANOVA).



Now we’re left with the “Model Summary” and the “Coefficients” table. Time for some analysis!

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Model Summary** | | | | |
| Model | R | **R Square** | Adjusted R Square | Std. Error of the Estimate |
| 1 | .284a | **.081** | .080 | 1.950 |
| a. Predictors: (Constant), Is Respondent white | | | | |

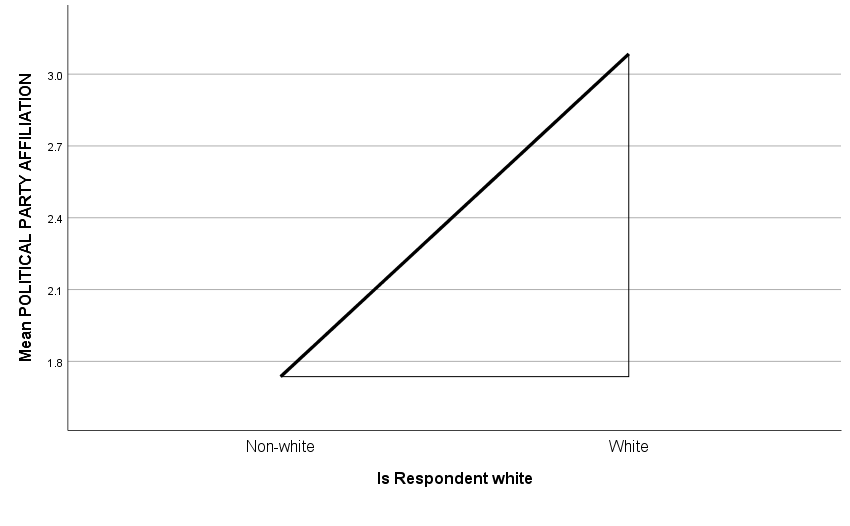
|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Coefficientsa** | | | | | | |
| Model | | Unstandardized Coefficients | | Standardized Coefficients | t | Sig. |
| B | Std. Error | Beta |
| 1 | (Constant) | **1.736** | .064 |  | 27.024 | .000 |
| Is Respondent white | **1.348** | .074 | .284 | 18.274 | .000 |
| a. Dependent Variable: POLITICAL PARTY AFFILIATION | | | | | | |

Let’s go in order of the information. First, we’ll look at the R² value in the Model Summary. As can be seen above, the R² is .081. This number should be read as a percentage, so in this case, R² is 8.1%. Now, how do we interpret this percentage? The R² percentage measures the amount of change in your dependent variable that can be explained by your independent variable. So, in this case, 8.1% of the change in an individual’s party affiliation can be explained by whether they are white or not. This isn’t that large, but that makes sense; there are an absurd number of factors that contribute to why someone affiliates with one party over another, and whether or not someone is white makes up a portion of that change; 8.1% of the change.

**Note: R² is NOT an indicator of statistical significance; no matter how small or large R² is does not affect whether or not the hypothesis is supported. Do NOT cite R² when discussing statistical significance.**

That’s the only figure we need from the Model Summary, so now we’ll move onto the Coefficients table, where the bulk of the analysis happens! First, let’s talk about the constant, located in the first column under the “B”. In technical terms, the constant represents where the linear regression line crosses the y axis, so the y-intercept. Basically, a linear regression table allows you to create a line graph representing the relationship between your variables (in this case, whiteness and party affiliation). So, the constant is 1.736 on the y-axis.

Go back to high school math for a minute. When graphing a relationship, the IV goes on the x-axis and the DV goes on the y-axis, because as you increase or decrease the x-value, the value for y will change as well. As the IV moves either up or down, it causes a shift in the DV. That means that the DV, party affiliation, is running on the y-axis. But how do we know which values to put on the y-axis? Go back to the coding. Partyid is coded from 0-7, strong democrat to strong republican. That means that the constant value is placed between “Not strong democrat” and “independent, near democrat”, but a bit closer to “independent, near democrat”, because the value of the constant is 1.736, closer to 2 than 1. So, if you were to graph the relationship between whiteness and party affiliation, you could put the first point at (0, 1.736). This will be the starting point for analyzing the coefficients, which can be found in the first column underneath the constant. Coefficients are similar to Pearson’s correlations (remember, those numbers in Bivariate Correlations), in that they provide the strength and direction of the relationship between two variables; however, coefficients are more specific than Pearson correlations, in that coefficients tell us exactly how the DV will move on the y-axis when the IV increases by a code of 1. For this regression, the coefficient for your IV, whiterace, is 1.348. That means that when the IV moves from 0 to 1 (from non-white to white), the DV increases on the partyid scale by exactly 1.348 units. Now, this is where the connection to the constant comes in. If your constant is the number when x=0, the coefficient tells you that moving to x=1, or the respondent is white, increases the partyid scale by 1.348 units. So, you add 1.348 to the constant, 1.736, to get a score on the partyid scale of 3.084. See the chart below for a visual representation of what happens to the relationship when the coefficient is considered.

****

A code of 3 on the partyid scale is independent, closer to Republican codes than when x=0 for non-white respondents. Although this does not seem to be a monumentally large shift in party affiliation, it is the direction that was hypothesized originally. **But, we cannot say that the hypothesis has support unless the relationship is statistically significant.**

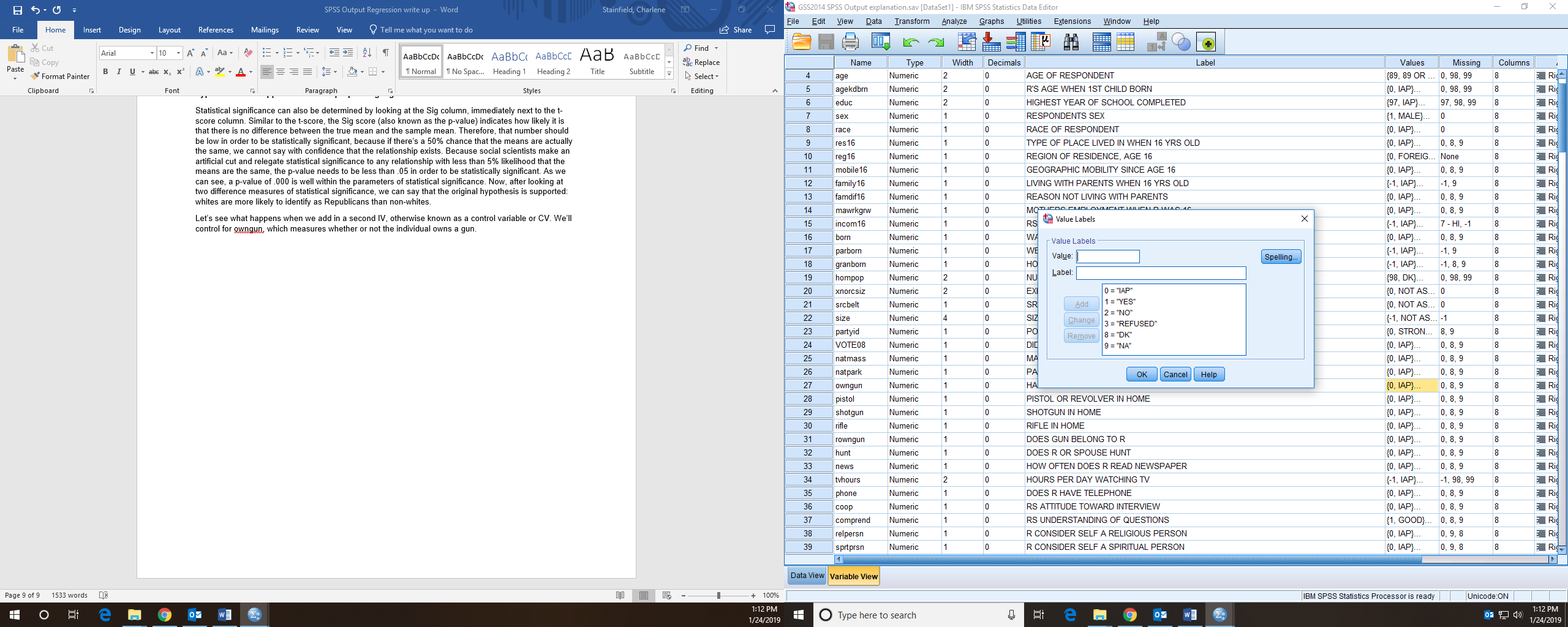
To determine statistical significance, there are two places we can look on the coefficient table. This first is at the t-score column to the right.

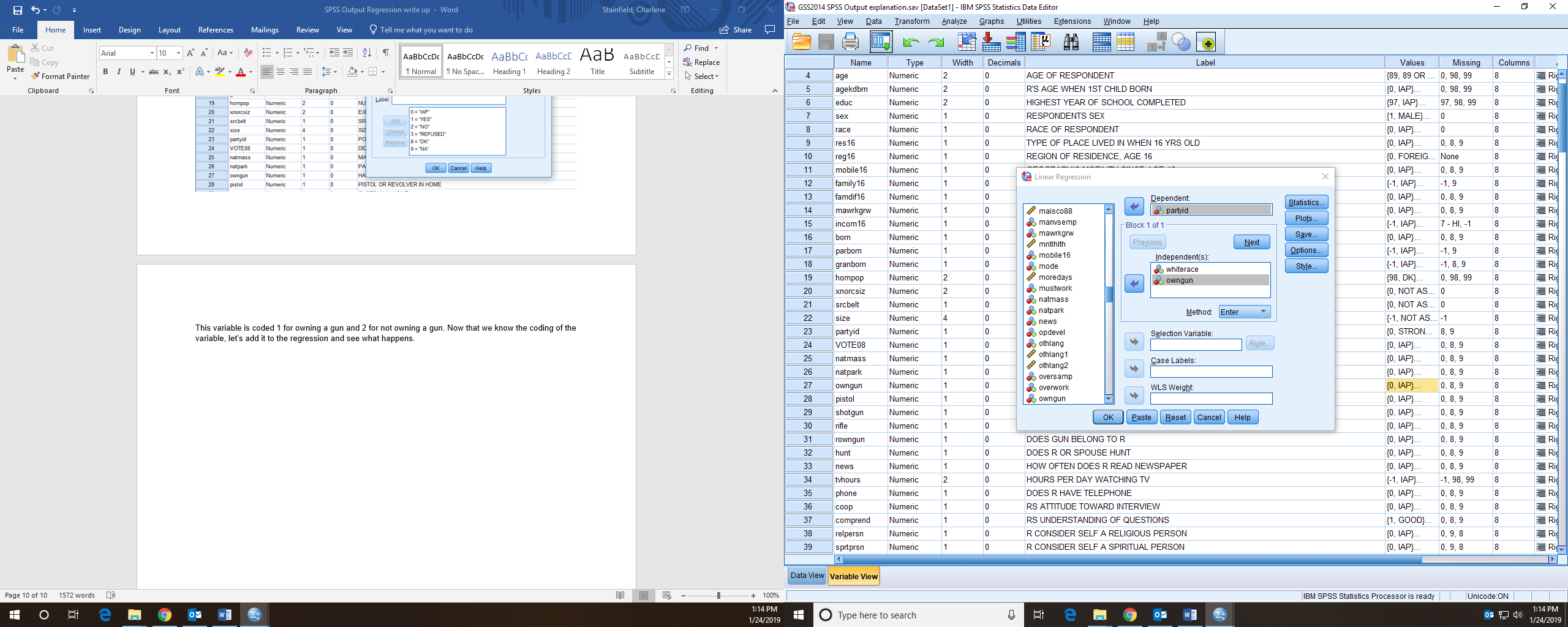
|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Coefficientsa** | | | | | | |
| Model | | Unstandardized Coefficients | | Standardized Coefficients | **t** | **Sig.** |
| B | Std. Error | Beta |
| 1 | (Constant) | 1.736 | .064 |  | **27.024** | **.000** |
| Is Respondent white | 1.348 | .074 | .284 | **18.274** | **.000** |
| a. Dependent Variable: POLITICAL PARTY AFFILIATION | | | | | | |

Think back to normal distributions and standard deviations. A majority of cases, 95%, fall within two standard deviations of the mean. If a sample mean falls outside of that 95%, on the 2.5% greater than or the 2.5% less than, it is considered statistically significant with 95% confidence. T-scores measures the standard deviations, so anything greater than the absolute value of 2 is considered statistically significant in this course. The t-score for whiterace is 18.274, indicating that the sample mean is a little more than 18 standard deviations above the average. Thus, this relationship is statistically significant, and we can say that the hypothesis has support.

**Note; we never “prove” or have “correct” hypotheses in this class. Your hypothesis was not “right”. What we have determined is that the null hypothesis can be rejected, and that the hypothesis has support. This is the proper language and should be used as such.**

Statistical significance can also be determined by looking at the Sig column, immediately next to the t-score column. Similar to the t-score, the Sig score (also known as the p-value) indicates how likely it is that there is no difference between the true mean and the sample mean. Therefore, that number should be low in order to be statistically significant, because if there’s a 50% chance that the means are actually the same, we cannot say with confidence that the relationship exists. Because social scientists make an artificial cut and relegate statistical significance to any relationship with less than 5% likelihood that the means are the same, the p-value needs to be less than .05 in order to be statistically significant. As we can see, a p-value of .000 is well within the parameters of statistical significance. Now, after looking at two difference measures of statistical significance, we can say that the original hypothesis is supported: whites are more likely to identify as Republicans than non-whites.

Let’s see what happens when we add in a second IV, otherwise known as a control variable or CV. We’ll control for owngun, which measures whether or not the individual owns a gun. First, check the codes.

This variable is coded 1 for owning a gun and 2 for not owning a gun. Now that we know the coding of the variable, let’s add it to the regression and see what happens. Again, place it in the “Independent(s)” box along with the main IV, whiterace.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Model Summaryb** | | | | |
| Model | R | R Square | Adjusted R Square | Std. Error of the Estimate |
| 1 | .303a | .092 | .091 | 1.948 |
| a. Predictors: (Constant), HAVE GUN IN HOME, Is Respondent white | | | | |
| b. Dependent Variable: POLITICAL PARTY AFFILIATION | | | | |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Coefficientsa** | | | | | | |
| Model | | Unstandardized Coefficients | | Standardized Coefficients | t | Sig. |
| B | Std. Error | Beta |
| 1 | (Constant) | 2.457 | .178 |  | 13.819 | .000 |
| Is Respondent white | 1.275 | .100 | .272 | 12.801 | .000 |
| HAVE GUN IN HOME | -.374 | .084 | -.094 | -4.448 | .000 |
| a. Dependent Variable: POLITICAL PARTY AFFILIATION | | | | | | |

Including owngun gives us a new Model Summary as well as a new Coefficient table. Now, the R² has increased to 9.2%, which makes sense. We’ve added another factor that could contribute to the change in an individual’s party affiliation. So, 9.2% of the change in someone’s party affiliation can be explained by the whether the respondent is white and has a gun in the home.

The constant and coefficients have also changed now that gun ownership has been considered. Now, the constant is 2.457, almost halfway between a weak Democrat and a leaning Democrat. The whiterace variable has a coefficient of 1.275, indicating that as x increases from non-white at 0 to white at 1, the party affiliation scale increases by 1.275 units from the constant to 3.732 on the scale, close to leaning Republican.

We also have a new coefficient for gun ownership at -.374. Why is it negative, when we would think that gun owners are more likely to be Republican? Let’s look back at the coding. A code of 1 is “Yes” and a code of 2 is “No”. So, as the variable moves up 1 from those with guns in the home to those without guns in the home, the DV decreases by .374 from the constant, to a value of 2.083, which is a weak Democrat on the DV scale. This makes sense, as those without guns are more likely to identify as Democrats. Variables that are coded inversely, like this one, require a careful analysis and thorough knowledge of how they are coded.

When examining statistical significance, we can see that the t-scores for each variable is greater than the absolute value of two, and their Sig scores are less than .05. Therefore, we have support for the hypothesis: When controlling for guns in the home, whites are still more likely to identify as Republicans.

You may have noticed that there are two columns in the Coefficients table that have not been addressed. Although we won’t use them much, it is still helpful to know what those numbers mean. The “Std. Error” column indicates the standard error around the coefficients. Because life is never so precise as numbers to the third decimal, the standard error is there to give a range that could be inclusive of the true coefficient. So, the coefficient for whiterace is 1.275 +/- .100, and the range for this coefficient is 1.175 to 1.375. The same concept can be applied to owngun, in which the range for the gun ownership coefficient is -.0458 to -.290.

As for the “Standardized Coefficients Beta”, there’s no need to deal with this column, so either leave it or delete it.

Now that you have a grasp on analyzing regression outputs, you can test hypotheses to your heart’s desire.