Model Metrics for Logistic Regression Models

- In order to understand how much variation in the dependent variable can be explained by the model (the equivalent of R^2 in multiple regression), you should consult ***Model Summary*** statistics.
- Although there is no close analogous statistic in logistic regression to the coefficient of determination R^2 the Model Summary Table provides some approximations.
- Logistic regression does not have an equivalent to the R-squared that is found in OLS regression; however, many researchers have tried to come up with one. There are a wide variety of pseudo-R-squared statistics (of which we discuss only two). Because this statistic does not mean what R-squared means in OLS regression (the proportion of variance explained by the predictors), interpreting this statistic must be done with great caution. The pseudo-R-squared statistics will have lower values than in multiple regression
- The SPSS output table below contains the *Cox & Snell R Square* and *Nagelkerke R Square* values, which are both methods of calculating the explained variation.

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	17.359[a]	.552	.737

a. Estimation terminated at iteration number 8 because parameter estimates changed by less than .001.

- Cox and Snells R-Square is an attempt to imitate the interpretation of multiple R-squared based on the likelihood, but its maximum can be (and usually is) less than 1.0, making it difficult to interpret.
- Here it is indicating that 55.2% of the variation in the Dependent Variable is explained by the logistic model.
- Nagelkerke's R^2 is part of SPSS output in the Model Summary table and is the most-reported of the R-squared estimates. Nagelkerkes R^2 will normally be higher than the Cox and Snell measure.
- In our case it is 0.737, indicating a moderately strong relationship of 73.7% between the predictors and the prediction.
- Nagelkerke R^2 is a modification of Cox & Snell R^2, the latter of which cannot achieve a value of 1. It is a enhancement of the Cox and Snell coefficient to assure that it can

vary from 0 to 1. Nagelkerke's R-Square will normally be higher than the Cox and Snell measure. For this reason, it is preferable to report the Nagelkerke R^2 value.

- However, they are interpreted in the same manner, but with more caution. Therefore, the explained variation in the dependent variable based on this model ranges from 55.2% to 73.7%, depending on whether you reference the Cox & Snell R^2 or Nagelkerke R^2 methods, respectively.