Nikhil Gopal

For this data, I would use cox-proportional hazards regression to model the relationship between failure time and health status. This intuitively makes sense as a healthier person should have a longer failure time, and we would want to see how much exactly being healthy can affect failure time. When using these survival methods, censored data can be included in the model.

The assumptions of the cox model include the hazard ratio assumption. The ratio of the hazards should be always proportional. Additionally, the model assumes that there are non-linear covariate relationships, and that there is not independent data.

Kaplan meir fits could also be used, but they require independence within the data. I’m assuming that failure time and transplant status are not independent, which would therefore mean we don’t have independent data. I still made the models in R to learn how to make them, but I will thus not be using them in my analysis.

If transplants did not exist, I would run a regression without including transplant time as a factor, I would simply run the regression with time as the outcome variable and health status as the predictor. If transplants did exist, I would still include them as a factor in my model even though there are some transplant times that happen after the failure time. In that case, the lack of a transplant could’ve affected the failure time, and thus it would be appropriate to include in the model. Thus, the failure time would still be the outcome, and health and transplant status would be parameters.

The analysis will be done in the R markdown file submitted for the computation set.