VetPath571 – Spring 2021
Final Exam

1) You have been provided with a dataset entitled 'whas.csv', containing patient information on 500 patients who were hospitalized for acute myocardial infarction (MI). Open this dataset, read the data dictionary, and answer the following questions:

- What is the average age of a male patient in the cohort? A female patient?
- What is the average length of stay in the hospital for patients?
- What percentage of patients survived their hospitalization?

2) Choose one of the outcomes in the dataset – either death during hospitalization, or death after all follow-up. Produce a Kaplan-Meier curve for this outcome.

3) Your PI is interested in the difference in survival for patients **experiencing their first vs. a recurrent MI.** For the same outcome you picked as above, estimate or visualize the difference in survival using any method you choose.

4) For that same outcome, conduct an exploratory data analysis looking at four variables of your choice. This may include looking at the distribution of the variables, their univariate association with the outcome, etc.

5) Your PI would now like you to revisit the first vs. recurrent MI question, adjusting for confounding. They are interested not only in whether a patient survives, but how long those who die took to die. You *must* use a survival analysis method, though which method is up to you.

Bonus (Adds 10% to your final score): Describe a problem in your area of interest, and how you would approach using some form of system science model (i.e. compartmental, network or agent-based model) to address this problem. Note: If you are currently using such a model as part of your research, choose something else.