

CHALLENGE

# Cox Proportional Hazard Regression

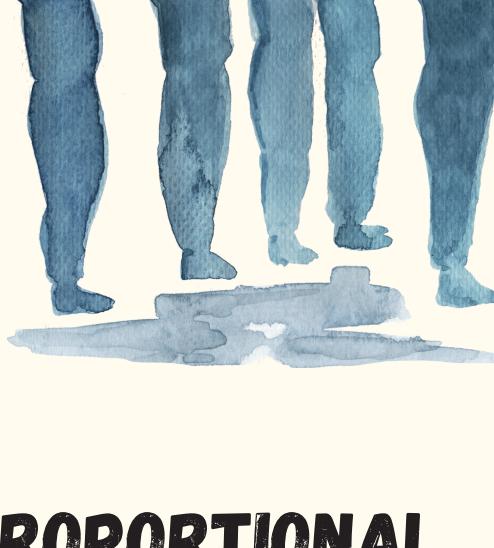
The Case Study is a randomized trial of two treatment regimens for lung cancer, a typical survival analysis data set. The patients were all Veterans.

The goal of the activity is to identify the drivers that improve the survival rate or life expectancy of the treatment, with a focus on a medical treatment.

## DUMMY VARIABLES

01

All variables must be in numeric format. The typical way is to transform categorical variables into dummy variables. However, we must be wary of the dummy variable trap, which is a multicollinearity issue that can arise.



## 02 COX PROPORTIONAL HAZARD REGRESSION



Straight to the modelling. You should have everything set up to do it. State the dependent variable, time and other drivers, and boom, you are done!

## 03 VISUALIZATION

Looking at the output of a regression model can be confusing and even overwhelming. Fortunately, there is an easy way to visualize the results, where we can see the impact and statistical significance.



Diogo Resende

