STAT460 – Homework 3
Due: Feb. 11 at the start of class.

1. These data record the level of atmospheric ozone concentration from eight daily meteorological measurements made in the Los Angeles basin in 1976. We have the 330 complete cases[1]. We want to find climate/weather factors that impact ozone readings. Ozone is a hazardous byproduct of burning fossil fuels and can harm lung function. The data set for this problem is:

Variable Name	Definition
ozone	Log Maximum Ozone
vh	Vandenberg 500 mb Height
wind	Wind Speed (mph)
humidity	Humidity (%)
temp	Sandburg AFB Temperature
ibh	Inversion Base Height
dpg	Daggot Pressure Gradient
ibt	Inversion Base Temperature
vis	Visibility (miles)
doy	Day of the Year

2. For the ozone data, find the ridge regression estimators for $\lambda = 1$ via

   (a) Least squares, solving $(\mathbb{X}^\top \mathbb{X} + \lambda I) = \mathbb{X}^\top Y$

   (b) `glmnet`

   Compare the two solutions. Don't forget about scaling $\mathbb{X}$ manually if needed.

---

[1]Note that this dataset violates some assumptions of linear regression. Do you know which one(s)? For this assignment, ignore this fact.