EDA Solar Power

Matt Oehler

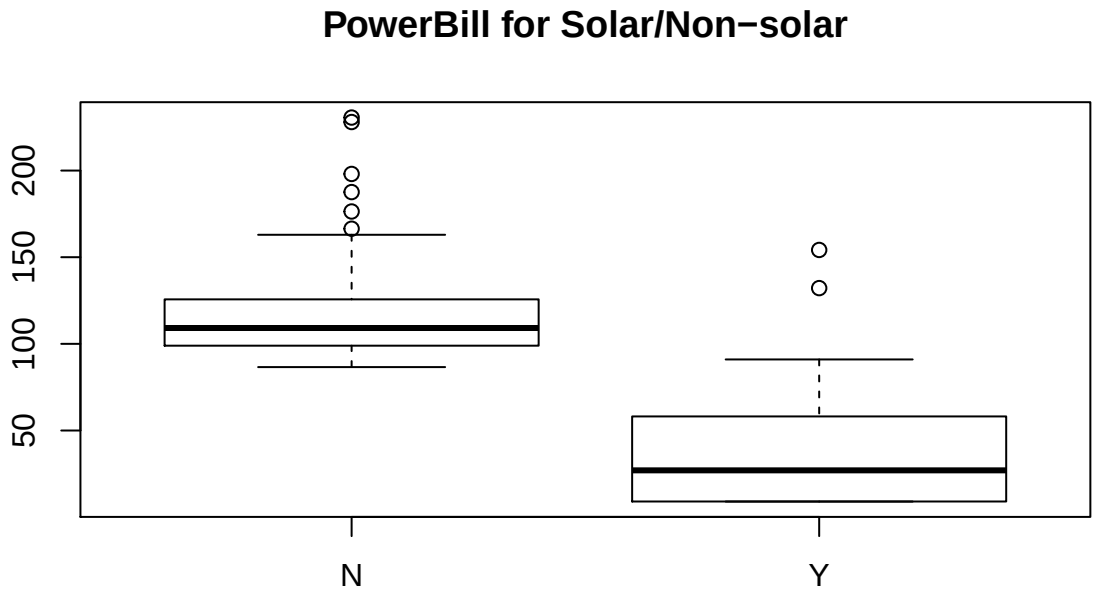
February 20, 2018

1. Description

Solar power is an environmentally friendly source of power as opposed to other sources such as fossil fuels. We are provided with the data of a customers power bills each month before and after they switched to solar power (51 months in total). The solar sales company is interested in how much this customer saves on average each month after switching to solar power, and how long it will take him to earn back the initial set up cost of installing solar panels based on his monthly savings.

2. Data Exploration

The dataset has 51 months worth of power bills for a customer (29 months of non-solar and 22 months of solar). Based on the boxplot shown below it appears that after switching to solar power, this customers power bill was significantly reduced.



3. Statistical Method

Using a standard linear regression model may be an appropriate way to approach this model. It will give us an estimate for the average amount saved per month by this customer. We can then use that monthly savings estimate to figure out approximately how long it would take the customer to recoup the money spent converting to solar power.

4. Things I don't know

Since our covariate isn't continuous, I don't know the best way to check the linearity assumption of our linear model. I'm also unsure if there is a clever way for us to use the date variable of the data set to enhance the analysis. Perhaps it is practical to look at seasonal trends in power bills.