Hi guys, John the Quant here. This is week 3 of the free data science bootcamp. If you didn’t start with week 1, a link should pop up here and there is one in the description as well.

Last week (look for the pop up here) we looked at the distribution, or shape of the data, talked about what we generally want the shape of the data to be, and figured out an intuitive way to make predictions from data and estimate how accurate those predictions would be. In general, we said that we want our data to be symmetrical, because then all the measures of center are close together and it removes a lot of ambiguity.

This week, we are going to figure out how to transform the data, to change its shape to make it more symmetrical. We’re going to learn a few ways to do that. Then, we’re going to do some more investigation of the data and we’ll talk about other properties we want the data to have and how we can further change the data to make it happen.

Let’s get started.

### NOTEBOOK

### END OF NOTEBOOK

Alright! We went through an entire dataset, column by column, and made each and every feature as, hopefully, informative as possible. We dealt with missing values either by dropping them when there weren’t very many, or by filling them in when there were just a few, or even using an Imputer model! We used square-root or log transformations or even Yeo-Johnson transformations to make features more symmetrical, and when we couldn’t make them symmetrical we bucketized them to give our model the best chance of learning from the data.

In this week’s mini-project, you’re gonna go through another dataset, making decisions on which transformations to apply to which columns, and more importantly, explaining your decisions with a story! You’ll find that explaining your decisions is often more important than making “perfect” decisions.

I hope you learned something this week! Leave your questions down below, and try to answer questions in the comments, too! You don’t have to know everything to help other people learn. If you learned something, please like this video, and if you want to continue learning with me, go ahead and subscribe. See you next week!