

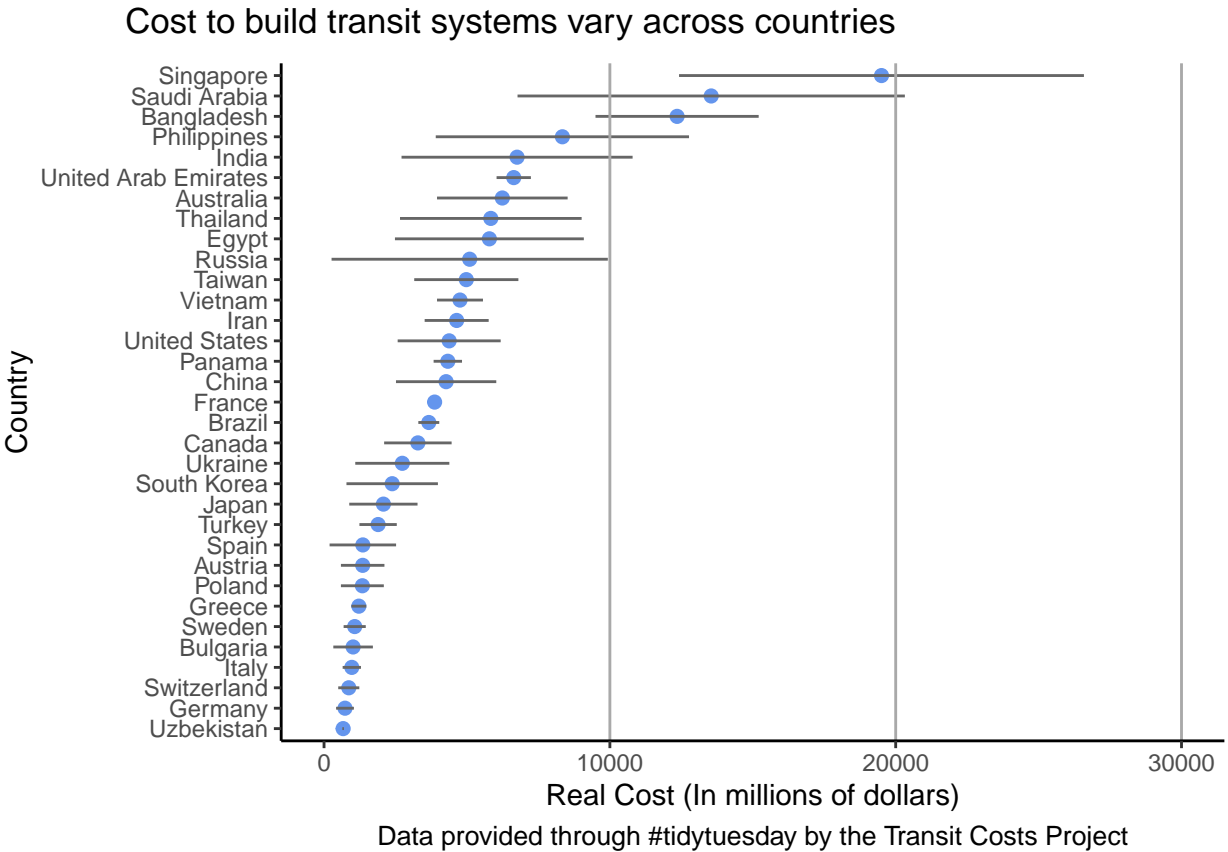
# EDLD 652 Homework

Elliott Doyle, Diana DeWald

2/8/2022

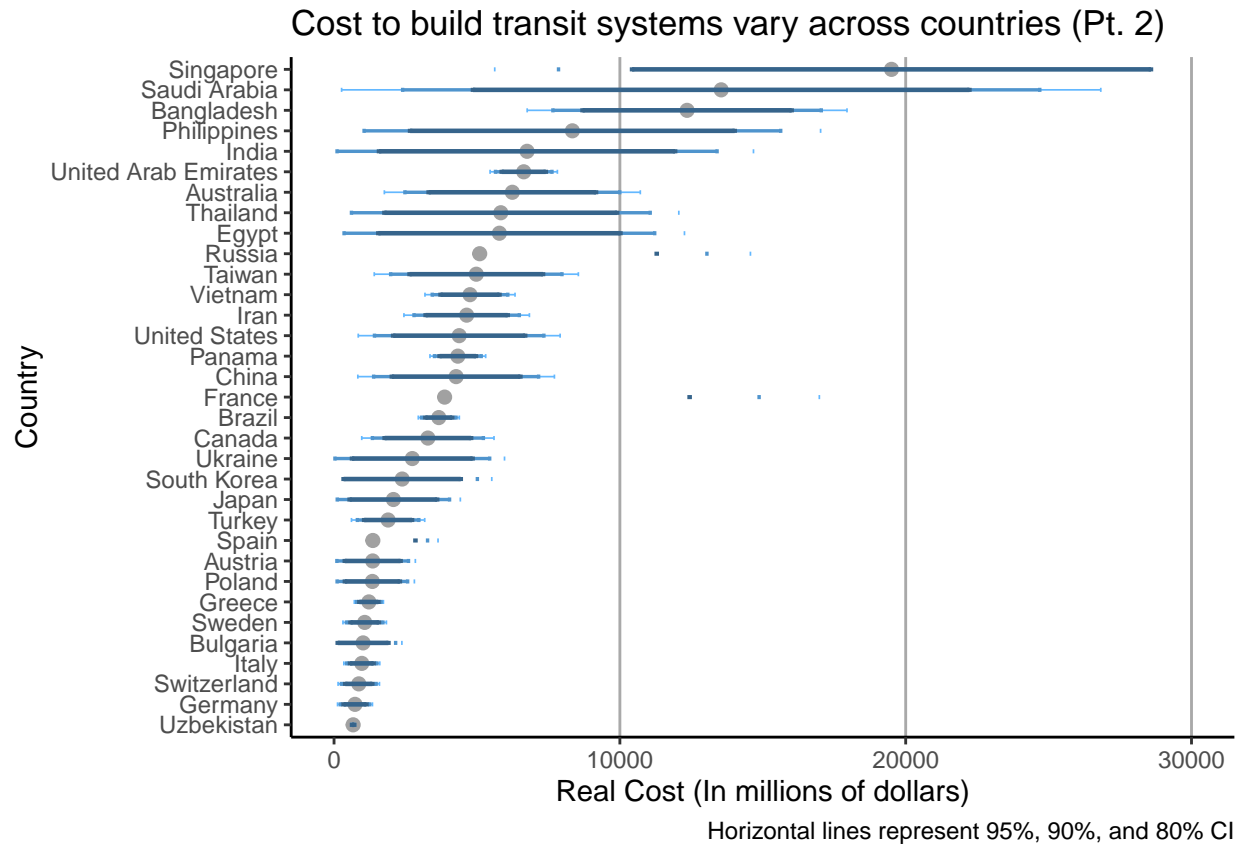
## Question 1:

Use the transit costs data to reproduce the following plot. To do so, you will need to do a small amount of data cleaning, then calculate the means and standard errors (of the mean) for each country.



## Question 2:

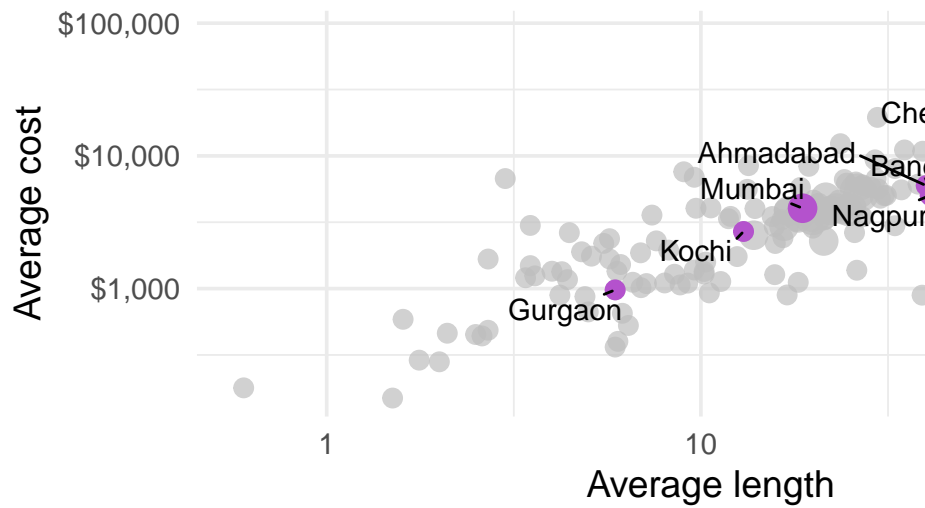
Visualize the same relation, but displaying the uncertainty using an alternative method of your choosing.



### Question 3:

Compute the mean length and real\_cost by city. Reproduce the following plot.

Longer transit systems tend to cost more  
India has among the most transit systems in the world



Number of transit systems

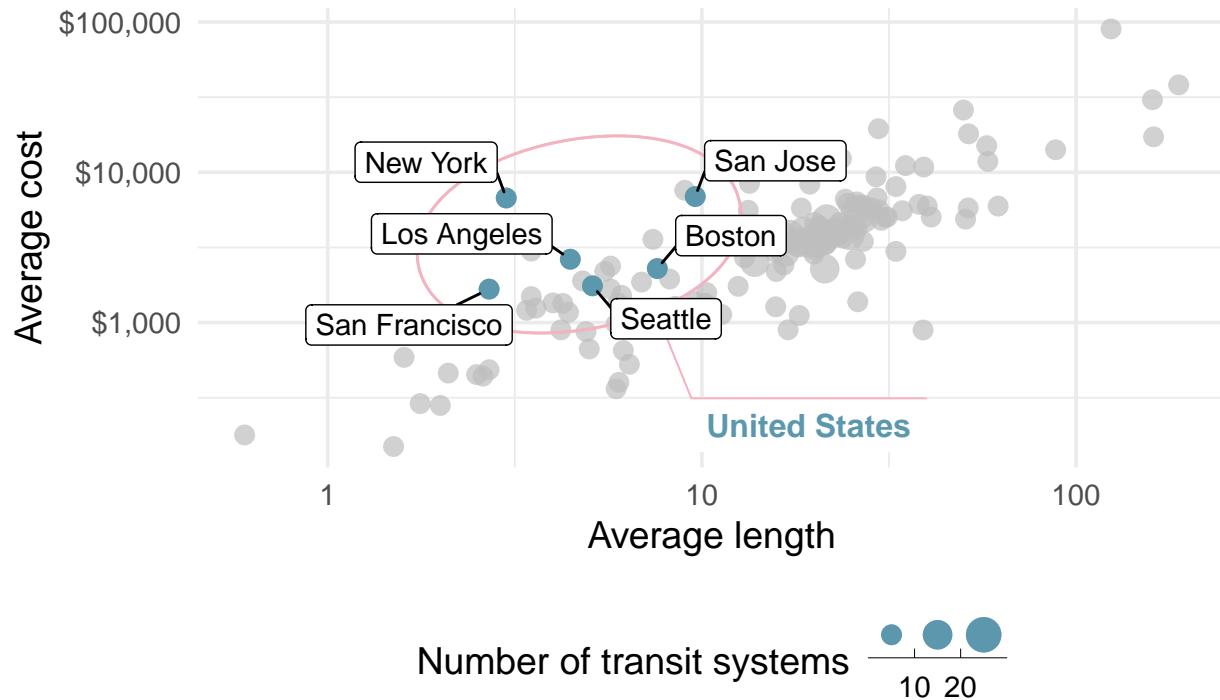
Note the log trans

Hint: Look into `scale_size_binned()`

#### Question 4:

Using basically the same data, reproduce the following plot. Note you'll need the `country_name` column in your dataset.

### Longer transit systems tend to cost more



Note the log transformations to the axes

#### Question 5:

Use the crime dataset to run the following code and fit the corresponding model. Note, it may take a moment to run.

The following plot displays a discretized version of the probability space for the differences in crime between the neighborhoods. Replicate this plot, but comparing the Barnum neighborhood to the Barnum-West neighborhood. Make sure to put the values in a data frame, and create a new variable stating whether the difference is greater than zero (which you will use to fill by). Note that you do not need to replicate the colors in the subtitle to match the balls, as I have, but if you'd like to you should likely use the `ggtext` package.

Note: Your probabilities will not directly correspond with the `p` values, which are essentially twice the probability you are displaying (because the test is two-tailed).

Probability of differential crime rates between neighborhood  
Regis compared to Barnum

