

# Problem Set 9

Your Name

Due date: Monday, November 16th

For this problem set we will use data from **New York City** that tested children under 6 years old for elevated blood lead levels (BLL). [You can read more about the data on their website].

About the data:

All NYC children are required to be tested for lead poisoning at around age 1 and age 2, and to be screened for risk of lead poisoning, and tested if at risk, up until age 6. These data are an indicator of children younger than 6 years of age tested in NYC in a given year with blood lead levels of 5 mcg/dL or greater. In 2012, CDC established that a blood lead level of 5 mcg/dL is the reference level for exposure to lead in children. This level is used to identify children who have blood lead levels higher than most children's levels. The reference level is determined by measuring the NHANES blood lead distribution in US children ages 1 to 5 years, and is reviewed every 4 years.

## Question 1

Recreate the below table with kable.

You will need to calculate the BLL per 1,000, filter for years 2015-2016, and rename the boroughs based on the following coding scheme:

- 1: Bronx
- 2: Brooklyn
- 3: Manhattan
- 4: Queens
- 5: Staten Island

## Question 2

Replicate the following bar chart. Since we want the graph to have an ascending order, we will need to factor borough\_id. <https://colorswall.com/palette/454/>:

- #ff6600: orange
- #003884: blue

### Question 3

Recreate the following line graph. You will need to combine the three measures of BLL to calculate a rate per borough per year. Your data frame, at it's smallest, will have 60 observations (12 years x 5 boroughs).