Feb 12 Readings

R: Handling Missing data

1. https://stats.idre.ucla.edu/r/faq/how-does-r-handle-missing-values/
This article talks about how missing data is treated differently in R as compared to other statistical computing languages. It also mentions the different options that come with NA.

2.http://r-statistics.co/Missing-Value-Treatment-With-R.html
This article talks about missing data treating and value imputations with some R code for examples.

3. Missingdata.pdf
I would like you to read the first 6 pages of this chapter (upto page 535). It is good to learn some basic theory behind missing-data imputations.

Feb 14 Readings

Multiple imputations using MICE.

1. https://datascienceplus.com/imputing-missing-data-with-r-mice-package/
Just like R has the tidyverse set of packages for data cleaning, it has the MICE package designed especially for data-imputations. Here is a quick tutorial with some R code.

2. http://dept.stat.lsa.umich.edu/~jerrick/courses/stat701/notes/mi.html
Some more mice (also 'with' and 'pool' functions) and multiple imputations.

3. http://web.maths.unsw.edu.au/~dwarton/missingDataLab.html
More missing data analysis with MICE. If you would like to follow along with this entire tutorial, you will first need to install the NHANES package which contains the NHANES dataset.

```
> install.packages("NHANES")
> library(NHANES)
```

EXTRA: Missingdata.pdf Finish reading this chapter. Some of the R code in there is not well written and looks pretty ugly. Just read more theory/concepts to understand why Data imputation can be complicated but is imperative for achieving accurate results (You will not be quizzed on this).