DATA SCIENCE COURSE PREWORK (~10 HOURS):

Congratulations on joining us at General Assembly for the Data Science course! We are so excited to have you and looking forward to working with you all soon!

In order to best prepare you for this course, we've gathered a list of resources that we'd like for your to complete <u>prior</u> to beginning the course on 6/25. So when you arrive on your first day, we will all be up to speed and be familiar with the vocabulary and programming languages necessary to succeed in this course.

Throughout this course, you will be exposed to both the R and the Python programming languages so in order to be most prepared we recommend going through the following tutorials starting with R and some review of the command line. This should take you around 10 hours to complete.

Please download/install the following:

> **R**
> Download link: http://cran.r-project.org/
> Online documentation: http://cran.r-project.org/manuals.html

> **ggplot2 library**
> Download link: http://ggplot2.org/

Please complete the following tutorials if any of these topics are new to you:

> **R**
> Code School Tutorial (~2 hours):
> http://tryr.codeschool.com/
>
> UCSB Tutorial (~2 hours):
> http://www.nceas.ucsb.edu/files/scicomp/Dloads/RProgramming/BestFirstRTutorial.pdf
>
> UPenn ggplot2 Tutorial (~2 hours):
> Complete up through "Layers" and read through the rest
> http://www.ling.upenn.edu/~joseff/avml2012/
>
> **Unix/Command line**
> Learn Command Line the Hard Way (~2 hours):
> http://cli.learncodethehardway.org/book/

Additional Optional Resources:

> http://cran.r-project.org/manuals.html
> http://data.princeton.edu/R/
> http://www.mac-terminal.com/

Please reach out if you have any questions about anything or need help.