**GALE - Data Science Assessment**

This challenge is designed to test your technical and analytics skills on a real-world data set.

We'll use the New York taxis data - in particular the Green taxis - as our experimental data set.

Please download Green taxi data for February 2016 from this site:

http://www.nyc.gov/html/tlc/html/about/trip\_record\_data.shtml

Please zip all code and outputs and email them back to (Recruiter)

1. Download the data, load it into your favorite statistical programing software or database.

Report the number of rows and columns that you've loaded.

2. Visualize trip distance by time of day in any way you see fit, any observations?

3. What are the most popular pickup locations on weekdays vs weekend?

4. I want to know where I can most easily get a cab. Recommend a pick up spot where i can

find a cab given my lat long

-------------------------------------------------------------------------------------------------------------------------------**My Answer to the Assessment Has Two Parts**:

For question 1-3, they are answered in *Green Taxi Case Study Part1.* I have attached the R markdown file in the submission and host the result via the link below.

<http://rpubs.com/couyang/379824>

For question 4, the question is answered in *Green Taxi Case Study Part2.* I have attached the R markdown (shiny) file in the submission. This part has the feature of inputting the longitude and latitude and it recommends 20 pick-up spots based on the input.