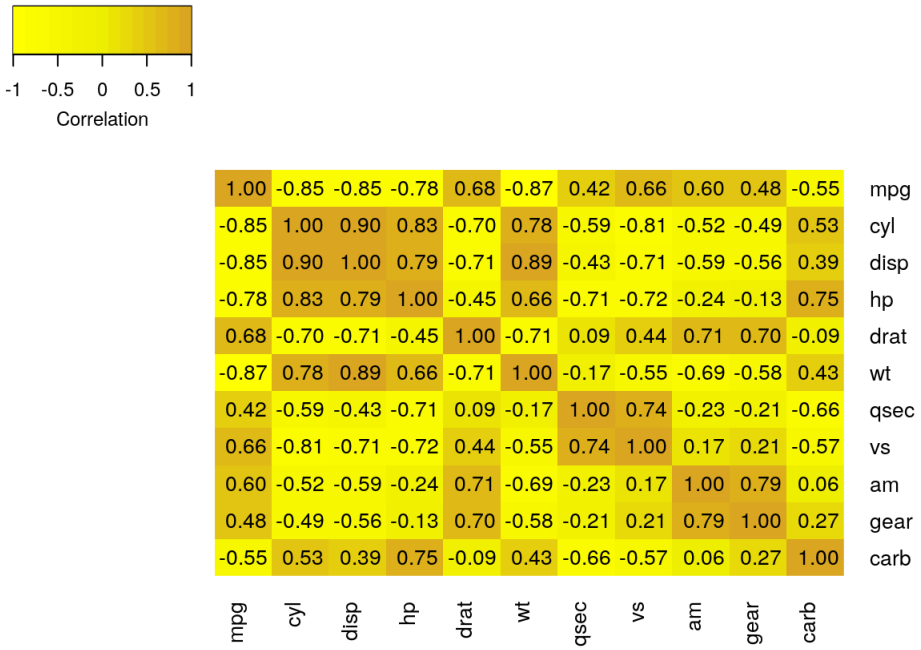
visualization

Qiuyi Lu

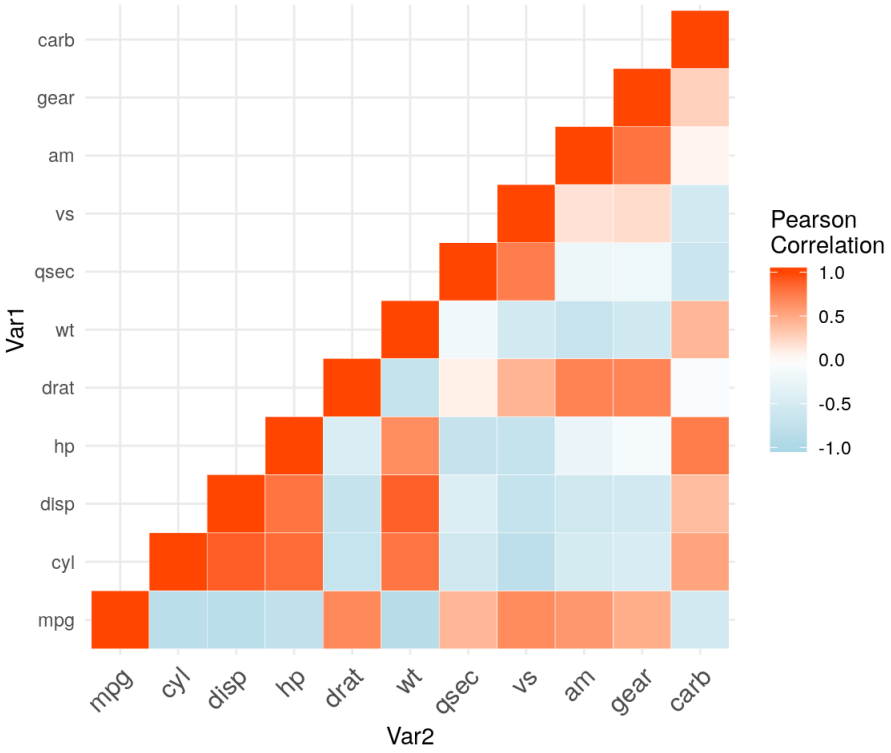
12/14/2019

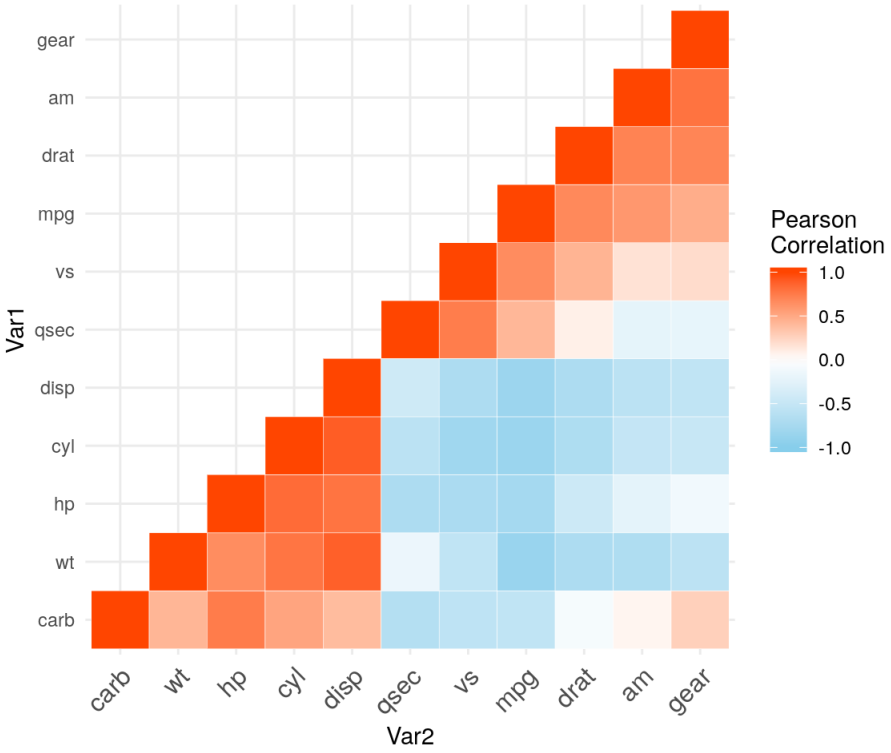
Heat Map:

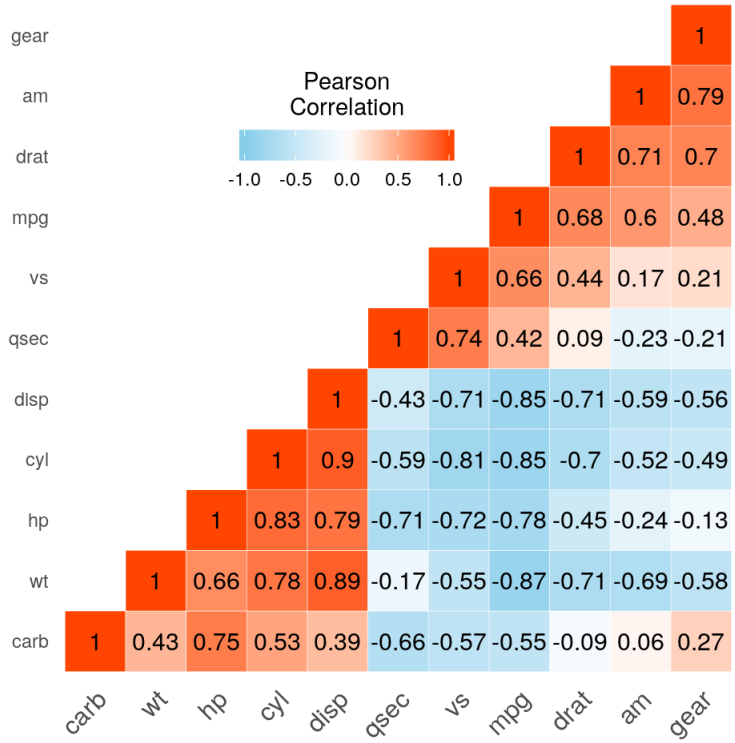
We can use heat map to see correlations among variables.
It's useful for selecting features when building models. For linear regression model, it can also help check multicollinearity.
Many bio companies use heat map to see the gene similarity among different species and could cluster species.




```
##       mpg       cyl       disp         hp       drat         wt        qsec
## mpg     1 -0.852162 -0.8475514 -0.7761684  0.6811719 -0.8676594  0.41868403
## cyl    NA  1.000000  0.9020329  0.8324475 -0.6999381  0.7824958 -0.59124207
## disp   NA        NA  1.0000000  0.7909486 -0.7102139  0.8879799 -0.43369788
## hp     NA        NA         NA  1.0000000 -0.4487591  0.6587479 -0.70822339
## drat   NA        NA         NA         NA  1.0000000 -0.7124406  0.09120476
## wt     NA        NA         NA         NA         NA  1.0000000 -0.17471588
## qsec   NA        NA         NA         NA         NA         NA  1.00000000
## vs     NA        NA         NA         NA         NA         NA          NA
## am     NA        NA         NA         NA         NA         NA          NA
## gear   NA        NA         NA         NA         NA         NA          NA
## carb   NA        NA         NA         NA         NA         NA          NA
##               vs         am       gear        carb
## mpg    0.6640389  0.5998324  0.4802848 -0.55092507
## cyl   -0.8108118 -0.5226070 -0.4926866  0.52698829
## disp  -0.7104159 -0.5912270 -0.5555692  0.39497686
## hp    -0.7230967 -0.2432043 -0.1257043  0.74981247
## drat   0.4402785  0.7127111  0.6996101 -0.09078980
## wt    -0.5549157 -0.6924953 -0.5832870  0.42760594
## qsec   0.7445354 -0.2298609 -0.2126822 -0.65624923
## vs     1.0000000  0.1683451  0.2060233 -0.56960714
## am            NA  1.0000000  0.7940588  0.05753435
## gear          NA         NA  1.0000000  0.27407284
## carb          NA         NA         NA  1.00000000
```

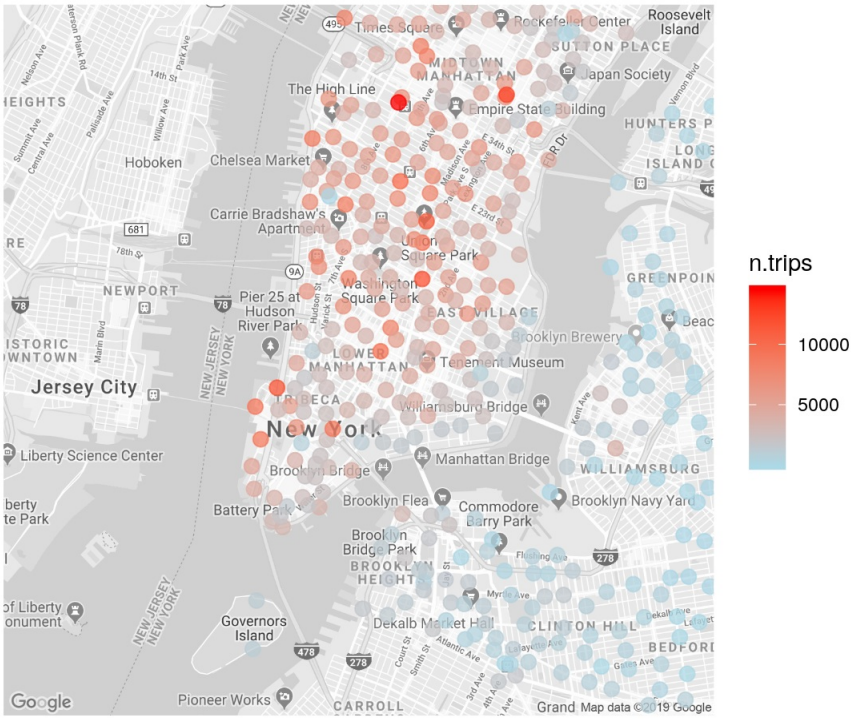

carb
gear
am
vs
qsec
wt
drat
hp
disp
cyl
mpg
Var1
mpg
cyl
disp
hp
drat
wt
qsec
vs
am
gear
carb
Var2
Pearson Correlation
1.0
0.5
0.0
-0.5
-1.0


gear
am
drat
mpg
vs
qsec
disp
cyl
hp
wt
carb
Var1
carb
wt
hp
cyl
disp
qsec
vs
mpg
drat
am
gear
Var2
Pearson Correlation
1.0
0.5
0.0
-0.5
-1.0



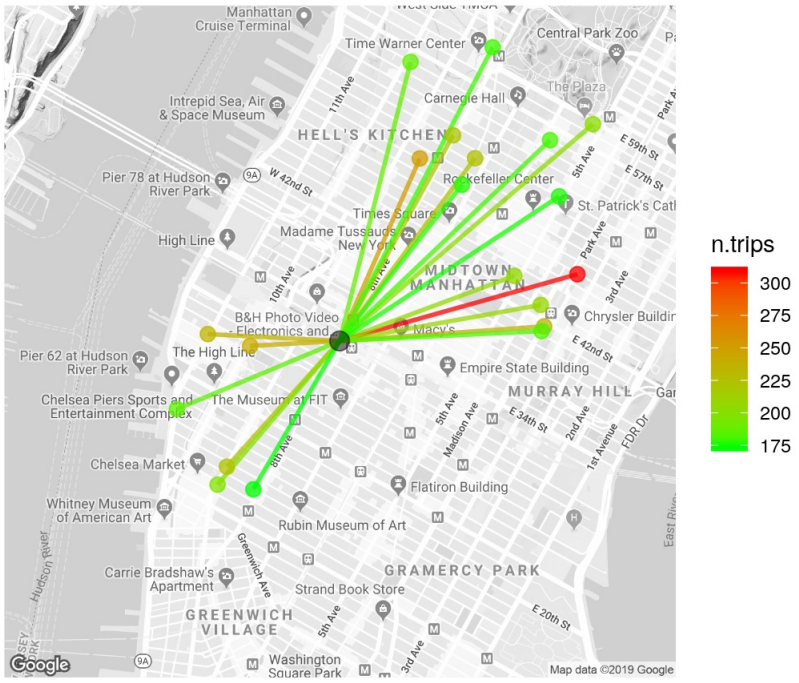
Geographic map

We can use map to find the popularity of some places. If we want to set a new station, we can check where do people take a lot of activities and set points there. Rideshare companies can use it to better allocate their drivers and pricing. Travelling/local service companies can recommend the place for users.

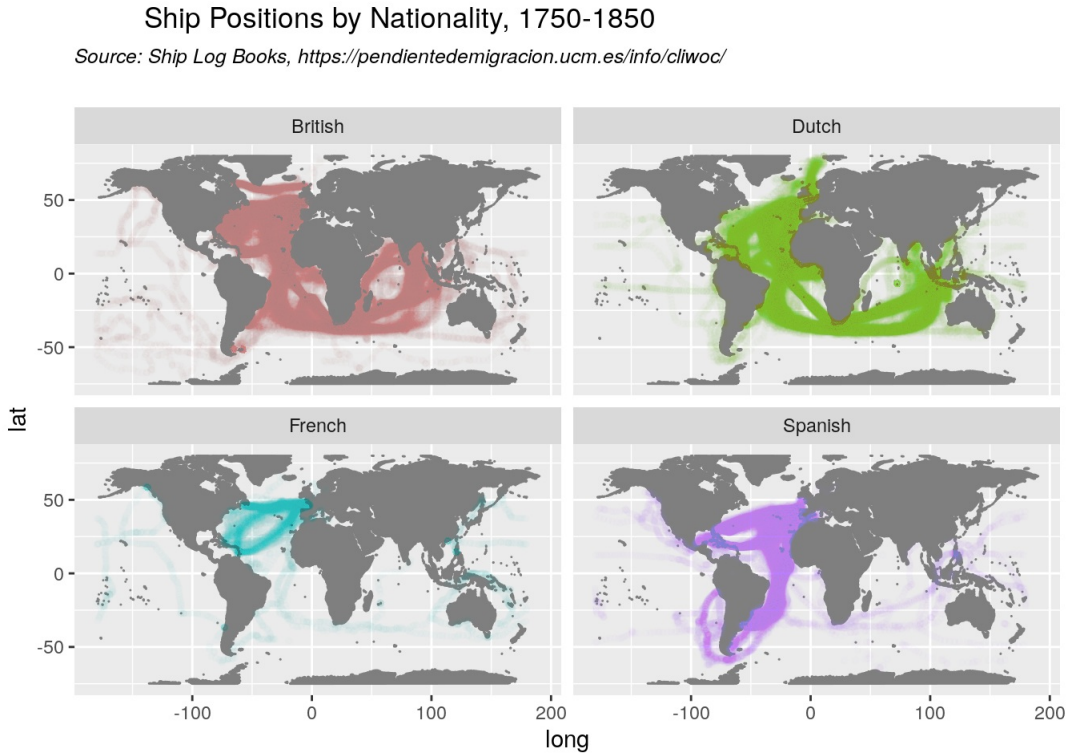



```
## # A tibble: 1 x 5
##   start.station.id   lat  long name            n.trips
##              <int> <dbl> <dbl> <chr>              <int>
## 1              521  40.8 -74.0 8 Ave & W 31 St   14498
```

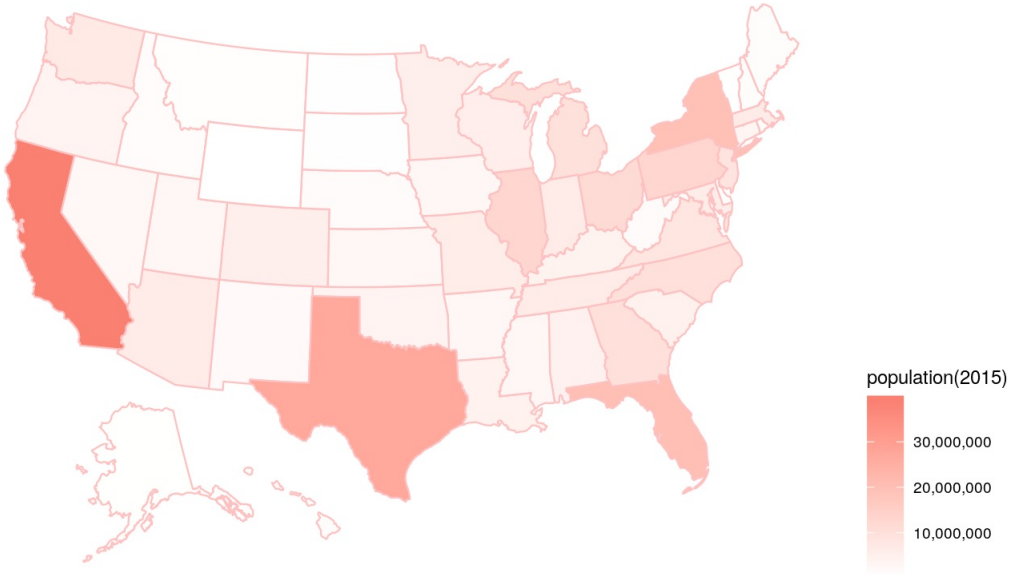
Top20 Trips starting at 8 Ave & W 31 St



World map:



US MAP



Plotly： A dynamic map. You can put your mouse on the area and get info.