```
In [94]:
             from IPython.core.display import display, HTML
             display(HTML("<style>.container { width:80% !important; }
          2
          3
             import pandas as pd
             import numpy as np
             import math
             import datetime as dt
          7
             import glob
          8
             import os
          9
         10 # pre setting
         pd.set_option('display.max_rows', 500)
         12 pd.set_option('display.max_columns', 500)
         13
             pd.set_option('display.width', 1000)
```

## Out[95]:

	Unnamed: 0	Date	Bus	Route_num	Route_road_lst	route_len
0	0	2018- 02-12	B297	1	['Hoyt St', 'Dorchester Ave', 'Columbia Rd', '	55
1	1	2018- 02-12	B306	1	['Industrial Dr', 'W Milton St', 'Vallaro Rd',	75
2	2	2018- 02-12	B308	1	['Industrial Dr', 'Milton St', 'Hyde Park Ave 	55
3	3	2018- 02-12	B310	1	['Industrial Dr', 'Neponset Valley Pkwy', 'Tur	49

```
In [57]:
              ## analysis on Q2
              df = df.groupby("Bus")['time_to_dist'].mean().reset_index()
              df = df.sort_values(by=['time_to_dist'],ascending=False).tage
              df
```

## Out[57]:

	Bus	time_to_dist
571	MS163	6.047476
594	MS186	5.766686
695	WB930	5.612657
558	MS150	5.494232
612	MS204	5.400399
573	MS165	5.220682
557	MS149	3.800818
580	MS172	3.645314
570	MS162	3.325651
569	MS161	3.180137

```
In [104]:
```

```
## analysis on Q1
 2
 3
   list q1 = []
   list_q1_2 = []
4
 5
    list_bus_q1 = []
    for file in os.listdir("X:/506Final/Q1_results"):
        all_files = glob.glob("X:/506Final/Q1_results/"+ file
 7
        for filename in all_files:
 8
9
            bus name = filename.replace(("X:/506Final/Q1 resul
            output = pd.read_csv(filename, index_col=None, hea
10
11
            list_q1.append(output[["road_name","avg_speed","sr
12
13
    average_df = pd.concat(list_q1)
14
15
```

12/9/2019 Q1\_Q2\_analysis

In [108]:

```
print(average_df.shape)
average_df = average_df.groupby("road_name")['avg_speed','
avg_speed_df = average_df[average_df['avg_speed']>0].sort_
# avg_speed_df = average_df.sort_values(by=['avg_speed'])
avg_speed_df
```

4124	Woodstock Ave	0.434200	104.000000
2148	May St	0.434200	104.000000
1679	I-295 N / US-1 N / US-1A N	0.868400	104.000000
1177	Fairfax Rd	0.434200	104.000000
4038	Wiget St	0.434200	104.000000
867	Dakin Rd	0.409456	104.091567
2554	Paul Revere Rd	0.354040	104.600000
2738	RT-107 / Sgt James Ayube Memorial Dr	0.701400	105.000000
1683	I-295 S / US-1 S / US-1A S	0.876750	105.000000
1945	Legends Hwy / Accolon Way	0.459632	105.474136
2752	RT-110 E / RT-113 E / Lowell St	0.528628	105.514603
1719	I-95 N / US-1 S	0.971169	105.734266