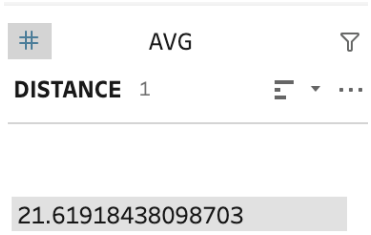
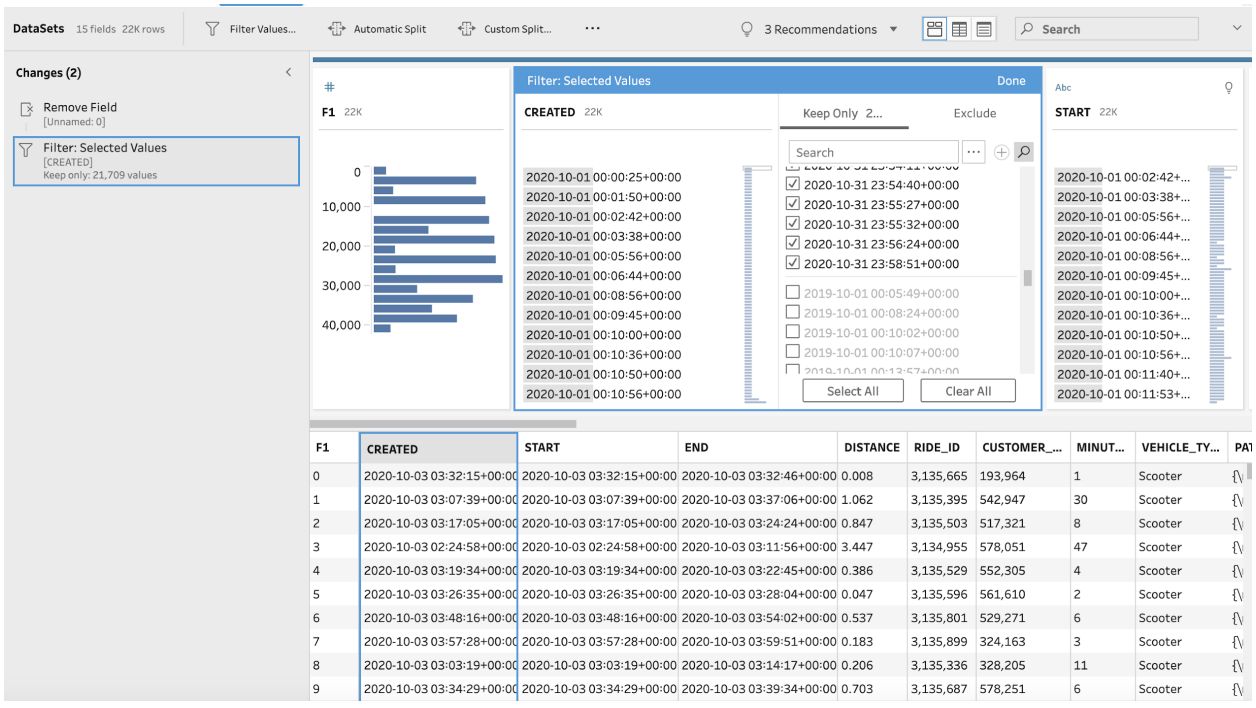
ABSTRACT DATA E-SCOOTER TEAM36

1 a) Using Tableau we were able to filter out dates to just October 2020 and calculate the average distance traveled during that month. As shown in the second image all october 2020 dates are selected when calculating avg distance.Our program and computation :
Issue: Does not find for just confined to campus, working on geopandas to figure an approach





F1	CREATED	START	END	DISTANCE	RIDE_ID	CUSTOMER_...	MINUT...	VEHICLE_TY...	PA
0	2020-10-03 03:32:15+00:00	2020-10-03 03:32:15+00:00	2020-10-03 03:32:46+00:00	0.008	3,135,665	193,964	1	Scooter	{\
1	2020-10-03 03:07:39+00:00	2020-10-03 03:07:39+00:00	2020-10-03 03:37:06+00:00	1.062	3,135,395	542,947	30	Scooter	{\
2	2020-10-03 03:17:05+00:00	2020-10-03 03:17:05+00:00	2020-10-03 03:24:24+00:00	0.847	3,135,503	517,321	8	Scooter	{\
3	2020-10-03 02:24:58+00:00	2020-10-03 02:24:58+00:00	2020-10-03 03:11:56+00:00	3.447	3,134,955	578,051	47	Scooter	{\
4	2020-10-03 03:19:34+00:00	2020-10-03 03:19:34+00:00	2020-10-03 03:22:45+00:00	0.386	3,135,529	552,305	4	Scooter	{\
5	2020-10-03 03:26:35+00:00	2020-10-03 03:26:35+00:00	2020-10-03 03:28:04+00:00	0.047	3,135,596	561,610	2	Scooter	{\
6	2020-10-03 03:48:16+00:00	2020-10-03 03:48:16+00:00	2020-10-03 03:54:02+00:00	0.537	3,135,801	529,271	6	Scooter	{\
7	2020-10-03 03:57:28+00:00	2020-10-03 03:57:28+00:00	2020-10-03 03:59:51+00:00	0.183	3,135,899	324,163	3	Scooter	{\
8	2020-10-03 03:03:19+00:00	2020-10-03 03:03:19+00:00	2020-10-03 03:14:17+00:00	0.206	3,135,336	328,205	11	Scooter	{\
9	2020-10-03 03:34:29+00:00	2020-10-03 03:34:29+00:00	2020-10-03 03:39:34+00:00	0.703	3,135,687	578,251	6	Scooter	{\

While working on Pandas:

How to find most common Starting points and Ending destinations

STARTING POINTS most common [38.992447, -76.949093] appeared 20 times in the dataset making it the most popular starting point.

In [104]:
```
import numpy as np
s=df[['START LAT','START LONG']]
print(s)
x=s.value_counts()
x
```

```
       START LAT  START LONG
0      38.990656  -76.934163
1      38.982087  -76.938181
2      38.979613  -76.937600
3      38.982470  -76.943553
4      38.982084  -76.938116
...          ...         ...
40320  38.988110  -76.943936
40321  38.986389  -76.935917
40322  38.996377  -76.940231
40323  38.991956  -76.944153
40324  38.990645  -76.934167

[40325 rows x 2 columns]
```

Out[104]:
```
START LAT  START LONG
38.992447  -76.949093    20
38.992699  -76.935208    17
38.988723  -76.940666    14
38.982561  -76.943550     7
38.984513  -76.941714     5
                        ..
38.991116  -76.942908     1
38.991114  -76.933588     1
           -76.946452     1
38.991113  -76.932574     1
38.955716  -76.939235     1
Length: 40039, dtype: int64
```

ENDING POINTS:
most common Ending Point [38.992447, -76.949093] appeared 20 times in the dataset making it the most popular starting point.

```
In [105]: c=df[['END LAT','END LONG']]
          print(c)
          z=c.value_counts()
          z
```

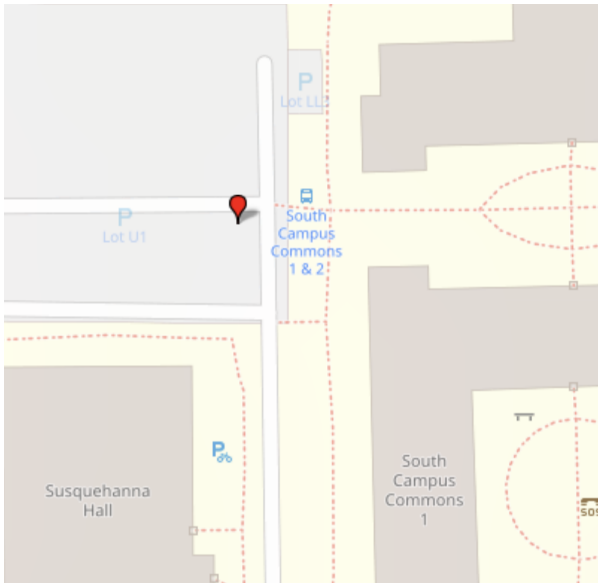
```
            END LAT   END LONG
0         38.990739 -76.934285
1         38.992849 -76.949165
2         38.992676 -76.933335
3         38.982476 -76.943546
4         38.986101 -76.943257
...             ...        ...
40320     38.992061 -76.934104
40321     38.992196 -76.933818
40322     38.992842 -76.945600
40323     38.987199 -76.940216
40324     38.982452 -76.936088

[40325 rows x 2 columns]

Out[105]: END LAT    END LONG
          38.992447  -76.949093    20
          38.992699  -76.935208    17
          38.988723  -76.940666    14
          38.982561  -76.943550     7
```

Destination of CUSTOMER ID: 578051
Map used with an online visualization (as we were having issues with geopandas and ipyleaflet libraries) https://www.gpsvisualizer.com/map?output_leaflet

```
In [56]: df.iloc[3:4]
```

Out[56]:

	VEHICLE_TYPE	PATH	TIMESTAMPS	START LONG	START LAT	END LONG	END LAT
	Scooter	{\n "coordinates": [\n [\n -76.943553...	[\n "2020-10-03T02:25:04Z",\n "2020-10-03T02...	-76.943553	38.98247	-76.943546	38.982476



START of travel

Issue: Ipyleaflet not showing map on Jupyter Notebook:

```
In [29]: import pandas as pd
         import geopandas as gpd
         from shapely.geometry import Point
         %matplotlib inline

 In [ ]: from shapely.geometry import Polygon

In [24]: import ipyleaflet

 In [ ]: from ipyleaflet import Map
         Map(center=[38.990656,-76.934163],zoom=10)

In [25]: from ipyleaflet import (Map,DrawControl)
         m=Map(center=[38.990656,-76.934163],zoom=10)
         d=DrawControl(circle={'shapeOptions':{'color':'#0000FF'}})
         m.add_control(d)
         display(m)
         #not displaying map

In [26]: from ipyleaflet import (Map,DrawControl)
         m=Map(center=[38.990656,-76.934163],zoom=10)


In [31]: from ipyleaflet import Map, Marker

         center = (52.204793, 360.121558)

         m = Map(center=center, zoom=15)

         marker = Marker(location=center, draggable=True)
         m.add_layer(marker);

         display(m)
         type(m)

Out[31]: ipyleaflet.leaflet.Map
```

Approach moving forward:

- We have a meeting with DSA tomorrow (Thursday 02/25/2021) to get further assistance with coding/data analysis. We plan to continue to work with Python and various libraries (Pandas, ipyleaflet) to filter and sort the data. And if needed, use other geo-visualization tools to actually graph and chart our data and show a map of exit/entry points.
- We plan to parse the data by year (2019/2020) first and explain the different trends found in each data set and see if there are any interesting finds.