For 7/10/2017 bi-weekly meeting

- What does t-SNE actually do?
  Our goal is to take a set of points in a high-dimensional space and find a faithful representation of those points in a lower-dimensional space, typically the 2D space. Essentially it is mainly a ***data exploration*** and ***visualization*** technique.

  Instead of trying to preserve

  t-SNE can be used in the process of classification and clustering by using its output as the input feature for other classification algorithms (similar to PCA).

- Where and when to use t-SNE?
  Use it for exploratory data analysis. It will give you a good sense of patterns hidden inside the data. It can also be used as an input parameter for other classification& clustering algorithms.

- Distances between well-separated clusters in a t-SNE plot might ***not mean anything***
  (You cannot see relative sizes of clusters in a t-SNE plot)

  What's going on?
  The t-SNE algorithm adapts its notion of "distance" to regional density variations in the data set. As a result, it naturally expands dense clusters and contracts sparse ones.

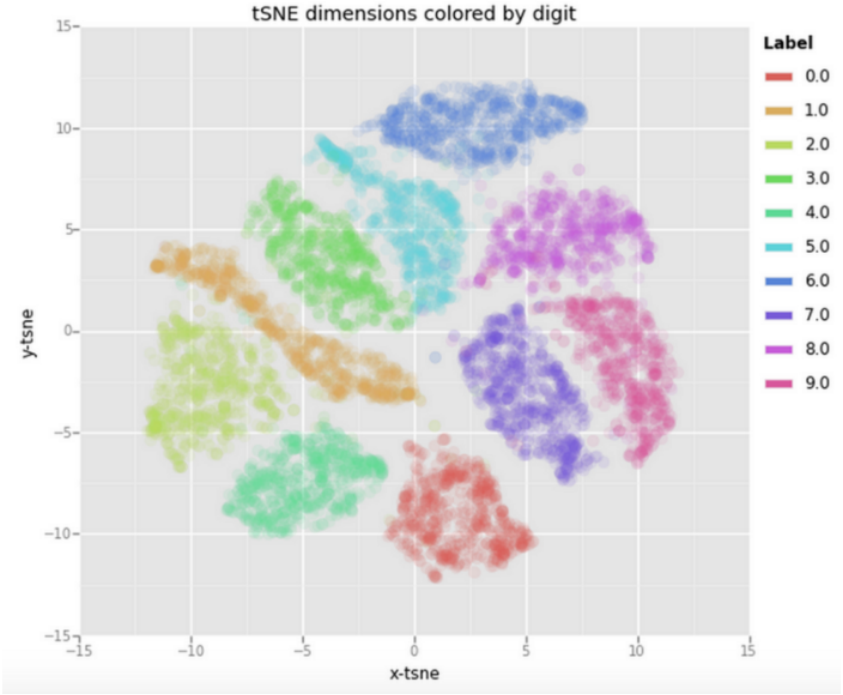
- The cloud of points was generated randomly: those "clumps" ***aren't meaningful***. Recognizing these clumps as random noise is an important part of reading t-SNE plots.
- For topology, we may need more than one plot.
- Its flexibility makes it tricky to interpret.

- You can sometimes see some shapes. T-SNE tends to expand denser regions of data.

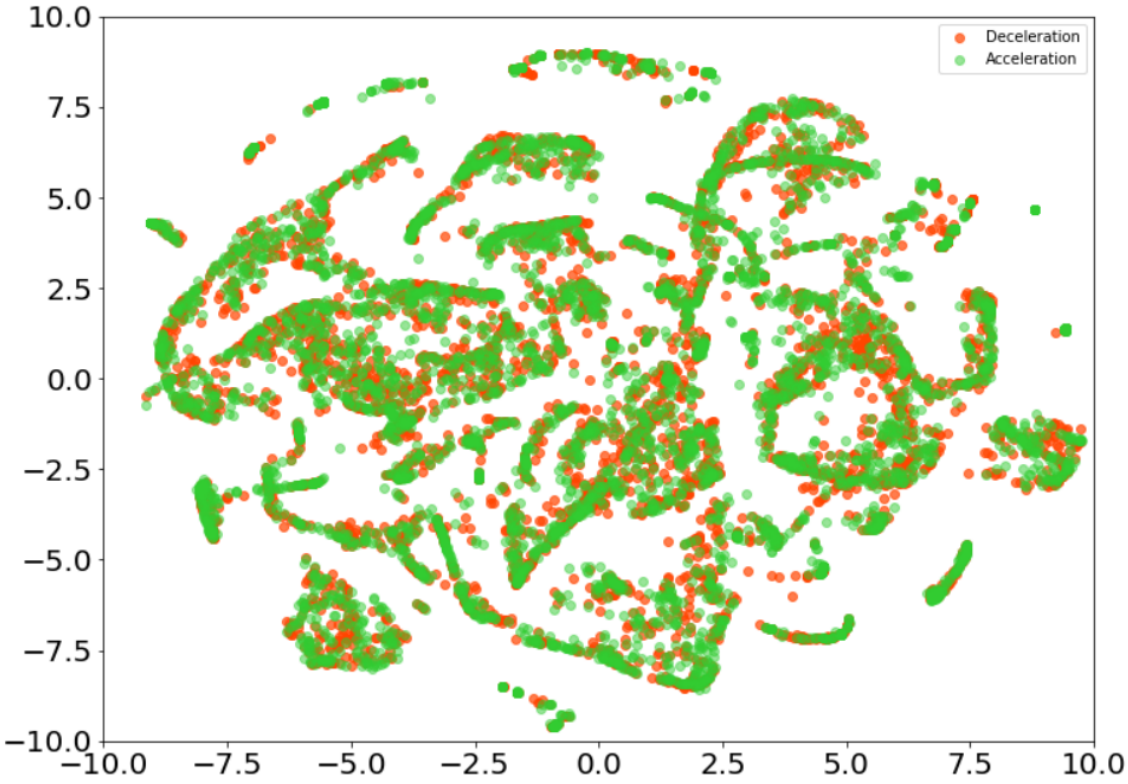
Variables used for speed:

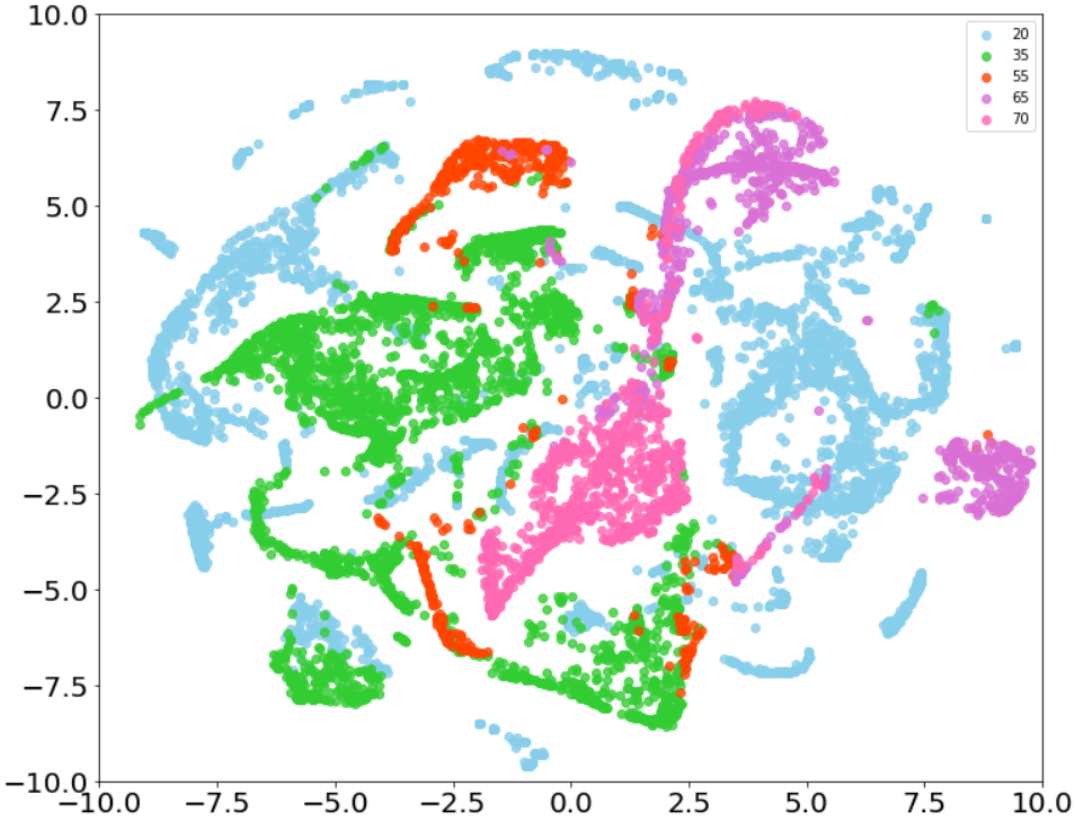
```
['traffic_den',
 'avg_speed',
 'road_typeBusiness_District',
 'road_typeFreeway',
 'road_typeResidential_Road',
 'stop_ind',
 'stop_grp_cnt',
 'latG',
 'lonG',
 'speed',
 'ang_speed_gyro',
 'lon_delta',
 'inc_mileage',
 'avg_latG',
 'avg_latG_mag',
 'algorithmCurve_in_the_road',
 'algorithmLane_Change',
 'algorithmNormal_Driving',
 'algorithmNormal_Turn']
```

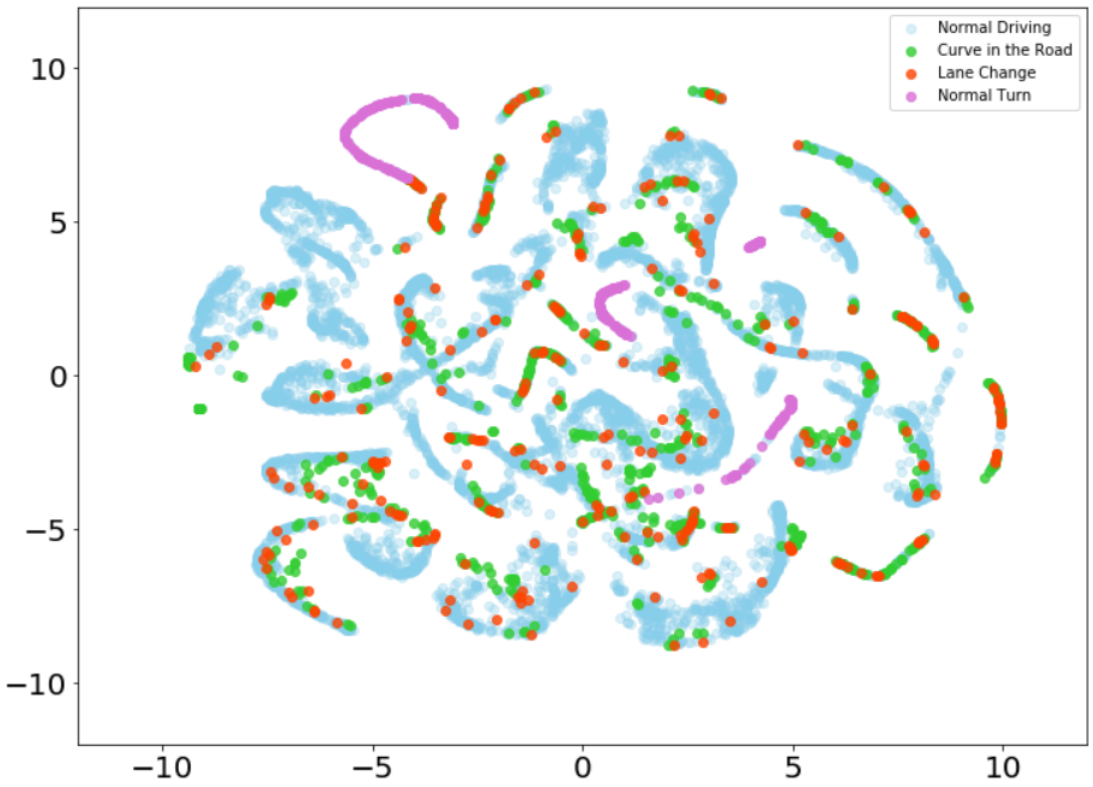
Ideal result:



Digits are very clearly clustered in their own little group. If we would now use a **clustering algorithm** to pick out the separate clusters, we could probably quite accurately assign new points to a label.


Deceleration
Acceleration
10.0
7.5
5.0
2.5
0.0
−2.5
−5.0
−7.5
−10.0
−10.0
−7.5
−5.0
−2.5
0.0
2.5
5.0
7.5
10.0


20
35
55
65
70
10.0
7.5
5.0
2.5
0.0
−2.5
−5.0
−7.5
−10.0
−10.0
−7.5
−5.0
−2.5
0.0
2.5
5.0
7.5
10.0




```
['latG', 'lonG', 'ang_speed_gyro', 'avg_latG', 'avg_latG_mag']
```

latG: G-Force on the car in the left-right directions at this second
lonG: G-Force in the forward-backward directions at this second
Ang_speed_gyro: the change in orientation of the car in that second, measured in degrees/second, angular speed for gyroscope
Avg_latG: rolling latG in 10 second (created in last semester)
Avg_latG_mag: rolling absolute value of latG in 5 second (created in last semester)