

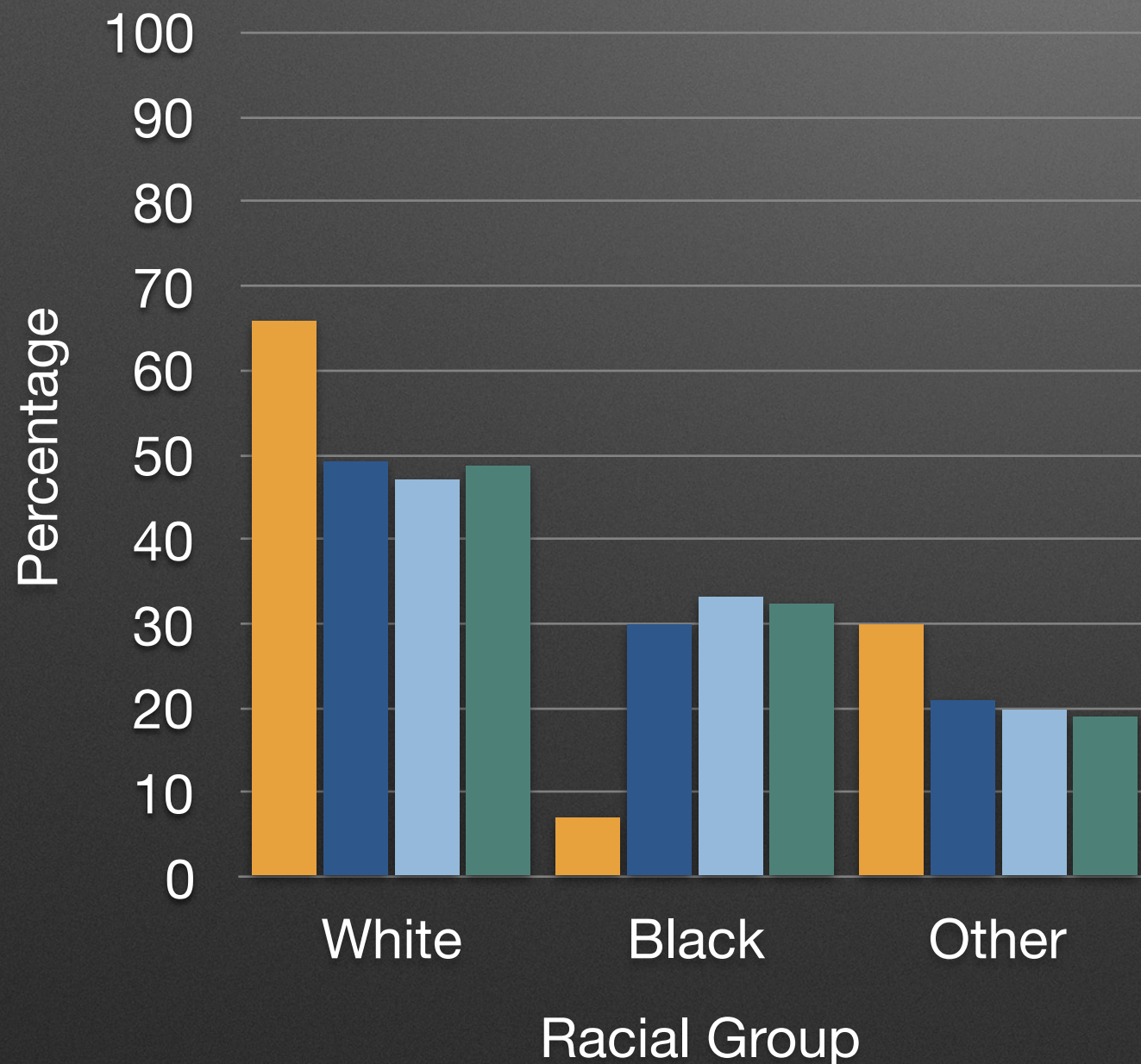
What Do You Mean, You Were Arrested?

A Look At Seattle Terry Stop Data

Coronet Consultants, LLC

Population vs Arrests

Population /
Stops
General Arrest
Physical Arrest

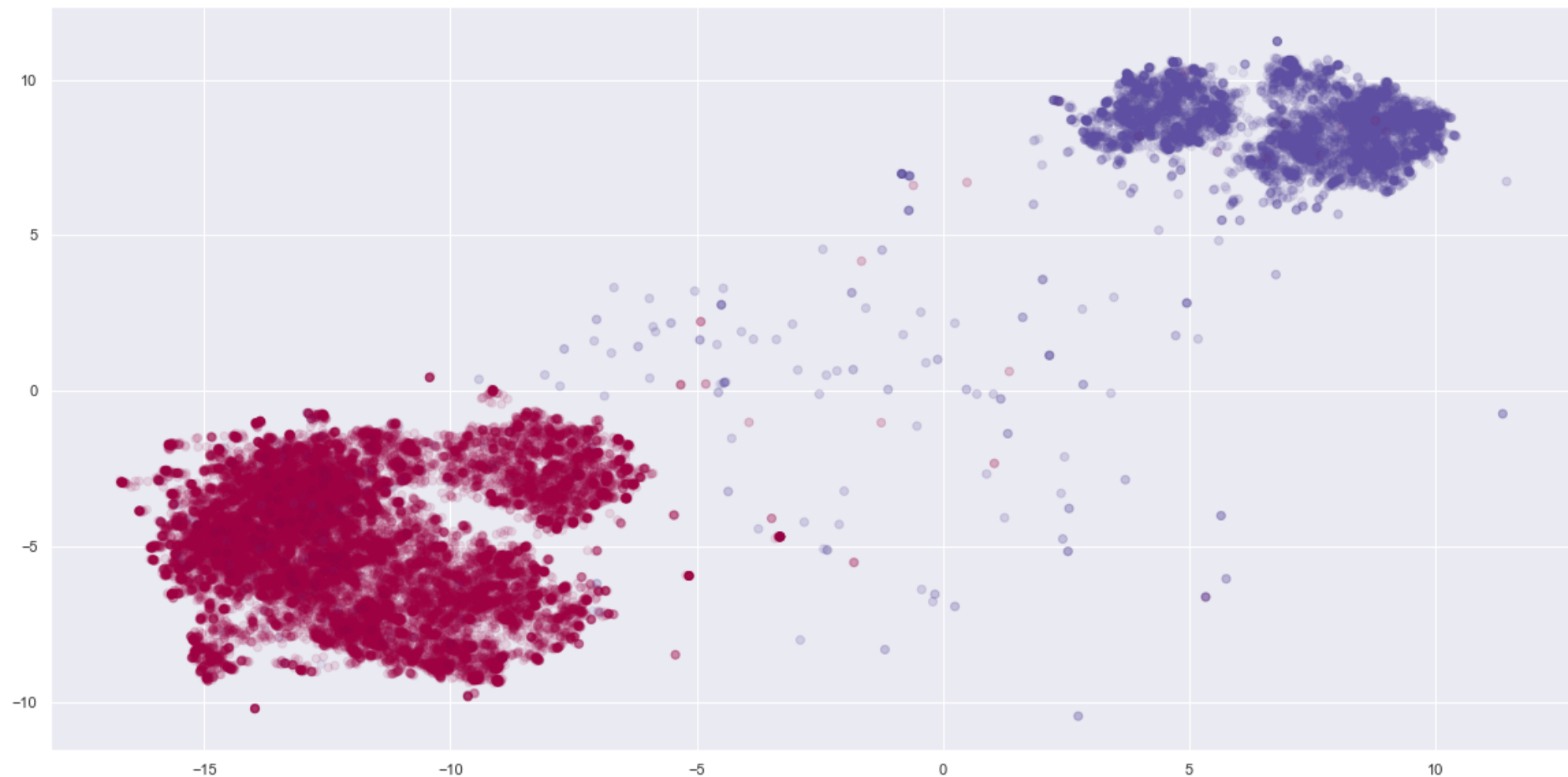


Total # of Stops = 39,406

Total General Arrests = 9578

Total Physical Arrests = 1369

Data covers the period of
March 2015 - January 2020



What Is Up With This Data?

Overlapping clusters indicate weak signal / noisy data.

Prediction Challenges

- Initial target was physical arrest
 - 91% accuracy from best model
 - 97% accuracy from predicting “no physical arrests were made”
- Final target was general arrest
 - 83.8% accuracy from best model
 - 80.4% accuracy from predicting “no general arrests were made”
- “Noisy” data/weak signal (complex/difficult problem)

Next Steps

We recommend the following steps:

1. Focus on separation of data into multiple subsets to reduce noise / increase signal strength
2. Using subsets, solve the general arrest problem
 - Subset target: $\geq 90\%$ accuracy
3. Once step 2 is successful, solve the physical arrest problem.

Estimates

Create data subsets and models for general arrest target

- Total cost \$40,000
 - ~4 weeks x 2 junior data scientists @ \$100/hour = \$32,000
 - ~1 week x 1 senior data scientist @ \$200/hour = \$8,000

Solving the physical arrest problem would incur additional costs that are not part of this estimate.

In Conclusion

- With the data we have, we struggle to accurately predict whether a Terry Stop will result in an arrest.
- Better data would make the problem easier to solve:
 - Subject economic status
 - Subject education level
 - Subject's previous criminal record, if any
 - Detailed stop location information
 - Length of officer service at time of stop