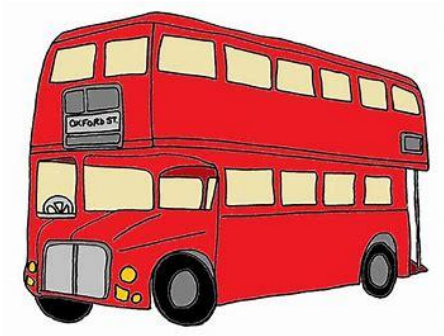
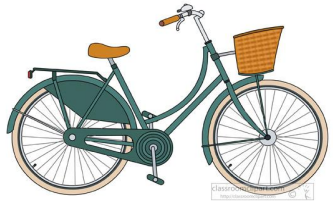
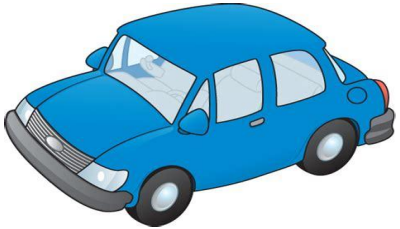
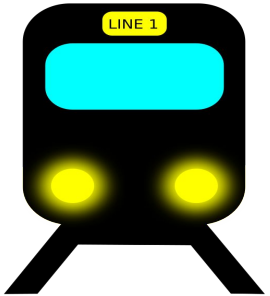
Changing the Commute:
Using census data to determine optimal targeting for commute preferences









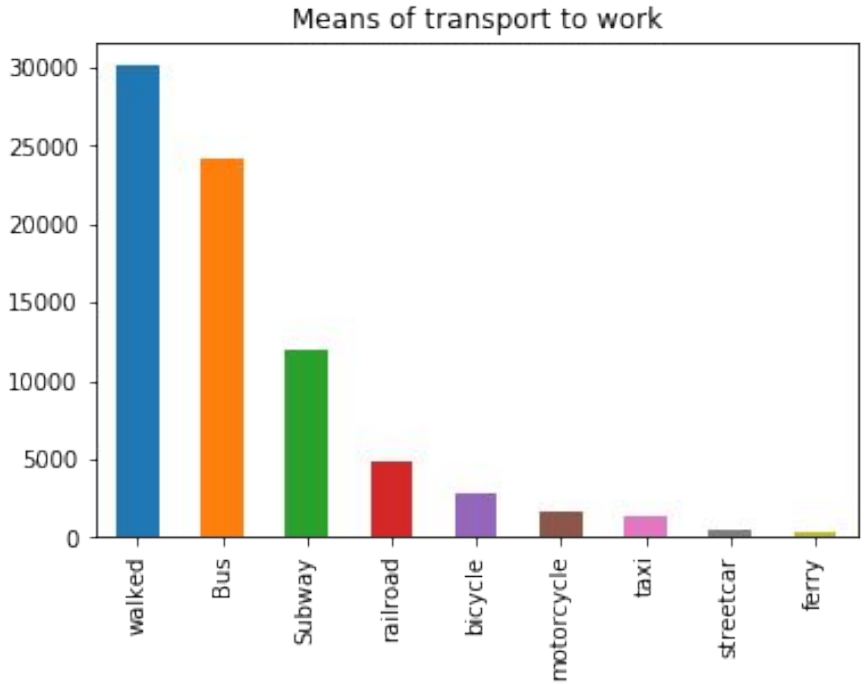
Jennifer Steele
08/31/2018

Who is most likely to adopt an alternate means of transport?

- Using 1 million rows of census data for employed adults, I use un-classified clustering algorithms to determine who is most likely to adopt an alternate means of commute
- Variables used are: age, US born, depart time, income, >40 hours week, Gender, Marital status, occupation, industry, children, years of schooling
- Of the 1 million rows, 86% commute to work by car

- Looking at who is likely to adopt an alternate means can help companies allocate resources, as well as government to build infrastructure and estimate ridership

86% of employed adults commute to work by car



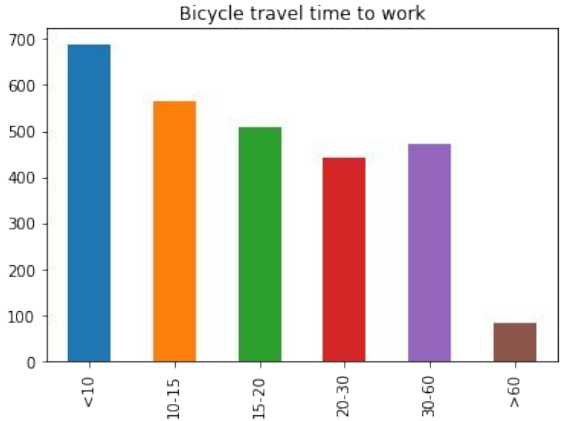
Of those who don't drive, walking, bus, subway and railroad are next in order of popularity

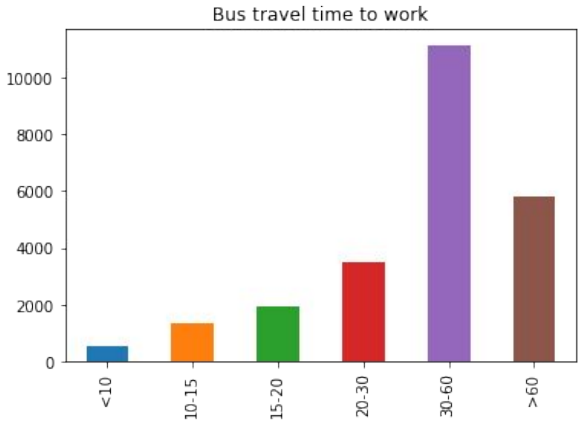
How can we determine who is most readily induced to move out of their cars and into some other means of transportation?

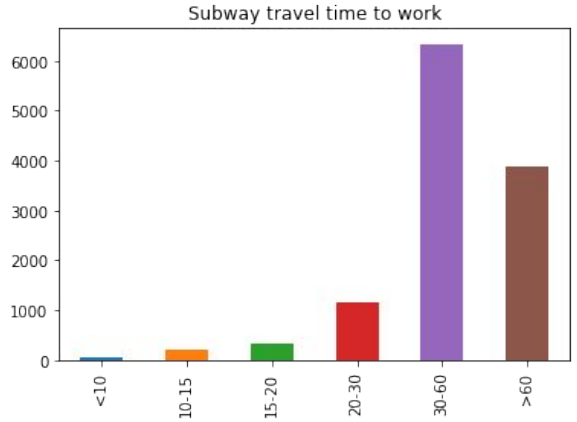
Distribution of travel time varies widely by means of transport

Walkers have the shortest commute, while subway riders have the longest on average.

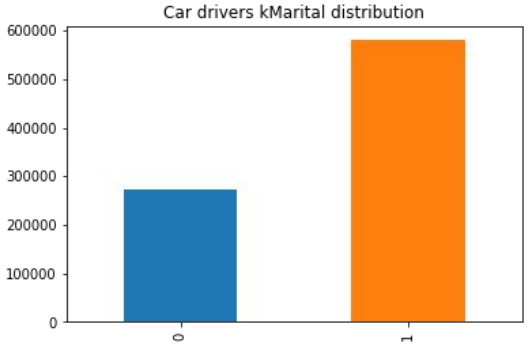
Cyclists have a fairly flat distribution across travel

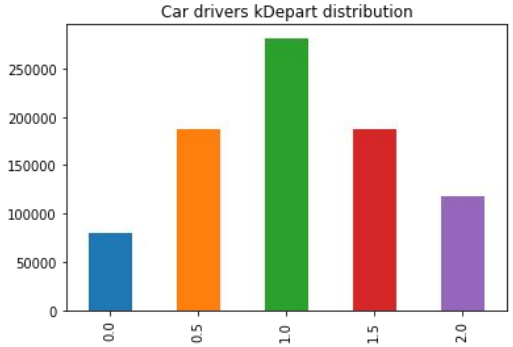


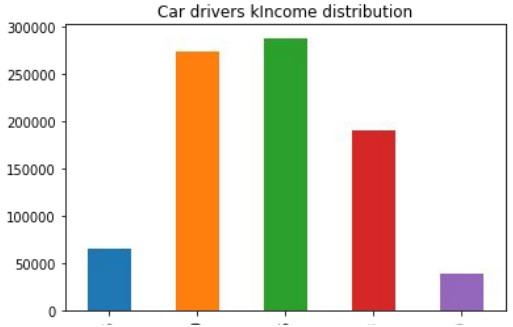


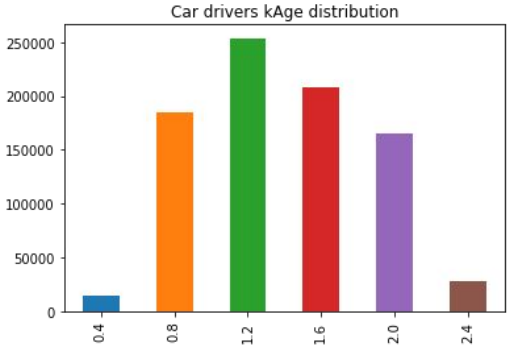


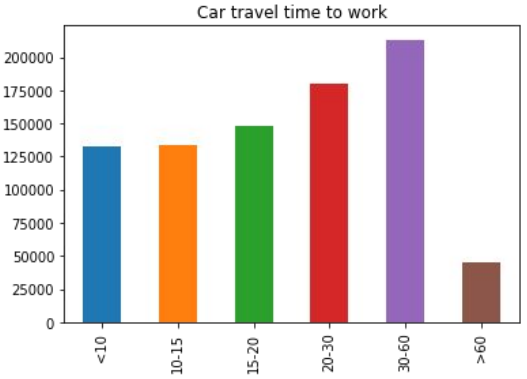
People who commute by car have a fairly flat distribution across travel times, and have the same distribution of characteristics as the general population:











- The majority are married
- ⅓ leave for work between 7-8am
- Income and age are normally distributed
- Travel time to work tends to be longer for car drivers

Who are potential non-car commuters?

- Focusing on a normalized subset of the census characteristics, I found each car driver's 20 nearest neighbors (k-Nearest Neighbors algorithm).
- Given these 20 nearest neighbors, I calculated how many of them commuted by bus, subway/railway, or bike
- For car drivers with at least 4 nearest neighbors in a category (2 for bike) I then clustered them into four clusters per category using a k-Means clustering algorithm.
- The general characteristics for each cluster are presented in the following slides.

Potential Bus Commuters: 15.5k

Young Female service workers: US born single females, they have some college, and work in either retail, or entertainment industries

Foreign born manufacturers: foreign born workers of all ages, working in precision production, craft and repair in the manufacturing industry

Female Professionals:: Predominantly single, early risers with high incomes, with nearly 75% leaving before 8am. They tend to work in technical, sales and admin support occupations and have university or associate degrees

Laborers: Foreign born with minimal education (less than high school)

Potential Train Commuters: 11k

Young professionals: Wealthy night owls working in managerial and professional occupations in the finance and insurance industries, the majority have masters degrees

Working Girls: Young, largely foreign born female night owls working in technical, sales and admin support in the finance and insurance industries, education levels range from college degrees to masters

Married Immigrants: Older (65% >50) foreign born low income married service workers with minimal schooling

Financiers: 30-50 year olds, high income married men working as managerial and professional occupations in finance and insurance industries. They generally have bachelors or masters degrees (80%)

Potential Bike Commuters (4.1k)

The Laborers: US born, largely single men working as operators, fabricators or laborers, they generally have some college

The Student: 20-30 year old night owls, men working in the retail industry with some college or an associate degree

The Family Man: Young, foreign born men who leave after 10am and are more likely to be married, working in the retail industry with minimal schooling.

MAMILs: 30-40 year old, US born, leave for work between 7-8am and have high incomes. They work long hours, are married, and in managerial and professional occupations and are very highly educated.

Takeaways

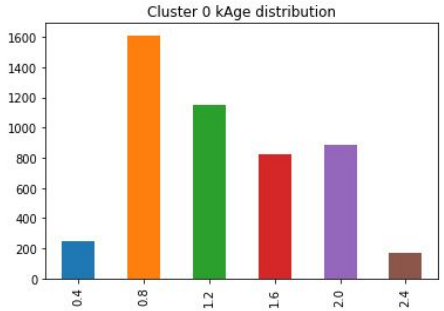
In determining who would most benefit from, and adopt, alternate forms of transportation it’s useful to see whether similar individuals have adopted alternate forms.

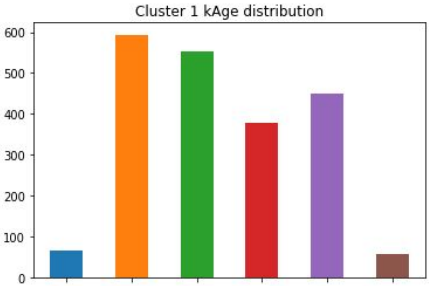
This is geographically-blind, so we are supposing that a finance professional in Delaware would be equally likely to take a subway as a finance professional in NYC, all else equal

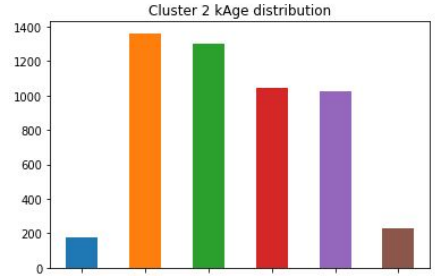
Likelihood of adoption varies across individuals, dependent on characteristics, but can be modelled through a clustering algorithm to find target groups.

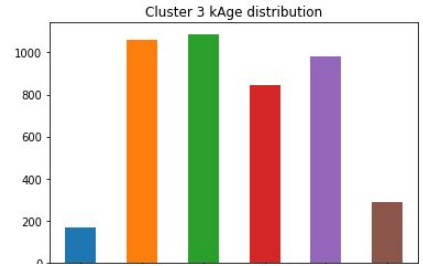
Appendix

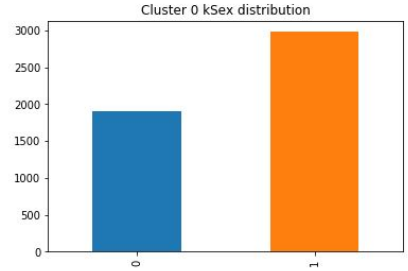
Potential Bus Clusters

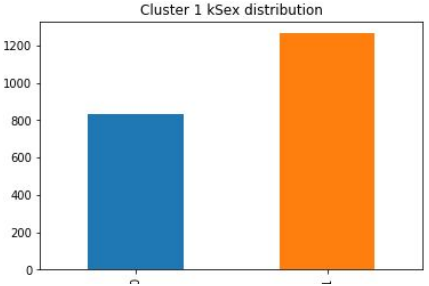


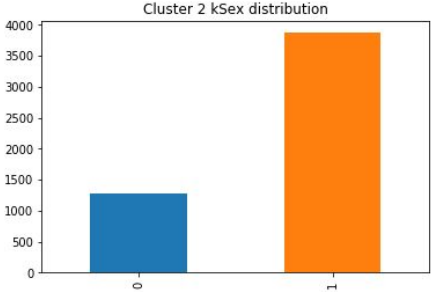


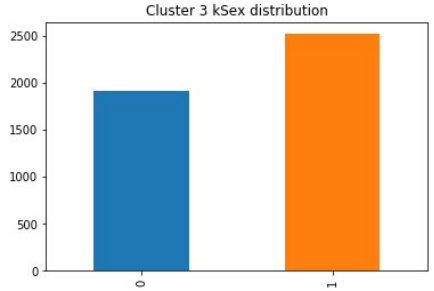


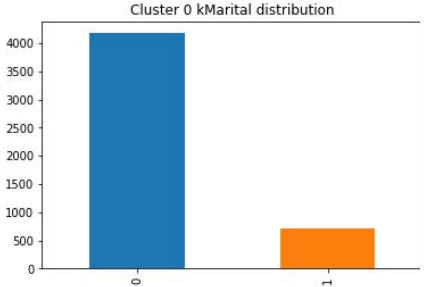


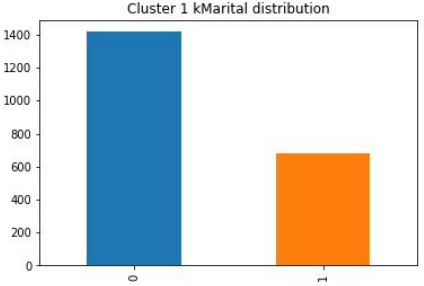


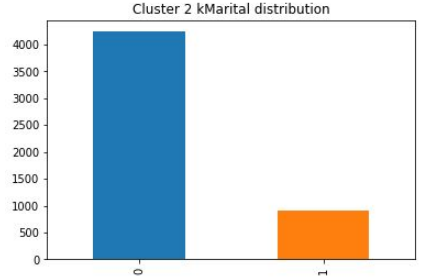


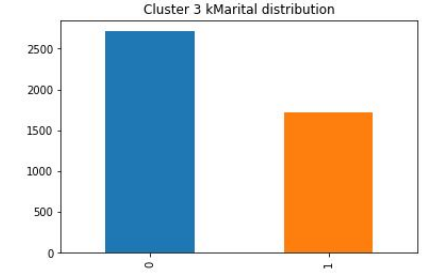




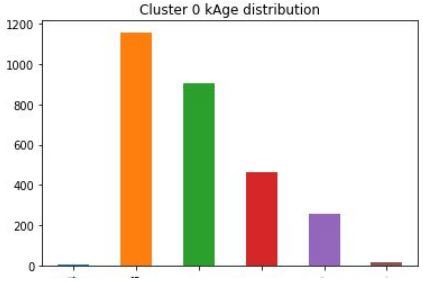


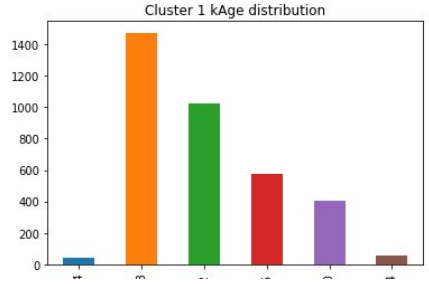


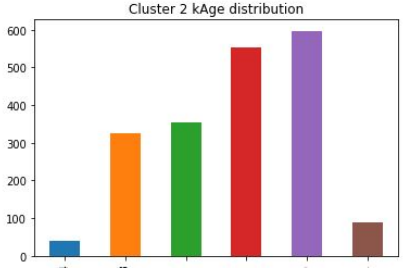


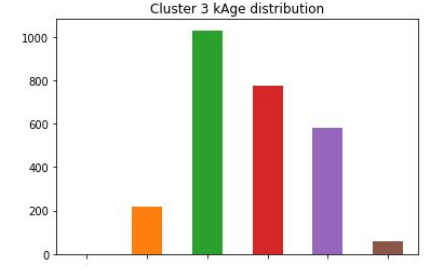


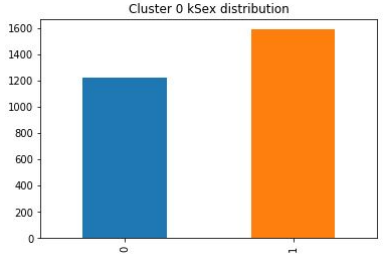
Potential Train Clusters

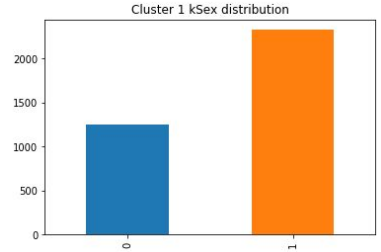


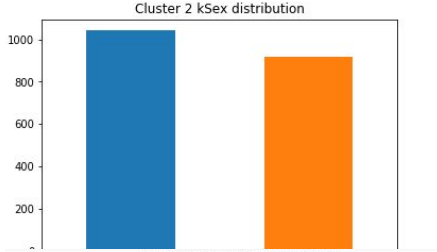


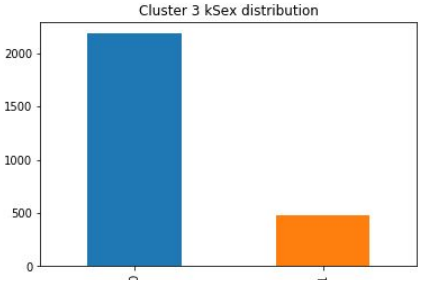


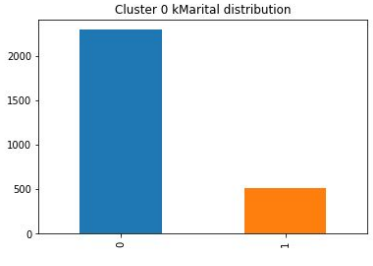


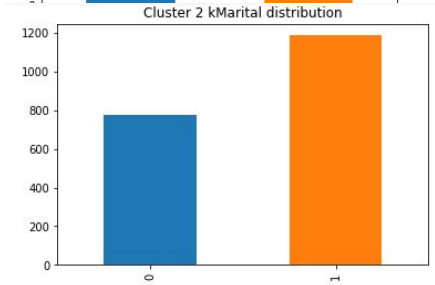


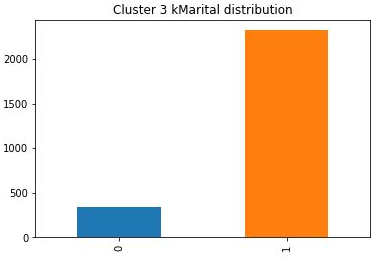




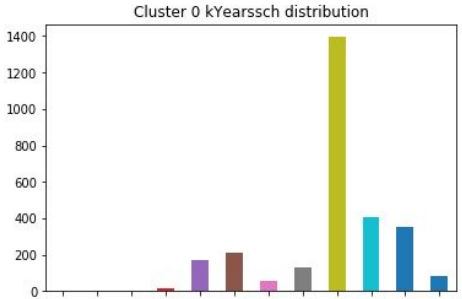


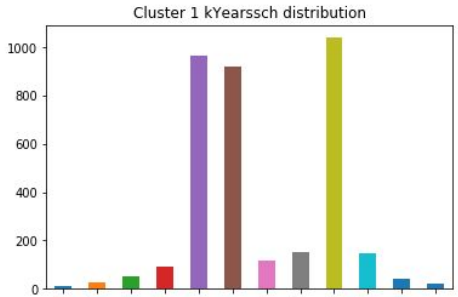


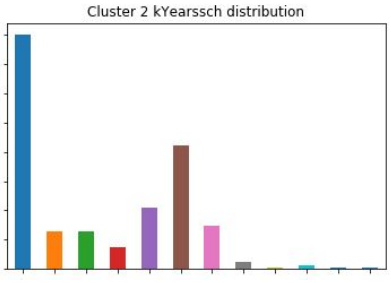


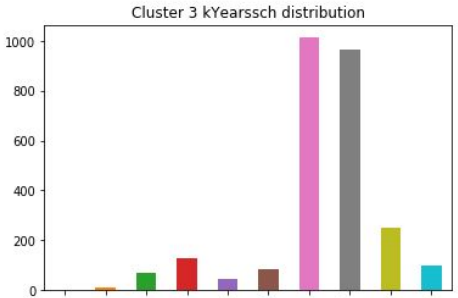


Potential Train Cluster (cont - schooling)









Potential Bike Clusters