

A person is riding a blue Citi Bike on a city street. The bike has "citi bike" written on the frame and the front fender. The number "15763" is visible on the rear fender. The background is a blurred city street with other vehicles.

Analiza Danych Rowerowych w Nowym Jorku

Łukasz Lepianka
Marta Szuwarska

Spis treści



Czyszczenie danych



Aplikacja w Shiny



Analiza wieku rowerzystów
- w powrównaniu z odległością
- w porównaniu z prędkością



Przewidywanie typu użytkownika
(Subscriber lub Customer) na
podstawie pozostałych danych

Z jakich danych korzystaliśmy?

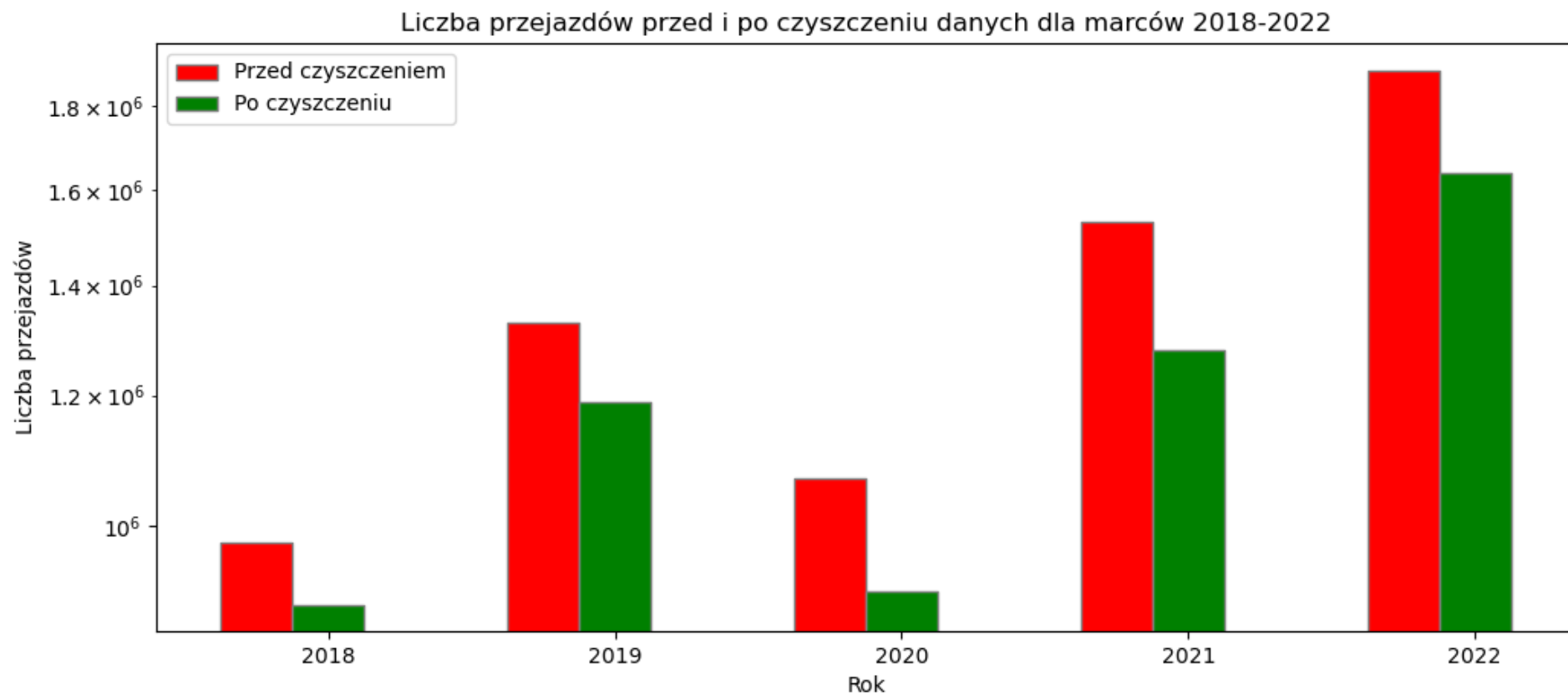
Marce
2018-2022

Cały rok
2019

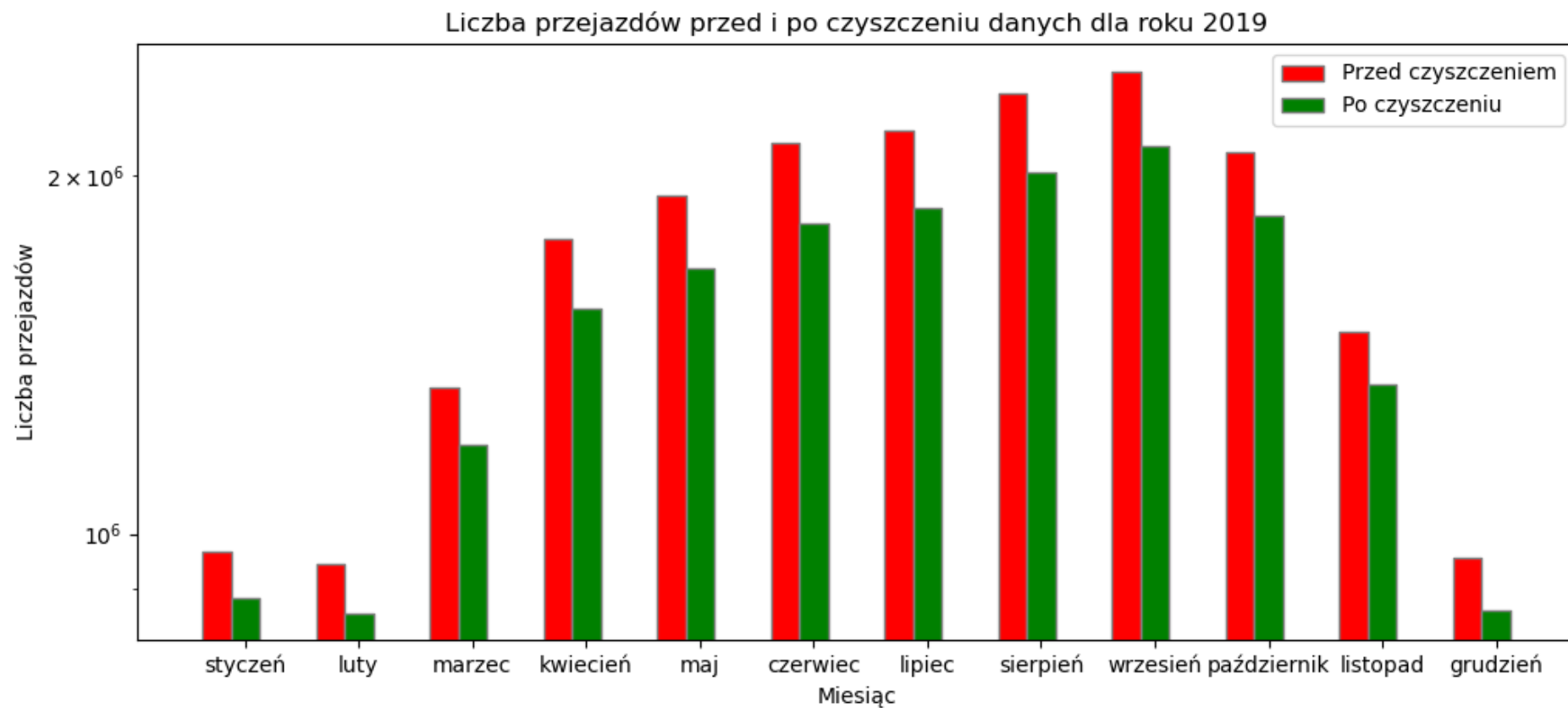
Cały rok
2016

Kwiecień
2022

Czyszczenie danych dla marców 2018-2019



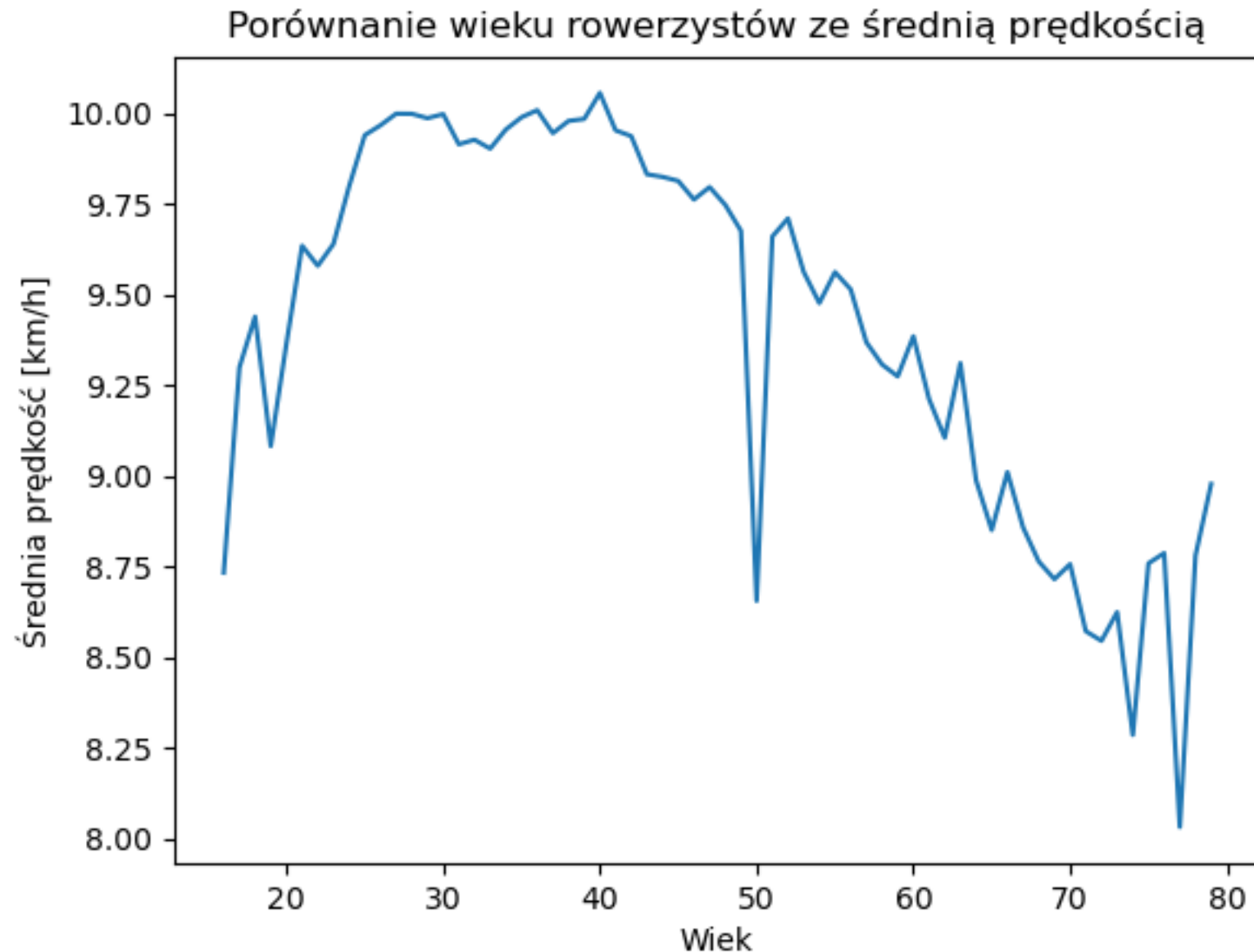
Czyszczenie danych dla roku 2019



Analiza wieku rowerzystów w porównaniu z odległością

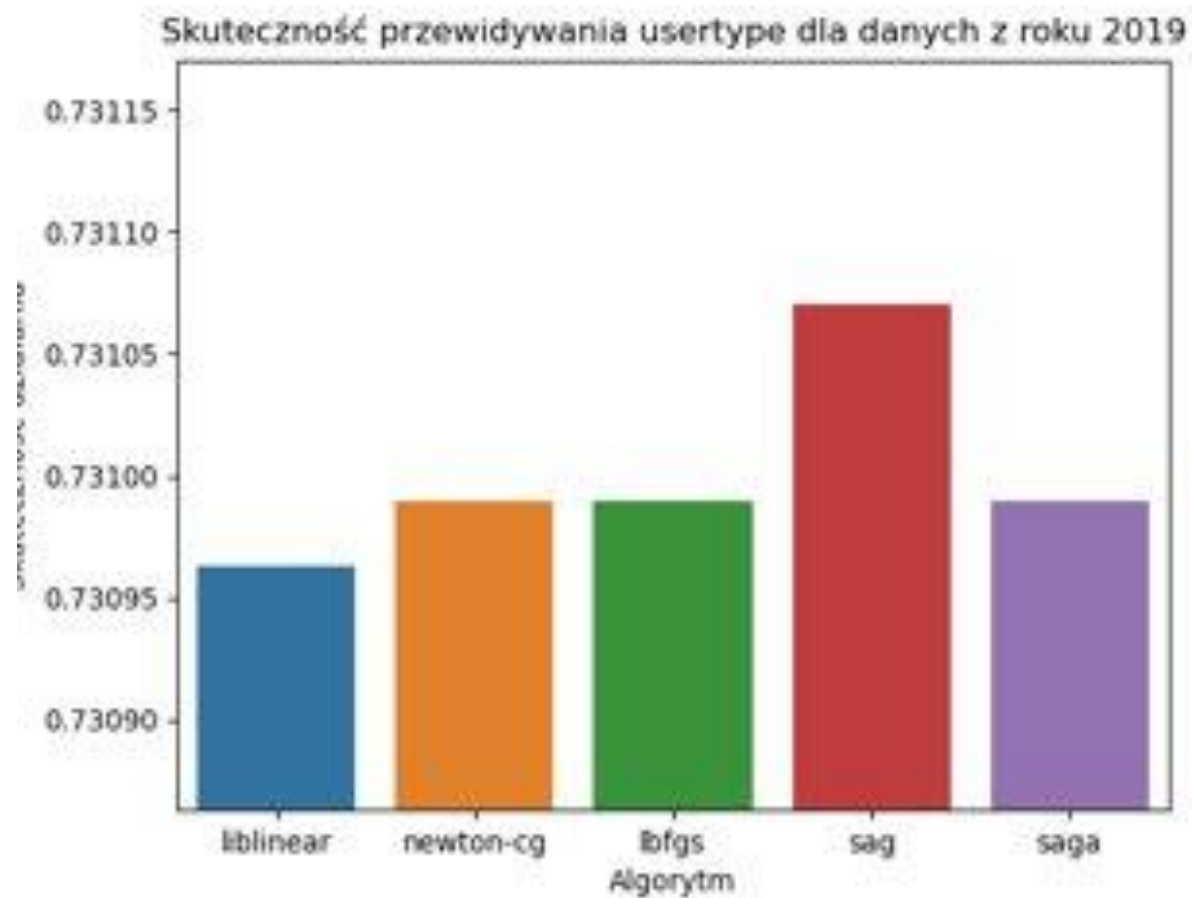


Analiza wieku rowerzystów w porównaniu z prędkością

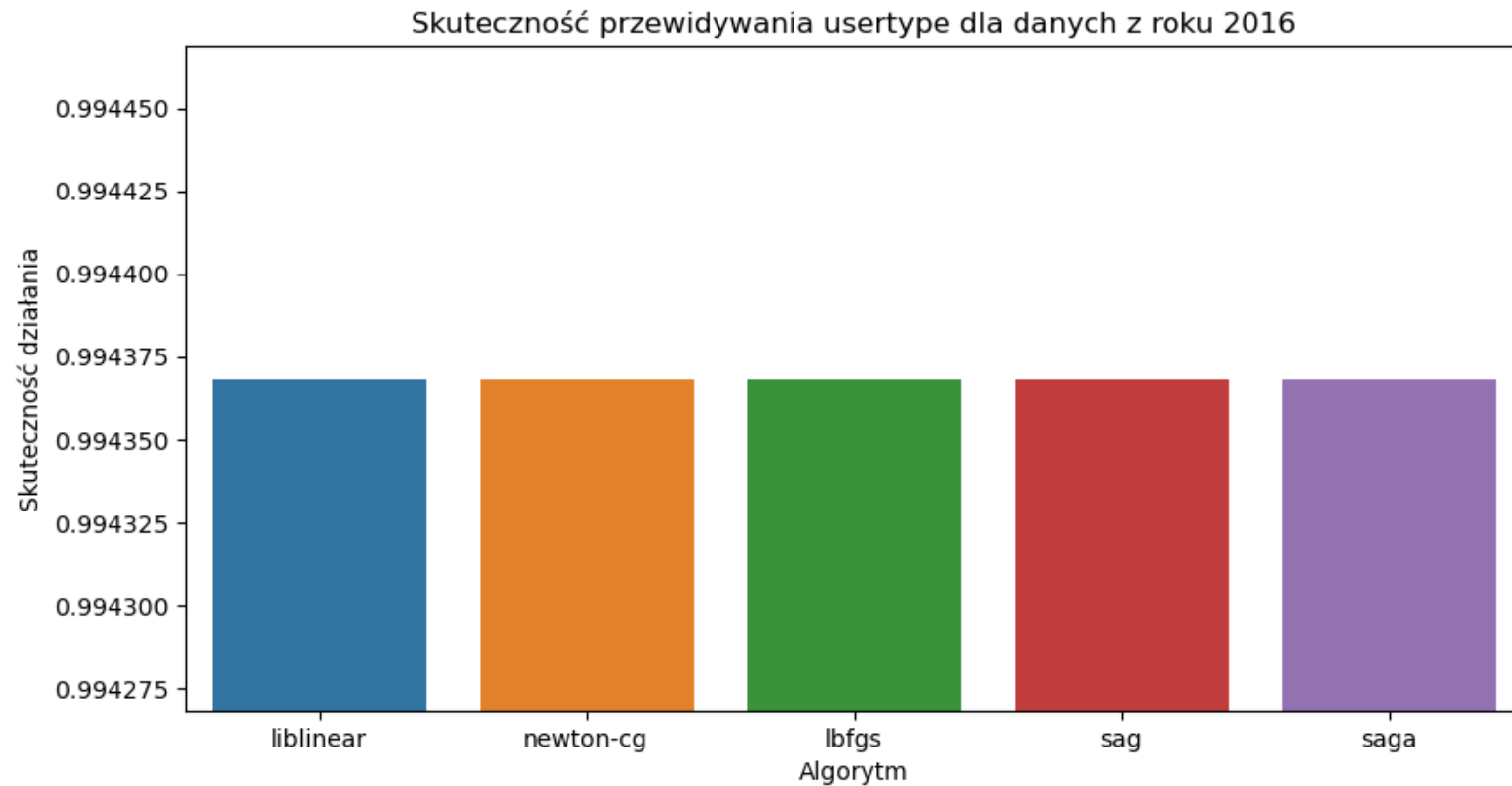


Przewidywanie typu użytkownika dla roku 2019

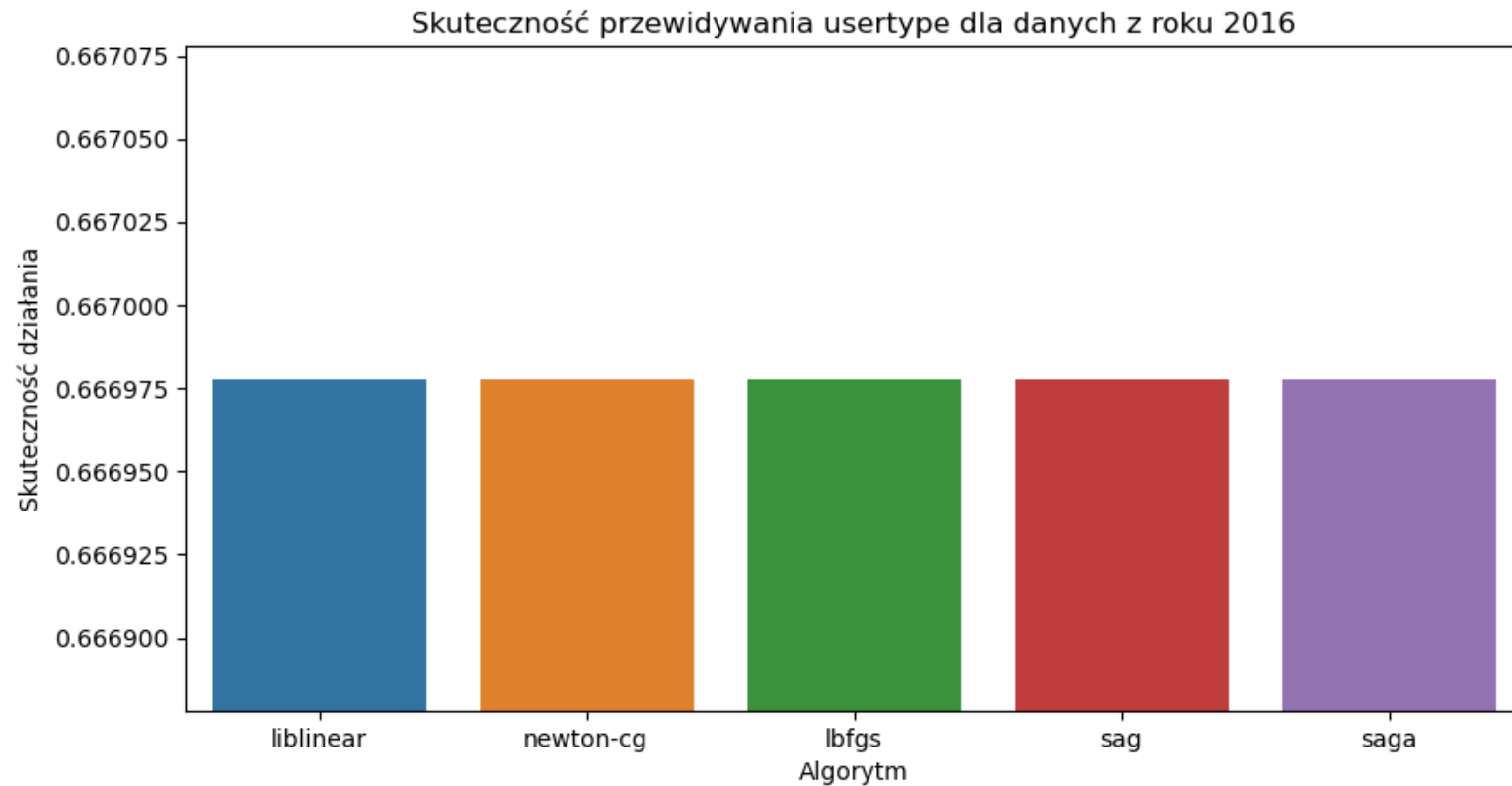
Przewidywanie typu użytkownika dla roku 2019



Przewidywanie typu użytkownika dla roku 2016 (dane przed czyszczeniem)



Przewidywanie typu użytkownika dla roku 2016 (dane po czyszczeniu)



Podczas
projektu
nauczyliśmy
się:

- Współpracowania z Gitem
- Analizowania danych w R
- Analizowania danych w Pythonie
- Tworzenia podstawowych aplikacji w Shiny
- Tworzenia interaktywnych map w Leaflet, a przy okazji wykorzystywanie do tego celu Spatial Objectów.
- Tworzenia podstawowych modeli predykcyjnych w Pythonie za pomocą regresji logistycznej (Pandas, Scikit-Learn)
- Obliczania odległości między współrzędnymi na kuli ziemskiej
- Tworzenia wykresów za pomocą matplotlib, ggplot2