

PDRPy2017

Analiza danych - tramwaje warszawskie

Łukasz Muszyński
Dawid Stelmach

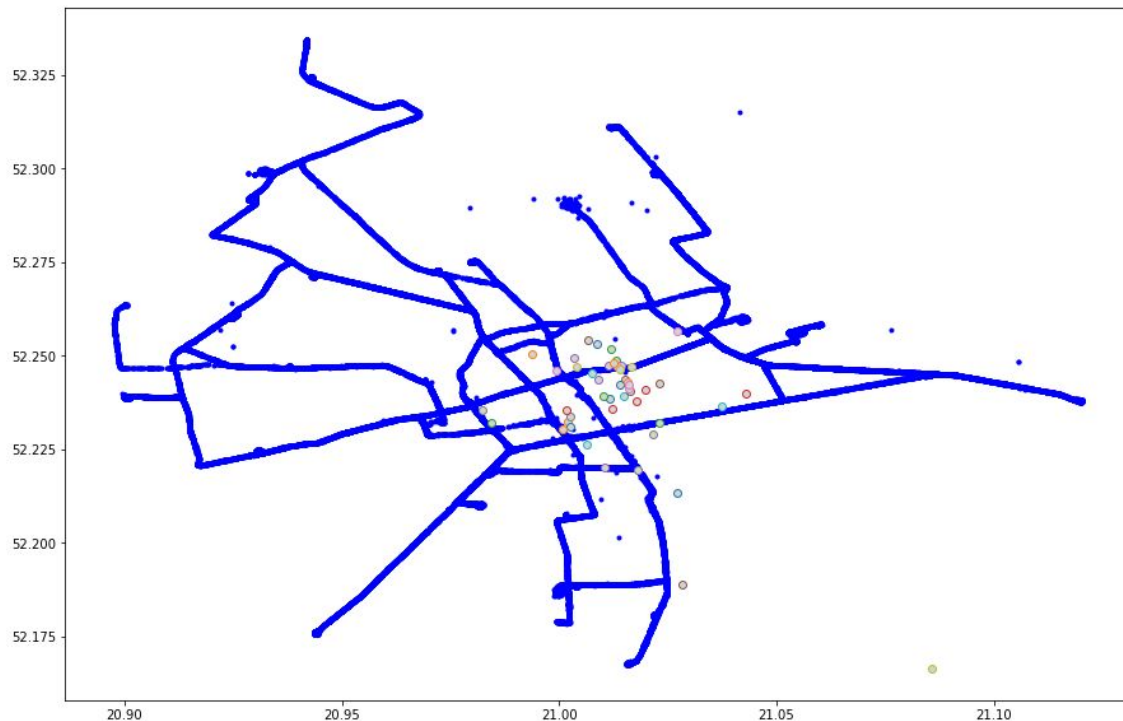
Pobieranie danych

- 16.05 - 29.05.2017
- skrypt Pythonowy + AWS
- dane pobierane co 3 min.
- 182MB danych

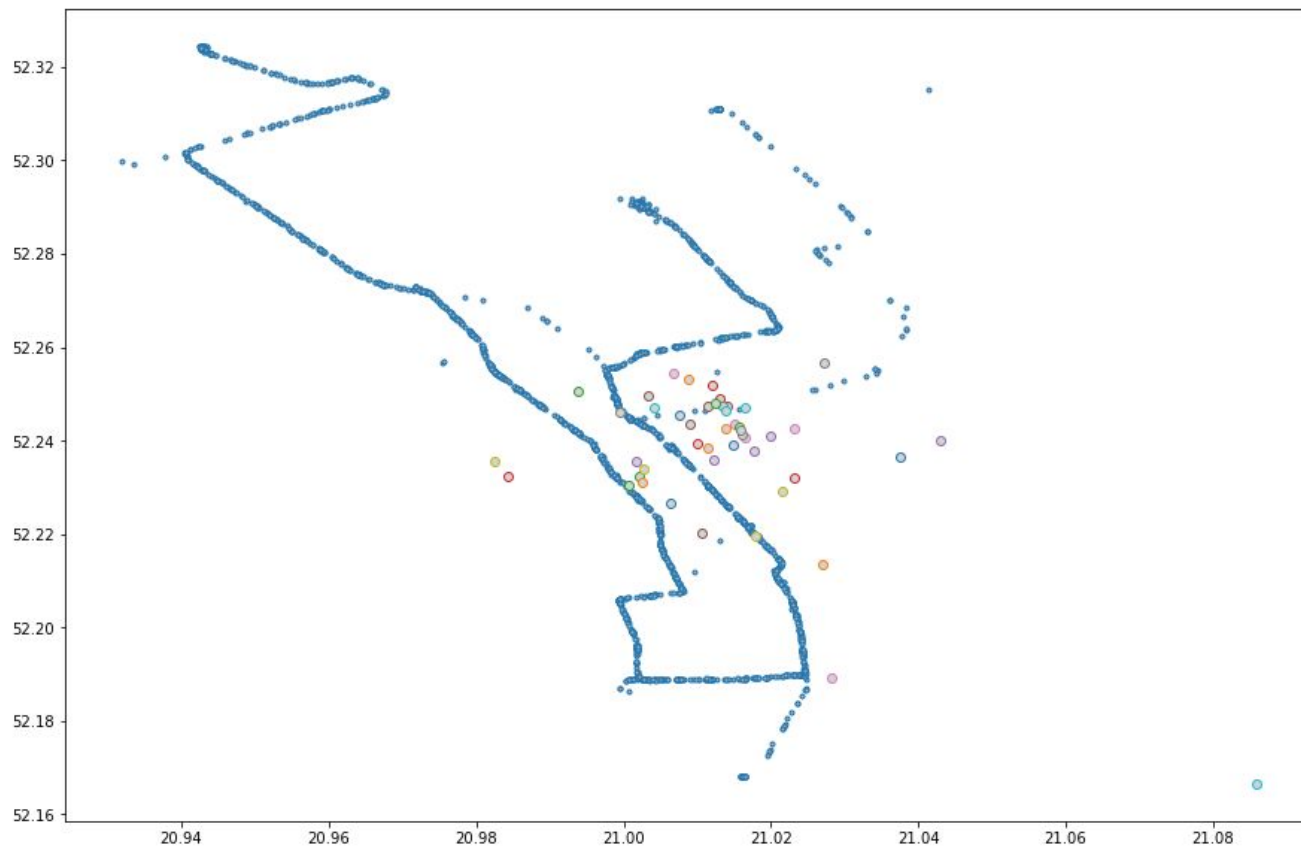


Najciekawsze trasy (budynki)

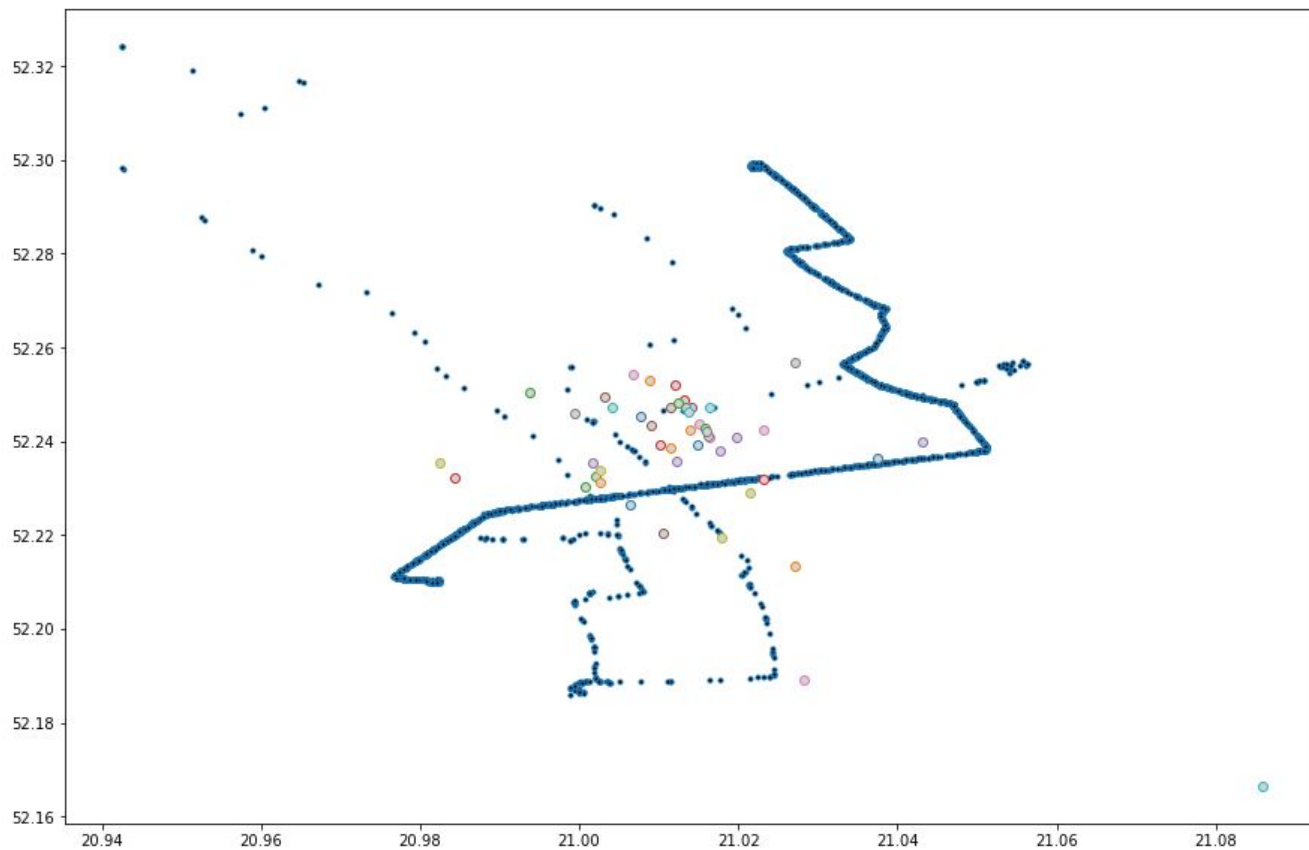
- 47 budynków na podstawie <http://www.ranker.com/list/warsaw-buildings-and-structures/reference>
- przejazd w odległości maksymalnie 300 m. od budynku



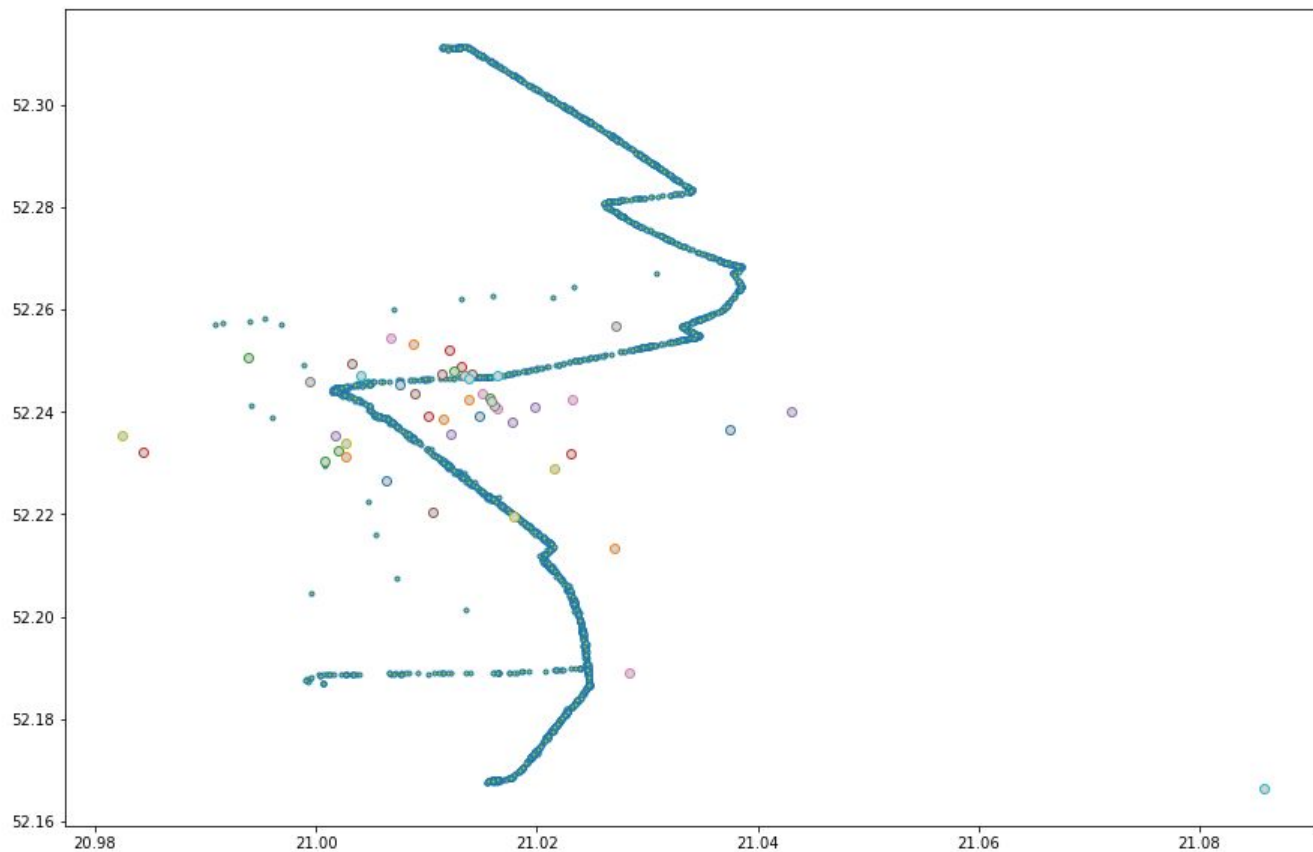
Liczba budynków dla linii 18: 25



Liczba budynków dla linii 25: 24

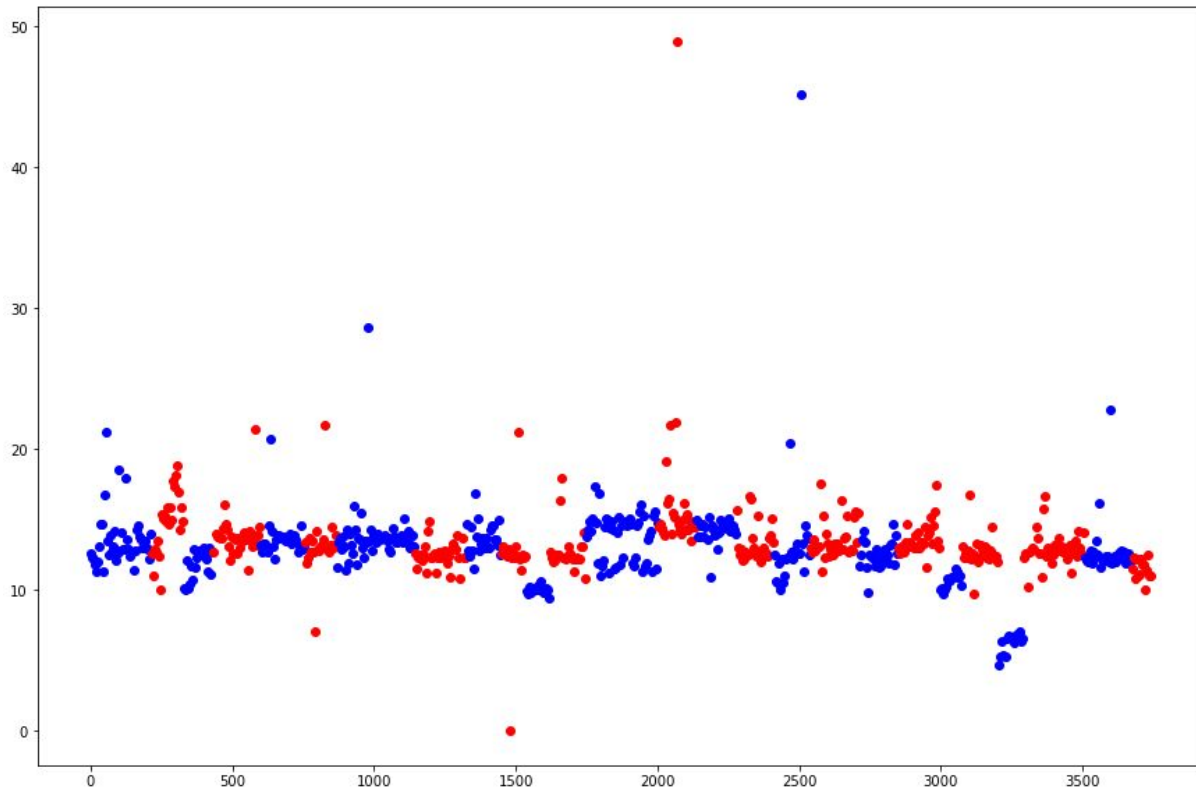


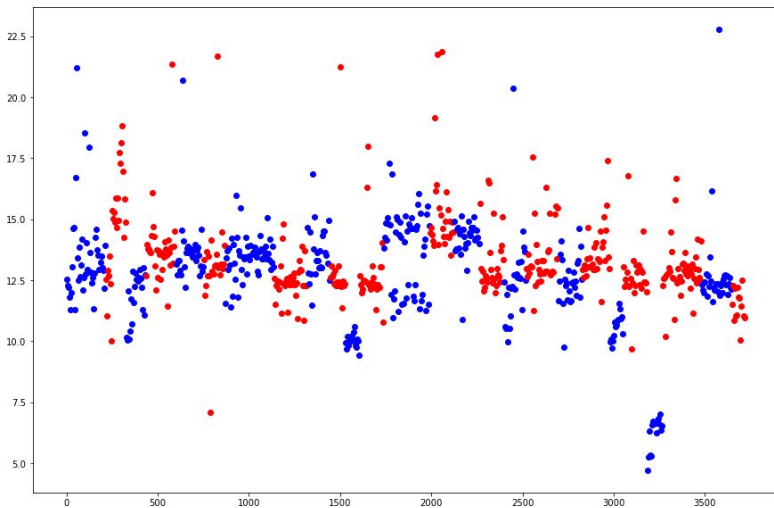
Liczba budynków dla linii 4: 19



Średnie prędkości

- summaryczne dane z 2 tygodni
- założenia dla średniej prędkości (zdroworozsądkowe)
 - max 25 km/h
 - min 2 km/h

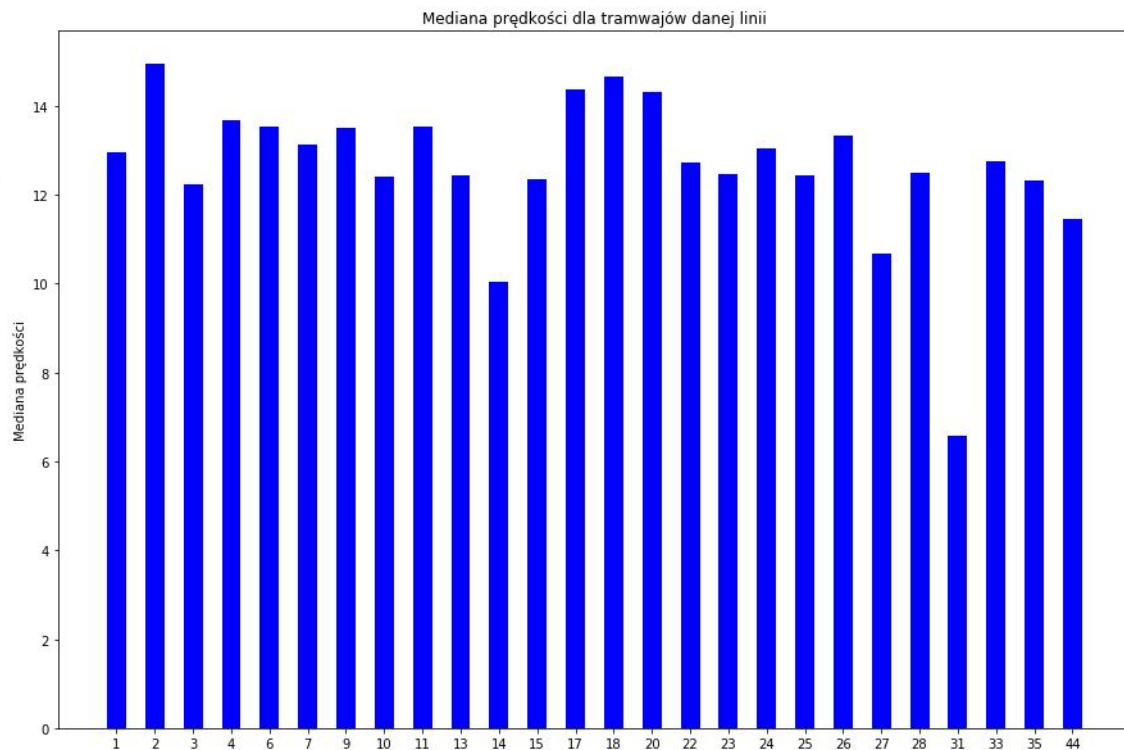




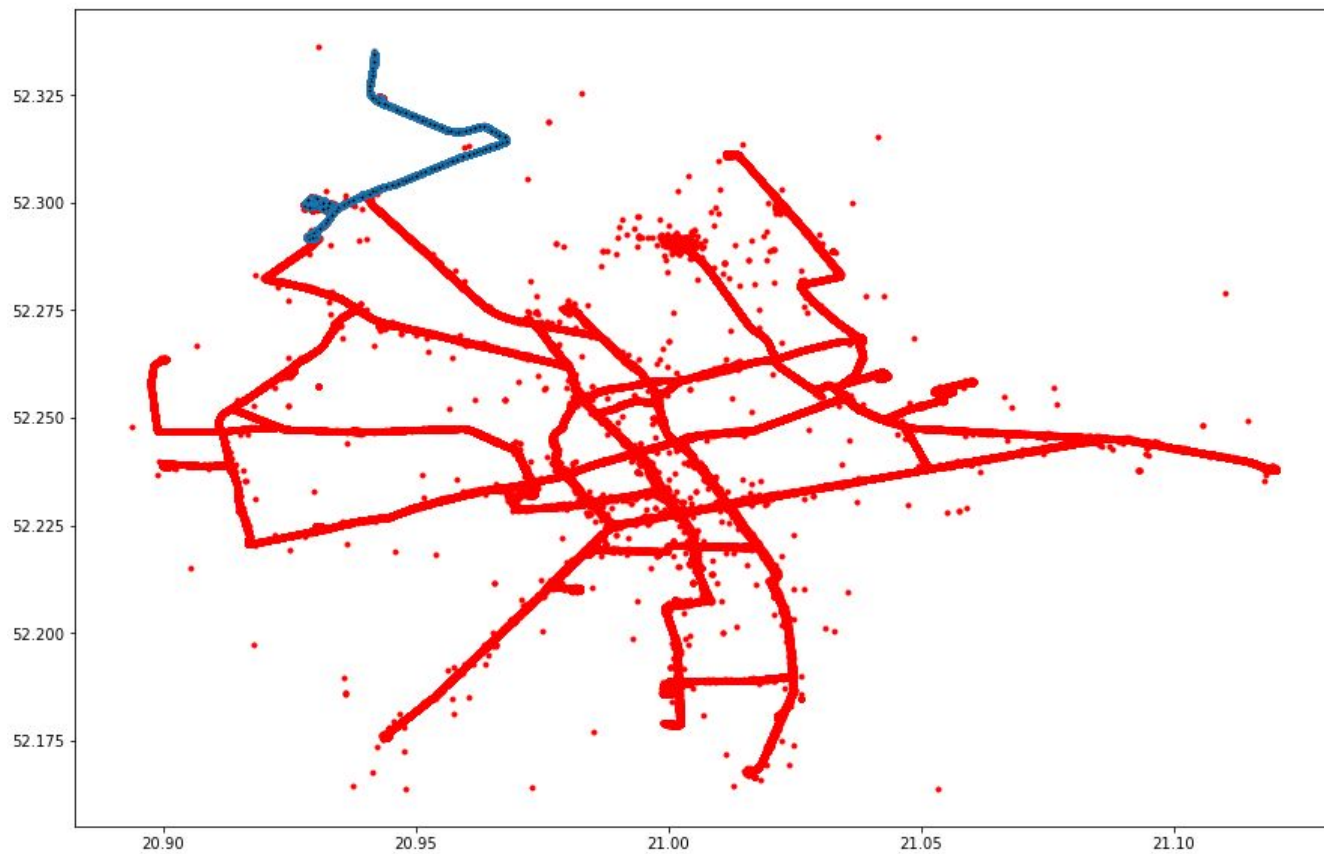
Po odfiltrowaniu

Mediana wydaje się
bezpieczniejszym estymatorem
niż średnia.

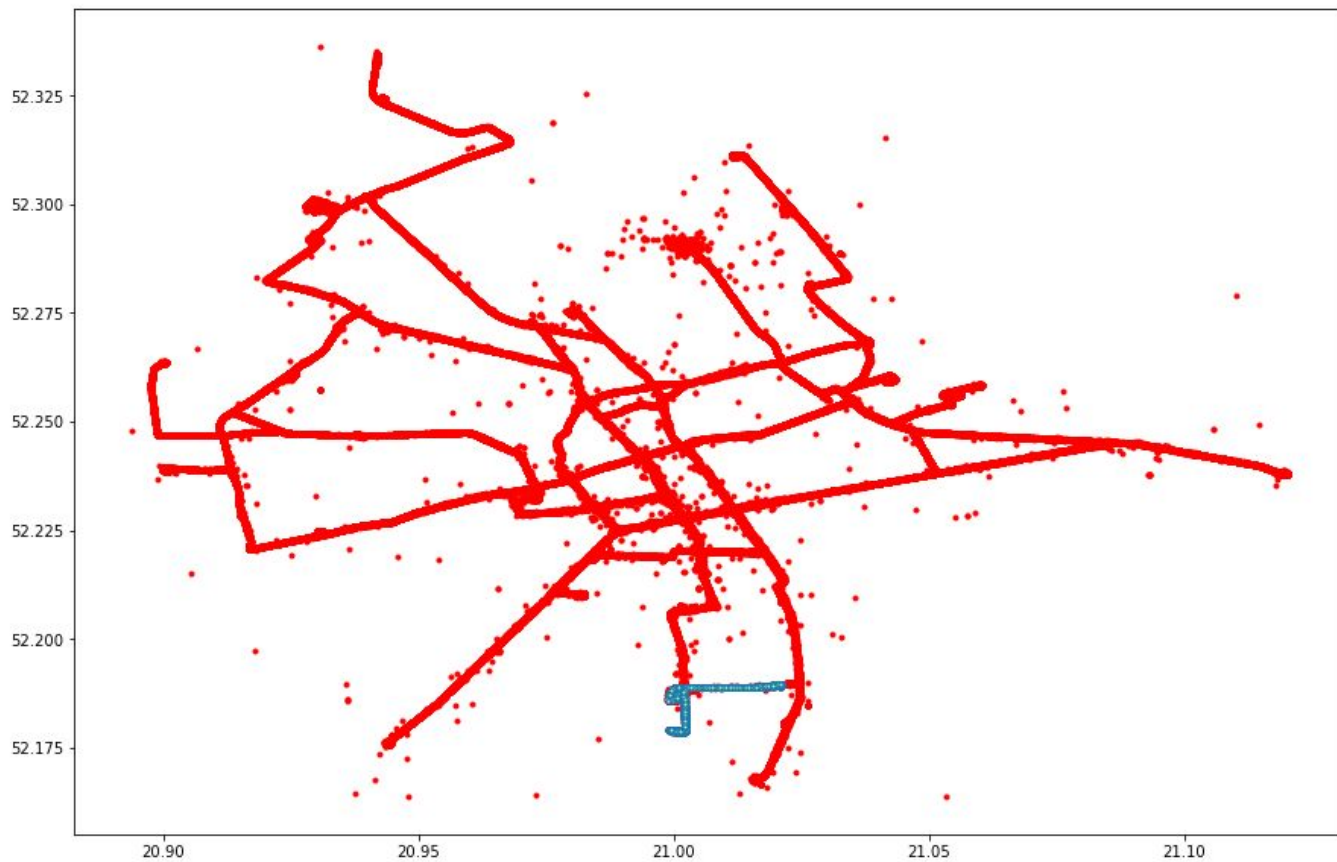
- min - 6.6 km/h
- max - 15 km/h



Linia nr. 2 - najszybsza

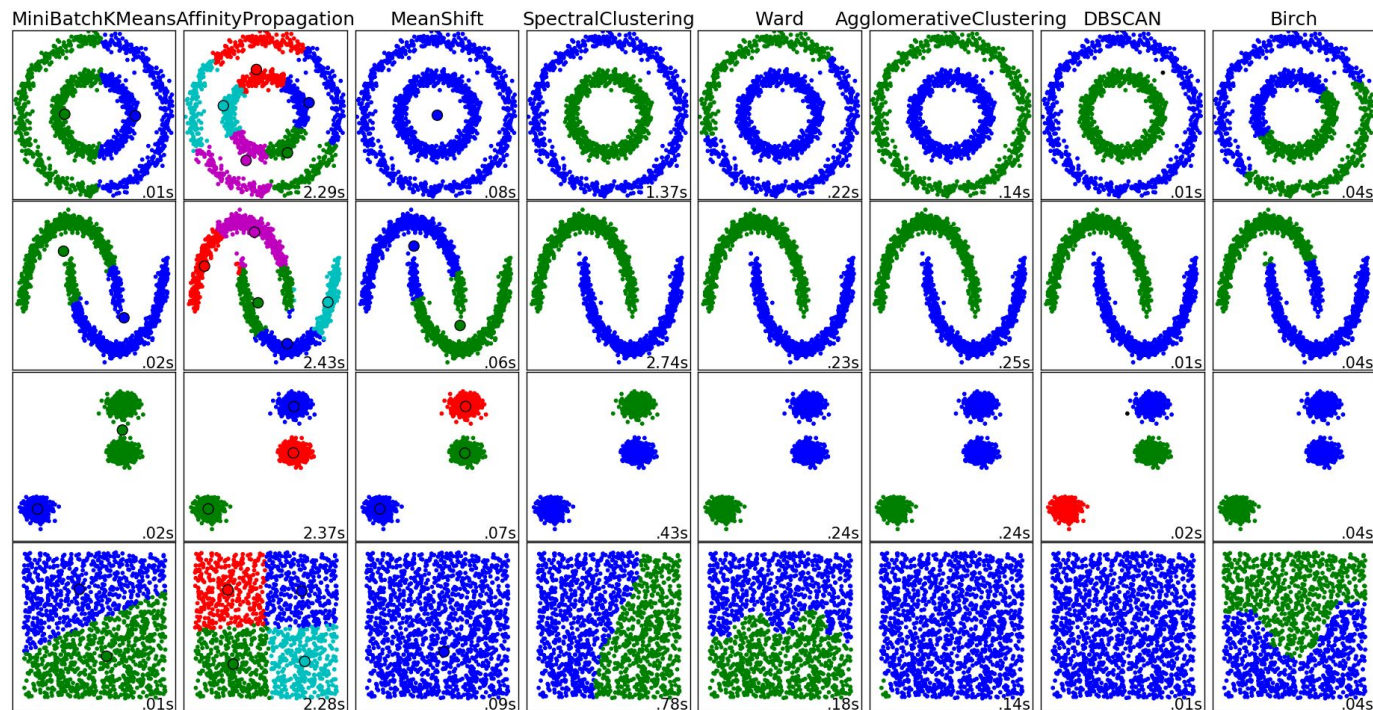


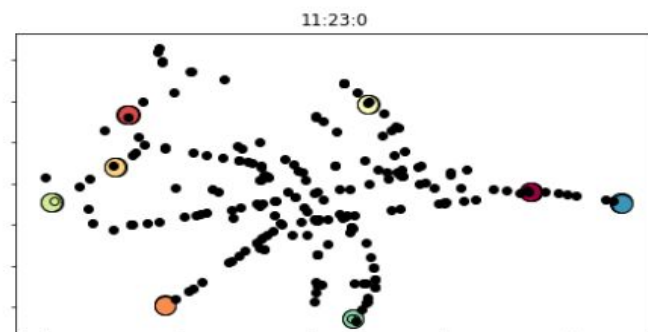
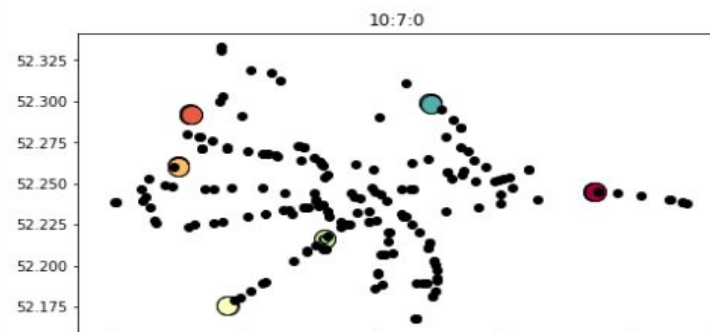
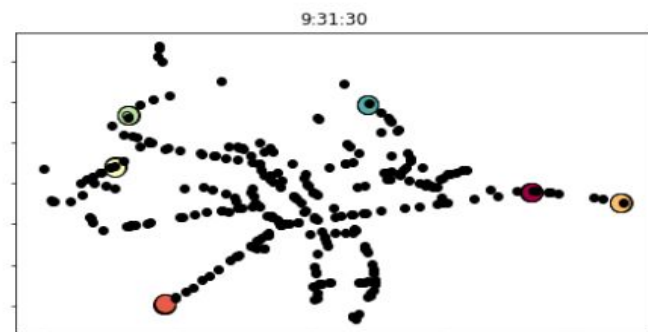
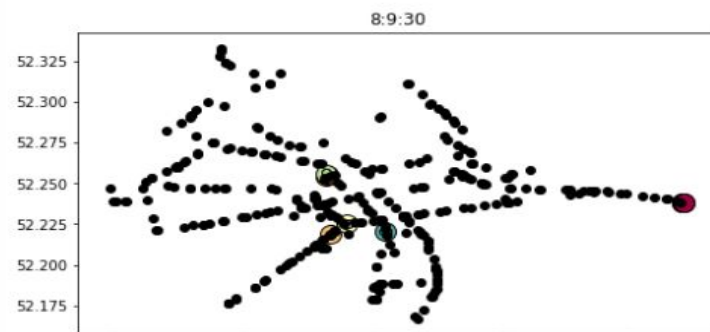
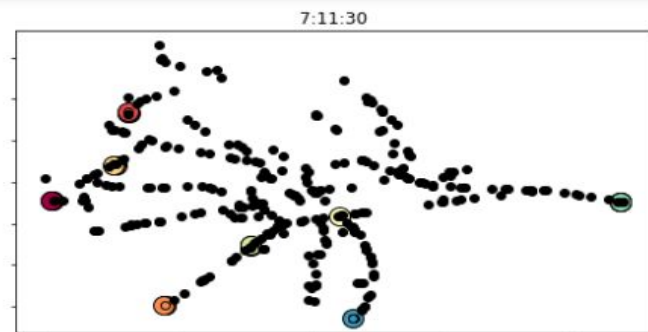
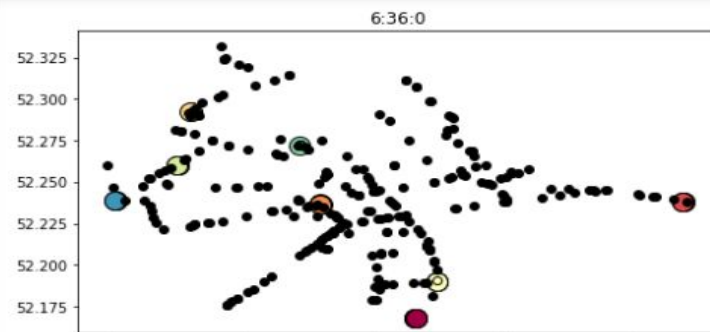
Linia nr. 31 - najwolniejsza



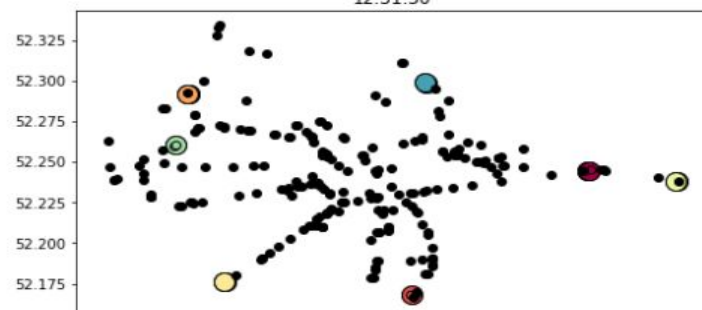
Clustering

- zajezdnie
- zatory
- skupiska przystanków
- Zastosowany algorytm - DBSCAN
- Minimalna liczność klastra - 4 tramwaje
- Maksymalna odległość - 50 m.

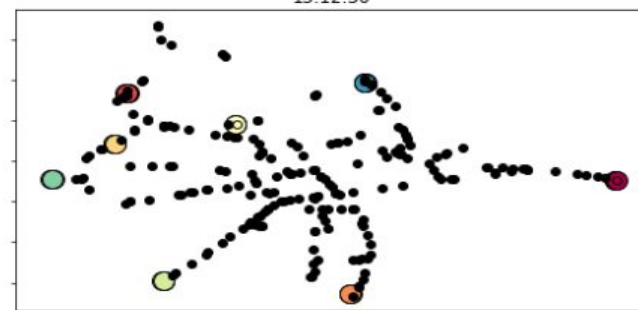




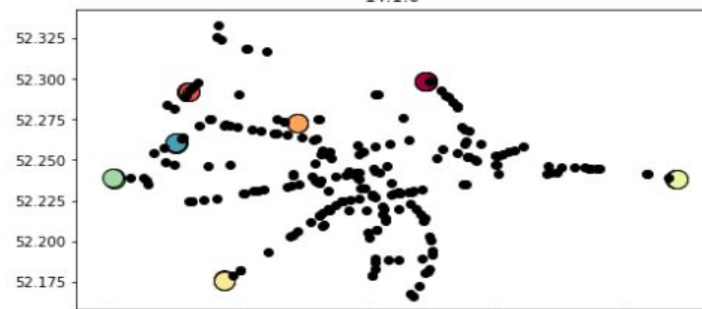
12:31:30



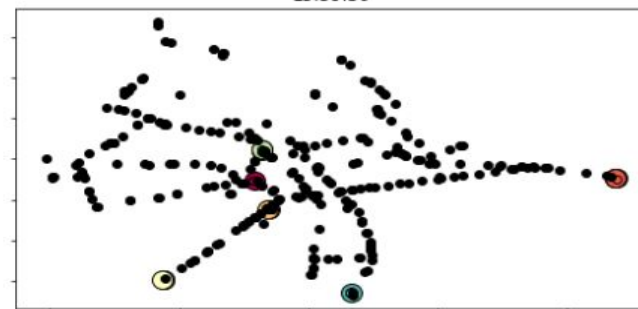
13:12:30



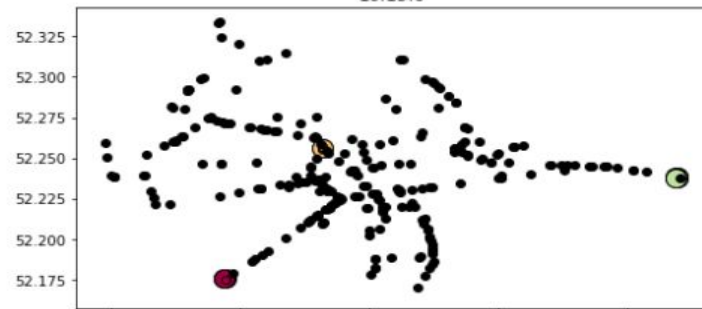
14:1:0



15:59:30



16:15:0



17:15:30

