

Wykrywanie korelacji pomiędzy aktywnością użytkowników Orange a opóźnieniami komunikacji miejskiej

Zespół “100”

**URBAN SENSORS
THE DATA SCIENCE HACKATHON
Warszawa, 26.09.2017**

Motywacja



Pomysł

Być może są korelacje między aktywnością użytkowników telefonii komórkowej a opóźnieniami komunikacji publicznej / korkami ?

Hipotezy:

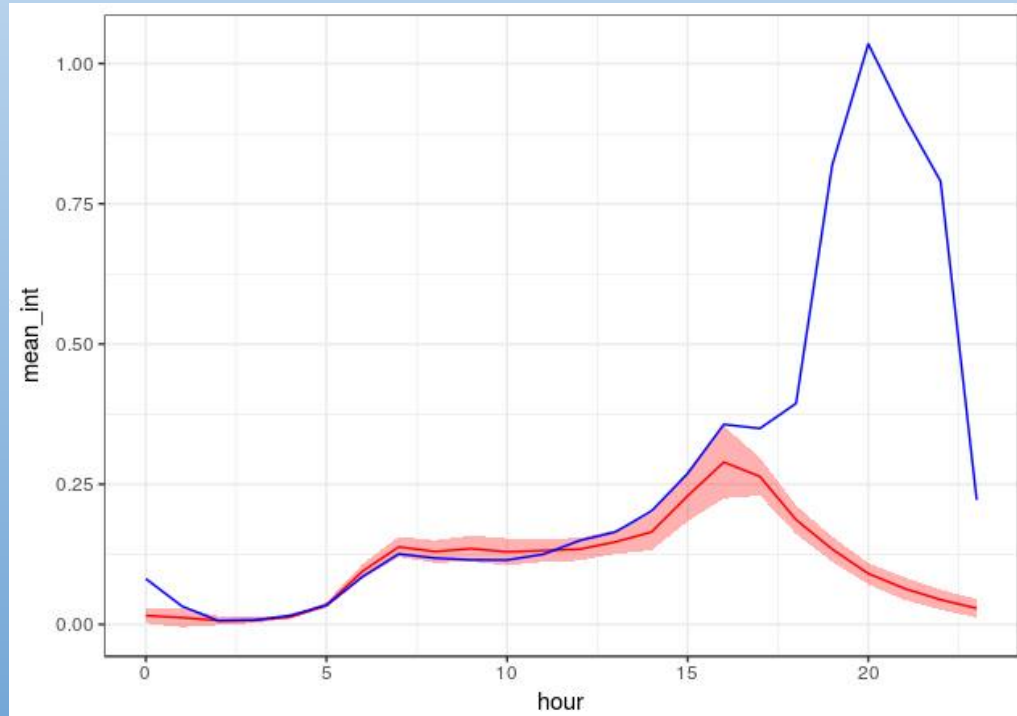
- Jest duża aktywność użytkowników na danym obszarze -> być może jest tam dużo pojazdów / pasażerów -> być może są wtedy korki / opóźnienia komunikacji ?
- Jest duża aktywność użytkowników na danym obszarze -> być może jest tam jakieś wydarzenie -> być może niedługo dużo osób będzie podróżować (i w pobliżu utworzą się korki) ?

Narzędzia i materiały

- Dane offline komunikacji miejskiej
- Dane offline z telefonii komórkowej
- R

Wyniki

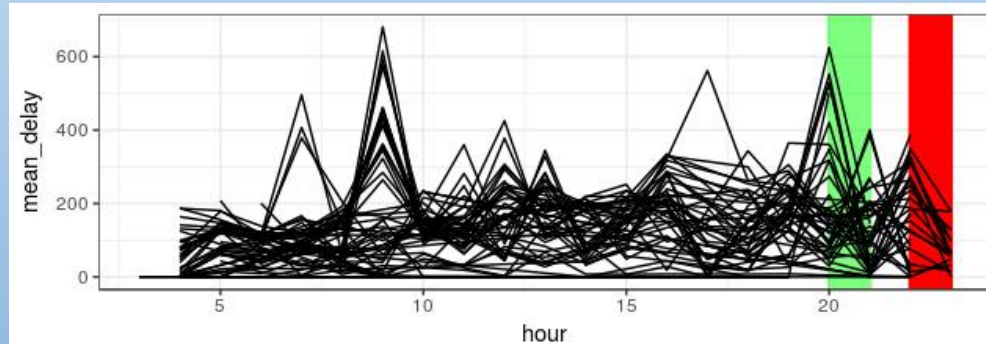
Zauważyliśmy korelację między wydarzeniem masowym (mecz Polska-Kazachstan, 4.09, g. 20:45) a aktywnością użytkowników Orange.



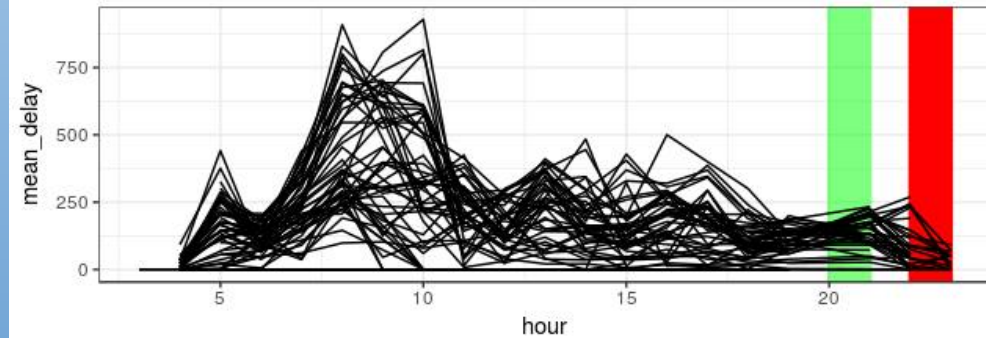
Wyniki

Zauważyliśmy korelację między wydarzeniem masowym (mecz Polska-Kazachstan, 4.09, g. 20:45) a opóźnieniami dojazdu (rys. autobus 162 - trasa blisko stadionu).

Poniedziałek 4.09



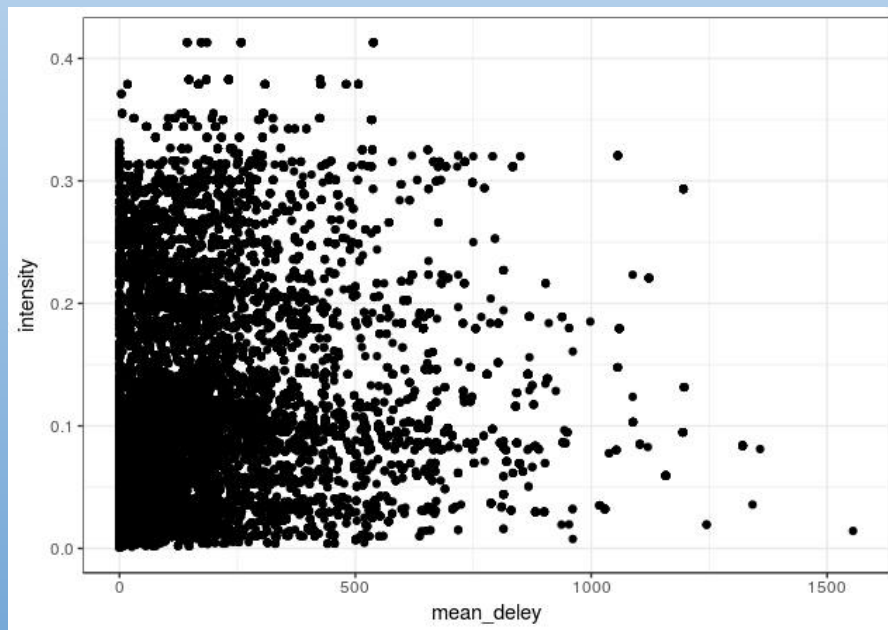
Wtorek, 5.09



Wyniki

Nie znaleźliśmy korelacji między aktywnością a opóźnieniem
(wsp. korelacji liniowej ~6%)

Linia 162



Wnioski

- Nie wszystkie hipotezy dotyczące korelacji zostały potwierdzone
- Ale wynik negatywny też może mieć wartość
- Być może aktywność uczestników ruchu jest zbyt mała
- Być może warto badać korelacje pomiędzy niedużymi obszarami

Dziękujemy za uwagę!