



#AlertMPK

... czyli dzień jak co dzień



Analiza utrudnień w ruchu komunikacji miejskiej na podstawie postów MPK Wrocław z Facebooka

Kamil Raczycki

Marcin Szymański

Yahor Yeliseyenka

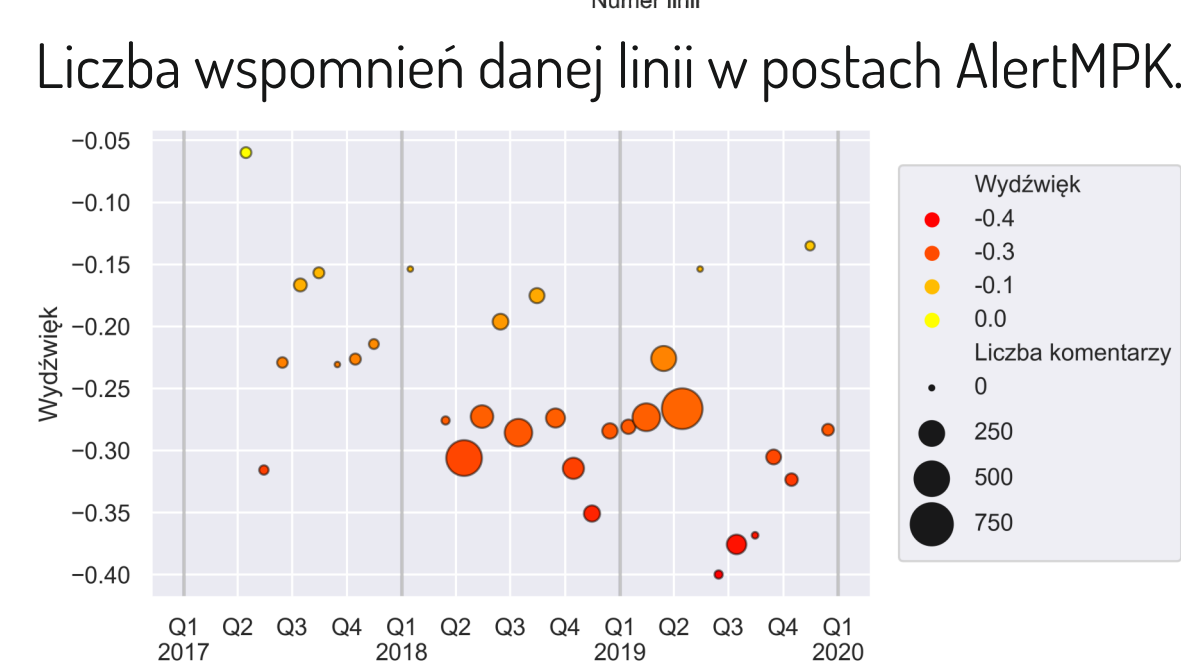
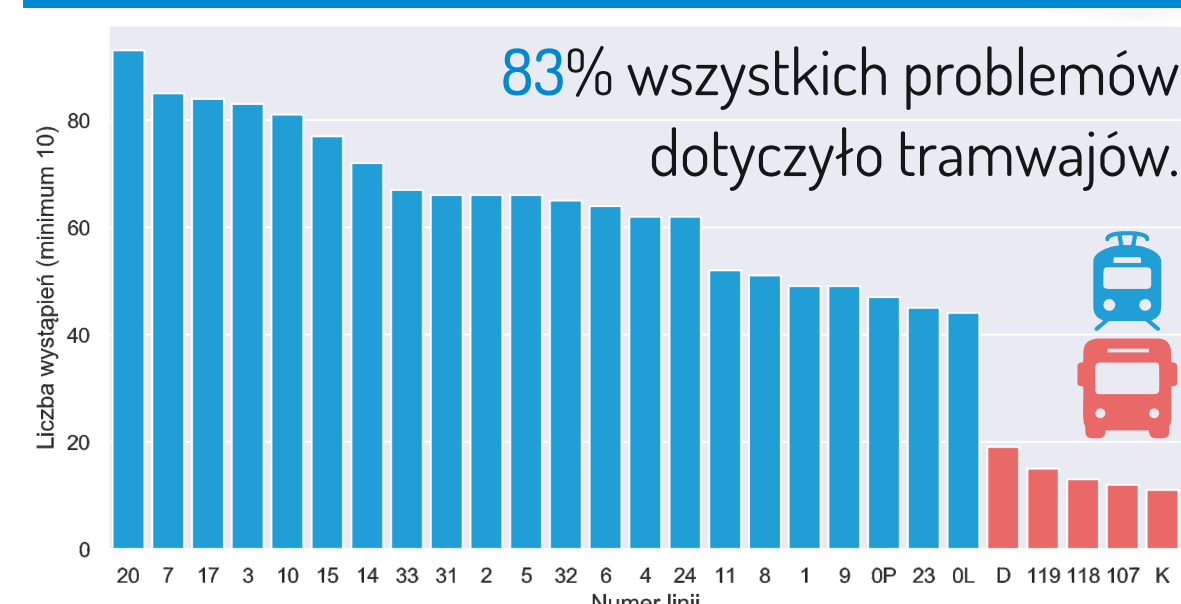
Wprowadzenie

MPK Wrocław przewozi około 200 milionów pasażerów rocznie, czyli około 550 tysięcy dziennie. W tak dużym mieście jak Wrocław utrudnienia w ruchu występują na porządku dziennym. Przewoźnik przekazuje informacje o wszelkich problemach w przejazdach, używając tablic informacyjnych, dedykowanej aplikacji oraz mediów społecznościowych (Facebook oraz Twitter).

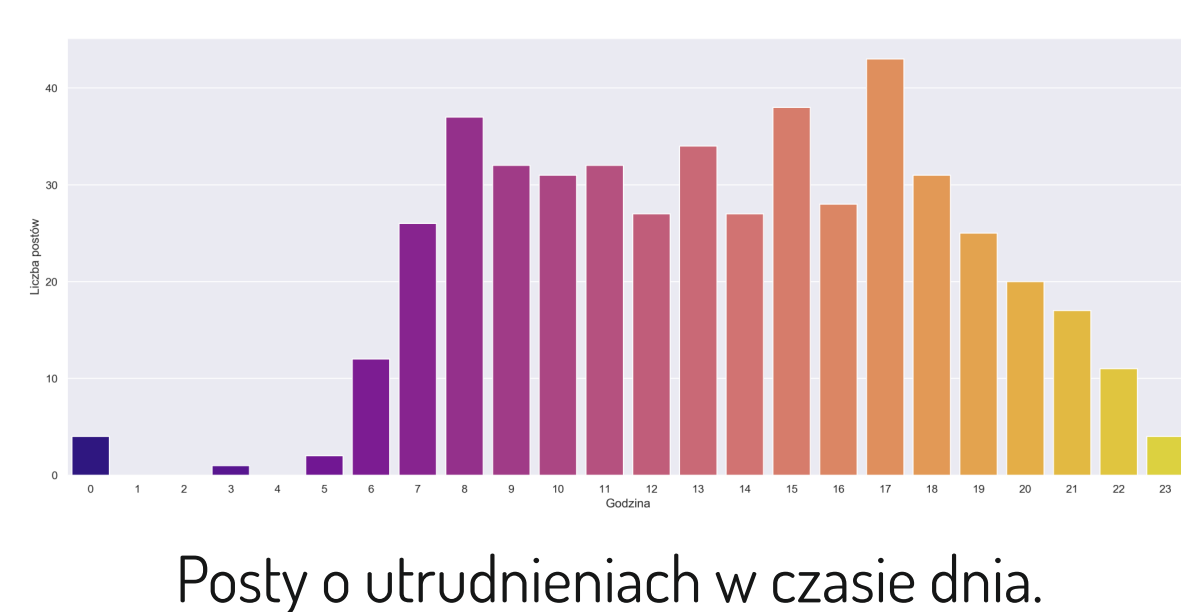
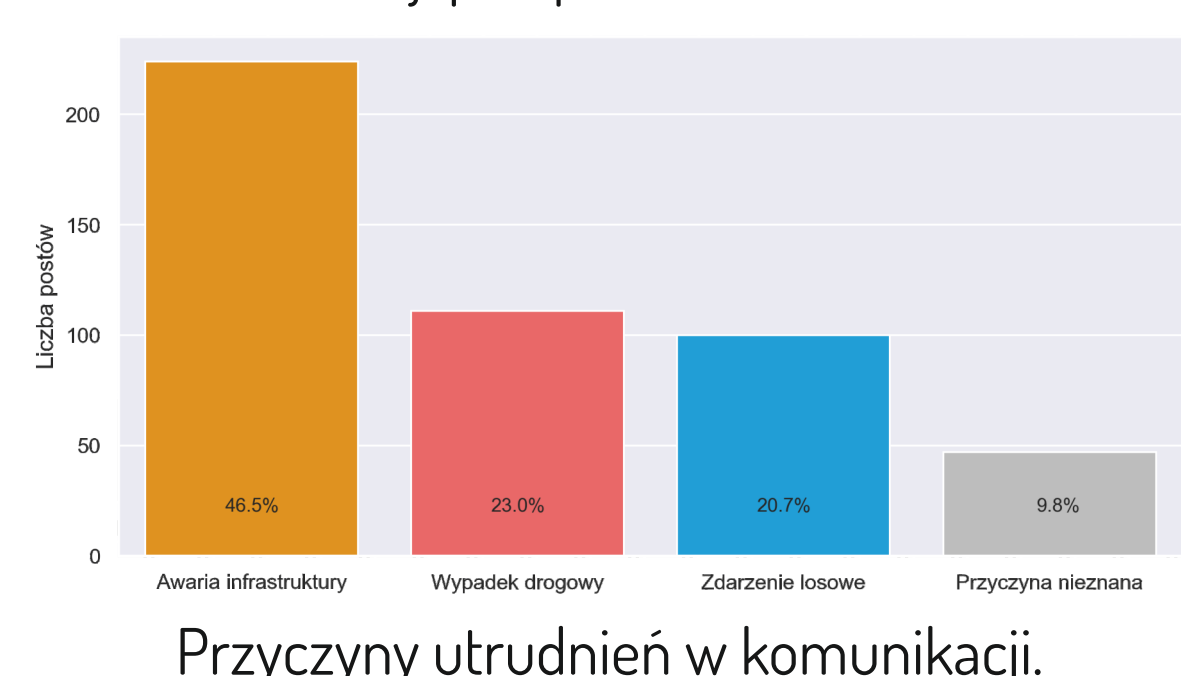
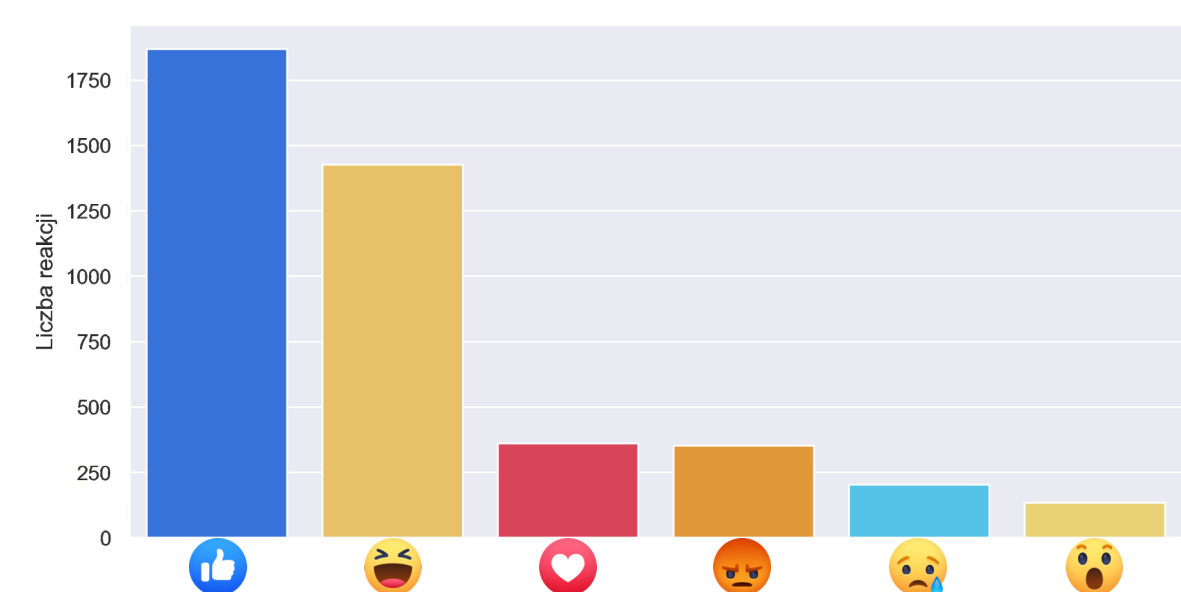
Celem projektu było zebranie postów i komentarzy z oficjalnej strony MPK Wrocław, a następnie analiza wpisów dotyczących utrudnień w ruchu oraz zbadanie wydźwięku komentarzy pod nimi.

479 postów o utrudnieniach.
Incydent średnio
co 2.28 dnia.

Wykresy



Wydźwięk komentarzy pogrupowanych miesięcznie w których zostały napisane.



Dane

Dane użyte w projekcie pochodziły z wielu źródeł:

- Posty oraz komentarze - Facebook (profil MPK Wrocław)
- Tweety - Twitter (profil AlertMPK)
- Historyczne dane pogodowe - OpenWeather
- Pozycje przystanków - Otwarte Dane Wrocław
- Lokalizacje ulic i skrzyżowań - Google Maps oraz OpenStreetMaps

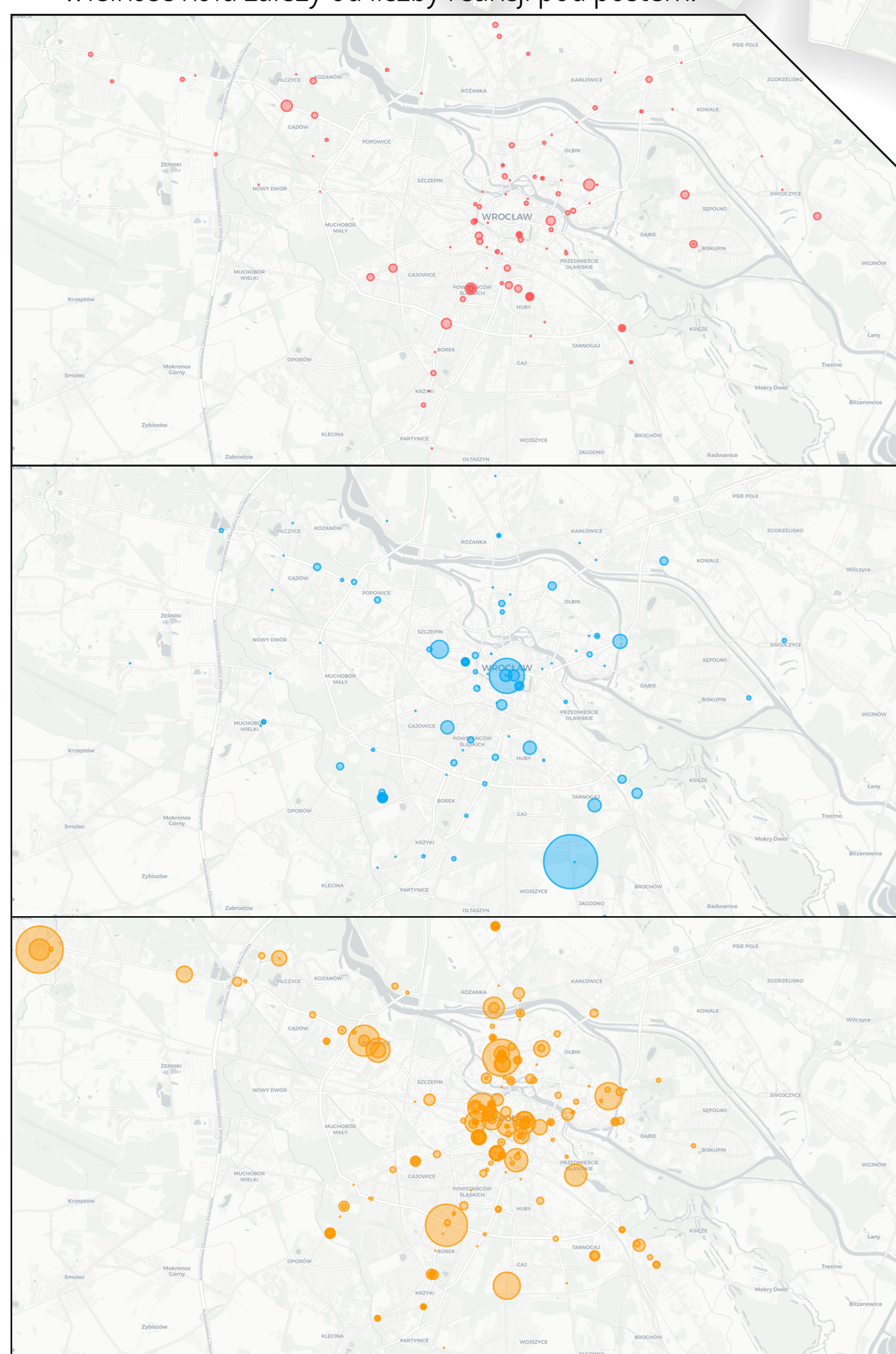
Łącznie zebrano 1680 postów, 24809 komentarzy i 4161 tweetów z okresu od 1 stycznia 2017 do 31 grudnia 2019.

Obszar z największą liczbą utrudnień: 20

Mapa z zaznaczonymi utrudnieniami w ruchu komunikacji miejskiej

- Wypadki drogowe (również niezależne od MPK)
- Awarie infrastruktury
- Zdarzenia losowe
- Przyczyna nieznana

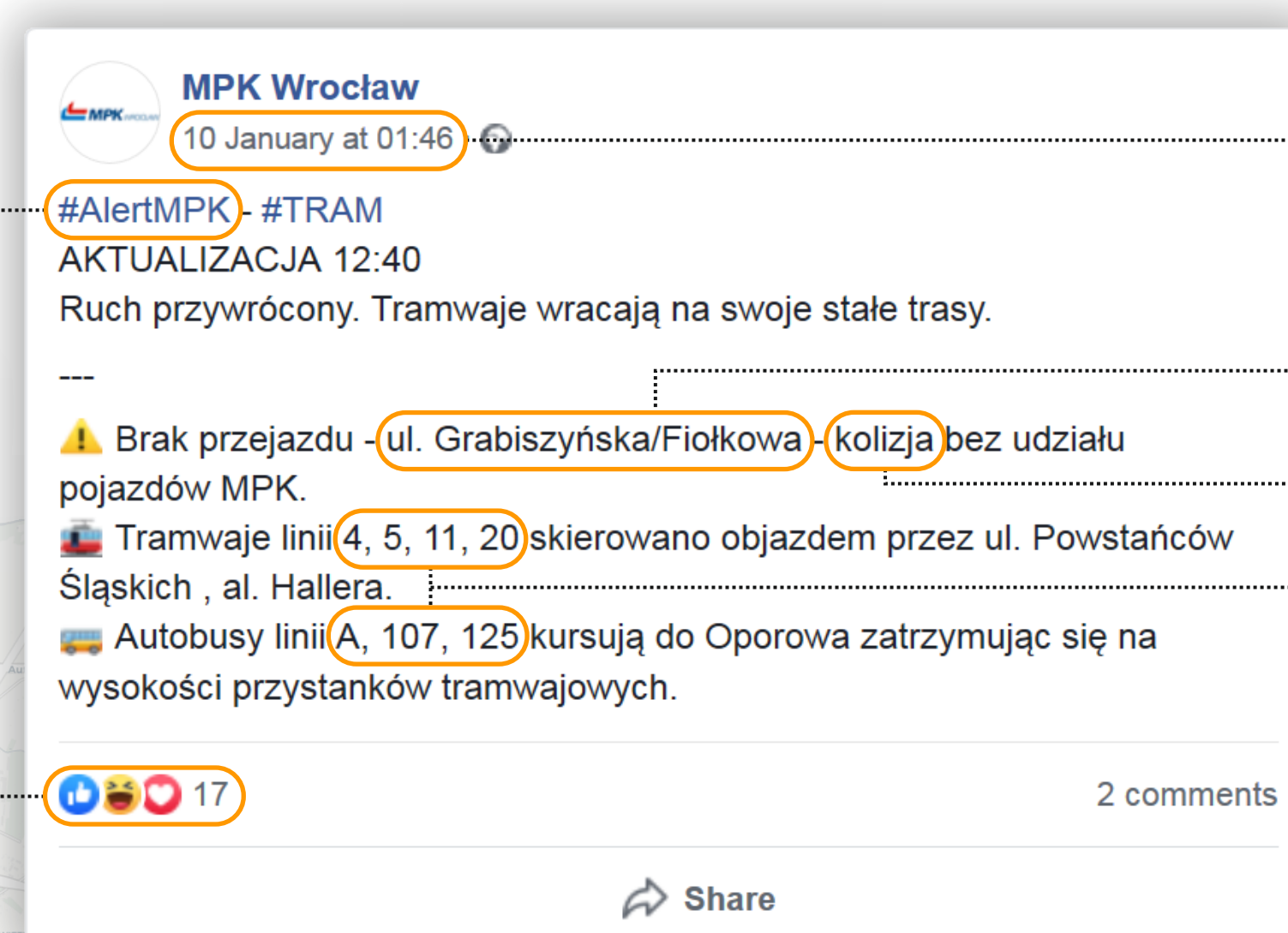
Mapy przedstawiają pozycje incydentów. Od góry: wypadki drogowe, zdarzenia losowe, awarie. Wielkość koła zależy od liczby reakcji pod postem.



Analiza danych

Hashtag

Reakcje



Data

Pozycja

Typ

Linie

Blisko połowa (46.5%) utrudnień wynika z awarii infrastruktury.

Post z największą liczbą komentarzy: 144
2018-03-02 6:00
Ul. Powstańców Śląskich - uszkodzona sieć trakcyjna.

Post z największą liczbą reakcji: 105
77 - 😂 14 - 🍷 13 - 🍷 1 - 🍷
2018-09-03 17:30
Ul. Buforowa/Konduktorska - zły stan nowo otwartej drogi.

Podsumowanie

- W badanej dziedzinie spodziewaną rzeczą jest brak pozytywnego wydźwięku w wypowiedziach. Przeważającą częścią są jednak wypowiedzi neutralne.
- Alarmującym stanem rzeczy jest przewaga awarii w ogólnym udziale wszystkich utrudnień w badanym okresie czasu.
- Badanie wydźwięku komentarzy okazało się być trudnym zadaniem m.in. ze względu na występujący w wypowiedziach sarkazm. Sens wielu komentarzy zależał mocno od kontekstu (od treści posta lub innych komentarzy), wypowiedzi miały charakter mocno potoczny, występowały liczne literówki a sama treść komentarza zwykle była krótka i niejednoznaczna.

Przetwarzanie danych

1. Pobranie danych z profilu MPK Wrocław.
2. Odfiltrowanie postów na podstawie hashtaga #AlertMPK, odrzucenie pustych komentarzy, wyciszczenie informacji o użytkownikach.
3. Kategoryzacja postów z użyciem modelu FastText (dokładność 87%) na 7 kategorii, z których 4 są dalej przetwarzane.
4. Odszukanie informacji o miejscu zdarzenia i odpytanie Open Street Maps o dokładną pozycję.
5. Kategoryzacja wydźwięku komentarzy z użyciem drugiego modelu FastText (dokładność 70 %) w podziale na klasy: negatywny, obojętny, pozytywny.

Projekt został zaimplementowany w Pythonie 3 z użyciem następujących bibliotek: fastText, Folium, GeoPy, NumPy, Pandas, Seaborn, Selenium.

{227961, 221060, 231740}@student.pwr.edu.pl