

Ćwiczenie 11

Cele

Podstawowym celem tego ćwiczenia było utworzenie interaktywnego dashboard'u w programie Power BI wykorzystując poniżej opisane dane. Ponadto przed utworzeniem dashboard'u zakres tego ćwiczenia objął również przygotowanie danych (ich zrozumienie, wyczyszczenie, wymodelowanie) do wizualizacji.

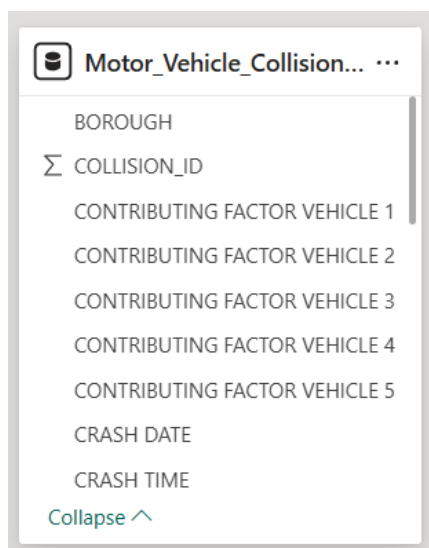
Opis danych

Dane zostały zaimportowane do pliku z rozszerzeniem .csv ze strony: https://data.cityofnewyork.us/Public-Safety/Motor-Vehicle-Collisions-Crashes/h9gi-nx95/about_data

W pobranym pliku zawarte są informacje na temat wypadków/kolizji drogowych w Nowym Yorku. Każdy rekord jest osobnym wypadkiem/kolizją. Dane są dosyć obszerne, zajmują około 500 megabajtów danych, co przekłada się na ponad 2 mln rekordów. Obejmują one

Realizacja

Na początku dane zostały zaimportowane do Power BI:



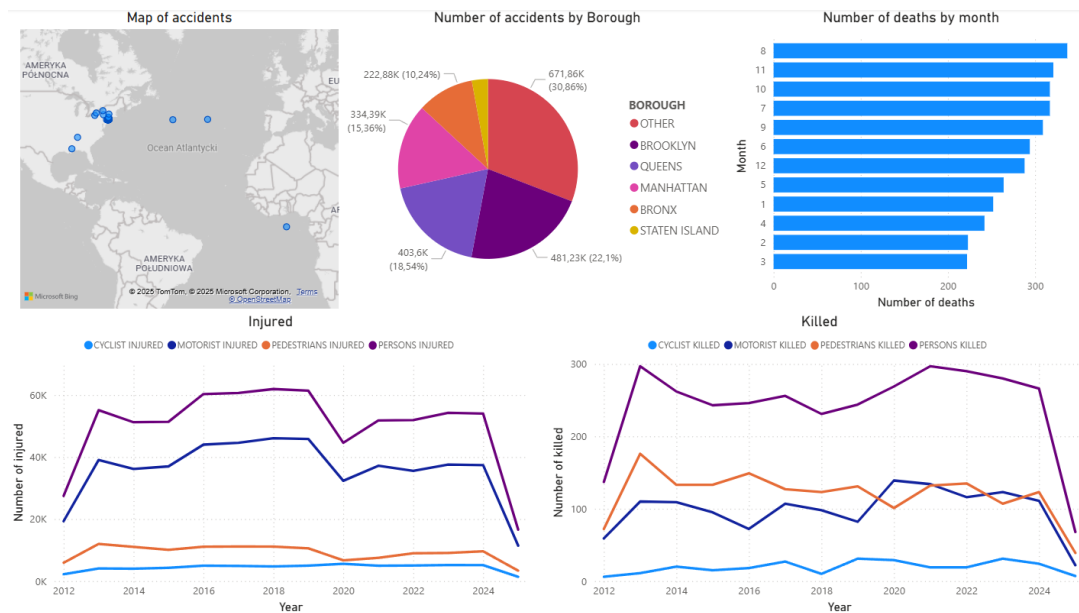
Następnie przeprowadzono proces zaznajamiania się z danymi, a także przygotowywania ich i modelowania w celu wizualizacji za pomocą dashboard'u. Wykonano następujące kroki:

- Przekonwertowano kolumnę CRASH DATE z formatu MM/DD/YYYY do formatu DD.MM.YYYY, a następnie zmieniono jej typ z string na date. Na jej podstawie została utworzona hierarchia czasowa.

- Zamieniono typy danych przedstawiających ilość ofiar śmiertelnych i osób poszkodowanych na wartości numeryczne.
- Zmieniono wartości brakujące i wartości „0” w kolumnie z nazwą dzielnicy na Other
- Wyciągnięto z danych o dacie incyduentu wartość miesiąca.
- Inne operacje przyczyniające się do jak największej liczby wartościowych informacji, jakie raport może za sobą nieść.

Podsumowanie

Screenshot utworzonego dashboard'u znajduje się poniżej:



Ponadto do sprawozdania dołączono nagranie pokazujące interaktywność dashboard'u.