

Analiza liczby śmierci w wypadkach drogowych w wybranych państwach Europy w 2019 roku

Michał Wasiński
Informatyka Stosowana
Politechnika Wrocławska
272674@student.pwr.edu.pl

I. WSTĘP

Samochody, choć towarzyszą nam stosunkowo niedługo, każdego roku przyczyniają się do wielu przedwczesnych śmierci na całym świecie. W 2019 roku w Europie odnotowano tysiące śmiertelnych wypadków drogowych, co stanowi poważny problem społeczny i zdrowotny. Różnice w liczbie śmierci w wypadkach samochodowych między krajami są znaczne, co sugeruje, że istnieje wiele różnorodnych czynników wpływających na bezpieczeństwo na drogach.

Celem niniejszego projektu jest zbadanie zależności liczby śmierci w wypadkach samochodowych na milion mieszkańców od różnych czynników w wybranych krajach Europy w 2019 roku. W szczególności, analiza koncentruje się na takich zmiennych, jak udział procentowy samochodów starszych niż 10 lat, liczba samochodów na 1000 mieszkańców, wysokość maksymalnych mandatów za przekroczenie prędkości o ponad 20 km/h w terenie niezabudowanym oraz przynależność państwa do określonego regionu geograficznego.

Przeprowadzone analizy pozwolą na identyfikację, które z tych czynników mają największy wpływ na liczbę śmierci w wypadkach drogowych. Dzięki temu możliwe będzie wskazanie obszarów, w których interwencje mogą przynieść największe korzyści w zakresie poprawy bezpieczeństwa drogowego.

II. ZBIÓR DANYCH I JEGO PRZETWARZANIE

A. Zbiór danych

Zgromadzone zbiory danych zostały pobrane z następujących źródeł:

- **Eurostat:** Liczba samochodów na tysiąc mieszkańców [1], liczba samochodów z podziałem na wiek i w sumie [2], populacja [3], liczba śmierci w wypadkach drogowych [4].
- **speedingeurope.com:** Wysokości mandatów za przekroczenie prędkości o ponad 20 km/h w terenie niezabudowanym [5].
- **worldbank.org:** Powierzchnie państw [6].

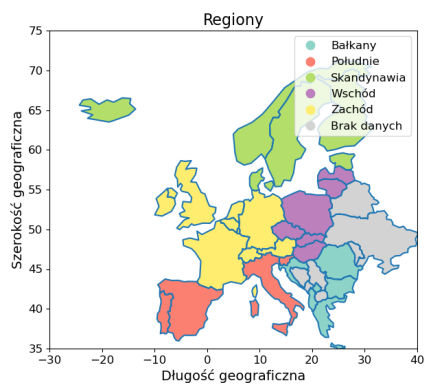
Dane dotyczące regionów zostały dodane przeze mnie. Mają grupować państwa, szczególnie ze względu na podobieństwo kulturowe.

B. Przetwarzanie wstępne

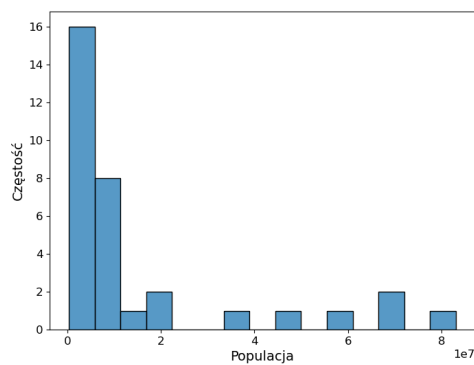
Przed przystąpieniem do analizy danych, wykonano następujące czynności:

- Uzupełniono brakujące dane:
 - Wysokości mandatów za przekroczenie prędkości dla Bułgarii [7] i Rumunii [8].
 - Liczby śmierci w wypadkach drogowych dla Albanii [9], Kosowa [10] i Wielkiej Brytanii [11].
 - Liczbę samochodów dla Wielkiej Brytanii [12].
 - Informację o powierzchni Kosowa [13].
- Pozmieniano nazwy kolumn dla lepszej czytelności i spójności.
- Obliczono liczbę śmierci w wypadkach samochodowych na milion mieszkańców oraz procentowy udział samochodów starszych niż 10 lat w ogólnej liczbie samochodów.
- Zmieniono kody regionów na ich pełne nazwy dla łatwiejszej interpretacji.
- Uzupełniono brakujące wartości dla procentowego udziału starych samochodów i wysokości mandatów średnimi wartościami dla danego regionu.
- Usunięto wiersze niezawierające współczynnika śmierci, aby zapewnić spójność danych.
- Usunięto kolumny zawierające informacje o liczbie starych samochodów, liczbie wszystkich samochodów oraz liczbie śmierci w wypadkach drogowych.

Dokonany podział krajów na regiony został przedstawiony na poniższej grafice (1). Państwa zaznaczone na szaro nie znajdują się w zbiorze danych.



Rysunek 1. Mapa regionów



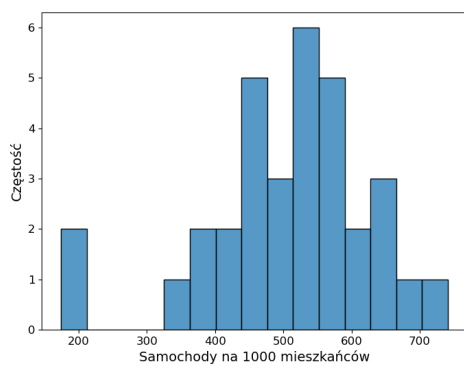
Rysunek 3. Histogram populacji

C. Analiza eksploracyjna

Poniżej przedstawiono statystyki i histogramy dla poszczególnych danych.

Tabela I
STATYSTYKI LICZBY SAMOCHODÓW NA TYSIĄC MIESZKAŃCÓW

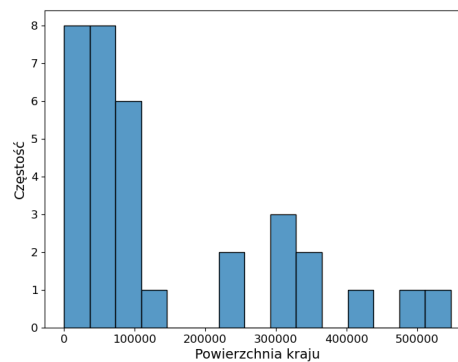
Minimum	Maximum	Średnia	Mediana
174.0	741.0	506.66	530.0



Rysunek 2. Histogram ilości samochodów na tysiąc mieszkańców

Tabela III
STATYSTYKI POWIERZCHNI

Minimum	Maximum	Średnia	Mediana
320.0	547557.0	144911.86	77205.5



Rysunek 4. Histogram powierzchni

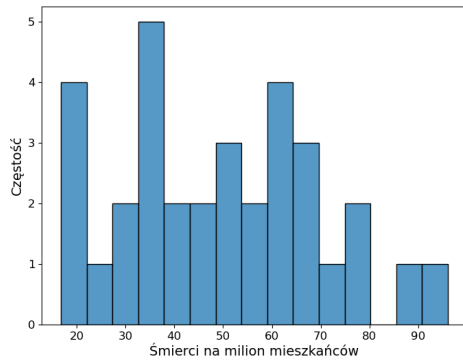
Tabela II
STATYSTYKI POPULACJI

Minimum	Maximum	Średnia	Mediana
356991	83019213	16124067.09	7000039.0

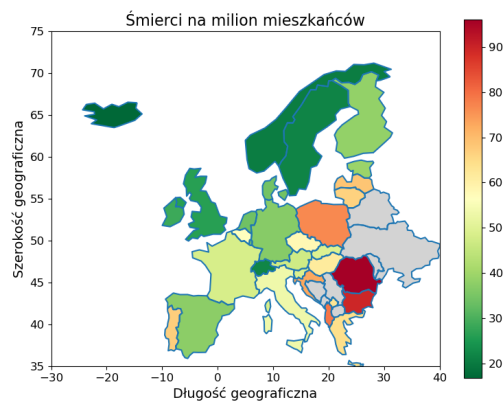
Tabela IV
STATYSTYKI ŚMIERCI W WYPADKACH SAMOCHODOWYCH NA MILION MIESZKAŃCÓW

Minimum	Maximum	Średnia	Mediana
16.81	96.01	49.98	49.02

W celu określenia zależności liczby śmierci od poszczególnych parametrów, stworzona została poniższa mapa korelacji.



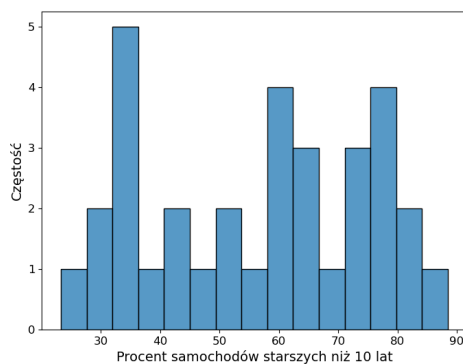
Rysunek 5. Histogram współczynnika śmierci



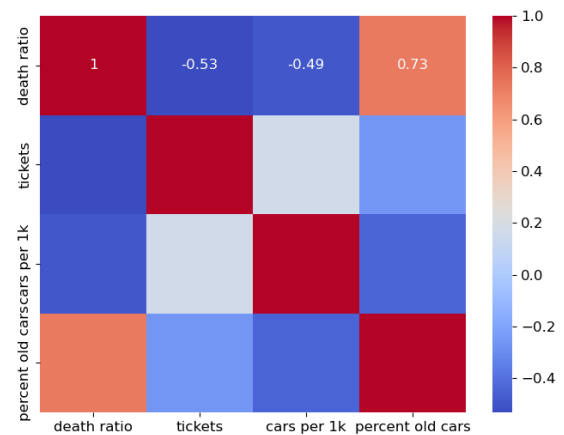
Rysunek 6. Mapa współczynnika śmierci

Tabela V
STATYSTYKI PROCENTOWEGO UDZIAŁU SAMOCHODÓW STARSZYCH NIŻ 10 LAT

Minimum	Maximum	Średnia	Mediana
23.31	88.47	57.23	61.38



Rysunek 7. Histogram procentowego udziału starych samochodów



Rysunek 8. Mapa korelacji

III. EKSPERYMENTY

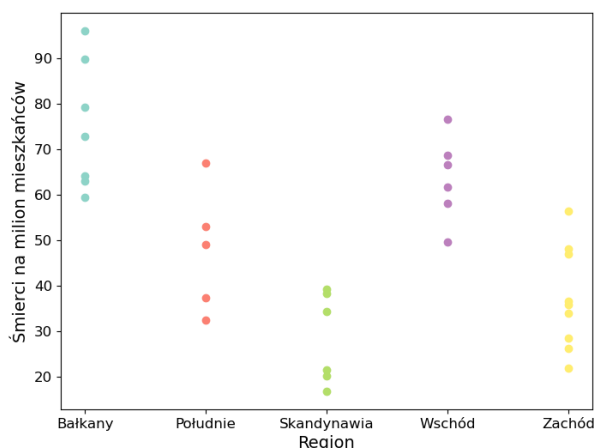
1) Eksperyment 1: Wpływ regionu na wskaźnik śmiertelności

Opis eksperymentu

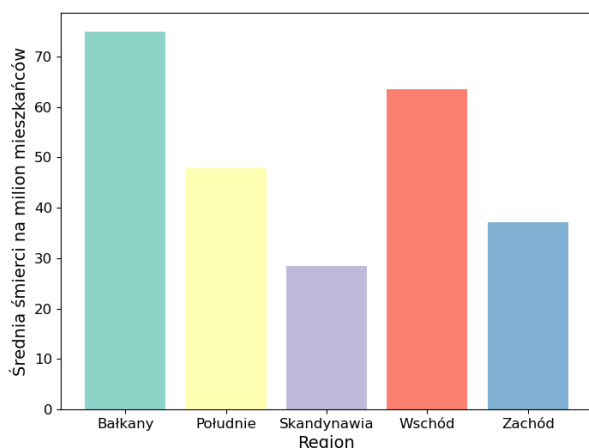
Celem eksperymentu było zbadanie, czy przynależność do określonego regionu geograficznego ma wpływ na wskaźnik śmiertelności w wypadkach drogowych. Analiza została przeprowadzona dla regionów: Zachód, Wschód, Południe, Skandynawia, Bałkany. Dane zostały podzielone na regiony, a następnie przeanalizowano wskaźniki śmiertelności dla każdego z nich. Do wizualizacji wyników użyto dwóch wykresów: wykresu punktowego (9), na którym wskaźniki śmiertelności są zaznaczone na osi pionowej, a regiony na osi poziomej, oraz wykresu słupkowego przedstawiającego średnie wartości wskaźników śmiertelności dla każdego regionu (10).

Wyniki

Wyniki wskazują na wyraźne różnice w wskaźniku śmiertelności pomiędzy regionami. Najmniejsza liczba śmiertelnych wypadków drogowych została zaobserwowana w regionie Skandynawii, natomiast największa na Bałkanach. Ponadto, dane z danego regionu wykazują dużą zgodność wewnątrz regionów, co sugeruje, że region geograficzny jest istotnym czynnikiem wpływającym na wskaźnik śmiertelności. Analiza danych sugeruje, że różnice w kulturze jazdy mogą mieć wpływ na wskaźnik śmiertelności w poszczególnych regionach.



Rysunek 9. Śmierci z podziałem na regiony



Rysunek 10. Średnie liczb śmierci z podziałem na regiony

2) Eksperyment 2: Zależność wysokości mandatów za przekroczenie prędkości a wskaźnik śmiertelności

Opis eksperymentu

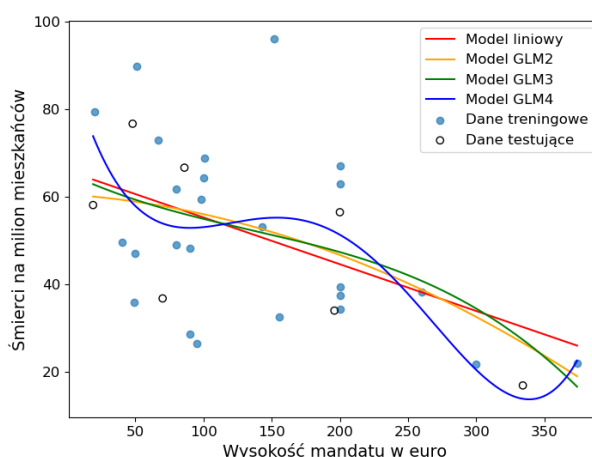
Celem eksperymentu było zbadanie, czy wysokość maksymalnych mandatów za przekroczenie prędkości o ponad 20 km/h w terenie niezabudowanym ma wpływ na liczbę śmierci w wypadkach samochodowych. W analizie wykorzystano modele regresji liniowej oraz generalizowanych modeli liniowych (GLM) o stopniach 2, 3 i 4. Norwegia została pominięta na wykresie, ze względu na znaczną odmienną wysokość mandatu w porównaniu do pozostałych krajów.

Wyniki

Eksperyment wykazał, że wyższe mandaty za przekroczenie prędkości są skorelowane z mniejszą liczbą śmierci w wypadkach samochodowych. Pomimo pewnych różnic, wszystkie modele - zarówno regresja liniowa, jak i generalizowane modele liniowe - potwierdziły tę zależność, co jest widoczne na poniższym wykresie (11). Wyniki błędów średniokwadratowych dla poszczególnych modeli wskazują na najlepsze dopasowanie do danych dla modelu GLM o stopniu 2.

Błędy średniokwadratowe:

- Model liniowy: 185.0
- Model GLM2: 174.4
- Model GLM3: 176.5
- Model GLM4: 202.8



Rysunek 11. Zależność współczynnika śmierci od wysokości mandatów

3) Eksperyment 3: Zależność liczby samochodów na 1000 mieszkańców a wskaźnik śmiertelności

Opis eksperymentu

W ramach eksperymentu zbadano związek pomiędzy liczbą samochodów na 1000 mieszkańców a wskaźnikiem śmiertelności w wypadkach drogowych. Przed przeprowadzeniem analizy pominięto outliery (2 najmniejsze wartości dla samochodów na 1000 mieszkańców). Następnie użyto modeli regresji liniowej oraz generalizowanych modeli liniowych (GLM) rzędu 2, 3 i 4. Dane testujące stanowiły 20% całości. Wyniki eksperymentu zostały przedstawione na wykresie (12).

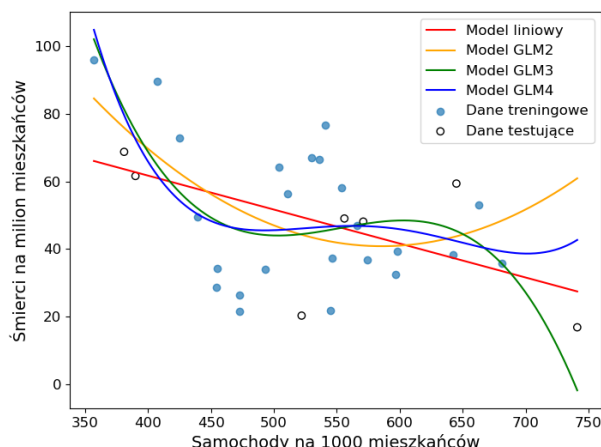
Wyniki

Eksperyment wykazał, że wraz ze wzrostem liczby samochodów na 1000 mieszkańców, powoli maleje liczba śmierci w wypadkach drogowych. Otrzymane wyniki błędów średniokwadratowych

dla poszczególnych modeli wskazują na najlepsze dopasowanie do danych dla modelu GLM rzędu 3.

Błędy średniokwadratowe:

- Model liniowy: 215.7
- Model GLM2: 435.2
- Model GLM3: 210.0
- Model GLM4: 266.8



Rysunek 12. Zależność współczynnika śmierci w zależności od liczby samochodów na tysiąc mieszkańców

4) Eksperyment 4: Zależność udziału samochodów starszych niż 10 lat a wskaźnik śmiertelności

Opis eksperymentu

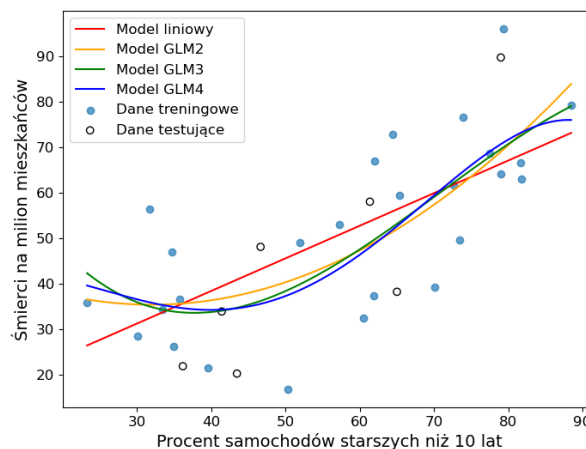
Eksperyment miał na celu zbadanie związku między udziałem samochodów starszych niż 10 lat w ogólnej liczbie samochodów a wskaźnikiem śmiertelności. Użyto modeli regresji liniowej oraz generalizowanych modeli liniowych (GLM) rzędu 2, 3 i 4. Dane testujące stanowiły 20% całego zbioru danych. Wyniki eksperymentu zostały przedstawione na wykresie (13).

Wyniki

Analiza wykazała, że ryzyko śmiertelności rośnie wraz z wiekiem samochodów. Modele GLM o stopniach 3 i 4 osiągnęły najniższy błąd średniokwadratowy, co sugeruje ich lepsze dopasowanie do danych w porównaniu do pozostałych modeli. Eksperyment potwierdza istotność wpływu wieku samochodów na ryzyko śmiertelności w wypadkach drogowych.

Błędy średniokwadratowe:

- Model liniowy: 223.1
- Model GLM2: 184.0
- Model GLM3: 171.2
- Model GLM4: 170.8



Rysunek 13. Zależność współczynnika śmierci od ilości starych samochodów

IV. WNIOSKI

Przeprowadzone badanie ukazało, że bezpieczeństwo drogowe jest wielowymiarowym problemem, na który wpływają różnorodne czynniki. Analiza danych wykazała istotny związek między wskaźnikiem śmiertelności w wypadkach drogowych a takimi zmiennymi jak wiek samochodów, liczba pojazdów na mieszkańca, wysokość mandatów za przekroczenie prędkości, a także region geograficzny.

Wyniki wskazały, że większa liczba samochodów na 1000 mieszkańców oraz wyższe mandaty za przekroczenie prędkości miały tendencję do malejącej liczby śmierci w wypadkach drogowych. Warto zatem rozważyć podniesienie kar za przekraczanie prędkości, nie tylko w celu zwiększania budżetu państwa. Natomiast, jak się można było spodziewać, udział samochodów starszych niż 10 lat często korelował z większą liczbą wypadków drogowych. Ważnym czynnikiem okazała się również przynależność do danego regionu geograficznego. Sugeruje to, że na liczbę śmierci w wypadkach samochodowych mocno wpływa również kultura jazdy.

Pomimo używania zaawansowanych modeli regresji liniowej i generalizowanych modeli liniowych, dokładność prognozowania wskaźnika śmiertelności okazała się zróżnicowana. Czasem modele te radziły sobie dobrze, ale w niektórych przypadkach ich skuteczność była ograniczona. Mała liczba danych sprawiła, że modele w dużym stopniu zależały od tego, jakie dane testujące zostały wylosowane. Można jednak zauważyć, że konkretne modele radziły sobie lepiej w danym przypadku.

Wnioski z projektu mogą być przydatne dla specjalistów ds. bezpieczeństwa drogowego. Identyfikacja czynników wpływających na bezpieczeństwo na drogach może prowadzić do skuteczniejszych działań mających na celu zmniejszenie liczby wypadków drogowych.

Mimo uzyskanych wyników, istnieje potrzeba dalszych badań nad zależnościami między czynnikami a wskaźnikiem śmiertelności w wypadkach drogowych. Problem jest złożony i wymaga analizy większej ilości danych.

LITERATURA

- [1] [dostęp: 2024-04-23]. Dostępny w Internecie: https://ec.europa.eu/eurostat/databrowser/view/road_eqs_carhab/default/table?lang=en&category=road.road_eqs.
- [2] [dostęp: 2024-05-08]. Dostępny w Internecie: https://ec.europa.eu/eurostat/databrowser/view/road_eqs_carage/default/table?lang=en&category=road.road_eqs.
- [3] [dostęp: 2024-04-23]. Dostępny w Internecie: https://ec.europa.eu/eurostat/databrowser/view/tps00001/default/table?lang=en&category=t_demo.t_demo_pop.
- [4] [dostęp: 2024-04-23]. Dostępny w Internecie: https://ec.europa.eu/eurostat/databrowser/view/tran_sf_roadro__custom_11036996/default/table.
- [5] [dostęp: 2024-04-23]. Dostępny w Internecie: <https://speedingeurope.com/>.
- [6] [dostęp: 2024-05-18]. Dostępny w Internecie: <https://data.worldbank.org/indicator/AG.LND.TOTL.K2?locations=1W>.
- [7] [dostęp: 2024-05-18]. Dostępny w Internecie: <https://www.kgmp-legal.com/en/traffic-offences-and-fines-in-bulgaria/>.
- [8] [dostęp: 2024-05-18]. Dostępny w Internecie: https://en.wikipedia.org/w/index.php?title=Speed_limits_in_Romania&oldid=883033884.
- [9] [dostęp: 2024-05-18]. Dostępny w Internecie: <https://irdeng.com/2024/04/16/ird-engineering-to-support-the-development-of-the-new-national-road-safety-strategy-in-the-republic-of-albania/>.
- [10] [dostęp: 2024-05-18]. Dostępny w Internecie: https://www.researchgate.net/publication/357972505_Role_of_Emergency_Medical_Services_in_pre-hospital_Management_of_Road_Accidents_appropriate_on-site_Actions.
- [11] [dostęp: 2024-05-18]. Dostępny w Internecie: <https://www.gov.uk/government/statistics/reported-road-casualties-great-britain-annual-report-2019>.
- [12] [dostęp: 2024-05-18]. Dostępny w Internecie: <https://assets.publishing.service.gov.uk/media/5e610a50d3bf7f108502ecaa/tsgb-2019.pdf>.
- [13] [dostęp: 2024-05-18]. Dostępny w Internecie: <https://en.wikipedia.org/wiki/Kosovo>.