**Statystyka (II) – Laboratorium II**

**(Tworzenie bazy danych ze źródeł Internetowych)**

**Zadanie 1**

Na podstawie bazy danych GUS stworzyć plik dotyczący infrastruktury transportowej
w województwach Polski w 2016 roku. (Plik danych w Statystyce)

Należy wejść na stronę Głównego Urzędu Statystycznego *www. stat. gov. pl* a następnie wybrać Banki i bazy danych 🡪Bank Danych Lokalnych. Do analizy pobrać zmienne:

X1 – Ludność (Ludność)\*

X2 – Powierzchnia (Podział terytorialny 🡪 Powierzchnia geodezyjna kraju [km2]

X3 – Pojazdy samochodowe i ciągniki (Transport i łączność 🡪Pojazdy🡪Pojazdy ogółem)

X4 – Drogi publiczne o nawierzchni twardej (Transport i łączność – drogi publiczne powiatowe wg. typu nawierzchni)

X5 – Długość ścieżek rowerowych (Transport i łączność – komunikacja miejska)

X6 – Wypadki drogowe (Transport i łączność – wypadki drogowe).

\*wchodząc do Banku Danych lokalnych wybrać opcję Dane według dziedzin a następnie odpowiednią zmienną np. Ludność dalej Stan Ludności i wybrać Ludność wg grup wieku i płci. Zatwierdzić na dole przyciskiem *Dale*j. Otworzy się okno, gdzie można określić kategorie względem, których pobiera się dane.

Lata 🡪2016

Wiek🡪Ogółem

Płeć 🡪Ogółem

Następnie zatwierdzić przyciskiem *Dalej*. Kolejno wybiera się jednostki terytorialne, wg których prezentowane są dane Wybór jednostek terytorialnych wg układu TERYT. Należy wybrać tylko województwa.

Po wyborze jednostki terytorialnej baza powinna zapamiętać wybór i kolejne dane będą pobierane tylko dla województw.

Pod okienkami wyboru podświetli się przycisk Dalej i po naciśnięciu otrzymujemy wybrane dane.

Należy skopiować je do Excela (lub pobrać już ze strony plik w Excelu) a potem przenieść do Statystyki.

Z wybranych zmiennych stworzyć wskaźniki (przeliczyć na liczbę ludności lub na powierzchnię np.:).

X7-wskaźnik 1 pojazdy samochodowe i ciągniki na drogi publiczne o twardej nawierzchni przypadające na 1 km

X8- wskaźnika 2 drogi publiczne o nawierzchni twardej na 100 km2

X9- wskaźnik 3 długość ścieżek rowerowych na tys. km2

X10- wskaźnik 4 liczba wypadków drogowych na 10 tys. ludności.

**Po stworzeniu wskaźników nadać nazwy województwom w kolumnie oznaczającej Liczbę porządkową**.

Na podstawie stworzonej bazy danych:

1. Wybrać jeden wskaźnik i dokonać charakterystyki za pomocą statystyk opisowych (średnia, mediana, modalna, odchylenie standardowe, typowy przedział zmienności, współczynnik zmienności, kwartyl dolny i górny), określić asymetrię i kurtozę. (wskazać województwo o najmniejszym i największym poziomie wskaźnika).
2. Wybrać dwa wskaźniki i sprawdzić czy istnieje zależność pomiędzy nimi. Policzyć współczynnik korelacji, podać równanie regresji i dokonać interpretacji.
3. Wskazać najlepsze i najgorsze województwa pod względem badanych wskaźników. **(Opcja Sortuj Dane przedstawić na wykresie słupkowym. Wykresy 🡪 Wykresy 2W🡪 Wykresy słupkowe)**