

## Ćwiczenia 8 & 9 (Wizualizacja i przetwarzanie danych)

- Zbiory danych dla zadań 1-8 znajdują się w pakiecie `PogromcyDanych` (wersja angielska).
- Z zadań 13-15 należy rozwiązać co najmniej 2 zadania i przesłać rozwiązania (kod jak i gotowe mapy) poprzez MS Teams. Brak rozwiązań będzie skutkowało **brakiem zaliczenia ćwiczeń!!!**

1. Dla zbioru danych `cats_birds` przedstaw zależność pomiędzy zmiennymi `length` oraz `speed` na wykresie rozrzutu. Dodaj trend liniowy do wykresu. Obie grupy zwierząt powinny mieć własną prostą regresji oraz typ i kolor punktów.

2. **(S)** W zbiorze danych `pearson` mamy informacje o wzroście ojców i synów. Przedstaw zależność pomiędzy tymi cechami na wykresie rozrzutu. Dodaj trend liniowy bez przedziału ufności.

3. W zbiorze danych `seriesIMDB` mamy informacje o odcinkach 198 seriali. Stwórz wykres, w którym na osi Y będzie nazwa serialu, a na osi X wykres pudełkowy ocen odcinków danego serialu. Popraw wykres tak, aby seriale były posortowane według median ocen.

4. W zbiorze danych `diagnosis` mamy zmienną `educ4_2013` z poziomami edukacji i zmienną `gp29` z odpowiedziami na pytanie co jest najważniejsze w życiu. Stwórz wykres słupkowy. Każdy słupek odpowiada jednemu poziomowi edukacji. Każdy słupek powinien być podzielony na części wypełnione różnymi kolorami w zależności od zmiennej `gp29`.

5. **(S)** Ze zbioru `auta2012` wybierz jedynie auta Volkswagen Passat. Użyj funkcji `geom_smooth()`, aby przedstawić zależność pomiędzy rokiem produkcji i ceną. Spróbuj zastosować różne szablony wykresu.

6. Zmień poniższy wykres:

```
ggplot(cats_birds, aes(x = weight, y = speed, size = lifespan, color = lifespan)) +  
  geom_point()
```

Zastąp skalę kolorów na gradient od koloru zielonego do czerwonego i zastąp kropki kwadratami. Dodaj tytuł i etykiety osi.

7. Ze zbioru danych `auta2012` wybierz wszystkie samochody marki Toyota. Na wykresie kołowym zwizualizuj liczbę pięciu najpopularniejszych modeli. Następnie te same dane przedstaw na wykresie słupkowym w kolejności malejących liczebności. Nie zapomnij o legendzie.

8. Dla wzrostu synów i ojców ze zbioru danych `pearson` narysuj na jednym wykresie histogram wraz z nieparametryczną oceną gęstości dla każdej cechy. Linie pomiędzy słupkami histogramu pokoloruj na biało. Oba wykresy powinny być obok siebie na jednym diagramie.

9. Dla zbioru danych `iris` przygotuj obrazek jak na Rysunku 1. Rozważ użycie funkcji `geom_density2d()` oraz `theme_light()`.

10. Dla zbioru danych `iris` przygotuj obrazek jak na Rysunku 2.

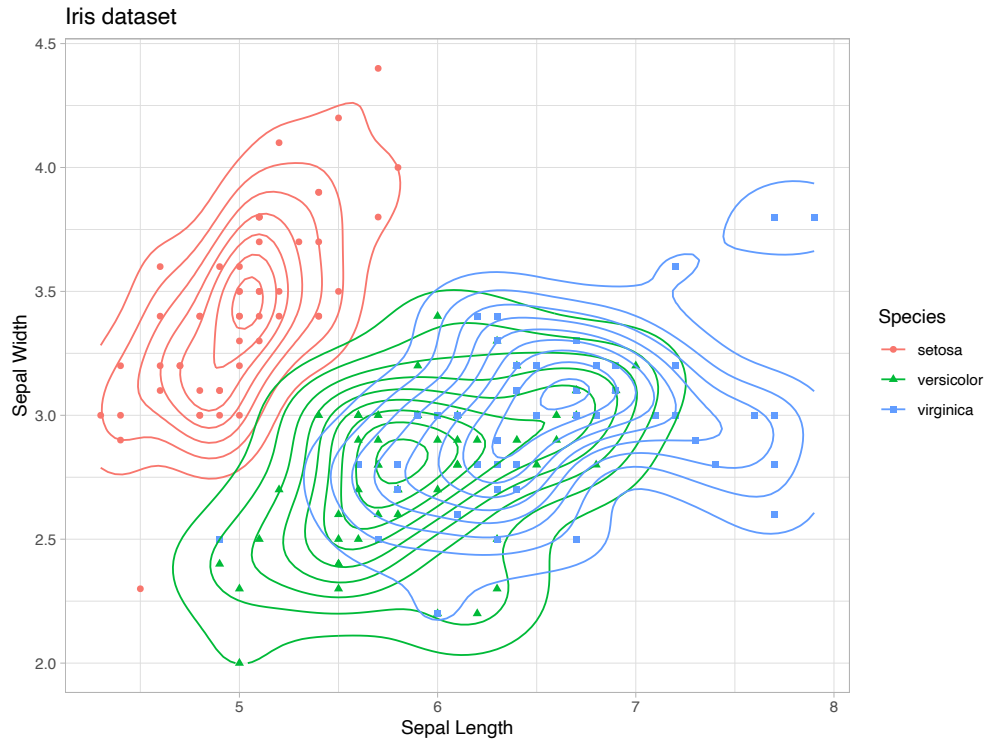
11. **(S)** Dla zbioru danych `mtcars` przygotuj obrazek jak na Rysunku 3.

12. Dla zbioru danych `economics` z pakietu `ggplot2` przygotuj obrazek jak na Rysunku 4. Wykorzystaj szablon `theme_economist_white()` z pakietu `ggthemes`.

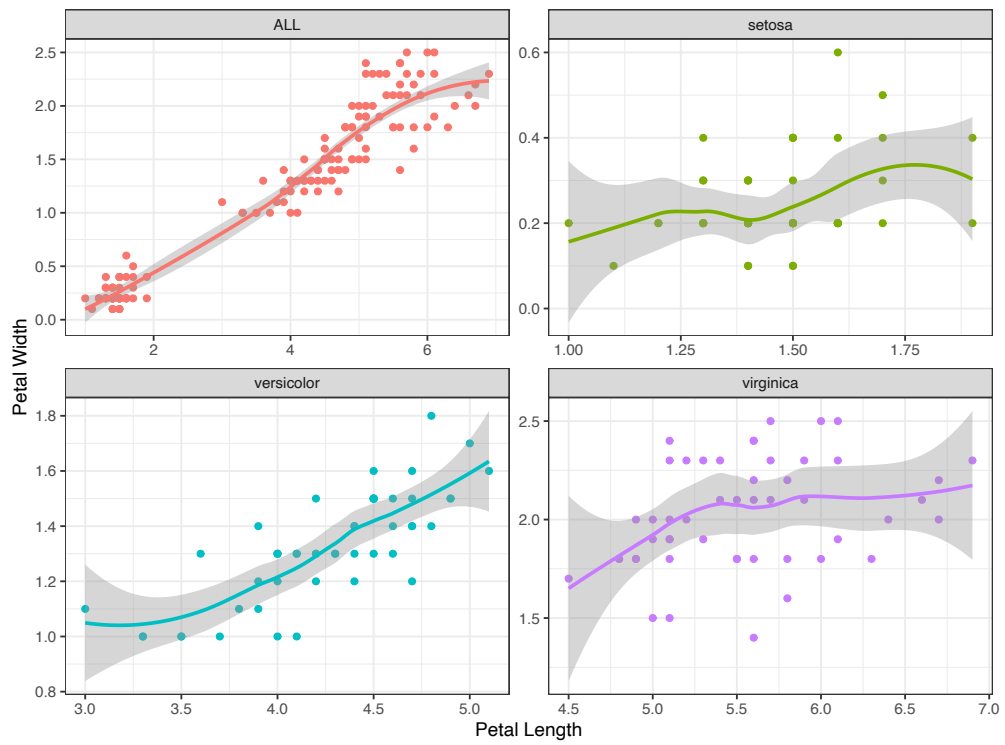
13. Przygotuj mapę basenu Morza Śródziemnego (podobna do mapy Zatoki Meksykańskiej z Wykładu).

14. Przygotuj mapę Alaski (podobna do mapy Florydy z Wykładu).

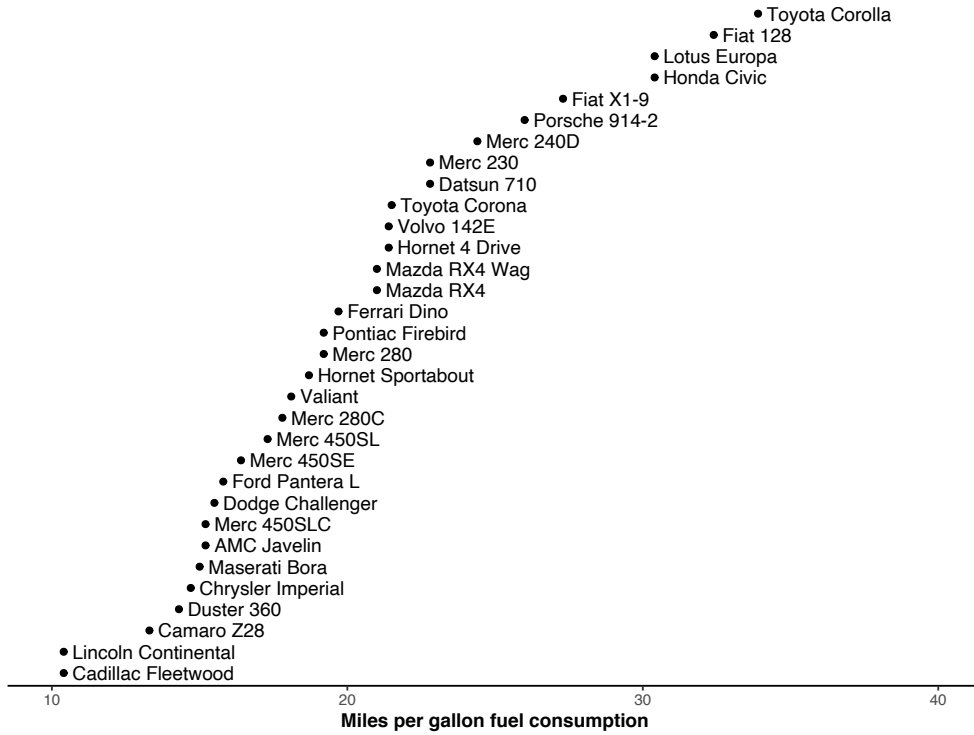
15. **(S)** Przygotuj mapę Danii wraz z Grenlandią (podobna do mapy USA z Wykładu).



RYSUNEK 1. Rysunek do Zadania 9



RYSUNEK 2. Rysunek do Zadania 10



RYSUNEK 3. Rysunek do Zadania 11



RYSUNEK 4. Rysunek do Zadania 12