Eksploracja danych dla Matematyki stosowanej

Ćwiczenia nr 2

**Zad. 1** W pliku *bank\_marketing\_training* będą nas interesować następujące predyktory:

* *age* – wiek,
* *previous\_outcome* – poprzednia odpowiedź,

oraz zmienna celu:

* *response* – odpowiedź.

Wykonaj poniższe polecenia kolejno w R i Pythonie:

1. Wczytaj plik.
2. Sporządź zestawiony wykres słupkowy dla zmiennej *previous\_outcome*, oznaczając kolorem wartości zmiennej *response* (odpowiedź). Następnie zrób znormalizowaną wersję tego wykresu.
3. Wykonaj tabelę krzyżową (kontyngencji) przedstawiającą łączny rozkład dwóch zmiennych: *previous\_outcome* (w kolumnach) i *response* (w wierszach)*.* W tabeli zamieść procenty liczone w kolumnach i zaokrąglone do 1 miejsca po przecinku.
4. Wykonaj zestawiony histogram dla zmiennej *age*, oznaczając kolorem wartości zmiennej *response* (odpowiedź). Następnie zrób znormalizowaną wersję tego wykresu.
5. Przekoduj zmienną *age* na nową zmienną *age\_binned* zgodnie ze schematem:
   1. *,,Under 27”*, jeśli *age<27,*
   2. *,,27 to 60”,* jeśli *27age60,*
   3. *,,Over 60”,* jeśli *age>60.*
6. Wykonaj zestawiony oraz zestawiony i znormalizowany wykres słupkowy dla zmiennej *age\_binned,* oznaczając kolorem wartości zmiennej *response* (odpowiedź).

**Zad. 2 Opis pliku Iris.csv**

Plik zawiera dane o 150 kwiatach irysów. Poszczególne zmienne oznaczają:

* *dlugosc\_kielich* – długość działki kielicha,
* *szerokosc\_kielich* – szerokość działki kielicha,
* *dlugosc\_platek* – długość płatka korony,
* *szerokość\_platek* – szerokość płatka korony,
* *klasa* – gatunek irysa (1 – iris setosa, 2 – iris versicolor, 3 – iris virginica).

Wykonaj poniższe polecenia kolejno w R i Pythonie:

* 1. Oblicz korelacje między wszystkimi zmiennymi. Z jaką zmienną najbardziej skorelowany jest gatunek irysa? Które z pozostałych zmiennych są ze sobą skorelowane?
  2. Wykonaj wykresy rozrzutu lub macierz wykresów, by zwizualizować zależność gatunku irysa od pozostałych zmiennych.
  3. Wykonaj trójwymiarowy wykres rozrzutu zależności gatunku irysa od długości kielicha i długości płatka.