Laboratorium 1

Środowisko pracy

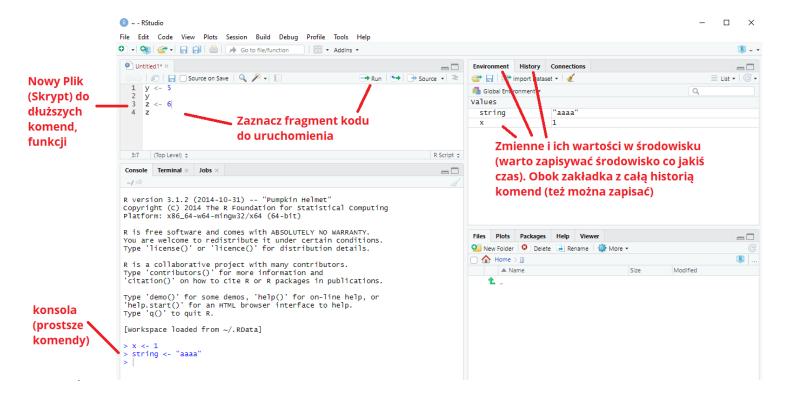
Na zajęciach wybieramy jeden z dwóch języków: **R** lub **Python**, i wykonujemy parę zadań, by zaznajomić się z danym środowiskiem. Wybór należy od Ciebie, spójrz jednak na tabelę niżej.

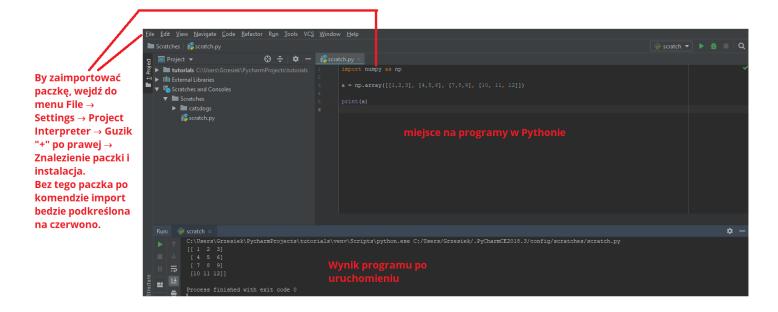
R	Python
Mniej pracy: ściągnij paczkę i	Bardziej elastyczny: ściągnij
uruchom komendę	paczkę i stwórz model
Dla naukowców, badaczy, R&D	Dla programistów, developerów
Mniej popularny, rzadziej w	Bardziej popularny, częściej w
ogłoszeniach o pracę	ogłoszeniach o pracę

Jeśli wybierzesz R, uruchom program R-Studio. Jeśli zdecydujesz się na Pythona, możesz uruchomić środowisko PyCharm (Community Edition). Można też wykorzystać środowisko Anaconda (Spyder do Pythona i RStudio do R).

Poniżej screenshoty z obu środowisk z najważniejszymi komponentami.

R-Studio





Zadanie 1

W tym zadaniu wykonamy kilka komend matematyczno-statystycznych:

- a) Zapisz pod zmienną a liczbę 123, a pod zmienną b liczbę 321. Następnie prostą komendą policz i zwróć wynik mnożenia liczb.
- b) Wczytaj dwa wektory z liczbami [3, 8, 9, 10, 12] oraz [8, 7, 7, 5, 6]. Następnie zwróć sumę tych wektorów oraz iloczyn (po współrzędnych) tych wektorów.
- c) Dla powyższych wektorów podaj iloczyn skalarny i długości euklidesowe.
- d) Stwórz dwie dowolne macierze 3 x 3 pomnóż je po współrzędnych i wyświetl wynik, a następnie pomnóż je macierzowo i wyświetl wynik.
- e) Stwórz wektor 50 losowych liczb z zakresu od 1 do 100.
- f) Dla wektora z punktu (e) policz średnią z wszystkich jego liczb, min, max oraz odchylenie standardowe.
- g) Dokonaj normalizacji wektora z podpunktu (e) (ściskamy wszystkie liczby do przedziału [0,1]) za pomocą poniższego wzoru (x_i to liczba w starym wektorze na pozycji i, a z_i to liczba w nowym wektorze na pozycji i)

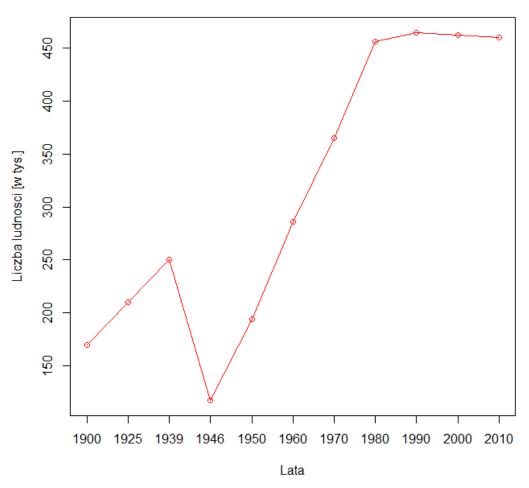
$$z_i = \frac{x_i - \min(x)}{\max(x) - \min(x)}$$

W oryginalnym wektorze jakie było max? Na której pozycji stało? Jaka liczba stoi na tej pozycji w nowym wektorze?

Zadanie 2 W tym zadaniu zrobimy parę operacji na bazie danych i stworzymy wykres.

- a) Załaduj plik miasta.csv do programu w Pythonie lub do środowiska w R. Wyświetl tabelę z danymi za pomocą odpowiedniej komendy.
- b) Dodaj za pomocą odpowiedniej instrukcji wiersz do tabeli z ludnością w 2010 roku: 2010,460,555,405
- c) Stwórz wykres dla ludności Gdańska (w R komendą plot, Pythonie skorzystaj z paczki matplotlib). Dodaj odpowiednie oznaczenia osi, tytuły. Wykres ma być liniowy z punktami i w kolorze czerwonym. Uwaga, w R ustawienie lat na osi x wymaga trochę wysiłku (komendy xaxt w plot i axis). Powinno wyjść coś następującego:

Ludnosc w miastach Polski



d) Stwórz dodatkowo wykres, który będzie zestawiał zmiany ludności wszystkich miasta na jednym wykresie w różnych kolorach. Dodaj legendę.