

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 uruchom komendę	Bardziej elastyczny: ściągnij paczkę i stwórz model
Dla naukowców, badaczy, R&D	Dla programistów, developerów
Mniej popularny, rzadziej w ogłoszeniach o pracę	Bardziej popularny, częściej w 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 screenshotsy z obu środowisk z najważniejszymi komponentami.

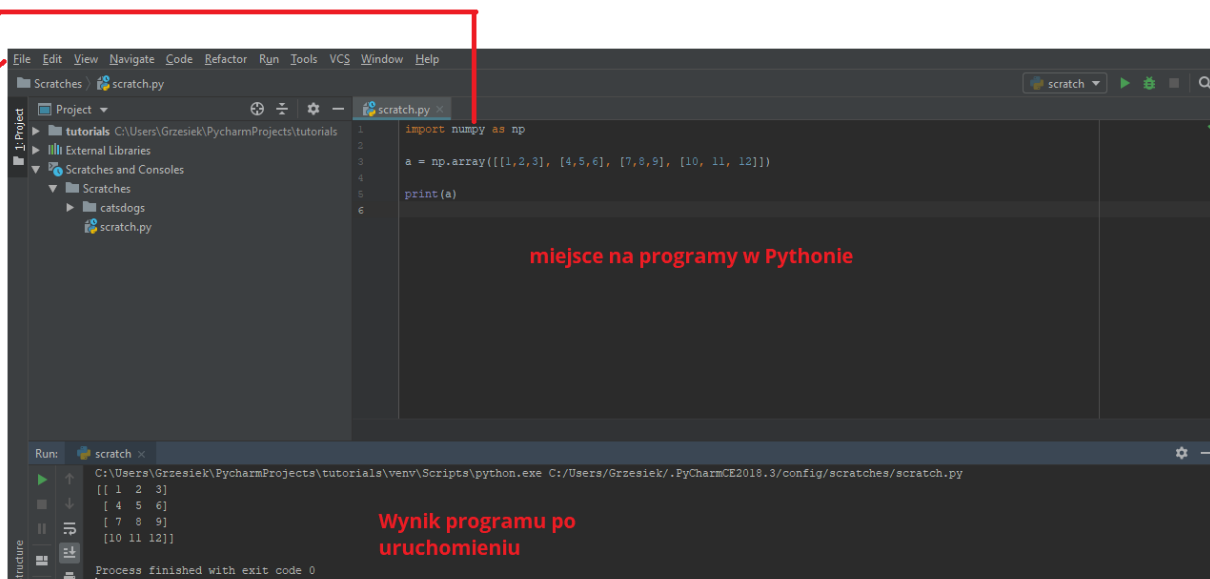
R-Studio

The screenshot shows the RStudio interface with several red annotations pointing to specific components:

- Nowy Plik (Skrypt) do dłuższych komend, funkcji**: Points to the "New File (Script)" button in the top-left toolbar.
- Zaznacz fragment kodu do uruchomienia**: Points to the "Run" button in the top toolbar.
- konsola (prostsze komendy)**: Points to the Console pane at the bottom left, which shows the R version and some initial commands.
- Zmienne i ich wartości w środowisku (warto zapisywać środowisko co jakiś czas). Obok zakładka z całą historią komend (też można zapisać)**: Points to the Environment pane on the right, which displays the current workspace variables (x, string) and their values.

PyCharm

By zaimportować paczkę, wejdź do menu File → Settings → Project Interpreter → Guzik "+" po prawej → Znalezienie paczki i instalacja. Bez tego paczka po komendzie import będzie podkreślona na czerwono.



Niezależnie od wyboru środowiska wykonaj poniższe zadania.

Zadanie 1

W tym zadaniu wykonamy kilka komend matematyczno-statystycznych:

- Zapisz pod zmienną **a** liczbę 123, a pod zmienną **b** liczbę 321. Następnie prostą komendą policz i zwróć wynik mnożenia liczb.
- 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.
- Dla powyższych wektorów podaj iloczyn skalarny i długości euklidesowe.
- Stwórz dwie dowolne macierze 3 x 3 pomnóż je po współrzędnych i wyświetl wynik, a następnie pomnóż je macierzowo i wyświetl wynik.
- Stwórz wektor 50 losowych liczb z zakresu od 1 do 100.
- 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)

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

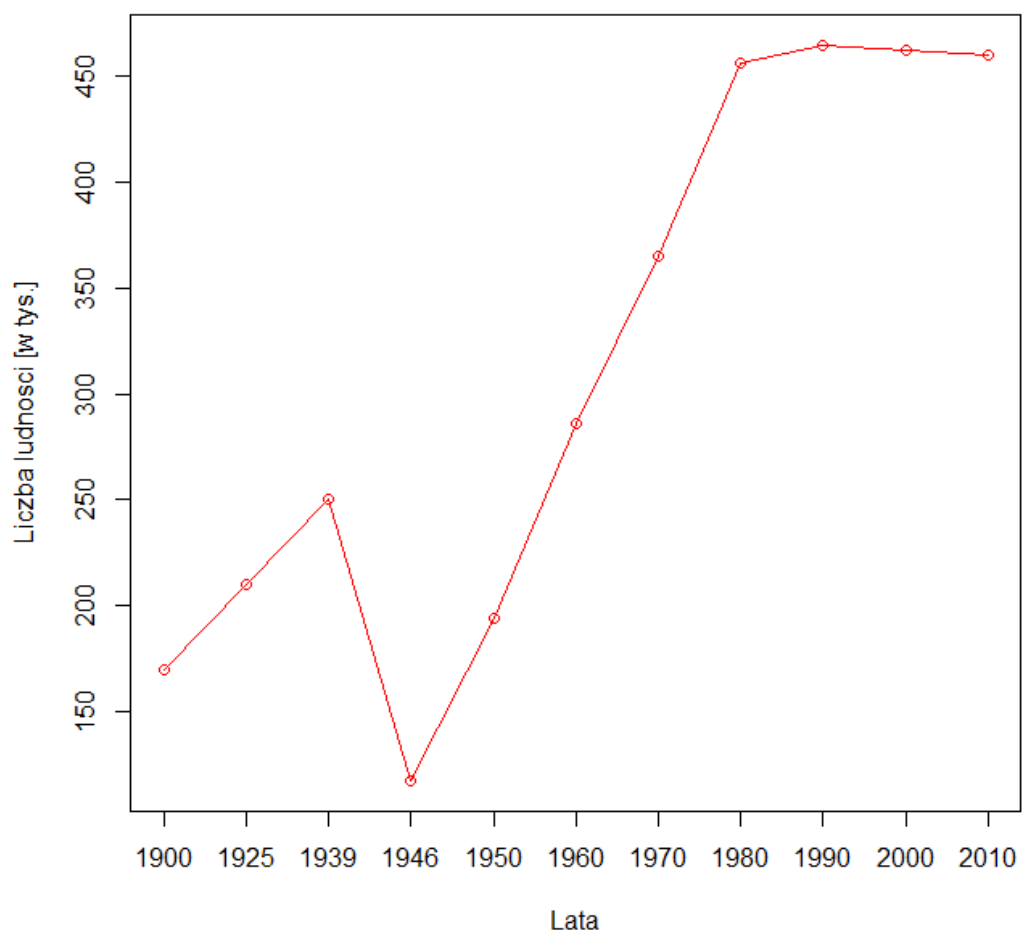
W oryginalnym wektorze jakie było max? Na której pozycji stało? Jaka liczba stoi na tej pozycji w nowym wektorze?

Zadanie 2

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

- Załaduj plik [miasta.csv](#) do programu w Pythonie lub do środowiska w R. Wyświetl tabelę z danymi za pomocą odpowiedniej komendy.
- Dodaj za pomocą odpowiedniej instrukcji wiersz do tabeli z ludnością w 2010 roku: 2010,460,555,405
- 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:

Ludność w miastach Polski



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