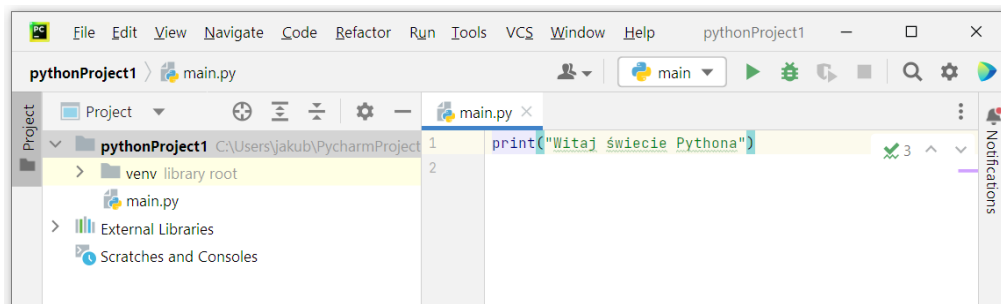


# Wstęp do programowania



## Lista 7

opracował dr inż. Jakub Długosz

### Tematyka:

Wyrażenia regularne.

Operacje na plikach tekstowych.

### Lista na ocenę.

#### Zadanie 1 – ochrona danych osobowych

Napisz program, który wczyta plik `nazwiska_jawne.txt` i z użyciem wyrażeń regularnych zamieni nazwisko (w tym dwuczłonowe, gdzie mogą również pojawiać się białe znaki) występujące po słowie nazwisko (uwzględnij deklinację) na odpowiedni ciąg symboli zawierający "\*\*\*". Wynik powinien zostać zapisany do pliku `nazwiska_ukryte.txt`.

##### We (plik `nazwiska_jawne.txt`):

Jest wielu bohaterów tego nietypowego zdarzenia. Pojawia się **nazwisko Malinowski**, ale nie tylko. Występuje również osoba z **nazwiskiem Kozak-Nowak**.

To dwuczłonowe nazwisko pojawia się w artykułach prasowych wyjątkowo często. Odsyłam do autora jednego z wpisów o **nazwisku Morawski - Dunin**.

##### Wy (plik `nazwiska_ukryte.txt`):

Jest wielu bohaterów tego nietypowego zdarzenia. Pojawia się **nazwisko \*\*\***, ale nie tylko. Występuje również osoba z **nazwiskiem \*\*\*\_\*\*\***.

To dwuczłonowe nazwisko pojawia się w artykułach prasowych wyjątkowo często. Odsyłam do autora jednego z wpisów o nazwisku **\*\*\* - \*\*\***.

##### Wskazówka:

Do zamiany użyj funkcji `re.sub()`.

#### Zadanie 2 – usunięcie znaczników html

Napisz program, który wczyta plik `stronaWWW.html` i z użyciem wyrażeń regularnych usunie wszystkie znaczniki html, pozostawiając tylko tekst zawarty wewnątrz znaczników. Wynik powinien zostać zapisany do pliku `stronaWWW.txt`.

**We (plik stronaWWW.html):**

```
<!DOCTYPE html>
<html lang="pl">
<head>
  <title>Moja strona WWW</title>
  <meta charset="utf-8" />
  <meta http-equiv="X-UA-Compatible" content="IE=edge" />
  <meta name="viewport" content="width=device-width, initial-scale=1, shrink-to-
fit=no" />
  <link rel="stylesheet" type="text/css" href="css/xxxxxxx.css" />
</head>
<body>
  <h1>Moja strona WWW</h1>
  <p>Ciekawy tekst</p>
  <script> </script>
</body>
</html>
```

**Wy (plik stronaWWW.txt):**

```
Moja strona WWW
Moja strona WWW
Ciekawy tekst
```

**Wskazówka:**

Do zamiany użyj funkcji `re.sub()`.