

Sprawozdanie

Analiza i Bazy Danych - Przygotowanie środowiska do pracy w DS

Adrian Gacek

Wstęp

Celem zajęć jest zapoznanie się ze środowiskiem pracy na przedmiocie Analiza i Bazy Danych.

Zadanie 1

Zainstaluj wirtualne środowisko conda na systemie Ubuntu.

1. Stworzenie środowiska wirtualnego, którego nazwa to imięnazwisko.
 2. Aktywacja środowiska wirtualnego.
 3. Instalacja pakietów Python z pliku requirements.txt.
 4. Wylistowanie zainstalowanych pakietów.
 5. Zrobienie zrzutu ekranu i dodanie do kursu na upel.agh.edu.pl
-

Zadanie 2

Utwórz własne repozytorium do przedmiotu na github.com.

1. Utworzenie własnego repozytorium.
 2. Sklonowanie repozytorium.
 3. Dodanie plików do repozytorium.
 4. Wstawienie linku do repozytorium do kursu na upel.agh.edu.pl.
-

Pierwsze zajęcia polegały na przygotowaniu wirtualnego środowiska oraz repozytorium. Pierwsze dwa zadania były czasochłonne, jednak z pomocą prowadzącej nie sprawiły problemów. Po wykonaniu drugiego zadania można było rozpocząć pisanie kodu.

Zadanie 3

Zdefiniuj poniższą funkcję i sporządź jej wykres dla argumentów z danego przedziału:

$$f(x) = x^2 + 5$$

$$x > -1 \text{ oraz } x < 1$$

$$x > -6 \text{ oraz } x < 6$$

$$x > 0 \text{ oraz } x < 5$$

Wspierając się dokumentacją Matplotlib. Dodaj do wykresu etykiety osi, tytuły wykresów i legendy.

```
# import bibliotek
import matplotlib.pyplot as plt
import numpy as np
import pandas as pd

# definicja funkcji i przedziałów
def f(x):
    return x**2+5

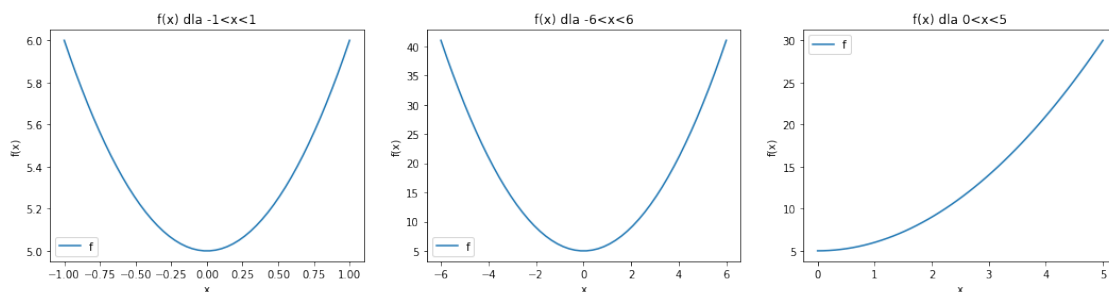
x1 = np.linspace(-1, 1)
x2 = np.linspace(-6, 6)
x3 = np.linspace(0, 5)

# Rysowanie wykresów
plt.figure(figsize=(18, 4))
plt.subplot(1, 3, 1)
plt.plot(x1, f(x1))
plt.title("f(x) dla  $-1 < x < 1$ ")
plt.xlabel("x")
plt.ylabel("f(x)")
plt.legend("f(x)")

plt.subplot(1, 3, 2)
plt.plot(x2, f(x2))
plt.title("f(x) dla  $-6 < x < 6$ ")
plt.xlabel("x")
plt.ylabel("f(x)")
plt.legend("f(x)")

plt.subplot(1, 3, 3)
plt.plot(x3, f(x3))
plt.legend("f(x)")
plt.title("f(x) dla  $0 < x < 5$ ")
plt.xlabel("x")
plt.ylabel("f(x)")

plt.show()
```



Zadanie 4

Utwórz dataframe, w którym kolumny mają nazwy: name, surname, age, sex. Uzupełnij pięcioma dowolnymi rekordami oraz wyświetl informacje o danych pandas info, opis danych pandas describe, wyświetl pierwsze trzy rekordy pandas head.

Plik z rozwiązaniem zadania 3 i zadania 4 opatrzonym komentarzami, dodaj do kursu na upel.agh.edu.pl

```
# Utworzenie dataframe
d = {"name": ["Adrian", "Mateusz", "Krzysztof", "Diana", "Elżbieta"],
      "surname": ["Hamera", "Kotlarz", "Gacek", "Marek", "Wójcik"], "age":
      [18, 21, 21, 108, 4], "sex": ["m", "m", "m", "f", "f"]}
df = pd.DataFrame(data=d)
# Wyświetlenie informacji nt. dataframe
print(df.info())

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 5 entries, 0 to 4
Data columns (total 4 columns):
#   Column      Non-Null Count  Dtype
---  -
0    name        5 non-null       object
1    surname      5 non-null       object
2    age          5 non-null       int64
3    sex          5 non-null       object
dtypes: int64(1), object(3)
memory usage: 288.0+ bytes
None

df.head(3)

   name  surname  age  sex
0  Adrian   Hamera   18   m
1  Mateusz  Kotlarz   21   m
2  Krzysztof  Gacek   21   m

df.describe()
```

```
count    age
mean    5.000000
std    34.400000
min    41.740867
25%    4.000000
50%    21.000000
75%    21.000000
max    108.000000
```

Funkcja `DataFrame.info()` jest metodą opisującą zawartość dataframe pod kątem ilości i typu zawartych danych oraz pamięci. `DataFrame.head()` wypisuje wybraną liczbę początkowych rekordów. `DataFrame.describe()` pozwala na obserwację podstawowych wskaźników statystycznych opisujących wartości liczbowe.

Wnioski

Zajęcia pozwoliły mi przygotować się do wykonywania trudniejszych zadań w przyszłości oraz wstępnie zapoznać się z klasą dataframe.