

Temat: Pandas – Analiza danych

Wykonaj zadanie zgodnie z poniższymi wymaganiami:

1. Zaimportuj bibliotekę pandas i nazwij ją jako pd.
2. Wczytaj dane temperatury z pliku o nazwie: temperature.csv, a następnie wyświetl **wszystkie** wczytane dane.
3. Wypisz na ekran ilość rekordów zawartych w obiekcie DataFrame, w postaci:
`Ilość rekordów w obiekcie DF wynosi: xxxx`
4. Wyświetl 100 pierwszych rekordów obiektu DataFrame.
5. Wyświetl 100 ostatnich rekordów obiektu DataFrame.
6. Wyświetl tylko kolumny zawierające nazwę Państwa i Miasta.
7. Wyświetl tylko dane temperatury z Polski.
8. Wyświetl tylko dane temperatury z Wrocławia od 1900 do 1999 (włącznie).
9. Oblicz średnią styczeniową temperaturę dla Warszawy z pełnego okresu danych.
10. Znajdź najwyższą i najniższą temperaturę dla Polski z pełnego okresu danych.

Dokumentacja.

Funkcje i metody, które mogą być konieczne do wykonania zadania:

1. **len(obiekt)** – zwraca ilość elementu obiektu (stringu, listy, innych zaawansowanych obiektów iterowalnych)
2. **read_csv** – zwraca obiekt DataFrame z danymi wczytanymi z pliku csv. Ścieżka do pliku jest argumentem obowiązkowym. Argument o nazwie **sep** umożliwia podanie innego niż standardowy separator pliku CSV (domyślnym separatorem jest przecinek).
3. **head()** i **tail()** – metody zwracające początkową i końcową część obiektu Series/DataFrame.
4. **mean()**, **min()**, **max()** – metody zwracające odpowiednio średnią, minimalną i maksymalną wartość liczbową, z kolumny (lub kolumn). Jeżeli kolumny posiadają wartości nie będące wartościami liczbowymi, metoda zwróci błąd.
5. **loc[], iloc[]** – atrybuty zwracające wiersz zgodnie z kryteriami. Odwołanie przez atrybuty **loc** i **iloc** zapewniają bezpieczny dostęp do danych niezależnie od zmiany indeksów (przeindeksowania DataFrame w trakcie filtrowania danych).

Filtrowanie danych obiektu DataFrame.

Przykładowe dane:

Id	A	B	C
1	23	PL	3.6
2	32	UK	4.5
3	23	UK	3.2
4	25	PL	4.6
5	12	US	21
6	33	US	3.6

```
# Wyświetlanie obiektu DataFrame o nazwie df, np. zaimportowanego z pliku CSV
print(df)

# Wyświetlanie tylko jednej kolumny np.:
print(df[["A"]])

# lub
print(df["A"])

# Wyświetlanie kilku kolumn:
print(df[["A", "B"]])

# Wyświetlanie tylko kolumny A dla wierszy, które spełniają określone kryteria np. kolumna B ==
# US i kolumna C < 5
print(df["A"][df["B"]=="US"][df["C"]<5])

#lub bezpieczniej:
print(df["A"].loc[df["B"]=="US"].loc[df["C"]<5])

# Obliczanie średniej dla kolumny C, gdy wiersze z kolumny B są równe PL
print(df[df["B"]=="PL"]["C"].mean())

#lub bezpieczniej:
print(df.loc[df["B"]=="PL"]["C"].mean())
```