SPRAWOZDANIE

Zajęcia: Nauka o danych I

Prowadzący: prof. dr hab. Vasyl Martsenyuk

|  |  |
| --- | --- |
| Laboratorium Nr 1  Data 27.09.2025  Temat: "Wprowadzenie do narzędzi i środowiska pracy w analizie danych"  Wariant 1 | Artur Rolak  Informatyka  II stopień, stacjonarne,  1 semestr, gr. 1b |

https://github.com/rlNkoo/Studia-Magisterskie/tree/main/Nauka%20o%20Danych%20L/Laboratorium%201

1. Polecenie:

Zadanie polegało na zapoznaniu się ze środowiskiem Jupyter Notebook oraz bibliotekami języka Python wykorzystywanymi w analizie danych, w szczególności pandas, numpy oraz matplotlib.  
Celem było:

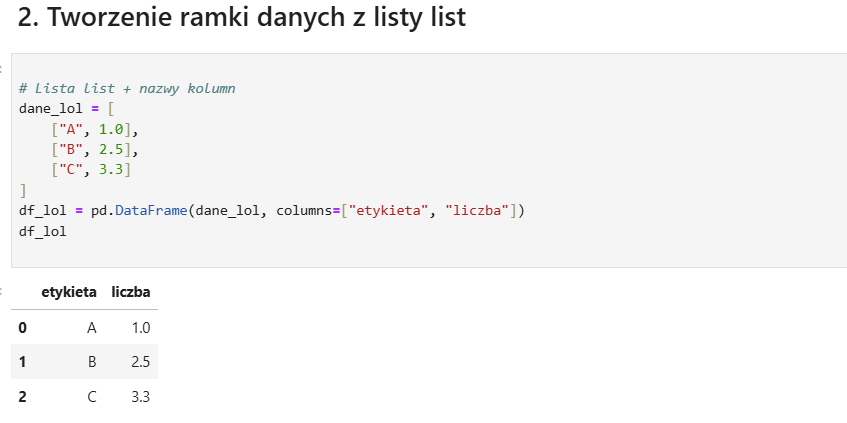
* stworzenie ramek danych (DataFrame) z różnych źródeł: słownik, lista list, plik CSV,
* wykonywanie podstawowych operacji: przegląd, filtrowanie, grupowanie, transpozycja, statystyki,
* tworzenie wykresów i zapisywanie danych do plików,
* przygotowanie kompletnego środowiska pracy do analizy danych.

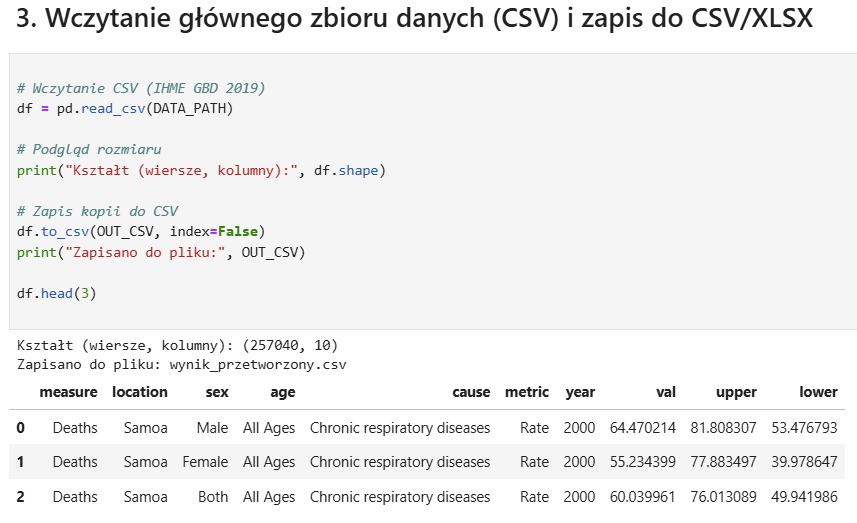
Do realizacji zadania wykorzystano dane z projektu **IHME Global Burden of Disease Study 2019 (GBD 2019)** w formacie .csv.

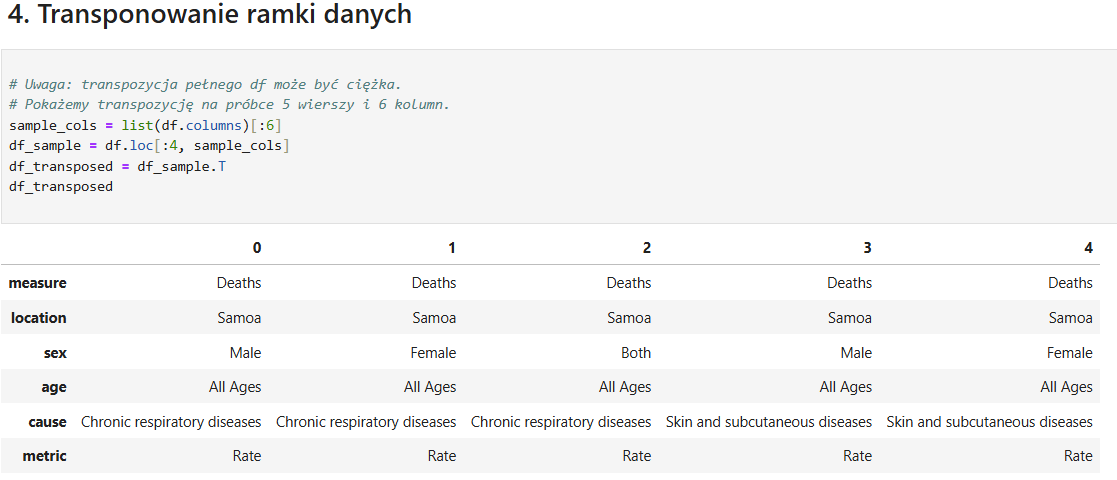
1. Opis programu opracowanego (kody źródłowe, rzuty ekranu)

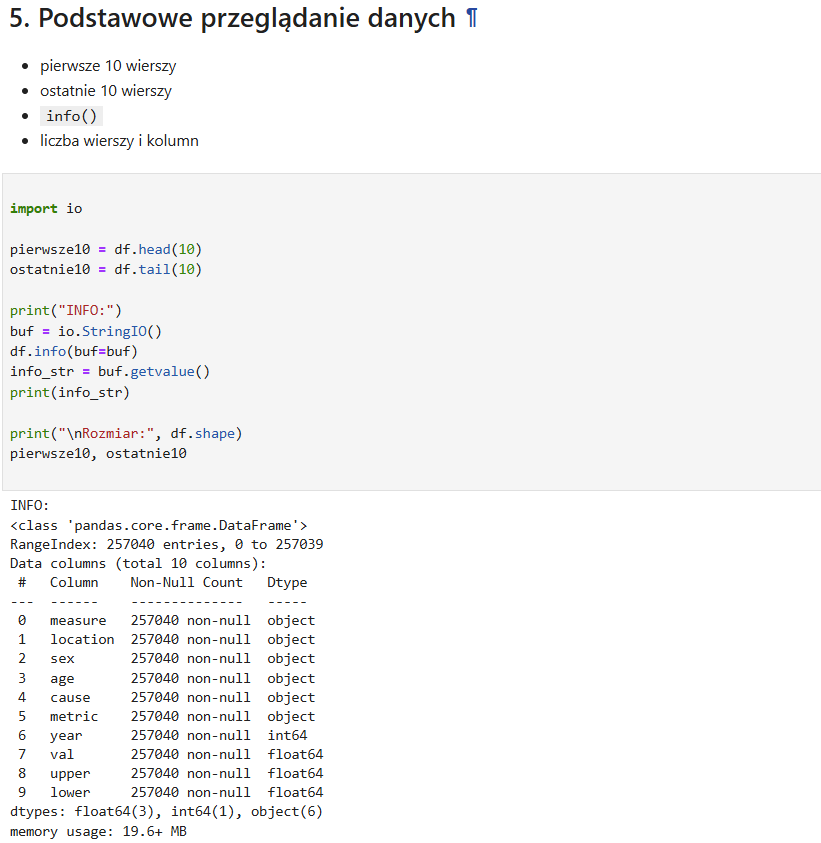


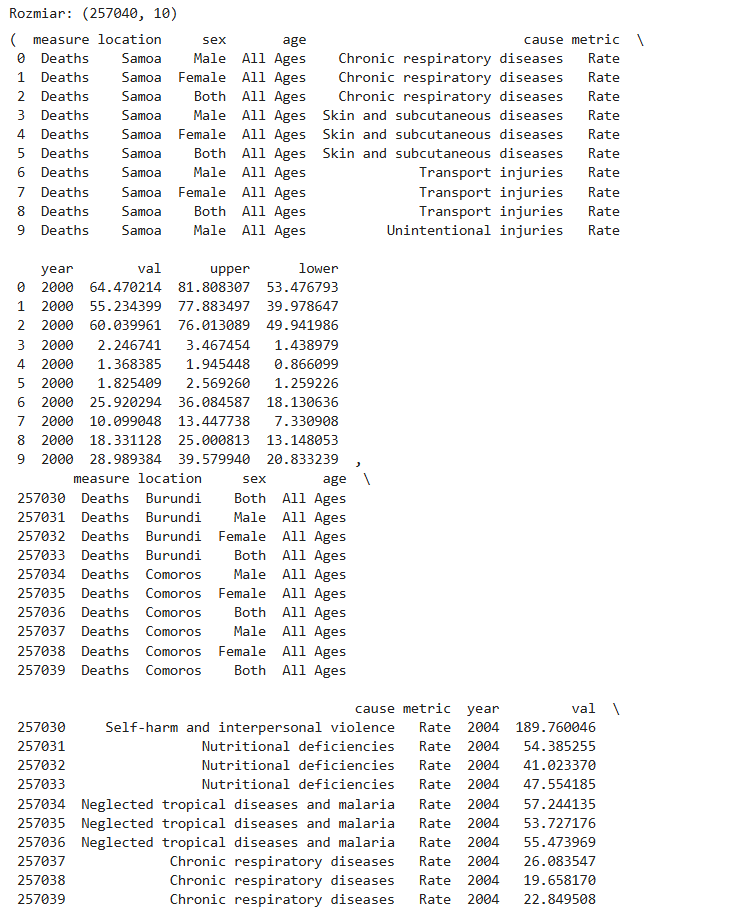


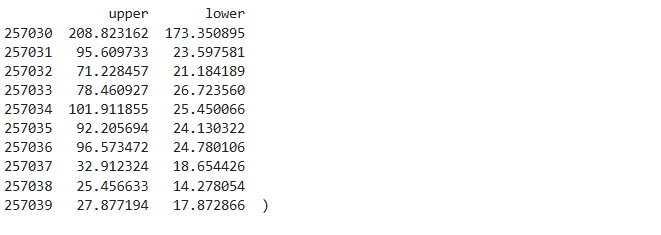


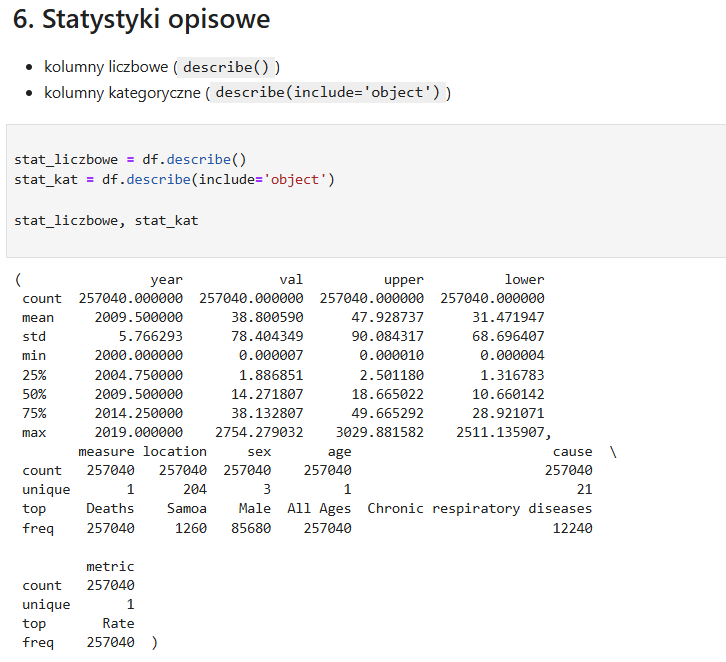


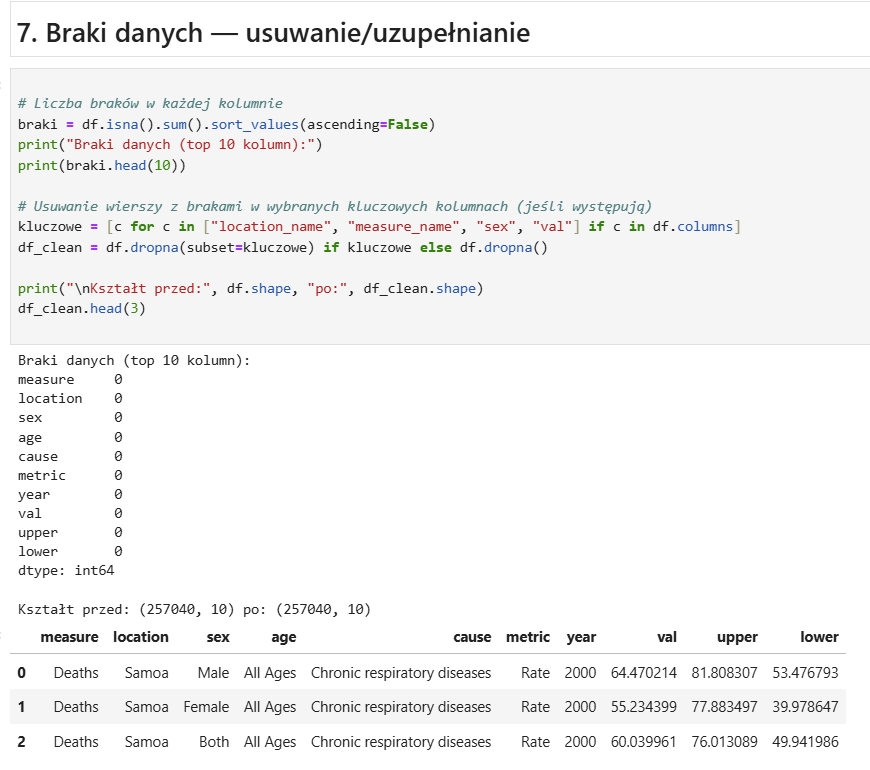


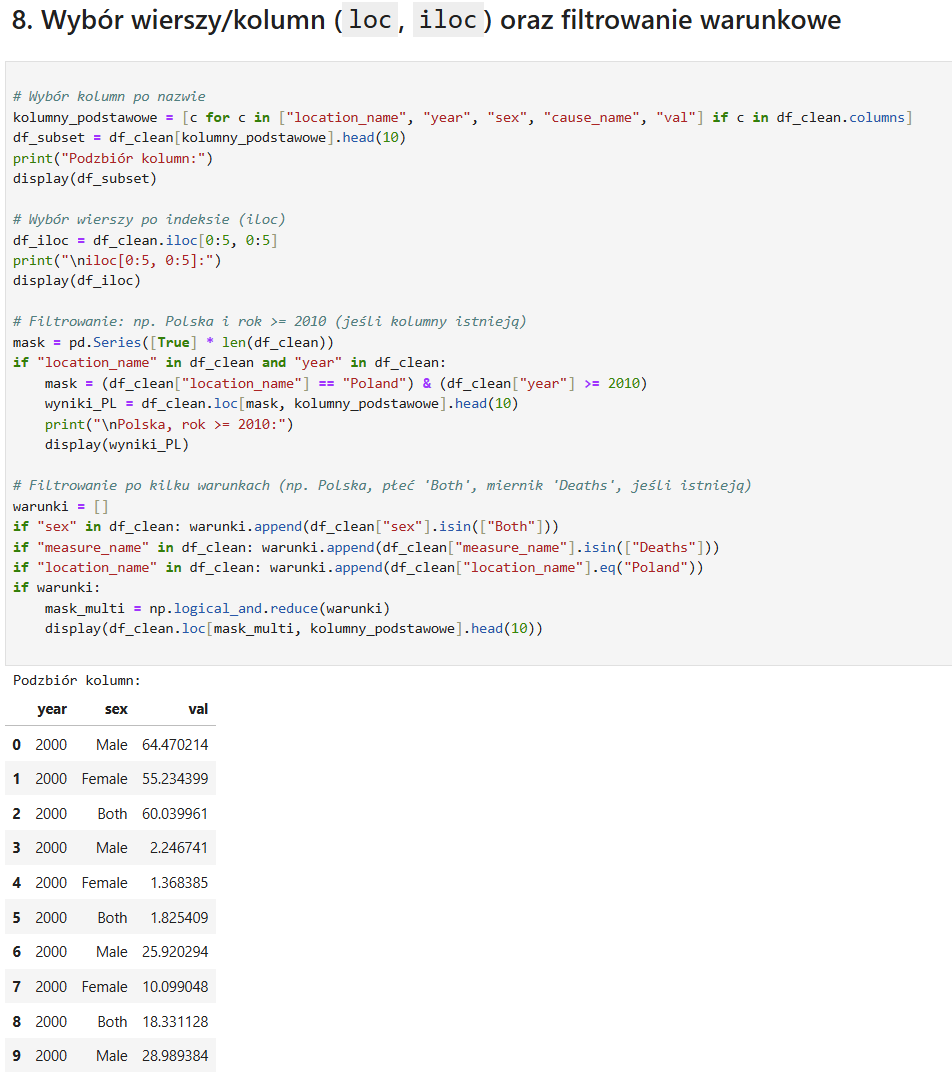


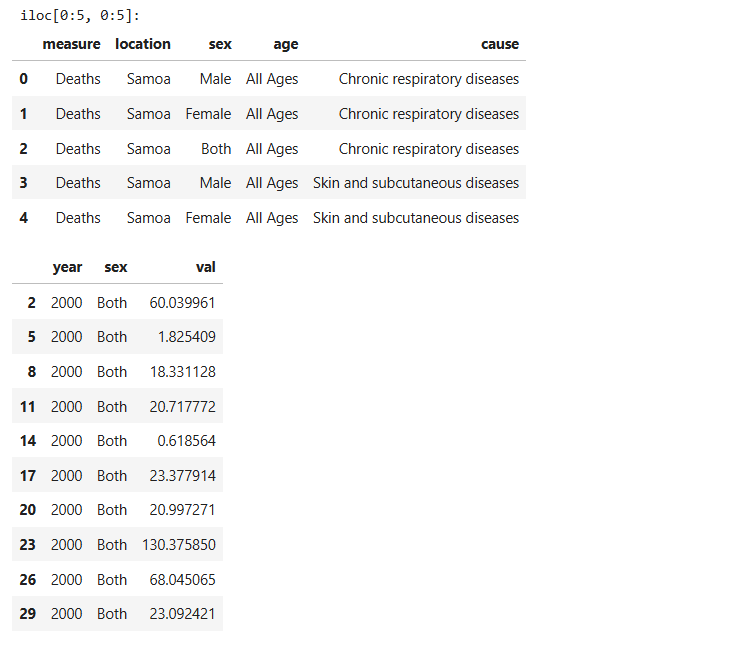








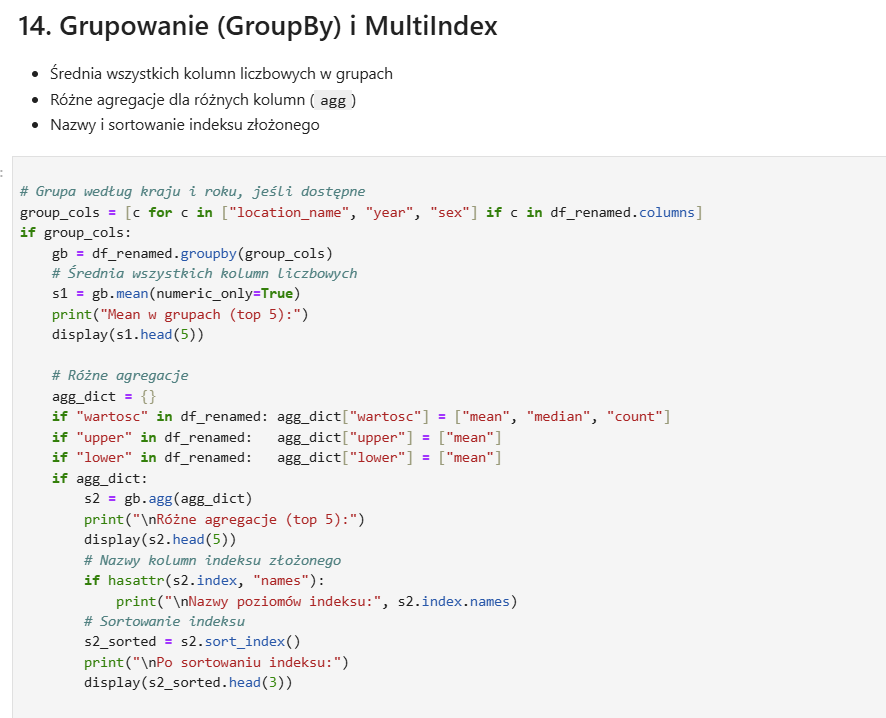


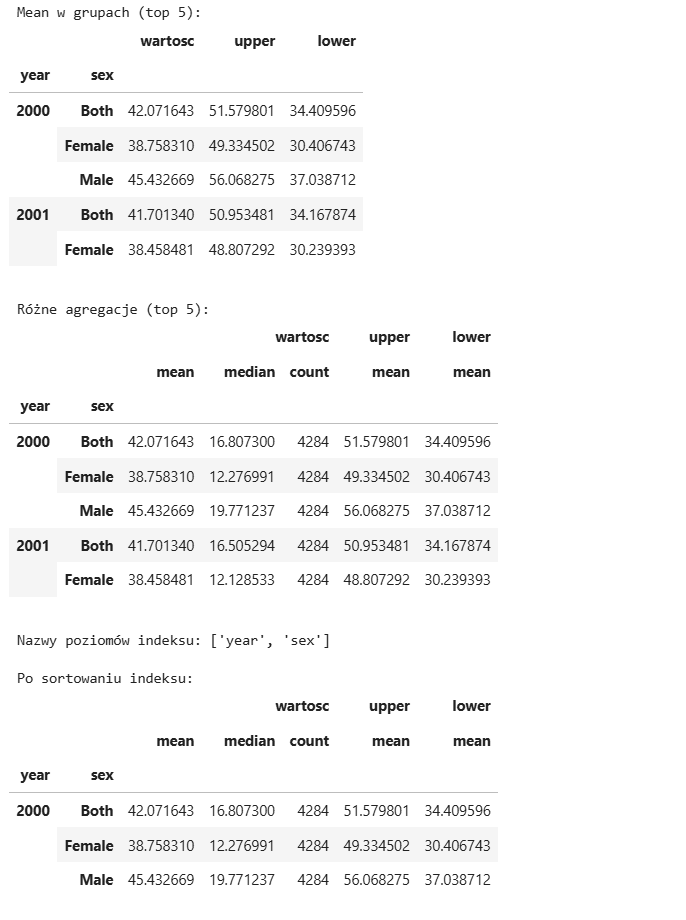




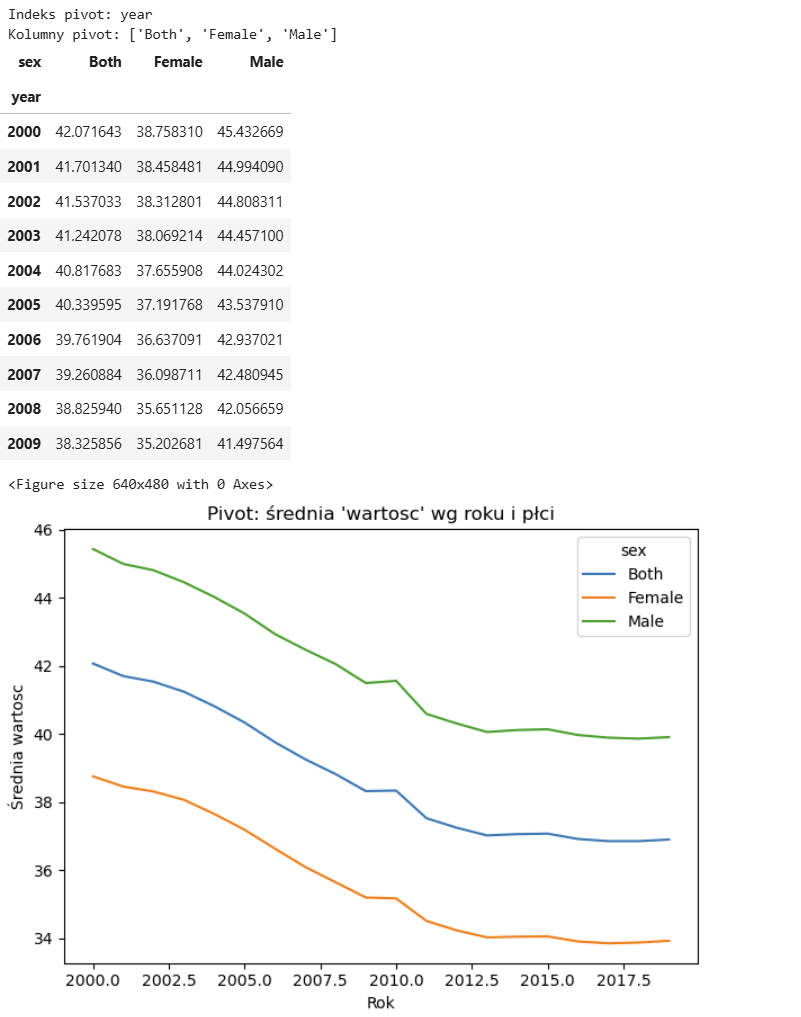


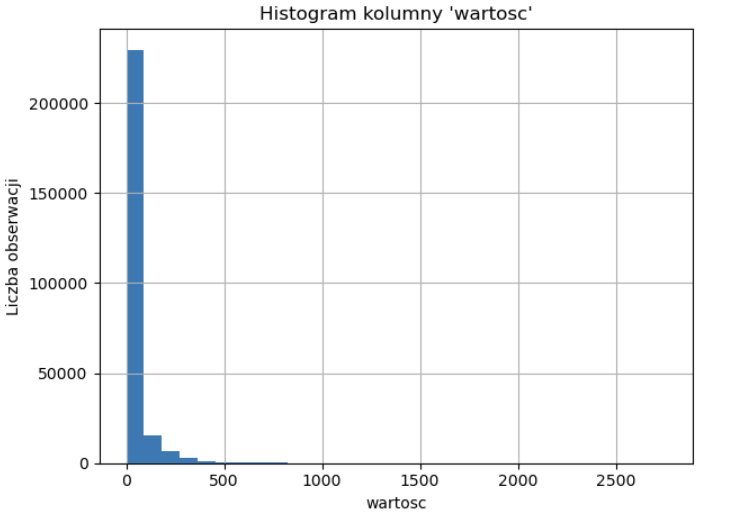
















1. Wnioski

Wykonane laboratorium pozwoliło mi zapoznać się z podstawowymi narzędziami wykorzystywanymi w analizie danych przy użyciu języka Python. Praca w środowisku JupyterLab okazała się bardzo intuicyjna i umożliwiła łączenie kodu, wyników oraz opisów w jednym dokumencie, co znacząco ułatwia proces analizy i dokumentowania pracy. Dzięki bibliotece pandas poznałem praktyczne sposoby wczytywania, przetwarzania i filtrowania danych, a także wykonywania na nich różnych operacji statystycznych. Z kolei wykorzystanie biblioteki matplotlib pozwoliło mi na stworzenie prostych, czytelnych wizualizacji, które ułatwiają interpretację wyników. Praca z rzeczywistym zbiorem danych IHME Global Burden of Disease 2019 umożliwiła mi zrozumienie typowego przepływu pracy analityka danych – od wczytania i przygotowania danych, poprzez analizę, aż po prezentację wyników. Uważam, że realizacja tego zadania wprowadziła mnie w praktyczne aspekty analizy danych w Pythonie i stanowi solidną podstawę do dalszego rozwijania umiejętności z zakresu nauki o danych.