SPRAWOZDANIEZajęcia: Eksploracja i wizualizacja danych  
Prowadzący: prof. dr hab. Vasyl Martsenyuk

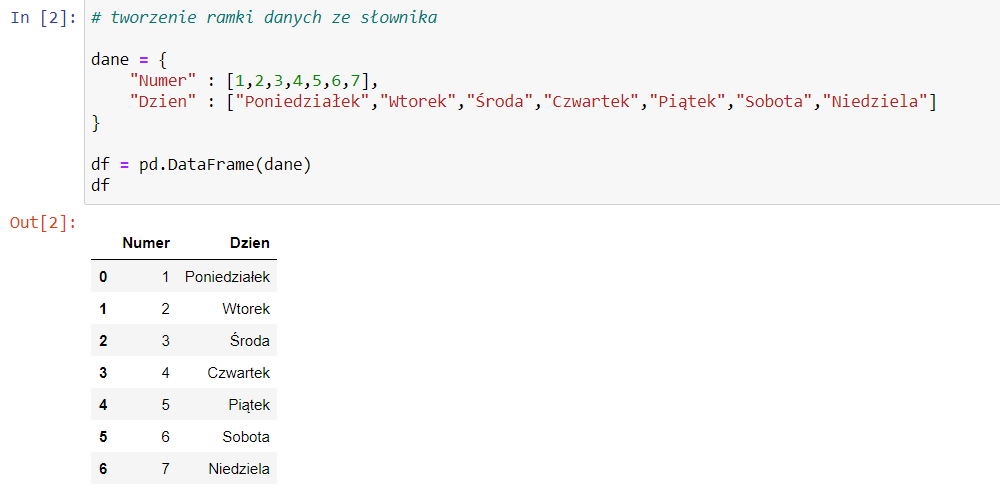
Laboratorium 130.10.2021  
Temat:​ "Wstęp do Python. Biblioteka Pandas"

Przemysław Dziki  
Informatyka II stopień,  
niestacjonarne (zaoczne),  
III semestr,  
https://github.com/pdziki84-student/Ewid.git

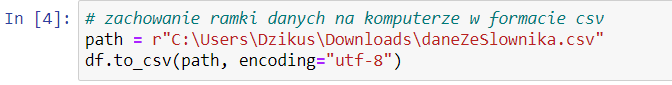
1. Ładowanie biblioteki Pandas



1. Tworzenie ramki danych ze słownika



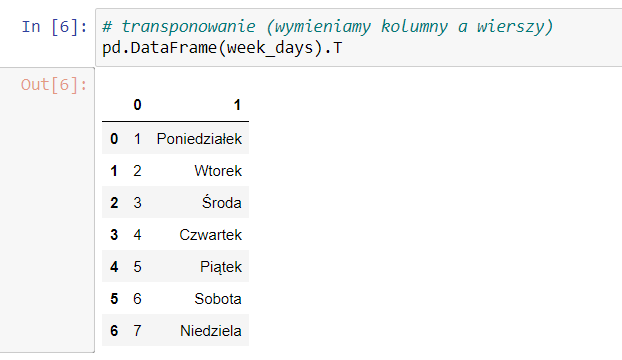
1. Zachowanie ramki danych pobranych z pliku w formacie csv (xlsx)



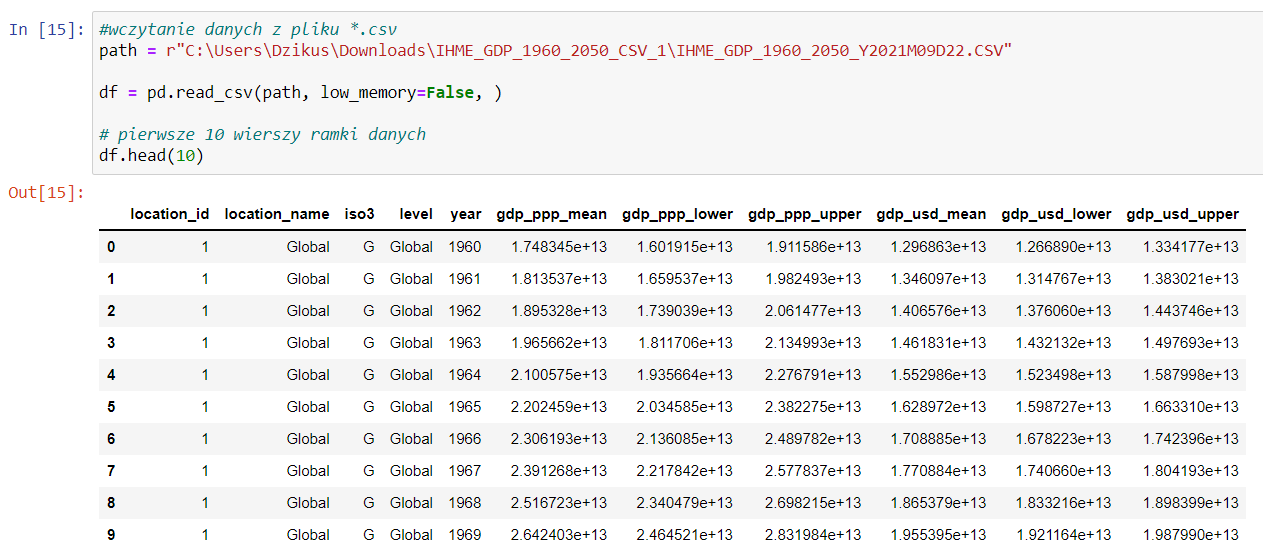
1. Tworzenie ramki danych z listy list



1. Transponowanie (wymieniamy kolumny a wierszy)



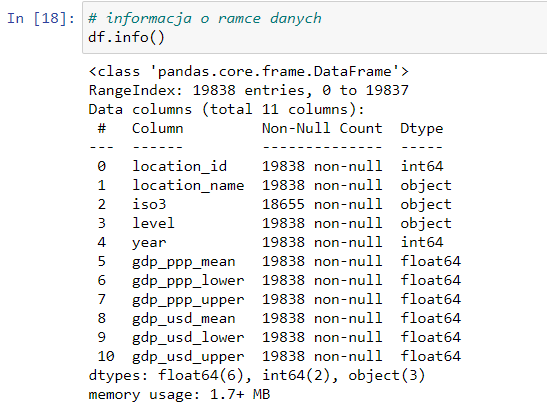
1. Wyświetlić pierwsze 10 wierszy ramki danych



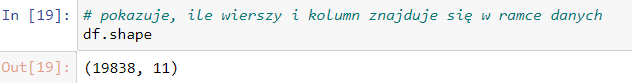
1. Wyświetlić ostatnie 10 wierszy ramki danych



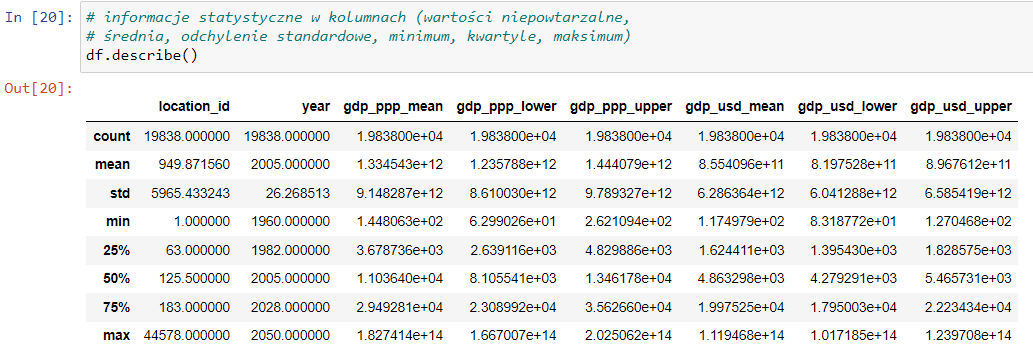
1. Wyświetlić informacje˛ o ramce danych



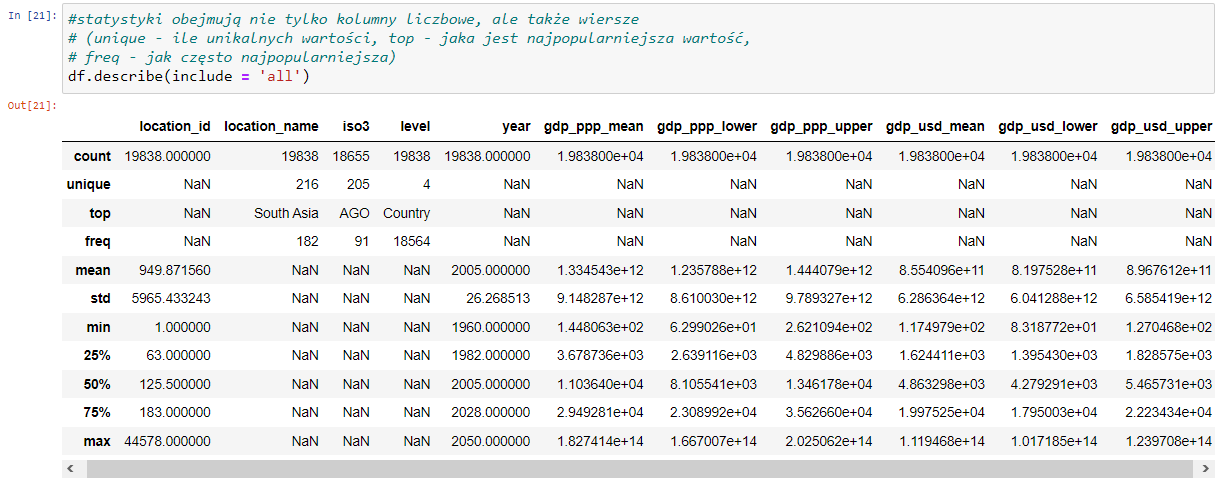
1. Wyświetlić, ile wierszy i kolumn znajduje się w ramce danych



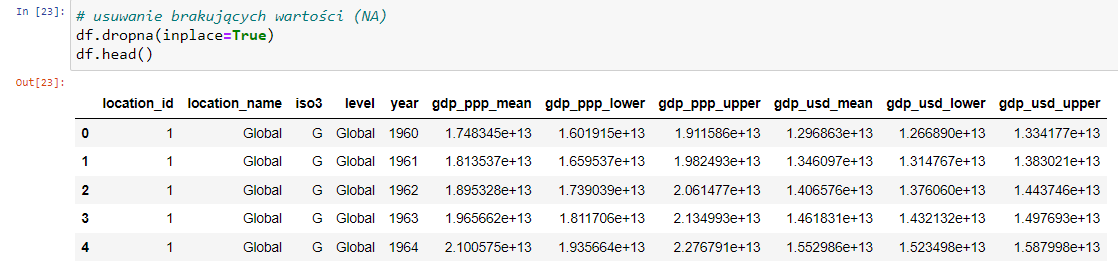
1. Wyświetlić informacje˛ statystyczna˛ o kolumnach liczbowych (wartości  
   niepowtarzalne, średnia, odchylenie standardowe, minimum, kwartyle,  
   maksimum)



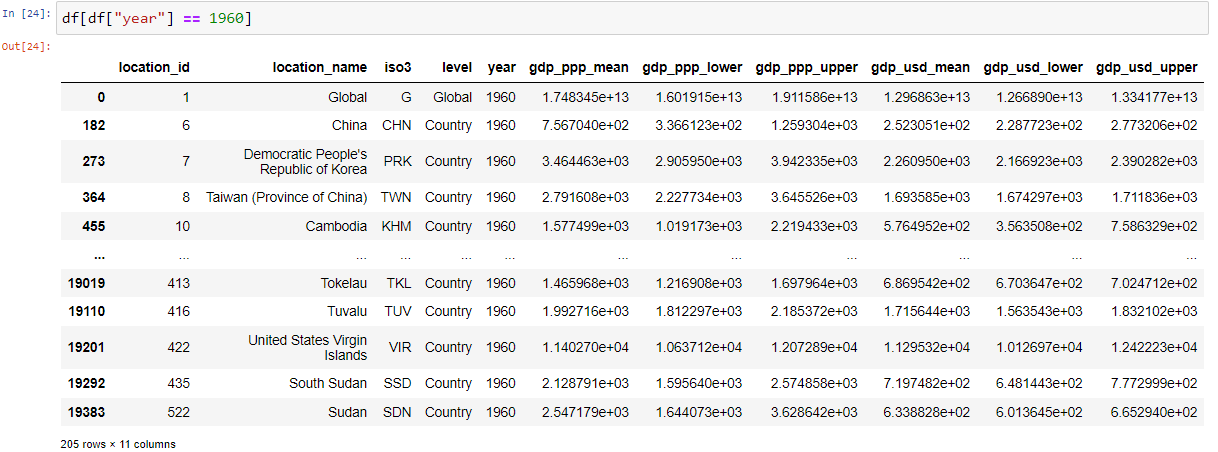
1. Wyświetlić informację statystyczną o kolumnach kategoryzowanych (ile  
   unikalnych wartości, top - jaka jest najpopularniejsza wartość, freq -  
   jak często najpopularniejsza)



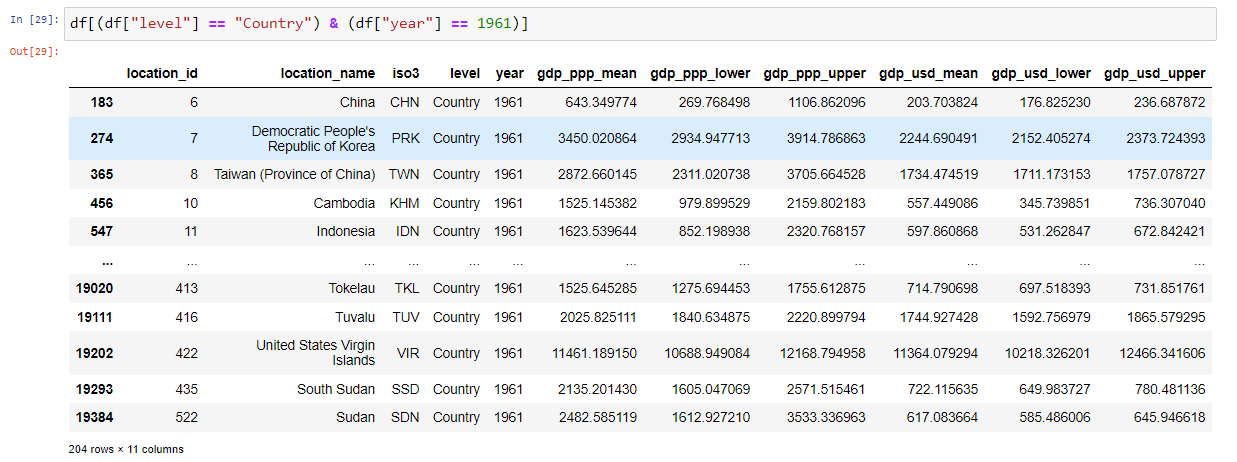
1. Usunąć brakujące wartości w ramce danych



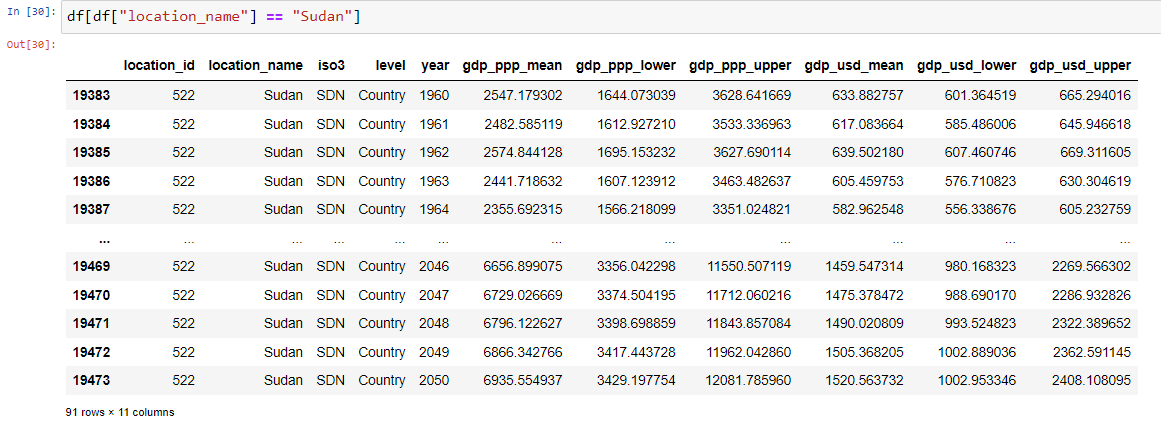
1. Przedstawić wybór wierszy z ramki danych pod warunkiem odnośnie  
   określonej wartości kolumny



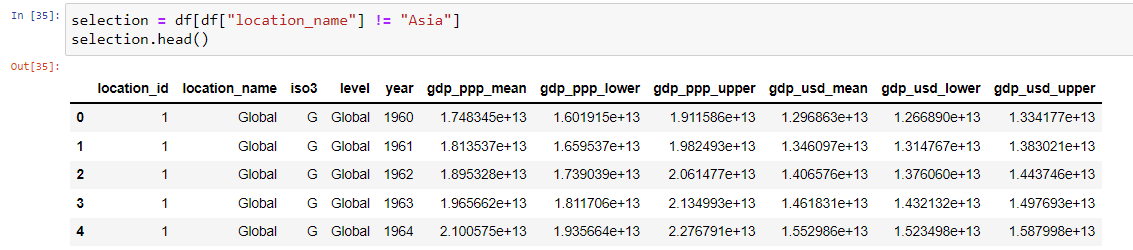
1. Przedstawić wybór wierszy z ramki danych pod warunkiem spełnienia  
   kilku warunków jednocześnie



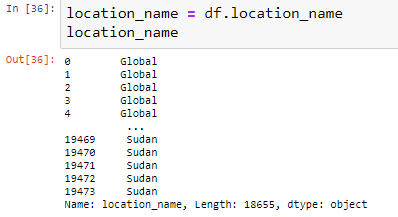
1. Wybrać wiersze które zawierają w kolumnie kategoryzowanej określone  
   słowo



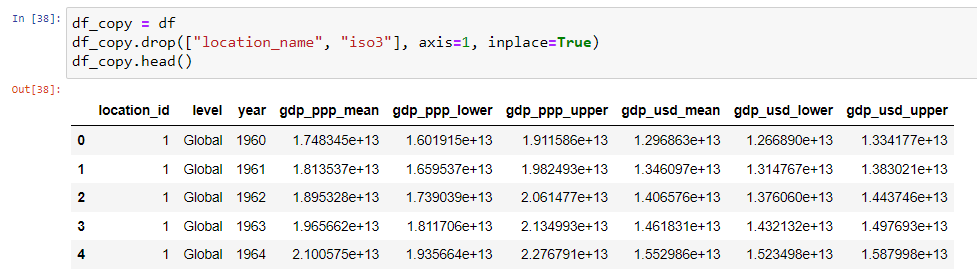
1. Wybrać wiersze które nie zawierają w kolumnie kategoryzowanej określone  
   słowo



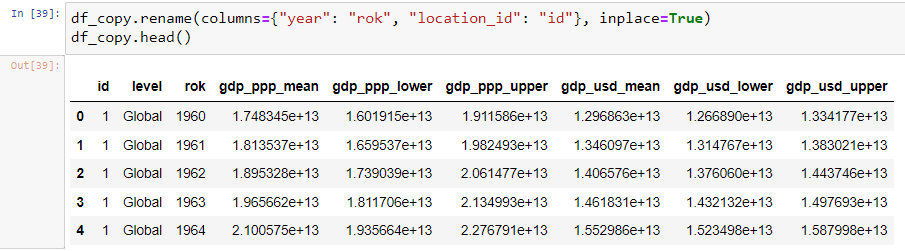
1. Utwórz kolumnę na podstawie istniejącej



1. Usuń kolumnę



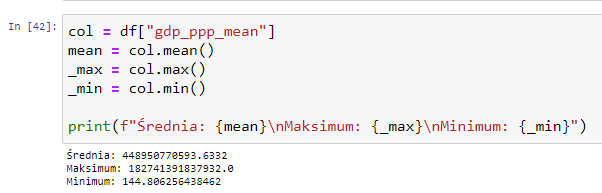
1. Zmień nazwę kolumny



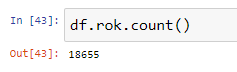
1. Zachowaj ramkę danych jako plik csv na komputerze



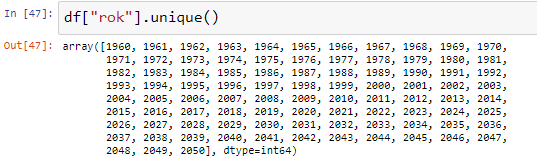
1. Wyświetlić średnia (maksymalną, minimalną) wartość z jednej kolumny



1. Wyświetlić liczbę wierszy



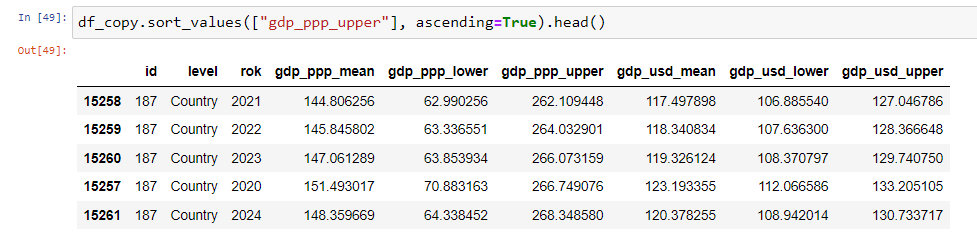
1. Wyświetlić wartości unikatowe w kolumnie



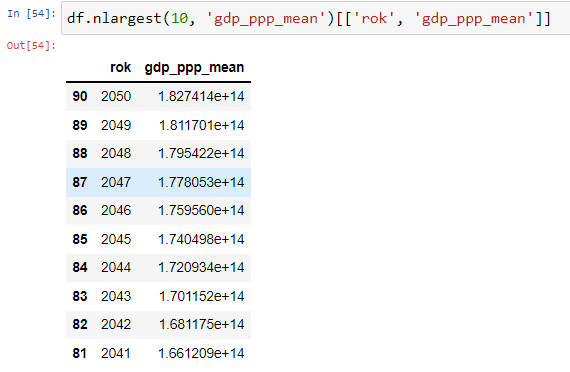
1. Wyświetlić liczby rekordów odpowiadających do wartości



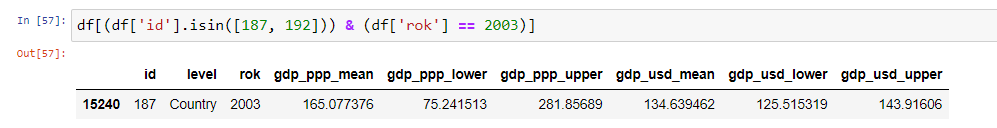
1. Sortowanie wierszy ramki danych według wartości określonej kolumny  
   (malejąco, rosnąco)



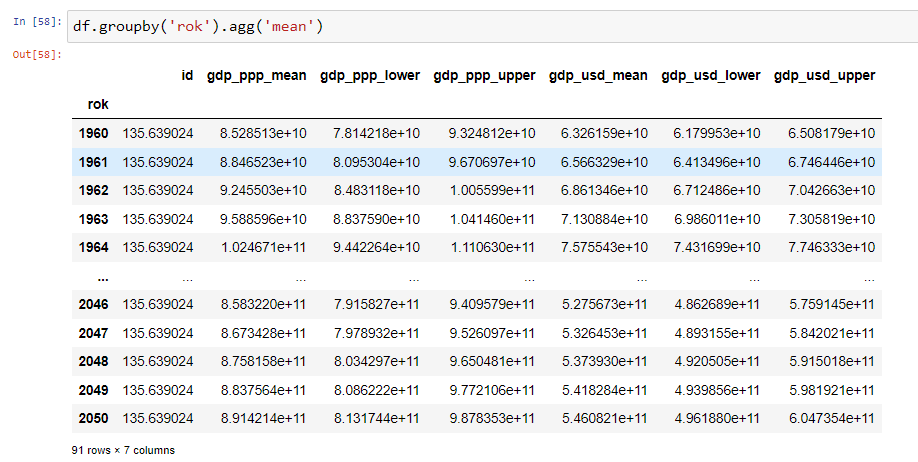
1. Wyświetlić wierszy dla 10 największych (najmniejszych) wartości określonej  
   kolumny



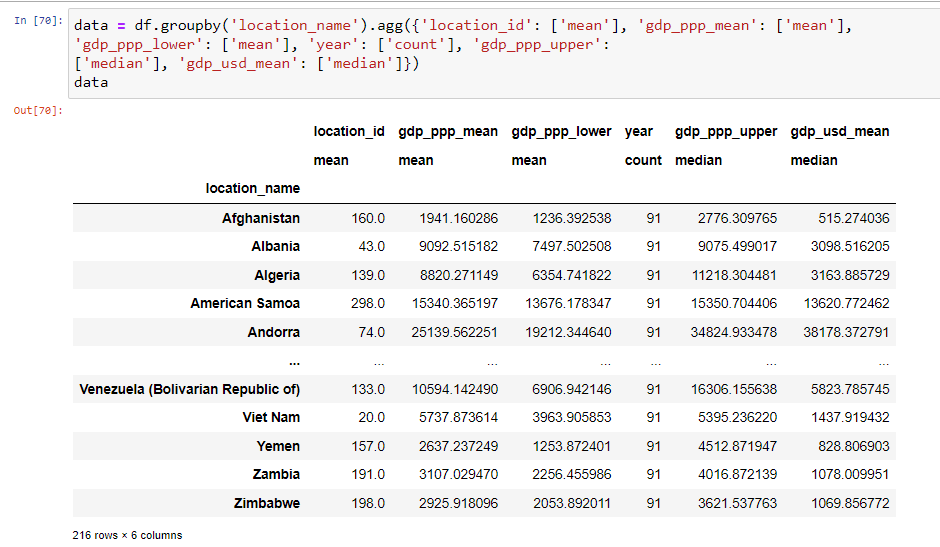
1. Wyświetlić wierszy dla 10 największych wartości określonej kolumny  
   pod warunkiem określonych wartości innej kolumny



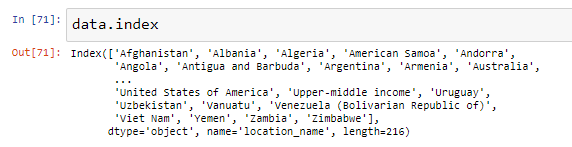
1. Grupowanie wierszy według wartości kolumny kategoryzowanej, potem  
   - uśrednienie wartości wszystkich kolumn w grupie – MultiIndex



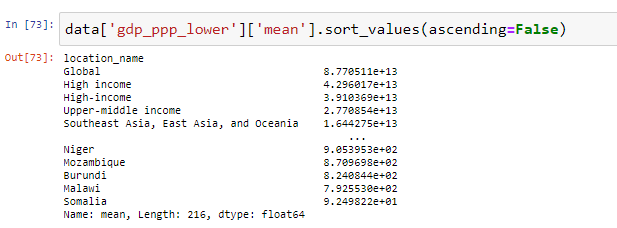
1. Grupowanie wierszy według wartości kolumny kategoryzowanej, potem  
   - uśrednienie wartości dla pewnych kolumn, liczba wartości i mediana  
   dla pozostałych kolumn w grupach



1. Wyświetlić nazwy kolumn indeksu złożonego



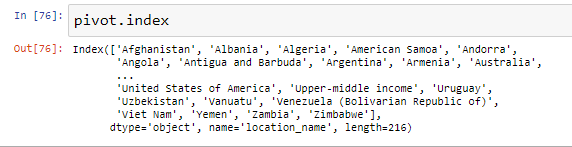
1. Sortować kolumnę indeksu złożonego



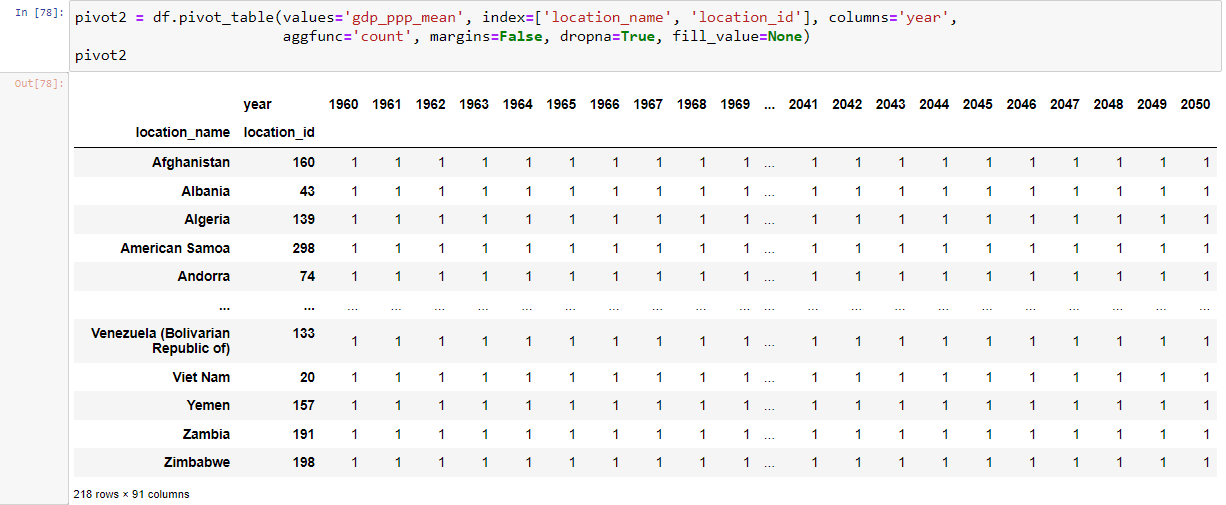
1. Stworzyć tabelę przystawną (pivot table) na podstawie ramki danych



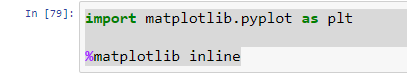
1. Wyświetlić indeksy i kolumny tabeli przystawnej



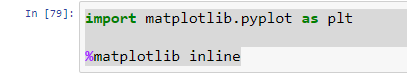
1. Utwórz indeks złożony tabeli przystawnej i wyświetl go



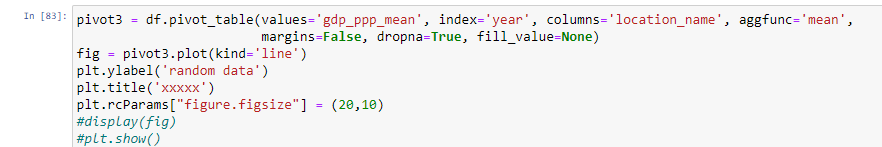
1. Zaimportuj moduł pyplot z biblioteki matplotlib

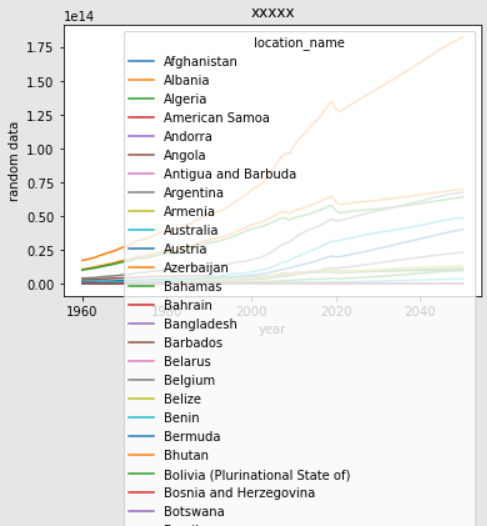


1. Wskazać, że wykresy należy rysować bezpośrednio w zeszycie, a nie w  
   osobnej zakładce

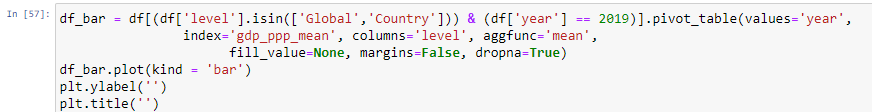


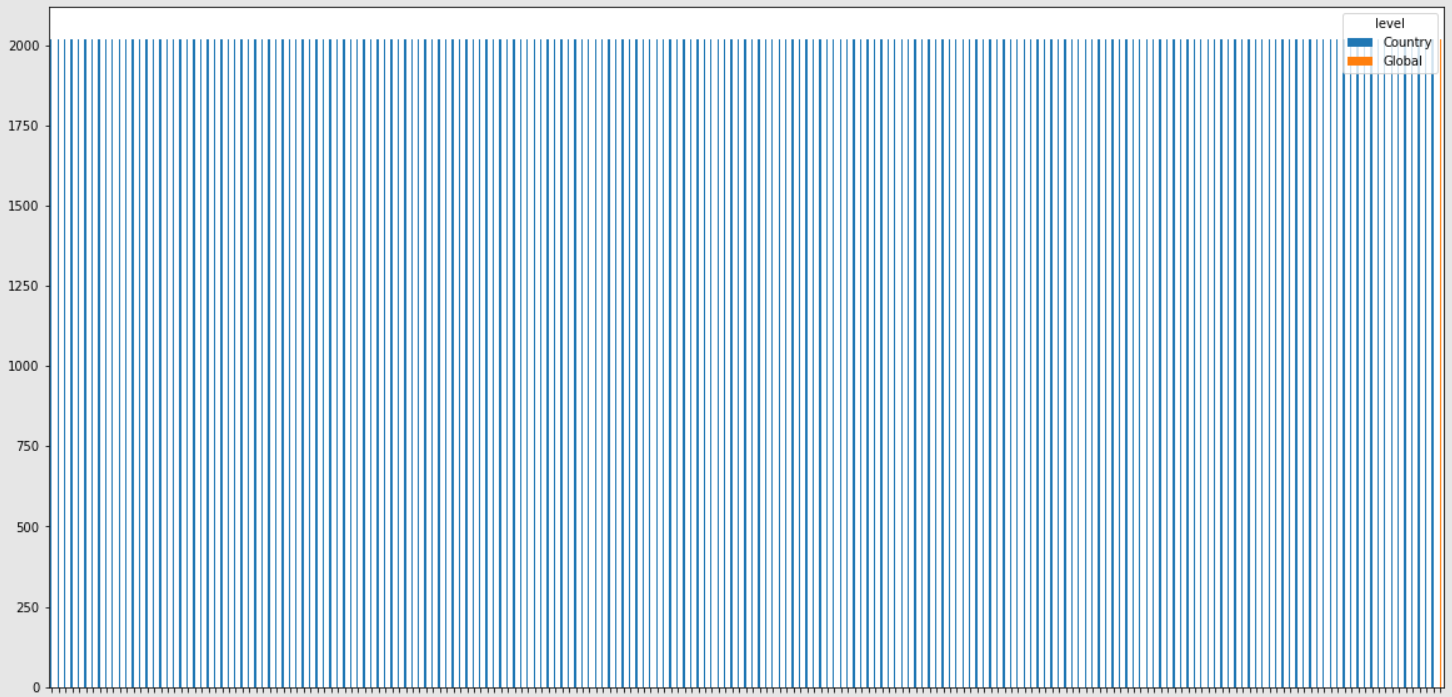
1. Wyświetlić wykres na podstawie tabeli przystawnej



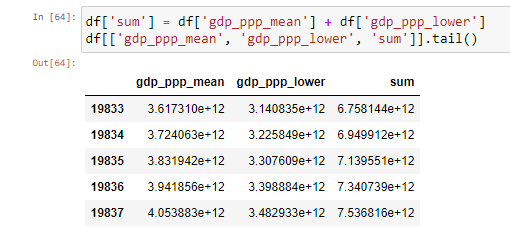


1. Narysować histogram na podstawie wartości kolumny

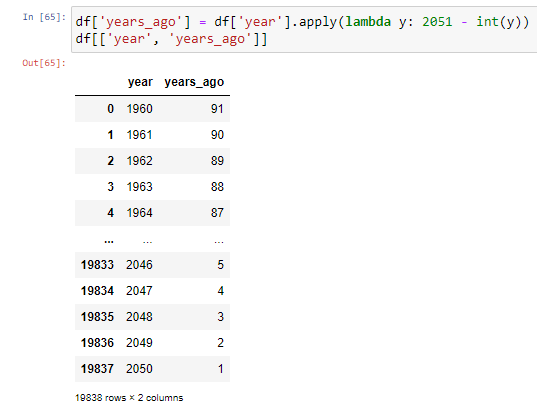




1. Pokazać dodawanie nowych kolumn za pomocą operacji matematycznych



1. Przedstawić na przykładzie dodawanie nowych kolumn z pomocą funkcji  
   lambda



1. Przedstawić możliwości pracy z dużymi plikami przy użyciu argumentu  
   chunksize

