

# **Sprawozdanie na temat nierówności dochodów w Europie.**

Karolina Bakalarz

Amelia Bieda

## **Wstęp**

### **Cel sprawozdania**

Celem niniejszego sprawozdania jest analiza trzech kluczowych zagadnień związanych z poziomem zamożności i nierówności dochodowych w Europie.

Przede wszystkim, badanie ma na celu zrozumienie, w jaki sposób położenie geograficzne krajów w Europie wpływa na ich poziom zamożności, mierzony poprzez PKB.

Zbadamy również wpływ przynależności do Unii Europejskiej na gospodarki krajów, z naciskiem na ewentualne zmiany w poziomie PKB po akcesji do UE.

Ostatecznie sprawdzimy nierówności dochodowe między poszczególnymi państwami europejskimi, poprzez wyniki wskaźnika Gini'ego.

Taka analiza pozwoli nam lepiej zrozumieć sytuację ekonomiczną w Europie, w kontekście zamożności i rozwoju gospodarczego.

### **Źródło danych**

Dane wykorzystane w analizie pochodzą z najnowszej wersji World Income Inequality Database (WIID), opublikowanej 28 listopada 2023 roku. WIID to największe dostępne źródło danych o nierównościach dochodowych na świecie, udostępniane bezpłatnie na otwartej licencji.

Źródła danych obejmują raporty krajowe i regionalne publikowane przez instytucje statystyczne i rządowe, międzynarodowe bazy danych, takie jak Światowy Bank, OECD i Eurostat, oraz ankiety i badania gospodarstw domowych dotyczące dochodów, konsumpcji i nierówności.

Wszystkie dane są standaryzowane, aby umożliwić porównania między krajami i okresami, co pozwala WIID dostarczać spójne serie czasowe, wskaźniki nierówności oraz pełne rozkłady dochodów na poziomie krajowym i globalnym.

## Opis danych

- Kraj - określa, do którego europejskiego państwa przynależą dane w badanym wierszu.
- Skróty - trzycyfrowy kod kraju w formacie ISO 3166-1.
- G - wskaźnik Giniego, czyli miara koncentracji rozkładu zmiennej losowej, w tym przypadku nierówności dochodowej, określona na przedziale 0-100 (w procentach). Wartość zerowa współczynnika wskazuje na pełną równomierność rozkładu.
- Zasoby - szczegółowy podział zasobów
  - IDK CZY TO UWZGLĘDNIMY, to narazie bardziej dla nas, w oryginalnych danych mamy podział: “Consumption”, “Income, net”, “Income, net/gross”, “Earnings, gross”, “Income, gross”, “Market income”, “Taxable income, gross”, “Taxable income, net”, “Earnings, net”, “Monetary income, net”, “Monetary income, gross”, “Factor income”, “Monetary income, net (excluding property income)”, “Earnings”, “Taxable income, excluding property income”, “Taxable income, gross (including deductions)”.
- Skala - szczegółowa skala normalizacji(?). ?przeskalowane dochody?
  - “Per capita”, “OECD-modified”, “Equivalized”, “No adjustment”, “Square root”, “OECD”, NA, “National scale”, “1977 McClements scale”.
- Region - grupowanie subregionalne, czyli podział na Europę północną, południową, zachodnią i wschodnią (oparte na geoschemacie Organizacji Narodów Zjednoczonych).
- G - wskaźnik Giniego, czyli miara koncentracji rozkładu zmiennej losowej, w tym przypadku nierówności dochodowej.
- UE - status należenia do Unii Europejskiej dla odpowiedniego roku.
- PKB - wartość produktu krajowego brutto (w dolarach amerykańskich?).

## Wczytanie i przygotowanie danych

- nadanie odpowiednich nazw zmiennych,
- zadbanie o odpowiednie typy kolumn,
- czyszczenie wartości, jeśli zachodzi taka konieczność,
- dyskusję lub obsługę braków danych;

Z bazy danych WIID będziemy w niniejszym raporcie skupiać się na analizie danych z Europy. Z tego powodu, pozostawiłyśmy wartości dotyczące jedynie analizowanego przez nas kontynentu. Zweryfikowałyśmy typy danych dla każdej kolumny, następnie, dla lepszej czytelności, nazwy kolumn zostały zmienione. Wiersze z brakami danych ("NA") usunęłyśmy, aby uniknąć nieciągłości. W krajach rozwiniętych dane o dochodach są powszechniej stosowane w badaniach, niż dane o wydatkach konsumpcyjnych, stąd zdecydowałyśmy się usunąć zmienną "Consumption" w kategorii "Zasoby".

Po przygotowaniu otrzymałyśmy zbiór około 13 tysięcy wartości w 9 kolumnach. Prezentuje się on następująco:

```
# A tibble: 6 x 9
  Kraj      Skróć Rok      G Zasoby      Skala      Region      UE      PKB
  <fct>    <fct> <dbl> <dbl> <fct>      <fct>    <fct>    <fct> <dbl>
1 Albania ALB    2017  38.6 Income, net/gross Per capita Southern E~ Non~~ 12771
2 Albania ALB    2018  37.1 Income, net/gross Per capita Southern E~ Non~~ 13317
3 Albania ALB    2019  36.0 Income, net/gross Per capita Southern E~ Non~~ 13653
4 Austria AUT    1987  23.0 Income, net      Per capita Western Eu~ EU    33985
5 Austria AUT    1994  30.8 Income, net/gross Per capita Western Eu~ EU    39435
6 Austria AUT    1995  30.8 Income, net      Per capita Western Eu~ EU    40425
```

Warning: pakiet 'e1071' został zbudowany w wersji R 4.4.2

Dołączanie pakietu: 'scales'

Następujący obiekt został zakryty z 'package:purrr':

```
discard
```

Następujący obiekt został zakryty z 'package:readr':

```
col_factor
```

```
# A tibble: 11 x 4
  Statystyka      PKB      Rok      G
  <chr>          <chr>   <chr> <chr>
1 Średnia        39 309   2 007 32
2 Mediana        38 120   2 009 32
3 Odchylenie_standardowe 19 421   11    5
4 Min            4 901    1 958 16
5 Max           120 648   2 022 54
6 Kwartyla_1     26 583   2 000 29
7 Kwartyla_3     48 557   2 015 35
8 IQR            21 974   15    6
9 Skala          115 747   64   38
10 Skośność       1      -1    0
11 Kurtosis       4      1    0
```

## Analiza danych

### Akapit

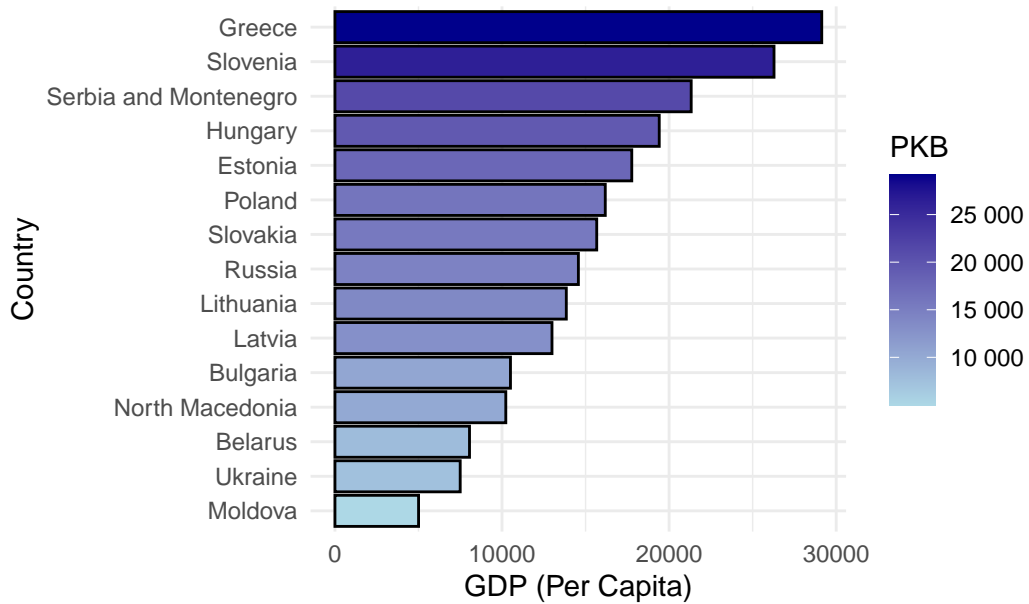
1. Czy jest zależność między położeniem geograficznym w Europie a zamożnością?

Może te akapity będą odpowiedziami na poszczególne pytania badawcze?

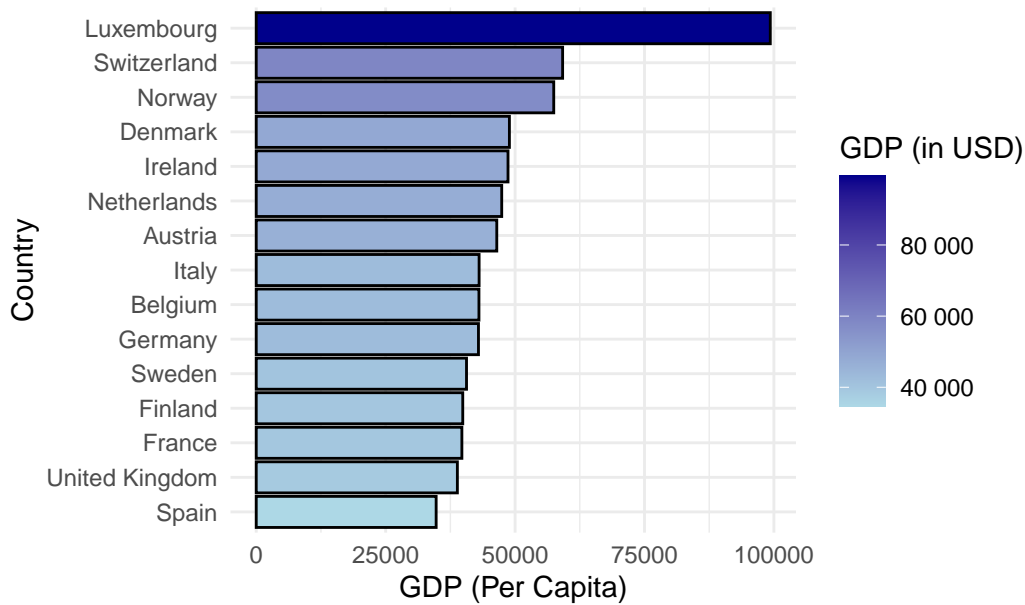
- Dla cech kategoriycznych: Analiza rozkładu, testy chi-kwadrat, kontyngencje.
- Dla cech ciągłych: Analiza korelacji, testy t-studenta, regresja liniowa.

Interpretacja wyników - Opisać wyniki analizy, odnosząc się do postawionych wcześniej pytań badawczych. Logicznie przedstawić wyniki, które prowadzą do odpowiedzi na pytania badawcze.

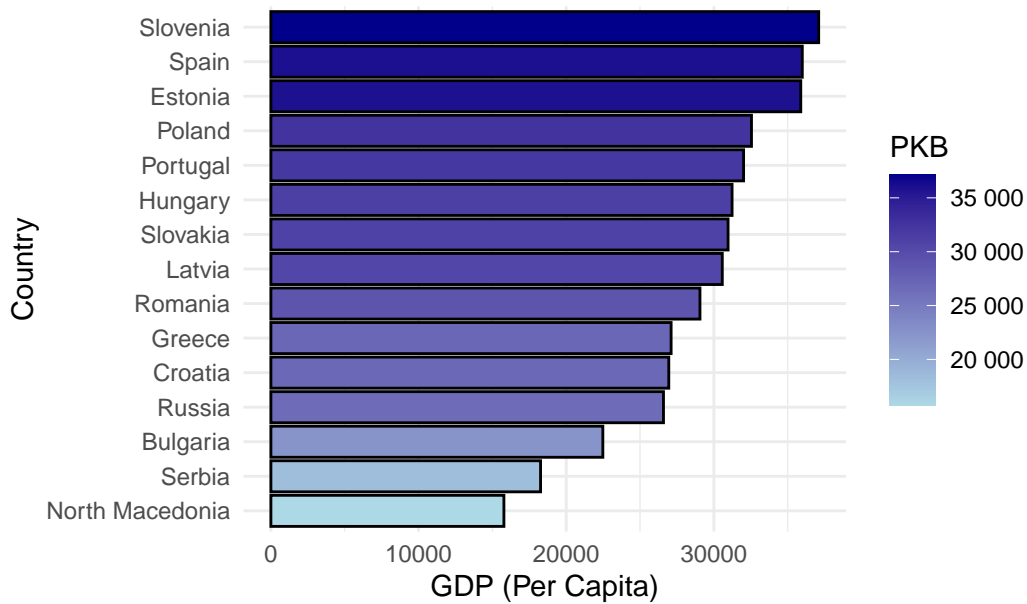
The 15 Countries with Lowest GDP in Europe in



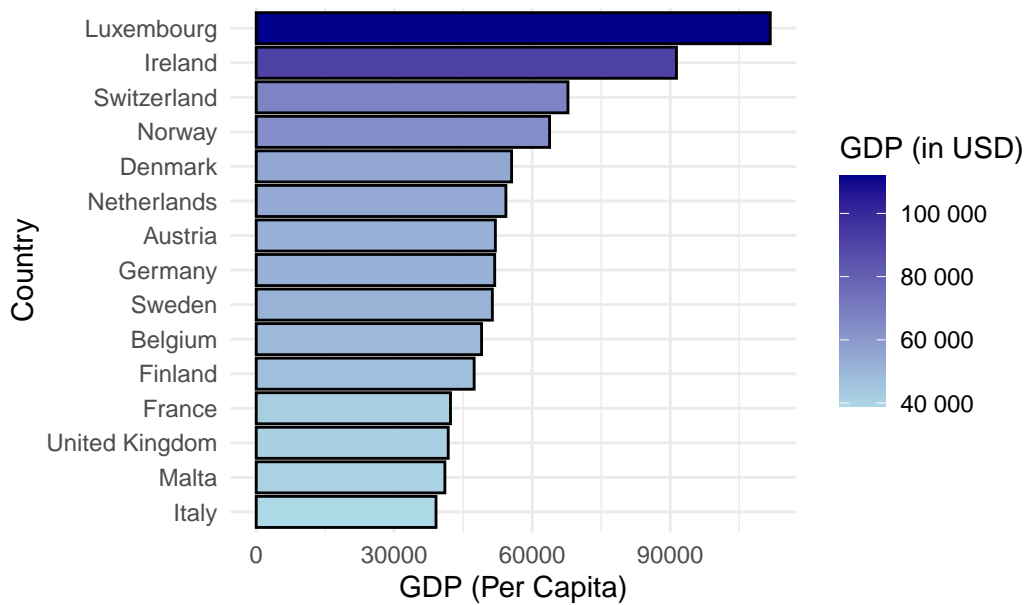
The 15 Countries with Highest GDP in Europe in 2000

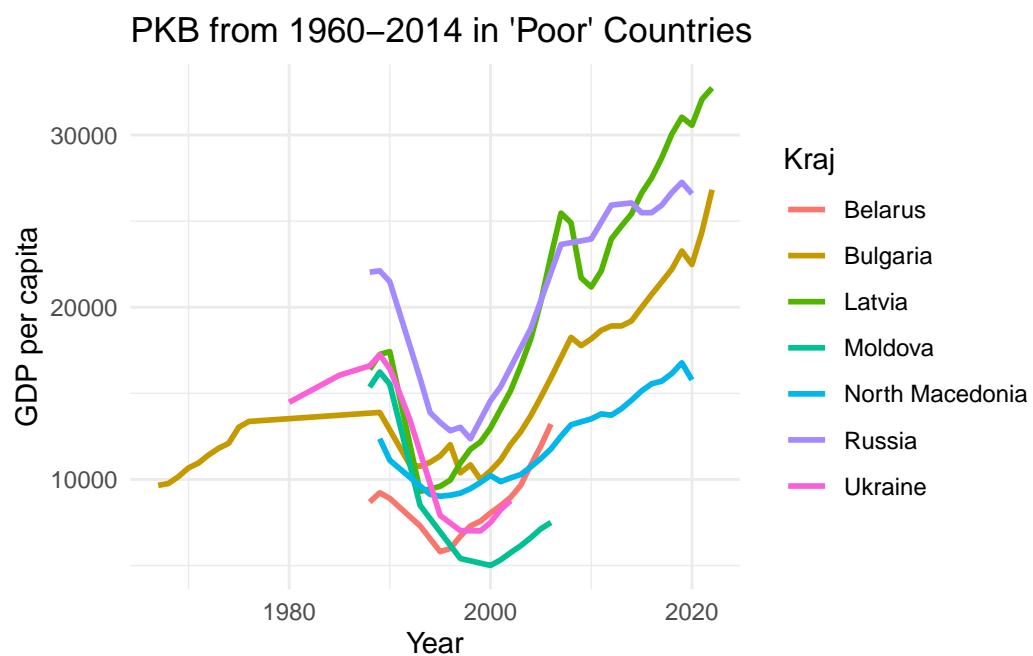
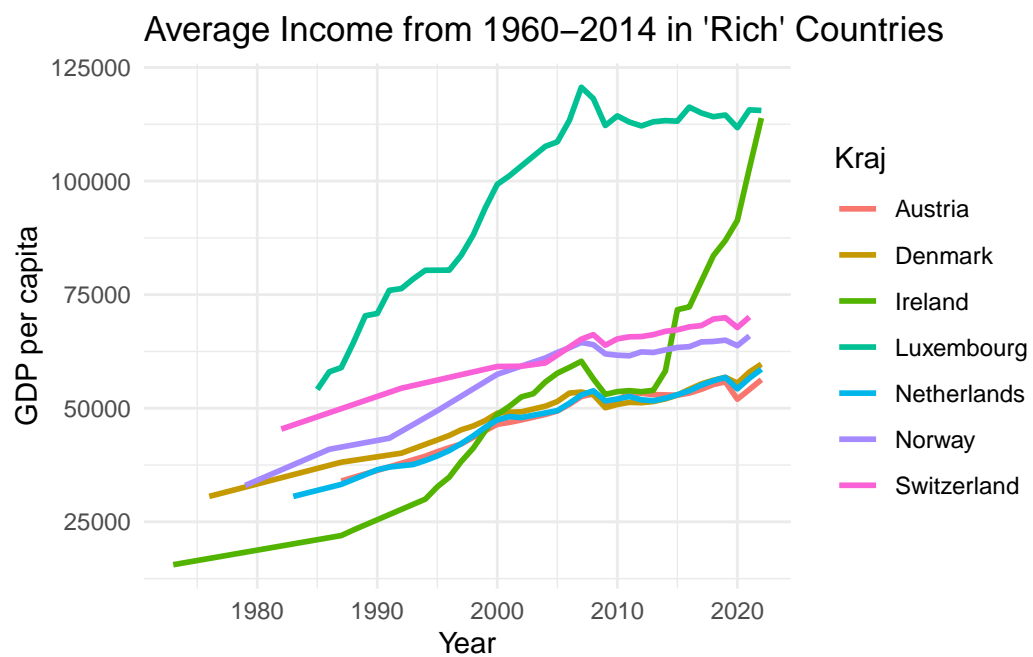


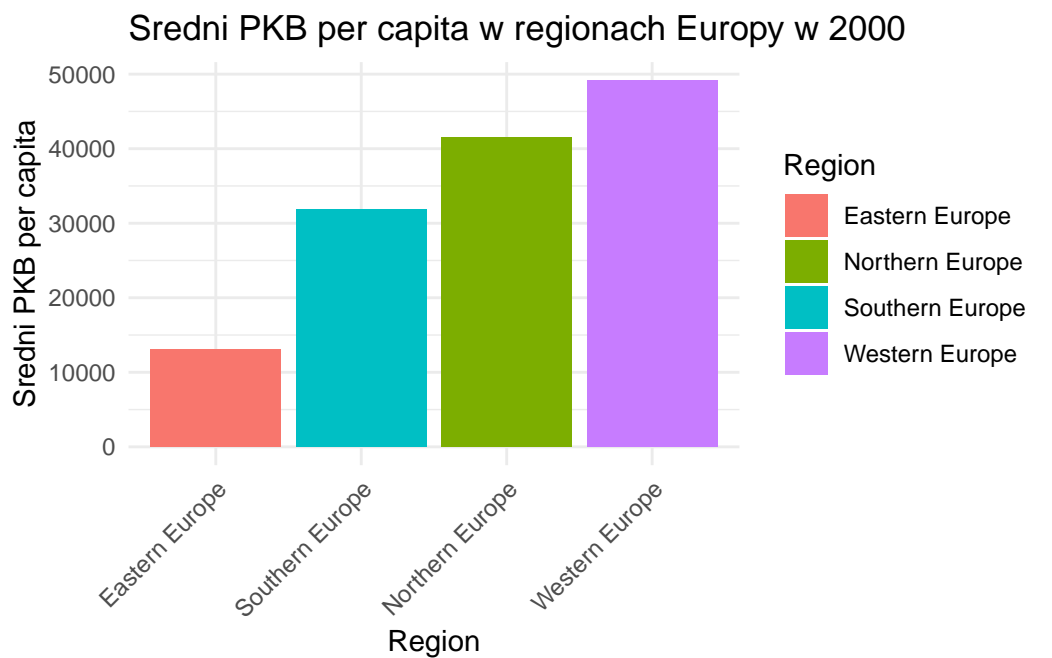
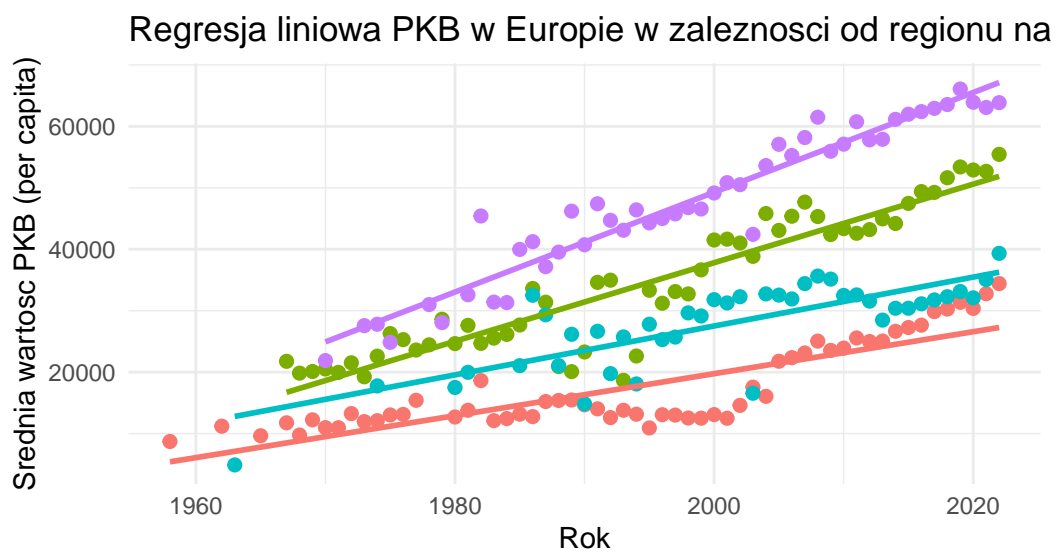
The 15 Countries with Lowest GDP in Europe in 2020



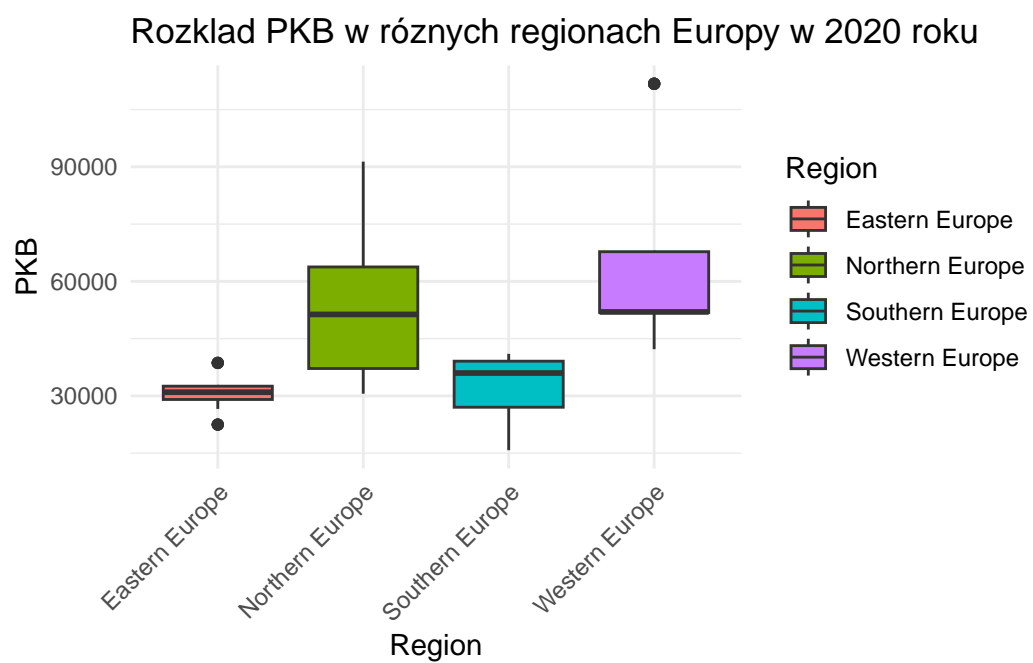
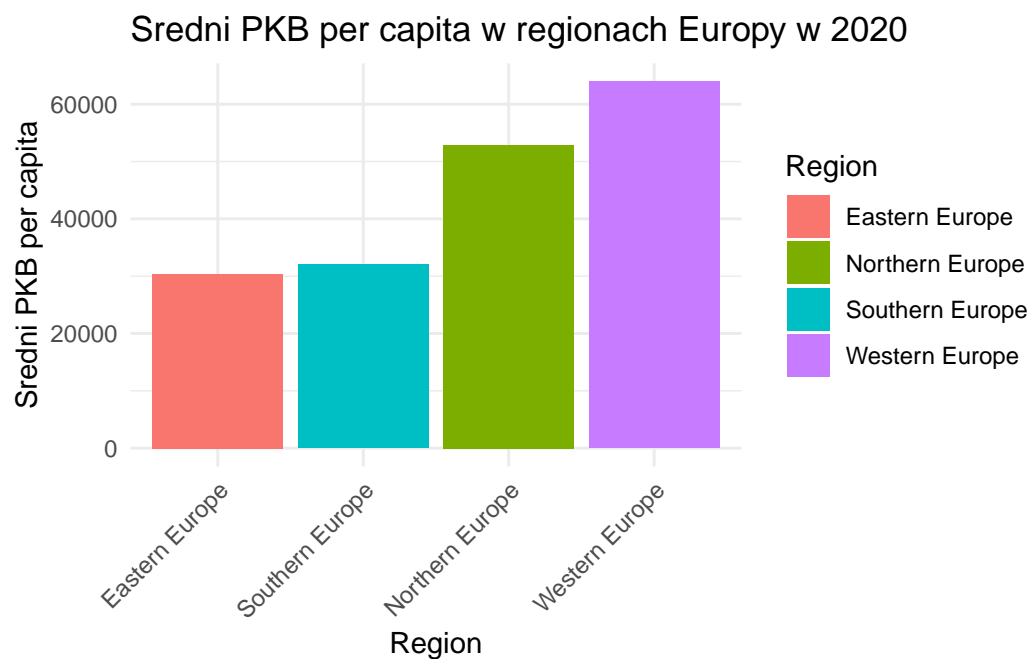
The 15 Countries with Highest GDP in Europe in 2000

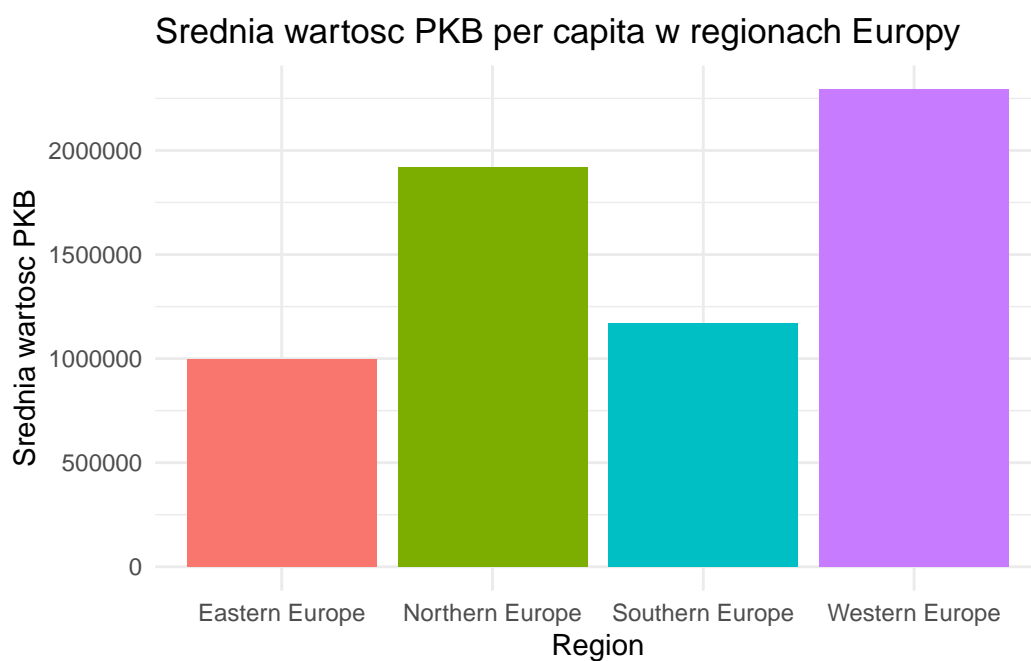
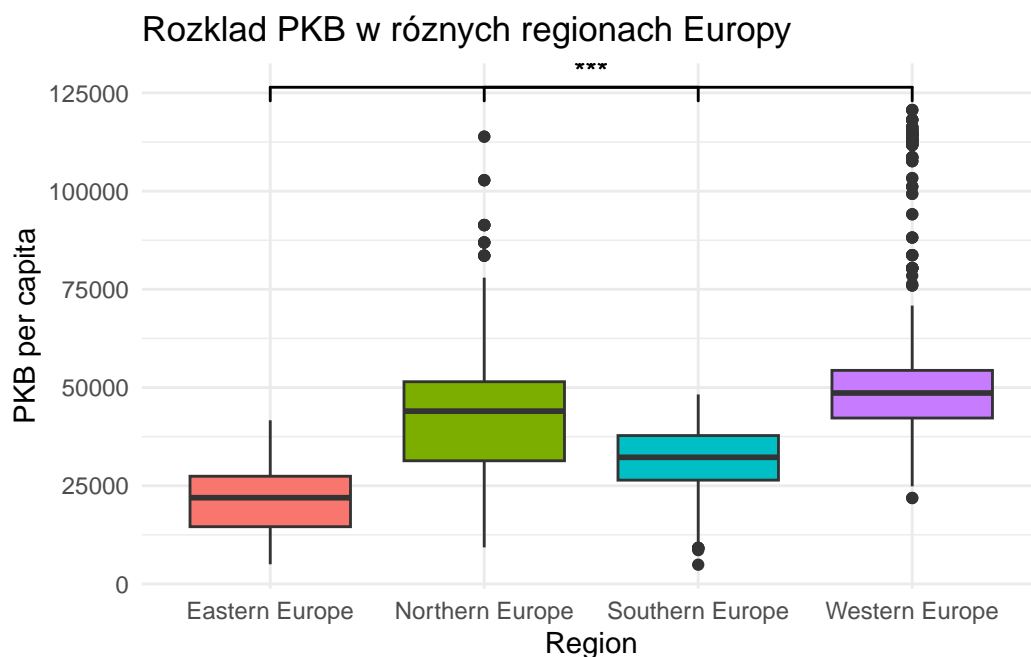






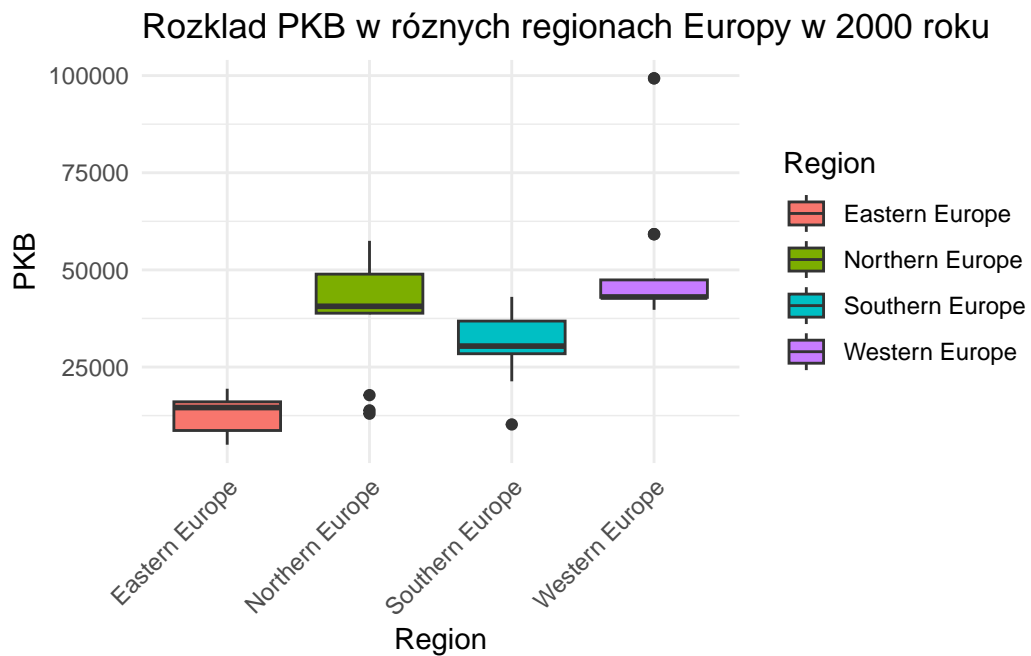






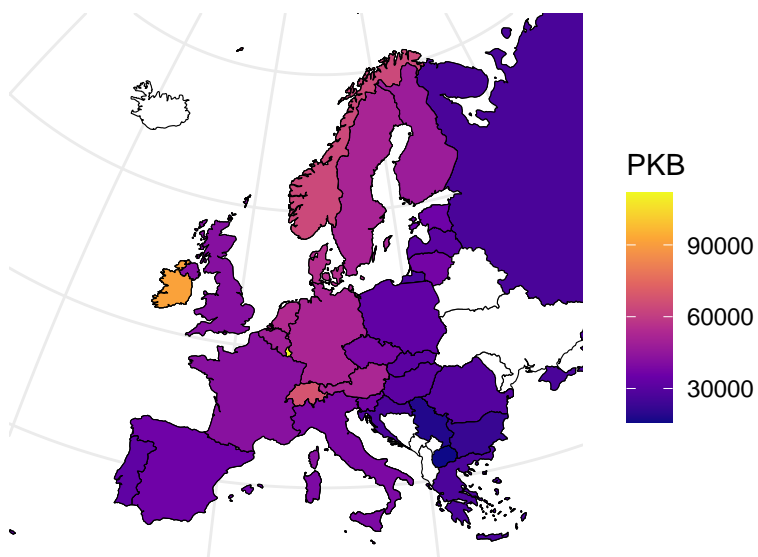
```
# A tibble: 12 x 4
# Rowwise:
  Region1      Region2      p_value p_adjusted
  <fct>      <fct>      <dbl>    <dbl>
1 Eastern Europe Western Europe 6.94e-239 6.94e-239
```

2	Western Europe	Eastern Europe	6.94e-239	6.94e-239
3	Northern Europe	Eastern Europe	1.05e-217	1.05e-217
4	Eastern Europe	Northern Europe	1.05e-217	1.05e-217
5	Western Europe	Southern Europe	3.14e-144	3.14e-144
6	Southern Europe	Western Europe	3.14e-144	3.14e-144
7	Northern Europe	Southern Europe	7.80e- 77	7.80e- 77
8	Southern Europe	Northern Europe	7.80e- 77	7.80e- 77
9	Eastern Europe	Southern Europe	3.11e- 74	3.11e- 74
10	Southern Europe	Eastern Europe	3.11e- 74	3.11e- 74
11	Northern Europe	Western Europe	4.75e- 40	4.75e- 40
12	Western Europe	Northern Europe	4.75e- 40	4.75e- 40



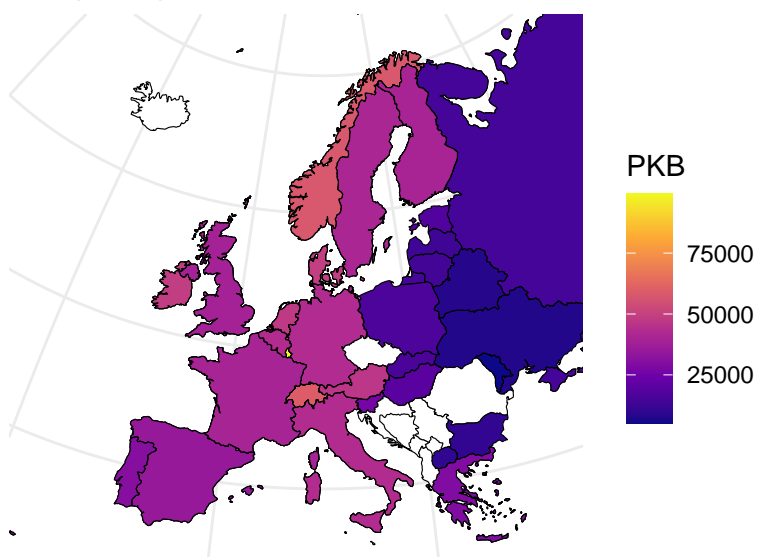
## Mapa PKB na świecie (Wybrane kraje, rok 2020)

PKB per capita w 2020 roku



## Mapa PKB na świecie (Wybrane kraje, rok 2000)

PKB per capita w 2000 roku



### Akapit

2. Czy dołączenie do UE miało wpływ na zmianę GDP?

## Akapit

3. Zbadanie współczynnika nierówności dochodowej w Europie.

## Wnioski

Na końcu tej sekcji przedstaw wnioski związane z odpowiedziami na pytania badawcze, np. "Z analizy wynika, że istnieje silna korelacja między PKB a poziomem nierówności dochodowych."

## Podsumowanie

Przypomnienie najważniejszych wniosków z analizy.

## Dyskusja

Myślę, że to bardzo fajny dodatkowy akapit, który warto uwzględnić na lepsze punkciki :)

- **Ograniczenia:** Omówienie wszelkich ograniczeń analizy, np. problem z jakością danych, brakującymi wartościami, czy ograniczonym zakresem badanych zmiennych.
- **Propozycje przyszłych badań:** Zasugerowanie, co można by zbadać w przyszłości, aby poszerzyć lub pogłębić wyniki.