**Rola przepływu towarów i dóbr oraz środków finansowych na kształtowanie PKB per capita w Polsce**

Wykonawcy

Michał Głąb (50%)

Mateusz Dzięgielewski (50%)

**Wprowadzenie**

W obecnych czasach przepływ zarówno różnego rodzaju towarów jak i środków pieniężnych jest znacznie ułatwiony. Żyjąc w erze globalizacji można uczestniczyć w wymianie handlowej na wielką skalę i to z najdalszymi zakątkami globu. Podobnie można czerpać finansowanie z  całego świata lub też dokonywać inwestycji alokując swoje środki za granicą. Poniższa analiza skupia się na zbadaniu roli jaką pełnią przepływy towarów i dóbr oraz środków finansowych na kształtowanie Produktu Krajowego Brutto per capita w Polsce. Czy niespotykana dotąd swoboda handlu, transportowania dóbr i dokonywania inwestycji powoduje, że Polska się bogaci, a jej mieszkańcom żyje się dostatniej? Czy te same przepływy, ale w obrębie granic naszego kraju też mają znaczenie? Które konkretne czynniki odgrywają największą rolę w  kształtowaniu się polskiego PKB per capita?

**Opis wykorzystanych danych**

Dane pochodzą z Banku Danych Lokalnych GUS oraz ze strony Gapminder.org. Wszystkie dane były kompletne, jednak ich różne pochodzenie i ilość potencjalnych zmiennych ograniczyły zaprezentowaną ilość obserwacji. Także ze względu na zróżnicowane rzędy wielkości lub ukazanie wartości jako % PKB, zostały dokonane odpowiednie zmiany aby doprowadzić dane źródłowe do ich ostatecznej postaci, w której będą najbardziej użyteczne. Potencjalne zmienne i ich obserwacje prezentuje tabela 1.

Tabela Dane Obraz zawierający tekst, zrzut ekranu, numer, linia

Opis wygenerowany automatycznie

Obliczenia własne

* PKB per capita – wskaźnik ten lepiej odwzorowuje zamożność obywateli niż zwykłe PKB, dlatego też lepiej pasuje do zdefiniowanych problemów badawczych.
* Drogi ekspresowe i autostrady [tys. km] – Rozbudowana infrastruktura drogowa jest kluczowa, jeśli chodzi o możliwości transportu dóbr na terenie kraju. Nie bez powodu nazywa się ją krwioobiegiem gospodarki, gdyż za jej pomocą przewozi się znaczną część dóbr. Dodatkowo, to właśnie w pobliżu dróg ekspresowych i autostrad powstają w dzisiejszych czasach zakłady przemysłowe i magazyny, których funkcjonowanie także przyczynia się do kreowania polskiego PKB.
* Kredyty i pożyczki [mld zł] – zmienna ta opisuje kredyty i pożyczki krajowe jako źródło finansowania nakładów inwestycyjnych przedsiębiorstw. Opisuje ona przepływ nie dóbr fizycznych, a kapitału, który jest równie ważny w pomnażaniu bogactwa i  tworzeniu dobrobytu. Dług jest powszechną formą finansowania inwestycji w  przedsiębiorstwach, a więc jego występowanie nie musi oznaczać kłopotów finansowych, ale też możliwość rozwoju, dzięki nowo pozyskanym środkom.
* Załadunek z tranzytem [tys. t] – Transport morski jest najczęściej używanym środkiem przewozu dużych ilości towarów na dalekich dystansach, a więc odgrywa kluczową rolę dla wymiany handlowej w czasach globalnej gospodarki.
* Inwestycje zagraniczne [mld $] – To czego nie mogą sfinansować krajowi inwestorzy, najczęściej finansowane jest środkami z zagranicy, a wraz z zagranicznym kapitałem, do kraju trafiają technologia i ‘know-how’, które często stanowią już wartość samą w  sobie, co również przyczynia się do kreowania dobrobytu.
* Bilans handlowy [mld $] – Tak jak już było to wspomniane, handel zagraniczny odgrywa znaczącą rolę w dzisiejszej gospodarce, a dodatni jego bilans, to oczywiście dodatkowy wzrost gospodarczy.

**Metody**

W celu oszacowania formy modelu i jego dalszej weryfikacji wykorzystano następujące metody i testy statystyczne:

1. Statystyka Opisowa:

- Średnie i Odchylenia Standardowe: Do obliczania średniej arytmetycznej (miara centralnego położenia danych) i odchyleń standardowych (miara rozproszenia danych wokół średniej) dla poszczególnych zmiennych.

- Macierz Korelacji i Wykres Korelacji: Użycie macierzy korelacji pozwala na ocenę związku między różnymi zmiennymi. Wykres korelacji, utworzony za pomocą funkcji `corrplot`, wizualizuje te zależności, ułatwiając identyfikację potencjalnych związków lub redundancji między zmiennymi.

2. Model Regresji Liniowej:

- Budowa Modelu Regresji Liniowej za pomocą KMNK: Zastosowanie modelu regresji liniowej umożliwia przewidywanie wartości zmiennej zależnej (w tym przypadku PKB per capita) na podstawie jednej lub więcej zmiennych niezależnych.

- Diagnostyka Modelu Regresji: Wykorzystanie funkcji `summary(model\_reg)` do oceny modelu, w tym analiza wartości p testu t dla poszczególnych zmiennych niezależnych, co pozwala ocenić ich znaczenie statystyczne w modelu.

- Ponowne szacowanie modelu przy użyciu UMNK: Jeden ze sposobów tworzenia Modelu Regresji, stosowany w przypadku zarówno samodzielnego jak i jednoczesnego występowania autokorelacji i heteroskedastyczności składnika losowego.

3. Analiza Reszt:

- Test Serii: Sprawdzenie, czy reszty są rozłożone losowo, co jest kluczowym założeniem w modelowaniu regresyjnym.

- Test Durbina-Watsona: Testowanie autokorelacji reszt, czyli sprawdzenie, czy błędy modelu są od siebie niezależne.

- Test Shapiro-Wilka: Testowanie normalności rozkładu reszt, co jest jednym z założeń modelu regresji liniowej.

- Test Goldfelda-Quandta: Sprawdzenie, czy wariancja składnika losowego jest stała, czy też zmienia się w czasie.

Każda zaprezentowana wyżej metoda oraz test odgrywają istotną rolę w procesie modelowania statystycznego i ekonometrycznego, od podstawowej analizy danych po bardziej zaawansowane techniki oceny modelu. Wszystkie metody użyte w naszych obliczeniach pochodzą z Dziechciarz Józef. (2003) *Ekonometria. Metody, przykłady, zadania.*

**Wstępna analiza danych**

Analizę warto rozpocząć od sprawdzenia podstawowych statystyk opisowych każdej ze zmiennych, co zaprezentowane zostało w tabelach 1 i 2.

Tabela 2 Statystyki Opisowe

Obraz zawierający tekst, zrzut ekranu, Czcionka

Opis wygenerowany automatycznie

Obliczenia własne

Tabela 3 Statystyki Opisowe c.d.

Obraz zawierający tekst, zrzut ekranu, Czcionka

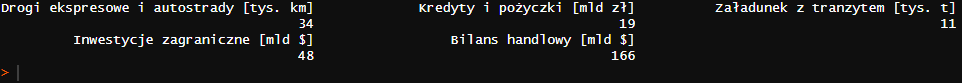
Opis wygenerowany automatycznie

Obliczenia własne

Na pierwszy rzut oka można wywnioskować, że średnie we wszystkich przypadkach niewiele różnią się od wartości mediany, co oznacza brak obserwacji odstających. Jeśli chodzi o zakres obserwacji, wyjątkiem jest bilans handlowy, którego minimalna wartość jest ujemna, bierze się to ze sposobu obliczania bilansu i oznacza, że w danym okresie import był wyższy od eksportu.

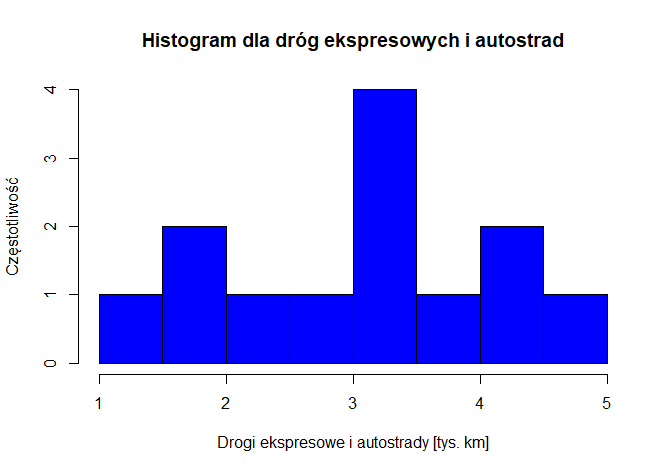
Ważne pod kątem dalszej budowy modelu regresji jest określenie współczynnika zmienności dla każdego z obserwowanych czynników. Jak widać w tabeli 3, każdy z nich przekracza granicę 10%, a więc zmienność jest wystarczająco duża aby dane zostały użyte w kolejnych analizach.

Tabela 4 Współczynniki zmienności



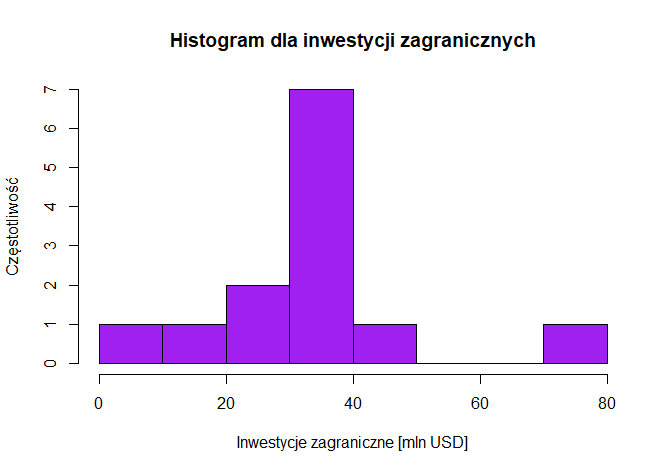
Obliczenia własne

Histogram dróg ekspresowych i autostrad, który jest jednomodalny i symetryczny, przypominający rozkład normalny, wskazuje na to, że większość tych dróg ma zbliżoną długość. Jednomodalność oznacza dominujący wzór w danych, w tym przypadku sugerując, że istnieje najczęściej występująca długość dróg. Symetria histogramu wokół tej centralnej wartości wskazuje, że długości krótsze i dłuższe od tej modalnej wartości pojawiają się z podobną częstotliwością.



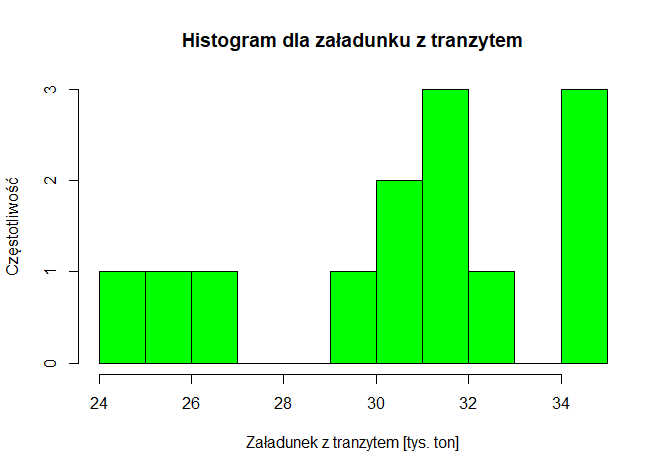
Rysunek Rozkład dróg ekspresowych i autostrad opracowanie własne

Histogram ilustrujący inwestycje zagraniczne z jedną modalną wartością i lewostronną asymetrią sugeruje, że większość inwestycji zagranicznych ma wartość bliską tej dominującej, modalnej wartości, ale równocześnie występuje długi ogon mniejszych inwestycji.



Rysunek Rozkład inwestycji zagranicznych, opracowanie własne

Histogram przedstawiający rozkład częstości załadunku z tranzytem, w którym największa liczba obserwacji znajduje się w przedziale powyżej 30 tysięcy ton, daje cenne informacje na temat charakterystyki transportu towarów. Wysoka częstość załadunków w tym przedziale wskazuje, że znaczna część transportu obejmuje duże ładunki, co może świadczyć o intensywności handlu i logistyki w obszarach, gdzie dokonuje się takie załadunki.



Rysunek rozkład załadunku z tranzytem, opracowanie własne

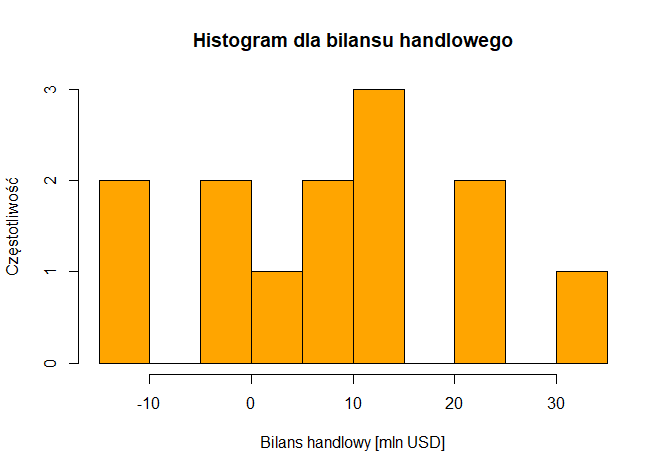
Histogram przedstawiający rozkład częstości kredytów i pożyczek w milionach złotych, z dominującymi wartościami w przedziale 30-40 milionów złotych, wskazuje na specyfikę zachowań kredytowych w badanym obszarze. Największa liczba kredytów i pożyczek skoncentrowana w tym przedziale świadczy o tym, że dla większości kredytobiorców, instytucji finansowych lub projektów inwestycyjnych, kwoty w zakresie 30-40 milionów złotych są najbardziej preferowane

Obraz zawierający tekst, diagram, zrzut ekranu, linia

Opis wygenerowany automatycznie

Rysunek Rozkład kredytów i pożyczek, opracowanie własne

Histogram prezentuje rozkład bilansu handlowego wyrażony w milionach dolarów amerykańskich (USD). Najczęściej występujący przedział wartości to około 10 milionów USD, gdzie częstotliwość wynosi 3.



Rysunek 5 Rozkład bilansu handlowego, opracowanie własne

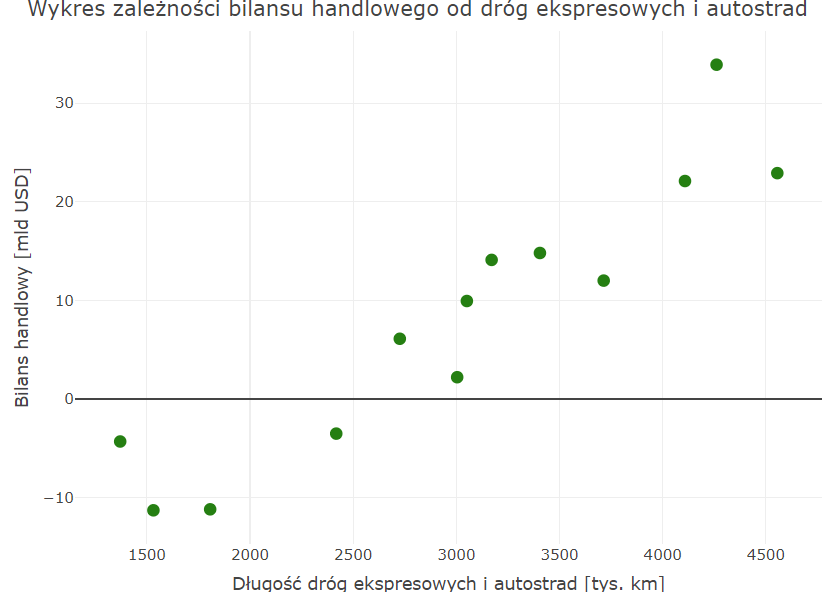
Znając rozkład zmiennych należy również sprawdzić ich wzajemne zależności, co przyda się później w procesie tworzenia modelu. Macierz korelacji dla zmiennych objaśniających została przedstawiona na rysunku 6.

Obraz zawierający tekst, zrzut ekranu, diagram, Wielobarwność

Opis wygenerowany automatycznie

Rysunek Macierz korelacji zmiennych objaśniających, opracowanie własne

Poszczególne zmienne na tworzą ze sobą grupy na podstawie ich wzajemnej korelacji. Wyróżnić można kredyty i pożyczki wraz z inwestycjami zagranicznymi, drogi ekspresowe i autostrady z bilansem handlowym, a także oddzielnie, załadunek z tranzytem. Dwie najsilniej skorelowane ze sobą zmienne, a więc drogi ekspresowe i autostrady oraz bilans handlowy dodatkowo zostały przedstawione na wykresie rozrzutu na rysunku 7.



Rysunek 7 Wykres zależności między bilansem handlowym, a drogami ekspresowymi i autostradami, opracowanie własne

Rozrzut poszczególnych punktów układa się w dodatnią, niemal liniową, zależność. Nie widać żadnych wyraźnie odstających punktów co jest potwierdzeniem silnego związku pomiędzy tymi zmiennymi.

**Model liniowy**

Za pomocą metody grafowej wyznaczyliśmy model w którym zmienną objaśnianą jest PKB per capita. A zmiennymi objaśniającymi są drogi ekspresowe i autostrady, inwestycje zagraniczne oraz załadunek z tranzytem. Oszacowanie parametrów strukturalnych, błędy szacunków, a także testy istotności zostały przedstawione w tabeli 4.

Tabela 5 Model regresji z trzema zmiennymi

Obraz zawierający tekst, zrzut ekranu, Czcionka

Opis wygenerowany automatycznie

Obliczenia własne

Skorygowany współczynnik determinacji na poziomie 94% świadczy o dobrym dopasowaniu danych do modelu. Również test Fischera wykazał, że wektor parametrów strukturalnych jest istotnie różny od zera. Niestety z testu Studenta wynika, że z modelu trzeba wykluczyć zmienną opisującą załadunek z tranzytem ze względu na mała istotność parametru. Ponownie oszacowany model uwzględniający już tylko dwie zmienne objaśniające opisuje tabela 5.

Tabela 6 Model regresji z dwoma zmiennymi, obliczenia własne

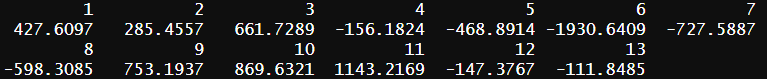
Obraz zawierający tekst, zrzut ekranu, Czcionka

Opis wygenerowany automatycznie

*Obliczenia własne*

Jak można zauważyć, tym razem wszystkie parametry są istotne, a skorygowany współczynnik determinacji nieznacznie wzrósł. Oznacza to, że model uwzględniający dwie zmienne lepiej opisuje badane zjawisko i właśnie w tej formie powinien zostać zastosowany. Mając już określone parametry strukturalne należy sprawdzić losowość reszt aby upewnić się, że postać liniowa modelu jest odpowiednią do opisu zmiennej objaśnianej. Tabela 6 zawiera reszty modelu na podstawie, których zostanie przeprowadzony test serii.

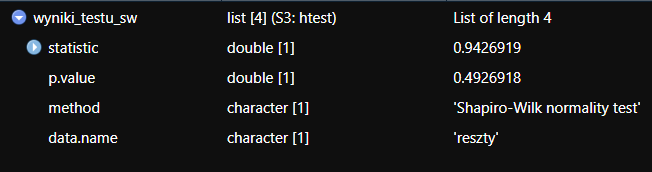
Tabela 7 Reszty modelu



*Obliczenia własne*

Dodatnim wartościom reszt przypisujemy literę A, a ujemnym B. Przez serię rozumiemy ciąg jednakowych symboli np. AAA lub BBBB. (Dziechciarz, 2003). Ilość serii wynosi oznaczona jako Ke 4. Liczebność reszt dodatnich to 6, a ujemnych 7. Liczebności te potrzebne są do odczytania wartości krytycznych K1 i K2 z tablic. Wynoszą one kolejno 4 i 10. Hipotezę zerową, mówiącą o losowości reszt odrzucamy, kiedy Ke jest nie zawiera się w zamkniętym przedziale pomiędzy K1 i K2. W tym przypadku, nie ma podstaw do odrzucenia hipotezy zerowej, a więc reszty są losowe, a postać linowa modelu jest prawidłowa.

Tabela 8 Wyniki testu Shapiro-Wilka



*Obliczenia własne*

Otrzymana wartość statystyki z testu Shapiro-Wilka wynosi 0,943, a p-value to 0,493. Analiza tych wyników wskazuje na to, że nie ma podstaw do odrzucenia hipotezy zerowej (H0), która zakłada, że rozkład danych jest normalny. W praktyce, wartość p-value równa 0,493 jest znacznie wyższa niż tradycyjny poziom istotności 0,05, co sugeruje brak podstaw do odrzucenia hipotezy o normalności rozkładu.

Tabela 9 Wyniki testu Goldfelda-Quandta

Obraz zawierający tekst, Czcionka, zrzut ekranu

Opis wygenerowany automatycznie

*Obliczenia własne*

Wyniki testu Goldfelda-Quandta to statystyka testowa o wartości 11,183 i p-value równe 0,0841. Ten test jest używany w analizie regresji do sprawdzenia hipotezy dotyczącej homoskedastyczności (równości wariancji) reszt regresji pomiędzy dwiema grupami. Otrzymane wyniki świadczą o tym , że nie ma podstaw do odrzucenia hipotezy zerowej. Zatem wariancja reszt nie różni się istotnie pomiędzy dwiema grupami.

Tabela 10 Wyniki testu Durbina-Watsona

Obraz zawierający tekst, zrzut ekranu, Czcionka, oprogramowanie

Opis wygenerowany automatycznie

*Obliczenia własne*

Statystyka testowa wynosi w zaokrągleniu 0,986 a p-value jest na poziomie 0,0048 co świadczy na rzecz przyjęcia hipotezy alternatywnej o występowaniu autokorelacji. Autokorelacja składnika losowego bardzo często występuje w modelach tworzonych na podstawie szeregów czasowych. W tym przypadku może ona również występować z powodu inercji zjawisk gospodarczych, czyli inaczej mówiąc, ich bezwładności. Zmiany w gospodarce mogą potrzebować czasu, aby wywołać odpowiedni efekt. Model zawierający autokorelację nie jest dokładny w swoich szacunkach, a więc należy go poprawić, a jednym ze sposobów jest zastosowanie UMNK przy jednoczesnym uchyleniu jednego z założeń Gausa-Markowa dotyczącego sferyczności składnika losowego. Tak oszacowany model pozwala dokładniej odwzorować badane zjawisko. Ostateczną postać modelu przedstawia równanie 1.

Równanie 1 Ostateczna postać równania regresji, obliczenia własne

Na tej podstawie można wnioskować, że wraz z każdym tysiącem kilometrów wybudowanych dróg ekspresowych i autostrad, PKB per capita w Polsce rośnie o 3330,32$ lub rośnie o 32,43$ wraz z każdym miliardem dolarów pochodzącym z inwestycji zagranicznych.

**Podsumowanie**

Z naszego badania wynika, że na wzrost PKB per capita, a więc ogólnego dobrobytu mieszkańców Polski, najmocniej wpływają infrastruktura drogowa oraz inwestycje zagraniczne. Ulepszenie infrastruktury drogowej przyczynia się do zwiększenia efektywności transportu, co z kolei może przyspieszyć rozwój gospodarczy poprzez ułatwienie dostępności rynków, zmniejszenie kosztów logistycznych dla przedsiębiorstw i poprawę ogólnej mobilności pracowników. Z drugiej strony, inwestycje zagraniczne przynoszą do Polski kapitał, nowoczesne technologie oraz know-how, wzmacniając sektor prywatny, tworząc miejsca pracy i zwiększając konkurencyjność kraju na arenie międzynarodowej. Te inwestycje często prowadzą do zwiększenia eksportu i integracji Polski z globalnymi łańcuchami wartości.

Jest istotne, aby zauważyć współzależność między rozwijającą się infrastrukturą drogową a  przyciąganiem inwestycji zagranicznych. Rozbudowana infrastruktura ułatwia zagranicznym inwestorom dostęp do lokalnych rynków i zasobów, czyniąc Polskę bardziej atrakcyjną jako miejsce inwestycji. Istotne jest również uwzględnienie potencjalnego ryzyka związanego z  nadmiernym poleganiem na kapitale zagranicznym, które może prowadzić do zależności gospodarczej, oraz konieczność zapewnienia, by rozwój infrastruktury drogowej był zrównoważony i uwzględniał aspekty społeczne i środowiskowe.

**Literatura**

Dziechciarz Józef. (2003) *Ekonometria. Metody, przykłady, zadania*. Wrocław: Wydawnictwo Akademii Ekonomicznej im. Oskara Langego we Wrocławiu

Gapminder. (2023). *GPD per capita (Price and inflation adjusted, in PPP$2027).* Pobrano z  http://gapm.io/dgdpcap\_cppp

Główny Urząd Statystyczny. (2023). *Drogi ekspresowe i autostrady.* Pobrano z  https://bdl.stat.gov.pl/bdl/metadane/cechy/1722

Główny Urząd Statystyczny. (2023). *Międzynarodowy obrót morski w portach morskich.* Pobrano z https://bdl.stat.gov.pl/bdl/metadane/cechy/2618

Główny Urząd Statystyczny. (2023). *Źródła finansowania nakładów inwestycyjnych w  przedsiębiorstwach.* Pobrano z https://bdl.stat.gov.pl/bdl/metadane/cechy/3445?back=True

The World Bank. (2023). *External balance on goods and services (current US$).* Pobrano z  https://data.worldbank.org/indicator/NE.RSB.GNFS.CD

The World Bank. (2023). *Foreign direct investment, net flows (% GDP).* Pobrano z  https://data.worldbank.org/indicator/BX.KLT.DINV.WD.GD.ZS