Rozdział III

3.1. Źródło danych i opis zmiennych

Dane przekrojowe wykorzystane w modelu ekonometrycznym dotyczą 380 powiatów i miast na prawach powiatu w Polsce i wyrażają stan na rok 2020. Do przeprowadzenia estymacji regresji liniowej za pomocą Metody Najmniejszych Kwadratów (MNK) i stworzenia wykresów i tabel znajdujących się w pracy wykorzystano język programowania R. Praca opiera się na danych pochodzących z trzech źródeł: Banku Danych Lokalnych, projektu Michała Rogalskiego – *COVID-19 w Polsce* oraz archiwalnych danych dla powiatów związanych z koronawirusem COVID-19 udostępnionych przez Ministerstwo Zdrowia.

Wskaźniki i dane dotyczące sytuacji gospodarczej, demograficznej, społecznej powiatów zostały pobrane z Banku Danych Lokalnych, za którego prowadzenie i rozwój odpowiedzialny jest Główny Urząd Statystyczny w Polsce.

Baza danych COVID-19 w Polsce stworzona przez Michała Rogalskiego na podstawie raportów Ministerstwa Zdrowia, Wojewódzkich Stacji Sanitarno-Epidemiologicznych (WSSE), Powiatowych Stacji Sanitarno-Epidemiologicznych (PSSE) i Urzędów Wojewódzki została wykorzystana do pobrania danych dotyczących sumy przypadków zakażeń koronawirusem i zgonów nim spowodowanych w powiatach. Okres sumowania zgonów jest różny dla różnych województw i zależy to od daty zaprzestania publikowania danych dotyczących pandemii przez PSSE i WSSE.

Dane archiwalne dotyczące zarażeń koronawirusem dla powiatów opublikowane przez Ministerstwo Zdrowia (MZ) posłużyły do oszacowania łącznej liczby osób przebywających na kwarantannie dla poszczególnych powiatów. W archiwum znajdują się statystyki dotyczące COVID-19 w przedziale od 24.11.2020 do 31.12.2020 roku, związane jest to ze zmianą metodologii liczenia przypadków zachorowań i zgonów spowodowanych koronawirusem przez MZ. Zostały one wykorzystane do oszacowania łącznej liczby osób przebywających na kwarantannie w 2020 roku. Sposób szacowania:

1. Na podstawie danych z okresu 23.11.2020 – 31.12.2020 wyliczono średnią dzienną liczbę osób objętych kwarantanną.

2. Zgodnie z rozporządzeniem Ministra Zdrowia, stan epidemii koronawirusa w Polsce obowiązuje od 20.03.2020 roku. Średnią dzienną liczbę osób objętych kwarantanną wyliczoną w punkcie 1. pomnożono przez liczbę dni objętych stanem epidemii (286 dni) i zaokrąglono do liczb całkowitych.

3. Jeżeli dany powiat nie został uwzględniony w zestawieniu Ministerstwa Zdrowia, to dane dotyczące kwarantann na jego obszarze zostały wyliczone na podstawie średniej dziennej liczby osób objętych kwarantanną dla całego kraju.

3.2. Opis zmiennych

3.2.1. Zmienna objaśniana

Zmienną zależną w szacowanym modelu jest współczynnik urodzeń (*birth.rate*), który określa liczbę żywych urodzeń na 1000 ludności.

3.2.2. Zmienne objaśniające

Wszystkie zmienne niezależne użyte w modelu są zmiennymi ciągłymi.

* *covid19\_cases* - suma przypadków zachorowań na chorobę COVID-19.
* *covid19\_deaths* – suma zgonów wywołanych chorobą COVID-19.
* *covid19\_quarantine* – suma osób przebywających na kwarantannie z powodu kontaktu z osobą chorą na COVID-19.
* *marriage\_rate* – współczynnik małżeństw określający liczbę zawartych małżeństw na 1000 ludności.
* *divorce\_rate* – współczynnik rozwodów, który jest stosunkiem liczby orzeczonych rozwodów na 1000 ludności.
* *budget\_reve\_pc* – dochody budżetów powiatów (lub miast na prawach powiatu) w przeliczeniu na jednego mieszkańca. Zmienna wyrażona w złotych polskich.
* *unemployment\_rate* – stopa bezrobocia rejestrowanego wyrażona w procentach.
* *women\_reproductive* – liczba kobiet w wieku reprodukcyjnym, czyli między 15. a 49. rokiem życia.
* *femininity\_ratio* – współczynnik feminizacji określający liczbę kobiet przypadającą na 100 mężczyzn.
* *avg\_salary* – średnie wynagrodzenie brutto. Zmienna wyrażona w złotych polskich.
* *women\_working* – odsetek kobiet pracujących jako stosunek liczby pracujących kobiet do liczby kobiet w danym powiecie.
* *men\_working* – odsetek mężczyzn pracujących jako stosunek liczby pracujących mężczyzn do liczby mężczyzn w powiecie.
* *median\_house\_price* – mediana cen za 1 m2 lokali mieszkaniowych sprzedanych w ramach transakcji rynkowych. Zmienna wyrażona w złotych polskich.
* *house\_ratio* – mieszkania oddane do użytkowania na 1000 ludności.
* *houses\_area\_pc* – stosunek powierzchni użytkowej mieszkań w danym powiecie do jego populacji. Zmienna wyrażona w metrach kwadratowych.
* *avg\_people\_per\_house* – średnia liczba osób przypadająca na jedno mieszkanie.
* *education\_expenditure* - wydatki budżetów powiatów (i miast na prawach powiatu) w dziale 801 – Oświata i wychowanie w przeliczeniu na jednego mieszkańca. Zmienna wyrażona w złotych polskich.
* *health\_expenditure* - wydatki budżetów powiatów (i miast na prawach powiatu) w dziale 851 – Ochrona zdrowia w przeliczeniu na jednego mieszkańca. Zmienna wyrażona w złotych polskich.
* *social\_expenditure* – wydatki budżetów powiatów (i miast na prawach powiatu) w dziale 852 – Pomoc społeczna w przeliczeniu na jednego mieszkańca. Zmienna wyrażona w złotych polskich.
* *family\_expenditure* – wydatki budżetów powiatów (i miast na prawach powiatu) w dziale 855 – Rodzina w przeliczeniu na jednego mieszkańca. Zmienna wyrażona w złotych polskich.
* *children* – liczba osób w przedziale wiekowym 0 – 2 w danym powiecie.
* *nursery\_places* – miejsca w żłobkach, oddziałach żłobków i klubach dziecięcych.
* *doctors* – personel pracujący w sektorze ochrony zdrowia na 10 tysięcy ludności.
* *urbanisation\_rate* – współczynnik urbanizacji określający udział mieszkańców miast w liczbie ludności powiatu (lub miasta na prawach powiatu).
* *bus\_stops* – liczba przystanków autobusowych (z trolejbusowymi) i tramwajowych.

Statystyczna analiza zmiennych

Przed zbudowaniem wstępnego modelu chcę bliżej przyjrzeć się zmiennym, tak aby móc zaproponować jak najlepszą formę funkcyjną.

1.1. Zmienna objaśniana – współczynnik urodzeń

bibliografia

<https://docs.google.com/spreadsheets/d/1Tv6jKMUYdK6ws6SxxAsHVxZbglZfisC8x_HZ1jacmBM/edit#gid=42091932>

<https://bdl.stat.gov.pl/BDL/start>

baza danych covidowa mz gov