

Projekt I

Panel z danymi o COVID-19

Sprawozdanie
"Geoportale i usługi danych
przestrzennych"
Rok akademicki 2020/2021

Wykonały:

Natalia Gadocha, nr 304165
Katarzyna Kowalczyk, nr 304174
Karolina Kuna, nr 304177

Wstęp	2
Cel projektu	2
Wykorzystane technologie	2
Dane	3
Wykorzystane dane	3
Problem aktualności danych	3
Szczegółowy przebieg projektu	4
Podział obowiązków	4
I faza - przygotowanie danych	4
II faza - odświeżanie danych	6
III faza - wykonanie dashboardu	7
Panel wynikowy	7
Założenia	7
Funkcjonalności	7

Podsumowanie	9
Problemy - dane	9
Problemy - ArcGIS Online	9

1. Wstęp

1.1. Cel projektu

Celem projektu jest utworzenie interaktywnego panelu, który przedstawia dane dotyczące wirusa Covid-19 dla Polski. Zamysłem jest, by na głównym ekranie pojawiła się mapa zawierająca podział na województwa, wraz z intuicyjnymi funkcjonalnościami, takimi jak wykresy oraz statystyki.

Ważne, by użytkownicy w szybki i bezpośredni sposób otrzymali dostęp do najnowszych danych. Przede wszystkim będą to dane dotyczące codziennej liczby przypadków i zgonów, całkowitej sumy przypadków oraz statystyki dla każdego z województw.

Panel jest narzędziem, którego obecnie brakuje na rynku i może stanowić rzetelne oraz szybkie źródło informacji dla każdego mieszkańca.

1.2. Wykorzystane technologie

Podczas pracy wykorzystamy narzędzie ArcGIS Online, które umożliwia tworzenie intuicyjnych panelów z danymi przestrzennymi. W początkowych fazach projektu zostanie także użyte funkcjonalności ArcGIS Desktop.

Aby rozwiązać problem aktualności danych, posłużymy się także narzędziem FME Desktop, które umożliwi sprawne przesyłanie uaktualnionych danych na serwer.

2. Dane

2.1. Wykorzystane dane

Dane - COVID-19

Dane, które zostaną wykorzystane w projekcie, pochodzą z jedynej w Polsce udostępnionej publicznie bazy danych dotyczących Covid-19, którą znajdziemy pod poniższym linkiem:

<http://bit.ly/covid19-poland>

Baza danych to arkusz w serwisie Google, w którym znajdziemy wyczerpujące informacje na temat aktualnego stanu choroby w naszym kraju. Wykorzystamy dane dotyczące:

- Suma zachorowań
- Zachorowania dzisiaj w województwach
- Zachorowania wczoraj w województwach
- Liczby zgonów
- Liczba osób wyleczonych
- Liczba przypadków aktywnych

Dane - geometria województw

Dane są aktualizowane codziennie.

Dane dotyczące geometrii województw zostały pozyskane ze strony:

<http://www.gugik.gov.pl/pzgik/dane-bez-oplat/dane-z-panstwowego-rejestru-granic-i-powierzchni-jednostek-podzialow-terytorialnych-kraju-prg>

Tabelę atrybutów zmodyfikowano tak, by zawierała jedynie geometrię oraz nazwę województwa.

2.2. Problem aktualności danych

Na początku projektu problemem okazało się zsynchronizowanie panelu z danymi w taki sposób, aby informacje odświeżały się automatycznie. Rozwiązaniem okazało się jednak wykorzystanie programu FME Desktop. Program wczytuje zarówno dane źródłowe z Dokumentów Google, jak i dane w tabeli w serwisie ArcGis Online, wykorzystywane do panelu użytkownika. Po odpowiedniej transformacji, odświeżone dane wracają z powrotem do panelu. Dane odświeżają się codziennie o godzinie 18.

3. Szczegółowy przebieg projektu

3.1. Podział obowiązków

Projekt został podzielony na trzy najważniejsze fazy. Każda z nich została powierzona jednej osobie, z czego podział prezentuje się następująco:

- faza I - Karolina Kuna
- faza II - Natalia Gadocha
- faza III - Katarzyna Kowalczyk

W trakcie każdego z etapów niezmiernie ważna była współpraca oraz zaangażowanie w sprawy i problemy bieżące, co można było dostrzec u każdego z członków grupy.

3.2. I faza - przygotowanie danych

Pierwsza część przygotowania danych polegała na zmodyfikowaniu tabeli województw tak, by zostały jedynie ich nazwy oraz geometria.

Druga część dotyczyła już właściwych danych.

1. Po przeanalizowaniu arkusza z bazą danych wybrano najważniejsze informacje.
2. W osobnym arkuszu Google stworzono szkielet tabeli, z kolumnami odpowiadającymi odpowiednim informacjom.
3. Połączono nowo powstałej tabeli z oryginalnym dokumentem przy pomocy odpowiedniej funkcji, stworzonej przez Google:

```
IMPORTRANGE ("https://docs.google.com/spreadsheets/d/1ierEhD6gcq51HAm433knjnVwey4ZE5DCnu1bW7PRG3E/edit#gid=0"; "Aktualna sytuacja w Polsce!C3:G18")
```

Funkcja ta umożliwiła także stałą aktualizację skopiowanych danych podczas modyfikacji oryginalnej tabeli.

Pozostałe dwie tabele zostały utworzone ręcznie.

Poniżej przedstawiono powstałe tabele.

Tabela I - testy

Testy_w (Features: 17, Selected: 0)

Nazwa	Suma_przebadanych_probek	Tygodniowa_liczba_przeba...	Suma_osob_z_wynikiem_p...	Tygodniowa_liczba_osob_z...	Liczba_testow_na_tysiac_mi...
Śląskie	949,707	17,395	75,884	4,474	205.60
Opolskie	68,020	1,298	15,832	1,127	67.60
Wielkopolskie	718,998	17,132	55,975	5,829	200.90
Świętokrzyskie	171,549	2,819	17,091	758	135.90
Pomorskie	637,030	19,724	31,799	4,865	264.50
Kujawsko-Pomorskie	417,628	14,160	34,688	5,150	194.20
Podlaskie	185,216	4,350	14,500	1,343	153.10
Zachodniopomorskie	336,410	11,882	17,643	4,484	190.80
Dolnośląskie	531,859	11,718	35,025	2,694	179.30
Podkarpackie	143,360	3,488	34,370	1,220	65.70
Małopolskie	508,900	7,990	63,878	1,817	147.30
Warmińsko-Mazurskie	179,924	4,472	14,613	3,644	122.80
Łódzkie	484,373	11,546	40,908	3,674	191.70
Mazowieckie	1,380,309	29,315	74,892	7,065	250.00
Lubelskie	287,416	6,760	30,117	2,545	132.50
Lubuskie	79,441	3,787	10,923	1,484	74.60

Tabela II - hospitalizacja

Hospitalizacja_W (obiekty: 16, wybrane: 0)

Nazwa	Liczba_osób_hospitalizowanych_	Liczba_łóżek_dla_pacjentów	Liczba_zajętych_respiratorów_	Liczba_respiratorów
Śląskie	1 580	3 573	155	288
Opolskie	384	1 064	27	72
Wielkopolskie	1 574	2 577	114	240
Świętokrzyskie	510	1 229	46	110
Pomorskie	1 104	1 835	86	132
Kujawsko-Pomorskie	1 316	1 929	132	168
Podlaskie	575	1 462	50	117
Zachodniopomorskie	1 077	1 578	91	121
Dolnośląskie	1 271	2 483	119	185
Podkarpackie	672	1 787	90	171
Małopolskie	811	2 342	95	236
Warmińsko-Mazurskie	804	1 159	84	107
Łódzkie	1 371	2 855	105	285
Mazowieckie	2 461	4 633	271	440
Lubelskie	1 144	2 266	117	216
Lubuskie	479	1 125	60	82

Tabela III - zachorowania

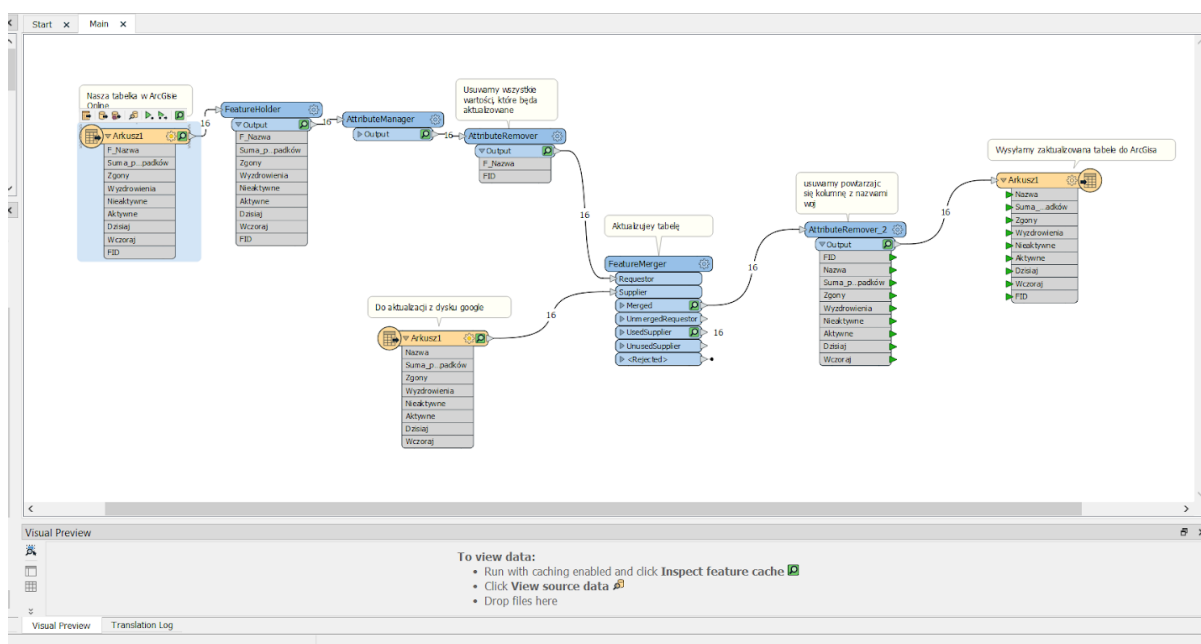
Arkusz1 (obiekty: 16, wybrane: 0)

_Nazwa	Suma_potwierdzone...	Zgony	Wyzdrowienia	Nieaktywne	Aktywne	Dzisiaj	Wczoraj
Dolnośląskie	89 528	2 000	35 525	37 525	52 003	115	246
Kujawsko-Pomorskie	87 221	1 562	16 234	17 796	69 425	484	673
Lubelskie	66 653	1 821	27 414	29 235	37 418	142	207
Lubuskie	31 838	688	10 608	11 296	20 542	96	169
Łódzkie	86 647	1 905	25 941	27 846	58 801	287	292
Małopolskie	109 558	2 682	64 233	66 915	42 643	194	200
Mazowieckie	160 625	3 480	69 307	72 787	87 838	746	903
Opolskie	37 059	1 057	12 053	13 110	23 949	84	100
Podkarpackie	59 960	1 861	28 904	30 765	29 195	116	118
Podlaskie	34 831	858	8 391	9 249	25 582	80	171
Pomorskie	84 563	1 588	23 013	24 601	59 962	623	705
Śląskie	157 262	3 448	63 784	67 232	90 030	282	288
Świętokrzyskie	35 480	930	12 414	13 344	22 136	59	77
Warmińsko-Mazurskie	54 869	1 208	11 877	13 085	41 784	139	376
Wielkopolskie	134 009	3 017	24 976	27 993	106 016	317	495
Zachodniopomorskie	64 253	1 051	13 790	14 841	49 412	436	518

3.3. II faza - odświeżanie danych

Problem odświeżania danych okazał się najpoważniejszy w całym projekcie. Próbowano różnych rozwiązań, jednak właściwym okazało się wykorzystanie programu FME Desktop. Dokładne połączenia danych obrazuje poniższa ilustracja.

FME - skrypt



1. Jako dane wejściowe zostaje dodana tabela, umieszczona na portalu ArcGIS Online. W procesie wczytywania, komórki za każdym razem są usuwane, co umożliwi późniejsze przesłanie uaktualnionych danych w miejsce poprzednich.
2. Przy pomocy narzędzia *Transformer - FeatureMerger* do tabeli wstawiane są nowe wartości. Tabele są łączone.
3. Dane zostają wysłane z powrotem na serwer.

Dodatkowo, do regularnego odświeżania danych została użyta aplikacja "Harmonogram zadań", znajdująca się w systemie Windows. Program uruchamia skrypt FME codziennie o godz. 18.

3.4. III faza - wykonanie dashboardu

Wykonanie ostatecznego widoku stanowiło równie ważną część projektu, ponieważ właśnie panel dotrze finalnie do odbiorców.

W wyniku ograniczeń programu ArcGIS - dane odświeżane za pomocą skryptu FME można było zapisać jedynie w swoich zasobach - panel został wykonany na jednym koncie użytkownika.

Podstawową kwestią było dobranie odpowiednich elementów i rozmieszczenie ich na stronie. Na podstawie wykonanych danych została wykonana mapa, wykresy, licznik zachorowań, a także boczny panel ze statystykami.

Na tym etapie powstało także końcowe sprawozdanie.

4. Panel wyników

4.1. Założenia

Podstawowym założeniem panelu jest minimalizm.

Uznajemy za najważniejsze czytelne i szybkie przekazywanie informacji, które umożliwi zapoznanie się z informacjami w intuicyjny sposób.

Panel przekazuje przede wszystkim te informacje, które są najbardziej pożądane, w drugiej kolejności - dane bardziej szczegółowe.

Drugim ważnym aspektem była dobra wizualizacja, a więc kontrast.

Kolory dobrano w taki sposób, by żółto-pomarańczowe wykresy oraz statystyki dobrze wyróżniały się na czarnym, stonowanym tle.

Również mapa została utrzymana w jasnych barwach.

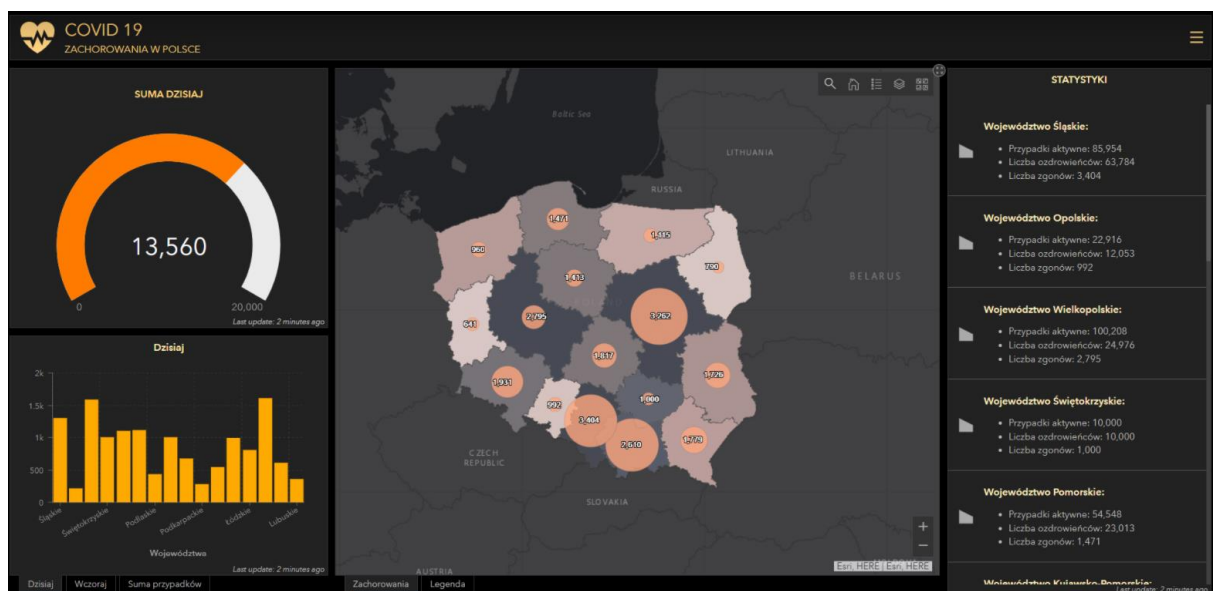
4.2. Funkcjonalności

Najważniejszym elementem panelu jest mapa. Umieszczona została w środkowej części i przedstawia zachorowania od początku pandemii w poszczególnych województwach. Po kliknięciu w konkretne województwo, pojawiają się szczegółowe dane.

Po lewej stronie znajdują się dwa elementy - licznik dzisiejszych przypadków w całej Polsce oraz panel z wykresami. Licznik umożliwia szybkie pozyskanie informacji o tym, jaki jest aktualnie stan zakażeń, wykresy pokazują stan w poszczególnych województwach dzisiaj, wczoraj i sumę od początku.

Po prawej stronie dostępne są bardziej szczegółowe statystyki - możemy odczytać liczbę aktualnych przypadków aktywnych, osób wyleczonych oraz sumaryczną liczbę zgonów.

DASHBOARD:



Dashboard składa się z następujących elementów:

- **Mapa** - bazuje na warstwach powstałych z połączenia tabel (zachorowania, testy, hospitalizacja) z plikiem, zawierającym geometrię województw. Sumy zachorowań i wyzdrowień są prezentowane przez skalę barw oraz przez symbole w kształcie kół, zawierające konkretne wartości liczbowe dla zgonów. Warstwa domyślna pokazuje liczbę zachorowań w województwie. Dane wyświetlane na mapie można zmieniać.
- **Legenda** - zmienia się w zależności od wyświetlanej warstwy.
- **Licznik zachorowań** - pokazuje ogólną sumę zachorowań w Polsce.
- **Wykresy słupkowe** - wizualizują ogólną sumę zachorowań dzisiaj, wczoraj i od początku pandemii.

- **Statystyki szczegółowe** - zawierają bardziej szczegółowe informacje z tych tabel, które nie są wyświetlane na mapie. Po kliknięciu elementu listy reprezentującego dane województwo, zostaje ono podświetlone na mapie.

5. Podsumowanie

Podczas pracy z projektem powstało kilka problemów.

Wiele z nich udało się rozwiązać, jednak niektóre pozostały bez odpowiedzi, a miliony pytań, które zrodziły się w naszej głowie, musi nadal podróżować przez otchłanie wszechświata.

5.1. Problemy - dane

Pierwszym problemem okazało się odświeżanie danych za pomocą programu OverWriterFS. Po połączeniu z kontem ArcGIS Online oraz arkuszem zawierającym odświeżającą się tabelę, przy próbie połączenia z samą warstwą, program wyświetlał komunikat o tym, że warstwa nie istnieje. Po kilku próbach naprawienia problemu i analizie kodu, pomysł został porzucony.

Drugim problemem okazało się połączenie tabeli z testami oraz hospitalizacją z arkuszem "COVID-19 w Polsce". Było to spowodowane sposobem, w jaki dane są dodawane do oryginału. Co tydzień, do arkusza dodawane są informacje w postaci kolejnej kolumny, a te, które są potrzebne, zmieniają położenie. Z kolei funkcja IMPORTRANGE(), która została użyta do stworzenia tabeli z zachorowaniami, przyjmuje statyczne indeksy komórek arkusza. Po przejrzeniu dokumentacji Google nie odnaleziono funkcji, która poradziłaby sobie z tym problemem bez własnej ingerencji programistycznej. W przypadku drugiej tabelki sytuacja przedstawiała się podobnie.

5.2. Problemy - ArcGIS Online

Jednym z problemów nierozwiązywalnych okazało się udostępnienie odświeżonego panelu całej grupie. ArcGIS miewa problemy z hojnością, a gościnność także nie jest jego najlepszą stroną.

Powiedzmy sobie szczerze - to introwertyk, który miewa różne humory. I trzeba być wdzięcznym, że po długich dyskusjach, udostępnił możliwość odświeżania danych chociaż jednej osobie.

Warto wiedzieć także, że ArcGIS, jako młody i jeszcze nie tak doświadczony przez życie program, wciąż uczy się języka polskiego. I o ile prezentacja naszej mowy ojczystej w oficjalnym świetle jest dla niego osiągalna, o tyle przemiana od wewnątrz wciąż pozostaje poza horyzontem możliwości. Gdy gra toczy się o najdelikatniejsze sprawy, a więc zmienne, pozostaje nieufny. Trwając w przestrachu przed krętymi ogonkami i liniami, postanawia udawać, że nie dostrzega problemu. Niczym jaszczurka, zastygająca w bezruchu, gdy dostrzeże drapieżnika.

I nawet jeśli udało się namówić go na małe ugody - dostosować nazwy kolumn do jego międzynarodowego słownictwa, to i tak, rana pozostała zbyt głęboka, a polskie nazwy zmiennych niezmiennie (tu wyjątkowo, bo zwykle zmienne jednak się zmieniają) pozostawały na swoich miejscach. I znikwały, znikwały w przestworzach świeżego dashboardu.

Lista ze statystykami dotyczącymi testów nigdy zatem nie powstała. Możemy tylko gdybać, jak potoczyłyby się losy świata, gdyby jednak miała możliwość zaistnieć...