



ČESKÉ VYSOKÉ UČENÍ TECHNICKÉ V PRAZE

Fakulta stavební

**Kartografické zpracování vybraných aspektů epidemie
nemoci Covid-19**

**Cartographic processing of selected aspects of the
Covid-19 pandemy**

BAKALÁŘSKÁ práce

Shvets Kateryna

Studijní obor: Geodézie a kartografie

Vedoucí bakalářské práce: Tomáš Janata

Praha, 2023



ZADÁNÍ BAKALÁŘSKÉ PRÁCE

I. OSOBNÍ A STUDIJNÍ ÚDAJE

Příjmení: **Shvets** Jméno: **Kateryna** Osobní číslo: **493675**
Fakulta/ústav: **Fakulta stavební**
Zadávací katedra/ústav: **Katedra geomatiky**
Studijní program: **Geodézie a kartografie**

II. ÚDAJE K BAKALÁŘSKÉ PRÁCI

Název bakalářské práce:

Kartografické zpracování vybraných aspektů epidemie nemoci Covid-19

Název bakalářské práce anglicky:

Catographic processing of selected aspects of the Covid-19 pandemy

Pokyny pro vypracování:

Seznam doporučené literatury:

Jméno a pracoviště vedoucí(ho) bakalářské práce:

Ing. Tomáš Janata, Ph.D. Katedra geomatiky FSv

Jméno a pracoviště druhé(ho) vedoucí(ho) nebo konzultanta(ky) bakalářské práce:

Datum zadání bakalářské práce: **01.03.2023**

Termín odevzdání bakalářské práce: **05.06.2023**

Platnost zadání bakalářské práce: _____

Ing. Tomáš Janata, Ph.D.
podpis vedoucí(ho) práce

prof. Dr. Ing. Karel Pavelka
podpis vedoucí(ho) ústavu/katedry

prof. Ing. Jiří Máca, CSc.
podpis děkana(ky)

III. PŘEVZETÍ ZADÁNÍ

Studentka bere na vědomí, že je povinna vypracovat bakalářskou práci samostatně, bez cizí pomoci, s výjimkou poskytnutých konzultací. Seznam použité literatury, jiných pramenů a jmen konzultantů je třeba uvést v bakalářské práci.

Datum převzetí zadání

Podpis studentky

Prohlášení

Prohlašuji, že jsem bakalářskou práci na téma “Kartografické zpracování vybraných aspektů epidemie nemoci Covid-19” zpracoval(a) samostatně za použití uvedené literatury a pramenů.

Dále prohlašuji, že nemám závažný důvod proti užití tohoto školního díla ve smyslu § 60 zákona č. 121/2000 Sb., o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon).

V Praze dne 02.05.2023

.....

Shvets Kateryna

Poděkování

Děkuji Ing. Tomáši Janatovi, Ph.D za její/jeho odborné vedení a užitečné rady při zpracování bakalářské práce a také za vstřícnost a poskytnuté konzultace. Děkuji také svým přátelům za podporu v průběhu psaní práce.

Abstrakt

Bakalářská práce se věnuje vizualizaci vybraných aspektů COVID-19 pomocí metod časového mapování. Konečným produktem bude 9 animací zobrazujících vývoj aspektů COVID-19 v čase ve střední Evropě, Africe a jedna pro celkový počet úmrtí na milion obyvatel na celém světě pro 10 země s nejvyššími hodnotami.

Klíčová slova:

COVID-19

Temporal Mapping

Spatial Mapping

Abstract

The bachelor's thesis is dedicated to visualising chosen aspects of COVID-19 with temporal mapping methods. The final product will be 9 animations displaying developing COVID-19 aspects over time in central Europe, Africa and one for total deaths per million worldwide for countries with top-10 highest values.

Key words:

COVID-19

Časové mapování

Prostorové mapování

Summary

	6
Introduction	8
1 History and development of COVID-19	9
2 Research	11
Best examples of COVID-19 maps and dashboards	11
Types of Maps used the most for mapping COVID-19	15
Mapping Methodologies mostly used for mapping COVID-19	16
3 Description a preparing the data	16
Source of Data	16
Data Coverage and Availability	17
Key Variables and Attributes	18
Data Preprocessing and Cleaning	24
4 ArcGIS Temporal Mapping techniques	26
Justification for Using ArcGIS Pro	26
Main Characteristics of ArcGIS Pro	27
Temporal Data Visualisation Methods in ArcGIS Pro	28
5 Visualising data with time slider and creating an animation	29
Preparing my data and importing to ArcGIS	29
Converting the table for joining	30
Joining table to the layer	31
Enabling time on the layer and symbolise the data	31
Creating an animation	32
6 Discussion	35
Conclusion	36
Literature and sources	37

Introduction

I chose this topic because of a few reasons. First is practical, I was interested in cartography and geomatics, and I see myself being involved in this in the future. Also, the opportunity to analyse, clean, arrange the big source of data and the ability to visualise it afterwards is a great skill to gain, which is possible during work on this project.

According to my research which I will describe further, most of the maps that were created visualising the spreading of the disease, total cases or deaths are basic, not updated, or professionally made. My idea was to prove that ArcGIS functionality, especially temporal mapping methods are good enough for creating

The final product of this project will be few animations, visualising certain COVID-19 aspects in time in central Europe, Africa, and

Also, this work will show how big-scale events like COVID-19 with this large amount of data can be visualised with methods I used, so these methods can be or be not considered for further possible projects. My personal interest in this was to test a few temporal mapping methods to visualise some of COVID-19 aspects.

By pursuing these research objectives, this thesis aims to advance our understanding of the COVID-19 pandemic's global dynamics and contribute to the understanding of cartographic visualisation methods and their limits in health crises like COVID-19. From professional side the aim is to test how close you can get with having ArcGIS and basic tools to create a project for

1 History and development of COVID-19

The emergence of COVID-19, caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), has brought about an unprecedented global health crisis. This article aims to shed light on the nature and development of COVID-19, exploring its origins, transmission, impact, and the scientific efforts undertaken to combat the pandemic.

The spread of COVID-19 across the world can be attributed to various factors, including globalisation, international travel, and the highly contagious nature of the virus. Here is an overview of how COVID-19 spread globally:

Initial Outbreak in Wuhan, China:

The first cases of COVID-19 were reported in December 2019 in Wuhan, Hubei Province, China. The virus likely originated from an animal source, possibly in a seafood market. The initial spread was primarily localised within Wuhan and surrounding regions.

- International Travel:

As Wuhan is a major transportation hub, the virus quickly spread beyond China's borders through international travel. Infected individuals unknowingly carried the virus to other countries, allowing it to gain a foothold in new locations. During the first days of COVID-19 pandemic the very first cases were confirmed afterwards in France and Italy.

- Community Transmission:

Once the virus reached other countries, it began spreading within local communities through close contact with infected individuals. Transmission occurred through respiratory droplets when people were in close proximity, such as in households, workplaces, or social gatherings.

- Global Travel and Tourism:

The interconnectedness of the modern world facilitated the rapid international spread of COVID-19. Air travel, in particular, played a significant role in carrying the virus across continents. Infected individuals could unknowingly transmit the virus during travel, leading to new outbreaks in previously unaffected regions.

- Asymptomatic and Pre-symptomatic Spread:

One challenging aspect of COVID-19 is that infected people can spread the virus even before experiencing symptoms or while having none. This feature allowed the virus to go undetected and contributed to its rapid transmission among people and all around the world.

- Local Outbreaks:

Certain events and settings, such as crowded gatherings, religious ceremonies, conferences, or social events, provided opportunities for large-scale transmission. These "superspreading events" facilitated the rapid spread of the virus within specific communities and sometimes resulted in subsequent outbreaks. Also, not following all instructions and ignoring basic safety rules caused hard detections and control over the situation.

- Border Controls and Containment Measures:

Countries implemented various measures to control the spread of the virus, such as travel restrictions, quarantine protocols, and social distancing measures. While these measures helped slow down the spread, community transmission and localised outbreaks still occurred.

- Global Pandemic Declaration:

In March 2020, the World Health Organization (WHO) declared COVID-19 a global pandemic due to its widespread and sustained transmission across multiple countries and continents. This declaration highlighted the urgent need for coordinated international efforts to combat the virus.

- Variants and Further Spread:

Over time, different variants of the SARS-CoV-2 virus emerged, some of which showed increased transmissibility. These variants, such as the Alpha, Beta, Gamma, and Delta variants, contributed to the continued spread of COVID-19 and presented additional challenges in controlling the pandemic.

- Vaccination Efforts:

The development and distribution of COVID-19 vaccines have played a crucial role in curbing the spread of the virus. Vaccination campaigns have been implemented globally, prioritising high-risk populations and aiming to achieve widespread immunity to limit further transmission.

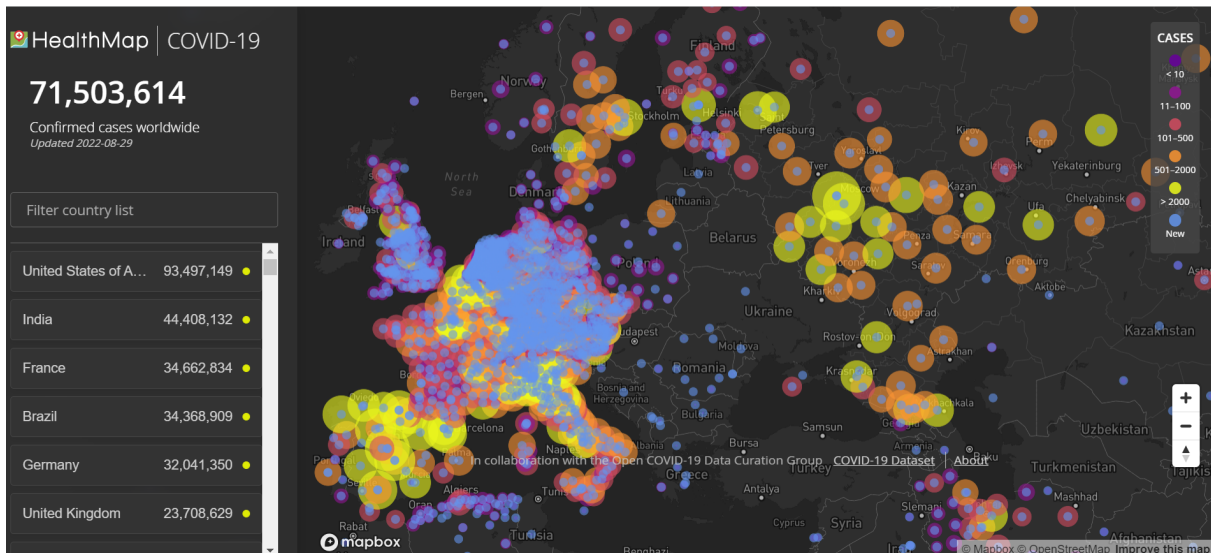
It's important to note that the specific trajectory of COVID-19 spread varied from country to country based on factors such as population density, healthcare capacity, public health response, and adherence to preventive measures. Monitoring and responding to the evolving situation have been vital in containing and mitigating the impact of the pandemic.

2 Research

Best examples of COVID-19 maps and dashboards

The COVID-19 pandemic, first declared by the World Health Organization in March 2020, triggered a global scramble for knowledge, understanding, and solutions. In the midst of this crisis, cartography emerged as an instrumental tool in visually communicating the patterns and progression of the disease. This chapter provides an in-depth analysis of various map types used during the pandemic, explores diverse mapping methodologies, their advantages and drawbacks, and elucidates each type and method with relevant examples.

Generally after research it is possible to see that most maps created for one country or dashboard for real-time displaying are not updated for years.

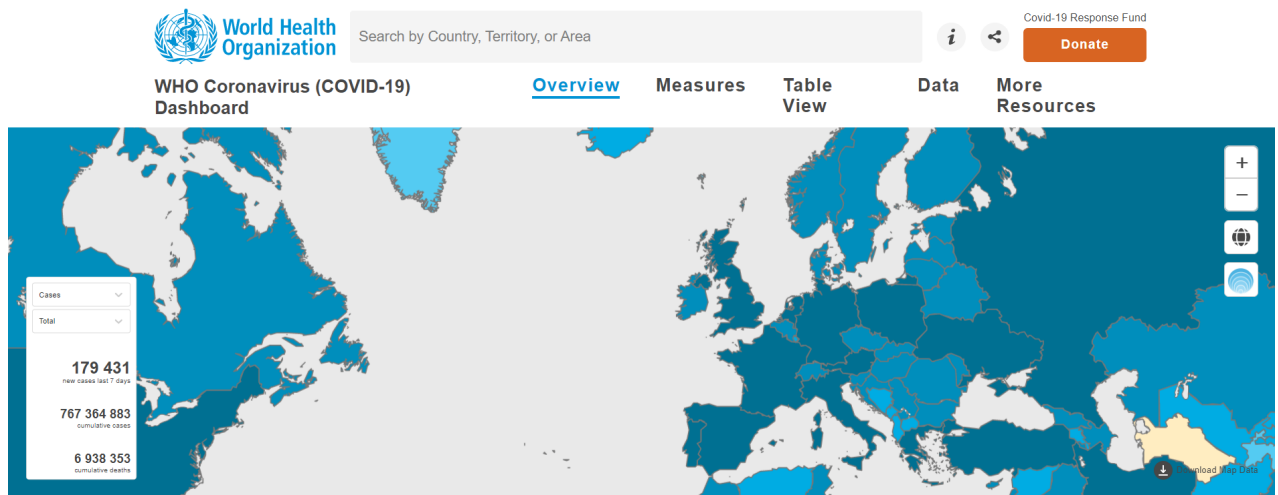


Img. 1 Confirmed cases worldwide
<https://www.healthmap.org/covid-19/>

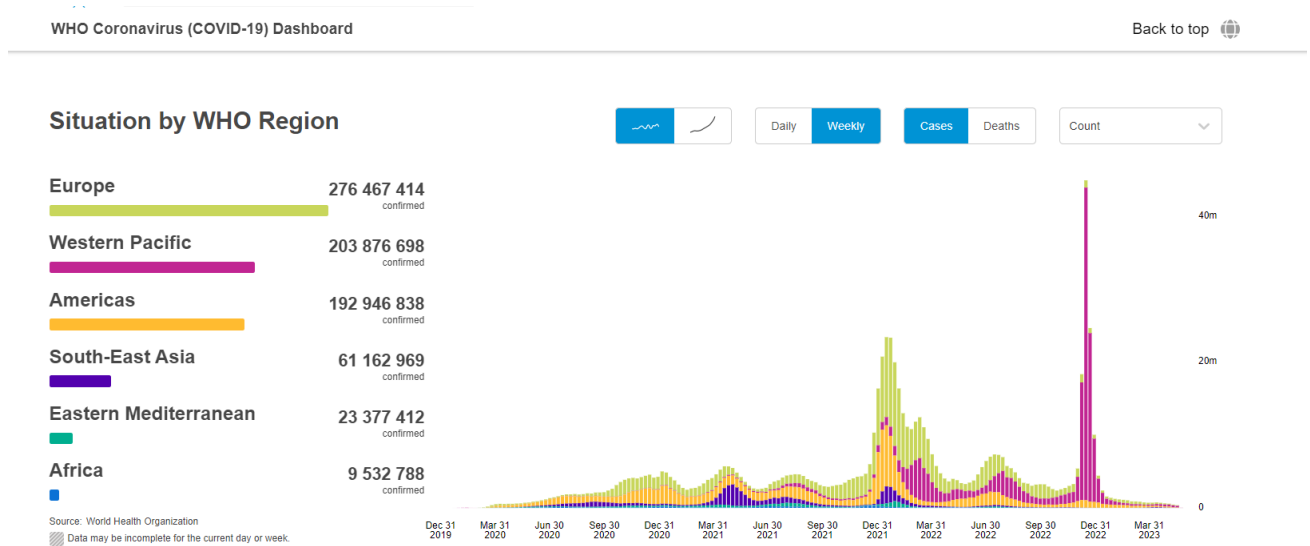
After researching I can highlight as the most complete one a dashboard provided by the World Health Organisation. Dashboard is daily updated, displaying 3 variables: cases, deaths and vaccinations. Furthermore, it allows you to select different regions (e.g., Africa, Americas, Eastern Mediterranean, Europe, South-East Asia, Western Pacific) for more localised statistics.

The centrepiece of the dashboard is a global geospatial map showing the distribution of COVID-19 cases. The map uses a choropleth design, with countries coloured based on the total number of confirmed cases. Darker shades of the colour represent higher case counts. Users can interact with the map by clicking on a country to see more specific data.

Also there are a series of graphs and charts that display trends over time, such as the global curve of cases and deaths, vaccination progress, and weekly epidemiological update. These charts use line graphs and bar graphs to represent the data in an easily digestible form.



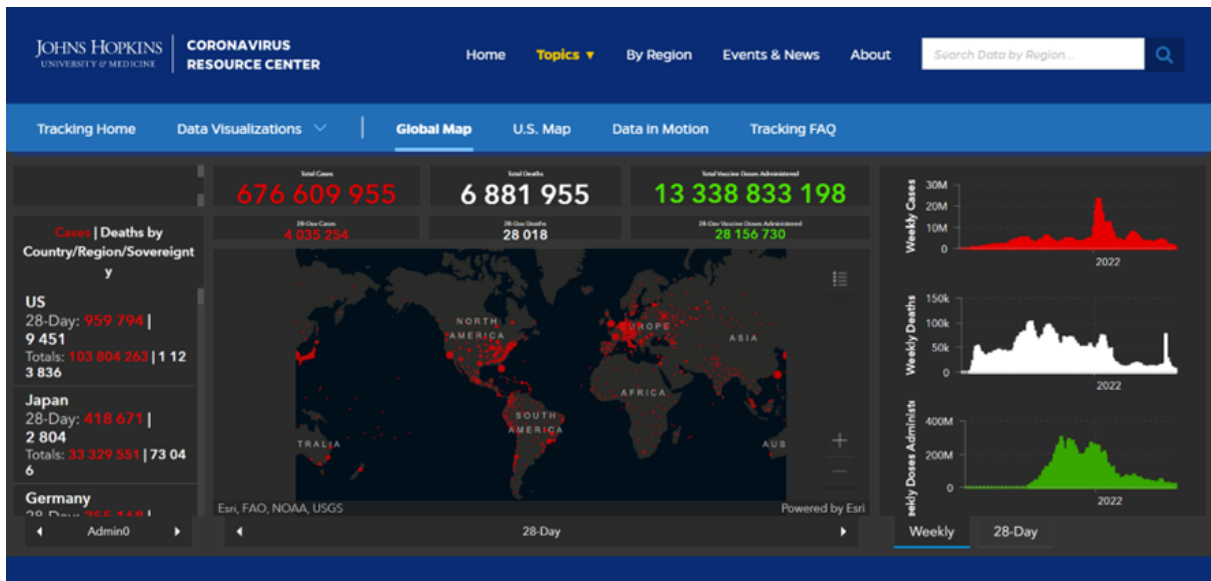
Img. 2



Img. 3

Another visually appealing and complete dashboard is one produced by The Johns Hopkins University (JHU). There are many COVID-19 related dashboards, but most of them are for the USA.

The central focus of the dashboard is a global map illustrating the COVID-19 cases worldwide. The dashboard uses proportional symbology based on the number of confirmed COVID-19 cases, with bigger symbols indicating a higher case count.



Img. 4

<https://coronavirus.jhu.edu/map.html>

The best was created by ESRI team - Mapping the spread of COVID-19 in Story Maps. It displays the story of COVID-19 with maps attached to every logical step – for example, map of the most popular flights from Wuhan:



Img. 5

<https://storymaps.arcgis.com/stories/4fdc0d03d3a34aa485de1fb0d2650ee0>

These projects were the inspiration for me to try to combine my skills and ArcGIS to create a visually appealing map of COVID-19 in the world.

Types of Maps used the most for mapping COVID-19

Choropleth Maps: Among the most used during the pandemic, choropleth maps present divided geographical areas coloured or patterned in relation to a data variable. This technique is especially effective in illustrating variable quantities, like COVID-19 infection rates across countries or states. Each region's colour intensity reflects the magnitude of the variable.

Heat Maps: Heat maps convey the intensity of COVID-19 cases through colour variations. Particularly useful for illustrating the concentration of cases in densely populated areas, these maps excel in spotlighting COVID-19 hotspots. For instance, the New York Times frequently used heat maps to depict the intensity of infection rates across the United States. Nevertheless, heat maps may sometimes be misinterpreted as representing geographic area rather than density or quantity of cases.

Dot Distribution Maps: These maps employ dots to signify the presence of COVID-19, where each dot represents a specific number of occurrences, such as cases, hospitalizations, or deaths. One powerful implementation was the map developed by University of Wisconsin-Madison, which used dots to display the density of COVID-19 cases in Wisconsin counties. These maps are, however, less effective in areas with high case numbers, where dots can overlap and obscure each other

Cartograms: Cartograms distort the geometry or space of the map to convey information, often showing a variable such as population size in relation to COVID-19 case numbers. For instance, Worldmapper published a series of cartograms demonstrating the global distribution of confirmed COVID-19 cases, adjusted for population size. Although providing valuable perspectives, cartograms can be

challenging to understand and hard to process visually when it comes to big amount of data.

Mapping Methodologies mostly used for mapping COVID-19

- Geospatial Mapping: Using geographic information systems (GIS), geospatial mapping visualises data related to physical locations. During the pandemic, this methodology was crucial for tracking the geographic and temporal spread of the virus. For instance, Esri's COVID-19 GIS Hub used geospatial mapping to track cases worldwide. However, the effectiveness of geospatial mapping is dependent on the quality and accuracy of the geographical data used, which can be a limitation.
- Data Aggregation: COVID-19 data was often collected at the local level before being aggregated to represent larger geographic areas (like counties, states, or countries). Aggregation allows for clearer, broader trends to be observed and understood. A notable example is the COVID-19 Data Repository by the Center for Systems Science and Engineering at Johns Hopkins University. Despite its usefulness, aggregated data can sometimes mask crucial details and anomalies occurring at a smaller scale.
- Real-Time Tracking: A cornerstone of COVID-19 mapping, real-time tracking required continuous data updates and immediate visualisation. Websites like "ncov2019.live," offered real-time tracking of cases, recoveries, and deaths globally. Real-time tracking, while highly beneficial for instant updates, is dependent on rapid, accurate data reporting, which can be challenging during a pandemic.

3 Description a preparing the data

Source of Data

The dataset used in this thesis is sourced from Our World in Data, a reputable and widely recognized organisation focused on global development data. Specifically, the dataset is obtained from their GitHub repository

(<https://github.com/owid/covid-19-data/tree/master/public/data>).

Our World in Data actively collects, verifies, and updates COVID-19 data from multiple sources, ensuring reliability and accuracy in their dataset.

Data Coverage and Availability

The dataset used in this thesis provides a comprehensive coverage of countries and regions worldwide, offering a global perspective on the COVID-19 pandemic. It includes data from various continents, countries of different sizes, populations, and levels of development. This broad coverage allows for a comprehensive analysis of the pandemic's global impact and facilitates cross-country comparisons.

While the dataset strives to include data from as many regions as possible, it is important to acknowledge that data availability and completeness can vary across countries. Some countries may have robust data reporting systems, providing detailed and timely information on COVID-19 cases, deaths, testing, and vaccinations. In contrast, other regions may face challenges in data collection and reporting, leading to potential data gaps or inconsistencies.

The availability of data can also be influenced by factors such as governmental policies, healthcare infrastructure, and resources allocated to data collection and reporting. Some countries may have established robust surveillance systems and well-defined reporting mechanisms, resulting in more comprehensive and reliable data. On the other hand, countries facing resource constraints or other operational challenges may experience limitations in data availability and quality.

To address these challenges, Our World in Data takes significant efforts to gather data from various sources, including official government reports, public health agencies, international organisations, and research institutions. The organisation aims to ensure the highest possible data quality and reliability by verifying and cross-referencing information from multiple sources. However, it is important to note that data discrepancies or variations in reporting standards may still occur.

In instances where data may be limited or missing for certain countries or time periods, it is crucial to interpret and analyse the dataset with caution. Researchers

should consider the potential impact of data gaps on the overall analysis and conclusions.

Furthermore, as the COVID-19 pandemic evolves and new data becomes available, it is advisable to periodically update the dataset to ensure the inclusion of the most recent information. Regular data updates will enhance the accuracy and reliability of the analysis and visualisation, enabling a more current understanding of the pandemic's dynamics across different regions.

Despite potential variations in data coverage and availability, the dataset sourced from Our World in Data's repository provides a solid foundation for examining and visualising the global COVID-19 pandemic.

Key Variables and Attributes

The Our World in Data (OWID) GitHub repository for COVID-19 data contains multiple datasets. Here, we'll be focusing on the primary `owid-covid-data.csv` file, which is a country-level dataset updated daily and includes various indicators related to the pandemic.

Below is a brief description of each variable in the dataset:

Variable	Description
<code>iso_code</code>	ISO 3166-1 alpha-3 - three-letter country codes.
<code>continent</code>	Geographical continent of the location.
<code>location</code>	Geographical location.

date	Date of observation.
total_cases	Total confirmed cases of COVID-19.
new_cases	New confirmed cases of COVID-19.
new_cases_smoothed	7-day smoothed average of new confirmed cases of COVID-19.
total_deaths	Total deaths attributed to COVID-19.
new_deaths	New deaths attributed to COVID-19.
new_deaths_smoothed	7-day smoothed average of new deaths attributed to COVID-19.
total_cases_per_million	Total confirmed cases of COVID-19 per 1,000,000 people.
new_cases_per_million	New confirmed cases of COVID-19 per 1,000,000 people.

<code>new_cases_smoothed_per_million</code>	7-day smoothed average of new confirmed cases of COVID-19 per 1,000,000 people.
<code>total_deaths_per_million</code>	Total deaths attributed to COVID-19 per 1,000,000 people.
<code>new_deaths_per_million</code>	New deaths attributed to COVID-19 per 1,000,000 people.
<code>new_deaths_smoothed_per_million</code>	7-day smoothed average of new deaths attributed to COVID-19 per 1,000,000 people.
<code>reproduction_rate</code>	Real-time estimate of the effective reproduction rate (R) of COVID-19.
<code>icu_patients</code>	Number of COVID-19 patients in intensive care units (ICUs) on a given day.
<code>icu_patients_per_million</code>	Number of COVID-19 patients in intensive care units (ICUs) per 1,000,000 people on a given day.
<code>hosp_patients</code>	Number of COVID-19 patients in hospital on a given day.

hosp_patients_per_million	Number of COVID-19 patients in hospital per 1,000,000 people on a given day.
weekly_icu_admissions	Number of COVID-19 patients newly admitted to intensive care units (ICUs) in a given week.
weekly_icu_admissions_per_million	Number of COVID-19 patients newly admitted to intensive care units (ICUs) in a given week per 1,000,000 people.
weekly_hosp_admissions	Number of COVID-19 patients newly admitted to hospitals in a given week.
weekly_hosp_admissions_per_million	Number of COVID-19 patients newly admitted to hospitals in a given week per 1,000,000 people.
new_tests	New tests for COVID-19.
total_tests	Total tests for COVID-19.
total_tests_per_thousand	Total tests for COVID-19 per 1,000 people.

<code>new_tests_per_thousand</code>	New tests for COVID-19 per 1,000 people.
<code>new_tests_smoothed</code>	7-day smoothed average of new tests for COVID-19.
<code>new_tests_smoothed_per_thousand</code>	7-day smoothed average of new tests for COVID-19 per 1,000 people.
<code>positive_rate</code>	The share of COVID-19 tests that are positive, given as a rolling 7-day average.
<code>tests_per_case</code>	Tests conducted per new confirmed case of COVID-19, given as a rolling 7-day average.
<code>tests_units</code>	Units used by the location to report its testing data.
<code>total_vaccinations</code>	Total number of COVID-19 vaccination doses administered.
<code>people_vaccinated</code>	Total number of people who received at least one vaccine dose.

people_fully_vaccinated	Total number of people who received all doses prescribed by the vaccination protocol.
total_boosters	Total number of COVID-19 vaccination booster doses administered.
new_vaccinations	New COVID-19 vaccination doses administered (only calculated for consecutive days).
new_vaccinations_smoothed	7-day smoothed average of new COVID-19 vaccination doses administered.
total_vaccinations_per_hundred	Total number of COVID-19 vaccination doses administered per 100 people in the total population.
people_vaccinated_per_hundred	Total number of people who received at least one vaccine dose per 100 people in the total population.
people_fully_vaccinated_per_hundred	Total number of people who received all doses prescribed by the vaccination protocol per 100 people in the total population.

total_boosters_per_hundred	Total number of COVID-19 vaccination booster doses administered per 100 people in the total population.
new_vaccinations_smoothed_per_million	7-day smoothed average of new COVID-19 vaccination doses administered per 1,000,000 people in the total population.

Tab 1

Manual processing

Data Preprocessing and Cleaning

In this thesis, the data preprocessing and cleaning procedures involved transforming the downloaded dataset, initially provided in an XML file format, into a format compatible with ArcGIS for visualisation and analysis. This process required manual cleaning and transformation steps to ensure data consistency and integrity.

Upon downloading the dataset, it consisted of data provided for every single date, which can be quite granular and may not be directly suitable for visualisation in ArcGIS.

In order to fix this, the first step was to choose which variables I keep for my work and which timing I want to have in my thesis. So my my dataset I chose to have next variables: iso_code

continent

location

date

total_cases

total_deaths

total_cases_per_million

total_deaths_per_million

reproduction_rate

It was decided after reviewing the whole dataset and finding out that most of the variables have no value, since many countries haven't reported it. Also, for my type of work it is enough to have just a few main variables.

Next step was to fix the time range and time steps for my dataset. I've got a recommendation that the optimal step is a week for each variable, so I needed to clean the data in some way. There were few attempts to do so: by Pivot tables, by using basic Excel tools and Google Sheets. So my solution was to find a tool which helps to choose every n-th row of a file, so i did the next steps: for each country separately i was choosing every 7th row of data, and while doing that checking that for each country I chose data rows for the same dates. Then it looked like this:

	A	B	C	D	E	F	G
1	iso_code	continent	location	date	total_cases	total_deaths	total_cases_per_million
2	AFG	Asia	Afghanistan	2020-01-09	0		
3	AFG	Asia	Afghanistan	2020-01-16	0		
4	AFG	Asia	Afghanistan	2020-01-23	0		
5	AFG	Asia	Afghanistan	2020-01-30	0		
6	AFG	Asia	Afghanistan	2020-02-06	0		
7	AFG	Asia	Afghanistan	2020-02-13	0		
8	AFG	Asia	Afghanistan	2020-02-20	0		
9	AFG	Asia	Afghanistan	2020-02-27	5		0,122
10	AFG	Asia	Afghanistan	2020-03-05	5		0,122
11	AFG	Asia	Afghanistan	2020-03-12	11		0,267
12	AFG	Asia	Afghanistan	2020-03-19	24		0,584
13	AFG	Asia	Afghanistan	2020-03-26	80	2	1,945
14	AFG	Asia	Afghanistan	2020-04-02	235	4	5,714
15	AFG	Asia	Afghanistan	2020-04-09	444	15	10,795
16	AFG	Asia	Afghanistan	2020-04-16	794	29	19,305

Img 6

Manual processing

By converting the data to a weekly format, it became more suitable for analysis and visualisation in ArcGIS, allowing for a more comprehensive exploration of the spatial and temporal patterns of COVID-19. This approach also facilitated smoother data handling and reduced complexity.

During the manual cleaning process, intermediate results could be observed and evaluated. This gave the opportunity to detect any anomalies or mistakes within the dataset. Manual control over the data transformation process allowed for the identification and rectification of any potential errors or inconsistencies that may have arisen during the cleaning and aggregation steps.

Next step was the data transforming and another Excel plugin was used to transform the table into a format compatible with ArcGIS. I created 2 datasets in different formats for opportunity to try both, and to be sure. This conversion involved adjusting the table structure, formatting the columns, and ensuring the appropriate data types for each attribute. By transforming the table into a compatible format, it became readily compatible with ArcGIS. Also it was done manually to control mid-results.

	A	B	C	D	E	F	G	H	I	J
1		tot_cas1	tot_cas2	tot_cas3	tot_cas4	tot_cas5	tot_cas6	tot_cas7	tot_cas8	tot_cas9
2	Afganistan	0	0	0	0	0	0	0	5	5
3	Africa							5	6	40
4	Albania									
5	Algeria								1	17
6	American Samoa									
7	Andorra									1
8	Angola									
9	Anguilla									
10	Antigua and Barbuda									
11	Argentina									10
12	Armenia									1
13	Aruba									
14	Asia	4	51	581	7808	28218	60091	75035	80901	90741
15	Australia				7	14	15	15	23	43
16	Austria								1	23
17	Azerbaijan									3
18	Bahamas									
19	Bahrain								37	53
20	Bangladesh									

Img 7

Manual processing

Both datasets were tested in ArcGIS and as the result the first one was needed for the type of work that was done, even though the second one was possible to be used too, but with using some ArcGIS tools like “Transport the table” which didn’t go well for me.

4 ArcGIS Temporal Mapping techniques

Justification for Using ArcGIS Pro

The choice of using ArcGIS Pro for visualising temporal data in this research hinges on several factors. Firstly, ArcGIS Pro offers an intricate yet user-friendly Geographic

Information System (GIS) interface that enables the efficient handling of spatial data. In terms of COVID-19 data, this can include geolocational attributes such as countries, cities, or even specific healthcare facilities.

Secondly, ArcGIS Pro is renowned for its temporal data capabilities. The platform allows for the creation of dynamic visualisations which can elegantly display how the COVID-19 pandemic has evolved over time. In a situation as fluid and rapidly changing as a global pandemic, this temporal component is crucial in grasping the full scope of the situation.

Finally, the software's interoperability makes it a versatile tool for any researcher. ArcGIS Pro can effectively communicate with various databases and data formats, enabling it to integrate seamlessly into diverse research methodologies.

Main Characteristics of ArcGIS Pro

ArcGIS Pro is a leading desktop GIS application, providing robust features that support data visualisation, analysis, and management. Its main characteristics include:

- 3D capabilities: ArcGIS Pro supports data visualisation in both 2D and 3D, offering users the ability to create multi-dimensional models and analyses. This is particularly beneficial when exploring complex phenomena like a global pandemic.
- Spatial and temporal analysis tools: This platform offers advanced analytical tools for both spatial and temporal data, allowing researchers to examine patterns and trends over time and space.
- Project-based workflow: In ArcGIS Pro, all resources are bundled into a project, which includes maps, layouts, data, and tools, making it easier to manage and organise resources effectively.
- Interoperability: It supports a variety of data formats, enabling users to work with different types of data, such as CSV, JSON, GeoJSON, etc., which are commonly used for storing COVID-19 data.
- Advanced symbology and representation: ArcGIS Pro has advanced symbology options, enabling the display of data in a more understandable and interpretable manner.

Temporal Data Visualisation Methods in ArcGIS Pro

The global dissemination of COVID-19 has necessitated a comprehensive understanding of how the virus evolves spatially and temporally. By harnessing the temporal data visualisation methods available in ArcGIS Pro, we can accurately chart the course of the pandemic over time, unravel complex patterns, and deliver actionable insights. These visualisations can help policymakers and health experts monitor the trajectory of the virus, discern trends, and accordingly plan targeted responses. Below are the primary methods of temporal data visualisation in ArcGIS Pro.

Introduction to Time-enabled Layers: One of the essential features of ArcGIS Pro for temporal data visualisation is the time-enabled layer. Time-enabled layers can manage and display geographic data at different moments. In the context of COVID-19, it means we can examine the same location at different periods, revealing how infection rates have changed over time.

Time Slider Tool: The Time Slider tool offers an intuitive interface for visualising and manipulating temporal data. By dragging the slider across a timeline, one can 'play' the evolution of the pandemic, revealing a cinematic portrayal of how the situation has unfolded. This dynamic temporal control enables the researchers to trace the proliferation of COVID-19 cases across the globe at varying speeds, facilitating in-depth temporal analysis.

Temporal Rendering: ArcGIS Pro provides options to apply temporal rendering to layers. This powerful feature allows users to represent the variation in intensity of an event over time. For instance, rendering COVID-19 data could involve colour-coding regions based on the severity of infections at any given point in time. This method would effectively illustrate peaks and troughs in the data, demonstrating the rise and fall in infection rates, and facilitating the identification of hotspots.

Creating Time Series Charts: Time Series Charts are an indispensable tool for revealing trends and patterns in data over specific periods. These charts offer a clear visual representation of data points indexed in time order, thus providing an overarching view of the pandemic's progression. For COVID-19, this could

encompass daily new cases, cumulative case counts, recoveries, or death rates over time.

Each of these temporal visualisation methods provides a unique lens through which to understand the evolving nature of the COVID-19 pandemic. The capacity to engage with the data in these different ways enhances our understanding of the virus's behaviour and allows us to plan and strategize effectively for the future.

5 Visualising data with time slider and creating an animation

For my work I decided to combine techniques of Temporal Mapping mentioned before, and key ones are: Time Slider Tool, Hot Spot Analysis for chosen locations and animated maps as the final result.

By incorporating the temporal dimension of your COVID-19 case data, you can visualise the progression of cases over time in an animated format.

Generally the whole process has a few steps, no matter which software will be used for this type of work:

- Preparing the data. This is the most important one. According to my experience and research, properly arranged data can make you avoid many mistakes while transforming it into some other software. Data should have time values with defined time steps. You have to pay attention to the size of the time step, the bigger the data file - more complications can occur.
- The raster layer should have the same location (or other values) names. When it comes to countries and territories it's important to compare names so it will be joined right.
- Join the temporal data file to your raster, do your calculations according to your needs, and pay attention

Here's how I created an animation map using ArcGIS:

Preparing my data and importing to ArcGIS

Ensure that your dataset includes the temporal component (e.g., week or date) and the corresponding total COVID-19 case values for each country.

For that purpose I used one of datasets I've prepared before:

	A	B	C	D	E	F	G
1	iso_code	continent	location	date	total_cases	total_deaths	total_cases_per_million
2	AFG	Asia	Afghanistan	2020-01-09	0		
3	AFG	Asia	Afghanistan	2020-01-16	0		
4	AFG	Asia	Afghanistan	2020-01-23	0		
5	AFG	Asia	Afghanistan	2020-01-30	0		
6	AFG	Asia	Afghanistan	2020-02-06	0		
7	AFG	Asia	Afghanistan	2020-02-13	0		
8	AFG	Asia	Afghanistan	2020-02-20	0		
9	AFG	Asia	Afghanistan	2020-02-27	5		0,122
10	AFG	Asia	Afghanistan	2020-03-05	5		0,122
11	AFG	Asia	Afghanistan	2020-03-12	11		0,267
12	AFG	Asia	Afghanistan	2020-03-19	24		0,584
13	AFG	Asia	Afghanistan	2020-03-26	80	2	1,945
14	AFG	Asia	Afghanistan	2020-04-02	235	4	5,714
15	AFG	Asia	Afghanistan	2020-04-09	444	15	10,795
16	AFG	Asia	Afghanistan	2020-04-16	794	29	19,305

Img 6 DataSet

Manual processing

Also I chose the world map dataset. I've tried a few, and the biggest issue was to find one that will include all independent territories also, so I can join my tables after.

The dataset that I chose was one that was provided by ArcGIS official website:

<https://hub.arcgis.com/datasets/esri::world-countries-generalized/explore>

So I added my Word map shapefile as a geodatabase, and manually compared my excel file with geodatabase from ArcGIS to make sure that country names are the same and also in World map gdb I joined all separated territories to authority country for example: Madeira to Portugal, Svalbard to Norway and etc.

After these steps I imported my prepared Excel file.

Converting the table for joining

Here I needed to convert my time values from text format to date format, by the "Convert Time Field" tool. This will be safer for my next steps and will help to avoid possible problems after.

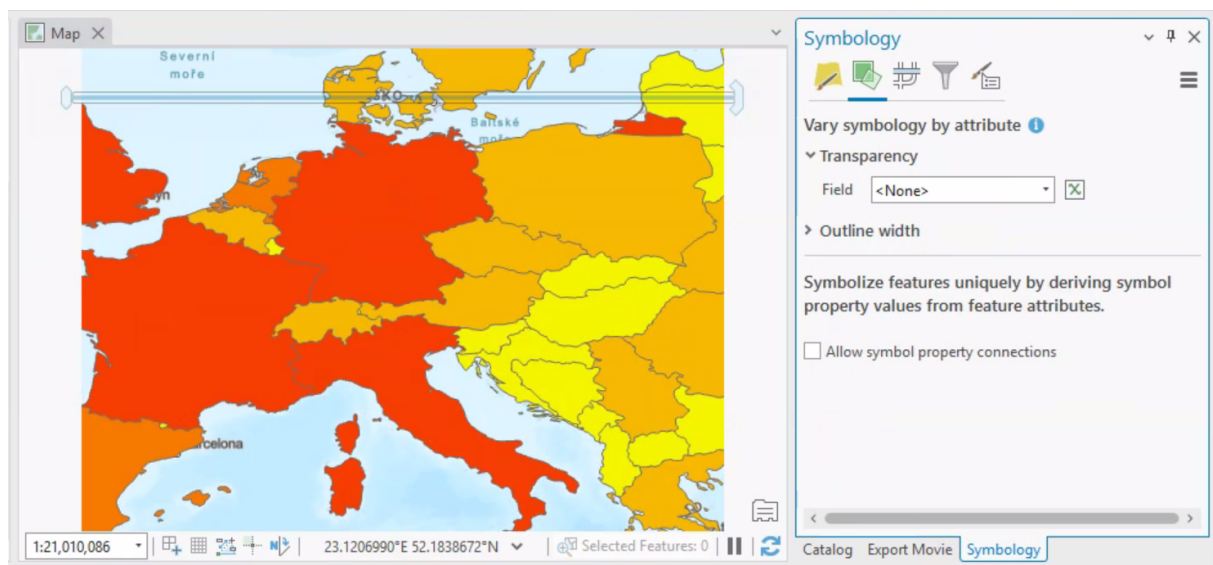
The conversion failed a few times, and the solution was to export my excel table to another one by "Excel to table" tool to make it editable.

Joining table to the layer

Here it was a basic Join tool, but the problem was the cardinality of the join: I needed one-to-many. It's an automatic process and I needed to join all my covid values for each date to certain polygons. There were few officially suggested methods - to join polygons to covid dataset (it's just wrong), to use Python method (very time-consuming) or to remove all joins (not suitable for my case). On some GIS forum I found an unobvious tip that helped - to have both datasets in one geodatabase, which made the join automatically one-to-many instead of one-to-one.

Enabling time on the layer and symbolise the data

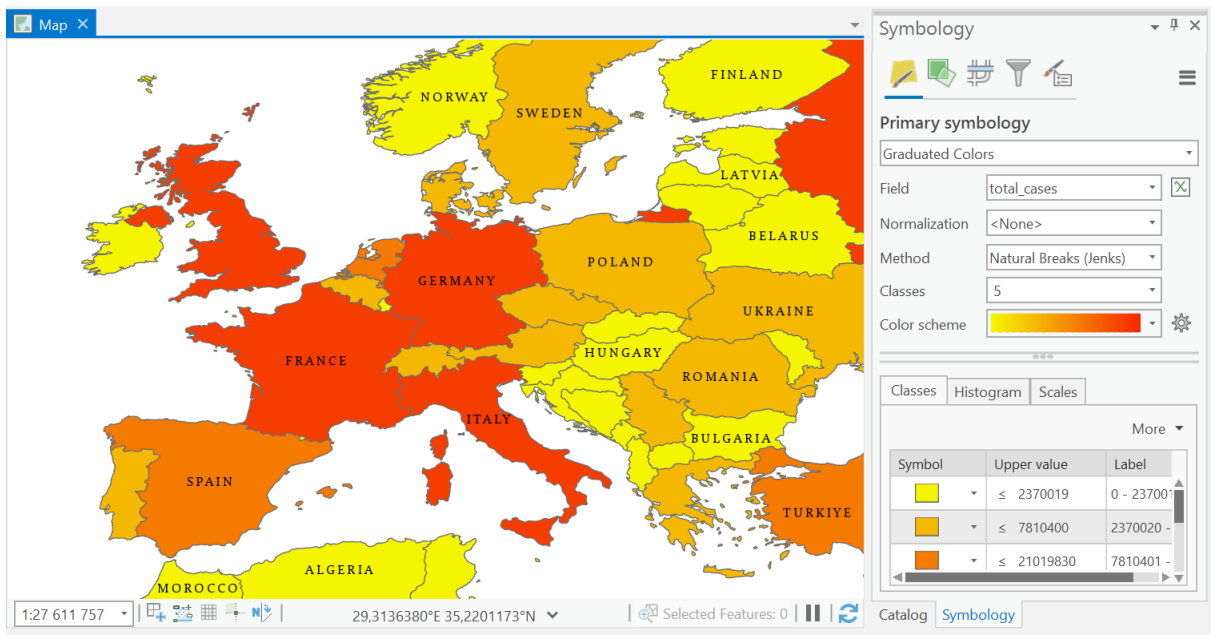
Symbolise the data: Choose an appropriate symbology to represent the total COVID-19 cases for each country. For this work, I chose graduated colours, the most intuitive was the yellow-red colour scheme. I chose the same one for all variables.



Img 8

Manual processing

After that I wanted to have country names on my map in English, so I needed to label my map. I wanted names to be only inside the polygons, so the amount of names are depending on zoom to the map.



Img 9

Manual processing

Then I enabled time for the layer by specifying the temporal field (dates).

Configure time settings: Adjust the time settings to define the time step 7 days (according to my data) and the desired speed of displaying.

The biggest problem was the speed of displaying the data on the timeline, so the best solution can be to do it locally.

In order to display data over time I decided to create an animation of COVID-19.

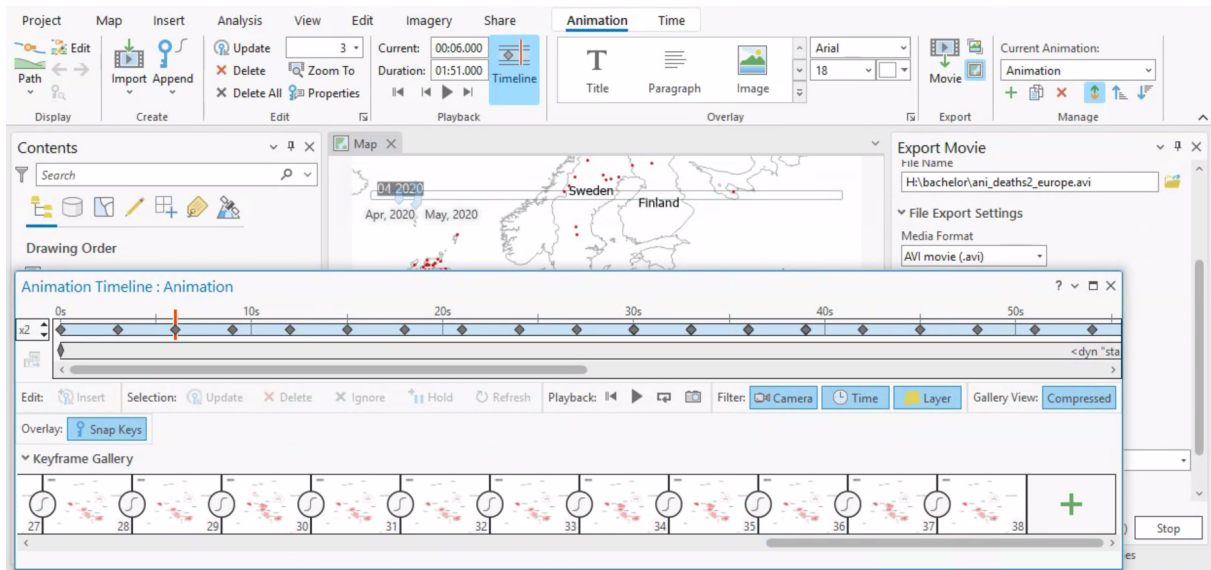
The problem of a big-data project is the drawing time of a map every time it's zoomed or when I play all time steps. So to avoid this I created an animation where I displayed COVID-19 cases over time in Europe.

First I selected and exported to a new layer the countries I was interested in. Then I enabled a time slider and visualised the features I was interested in (total cases, total deaths etc.)

Creating an animation

Next step is creating an animation. Newer ArcGIS version offers to create more animation types and for certain platforms like YouTube, but the basic ones are .avi, .gif, .mp3. I chose to create .avi videos that display the development of chosen features over time.

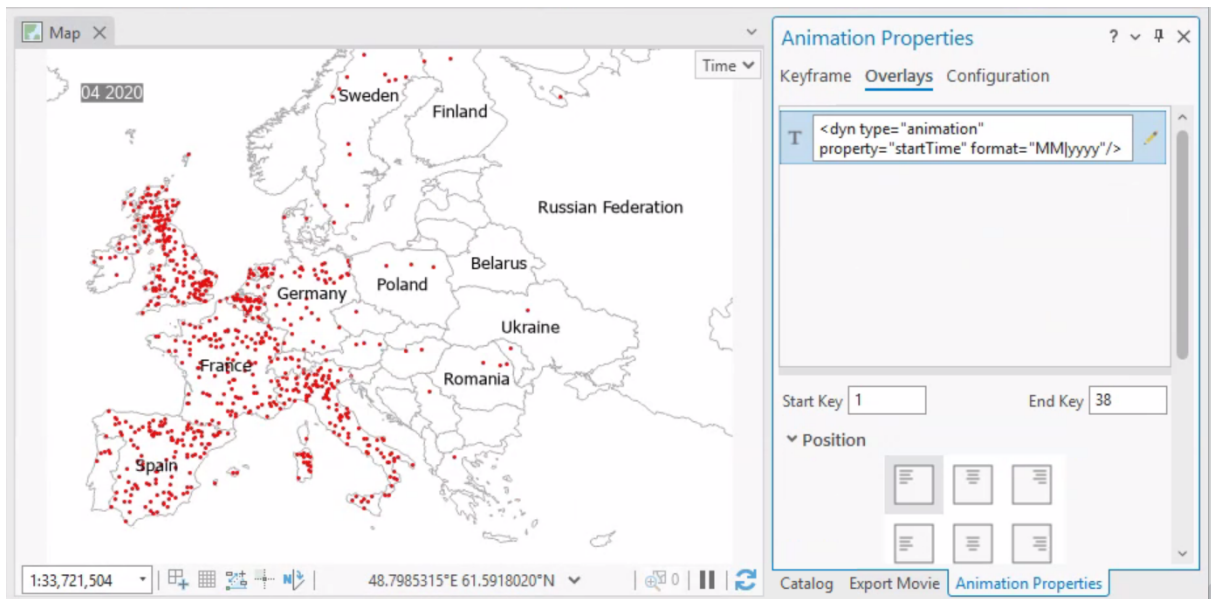
The animation is made manually by creating snaps from the map:



Img 10

Manual processing

I decided to make monthly snaps to keep the animation easier to create and for this case it is enough. Then the problem was to add a legend and a time slider to the animation, but ArcGIS documented some solutions for adding map time and dynamic texts to be displayed in a desired way.



Img 11

Manual processing

After creating a suitable snaps combination the animation can finally be exported in a desired format.

It was impossible for me to add a legend to an animation in ArcGIS, so I did it manually by adding a legend with the help of a video editor to every animation separately.

By following these steps, you can leverage ArcGIS's time-enabled capabilities to create an engaging and informative animation map that showcases the temporal dynamics of COVID-19 cases across countries.

The final animations could not be attached neither to the file nor to the whole work, so I created a One Drive folder with a free access:

<https://1drv.ms/f/s!AiJMCjdZoexOgQOmZOabkuwjqlq?e=waqLhN>

6 Discussion

For this type of work I considered more methods to try but some of them were just hard to use. For example Time Space Cube was a good idea for this project, but after a few attempts it was just displaying some mistakes that were just nearly impossible to fix.

Also during the work I had many troubles following instructions on how to use functions in ArcGIS provided by the official website because many steps were missing. So from my experience, forums and commentaries can provide more information than expected and also offer solutions that are not obvious.

Besides all that, I find temporal mapping very useful for these types of events like COVID-19.

Conclusion

The aim of this work was to test spatial temporal mapping methods to create a map of a project that visualises the big-scale event, and the perfect one is COVID-19.

The biggest problem was cleaning and preparing the data for the ArcGIS. Most ArcGIS spatial functions demand specific types of data in files, so for trying methods it was a need to create the optimal combination of tables, using join functions, etc.

As a final work, there were created animations that display the spreading of COVID-19, confirmed cases, total deaths, cases per million and deaths per million.

All mapping methods that were used in this work are simple and yet need a lot of work because of data preparation. While working on my thesis most of the time I've spent while looking up how to solve some function mistakes or how to make Join function work.

Using methods I used in my thesis can significantly help with improving healthcare development in cases like this. With simple ArcGIS tools it is possible to create trends and track some patterns to prevent them or to react.

All final animations I created are proof that ArcGIS is a powerful tool even for beginners to create projects from unprocessed data with a visually appealing and cartographically right outcome.

Literature and sources

Esri (2020). COVID-19 GIS Hub.

<https://coronavirus-resources.esri.com/>

Our World in Data (2020). Coronavirus Pandemic (COVID-19) – the data.

<https://ourworldindata.org/coronavirus-data>

World Health Organization (2020). WHO Coronavirus Disease (COVID-19) Dashboard.

<https://covid19.who.int/>

Johns Hopkins University & Medicine (2020). Coronavirus Resource Center.

<https://coronavirus.jhu.edu/map.html>

Centres of Disease Control and Prevention

<https://www.cdc.gov/coronavirus/2019-ncov/index.html>

National Institutes of Health

<https://covid19.nih.gov/>