

Map Visualization using Spatial and Spatio-Temporal Data: Application to COVID-19 Data

Mohammad A. Shaito

Department of Computer Science and Engineering, University of Texas at Arlington, Texas, USA mohammad.shaito@mavs.uta.edu

ABSTRACT

Currently, spatial geographic data can be collected for many applications that involve data on the planet earth. These collected data typically have coordinates (x, y), or longitude and latitude in map space, and thus can be located and displayed on maps. Data alone represents facts and has no meaning on its own but becomes meaningful when it is associated with application knowledge, such as elections, crimes, disease, etc. For example, there is no meaning behind those numbers (1, 23, 125, 355, ...), yet they are data that can get meaning when correlated with the total number of cases of COVID-19 in Texas per day starting on a certain date. Many devices can provide sequences of object location data over time (GPS in vehicles or mobile devices, etc.). However, no device can visualize or display them on its own without a visualization App. Both numeric and location data are raw data that need to be pre-processed and cleaned to become meaningful. Currently, collected data is a very valuable source of information which, after collection, can be processed, stored, analyzed, and visualized. In this paper, the available techniques of spatial data visualization will be overviewed. Moreover, a case study of COVID-19 spatio-temporal data visualization, using one of the techniques will be demonstrated. The COVID-19 data will be spatially visualized when data on a specific date is queried for analysis. On the other hand, spatio-temporal visualization will be displayed when a time series of COVID-19 data is queried for analysis.

KEYWORDS

Spatial Data, Spatio-Temporal Data, Visualization Techniques

ACM Reference Format:

Mohammad A. Shaito and Ramez Elmasri. 2021. Map Visualization using Spatial and Spatio-Temporal Data: Application to COVID-19 Data. In *The 14th PErvasive Technologies Related to Assistive Environments Conference (PETRA 2021), June 29–July 02, 2021, Corfu, Greece.* ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/3453892.3461336

PETRA 2021, June 29–July 02, 2021, Corfu, Greece

© 2021 Copyright held by the owner/author(s). Publication rights licensed to ACM. ACM ISBN 978-1-4503-8792-7/21/06...\$15.00 https://doi.org/10.1145/3453892.3461336 Ramez Elmasri

Department of Computer Science and Engineering, University of Texas at Arlington, Texas, USA elmasri@uta.edu

1 INTRODUCTION

Currently all matters pertaining to daily human life, like daily errands, weather conditions, medical systems, communication patterns, economics, social media, politics, scientific research, etc., are becoming highly technology dependent. Daily Collection of data from such human activities is an important source of information and is increasing dramatically. Data collection and the information that can be gained out of them is often recognized as the most valuable resource in the world, not oil [1]. For end-users to properly interact with and analyze such data, the best way is to have the collected data organized, processed, cleaned, analyzed, and visualized. Despite its importance, data remains facts with no meaning. Raw data like the numbers (1, 14, 100, 1500, etc.) must be examined and cleaned to obtain meaning. Once raw data is given a meaning it becomes useful information. For example, 100 m above sea level, here we know what the number 100 means, so now we can analyze the collected raw data and obtain new information out of it. In the real world, there exists a lot of geographic (spatial) data, which can be grouped based on features. This grouping results in different types of layers (customers, roads, buildings, elevation, land use, etc.). In fact, each layer will correspond to different types of objects.

For the end-user to have a correct explanation of the data he should be able to have an easy and clear communication with the data. This is best achieved through data visualization. Presenting data in tables or maps are examples of data visualization. Using maps will give the end-user more information than when using tables. Presenting collected spatial data graphically (visualizing it on a map) makes it easier to interpret and analyze and allows the user to identify areas of interest in the data to perform additional computation and further analysis.

In this paper, the presentation will mainly focus on spatial data visualization. There are several techniques that can be used to visualize spatial data. Six techniques that are commonly used to visualize spatial data will be discussed namely: Choropleth maps, Heat maps, Hexagonal binning, Dot maps, Bubble maps, and Cartogram maps.

1.1 Spatial Data

Spatial data, also known as geospatial or geographic information, is the data or information that identifies the geographic location of features and boundaries on Earth, including natural (oceans, lakes, mountains, *etc.*) or constructed features (boundaries of countries, roads, buildings, *etc.*). Spatial data is usually stored as coordinates and topology, which are data that can be displayed on a map. Spatial data is often accessed, manipulated, or analyzed through several Geographic Information Systems (GIS) [2].

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

1.2 Spatio-Temporal Data

Spatial refers to space and temporal refers to time. Spatial-temporal, or spatio-temporal, is used in data analysis when the collected data has both space and time features. Spatio-temporal data describes phenomena that occur in certain locations and at certain time points or periods such as movements of shipping vessels across a geographic sea location or ocean area over time. Spatio-temporal data can be displayed using special references in various reference systems, but the most common way is to display them on maps. On these maps, temporal features are typically represented via a calendar system. A person can use spatial-temporal reasoning to solve multi-step problems by envisioning how objects move in space and time [3].

Although most data are spatio-temporal, parts of the data become spatial when considered on a particular time point or time period. The second section focuses on spatial data visualization, whereas the third section gives a case study of spatio-temporal data visualization.

2 SPATIAL VISUALIZATION TECHNIQUES

As we discussed in the Introduction section, there is an abundance of collected data that are available for researchers and other endusers in different areas of knowledge (medical, weather, scientific, communication, economics, social media, politics, *etc.*). Such collected raw data are typically pre-processed and cleaned, so that exact information can be extracted out of them. Analysis of these data requires that researchers and data analysts have very good computational skills. Data analysts will find data analysis easier if collected data can be displayed on maps. Presenting collected spatial data graphically by visualizing it on a map makes it easier to interpret and analyze and allows the user to identify areas of interest in the data to do additional computation and further analyses.

Maps are a common way of data visualization. Features of relevant objects and areas are easily visualized through different kinds of maps (geographical maps, building plan maps, website layouts, *etc.*). Several tools and applications for spatial data visualization have been developed and are widely used. Among the most popular map-based visualization techniques are the choropleth map, heat maps, dot maps, and cartograms [4]. In this survey, we will overview six popular, yet simple tools that visualize spatial data, namely: Choropleth maps, Heat maps, Hexagonal binning, Dot maps, Bubble maps, and Cartogram maps.

Currently, the world is witnessing the spread of a virally transmitted infectious disease known as COVID-19 "the Corona Virus Disease 2019". A huge number of individuals were infected by COVID-19, until March 2021 a significant number of the infected people have died (2.5 million deaths according to WHO). The World Health Organization (WHO) has declared a sudden increase in the occurrence of COVID-19 cases in different countries starting the 22^{nd} of January 2020. Since then, there was a daily increase in the number of COVID-19 confirmed cases and deaths. This has resulted in big datasets that can be analyzed and studied. There are lots of personal, professional, and academic sites and agencies that provide access to COVID-19 datasets. A well-known, easy to access, daily updated dataset is maintained by the Centre for Systems Science and Engineering at John Hopkins University (CSSE at JHU) [10].

The six visualization techniques will be illustrated using COVID-19 data in the state of Texas. For each technique, the data will be processed, analyzed, and visualized on a map of Texas counties. The state is divided geographically into 254 counties on the map, and we display the COVID-19 data by county over the whole state of Texas map.

2.1 Choropleth Map

Choropleth map is a thematic map that uses different colorings, color shades, or pattern fills in various areas of a map to represent the value or range of values of some variable. Color shades are patterned differently for the represented geographic areas. The shaded areas are patterned according to statistics and values that are displayed on the map. For instance, a map of the USA might show population density for each state with a darker or lighter shade.

Choropleth maps are used for data presentation and visualization. For example, on the map of a certain county, population density, women participation in labor force, per capita income, presidential election vote percentages, *etc.*, are good examples of such statistical values that can be displayed and visualized using color shades. Significant efforts have been invested into development of various methods of data classification using choropleth mapping [5].

Some advantages of choropleth maps are that they are in widespread use, can be created easily, and they are easy to read and understand by different audiences. Regardless of these advantages, choropleth maps don't give a good representation of statistical values on the boundaries of the shaded areas. In a choropleth map, colors and shading patterns are chosen according to a prearranged key, and as a result the choropleth maps cannot show total values very well. Instead, they show ranges of values, where each range is assigned to a different color or shade. They also cannot show variation within the boundary of each bordered area (county in a state, for example) on the map.

Figure 1 shows COVID-19 confirmed cases per county in Texas using the choropleth map technique. Each county has a different number of cases, so it is colored differently based on the legend found on the top right of the figure. In Figure 1, dark colors represent counties with more confirmed cases by a particular date, in this case September 19, 2020. Lighter colored counties represent counties with fewer cases by that date.

2.2 Heat Maps

Heat maps use colors to communicate numeric data by varying the underlying values. Heat maps can be used to represent large sets of spatial data. Data values are represented red, green, and blue (RGB) as a linear function of the data values [6]. In this way, the color variations represent continuous changes in the values on the map rather than discrete changes by boundaries as in choropleth map.

Using heat map spatial representation makes it easier to understand relationships between data points and draw conclusions about overall trends. The viewer can interpret the chart at multiple scales, from the individual data point to the entire data set, due to the use of color gradients and clear organization. At the other end, Visualization Technique : " Choropleth MAP"



Figure 1: Choropleth Map showing the total COVID-19 cases in Texas by county until September 19, 2020.



Figure 2: Output for python code, showing heat map visualization for COVID-19 data in the state of Texas by county until September 19, 2020. Panel C, is a zoomed-in heat map of North Texas counties

the excessive use of colors can negatively affect the legibility of the visualization. Also, the color schemes must factor in cultural associations with various traits, such as red, which has connotations of heat, anger, passion, lust, as well as various political interpretations. Furthermore, color gradients transcend boundaries on the map so it can be difficult to visualize the information by each bordered area, like the counties in Figure 2

In Figure 2, confirmed cases of COVID-19 in the state of Texas per county are presented. The visualization used python coding of COVID-19 data to obtain the heat maps. From the figure, it can only be determined which counties are hot spots of COVID-19 infection. However, the actual number of confirmed cases cannot be determined. Figure 2C, shows that even when the heat map image is magnified to show the counties of North Texas only, still no statistical numbers can be inferred.

2.3 Dot map

Dot maps, alternatively known as point maps, dot density maps, or dot distribution maps, use dots to represent the data. Dots are placed into an approximate location where the data value was obtained. In general, dots are equally sized and are placed based on the geographical location. The user can obtain information by hovering around the dot on the map.

Dot maps are good at showing spatial patterns, since dots are distributed over the map in different geographic locations. They are best used to display locations of interest of a particular type at a certain location. The data for the distribution of population, dairy farms, shopping centers, coffee shops, schools, universities, mineral deposits, *etc.* can be displayed well using dot maps.

While developing a dot map, two parameters need to be considered: number of dots at a certain location and the value they represent. So, on the map some areas may have a higher number of dots, indicating that these areas have a high concentration of the chosen and less dense areas have low concentration.

Visualization Technique : " Dot MAP"

COVID19 Confirmed Cases in TEXAS by the County on 09-19-2020



Figure 3: Dot map visualization of confirmed cases in Texas counties by September 19, 2020.

Dot maps are optimal for showing the geographical distribution of things and can reveal spatial patterns on the map. They are easy to read, construct, and one can easily interpret total value of the presented data. Yet, it is time consuming to construct such dot maps, as it requires a lot of time for determining the number of dots, and additional time for displaying the georeferenced data.

Figure 3 shows a dot map of the state of Texas with nearly 600,000 confirmed COVID-19 cases until September 19, 2020, spread in different counties. A single blue dot was assigned to each county, yet the actual confirmed case in that county cannot be interpreted from the size of the dot as they are equally sized. The statistics can be displayed by hovering over the dot on the map, where a small pop-up window will show the name of the county and how many confirmed COVID-19 are present. In Figure 3, the popup window illustrates that the confirmed cases of COVID-19 in Dallas County is 77118.

2.4 Hexagonal Binning

To understand the spatial distribution of a large set of data, one can use a point distribution map. But there will not be enough space between displayed points to show where points of the same data value set end and another set of values start. A better way to explore the distribution of such big data on a map, of the same size, is by using regular polygonal bins such as hexagons, circles, or squares. Using hexagons one can form continuous bins, unlike circles and squares.

Hexagonal binning is a visualization technique that creates hexagons as a framework in the map to display the needed information [4]. Different colors and shades can be used for the created hexagons, like what is used in choropleth map. Hexagonal binning has been used to show pick-location of certain taxi cabs, uber clients, amazon delivery locations, election results, *etc.*

Using hexagonal bins has some advantages because patterns in data (like the edge of field) can be seen more easily than if using squares, since a hexagon can have more neighbors on its edge boundaries. Hexagons offer more accuracy in sampling than squares. However, when using hexagons, the cost connectivity analysis is less precise than when using squares. Hexagons have six adjacent neighbors versus eight with the square, but the types of square neighbors are not identical since four have a common line boundary and the other four have a common dot only. In addition, in case of upscaling or downscaling of a data sample with hexagons, data values can't be easily subdivided like what can be done with a square by dividing it evenly into 4 smaller squares.

The map in Figure 4 represents the confirmed COVID-19 cases in Texas counties by 09-20-2020. Each county is represented by a hexagon, using different colors based on the number of confirmed COVID-19 cases in each county.

2.5 Bubble map

In a bubble map, a data source is properly represented with a bubble (circle) of variable. The size of the bubble represents the number of cases at a particular location. Bubbles are distributed over the map according to their geographical coordinates.

Bubble maps are simple and easy to read. From the size of the bubble, one can infer the how high the data it represents is. Also, bubble maps can show several pieces of information in a single view by using a different color for each type of information. The size of the bubble shows the value of the data. A color gradient can be added to show the type of the corresponding value according to map legend. Other advantages of using a bubble map is it can be created from Excel data, different geographic regions can be summarized, the used icons and colors can be customized, and data values can be visualized by region. However, when the size of data varies with in high ranges, then large bubbles need to be displayed on the map. Large bubbles have the problem of overlap with other bubbles and regions [7].

Figure 5 shows a map of Texas. The size of each bubble represents the number of death cases from COVID-19 in different Texas counties until September 19, 2020. From the figure, it can be inferred that bubble maps visually display data better than dot maps. The county with higher number of COVID-19 cases can be easily obtained. Additionally, a popup window tells the statistical number





Visualization Technique : "Bubble MAP"

COVID19 Death Cases in TEXAS by the County till 09-19-2020



Figure 5: Bubble Map showing COVID-19 death cases in Texas counties.

of COVID-19 cases in each county. Dallas county has 1083 death from COVID-19 till September 19, 2020.

2.6 Cartogram map

Cartograms are maps that deal with geographical regions (countries, states) with respect to a certain variable of interest (population, income, number of confirmed COVID-19 cases, *etc.*). They have been used as a visualization means for geo-referenced data for over a century. Cartograms are widely used by newspapers, magazines, textbooks, blogs, and presentations to display election results, popularity, and regional patterns in general [8]. To convey related information clearly, the real size and shape of geographical regions are distorted in cartogram maps [9] Cartograms contain both statistical and geographical information. The confirmed and death cases of COVID-19 by August 11, 2020, are shown in the cartograms of Figure 6. The maps show accurate shape of the states of the USA, yet the size of each state is distorted to show corresponding COVID-19 statistics. This causes a change in the borders of the states of the real map. These are known as non-contiguous cartograms. The geometrical regional shapes were distorted by scaling down the size in the original maps.

2.7 Default "Traditional" map

Certain features of data values cannot be visualized using the techniques discussed previously. When data is collected and is available to be analyzed and visualized, traditional map can be a good choice for visualizing the analyzed data. Traditional maps offer a simple method of visualization where shading/coloring/highlighting the area of interest for a particular feature is used. The analyzed data



Figure 6: Cartogram Map showing the confirmed and death cases from COVID-19 in the US by state.

Visualization Technique : " Traditional MAP"

COVID19 Confirmed Cases in TEXAS by the County on 09-19-2020



Figure 7: Traditional Map showing COVID-19 cases in Texas counties by September 19, 2020.

is visualized individually without being compared to any other feature or data point. Google maps is a good example of traditional maps. New locations, routes and landmarks can be added as overlays for visualization. Also new shapes can be drawn to show the details of the data to be visualized. Most of the techniques defined by spatial data visualization researchers have been listed, yet some users may not follow any of these techniques. They will use their own technique of visualizing the data resulting from a spatial query. Figure 7 shows traditional map visualization technique, where a green icon is chosen to show the geo-location of the county, and when a user hovers on the icon, it will pop-up with the available data.

Other traditional visualization techniques include highlighting a boundary of a selected area such as a county or highlighting a path such as a suggested route on the map between two locations.

3 CASE STUDY: VISUALIZING SPATIO-TEMPORAL DATA BY EXTENDING BUBBLE MAPS

Choropleth maps, Heat maps, Dot maps, Bubble maps, Hexagonal binning, Cartograms and Traditional maps were used to visualize data of COVID-19 for a specific date (time point) as provided by CSSE at JHU. Moreover, some of these visualization techniques can be extended to visualize spatio-temporal data. In this section, we describe the implementation of extended bubble map visualization to display spatial data as it changes over time. Bubble maps will be used as a proof of principle example, but we will discuss how other techniques can be extended for spatio-temporal data visualization in future work.

The extended bubble maps will show a different visualization when a different date is chosen. COVID-19 datasets change daily [10]; there is a daily increase in the number of confirmed and death cases. In our example, bubble map visualization will be used to visual COVID-19 datasets for counties of Texas. The spatiotemporal visualization will be initiated from the start of the COVID-19 pandemic on January, 2020 and will include the following details:

- Dataset: time_series_county_JH.csv
- Programming language: Python.
- Libraries used: pandas, glob, os, plotly.
- Jupyter notebook will be used.

The dataset will be loaded from "John Hopkins CSSE Department" for confirmed COVID-19 cases in the US Figure 8. The loaded data contains information about the confirmed cases in all the states. It contains lot of columns that will not be used in the visualization. Data related to Texas state counties will be chosen. Data will be cleaned and transposed so that it can be visualized.

The selected data frames are saved to a .csv file, so that data can be accessed locally when needed. The data will be cleaned leaving the columns containing the counties of Texas and the dates of the COVID-19 cases only. Figure 9 shows the code used for data cleaning.

Now the data frames are ready to be transposed where the columns will be the name of the counties and the rows will contain the number of confirmed cases according to the row name which will be a date. Map Visualization using Spatial and Spatio-Temporal Data: Application to COVID-19 Data

.. ...

[1]: f ii ii ii ii	from pandas import Series, DataFrame import pandas as pd import glob import plotly.express as px															
L	Load D	ata														
[2]: #	‡ upto-da	te Da	ta, l	loaded	onlin	e right	from the se	ource "john H	Hopkins (SSE Depa	rtm	ent" fo	r confin	med Cov	rid 19 c id 19 da	ases i ta/css
0	confirmed	_df =	pd.r	read_c	5V(IIL	(µ5://1	awigitinous	erconcentreo	, CJJC013			10 13710				
0 [3]: 0	confirmed	_df =	pd.r ead()	read_c	54(110	(ps://1	aw.grtnoods		,			10 137				
(3): 0 t(3):	confirmed confirmed	_df = _df.h iso2	pd.r ead() iso3	code3	FIPS	Admin2	Province_State	Country_Region	Lat	Long_		10/19/20	10/20/20	10/21/20	10/22/20	10/23/2
(3): co t(3):	confirmed confirmed UDD 0 84001001	_df = _df.h iso2 US	pd.r ead() iso3 USA	code3	FIPS 1001.0	Admin2 Autauga	Province_State Alabama	Country_Region	Lat 32.539527	Long_ -86.644082		10/19/20 1999	10/20/20 2010	10/21/20 2021	10/22/20 2023	10/23/21
[3]: 0 :[3]: - (confirmed confirmed UID 0 84001001 1 84001003	_df = _df.h iso2 US	ead() iso3 USA USA	code3 840 840	FIPS 1001.0 1003.0	Admin2 Autauga Baldwin	Province_State Alabama Alabama	Country_Region US US	Lat 32.539527 30.727750	Long_ -86.644082 -87.722071		10/19/20 1999 6375	10/20/20 2010 6405	10/21/20 2021 6443	10/22/20 2023 6475	10/23/2 203 661
[3]: 0 t(3]: 1	confirmed, UID 0 84001001 1 84001003 2 84001005	_df = _df.h _us _us _us	ead() iso3 USA USA USA	code3 840 840	FIPS 1001.0 1003.0 1005.0	Admin2 Autauga Baldwin Barbour	Province_State Alabama Alabama	Country_Region US US US	Lat 32.539527 30.727750 31.868263	Long_ -86.644082 -87.722071 -85.387129		10/19/20 1999 6375 981	10/20/20 2010 6405 988	10/21/20 2021 6443 996	10/22/20 2023 6475 997	10/23/2 203 661 101
(3): 0 t(3): t t	confirmed confirmed 0 84001001 1 84001005 2 84001005 3 84001007	_df.h iso2 US US US	ead() iso3 USA USA USA USA	code3 840 840 840 840	FIPS 1001.0 1003.0 1005.0 1007.0	Admin2 Autauga Baldwin Barbour Bibb	Province_State Alabama Alabama Alabama	Country_Region US US US US	Lat 32.539527 30.727750 31.868263 32.996421	Long_ -86.644082 -87.722071 -85.387129 -87.125115		10/19/20 1999 6375 981 789	10/20/20 2010 6405 988 791	10/21/20 2021 6443 996 801	10/22/20 2023 6475 997 811	10/23/20 2030 6615 1012 825

Figure 8: Python code used to load COVID-19 data sets.

In [15]:	<pre>.5]: # leave data of counties and dates# df_confirmed_tx.drop(df_confirmed_tx.columns[[0,2,3,4]], axis = 1, inplace = True) df_confirmed_tx</pre>																	
Out[15]:		County	1/22/20	1/23/20	1/24/20	1/25/20	1/26/20	1/27/20	1/28/20	1/29/20	1/30/20		10/19/20	10/20/20	10/21/20	10/22/20	10/23/20	10/24/20
	2698	Anderson	0	0	0	0	0	0	0	0	0		2909	2920	2922	2922	2934	2941
	2699	Andrews	0	0	0	0	0	0	0	0	0		547	549	567	567	567	584
	2700	Angelina	0	0	0	0	0	0	0	0	0		2251	2251	2277	2277	2277	2277
	2701	Aransas	0	0	0	0	0	0	0	0	0		341	340	340	340	340	346
	2702	Archer	0	0	0	0	0	0	0	0	0		114	114	133	133	133	137
		-																
	2949	Wood	0	0	0	0	0	0	0	0	0		590	610	621	621	623	626
	2950	Yoakum	0	0	0	0	0	0	0	0	0		284	289	289	289	289	289
	2951	Young	0	0	0	0	0	0	0	0	0		470	470	497	497	507	535
	2952	Zapata	0	0	0	0	0	0	0	0	0		342	342	342	342	342	345
	2953	Zavala	0	0	0	0	0	0	0	0	0		438	443	452	452	458	460
	256 rd	ws × 282	columns															

Figure 9: Python Code used for data cleaning.

The transposed data frames will be organized into three columns (Date, Confirmed, and County). The organized data frames can be visualized only if "plotly" bars are used. Yet, they need to be visualized on the map. To do so, geographic coordinates need to be added to the dataset. Such data is available online and can be downloaded. It contains the geographic coordinates for all the counties in all the US states. The data related to "Texas" will be selected only.

The Lat. and Long. dataset are now merged with the organized dataset. The resulting data frames are then saved into a .csv file "Time_series_county_JH.csv" to be accessed later in the visualization step. In this step, Plotly Express and some of its plugins are used.

Figure 10 is the output of the code used. It shows an animated time series for COVID-19 data from Texas counties by applying bubble map technique. The Mapbox plugin which is supported by Plotly Express is used in the visualization. The size of the bubble corresponds to the number of the confirmed COVID-19 cases in each county. As can be seen, the color of the bubble matches the actual corresponding numeric value according to the legend bar in the map.

4 CONCLUSION

In this paper, six spatial data visualization techniques were illustrated by displaying COVID-19 data. The techniques were presented and discussed based on the spatial feature of the data. Some of these visualization techniques can be extended to visualize spatiotemporal data when the temporal attribute is added. A case study was presented where the bubble map method was successfully extended to visualize spatio-temporal COVID-19 data.

For future work, we plan to implement a tool that can perform spatio-temporal visualization using several of the visualization techniques we discussed and presented. The planned tool will allow the user to choose one of the techniques to visualize his or her data. We plan to develop guidelines that help the user determine which technique to use for their particular application. We will also try to apply one of the spatial clustering algorithms on COVID-19 data so that the reported data can be visually clustered on the map.

Mohammad Shaito and Ramez Elmasri

PETRA 2021, June 29-July 02, 2021, Corfu, Greece



Figure 10: Final output of the code showing the time series animation for confirmed COVID-19 case in Texas

REFERENCES

- lliott W., 2018. Paysafe Insights, The world's most valuable resource: Not oil, but data, https://www.paysafe.com/en/blog/the-worlds-most valuable-resource-notoil-but-data/ (accessed August 1, 2020)
- Beal, V., 2018. Spatial Data. Retrieved May 01, 2020, from http://webopedia.com/ TERM/S/spatial_data.html
- [3] OmniSci. 2020. Spatiotemporal. Retrieved August 03, 2020, from OmniSci: https: //www.omnisci.com/technical-glossary/spatial-temporal
- [4] Vasavi A., 2016. 7 Techniques to Visualize Geospatial Data Atlan: Humans of Data. Retrieved August 18, 2020, from https://humansofdata.atlan.com/2016/10/7techniques-to-visualize-geospatial-data
- [5] Andrienko, G., Andrienko, N., Savinov, A., 2001. A: Choropleth Maps: classification revisited. Proceedings of ICC 2001: 6-10 August 2001. 2001, Beijing. Edited

by Scientific and Technical Program Committee LOC, 1209-1219

- [6] Kuhfled, W. F., 2017. Heat Maps: Graphically Displaying Big Data and Small Tables. SAS Institute Inc., Cary, North Carolina, USA. 2017., https://support. sas.com/resources/papers/proceedings17/SAS0312-2017.pdf (accessed on July 5 2020).
- [7] Severino, R., 2017. Bubble map, https://datavizcatalogue.com/methods/bubble_ map.html (accessed on June 30, 2020)
- [8] Nusrat, S., and S. Kobourov., 2015. Visualizing cartograms: Goals and task taxonomy. https://arxiv.org/abs/1502.07792
- [9] GISGeography, 2020. Cartogram Maps: Data Visualization with Exaggeration, https://gisgeography.com/cartogram-maps/ (accessed on October 30, 2020).
- [10] CSSE at John Hopkins University. 2019. CSSEGISandData/COVID-19. Retrieved September 20, 2020, from https://github.com/CSSEGISandData/COVID-19