Big Data Visualization and Visual Analytics of COVID-19 Data

Carson K. Leung[1]*, Yubo Chen[1], Calvin S.H. Hoi[1], Siyuan Shang[1], Yan Wen[1], Alfredo Cuzzocrea[2]

[1] *Department of Computer Science, University of Manitoba*, Winnipeg, MB, Canada
[2] *iDEA Lab, University of Calabria*, Rende, CS, Italy
* kleung@cs.umanitoba.ca

Abstract—In the current era of big data, a huge amount of data has been generated and collected from a wide variety of rich data sources. Embedded in these big data are useful information and valuable knowledge. An example is healthcare and epidemiological data such as data related to patients who suffered from epidemic diseases like the coronavirus disease 2019 (COVID-19). Knowledge discovered from these epidemiological data helps researchers, epidemiologists and policy makers to get a better understanding of the disease, which may inspire them to come up ways to detect, control and combat the disease. As "a picture is worth a thousand words", having methods to visualize and visually analyze these big data makes it easily to comprehend the data and the discovered knowledge. In this paper, we present a big data visualization and visual analytics tool for visualizing and analyzing COVID-19 epidemiological data. The tool helps users to get a better understanding of information about the confirmed cases of COVID-19. Although this tool is designed for visualization and visual analytics of epidemiological data, it is applicable to visualization and visual analytics of big data from many other real-life applications and services.

Keywords—big data, visualization, visual analytics, COVID-19, epidemiological data

I. INTRODUCTION

As we are living an era of big data, big data are everywhere. To elaborate, with advances in technology, a huge amount of data has been easily generated and collected from a wide variety of rich data sources at a rapid rate. These big data can be of different levels of veracity (e.g., precise data, imprecise and uncertain data [1-3]). Examples of big data include social network data [4-8], financial time series [9-11], omic data (e.g., genomic data) [1, 12], disease reports [13, 14], as well as epidemic data and statistics.

Embedded in these big data are useful information and valuable knowledge. This calls for data science, which aims to discover knowledge from these big data via data mining algorithms, machine learning tools, mathematical and statistical models, data analytics, and visual analytics. The discovered knowledge is useful. For instance, knowledge discovered from these epidemiological data helps researchers, epidemiologists and policy makers to get a better understanding of the disease, which may inspire them to come up ways to detect, prevent, and/or control diseases such as viral diseases. Examples of viral diseases include:

- severe acute respiratory syndrome (SARS), with outbreak in 2002–2004;

- Middle East respiratory syndrome (MERS), with outbreak in 2012–2015; and

- coronavirus disease 2019 (COVID-19), with outbreak started in 2019 and became pandemic in 2020.

Due to the COVID-19 pandemic, many researchers have focused on different aspects of the COVID-19 disease. These include clinical and treatment information [15, 16], as well as drug discovery [17], related on research medical and health sciences. In contrast, as computer scientists, we focus on other aspects of COVID-19 data—namely, epidemiological data.

As "a picture is worth a thousand words", having methods to visualize and visually analyze these big data makes it easily to comprehend the data and the discovered knowledge [18]. A majority of the existing visualizers on the COVID-19 epidemiological data focused on showing the numbers of confirmed cases and mortality spatially and/or temporally. In other words, they show

- spatial differences among different continents, countries, regions, or sovereignties; and/or

- temporal differences among weeks or days along the timeline—e.g., to show the effects of public health strategies and mitigation techniques such as social/physical distancing or stay-at-home orders in "flattening the (epidemic) curve".

As the numbers of inhabitants and tests both play roles in the data and their analyses, they help in the computation of figures like (a) the numbers of confirmed cases and mortality per thousand/million inhabitants and (b) the number of tests per thousand inhabitants.

However, in addition to the number of cases or mortality, there are other important knowledge that can be discovered from the epidemiological data via data mining. For instance, frequent pattern mining finds relationships among attributes (or features) associated with confirmed COVID-19 cases. Moreover, a visual representation of this discovered knowledge gives a more comprehensive representation, which in turn leads to a better insight and understanding of the data and discovered knowledge. Hence, in this paper, we present a tool for big data visualization and visual analytics of COVID-19 epidemiological data. Due to the nature of these data, it is not unusual to have NULL values for some of the attributes (e.g., unstated transmission methods of disease). We tool provides users with flexibility to visualize the data with and without stated values for attributes.

Our *key contributions* of this paper include our design and development of a visualization and visual analytics tool for COVID-19 epidemiological data. Our tool incorporates (a) data analytics (especially, frequent pattern mining), (b) data visualization, and (c) visual analytics (especially, visualization and analysis of frequent patterns). Moreover, although the tool is designed for visualization and visual analytics of the COVID-19 epidemiological data, it is applicable to visualization and visual analytics of big data from many other real-life applications and services.

The remainder of this paper is organized as follows. Next section discusses some related work. Section III presents our visualization and visual analytics tools. Section IV shows evaluation results, and Section V draws the conclusions.

II. RELATED WORKS

A. COVID-19 Visualizers

Due to the COVID-19 pandemic, many visualizers and dashboards have been developed over the past few months. Some of them [19-21] visualized literatures related to COVID-19 research, and some others [22] visualized economic impact of COVID-19. However, a majority of them [23] focused on the actual COVID-19 cases. Globally notable visualizers include (a) World Health Organization (WHO) Coronavirus Disease 2019 (COVID-19) Dashboard [24], (b) COVID-19 Dashboard by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University (JHU)[1], and (c) COVID-19 dashboard by European Center for Disease Prevention and Control (ECDC)[2]. In addition, Government of Canada[3], provincial and territorial governments, major news channels/media/networks (e.g., TV[4], newspaper), and Wikipedia[5] capture data and provide dashboards on information about the COVID-19 pandemic situations in Canada. There are several commonality among these visualizers and dashboards. For instance, they mostly focused on the total numbers of new cases, confirmed cases, and deaths.

The *spatial information* about the total of numbers of confirmed cases and deaths in different countries (or regions/sovereignties) is usually represented by (a) the bubble map or (b) the choropleth map. To elaborate, in a *bubble map*, the total number of confirmed cases for each country is indicated by the radius of the bubble representing the country. See Fig. 1. Similarly, the bubble map can also show the number of new cases or deaths, in an absolute value (e.g., N_1 newly reported cases) or a relative figure with respect to population (e.g., N_2 deaths per one million population). See Fig. 2.

While the severity of COVID-19 in many countries can be representing by the sizes or radii of the bubbles representing these countries, many bubbles overlap. The overlapping and/or containment of bubbles makes it challenging to visualize the severity of the disease in countries in dense regions such as Eastern Caribbean and Southeastern Europe as shown in Fig. 2.
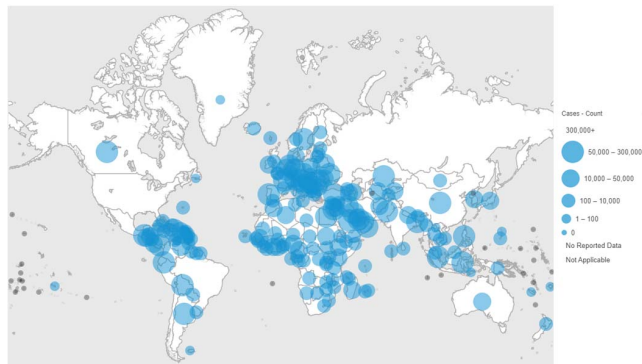

Fig. 1. A snapshot of a bubble map [24] showing the total number of confirmed cases among different countries in the world as of August 07, 2020.
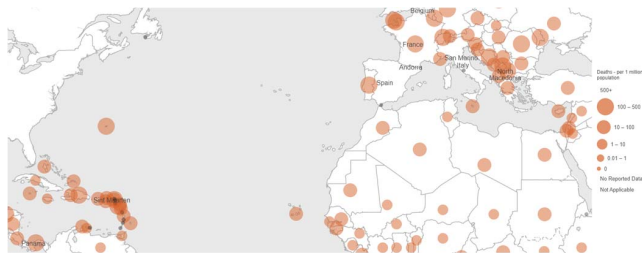

Fig. 2. A zoom-in view of a bubble map [24]—with overlapping bubles (especially in Eastern Caribbean and Southeastern Europe)—showing the total number of deaths per 1M population among different countries in the world as of August 07, 2020.
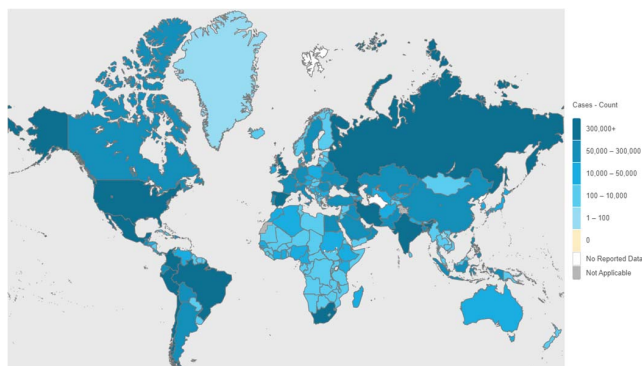

Fig. 3. A snapshot of a choropleth map [24] showing the total number of confirmed cases among different countries in the world as of August 07, 2020.

As an alternative, a *choropleth map* uses differences in shading, coloring, or the placing of symbols within predefined areas to indicate the average values of a property or quantity in those areas. For instance, Fig. 3 shows a choropleth map for the COVID-19 epidemiological data, in which differences in shading to represent the total number of confirmed cases. The darker the shading, the more severity is the COVID-19 situation, which means the higher is the number of confirmed cases. Note

---

that the choropleth map in Fig. 3 shows the same information as the bubble map in Fig. 1 but with an alternative representation.

While the use of choropleth map solves the problems associated with the overlapping and containment of bubbles, it leads to another problem. For example, small countries in terms of geographic areas or sizes (e.g., Monaco, San Marino, Liechtenstein) may not be easily visible on the map. See Fig. 4.

Fig. 4. A zoom-in view of a choropleth map [24]—covering some small countries (e.g., Monaco, San Marino, Liechtenstein) that may not be easily spotted—showing the newly reported cases in last 7 days among different countries in the world as of August 07, 2020.
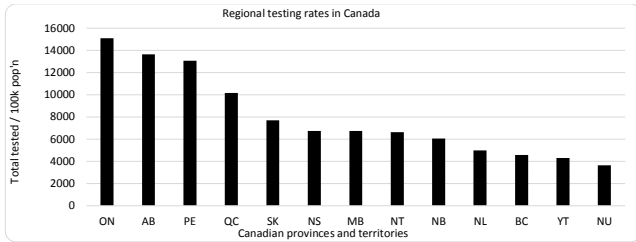

Fig. 5. A column chart showing the total number of tested per 100,000 inhabitants among provinces and territories in Canada as of August 06, 2020.

Moreover, *column charts* also provide spatial information from the COVID-19 epidemiological data. For example, Fig. 5 shows a column chart for comparing regional testing rates among the 10 provinces and 3 territories of Canada.

Besides spatial information, *temporal information* from the COVID-19 epidemiological data is also important as it shows the trends. Temporal information (e.g., daily or cumulative numbers of new cases, confirmed cases and deaths) is usually represented by *line graphs*, *column charts* (or stacked column charts), and *areas under curve* (or stacked areas under curve). Examples are shown in Figs. 6-7.
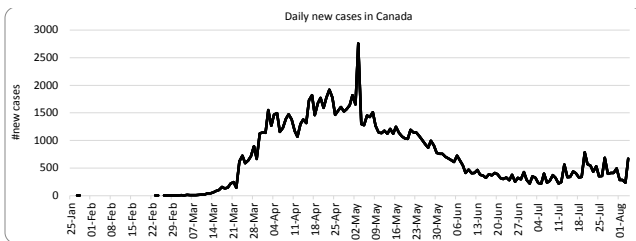

Fig. 6. A line graph showing the daily new confirmed cases in Canada from January 25 (for the first confirmed case in Canada) to August 04, 2020.
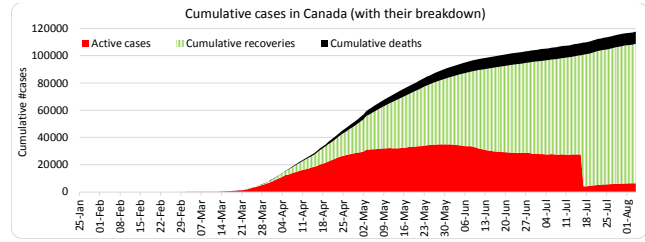

Fig. 7. A stacked area under curve showing the cumulative cases in Canada and their breakdown (i.e., active cases, recoveries, and deaths) from January 25 to August 04, 2020.

To a further extent, *combined spatial and temporal information* from the COVID-19 epidemiological data can be represented by *line graphs*. For example, daily new cases over time for all provinces and territories can be represented by a line graph with 13 lines, and each represents the trend of a province or territory.

B. Frequent Pattern Visualizers

In addition to the aforementioned attributes or features (e.g., numbers of new or active cases, recoveries, deaths), there are other attributes that are also important. These include administrative information, case details, symptoms, clinical course and outcomes, as well as exposures. Occasionally, some of this information (e.g., gender and age of confirmed cases) is available and reported in textual form. As "a picture is worth a thousand words", having a visual representation of this information can be helpful. Moreover, knowing the relationships among them can also be useful. This calls for data mining and visual analytics. For instance, the data-mining task of *frequent pattern mining* [25, 26] discovers frequently occurring ⟨attribute, value⟩-pairs and reveals relationships among these pairs.

Several platforms and visualizers [27-29] have been built to visualize knowledge discovered from different data mining tasks. Among them, some [30, 31] were designed and developed to visualize the discovered frequent patterns. For example, FIsViz [32] displays each frequent pattern containing k items (i.e., k-itemset) in the form a polyline that connects k nodes in a 2-dimensional space. Crucial information about any frequent k-itemset is captured by its associated (x, y)-coordinates in this space. As another example, observing that polylines in FIsViz may not be easily distinguishable from one another due to their potential bending and crossing-over for frequent patterns, FpVAT [33] displays each frequent k-itemset in a horizontal line that connects k nodes in a 2-dimensional space in a wiring-type diagram (i.e., orthogonal graph). See Fig. 8.

As an alternative to frequent pattern visualization via polylines or orthogonal wires, PyramidViz [34] visualizes frequent patterns in a hierarchical layout—namely, a building block layout. To elaborate, short patterns are put on the bottom of the pyramid, whereas longest related patterns (which are extensions of short patterns) are put on the top of the pyramid. In contrast to PyramidViz (which visualizes frequent patterns with a *side-view* of the pyramid), FpMapViz [35] visualizes frequent patterns with a *top-view*. To elaborate, short patterns are put on the background, whereas longest related patterns (which are extensions of short patterns) are put on the

foreground. See Fig. 8. However, these visualizers were not designed for visualizing COVID-19 epidemiological data.
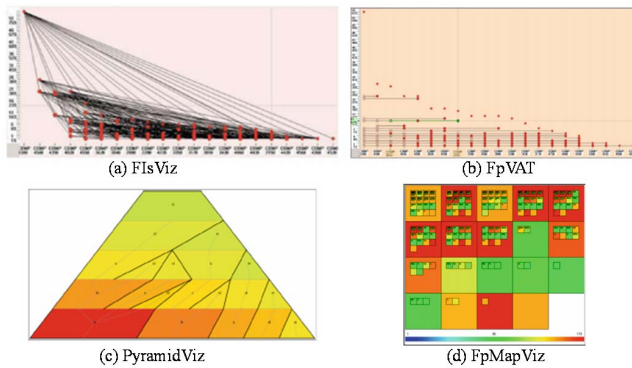

Fig. 8. Frequent pattern visualizers: (a) FIsViz [32], (b) FpViz [33], (c) PyramidViz [34], and (d) FpMapViz [35].

III. OUR VISUALIZER FOR ANALYZING COVID-19 EPIDEMIOLOGICAL DATA

To help users and researchers to get a better understanding of COVID-19 disease, we design and develop a big data visualization and visual analytics tools for COVID-19 epidemiological data. In this section, we illustrate our idea on COVID-19 epidemiological data for Canada [36].

A. Data Collection

Our tool first collects different categories of big COVID-19 epidemiological data from different sources (e.g., federal and provincial/territorial governments). These include:

- administrative information, which includes an unique privacy-preserving identifier for each case, its region and episode week (i.e., symptom onset week or its closest week);

- case details, which include gender, age group, and occupation of the cases (e.g., health care workers, school or daycare workers, long-term care residents);

- symptom-related data, which include additional information for the case who is not asymptomatic (i.e., symptomatic case) such as onset week of symptoms, as well as a collection of symptoms (e.g., cough, fever, chills, sore throat, runny nose, shortness of breath, nausea, headache, weakness, pain, irritability, diarrhea, and other symptoms);

- clinical course and outcomes, which include hospital status (e.g., hospitalized in the intensive care unit (ICU), non-ICU hospitalized, not hospitalized, unstated). For recovered case, it also includes additional information such as the recovery week. For the case who has not recovered, it indicates that the case died while infected by COVID-19; as well as

- exposures, which include transmission methods (e.g., domestic acquisition via contact of COVID-19 case or contact with traveller; international travel).

B. Data Preprocessing

After gathering related data from heterogeneous sources, our tool preprocesses the data. Given the nature of these cases, it is not unusual to have missing, unstated or unknown information (i.e., NULL values). For example, for some "Boolean" attributes (e.g., attribute "asymptomatic"), we observed three values: asymptomatic, symptomatic, and unstated (i.e., NULL value). Our tool links all values (including NULL) for each attribute. Moreover, our tool also detects and flags any data inconsistency.

C. Visualization of Frequent Patterns of Cardinality 1

Once the data are preprocessed, we integrate the data mining capability into our tool. Specifically, via frequent pattern mining, our tool first discovers frequently occurring domain attributes (i.e., 1-itemsets). Our tool uses a *pie chart* or sunburst diagram to represent values for each frequently occurring domain attribute.

Recall that existing frequent pattern visualizers mentioned in Section II-B display all frequent patterns. In contrast, our tool displays frequent patterns one-by-one. For a frequent 1-itemset, we show not only the frequent ⟨*attribute, value*⟩-pair but also pairs associated with other values for that *attribute*. This enables users to get insight about the portion of that *value* for the *attribute*.

Example 1 As a preview, when applying our tool to the COVID-19 epidemiological data for Canada up to August 06, 2020, our tool finds a frequent 1-itemset ⟨transmission, domestic acquisition⟩-pair with a frequency of 97,052 out of 107,916 cases (90%). Our tool also displays other pairs for attribute "transmission" in the pie chart: ⟨transmission, international travel⟩ (4%) and ⟨transmission, NULL⟩ (6%).

In addition, our tool gives users an option to ignore the NULL value for any attribute. If the user select this option, our tool displays not only the frequent ⟨*attribute, value*⟩-pair but also pairs associated with other non-NULL values for that *attribute*. This enables users to get insight about the portions of all non-NULL values for the *attribute*.

Example 2 Continue with our preview, without NULL values, our tool finds a frequent 1-itemset ⟨transmission, domestic acquisition⟩-pair with a frequency of 96% of cases with stated transmission methods. Our tool also finds and displays another pair with a non-NULL value for attribute "transmission": ⟨transmission, international travel⟩ (4%).

D. Visualization of Frequent Patterns of Higher Cardinality

In addition to showing each frequent 1-itemset ⟨*attribute, value*⟩-pair and its related pairs for other values of the *attribute* one-by-one, our big data visualization and visual analytic tool also shows each frequent *k*-itemset—i.e., frequent pattern of higher cardinality *k* (where *k* > 1)—by using a *sunburst diagram*. The diagram consists of *k* rings, one for each attribute (i.e., item) within the *k*-itemset. Again, our tool gives users an option to ignore the NULL value for any attribute.

Example 3 As another preview, our tool finds a frequent 3-itemset {⟨transmission, domestic acquisition⟩, {⟨death, no⟩, {⟨hospital status, not hospitalized⟩} with a frequency of 51,523 out of 107,916 cases (48%). Our tool also displays its related

itemsets: (a) $\{\langle$transmission, domestic acquisition$\rangle$, $\{\langle$death, no$\rangle$, $\{\langle$hospital status, $value_3\rangle\}$'s where $value_3 \in \{$non-ICU hospitalized, ICU, NULL$\}$, (b) $\{\langle$transmission, domestic acquisition$\rangle$, $\{\langle$death, $value_2\rangle\}$'s where $value_2$ is "yes" or NULL, and (c) $\{\langle$transmission, $value_1\rangle\}$'s where $value_1$ is "international travel" or NULL, in the sunburst diagram.

Example 4 Continue with our preview, without NULL values, our tool displays the frequent 3-itemset $\langle$transmission, domestic acquisition$\rangle$-pair in Example 3 with a frequency of 74% of cases with stated transmission methods, death status and hospital status.

IV. EVALUATION

A. Functionality Check with Related Works

To evaluate our tool, we applied it to different COVID-19 epidemiological data including the Canada cases from Statistics Canada [36]. In terms of functionality, as recall from Section II-A, many existing COVID-19 visualizers or dashboards mostly focus on single attribute (e.g., numbers of cases, recoveries, deaths) at any instance. In contrast, our big data visualization and visual analytics tool reveals relationships among different attributes related to epidemiological data.

Moreover, as recall from Section II-B, existing frequent pattern visualizers aim to display all and only frequent k-itemsets. In contrast, our tool visualizes not only each frequent k-itemset one-by-one but also its related $\langle$attribute, value$\rangle$-pairs within the k-itemset to give users an insight about the relationships (e.g., percentage) of the frequent k-itemset with respect to its related itemsets.

Furthermore, our tool provides users with options to include or exclude NULL values (which are common in COVID-19 data and many other datasets). By doing so, the users could get a better understanding of the distribution of the data and the discovered patterns, with respect to all values (including NULL) and all known/stated values (i.e., non-NULL values).

B. Case Study with Real-Life COVID-19 Epidemiological Data

To further evaluate our tool, we applied it to COVID-19 epidemiological data for Canada [36] captured by Public Health Agency of Canada (PHAC) and Statistics Canada. The data contain administrative information, case details, symptom-related information, clinical course and outcomes, as well exposure methods, for all 107,916 captured cases from January 25 (when the first case confirmed in Canada) to August 06, 2020.

Our tool conducts visual analytics on the data to discover frequent patterns and visualizes the discovered knowledge by displaying interesting information in the form of a pie chart for each frequent 1-itemset (and its related information) and a sunburst diagram for each frequent k-itemset (and its related information, for $k > 1$). As previewed in Example 1, our tool discovers that 90% of cases were transmitted through domestic acquisition (i.e., community exposures). See Fig. 9(a), which also shows that, among the remaining 10% of cases, 4% were transmitted through international travel (i.e., travel exposures) and 6% were unstated transmission (i.e., NULL).

To avoid distraction from NULL values, our tool provides users with flexibility of visualizing non-NULL values. See Fig. 9(b), which focuses on the 90%+4% = 94% of cases (i.e., those with stated/known values). As previewed in Example 2, our tool reveals that $90/94 \approx 96\%$ of cases with stated/known transmission methods were transmitted through domestic acquisition, whereas the remaining $4/94 \approx 4\%$ were transmitted through international travel.
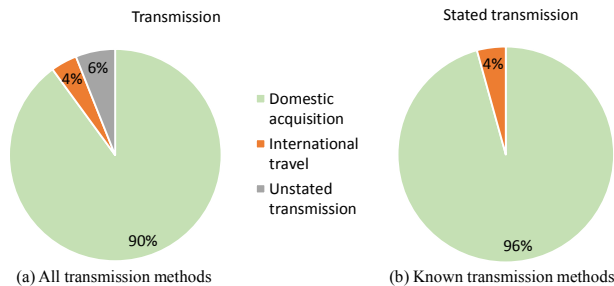
Fig. 9. Our visualization and visual anlytics tool shows a frequent 1-itemset $\{\langle$transmission, domestic acquisition$\rangle\}$ and its related values for attribute "transmission" by (a) inlcuding and (b) excluding the NULL value.

As previewed in Examples 3 and 4, our tool displays that 48% of cases who were transmitted through domestic acquisition recovered without being hospitalized. See Fig. 10(a), which also shows that, 5% required non-ICU hospitalization, 1% admitted to ICU, and 20% with unstated hospitalization status, for a total of 48+5+1+20 = 74% of cases survived. Unfortunately, 7% of cases deceased and 9% with unstated death status, for a total of 74+7+9 = 90% of cases transmitted through domestic acquisition. The remaining 4% and 6% were transmitted through international travel exposures) and with unstated transmission (i.e., NULL). By ignoring NULL values, Fig. 10(b) reveals that, among cases with known transmission methods, death status and hospitalization, 74% of them were transmitted through domestic acquisition recovered without being hospitalized.
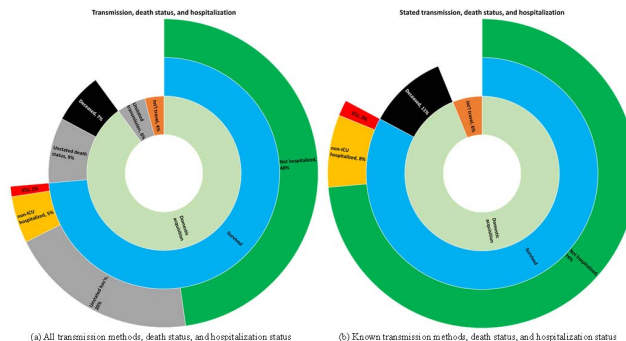
Fig. 10. Our visualization and visual anlytics tool shows a frequent 3-itemset $\{\langle$transmission, domestic acquisition$\rangle$, $\langle$death, no$\rangle$, $\{\langle$hospital status, not hospitalized$\rangle\}\}$ and its values for related attributes by (a) inlcuding and (b) excluding NULL values.

V. CONCLUSIONS

Over the past few months, there have been works on visualizing and analyzing different aspects of big data related to COVID-19. However, a majority of existing visualizers focus

on showing the temporal and/or spatial trends on the numbers of cases and mortality. In this paper, we focus on epidemiological aspects of COVID-19 data. Our key contributions include our design and development of a big data visualization and visual analytics tool for COVID-19 epidemiological data. By incorporating big data mining and data analytics into our tool, we discover frequent patterns, together with their related patterns. We provide users with flexibility to include or exclude NULL values. Evaluation results show benefits of our tool in providing a comprehensive and effective visual representation of these important patterns, which in turn helps researchers to get a better understanding of the COVID-19 disease and thus enables them to combat the disease. Moreover, our tool can be applicable to other real-life applications with NULL values. As ongoing and future work, we explore possibility to incorporate other techniques [37-40] into our tool.

REFERENCES

[1] F. Jeanquartier, C. Jean-Quartier, A. Holzinger, "Visualizing uncertainty for comparing genomic pediatric brain cancer data," in IV 2019, Part I, pp. 388-391.

[2] C.K. Leung, "Uncertain frequent pattern mining," in Frequent Pattern Mining, 2014, pp. 417-453.

[3] C.K. Leung, M.A.F. Mateo, D.A. Brajczuk, "A tree-based approach for frequent pattern mining from uncertain data," in PAKDD 2008, pp. 653-661.

[4] S. Fernando, et al., "Compositional microservices for immersive social visual analytics," in IV 2019, Part I, pp. 216-223.

[5] F. Jiang, C.K. Leung, S.K. Tanbeer, "Finding popular friends in social networks," in CGC 2012, pp. 501-508.

[6] C.K. Leung, S.K. Tanbeer, J.J. Cameron, "Interactive discovery of influential friends from social networks," SNAM 4(1), 2014, pp. 154:1-154:13.

[7] M. Mai, et al., "Big data analytics of Twitter data and its application for physician assistants: who is talking about your profession in Twitter?" in Data Management and Analysis, 2020, pp. 17-32.

[8] S.K. Tanbeer, C.K. Leung, J.J. Cameron, "Interactive mining of strong friends from social networks and its applications in e-commerce," JOCEC 24(2-3), 2014, pp. 157-173.

[9] R.C. Camara, et al., "Fuzzy logic-based data analytics on predicting the effect of hurricanes on the stock market," in FUZZ-IEEE 2018, pp. 576-583.

[10] D. Jonker, R. Brath, S. Langevin, "Industry-driven visual analytics for understanding financial timeseries models," in IV 2019, Part I, pp. 210-215.

[11] C.K. Leung, R.K. MacKinnon, Y. Wang, "A machine learning approach for stock price prediction," in IDEAS 2014, pp. 274-277.

[12] O.A. Sarumi, C.K. Leung, "Scalable data science and machine learning algorithm for gene prediction," in BigDAS 2019, pp. 118-126.

[13] A. Diallo, et al., "Proportional visualization of genotypes and phenotypes with rainbow boxes: methods and application to sickle cell disease," in IV 2019, Part I, pp. 1-6.

[14] J. Souza, C.K. Leung, A. Cuzzocrea, "An innovative big data predictive analytics framework over hybrid big data sources with an application for disease analytics," in AINA 2020, pp. 669-680.

[15] A.A. Ardakani, et al., "Application of deep learning technique to manage COVID-19 in routine clinical practice using CT images: results of 10 convolutional neural networks," Comp. Bio. Med. 121, 2020, pp. 103795:1-103795:9.

[16] M.B. Jamshidi, et al., "Artificial intelligence and COVID-19: deep learning approaches for diagnosis and treatment," IEEE Access 8, 2020, pp. 109581-109595.

[17] B. Robson, "COVID-19 coronavirus spike protein analysis for synthetic vaccines, a peptidomimetic antagonist, and therapeutic drugs, and analysis of a proposed achilles' heel conserved region to minimize probability of escape mutations and drug resistance," Comp. Bio. Med. 121, 2020, pp. 103749:1-103749:28.

[18] L. Caruccio, V. Deufemia, G. Polese, "Visualization of (multimedia) dependencies from big data," MTAP 78(23), 2019, pp. 33151-33167.

[19] P. Le Bras, et al., "Visualising COVID-19 research," 2020, CoRR abs/2005.06380

[20] J. Tu, M. Verhagen, B. Cochran, J. Pustejovsky, "Exploration and discovery of the COVID-19 literature through semantic visualization," 2020, CoRR abs/2007.01800

[21] F. Wolinski, "Visualization of diseases at risk in the COVID-19 literature," 2020, CoRR abs/2005.00848

[22] F. Zuo, et al., "An interactive data visualization and analytics tool to evaluate mobility and sociability trends during COVID-19," in ACM KDD Workshop on UrbComp 2020, pp. 5:1-5:5

[23] S. Zhang, Y. Cai, J. Li, "Visualization of COVID-19 spread based on spread and extinction indexes," Sci. China Inf. Sci. 63(6), 2020, pp. 164102:1-164102:3

[24] World Health Organization (WHO), WHO coronavirus disease (COVID-19) dashboard. https://covid19.who.int/

[25] A. Fariha, et al., "Mining frequent patterns from human interactions in meetings using directed acyclic graphs," in PAKDD 2013, Part I, pp. 38-49.

[26] C.K. Leung, "Frequent itemset mining with constraints," in Encyclopedia of Database Systems, 2nd edn, 2018, pp. 1531-1536.

[27] P. Buono, P. Carella, "Towards secure mobile learning. visual discovery of malware patterns in Android apps," in IV 2019, Part I, pp. 364-369.

[28] K.A. Hassan, et al., "A study on 2D and 3D parallel coordinates for pattern identification in temporal multivariate data," in IV 2019, Part I, pp. 145-150.

[29] M.L. Huang, et al., "Stroke data analysis through a HVN visual mining platform," in IV 2019, Part II, pp. 1-6.

[30] K.E. Barkwell, et al., "Big data visualisation and visual analytics for music data mining," in IV 2018, pp. 235-240.

[31] W. Jentner, D.A. Keim, "Visualization and visual analytic techniques for patterns," in High-Utility Pattern Mining, 2019, pp. 303-337.

[32] C.K. Leung, P.P. Irani, C.L. Carmichael, "FIsViz: a frequent itemset visualizer," in PAKDD 2008, pp. 644-652.

[33] C.K. Leung, C.L. Carmichael, "FpVAT: a visual analytic tool for supporting frequent pattern mining," ACM SIGKDD Explor. 11(2), 2009, pp. 39-48.

[34] C.K. Leung, V.V. Kononov, A.G.M. Pazdor, "PyramidViz: visual analytics and big data visualization of frequent patterns," in IEEE DASC-PICom-DataCom-CyberSciTech 2016, pp. 913-916.

[35] C.K. Leung, F. Jiang, P.P. Irani, "FpMapViz: a space-filling visualization for frequent patterns," in: IEEE ICDM 2011 Workshops, pp. 804-811.

[36] Statistics Canada, Table 13-10-0781-01 detailed preliminary information on confirmed cases of COVID-19 (revised), Public Health Agency of Canada. doi:10.25318/1310078101-eng

[37] A. Cuzzocrea, "Combining multidimensional user models and knowledge representation and management techniques for making web services knowledge-aware," WIAS 4(3), 2006, 289-312.

[38] A. Cuzzocrea, et al., "An innovative framework for supporting big atmospheric data analytics via clustering-based spatio-temporal analysis," JAIHC 10(9), 2019, pp. 3383-3398.

[39] A. Cuzzocrea, et al., "OLAP analysis of multidimensional tweet streams for supporting advanced analytics," in ACM SAC 2016, pp. 992-999.

[40] A. Cuzzocrea, I. Song, "Big graph analytics: the state of the art and future research agenda," in DOLAP 2014, pp. 99-101.