**CIND820 – Capstone Project**

**Literature Review and Data Description and Review**

**Katie Schilling**

**501130072**

**Supervisor: Tamer**

**February 14, 2022**

**Question #1 – What do you already know about this topic.**

Vital events are a key piece of information used when making decisions when it comes to infrastructure, resources, housing, health care etc. Understanding, and being able to determine trends, future population growth and declines, can allow officials properly, or reasonable allocate the resources needed to handle whatever the situation might be. Where we may see an increase in deaths, births or still births, we may find a need for added funds, and resources allocated to hospital and health care sectors overall, whereas if the trend is the opposite, the resources can be allocated elsewhere. Vital event statistics for Births, Stillbirths, deaths and marriages are the responsibility of each province and territory. Working in collaboration with Stats Canada, alongside each province and territory, as well as vital statistics registrars, vital events data has been complied, analyzed and published since 1921. Each statistic sheds light on its ability to trend, forecast and use the data to understand how its impacting public health, and resources. Birth statistics are used to monitor the number of new births, family sizes, as well as for target prevention programs for women who are at a high risk of having a baby with low birth weight, as well as gestation period.

Death statistics while can provide an overall outlook on population trends, it can also be a good statistic to monitor trends in public health, such as a health crisis, infectious diseases, unintentional injuries and suicides.

Common public data in regard to my second dataset, which is covid-19 positive cases, is popular, easy to find, and multi-faceted. On a single search you can find articles that compare vaccinated against unvaccinated, deaths, as well as deaths by vaccination status. Hospitalizations based on vaccination status, ICU admissions, again over all as well as by vaccination status. Age groups for all of these variables. The current case counts as well as predicted best-case and worst-case scenarios. Projected outcomes based on public health measures. Visualizations of data for each year, month and day, as well as even more knowledge regarding the different variants, rate of transmission, positivity rates, and even positivity rates by regions/postal code. There is no shortage of information when it comes to both topics. Almost too much information, making it hard to digest, compare, contrast and fully understand the broader picture. Currently, it does seem as we are through the worst of Covid-19. While the trends show the case counts, and waves getting larger each time, the mortality in each wave, which also seems to coincide with a new variant, seems to be decreasing. It would seem at this point we are in more of a stable position than we were a year or 2 ago.

**Question #2 – What do you have to say critically about what is already know?**

In my opinion, the largest goal of understanding vital events data and trends is to always ensure we are prepared. As we have seen over the past 2 years, with each wave of Covid-19, our hospitals, and other health care settings have been on the brink of collapse. Each wave created a need for increased help, and then each subsequent wave, the current resources were inefficient to keep the system from on the brink of collapse. Had sufficient funding and resources been allocated, and hospital infrastructure increased prior to the arrival of Covid-19, both our hospitals, and our healthcare workers would have been in a better position to manage the public health crisis. Additionally, with increased births, we also see an increased number of resources, primarily funds, being sent to middle- and lower-income families to help them support the financial burden of supporting children. The proper resources allocated for Child Tax Benefit payments, and funding for the subsidized daycare program for all provinces also helps to lower the burden in supporting children. Middle to low-income families often spend a long time on a waiting list for subsidized daycare funding to become available. This means many pay out of pocket or are unable to work due to the cost of childcare, the result is then the income level for those families shrinking, and the burden on the government greater as they rely on funding such as welfare to manage day to day expenses. Recognizing, and addressing an increased need for childcare, or the high rate in which the cost of living is increasing, could have created an opportunity to allocate necessary funds and resources to attempt to offset this issue, prior to it becoming a much larger issue. When looking at the structure and allocation of funds to each province for health care, there is a lot of ups and downs. There are increases, and the reverse course situations, where the total funding going to each province is adjusted and reduced. When looking at the cash component of the CHT (Canada Health Transfer), from 2013-14 to 2016-17 years, there was an annual growth of 6%, however from 2017-18 to date, the annual increase had fluctuated between 3% and 4.6%. Looking at these numbers, and the health crisis and pressure we have put on our healthcare systems and wondering if the trend in vital statistics coincides with the decrease in CHT to provinces for their healthcare infrastructure, or if we were in a position to be able to better prepare our healthcare system by allowing them to increase resources while not in the middle of a health care crisis.

**Question #3 – Has anyone else ever done the exact same thing?**

Looking at the analyses I have found, I have been unable to find an analysis that has been done with the same information. There are many analyses done on the Vital Events data, primarily based on the provinces specific data rather than data for Canada as a whole, and many analyses done on Covid-19 in reference to many different faucets of vitals events, along with many other statistics, there does not seem to be one that encompasses both. A lot of the data used is more enriched in features than the broad information. Many of the features analyzed are breakdowns, or subgroups of specific types of data such as birth weights, marriages, divorces, death by cause, ages, sex, province, city and many more variables. Finding a single article or analysis that looks at the broader picture of the data, prior to breaking it down into smaller sub-groups and analysis was not something I was able to accomplish, nor was I able to find one that took into account the Covid-19 positive cases, or the vital events as a whole alongside any covid-19 data.

**Question #4 – Has anyone else done anything that is related?**

In my research I have found many related analyses. One of which is comparing the Covid-19 pandemic with preterm and stillbirth rates using historical data for stillbirths at term and preterm gestation. Methodology used for the analysis utilized Laney control P` charts, and interrupted time-series analyses for the most recent 4 years. The overall results were that from January 2020- December 2020, the preterm birth rate in Ontario was 7.87% with no special cause variation, resulting in the interpretation that overall, there was no unusual change, overall or by subgroup, during the first 12 months of the pandemic vs the previous 17.5 years. I was also able to find a government produced population projection for the province of Ontario. Within this population projection, it analyzes what the actual growth was, vs. the projected growth and then looks at future growth, all the way up to 2046. Currently, based on the impact of Covid-19, they are projecting the growth to drop from 1.3% for 2019-2020 to 0.6% for 2020-2021, before increasing again to roughly 2.1% for the 2021-2022 year. From that point on, it is projected that the population increase will have a steady growth, reaching 1.0% by 2046. From what I gather based on this information and the historical trend line, it does appear that it will take approximately 25 years for the province of Ontario to get back close to the 1.3% increase in 2019-2022. I’m not entirely sure what economic or housing impact that growth, or lack thereof will have, but with the large proportion of Ontario’s population being within the 15-64 age bracket, we have a primarily working class, younger population, which may lead to a higher increase in population, and considerations may need to be taken into account on the impact of this, with the vital events data for all of Canada, and the allocation of funds and resources to support a potentially fast rate of growth.

**Question #5 – Where does your work fit in with what has gone before?**

From my research, I am seeing a lot of analysis on Covid-19 itself. Analysis on the number of positive cases, number of deaths, hospitalizations, both in ICU and regular hospitalizations. I am also seeing analysis of the number of vaccinated, and unvaccinated. There is a lot of data analysis done around Covid-19 itself. While I think it is important to look at the impact Covid has had on our country as a whole, I also think it’s important to take those numbers into account, as well as the other Vital statistics we track every year, to see how the overall image has changed with the introduction of covid-19, vs where we anticipated seeing these trend lines go over time. While the data may seem daunting to look at individually, I do feel that the larger picture, of comparing covid cases against our other vital stats, it may paint a picture that does not paint as harsh of a picture as just looking at the covid numbers. We’ve put a lot of emphasis on vital statistics, and on pandemic statistics, however, combining the 2, and comparing for causality, trends, forecasting and time series analysis, we can see the actual impact on these stats.

**Question #6 – Why is your research worth doing in the light of what has already been done?**

While researching to see if anyone has completed an analysis the same as mine, or similar, I have found that there are not a lot of articles, papers and analysis to use in comparison. There are a lot of time series datasets, and analysis. There are many options for each province or territory. Many take into account more robust variables such as province/territory, age, sex, weight, cause of death etc. These variables are highly valuable if you want to find trends within each area. Such as deaths. When analyzing deaths by location, age and cause, it allows you to see trends in certain causes of death, which age groups are being most affected, and whether it is affecting men or women more, or both equally. When looking at a specific area of the Vital events dataset, this could be a great asset, however I was unable to find a trend analysis, or time series forecast of the vital events data provided by Stats Canada. It may be that the data itself is too broad, and those who are researching are looking for more refined results, however I do believe looking at the analysis of the broader variables, before breaking them down into subgroups, can help to identify which variable to look at first. If there is a trend that’s alarming, rapidly changing, or even not changing at all like historical data, it may be a clue that there is something going on there, and it warrants a deeper look into what that might be. I feel the analysis I am doing, is beneficial, as in the larger scheme of things it’s important to not only look at the finer details, but to also be able to step back and see the big picture. When looking for research materials, there were a lot of articles that came up with information regarding Covid-19 and deaths. While I think it’s an important statistic, I also think its import to step back, and look at deaths as a whole. While we may have seen a large increase in deaths due to Covid-19, deaths attributed to other factors and health conditions may have decrease, essentially continuing the trend from historical data, rather than seeing a large jump in deaths. From reading articles regarding Covid-19 and deaths, you get this impression that we have our usual annual deaths, and a large spike in deaths from covid, which leads to the assumption that as a whole, we have had a drastic increase in deaths. In my analysis, I will be able to see deaths, compared to the number of covid positive cases, and see if there is a direct correlation between the 2. Determine that if we know the value for one, we can use it to better forecast the other.

**Reference List**

Elias, B. (2014, October 30). *Moving beyond the historical quagmire of measuring infant mortality for the first nations population in Canada*. Social Science & Medicine. Retrieved February 14, 2022, from https://www.sciencedirect.com/science/article/pii/S0277953614007096

Government of Canada, S. C. (2020, April 8). *Frequently asked questions on vital statistics*. Government of Canada, Statistics Canada. Retrieved February 14, 2022, from https://www.statcan.gc.ca/en/about/relevant/vscc/faq

*What now?: Canada health transfer: Background and future: Canada west foundation*. Canada West Foundation |. (2021, September 8). Retrieved February 14, 2022, from https://cwf.ca/research/publications/what-now-canada-health-transfer-background-and-future/

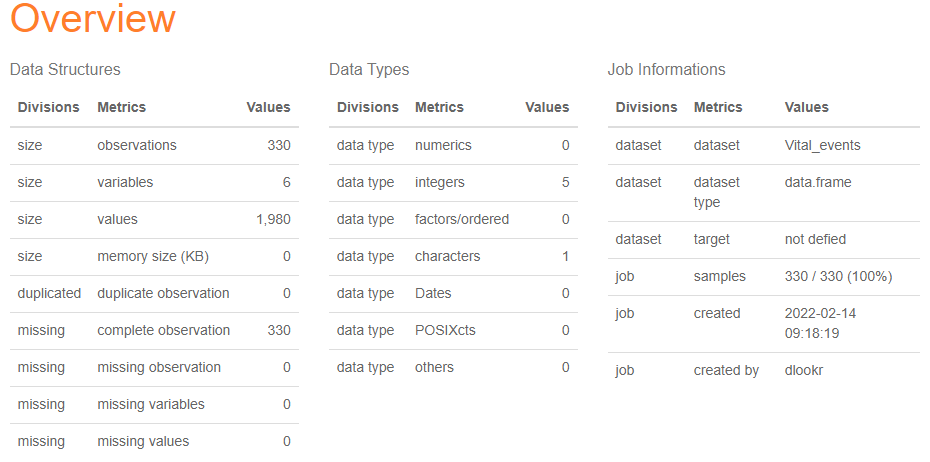
Shah, P. S., Ye, X. Y., Yang, J., & Campitelli, M. A. (2021, August 3). *Preterm birth and stillbirth rates during the COVID-19 pandemic: A population-based Cohort Study*. CMAJ. Retrieved February 14, 2022, from <https://www.cmaj.ca/content/193/30/E1164>

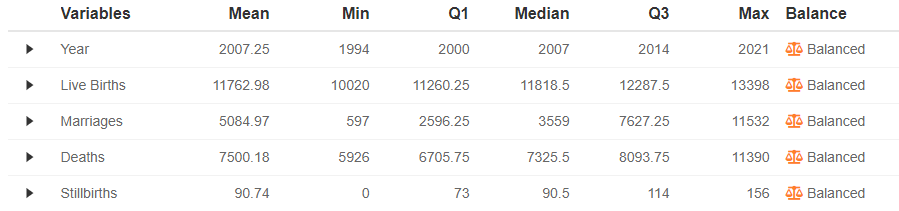
*Ontario population projections*. ontario.ca. (n.d.). Retrieved February 14, 2022, from https://www.ontario.ca/page/ontario-population-projections

Canada, P. H. A. of. (2021, May 28). *Covid-19 Daily Epidemiology update*. Canada.ca. Retrieved February 14, 2022, from https://health-infobase.canada.ca/covid-19/epidemiological-summary-covid-19-cases.html

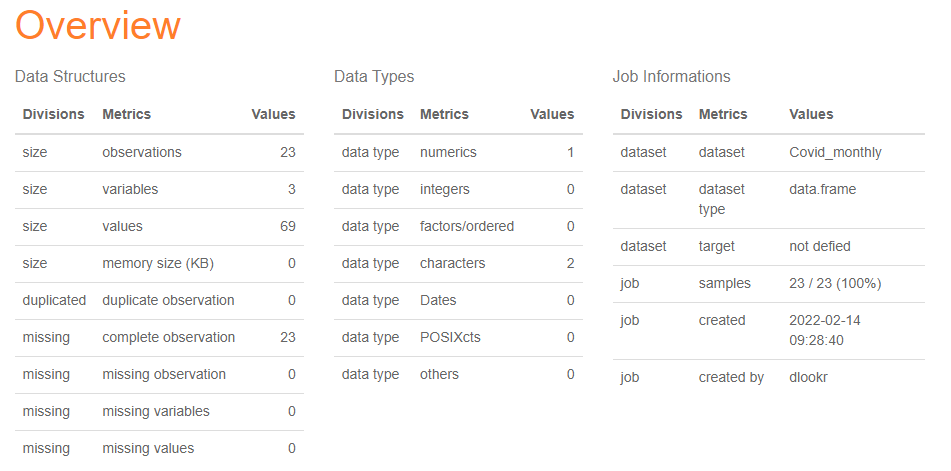
Government of Canada, Statistics Canada. (2021, December 22). *This summary provides an update on the work the agency has and is undertaking using existing and new data sources to provide critical insights on the social and economic impacts of covid-19 on Canadians. it includes data from fall 2021.* Government of Canada, Statistics Canada. Retrieved February 14, 2022, from https://www150.statcan.gc.ca/n1/pub/11-631-x/11-631-x2021003-eng.htm

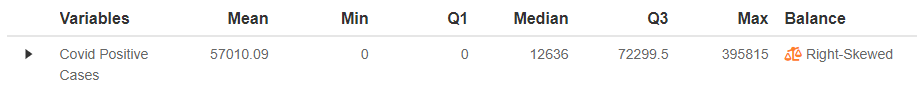
**Descriptive Statistics for Vital Events Dataset**

****



**Descriptive Statistics for Covid Monthly Dataset**

****



**Github repository Link**

[**https://github.com/Kschilling13/CIND820-Capstone-Project/**](https://github.com/Kschilling13/CIND820-Capstone-Project/)