A4: Common Analysis

Covid-19 and Mask Policy in Hartford, Connecticut

**Visualization:** the changes in the derivative function of the rate of infection.

图表, 直方图

描述已自动生成

**Explanation**

Mathematically, the curve of new infection cases per day during the Covid-19 epidemic can be considered as the derivative function of the total number of confirmed cases’ curve. The figure above shows the rate of confirmed infection cases of Covid-19 in Hartford County, Connecticut from January 22, 2020 to November 1, 2021 on a daily basis (0~0.25%). The red area below the curve represents the period during which the state's official mask policy was in effect, from April 20,2020 to August 15, 2021.

(Mask Mandates by County from CDC <https://data.cdc.gov/Policy-Surveillance/U-S-State-and-Territorial-Public-Mask-Mandates-Fro/62d6-pm5i>)

It is clear from the chart that, since the implementation of the mask mandate, the rate of new infections per day has dropped significantly until October 2020. In the following months, however, the daily rate of new infections rebounded rapidly, even though Hartford County still enforced the mask mandate. This may be related to seasonal factors, as studies have shown that cold weather can exacerbate the spread of Covid-19 virus. It may also be related to people's relaxation of vigilance and the influence of holiday season. According to the mask compliance survey data from the *New York Times*, around 75% of the population in this county said they would always wear a mask:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| COUNTYFP | NEVER | RARELY | SOMETIMES | FREQUENTLY | ALWAYS | Population |
| **09003** | **0.015** | **0.023** | **0.065** | **0.13** | **0.767** | **899,498** |

(data source: <https://github.com/nytimes/covid-19-data/blob/master/mask-use/mask-use-by-county.csv>)

The daily rate of new infections has dropped again since April 2021, remaining below 0.03%. On the one hand, I suppose this phenomenon is still related to the warming weather, and on the other hand, it is also because of the promotion of vaccines.

**Reflection**

For this assignment, I learned a lot from the collaborative activities, especially the Slack discussion of classmates. For instance, Aaliyah Hanni posted a question on Slack asking that ‘how are people showing mask-use in their visualization?’. I got a lot of inspiration from that thread, including how to use the background color to represent the period of Mask Mandate. Although my visualization is not exactly the same, classmates’ advice helped me to broaden my mind. In addition, the discussion on Slack also reminded me to pay attention to the rate of infections and not just confirmed cases of Covid-19. Patrick Peng pointed out that the definition of infection rate should be ‘daily\_active\_cases/daily\_pop\_at\_risk’, and our teaching assistant clarified that we could ‘simply the question and assume that population at risk ~= total population’. Thanks to this clarification which save me a lot of time. Sai Muktevi and Stacey Wheeler suggested ‘Convenient\_us\_\_confirmed\_cases.csv’ as the source dataset which also reminded me to explore other tables for this assignment.

Through this assignment, I also reviewed and consolidated the knowledge of pandas and matplotlib in Python. Such as using ‘df.loc.diff()’ to calculate the daily confirmed cases, ‘pd.merge()’ to combine two dataframes, ‘df.dt.strftime('%m/%d/%y')’ to convert datetime format, and ‘plt.fill\_between’ to fill the area between curves, which I used to plot the final visualization of this assignment.