

Homework 5: Interactive Dashboard

This is an individual assignment. You are welcome to discuss the assignment with your colleagues, but the work you turn in must be your own.

The purpose of this assignment is to build an interactive dashboard using Tableau. Submit the Tableau graphics and key findings/messages of each graphic in one PDF document. Moreover, save your dashboard along with all its worksheets as a Tableau packaged workbook (.twbx), rather than a Tableau workbook (.twb), and submit it on Canvas.

This exercise focuses on the phenomenon of *COVID-19 vaccination in the United States in 2021*. With the data provided (see pages 2-3 for details), create an interactive dashboard on the *vaccination in 2021 across regions* (i.e., Northeast, Midwest, South, and West) in the United States. Please note the following requirements:

- (1) The dashboard should include at least one data map (30 points), one time series plot (30 points), one bar chart (30 points), and a data/text table (30 points). If additional plots are needed, you will pick the type of plot appropriate for the intended visualization.
- (2) The dashboard should show at least the following aspects of the phenomenon: daily vaccination, vaccinated population, fully vaccinated population, distributed vaccination, and booster (missing each aspect will lead to a 20-point deduction). Additional aspects are allowed.

The aspects listed here do not necessarily pinpoint the specific variables in the data. You will need to identify the appropriate variables for visualization purpose (ineffective choice of variables will lead to deduction of up to 20 points). For example, while `people_vaccinated_per_hundred` and `people_vaccinated` both capture information related to vaccinated population (one of the aspects listed above), you need to evaluate these two variables and pick the one more suitable for your purpose. When comparing vaccination across regions, you should use normalized variables (e.g., `people_vaccinated_per_hundred`) to make the comparison more meaningful and easier to interpret. Do not rely on tooltips for presenting key information.

- (3) The dashboard should include a Region parameter for users to switch between regions (i.e., Northeast, Midwest, South, and West) in *at least half* of the plots (including data/text table) in the dashboard. For example, if there are five plots in the dashboard, the parameter should apply to at least three plots; if there are four plots in the dashboard, the parameter should apply to at least two plots. (40 points)
- (4) Duplicate one of the plots created for the dashboard and make it as ugly as possible by going against the design principles introduced earlier this semester. List the things you have done to make it ugly. Don't include this plot in your dashboard. (10 points)

Follow the proper design techniques (e.g., a meaningful title for the dashboard, show/hide the legend/title for each plot, effective use of color/size, effective choice of map projection, effective use of labels, effective use of padding and space in the dashboard, etc.) that we have gone over in class.

Data can be downloaded from Canvas.

United States vaccination data

Description: State-by-state data on United States COVID-19 vaccinations, based on the data updated daily by the United States Centers for Disease Control and Prevention

Source: Mathieu, E., Ritchie, H., Ortiz-Ospina, E. et al. A global database of COVID-19 vaccinations. Nat Hum Behav (2021)

Variables:

- date: date of the observation.
- region: based on Census Regions and Divisions of the United States (https://www2.census.gov/geo/pdfs/maps-data/maps/reference/us_regdiv.pdf)
- location: name of the state or federal entity.
- total_vaccinations: total number of doses administered. This is counted as a single dose, and may not equal the total number of people vaccinated, depending on the specific dose regime (e.g., people receive multiple doses). If a person receives one dose of the vaccine, this metric goes up by 1. If they receive a second dose, it goes up by 1 again.
- total_vaccinations_per_hundred: total_vaccinations per 100 people in the total population of the state.
- daily_vaccinations_raw: daily change in the total number of doses administered. It is only calculated for consecutive days. This is a raw measure provided for data checks and transparency, but we strongly recommend that any analysis on daily vaccination rates be conducted using daily_vaccinations instead.
- daily_vaccinations: new doses administered per day (7-day smoothed). For countries that don't report data on a daily basis, we assume that doses changed equally on a daily basis over any periods in which no data was reported. This produces a complete series of daily figures, which is then averaged over a rolling 7-day window.
- daily_vaccinations_per_million: daily_vaccinations per 1,000,000 people in the total population of the state.
- people_vaccinated: total number of people who received at least one vaccine dose. If a person receives the first dose of a 2-dose vaccine, this metric goes up by 1. If they receive the second dose, the metric stays the same.
- people_vaccinated_per_hundred: people_vaccinated per 100 people in the total population of the state.
- people_fully_vaccinated: total number of people who received all doses prescribed by the initial vaccination protocol. If a person receives the first dose of a 2-dose vaccine, this metric stays the same. If they receive the second dose, the metric goes up by 1.
- people_fully_vaccinated_per_hundred: people_fully_vaccinated per 100 people in the total population of the state.
- total_distributed: cumulative counts of COVID-19 vaccine doses recorded as shipped in CDC's Vaccine Tracking System.
- total_distributed_per_hundred: cumulative counts of COVID-19 vaccine doses recorded as shipped in CDC's Vaccine Tracking System per 100 people in the total population of the state.
- share_doses_used: share of vaccination doses administered among those recorded as shipped in CDC's Vaccine Tracking System.
- total_boosters: total number of COVID-19 vaccination booster doses administered (doses administered beyond the number prescribed by the initial vaccination protocol)
- total_boosters_per_hundred: total_boosters per 100 people in the total population.

An example of how the metrics is calculated:

Five people take part in a vaccination program, to be given a vaccine that requires 2 doses to be effective against the disease.

- Dina has received 2 doses, then a 3rd (booster) dose.
- Tommy has received 2 doses.
- Joel has received 1 dose.
- Ellie has not received any dose.

In the data:

- The total number of doses administered (total_vaccinations) will be equal to 6 ($3 + 2 + 1$)
- The total number of people vaccinated (people_vaccinated) will be equal to 3 (Dina, Joel, Tommy)
- The total number of people with a complete initial protocol (people_fully_vaccinated) will be equal to 2 (Dina, Tommy)
- The total number of boosters administered (total_boosters) will be equal to 1 (Dina)