

Robert Cheadle

1.03 Designing a Data Research Project

1. Create a project management plan including all the components listed in the Exercise. Your plan will be relevant for both Achievements 1 and 2 of your course. You can refer to your project brief for more specific information on what is required for each section:

Stakeholder communication:

- Stakeholder Identification
 - Medical agency frontline staff (nurses, physician assistants, and doctors)
 - Hospitals and clinics using the staffing agency's services
 - Influenza patients
 - Staffing agency administrators
- Stakeholder Communication:
 - Meetings (with all stakeholders): At the end of the first week a meeting with the relevant stakeholders stated in the previous section will be held. During the meeting, business requirements will be discussed, and any clarifying, funneling, adjoining, elevating, and privacy- and ethics-related questions will be presented. Finally, the project communication plan will be presented. At the end of week 5 a meeting with relevant stakeholders will be held to review and discuss the interim project report.
 - Calls: During the analysis phase, weekly calls will be held to update stakeholders on the status of the project milestones and to answer any questions they may have.

- Written Communication: Weekly emails with updates on the project progress and topics discussed in weekly call will be provided throughout the duration of the project.
- Emergency/Contingency Plan: Any urgent issues will be communicated via email with a follow-up call scheduled within three days.

Schedule and Milestones:

Week	Milestone	Tasks
1	1.2: Starting with Requirements	<ul style="list-style-type: none"> ● Create a list of the data questions you need to answer for your analysis.
1	1.3: Designing a Data Research Project	<ul style="list-style-type: none"> ● Design your data research project. ● Formulate a research hypothesis. ● Present Project Management Plan at meeting with stakeholders.
2	1.4: Sourcing the Right Data	<ul style="list-style-type: none"> ● Describe the data sets you have access to for your project. ● Explain the relevance and limitations of each data set to your project.
2	1.5: Data Profiling & Integrity	<ul style="list-style-type: none"> ● Create a data profile for each of the data sets in your analysis. ● Include information on data types, data integrity issues (accuracy and consistency), any

		cleaning you conducted, as well as summary statistics in each profile.
3	1.6: Data Quality Measures	<ul style="list-style-type: none"> ● Implement additional data quality measures to your data profiles related to completeness, uniqueness, and timeliness.
3	1.7: Data Transformation & Integration	<ul style="list-style-type: none"> ● Integrate data from two sources into one cohesive data set using data transformations.
4	1.8: Conducting Statistical Analyses	<ul style="list-style-type: none"> ● Calculate the variance and standard deviation for key variables. ● Identify variables with a potential relationship and test for a correlation.
4	1.9: Statistical Hypothesis Testing	<ul style="list-style-type: none"> ● Formulate a statistical hypothesis regarding an outcome of interest around two groups in your data. ● Conduct hypothesis testing and interpret the results.
5	1.10: Consolidating Analytical Insights	<ul style="list-style-type: none"> ● Create an interim report consolidating the findings of your analysis. ● Meet with relevant stakeholders to discuss interim report.
6	2.1: Intro to Data Visualization	<ul style="list-style-type: none"> ● Explain how data visualizations can be used in your project. ● Install Tableau.

6	2.2: Visual Design Basics & Tableau	<ul style="list-style-type: none"> ● Create a data visualization design checklist. ● Explain how the visualizations in each example can be improved. ● Connect your project data to Tableau.
7	2.3: Composition & Comparison Charts	<ul style="list-style-type: none"> ● Create a pie, bar, or column chart, as well as a tree map in Tableau. ● Use your visualization design checklist to design your charts.
7	2.4: Temporal Visualizations & Forecasting	<ul style="list-style-type: none"> ● Create a time forecast for a variable and display it in Tableau. ● Use your visualization design checklist to design your chart.
8	2.5: Statistical Visualizations: Histograms & Box Plots	<ul style="list-style-type: none"> ● Create visualizations that look at the distribution of a variable. ● Use your visualization design checklist to design your charts.
8	2.6: Statistical Visualizations: Scatter Plots & Bubble Charts	<ul style="list-style-type: none"> ● Create visualizations that look at the correlation between variables. ● Use your visualization design checklist to design your chart.
9	2.7: Spatial Analysis	<ul style="list-style-type: none"> ● Map a variable and justify your spatial visualization choice (heat, density, or choropleth).

		<ul style="list-style-type: none"> ● Use your visualization design checklist to design your chart.
9	2.8: Textual Analysis	<ul style="list-style-type: none"> ● Create a word cloud using qualitative data. ● Use your visualization design checklist to design your chart.
10	2.9: Storytelling with Data Presentations	<ul style="list-style-type: none"> ● Create a narrative to communicate your research findings and insights in relation to your research goals. ● Publish your analysis as a Tableau Storyboard.
10	2.10: Presenting Findings to Stakeholders	<ul style="list-style-type: none"> ● Record a video presentation for your stakeholders.

Project deliverables:

- At the end of week 5 an interim report consolidating the findings of the analysis will be created and presented to relevant stakeholders.
- At the end of week 10 a video presentation will be completed and provided to all relevant stakeholders.

Audience definition:

- Who is your audience?
 - Staffing Agency Administrators
 - Administration staff at hospitals and clinics using the staffing agency's services
 - Medical agency frontline staff (nurses, physician assistants, and doctors)

- Influenza patients
- Who will you be presenting to?
 - Staffing Agency Administrators
 - Administration staff at hospitals and clinics using the staffing agency's services
 - Medical agency frontline staff (nurses, physician assistants, and doctors)
 - Influenza patients
- Who'll be reading your report?
 - Staffing Agency Administrators
 - Administration staff at hospitals and clinics using the staffing agency's services
- Who'll be using your dashboard?
 - Staffing Agency Administrators

2. Refer to the “Context” section of your project brief to begin fielding some answers to your questions from Exercise 1.2. You can also conduct some independent research:

- [Background reading about hospital planning for influenza](#)
- [Background reading about influenza](#)

Who is at a higher risk being hospitalized due to flu complications?

Following is a list of all the health and age factors that are known to increase a person's risk of getting serious complications from flu:

- Adults 65 years and older
- Children younger than 2 years old¹

- Asthma
- Neurologic and neurodevelopment conditions
- Blood disorders (such as sickle cell disease)
- Chronic lung disease (such as chronic obstructive pulmonary disease [COPD] and cystic fibrosis)
- Endocrine disorders (such as diabetes mellitus)
- Heart disease (such as congenital heart disease, congestive heart failure and coronary artery disease)
- Kidney diseases
- Liver disorders
- Metabolic disorders (such as inherited metabolic disorders and mitochondrial disorders)
- People who are obese with a body mass index [BMI] of 40 or higher
- People younger than 19 years old on long-term aspirin- or salicylate-containing medications.
- People with a weakened immune system due to disease (such as people with HIV or AIDS, or some cancers such as leukemia) or medications (such as those receiving chemotherapy or radiation treatment for cancer, or persons with chronic conditions requiring chronic corticosteroids or other drugs that suppress the immune system)
- People who have had a stroke

Other people at higher risk from flu:

- Pregnant people and people up to 2 weeks after the end of pregnancy
- People who live in nursing homes and other long-term care facilities

- People from certain racial and ethnic minority groups are at increased risk for hospitalization with flu, including non-Hispanic Black persons, Hispanic or Latino persons, and American Indian or Alaska Native persons
- ¹ Although all children younger than 5 years old are considered at higher risk of serious flu complications, the highest risk is for those younger than 2 years old, with the highest hospitalization and death rates among infants younger than 6 months old.





































































Source: <https://www.cdc.gov/flu/highrisk/index.htm>

Which states are most affected by influenza?

YEAR

2020 ▾

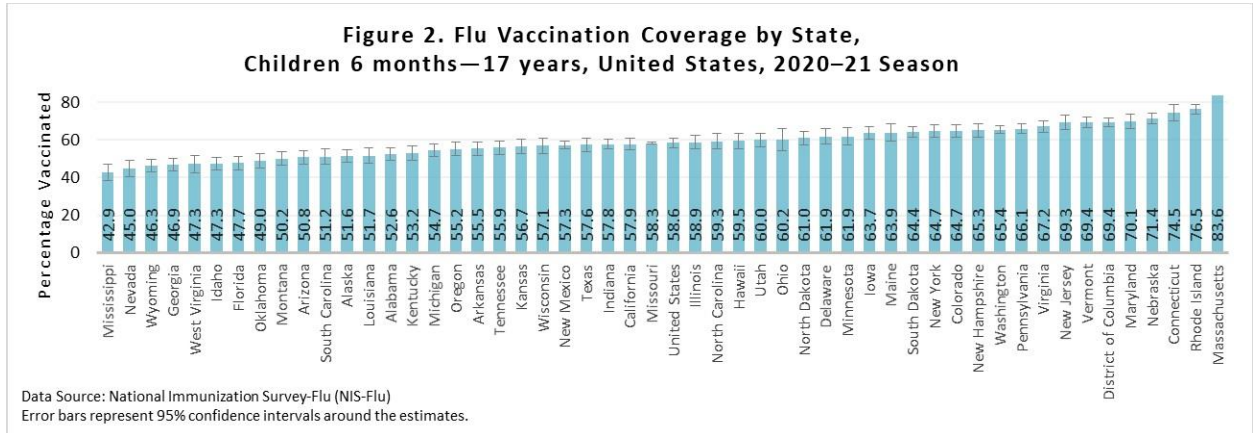
Data Table

Location	Death Rate (Click for...	Deaths
 Alabama 	17.7	1,114
 Alaska 	8	52
 Arizona 	11.6	1,113
 Arkansas 	18.1	691
 California 	13.2	6,062
 Colorado 	7.3	448
 Connecticut 	10.3	527
 Delaware 	11.6	156
 Florida 	9.5	3,191
 Georgia 	13.8	1,575
 Hawaii 	11.1	241
 Idaho 	8.6	179
 Illinois 	15.4	2,428
 Indiana 	12.6	1,021
 Iowa 	12.4	538
 Kansas 	13.8	517
 Kentucky 	16.3	898
 Louisiana 	14.2	778
 Maine 	11.5	237
 Maryland 	12.3	907
 Massachusetts 	14.5	1,331
 Michigan 	14.4	1,886
 Minnesota 	7.9	559
 Mississippi 	25.1	904
 Missouri 	14.3	1,167
 Montana 	8	120
 Nebraska 	12.1	298
 Nevada 	13.6	494
 New Hampshire 	9.6	182
 New Jersey 	14.1	1,638
 New Mexico 	13.4	356
 New York 	18.2	4,756
 North Carolina 	14	1,804
 North Dakota 	15.2	147
 Ohio 	13.3	2,043
 Oklahoma 	13.9	657
 Oregon 	7.1	390
 Pennsylvania 	12.9	2,368
 Rhode Island 	9.3	142
 South Carolina 	11.6	755
 South Dakota 	12.3	143
 Tennessee 	18.5	1,549
 Texas 	12.2	3,541
 Utah 	10.1	280
 Vermont 	6.2	56
 Virginia 	11.5	1,177
 Washington 	8.7	778
 West Virginia 	17.9	447
 Wisconsin 	9.7	733
 Wyoming 	11.6	81

The tables display the mortality rate and total deaths for each state due to Influenza/Pneumonia in 2020

Source: [Stats of the States - Influenza/Pneumonia Mortality \(cdc.gov\)](https://www.cdc.gov/nchs/data/flu/pneumonia_mortality/2020/2020_influenza_pneumonia_mortality_tables.html)

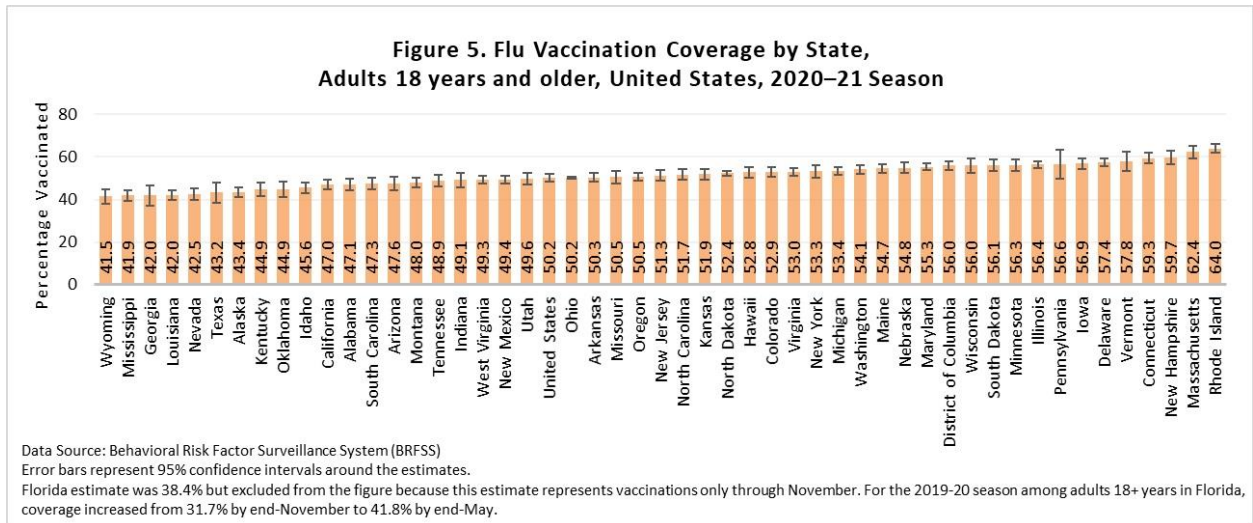
What percent of the population is vaccinated in each state?



The visual displays flu vaccination coverage by state for children 6 months – 17 years old during the 2020-2021 flu season.

Source: [Flu Vaccination Coverage, United States, 2020–21 Influenza Season | FluVaxView |](https://www.cdc.gov/flu/vax-view/)

[Seasonal Influenza \(Flu\) | CDC](https://www.cdc.gov/flu/)



The visual displays flu vaccination coverage by state for children 6 months – 17 years old during the 2020-2021 flu season.

Source: [Flu Vaccination Coverage, United States, 2020–21 Influenza Season | FluVaxView | Seasonal Influenza \(Flu\) | CDC](#)

I have noticed that a lot of data is already available on the topic of Influenza in the US. This has allowed me to gain a better understanding of which questions will drive my analysis.

3. Look through the data sets in the project brief and write out 3 to 5 hypotheses based on the variables in your data sets. You can use the examples in the task description to help you.

- If a person is less than or equal to 4 years old or greater than or equal to 65 years old than they are more likely to die from the flu.
- The flu season in the US is known to peak between December and February. If the peak flu season is from December to February than mortalities will increase across all states from December to February.
- If a state has a higher number of citizens considered vulnerable to complications from the flu than that state will have a greater proportion of deaths due to flu.
- If a state has a higher percentage of its population vaccinated than that state will have less instances of visits to hospitals or clinics due to flu.

4. Create a data wish list for your age-influenza mortality hypothesis. You'll learn more about sourcing data in the next Exercise, but for now, simply focus on determining what data you'll need to test your hypothesis.

- I would like data on vaccination numbers across all states.

- I would like data displaying vaccination rates across different age groups by state.
- I would like data showing healthcare worker vaccination rates by state.
- I would like data showing total number of clinical visits and hospitalizations annually due to flu.
- I would like data showing staffing agency employee residency information.