StatsVille Game Instructions

Go to the web site: https://stat2games.sites.grinnell.edu and select the **StatsVille** tab.

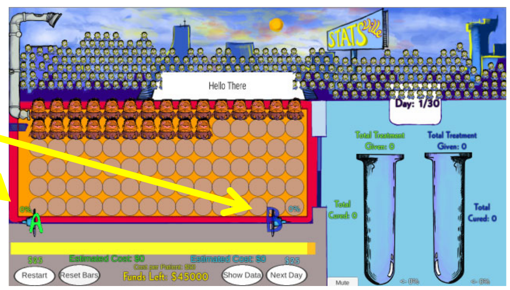
- Watch the **How to Play StatsVille** video.
- Click **Play StatsVille** (Note: *This site may take a few minutes to load.*)

Enter a **Player ID** and a **Group ID**. Any combination of alpha-numeric characters will work. *Note that this ID will be public on the web*.

Your task is to cure the creatures of StatsVille. Sick patients enter the clinic and you provide each patient with one of two potential treatments that are currently being tested, **A**lignItUp and **B**eHealthy. Often these are simply referred to as Treatment A and Treatment B, respectively. Both have shown some evidence of success when compared against a placebo. However, more testing is needed to determine which is most effective. Your task is to determine the best treatment strategy for the StatsVille community.

1) Select **Level 1**, check the **Show p-values** tab, and then click **Start**. In each level, 300 creatures live in StatsVille, and every day, the patients entering the clinic can be treated as a simple random sample from this city.

Goal: Move the A and B sliders to provide the most effective treatment combination to the patients in the clinic. Click the **Next Day** button to apply the treatment and record your results for each day.



2) Rules for the game:

a. **You win** when no more than 5 patients enter the hospital on a particular day, indicating you have controlled the spread of the epidemic.

b. You have a limited amount of funds. If you run out of money, you will lose.
   - Each treatment costs money.
   - Each time a patient enters the hospital, it costs money.

c. If too many patients catch the disease you will lose.

d. If you do not control the spread of the disease in less than 30 days, you will lose.

3) Administer a treatment to the patients that enter your clinic. Determine how many patients you wish to treat each day, then click "Next Day".

a. You can give each patient that enters the clinic Treatment A, Treatment B, or No Treatment.

b. If a patient is cured, they change color and will not get sick again.

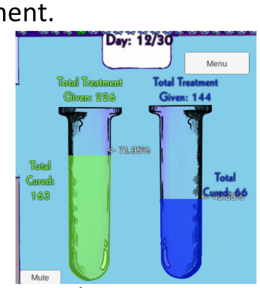
c. The test tubes show summary data for the overall effectiveness of each Treatment..

d. All patients are similar in the way they respond to each treatment. However, there is natural variation from day to day. For example, if Treatment A is 50% effective in this population, on a particular day 40% or 60% of patients may be cured.

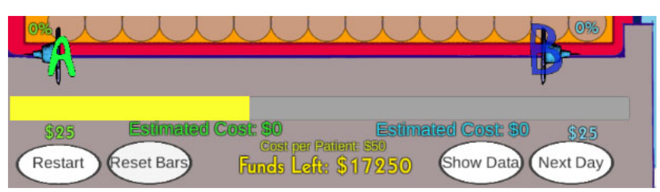
e. A patient can have no more than one treatment a day.

f. If a patient is not cured, they will return to the clinic the next day. Previous treatments have no impact on the current day's effectiveness of each treatment. In other words, time and previous treatments are uncorrelated with the effectiveness of a treatment.

g. Neither drug is expected to have complications as they are both made completely of natural ingredients.



4) The **Show Data** button provides daily statistics (including a p-value) for a two-proportion/chi-square test. The costs and effectiveness of each treatment changes for each level of the game.



Additional Instructor Notes:

Discussion Ideas after Playing the Game:

1. Read about the ASA's statement on p-values at https://amstat.tandfonline.com/doi/pdf/10.1080/00031305.2016.1154108?needAccess=true. Note you can download the pdf for free. Pick one of the six principles and discuss how it relates to this StatsVille activity. More than one of the principles can apply to this lab.
2. Address the challenge of multiple hypothesis tests conducted on the same population.
   - The game is designed so that each day is independent of other days. The effectiveness of Treatment A and Treatment B do not depend on the day. However, what cautions need to be considered when multiple tests are conducted from the same population?
   - While day-to-day samples are independent, explain why using cumulative samples each day is not independent.
3. Discuss the use of graphs and statistics
   - Are counts (number cured) or percentages (percentage cured) a better way to evaluate a treatment's effectiveness?
4. Should a two-proportion test be used with this data? Could we use a Chi-Square test or t-test instead?
5. There is daily variability in the effectiveness of each treatment. Even though there is variation between days, an overall pattern exists that can be modeled with a binomial distribution. Notice that the percentage of effectiveness stabilizes with larger samples. For example, assume n is the number treated and p is the effectiveness of a treatment
   - If $X \sim B(n = 5 \text{ and } p = 0.6)$, what are the most likely outcomes (values of X)? What are unlikely values of X?
   - Similarly if $X \sim B(n = 500 \text{ and } p = 0.6)$, what are the most likely outcomes (values of X)? What are unlikely values of X?
   - Create confidence intervals for $\hat{p}$ using the two examples above. Explain how sample size influences the width of these intervals.
   - This game is designed so that each daily sample of patients can be treated as independent samples.

Statistical Modeling: The spread of the disease follows a basic SIR (susceptible, infected, and recovered) differential equation model. We then use an underlying binomial model to simulate natural variability in the cure rate. Knowledge or discussion of the SIR model is not needed in any part of this lab. However, it may be useful to understand the binomial distribution.

Students quickly see that the cure rate for each treatment appears to fluctuate and small samples tend to produce unreliable results (i.e. results that do not necessarily accurately represent the entire population). For example, if the effectiveness of Treatment A is 70%, but only 5 patients are treated, it is likely that 2, 3, 4, or 5 of the patients will be cured. However, if 100 patients are treated, the law of large numbers ensures that the outcome (percent cured) will be much closer to the true population parameter, so the percent cured is likely to be between 66% and 74%.

Multiple levels of the game are developed by varying model parameters, such as the costs and effectiveness of each treatment. In the initial stages, the game is fairly easy to win, but at each level the game gradually advances in difficulty and requires students to build upon strategies used in previous levels.