UFO Explorer Tutorial

Matt Thompson

12/3/2021

UFO Explorer Tutorial

Table of Contents

1. About
2. Requirements
3. How to run software
4. Features
- 4.1. Data
- 4.2. Filters
5. Descriptive Statistics Tab
- 5.1 Descriptive Statistics
- 5.2 Summary Statistics
6. Visualization Tab
- 6.1 Chart
- 6.2 Scatter Plot
- 6.3 Interactive Map
7. Statistical Analysis Tab
- 7.1 Wilcoxon Rank Sum Test
- 7.2 Simple Linear Regression
- 7.3 Frequency Comparison
8. Contact

About

UFO Explorer is an interactive web based app that allows users to explore data from reported UFO sightings in North America.

Requirements

- R 4.0.1 or later
- Shiny 1.2.0 or later

To Run Software

If you already have R on your computer you can un-comment and run the below code chunk.

```
#install.packages("shiny")
#library(shiny)
#shiny::runGitHub("ufo-explorer", "blunderfist", ref = "main")
```

If you wish to visit the app using the web you can visit https://matt2021.shinyapps.io/UFO-Explorer/.

Edit: The app is currently not running on shinyapps.io due to memory limits being exceeded. The only option is to run on your machine using the code chunk above.

Data

The data consists of 12 variables all of which are pretty self explanatory. One that may not be immediately clear is duration_s. This is the duration of the sighting measured in seconds. The choice to use seconds was made because that was the most granular level available.

Some of the shapes may seem to be similar in description. These were left as originally reported to not maintain the original data. To learn more about how the data was modified for use in the app see ufo.Rmd.

Features

The following explains the features found in the app and how they work.

Filters

Filters are found on most pages. They allow the user to filter only the data they wish to view. Very specific filtering will likely result in few or zero results. In some cases it will result in an error being displayed. If this happens simply broaden the filter and try again.

Descriptive Statistics Tab

Descriptive Statistics

Displays interesting descriptive statistics for each variable. User can select a single variable, or a specific location. Information displayed will vary based on chosen variable. Filters can be applied to limit results.

For a single variable:

1. Select “Single Variable” radio button.
2. Apply any filters.
3. Click the **Update** button.

For a single variable:

1. Select “Location Specific” radio button.
2. Select the country, state, or city depending on the detail desired.
3. Apply any filters.
4. Click the **Update** button.

Some of the information includes measures of least and most for different categories. Ties are common especially when measuring least. Usually there are multiple results showing 1 as the measure. These ties are not shown together, and are broken by whichever sorting method was used on that particular operation.

Summary Statistics

Shows simple summary statistics for each variable in the data set. Only one variable can be selected at a time and will be instantly updated when selected. Filters are not available for this page.

Visualization Tab

Chart

User can select either a bar chart or a histogram. The bar chart allows the user to plot all variable types, whereas the histogram is restricted to only numeric variables.

Bar Chart

Displays a bar chart showing selected variable.

Steps:

1. Select the **bar chart** radio button.
2. Choose a variable for the x axis.
3. Apply any filters.
4. Click **Plot**.

Some options for the x axis will result in unreadable, overlapping labels on the x axis. One example is choosing **city**. If the variable selected is the same as one of the filter selections the results will not be interesting (ex: selecting state as the variable and filtering to a specific state will show just one bar).

Histogram

Displays a histogram showing selected variable.

Steps:

1. Select the **histogram** radio button.
2. Choose a variable for the x axis.
3. Choose the number of bins and desired bin width.
4. Apply any filters.
5. Click **Plot**.

The number of bins available was chosen to be enough to show for each year.

Scatter Plot

Displays a scatter plot for selected variables. Additional options include using colors and size to aid in categorizing variables on the plot. If **Jitter** is checked the plot will use jitter to add random noise to each point to reduce overlapping points.

Steps:

1. Choose a variable for the x axis.
2. Choose a variabel for the y axis.
3. Choose a color (optional).
4. Choose a size (optional).
5. Check jitter or leave blank.
6. Apply filters.
7. Click **Plot**.

Some options for the x or y will result in unreadable, overlapping labels on the respective axis. One example is choosing **city**.

Interactive Map

Displays a map showing blue points for the location of each report. Overlap can be a problem and filters should be used to improve readability. After selection is made click **Update** to view new map.

If the state selection is blank and not set to “All” the result will be a blank map. This can be an issue when switching from one country to another. The state selection does not automatically reset to “All” and must be manually set.

Statistical Analysis Tab

Wilcoxon Rank Sum Test

Allows user to compare the amount of reports between two geographic locations using the Wilcoxon rank sum test. Filters can be applied.

Steps:

1. Choose the level of detail for the locations (Country, State, or City).
2. Choose the two locations from the drop down menu.
3. Select either the **year** or **month** radio button.
4. Apply any filters.
5. Click **Update** to view results.
6. If the **Box Plot** box is checked, a box plot of the results will be displayed beneath the results table.

The results of the test will be displayed in a table and a brief summary of text is displayed beneath the table. The p-value is used to determine whether the two locations have a similar number of reports or not.

Simple Linear Regression

Displays a linear regression plot, model, and simple interpretation for a selected time period and reported shape.

Steps:

1. Choose a shape and unit of time from the drop down menu.
2. Apply any filters.
3. Click **Update**.

Results:

The results are plotted with a regression line. A table of the results is displayed beneath the plot, and beneath the table is a short, simple summary. Depending on the variables chosen and the filters applied there may not be enough data to show a useful results.

Note: The data does not meet the assumptions necessary for linear regression, but this tab was included non the less.

Note: There is currently a bug related to selecting a state. The state drop down selection is fixed to “All” until I have time to fix this issue.

Comparing Frequencies of Reports

Shows the frequency of reports between either two or three states over a given time period.

Steps:

1. Choose at least two but up to three states to compare.
2. Choose a unit of time for comparison during.
3. Apply filters.
4. Click **Update**.

Results:

The results will show the number of reports for each state over the time period. For example, if **hour** is chosen, the lines will show the number of reports on each hour of the day, allowing the user to compare the number of reported sightings between up to three states over the course of a day. If **day** is chosen, the results will show the number of reports daily over the course of a month. Using filters can find specific time periods.

Example:

Compare daily UFO reports between Florida and California during the month of January during 2020.

1. Select FL for first state.
2. Select CA for second state.
3. Select day for time period.
4. Filter months to only show 1 (Jan).
5. Filter years to only show 2020.
6. Click **Update**.

If no filters are selected the results will include the entire data set.

Contact

Comments are welcome. If you notice something that needs improvement reach out and let me know.

If you wish to use any portion of this app as part of your own project or work please give the proper credit and citation.