The "Data Science" Specialization   Learn More (https://www.coursera.org/specialization/jhudatascience/1?   ✕
(https://learner.coursera.help/hc/requests/new)   utm_source=spark&utm_medium=banner)

Peer Assessments (https://class.coursera.org/repdata-033/human_grading/) / Peer Assessment 2

Help Center (https://accounts.coursera.org/i/zendesk/courserahelp?return_to=https://learner.coursera.help/hc)

---

due in 1wk 4d

Submission Phase

1. Do assignment ☐ (/repdata-033/human_grading/view/courses/975146/assessments/4/submissions)

Evaluation Phase

2. Evaluate peers 🔒 (/repdata-033/human_grading/view/courses/975146/assessments/4/peerGradingSets)

Results Phase

3. See results 🔒 (/repdata-033/human_grading/view/courses/975146/assessments/4/results/mine)

☐ In accordance with the Honor Code, I certify that my answers here are my own work, and that I have appropriately acknowledged all external sources (if any) that were used in this work.

Save draft   Submit for grading

Introduction

Storms and other severe weather events can cause both public health and economic problems for communities and municipalities. Many severe events can result in fatalities, injuries, and property damage, and preventing such outcomes to the extent possible is a key concern.

This project involves exploring the U.S. National Oceanic and Atmospheric Administration's (NOAA) storm database. This database tracks characteristics of major storms and weather events in the United States, including when and where they occur, as well as estimates of any fatalities, injuries, and property damage.

Data

The data for this assignment come in the form of a comma-separated-value file compressed via the bzip2 algorithm to reduce its size. You can download the file from the course web site:

- Storm Data (https://d396qusza40orc.cloudfront.net/repdata%2Fdata%2FStormData.csv.bz2) [47Mb]

There is also some documentation of the database available. Here you will find how some of the variables are constructed/defined.

- National Weather Service Storm Data Documentation (https://d396qusza40orc.cloudfront.net/repdata%2Fpeer2_doc%2Fpd01016005curr.pdf)
- National Climatic Data Center Storm Events FAQ (https://d396qusza40orc.cloudfront.net/repdata%2Fpeer2_doc%2FNCDC%20Storm%20Events-FAQ%20Page.pdf)

The events in the database start in the year 1950 and end in November 2011. In the earlier years of the database there are generally fewer events recorded, most likely due to a lack of good records. More recent years should be considered more complete.

Assignment

The basic goal of this assignment is to explore the NOAA Storm Database and answer some basic questions about severe weather events. You must use the database to answer the questions below and show the code for your entire analysis. Your analysis can consist of tables, figures, or other summaries. You may use any R package you want to support your analysis.

Questions

Your data analysis must address the following questions:

1. Across the United States, which types of events (as indicated in the `EVTYPE` variable) are most harmful with respect to population health?

2. Across the United States, which types of events have the greatest economic consequences?

Consider writing your report as if it were to be read by a government or municipal manager who might be responsible for preparing for severe weather events and will need to prioritize resources for different types of events. However, there is no need to make any specific recommendations in your report.

Requirements

For this assignment you will need some specific tools

- RStudio: You will need RStudio to publish your completed analysis document to RPubs. You can also use RStudio to edit/write your analysis.

- knitr: You will need the knitr package in order to compile your R Markdown document and convert it to HTML

Document Layout

- Language: Your document should be written in English.

- Title: Your document should have a title that **briefly** summarizes your data analysis

- Synopsis: Immediately after the title, there should be a **synopsis** which describes and summarizes your analysis in at **most 10 complete sentences**.

- There should be a section titled **Data Processing** which describes (in words and code) how the data were loaded into R and processed for analysis. In particular, your analysis *must* start from the raw CSV file containing the data. You cannot do any preprocessing outside the document. If preprocessing is

time-consuming you may consider using the `cache = TRUE` option for certain code chunks.

- There should be a section titled **Results** in which your results are presented.

- You may have other sections in your analysis, but Data Processing and Results are **required**.

- The analysis document must have **at least one figure containing a plot**.

- Your analyis must have **no more than three figures**. Figures may have multiple plots in them (i.e. panel plots), but there cannot be more than three figures total.

- You must show all your code for the work in your analysis document. This may make the document a bit verbose, but that is okay. In general, you should ensure that `echo = TRUE` for every code chunk (this is the default setting in knitr).

Publishing Your Analysis

For this assignment you will need to publish your analysis on RPubs.com (http://rpubs.com). If you do not already have an account, then you will have to create a new account. After you have completed writing your analysis in RStudio, you can publish it to RPubs by doing the following:

1. In RStudio, make sure your R Markdown document (`.Rmd`) document is loaded in the editor

2. Click the `Knit HTML` button in the doc toolbar to preview your document.

3. In the preview window, click the `Publish` button.

Once your document is published to RPubs, you should get a unique URL to that document. **Make a note of this URL** as you will need it to submit your assignment.

NOTE: If you are having trouble connecting with RPubs due to proxy-related or other issues, you can upload your final analysis document file as a PDF to Coursera instead.

Submitting Your Assignment

In order to submit this assignment, you must copy the RPubs URL for your completed data analysis document in to the peer assessment question.

Please submit the URL from RPubs that points to your full report for this assignment.

NOTE: The URL for the RPubs document should begin with http:// (not http**s**://)

NOTE: If you are having trouble connecting with RPubs due to proxy-related or other issues, you can upload your final analysis document file as a PDF to Coursera instead.

B	*I*	☰	☰	% Link	<code>	Math		Edit: Rich ▼	Preview

Attach a file | *(supports: txt, png, jpg, gif, pdf)*

Evaluation/feedback on the above work

Note: this section can only be filled out during the evaluation phase.

Has either a (1) valid RPubs URL pointing to a data analysis document for this assignment been submitted; or (2) a complete PDF file presenting the data analysis been uploaded?

| ⇕ |

Overall evaluation/feedback

Note: this section can only be filled out during the evaluation phase.
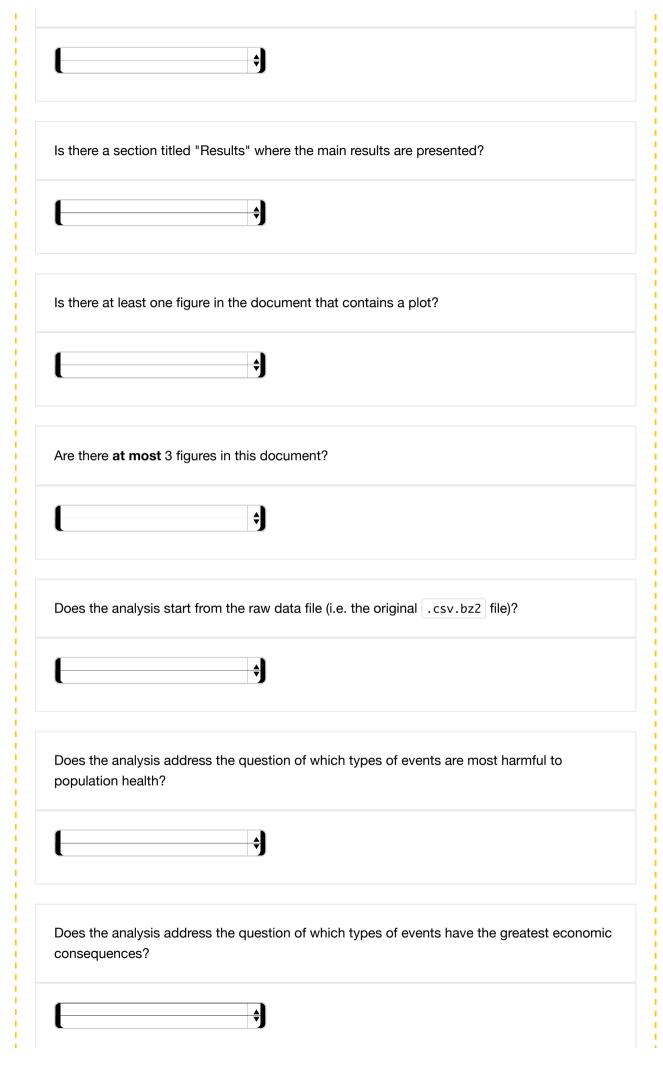
Is the document written in English?

| ⇕ |

Does the document have a title that briefly summarizes the data analysis?

| ⇕ |

Does the document have a synopsis that describes and summarizes the data analysis in less than 10 sentences?

| ⇕ |

Is there a section titled "Data Processing" that describes how the data were loaded into R and processed for analysis?

[▲▼]

Is there a section titled "Results" where the main results are presented?

[▲▼]

Is there at least one figure in the document that contains a plot?

[▲▼]

Are there **at most** 3 figures in this document?

[▲▼]

Does the analysis start from the raw data file (i.e. the original `.csv.bz2` file)?

[▲▼]

Does the analysis address the question of which types of events are most harmful to population health?

[▲▼]

Does the analysis address the question of which types of events have the greatest economic consequences?

[▲▼]

Do all the results of the analysis (i.e. figures, tables, numerical summaries) appear to be reproducible?

[▼]

Do the figure(s) have descriptive captions (i.e. there is a description near the figure of what is happening in the figure)?

[▼]

Does the analysis include description and justification for any data transformations?

[▼]

As far as you can determine, does it appear that the work submitted for this project is the work of the student who submitted it?

[▼]

Use this space to provide constructive feedback to the student who submitted the work. Point out both strengths and weaknesses in the submission and provide advice about how the work could be improved in the future.

[   ]

You've written 0 words

☐ In accordance with the Honor Code, I certify that my answers here are my own work, and that I have appropriately acknowledged all external sources (if any) that were used in this work.

Save draft

Submit for grading

Save draft

Submit for grading