



ANALYZING HAUNTED PLACES USING R

A Library Trick {R} Treat



SESSION STRUCTURE

1. Slides
2. Demo
 1. Aggregation
 2. Data visualization
 3. Data cleaning
 4. Text analysis
3. Design or conduct your own analysis



HAUNTED LOCATIONS DATA

The Shadowlands Haunted Places index:

<http://www.theshadowlands.net/places/>

- Data collection method: crowd sourcing
- Data access point: <https://data.world/timothyrenner/haunted-places>
- Data extraction methods and cleaning methods:
<https://github.com/timothyrenner/shadowlands-haunted-places>

ANALYSIS DESIGN EXERCISE

- What other data cleaning opportunities do you see in this dataset?
- Which states seem have the most violent hauntings?
- What are common words that appear in the Walmart hauntings?
- What are the most haunted cities within a particular state?
- Which state has the most haunted universities?
- Or, develop your own research question

WHICH STATES SEEM TO HAVE THE MOST VIOLENT HAUNTINGS

1. Identify how you will qualify violence
2. Quantify the occurrence of violence
 - Is violence a binary or scaled?
3. Determine the level of aggregation you would like to perform
 - Are you analyzing violence per location, per state, or per region of the country?