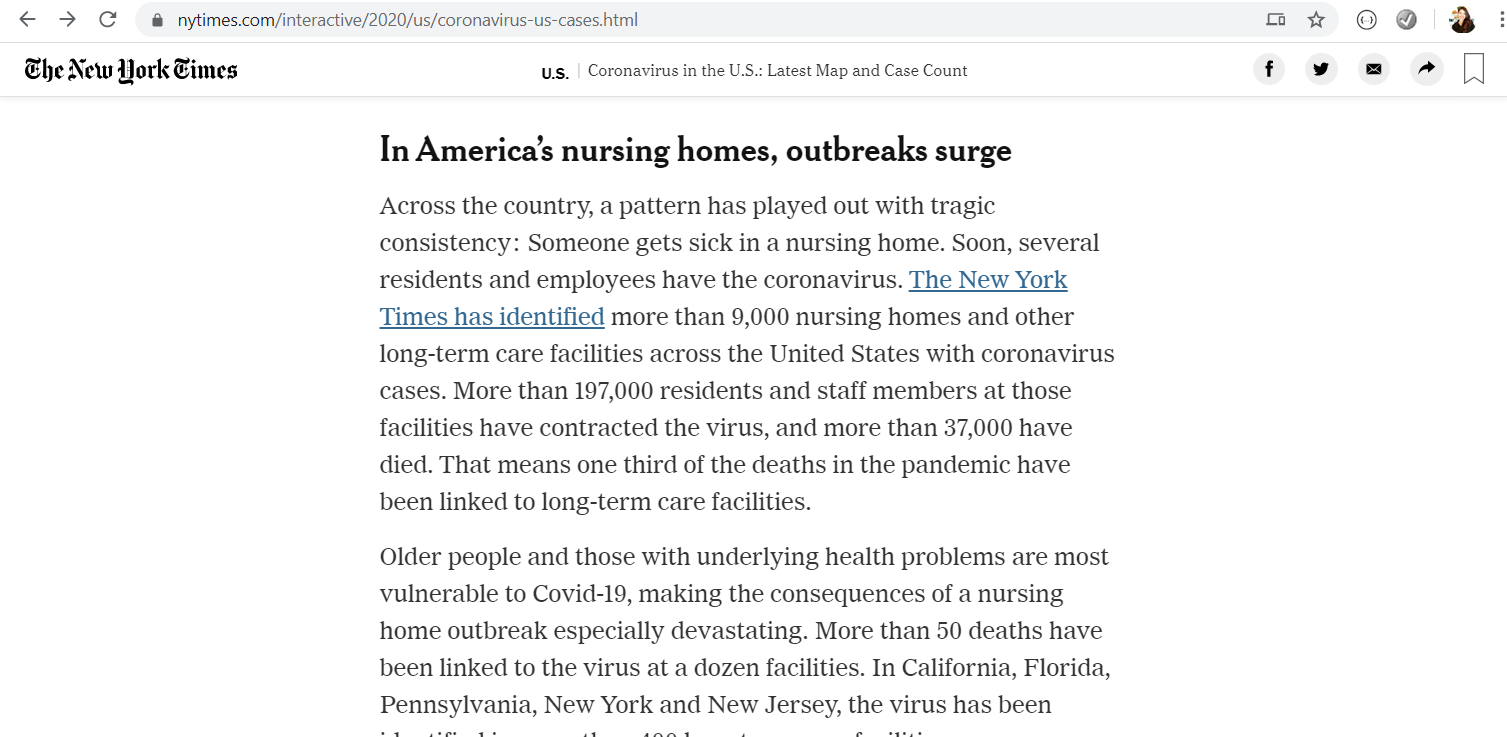
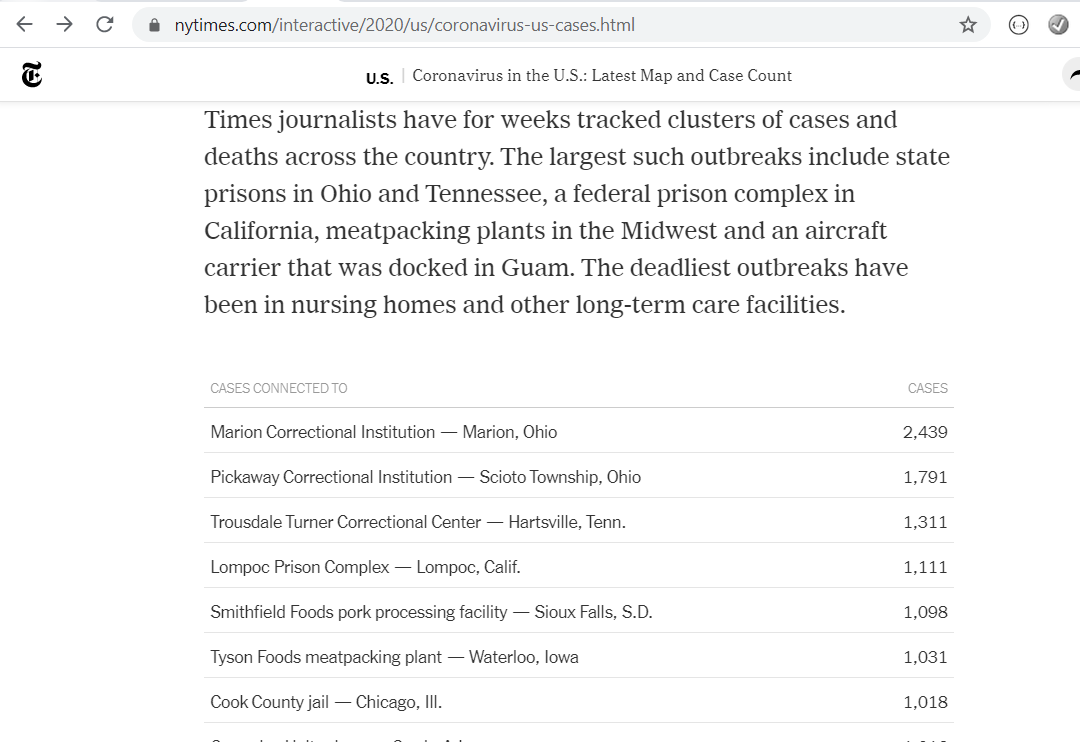
# **Extract:**

1. We have explored the NY Times data from <https://www.nytimes.com/interactive/2020/us/coronavirus-us-cases.html> for America’s nursing homes outbreaks surge.



As we Inspected through the data, applied some parsing of the data tables, we are only getting 10 out of 732 rows from the HTML site. It is not pulling the whole data table.



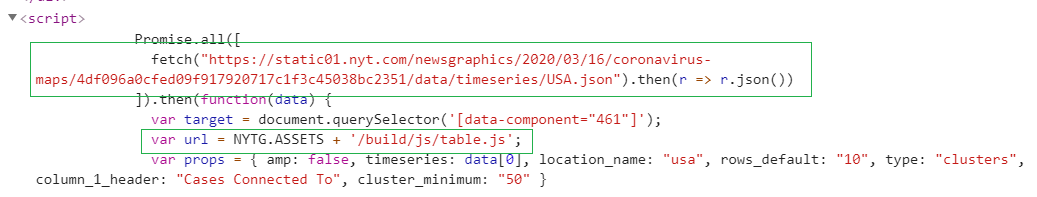


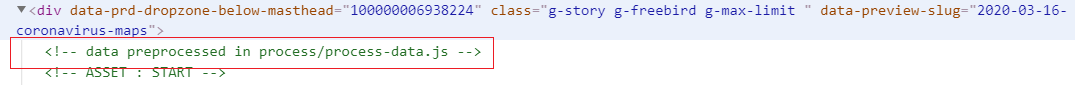
We checked using the View Page Source, and saw that after the 10th row, the table ends, and a JSON reference is nested within a script. We explored the JSON reference and it was a dump that contains too much data and it’s really hard to figure out how to extract our target.





Then we went back to Inspect the page and seen that part of the script for the tables is Java Script. Hence, we think the process to scrape the data from this website requires further knowledge on Java Script.





# **Transform & Load:**

We were not able to Transform and Load the data.

## This type of data was chosen due to f uprising numbers of COVID-19 deaths and infections for Homes and Nurses. Being part of a family of nurses, I proposed this type of data to learn more on how to find the right sources that can provide correct and organized data.