Data Appendix

I. HTML_clean_InmateInfo.csv
The unit of observation for this particular dataset is each inmate that underwent death row. For each inmate, there are several variables included that give identifying information about them:
   A. Execution Number
      a. The number identifying the inmate.
   B. Name
      a. The last name of the inmate.
   C. Date Received
      a. The date the inmate entered prison.
   D. Education Level
      a. The grade level completed.
   E. Date of Offense
      a. The date the crime was committed.
   F. Prior Occupation
      a. The occupation of the inmate prior to being incarcerated.
   G. Prior Prison Record
      a. Whether the inmate had a prior sentence in prison.
   H. Summary of Incident
      a. A description of the crime the inmate committed.

II. Cleaned_final_statements.csv
The unit of observation of particular interest for this dataset is the last words spoken by each inmate. The variables of this dataset are as follows:
   A. Execution Number
      a. The number identifying the inmate.
   B. Inmate Name
      a. The first, middle (when applicable), and last name of the inmate.
   C. Last Statement
      a. The last words spoken by the inmate before their death.

III. deathRowInitial.csv
This is the csv scraped from the original website's html. The unit of observation for this dataset is each inmate. The variables of this dataset are as follows:
   A. Execution#
      a. The number identifying the inmate.
   B. Last Name
      a. The last name of the inmate.
   C. First Name
      a. The first name of the inmate.
   D. TDCJ Number
      a. The Texas Department of Criminal Justice identification number of the inmate.
   E. Age

a.   The age of the inmate.
    F. Date
           a.   The date of death of the inmate.
    G. Race
           a.   The race of the inmate.
    H. County
           a.   The inmate's county.
    I. URL
           a.   The URL link to further information about each inmate and the crime they committed,
                which can be seen in HTML_clean_InmateInfo.csv**.**

 IV.   deathRow.html
This file is the original data scraped from the website that was then converted into a csv. The variables
from this are described in deathRowInitial.csv.

 V.   inmate_infoInitial.csv
This file is the original further inmate information csv prior to cleaning. The cleaned version, as well as
the variables, can be seen in HTML_clean_InmateInfo.csv.

 VI.   last_statement_links.csv
This file contains the links to the last words spoken by each inmate. The cleaned version, as well as the
variables, can be seen in Cleaned_final_statements.csv.

VII.   negative-words.txt
Words associated with negativity that are used for sentiment analysis from
http://www.cs.uic.edu/~liub/FBS/sentiment-analysis.html.