Readability of Presedential Farewell Addresses

by Daniel Cheslo

This program takes the Farewell Addresses of some of the past presidents (seen here), and using the Readability command included in the StringTools package, calculates the readability of the various Addresses using multiple methods, as will be outlined below. It then takes the different scores, and plots them on a heat map, to show how each president compares to one another.

- So first off is the SMOG method: https://en.wikipedia.org/wiki/SMOG.
  This method calculates its score based on the number of words of 3 syllables or greater included in the passage, and outputs a value which is the grade in which you should be able to read it.

- Next is the Gunning Fog method: https://en.wikipedia.org/wiki/Gunning_fog_index
  This method uses the same idea as the SMOG method, but uses a different formula to calculate its score.

- Next is the FRES method: https://en.wikipedia.org/wiki/Flesch%E2%80%93Kincaid_readability_tests#Flesch_reading_ease
  This method uses total syllables and total words to calculate a value, where the lower the number, the harder it is to read the passage. Due to this, it has been multiplied by -1 to better show values on the heat map.

- Then the FKGLF method: https://en.wikipedia.org/wiki/Flesch%E2%80%93Kincaid_readability_tests#Flesch.E2.80.93Kincaid_grade_level
  This method is more or less the same as FRES, but uses a different scale, and like the SMOG method and Gunning Fog method, uses a lower score to represent better readability.

- Next is the ARI method: https://en.wikipedia.org/wiki/Automated_readability_index
  This method uses number of characters and words, which translates to basing the score based on the number of longer words used.

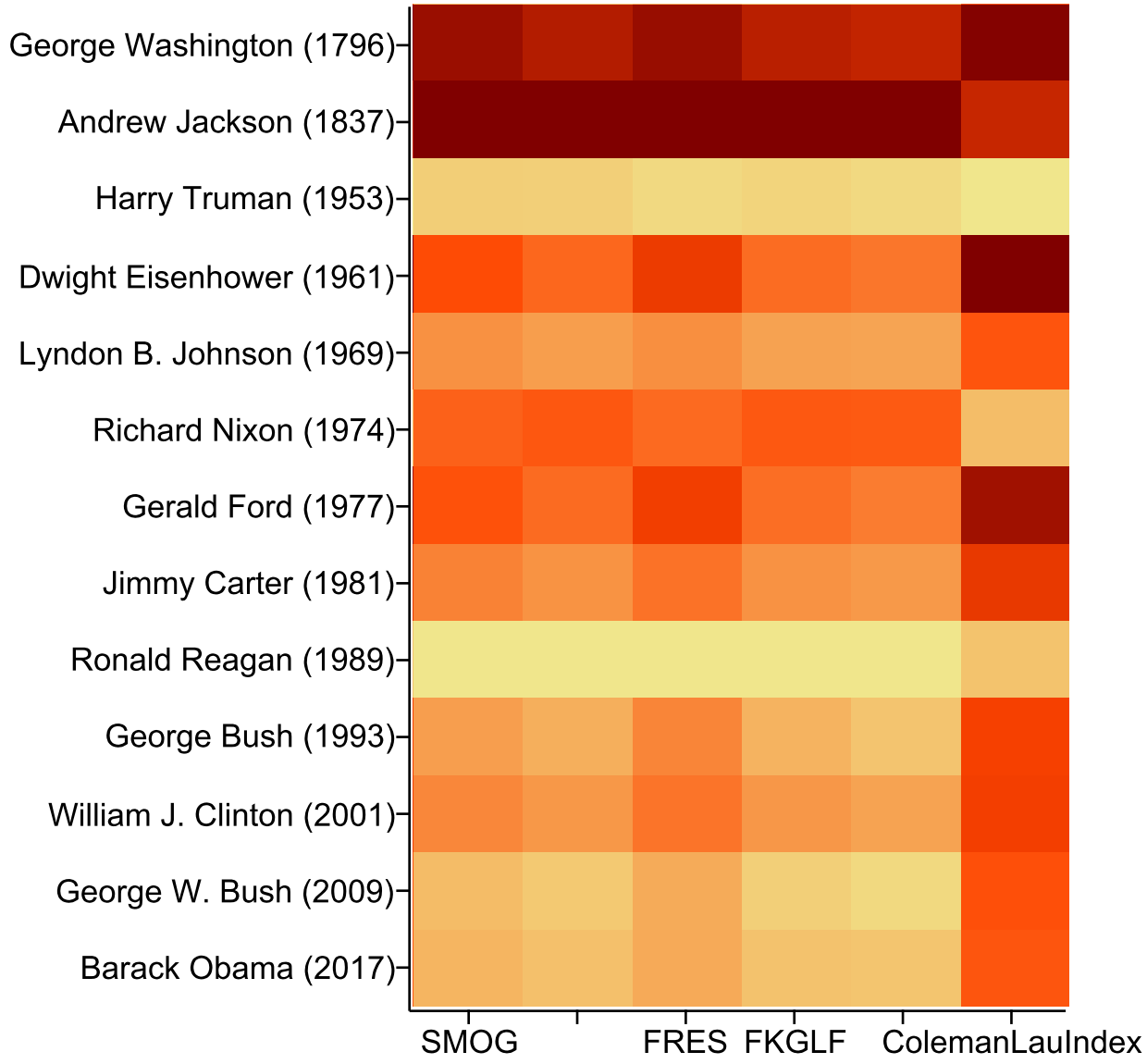
- Finally, the Coleman-Lau Index: https://en.wikipedia.org/wiki/Coleman%E2%80%93Liau_index
  This method uses the number of characters per 100 words, and number of sentences per 100 words to calculate its score.

As we can see from the heat map, Harry Truman and Ronald Reagan seem to be the most readable overall, throughout all the methods. Whether this is a good thing or not, is really up to your interpretation. Is using bigger words a bad thing?

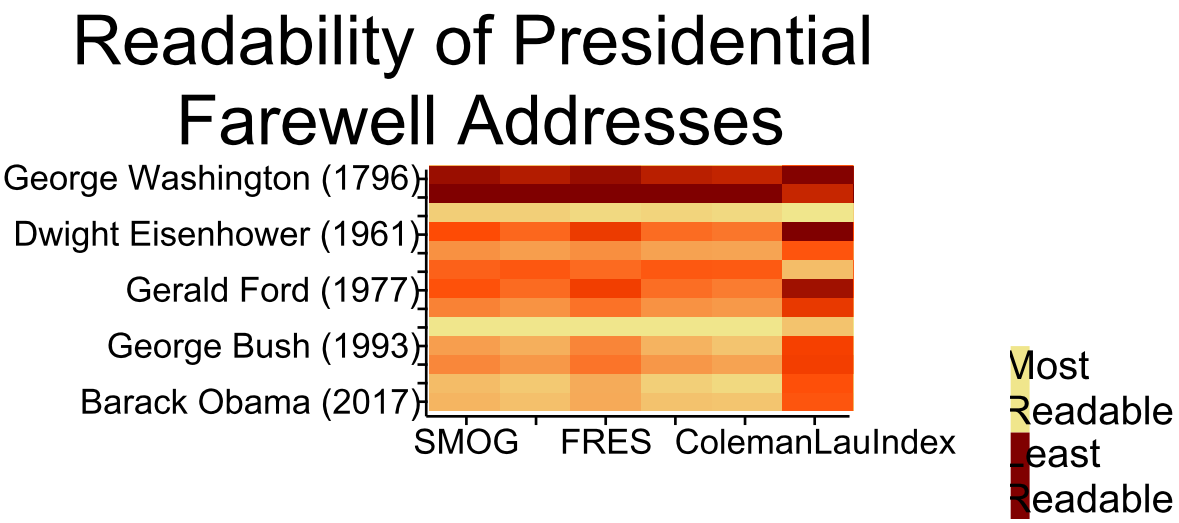
George Washington and Andrew Jackson have the worst readabiltiy scores, although this is probably in part due to the time of the addresses, and the change in language we've had since then.

```
> with(StringTools):
  with(plots):
  with(plottools):
> AJ := FileTools[Text][ReadFile]("this:///Andrew Jackson.txt"):
  BO := FileTools[Text][ReadFile]("this:///Barack Obama.txt"):
  DE := FileTools[Text][ReadFile]("this:///Dwight Eisenhower.txt"):
  GB := FileTools[Text][ReadFile]("this:///George Bush.txt"):
  GWB := FileTools[Text][ReadFile]("this:///George W. Bush.txt"):
  GW := FileTools[Text][ReadFile]("this:///George Washington.txt"):
  GRF := FileTools[Text][ReadFile]("this:///Gerald R. Ford.txt"):
  HT := FileTools[Text][ReadFile]("this:///Harry Truman.txt"):
  JC := FileTools[Text][ReadFile]("this:///Jimmy Carter.txt"):
  LJ := FileTools[Text][ReadFile]("this:///Lyndon B. Johnson.txt"):
  RN := FileTools[Text][ReadFile]("this:///Richard Nixon.txt"):
  RR := FileTools[Text][ReadFile]("this:///Ronald Reagan.txt"):
  WC := FileTools[Text][ReadFile]("this:///William J. Clinton.txt"):

> smog := [seq(Readability(i, method = SMOG), i in [GW, AJ, HT, DE, LJ, RN, GRF, JC, RR, GB, WC, GWB,
      BO])]:
  GunFog := [seq(Readability(i, method = GunningFog), i in [GW, AJ, HT, DE, LJ, RN, GRF, JC, RR, GB,
      WC, GWB, BO])]:
  fres := -1·~[seq(Readability(i, method = FRES), i in [GW, AJ, HT, DE, LJ, RN, GRF, JC, RR, GB, WC,
      GWB, BO])]:
  fkglf := [seq(Readability(i, method = FKGLF), i in [GW, AJ, HT, DE, LJ, RN, GRF, JC, RR, GB, WC, GWB,
      BO])]:
  ari := [seq(Readability(i, method = ARI), i in [GW, AJ, HT, DE, LJ, RN, GRF, JC, RR, GB, WC, GWB,
      BO])]:
  CLI := [seq(Readability(i, method = ColemanLauIndex), i in [GW, AJ, HT, DE, LJ, RN, GRF, JC, RR, GB,
      WC, GWB, BO])]:
> Addresses := DataFrame(smog, GunFog, fres, fkglf, ari, CLI, rows = [`George Washington (1796)`,
      `Andrew Jackson (1837)`, `Harry Truman (1953)`, `Dwight Eisenhower (1961)`,
      `Lyndon B. Johnson (1969)`, `Richard Nixon (1974)`, `Gerald Ford (1977)`, `Jimmy Carter (1981)`,
      `Ronald Reagan (1989)`, `George Bush (1993)`, `William J. Clinton (2001)`,
      `George W. Bush (2009)`, `Barack Obama (2017)`], columns = [`SMOG`, `GunningFog`, `FRES`,
      `FKGLF`, `ARI`, `ColemanLauIndex`]):
> p1 := Statistics:-HeatMap(Addresses, color = ["Khaki", "OrangeRed", "Maroon"], size = [1000, 500],
      font = [Arial, 10], title = "Readability of Presidential Farewell Addresses", titlefont = [Arial, 20])
      ;
```

Readability of Presidential Farewell Addresses




```
> p2 := display( rectangle( [0,-5], [1,-6], color = "Maroon", style = surface ), textplot( [4,-5.5,
      "Least Readable" ], font = [Arial] ), rectangle( [0,-5], [1,-4], color = "Khaki", style = surface ),
      textplot( [4,-4.5, "Most Readable" ], font = [Arial] ), view = [0 ..10, 0 .. -12 ], axes = none ) :

> display( Array( [p1, p2] ) )
```


Readability of Presidential Farewell Addresses
George Washington (1796)
Dwight Eisenhower (1961)
Gerald Ford (1977)
George Bush (1993)
Barack Obama (2017)
SMOG
FRES
ColemanLau
Index
Most Readable
Least Readable

>