

Mini Essay 5a*

Hritik Shukla

2024-02-06

Getting HTML data and saving it locally

Reading the locally saved HTML file into a tibble

```
# A tibble: 6 x 12
  No. Portrait Name(Birth-Death)Con~1 `Election(Parliament)` `Term of office`
  <chr> <chr>      <chr>                        <chr>                <chr>
1 No.   "Portrai~ Name(Birth-Death)Cons~ Election(Parliament) Took office
2 1     ""      Edmund Barton(1849-19~ 1901 (1st)          1 January1901
3 1     ""      Edmund Barton(1849-19~ 1901 (1st)          1 January1901
4 1     ""      Edmund Barton(1849-19~ 1901 (1st)          1 January1901
5 2     ""      Alfred Deakin(1856-19~ - (1st)             24 September1903
6 2     ""      Alfred Deakin(1856-19~ 1903 (2nd)          24 September1903
# i abbreviated name: 1: `Name(Birth-Death)Constituency`
# i 7 more variables: `Term of office` <chr>, `Term of office` <chr>,
#   Politicalparty <chr>, Ministry <chr>, Monarch <chr>,
#   `Governor-General` <chr>, Ref. <chr>
```

Cleaning column names and keeping only the required column

```
# A tibble: 6 x 1
  raw_text
  <chr>
1 Edmund Barton(1849-1920)MP for Hunter, NSW
2 Alfred Deakin(1856-1919)MP for Ballaarat, Vic[a]
3 Chris Watson(1867-1941)MP for Bland, NSW
4 George Reid(1845-1918)MP for East Sydney, NSW
5 Andrew Fisher(1862-1928)MP for Wide Bay, Qld
6 Joseph Cook(1860-1947)MP for Parramatta, NSW
```

*Code and data are available at: <https://github.com/hritikshuklas/miniessay5a>

Table 1: Lifespan Data

Prime Minister	Birth year	Death year	Age at death
Edmund Barton	1849	1920	71
Alfred Deakin	1856	1919	63
Chris Watson	1867	1941	74
George Reid	1845	1918	73
Andrew Fisher	1862	1928	66
Joseph Cook	1860	1947	87

Separating data into three columns - name, date (birth-death), or born (if currently alive).

```
# A tibble: 6 x 3
  name      date      born
  <chr>    <chr>    <chr>
1 Edmund Barton 1849-1920 <NA>
2 Alfred Deakin 1856-1919 <NA>
3 Chris Watson  1867-1941 <NA>
4 George Reid   1845-1918 <NA>
5 Andrew Fisher 1862-1928 <NA>
6 Joseph Cook   1860-1947 <NA>
```

Separating birth years and death years(if there is one) for both currently alive and dead PMs.

Giving columns more readable headers.

Plotting the lifespans of all prime ministers:

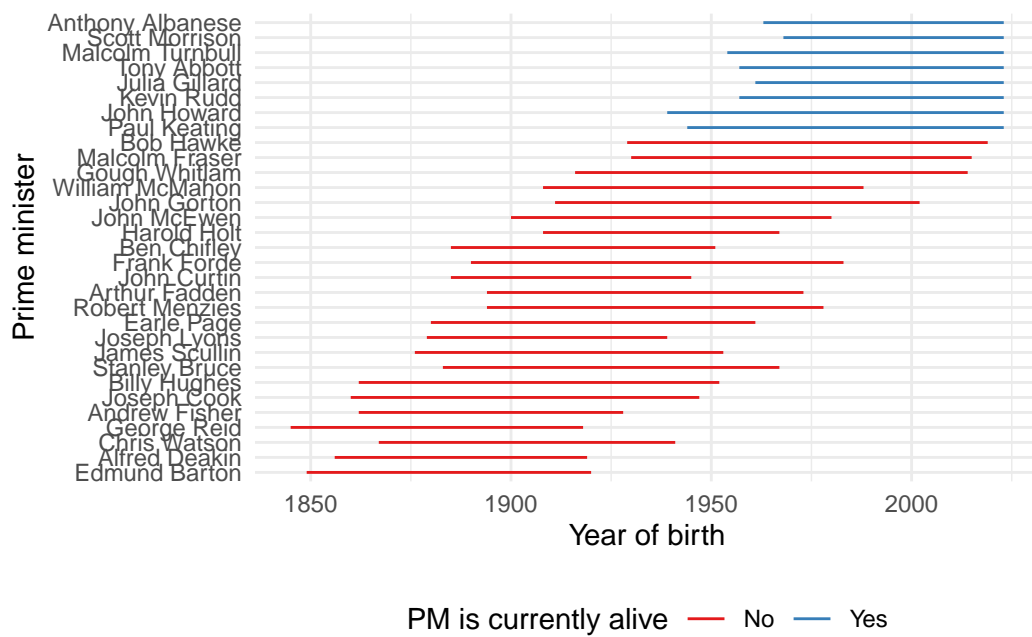


Figure 1: Lifespans of Australian Prime Ministers