Mini Essay 5: A visualized study of life span of prime ministers in Australia

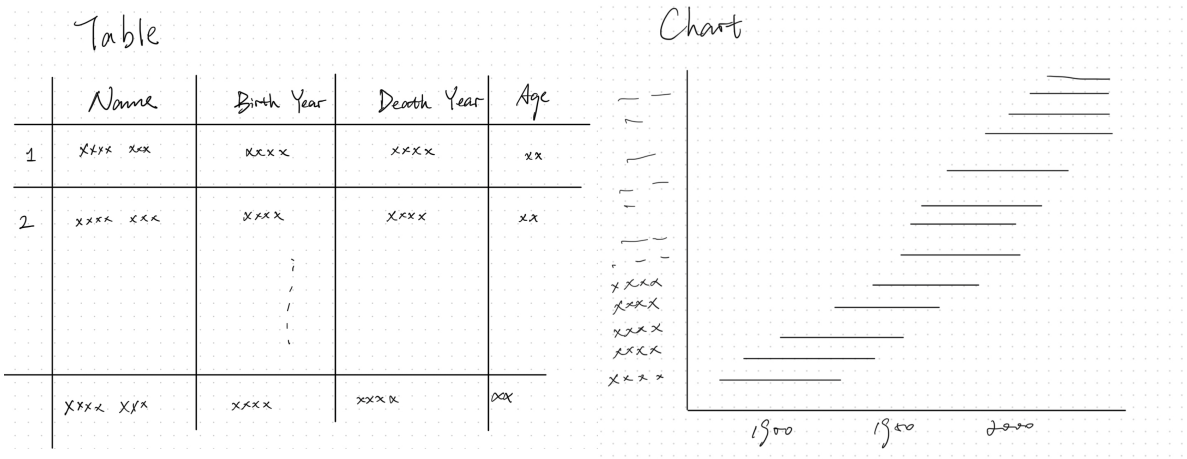
Yingzhi Zhang

February 6, 2024

1 Introduction

Australia has a parliamentary democratic government under its constitutional monarch system. The prime minister of Australia, set up by its constitution, is the head of the government. In the history of Australia, a total of 31 prime ministers have served for the nation since the establishment of the cabinet in 1901. The first prime minister was elected in March 1901, and the winner was Edmund Barton. The most recent prime minister is Anthony Albanese, who won the election in May 2022.

In this research, we will study the life span of each prime minister. We can study the life span by discovering the birth year and death year of the prime ministers. Calculating the difference by subtracting birth from death year, we can generate a table to show the age at death of all prime ministers, except for those who are still alive. Furthermore, we can draw a graph to visualize our study.



(a) Planning Table

(b) Planning graph

Figure 1: Plan of data

2 Data

We can first gather the data from open resources on the internet. Wikipedia.com provides a comprehensive table regarding the informatioin of prime ministers of Australia. Using web scraping procedure, we can gather the data we need from ths website. Here we use the programming language R(R Core Team 2022), and we also use packages "tidyverse"(Wickham et al. 2019), "httr"(Wickham 2023), "janitor"(Firke 2023), "knitr"(Xie 2014). In specifically, we will use "rvest"(Wickham 2022) under the "tidyverse" package to process html websites.

We can gather the data needed using "rvest" and SelectorGadget tool. Using the read_html function under "rvest" package, we can read the information on the Wikipedia website of "List of Prime Minister of Australia". However, not all the information on the website are in need. Therefore, we can use SelectorGadget, a javascript tool, to select the table that contains the detailed information of all prime ministers.

After gathering the data we need, we can clean the data to focus on the birth year and death year of each prime ministers. We will first select the column we need, and then select the useful content in the column. We can mutate two new columns to store the data of birth year and death year. We would leave "NA" in the death year for the prime ministers who are still alive. We can further calculate the age at death be subtracting the two numbers and mutating a new columns with the name "Age at death". Using the kable() tool, we can generate a table which shows the cleaned version of data that we get from the previous steps.

Table 1: Table of cleaned data regarding life span of Australian prime ministers

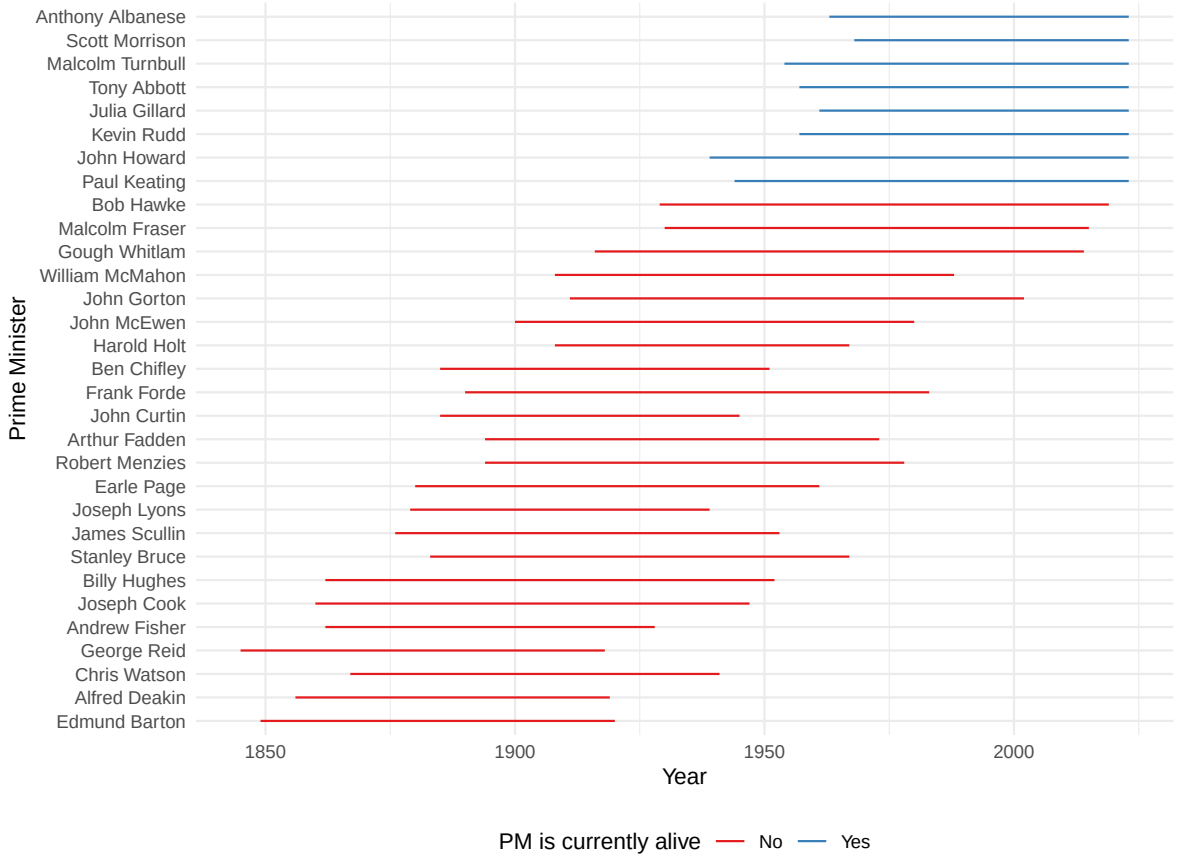
Prime Minister	Birth year	Death year	Age at death
Edmund Barton	1849	1920	71
Alfred Deakin	1856	1919	63
Chris Watson	1867	1941	74
George Reid	1845	1918	73
Andrew Fisher	1862	1928	66
Joseph Cook	1860	1947	87
Billy Hughes	1862	1952	90
Stanley Bruce	1883	1967	84
James Scullin	1876	1953	77
Joseph Lyons	1879	1939	60
Earle Page	1880	1961	81
Robert Menzies	1894	1978	84
Arthur Fadden	1894	1973	79
John Curtin	1885	1945	60
Frank Forde	1890	1983	93
Ben Chifley	1885	1951	66
Harold Holt	1908	1967	59

Table 1: Table of cleaned data regarding life span of Australian prime ministers

Prime Minister	Birth year	Death year	Age at death
John McEwen	1900	1980	80
John Gorton	1911	2002	91
William McMahon	1908	1988	80
Gough Whitlam	1916	2014	98
Malcolm Fraser	1930	2015	85
Bob Hawke	1929	2019	90
Paul Keating	1944	NA	NA
John Howard	1939	NA	NA
Kevin Rudd	1957	NA	NA
Julia Gillard	1961	NA	NA
Tony Abbott	1957	NA	NA
Malcolm Turnbull	1954	NA	NA
Scott Morrison	1968	NA	NA
Anthony Albanese	1963	NA	NA

For a better view of the data, we can use ggplot2() to get a figure of the life span of all Australian Prime Ministers. A bar chart is generated for better visualization effect. We can put the year on the x-axis and the name of prime ministers on the y-axis. To distinguish the prime ministers who are still alive, we can separate them using the color blue instead of red. As shown in Figure 2, each bar represents the life span of the prime minister.

Figure 2: Life span of Australian prime ministers



3 Discussion

As presented in Figure 2, there are 8 prime ministers alive and 23 prime ministers who have passed away. Most prime minister who are still alive were born after 1950, except John Howard and Paul Keating who were born in 1944 and 1939 respectively. While Anthony Albanese is the latest prime minister who is currently serving in the office, the previous prime minister, Scott Morrison is the yongest among all 31 prime ministers. Born in 1968, Mr. Morrison is only 56 years old this year. The most recent year in which a prime minister passed away is 2019. Bob Hawke died at the age of 90.

The first prime minister, Edmund Barton, was born in 1849 and passed away in 1920. The fourth prime minister in the history of Australia, George Reid, is the one who was born in the earliest year. Mr. Reid was born in 1845 and died in 1918 at the age of 73. Among the prime ministers who have passed away, the most long-lived prime minister is Gough Whitlam who was born in 1916 and died in 2014 at the age

of 98. On the other hand, Harold Holt is the prime minister who has the shortest life span. In general, the life span of Australian prime ministers is getting longer and longer among the years.

References

Firke, Sam. 2023. *Janitor: Simple Tools for Examining and Cleaning Dirty Data*. https://CRAN.R-project.org/package=janitor.

R Core Team. 2022. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Wickham, Hadley. 2022. *Rvest: Easily Harvest (Scrape) Web Pages*. https://CRAN.R-project.org/package=rvest.

———. 2023. *Httr: Tools for Working with URLs and HTTP*. https://CRAN.R-project.org/package=httr.

Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D'Agostino McGowan, Romain François, Garrett Grolemund, et al. 2019. "Welcome to the tidyverse." *Journal of Open Source Software* 4 (43): 1686. https://doi.org/10.21105/joss.01686.

Xie, Yihui. 2014. "Knitr: A Comprehensive Tool for Reproducible Research in R." In *Implementing Reproducible Computational Research*, edited by Victoria Stodden, Friedrich Leisch, and Roger D. Peng. Chapman; Hall/CRC.