

# R Notebook

```
library(tidyverse)
```

```
## -- Attaching packages ----- tidyverse 1.3.0 --
```

```
## v ggplot2 3.3.3      v purrr   0.3.4
## v tibble  3.0.6      v dplyr  1.0.4
## v tidyr   1.1.2      v stringr 1.4.0
## v readr   1.4.0      v forcats 0.5.1
```

```
## -- Conflicts ----- tidyverse_conflicts() --
```

```
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()     masks stats::lag()
```

```
library(mdsr)
presidential
```

```
## # A tibble: 11 x 4
##   name      start      end      party
##   <chr>    <date>    <date>    <chr>
## 1 Eisenhower 1953-01-20 1961-01-20 Republican
## 2 Kennedy    1961-01-20 1963-11-22 Democratic
## 3 Johnson    1963-11-22 1969-01-20 Democratic
## 4 Nixon      1969-01-20 1974-08-09 Republican
## 5 Ford       1974-08-09 1977-01-20 Republican
## 6 Carter     1977-01-20 1981-01-20 Democratic
## 7 Reagan     1981-01-20 1989-01-20 Republican
## 8 Bush       1989-01-20 1993-01-20 Republican
## 9 Clinton    1993-01-20 2001-01-20 Democratic
## 10 Bush      2001-01-20 2009-01-20 Republican
## 11 Obama     2009-01-20 2017-01-20 Democratic
```

```
select(presidential, name, party)
```

```
## # A tibble: 11 x 2
##   name      party
##   <chr>    <chr>
## 1 Eisenhower Republican
## 2 Kennedy    Democratic
## 3 Johnson    Democratic
## 4 Nixon      Republican
## 5 Ford       Republican
## 6 Carter     Democratic
## 7 Reagan     Republican
```

```
## 8 Bush      Republican
## 9 Clinton   Democratic
## 10 Bush     Republican
## 11 Obama    Democratic
```

```
filter(presidential, party == "Republican")
```

```
## # A tibble: 6 x 4
##   name      start      end      party
##   <chr>    <date>    <date>   <chr>
## 1 Eisenhower 1953-01-20 1961-01-20 Republican
## 2 Nixon      1969-01-20 1974-08-09 Republican
## 3 Ford       1974-08-09 1977-01-20 Republican
## 4 Reagan     1981-01-20 1989-01-20 Republican
## 5 Bush       1989-01-20 1993-01-20 Republican
## 6 Bush       2001-01-20 2009-01-20 Republican
```

*#Note that the == is a test for equality*

```
select(
  filter(presidential, lubridate::year(start) > 1973 & party == "Democratic"),
  name
)
```

```
## # A tibble: 3 x 1
##   name
##   <chr>
## 1 Carter
## 2 Clinton
## 3 Obama
```

*#For example, we can find which Democratic presidents served since Watergate.*

```
presidential %>%
  filter(lubridate::year(start) > 1973 & party == "Democratic") %>%
  select(name)
```

```
## # A tibble: 3 x 1
##   name
##   <chr>
## 1 Carter
## 2 Clinton
## 3 Obama
```

*#This expression is called a pipeline, that is the use of %>%*

```
library(lubridate)
```

```
##
## Attaching package: 'lubridate'
```

```
## The following objects are masked from 'package:base':
##
##   date, intersect, setdiff, union
```

```
my_presidents <- presidential %>%
  mutate(term.length = interval(start, end) / dyears(1))
my_presidents
```

```
## # A tibble: 11 x 5
##   name      start      end      party      term.length
##   <chr>    <date>    <date>    <chr>      <dbl>
## 1 Eisenhower 1953-01-20 1961-01-20 Republican      8
## 2 Kennedy    1961-01-20 1963-11-22 Democratic    2.84
## 3 Johnson    1963-11-22 1969-01-20 Democratic    5.16
## 4 Nixon      1969-01-20 1974-08-09 Republican    5.55
## 5 Ford       1974-08-09 1977-01-20 Republican    2.45
## 6 Carter     1977-01-20 1981-01-20 Democratic      4
## 7 Reagan     1981-01-20 1989-01-20 Republican      8
## 8 Bush       1989-01-20 1993-01-20 Republican      4
## 9 Clinton    1993-01-20 2001-01-20 Democratic      8
## 10 Bush      2001-01-20 2009-01-20 Republican      8
## 11 Obama     2009-01-20 2017-01-20 Democratic      8
```

```
my_presidents <- my_presidents %>%
  mutate(elected = year(start) - 1)
my_presidents
```

```
## # A tibble: 11 x 6
##   name      start      end      party      term.length elected
##   <chr>    <date>    <date>    <chr>      <dbl>    <dbl>
## 1 Eisenhower 1953-01-20 1961-01-20 Republican      8      1952
## 2 Kennedy    1961-01-20 1963-11-22 Democratic    2.84    1960
## 3 Johnson    1963-11-22 1969-01-20 Democratic    5.16    1962
## 4 Nixon      1969-01-20 1974-08-09 Republican    5.55    1968
## 5 Ford       1974-08-09 1977-01-20 Republican    2.45    1973
## 6 Carter     1977-01-20 1981-01-20 Democratic      4      1976
## 7 Reagan     1981-01-20 1989-01-20 Republican      8      1980
## 8 Bush       1989-01-20 1993-01-20 Republican      4      1988
## 9 Clinton    1993-01-20 2001-01-20 Democratic      8      1992
## 10 Bush      2001-01-20 2009-01-20 Republican      8      2000
## 11 Obama     2009-01-20 2017-01-20 Democratic      8      2008
```

```
my_presidents <- my_presidents %>%
  mutate(elected = ifelse(elected %in% c(1962, 1973), NA, elected))
my_presidents
```

```
## # A tibble: 11 x 6
##   name      start      end      party      term.length elected
##   <chr>    <date>    <date>    <chr>      <dbl>    <dbl>
## 1 Eisenhower 1953-01-20 1961-01-20 Republican      8      1952
## 2 Kennedy    1961-01-20 1963-11-22 Democratic    2.84    1960
## 3 Johnson    1963-11-22 1969-01-20 Democratic    5.16     NA
```

```
## 4 Nixon      1969-01-20 1974-08-09 Republican      5.55      1968
## 5 Ford       1974-08-09 1977-01-20 Republican      2.45       NA
## 6 Carter     1977-01-20 1981-01-20 Democratic       4        1976
## 7 Reagan     1981-01-20 1989-01-20 Republican       8        1980
## 8 Bush       1989-01-20 1993-01-20 Republican       4        1988
## 9 Clinton    1993-01-20 2001-01-20 Democratic       8        1992
## 10 Bush      2001-01-20 2009-01-20 Republican       8        2000
## 11 Obama     2009-01-20 2017-01-20 Democratic       8        2008
```

```
my_presidents <- my_presidents %>%
  rename(term_length = term.length)
my_presidents
```

```
## # A tibble: 11 x 6
##   name      start      end      party      term_length elected
##   <chr>    <date>    <date>    <chr>      <dbl>    <dbl>
## 1 Eisenhower 1953-01-20 1961-01-20 Republican      8      1952
## 2 Kennedy    1961-01-20 1963-11-22 Democratic    2.84    1960
## 3 Johnson    1963-11-22 1969-01-20 Democratic    5.16     NA
## 4 Nixon      1969-01-20 1974-08-09 Republican    5.55    1968
## 5 Ford       1974-08-09 1977-01-20 Republican    2.45     NA
## 6 Carter     1977-01-20 1981-01-20 Democratic      4     1976
## 7 Reagan     1981-01-20 1989-01-20 Republican      8     1980
## 8 Bush       1989-01-20 1993-01-20 Republican      4     1988
## 9 Clinton    1993-01-20 2001-01-20 Democratic      8     1992
## 10 Bush      2001-01-20 2009-01-20 Republican      8     2000
## 11 Obama     2009-01-20 2017-01-20 Democratic      8     2008
```

```
my_presidents %>%
  arrange(desc(term_length))
```

```
## # A tibble: 11 x 6
##   name      start      end      party      term_length elected
##   <chr>    <date>    <date>    <chr>      <dbl>    <dbl>
## 1 Eisenhower 1953-01-20 1961-01-20 Republican      8      1952
## 2 Reagan     1981-01-20 1989-01-20 Republican      8      1980
## 3 Clinton    1993-01-20 2001-01-20 Democratic      8      1992
## 4 Bush       2001-01-20 2009-01-20 Republican      8      2000
## 5 Obama     2009-01-20 2017-01-20 Democratic      8      2008
## 6 Nixon      1969-01-20 1974-08-09 Republican    5.55    1968
## 7 Johnson    1963-11-22 1969-01-20 Democratic    5.16     NA
## 8 Carter     1977-01-20 1981-01-20 Democratic      4     1976
## 9 Bush       1989-01-20 1993-01-20 Republican      4     1988
## 10 Kennedy    1961-01-20 1963-11-22 Democratic    2.84    1960
## 11 Ford       1974-08-09 1977-01-20 Republican    2.45     NA
```

```
my_presidents %>%
  arrange(desc(term_length), party, elected)
```

```
## # A tibble: 11 x 6
##   name      start      end      party      term_length elected
##   <chr>    <date>    <date>    <chr>      <dbl>    <dbl>
```

```
## 1 Clinton      1993-01-20 2001-01-20 Democratic      8      1992
## 2 Obama        2009-01-20 2017-01-20 Democratic      8      2008
## 3 Eisenhower   1953-01-20 1961-01-20 Republican      8      1952
## 4 Reagan       1981-01-20 1989-01-20 Republican      8      1980
## 5 Bush         2001-01-20 2009-01-20 Republican      8      2000
## 6 Nixon        1969-01-20 1974-08-09 Republican     5.55    1968
## 7 Johnson      1963-11-22 1969-01-20 Democratic     5.16     NA
## 8 Carter       1977-01-20 1981-01-20 Democratic      4      1976
## 9 Bush         1989-01-20 1993-01-20 Republican      4      1988
## 10 Kennedy      1961-01-20 1963-11-22 Democratic     2.84    1960
## 11 Ford        1974-08-09 1977-01-20 Republican     2.45     NA
```

```
my_presidents %>%
  summarize(
    N = n(), #Just to confirm the correct number of rows is being used
    first_year = min(year(start)),
    last_year = max(year(end)),
    num_dems = sum(party == "Democratic"),
    years = sum(term_length),
    avg_term_length = mean(term_length)
  )
```

```
## # A tibble: 1 x 6
##       N first_year last_year num_dems years avg_term_length
##   <int>      <dbl>      <dbl>    <int> <dbl>         <dbl>
## 1    11      1953      2017        5    64          5.82
```

```
my_presidents %>%
  group_by(party) %>%
  summarize(
    N = n(),
    first_year = min(year(start)),
    last_year = max(year(end)),
    num_dems = sum(party == "Democratic"),
    years = sum(term_length),
    avg_term_length = mean(term_length)
  )
```

```
## # A tibble: 2 x 7
##   party      N first_year last_year num_dems years avg_term_length
## * <chr>    <int>      <dbl>      <dbl>    <int> <dbl>         <dbl>
## 1 Democratic    5      1961      2017        5    28          5.6
## 2 Republican    6      1953      2009        0    36           6
```

## Extended example: Ben's time with the Mets