Introduction to the space launches data

INTERMEDIATE INTERACTIVE DATA VISUALIZATION WITH PLOTLY IN R

Adam Loy
Statistician, Carleton College

datacamp

Launches data set

Image credit: **Airforce Space Command, 45th Space Wing Public Affairs**

Data source: **https://github.com/rfordatascience/tidytuesday**

Launches data set

```
dplyr::glimpse(launches)
```

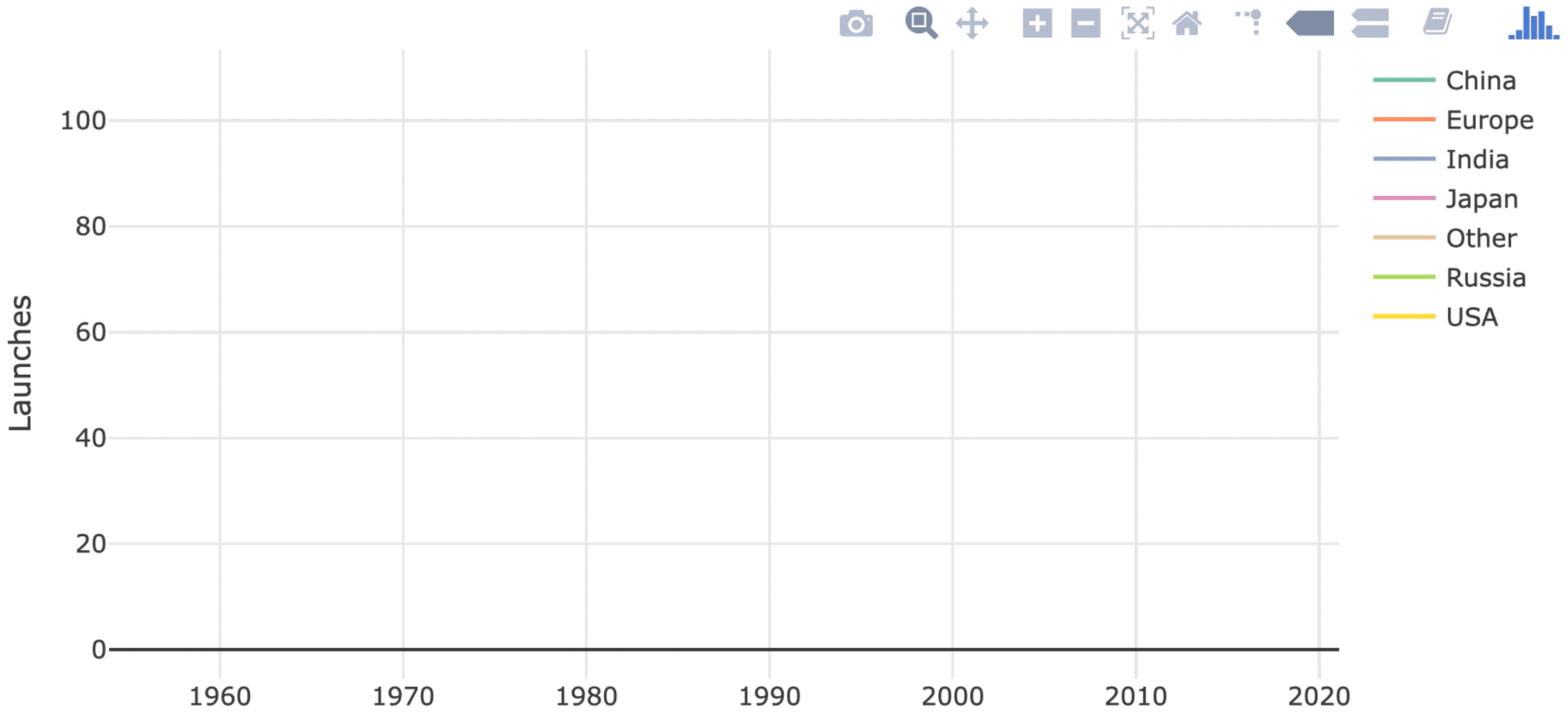
```
Observations: 5,726
Variables: 11
$ tag         <chr> "1967-065", "1967-080", "1967-096", "1968-042", …
$ JD          <dbl> 2439671, 2439726, 2439775, 2440000, 2440153, 244…
$ launch_date <date> 1967-06-29, 1967-08-23, 1967-10-11, 1968-05-23,…
$ launch_year <dbl> 1967, 1967, 1967, 1968, 1968, 1969, 1970, 1970, …
$ type        <chr> "Thor Burner 2", "Thor Burner 2", "Thor Burner 2"…
$ variant     <chr> NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, NA, …
$ mission     <chr> "Secor Type II S/N 10", "DAPP 3419", "DAPP 4417"…
$ agency      <chr> "US", "US", "US", "US", "US", "US", "US", "US", …
$ state_code  <chr> "US", "US", "US", "US", "US", "US", "US", "US", …
$ category    <chr> "O", "O", "O", "O", "O", "O", "O", "O", "O", "O"…
$ agency_type <chr> "state", "state", "state", "state", "state", "state"…
```

Variables

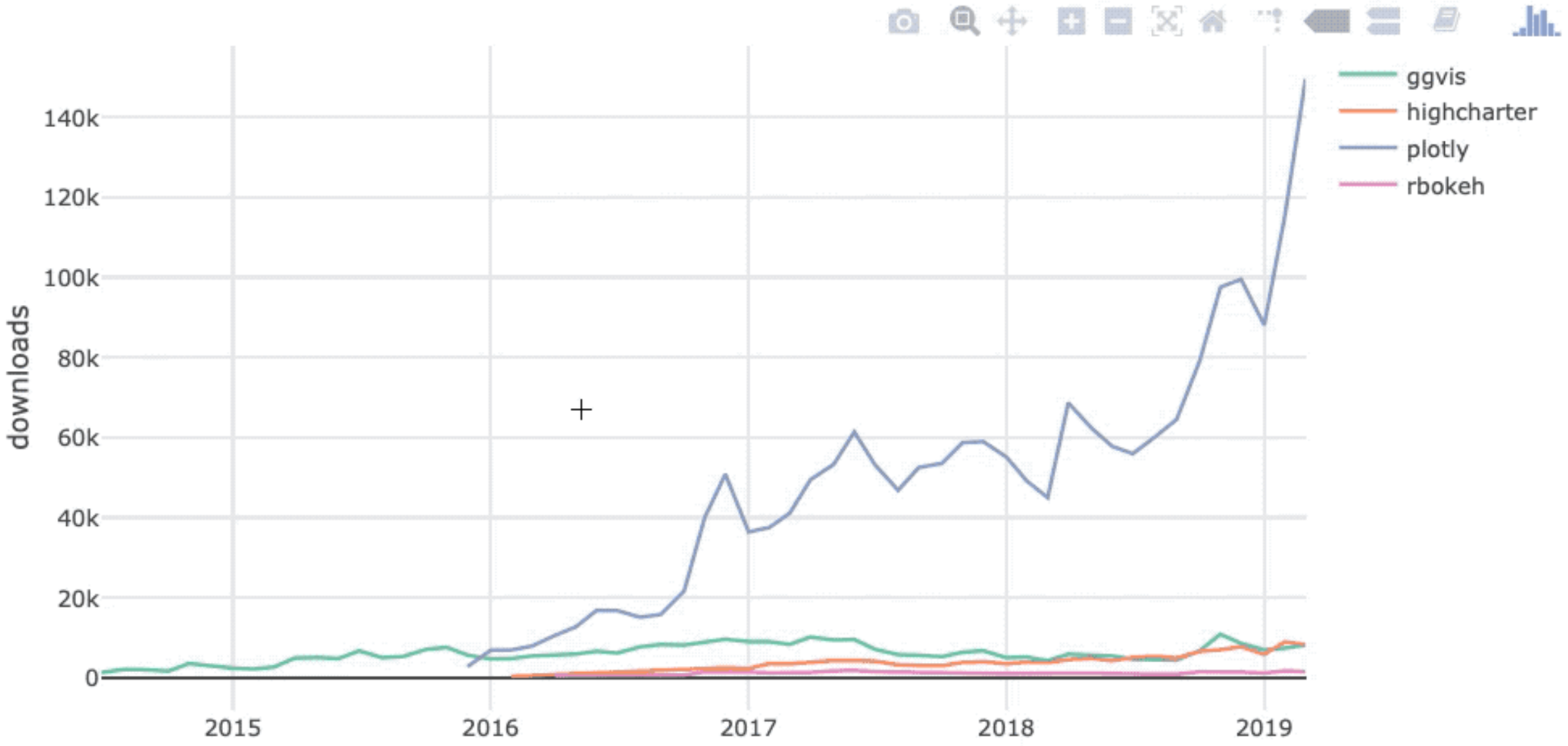
variable	definition
launch_year	year of launch
agency	launching agency
state_code	launching agency's state
category	success (O) or failure (F)
agency_type	type of agency

variable	definition
tag	Harvard or **COSPAR** id of launch
JD	**Julian Date** of launch
launch_date	date of launch
type	type of launch vehicle
variant	variant of launch vehicle
mission	mission name

Exploring the space race



R package downloads



CRAN download logs

```
dplyr::glimpse(monthly_logs)
```

```
Observations: 171
Variables: 4
$ package   <chr> "ggvis", "ggvis", "ggvis", "ggvis", "ggvis", "ggvis",…
$ date      <date> 2014-06-30, 2014-07-31, 2014-08-31, 2014-09-30, 2014…
$ dec_date  <dbl> 2014.49, 2014.58, 2014.66, 2014.75, 2014.83, 2014.91,…
$ downloads <dbl> 1344, 2120, 2035, 1702, 3590, 2899, 2427, 2227, 2708,…
```

Grouped brushing

```
shared_logs <- monthly_logs %>% SharedData$new(key = ~package)
shared_logs %>%
  plot_ly(x = ~dec_date, y = ~downloads, color = ~package) %>%
  add_lines() %>%
  highlight()
```

Let's practice!

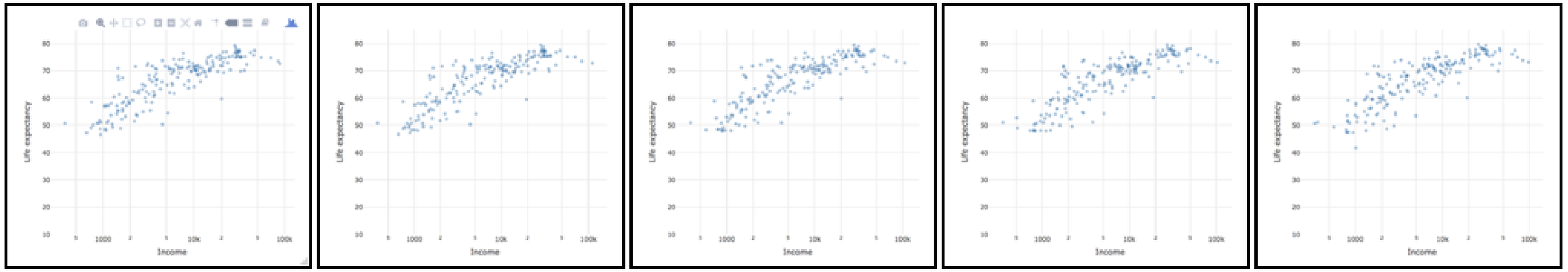
INTERMEDIATE INTERACTIVE DATA VISUALIZATION WITH PLOTLY IN R

Recap: Animation

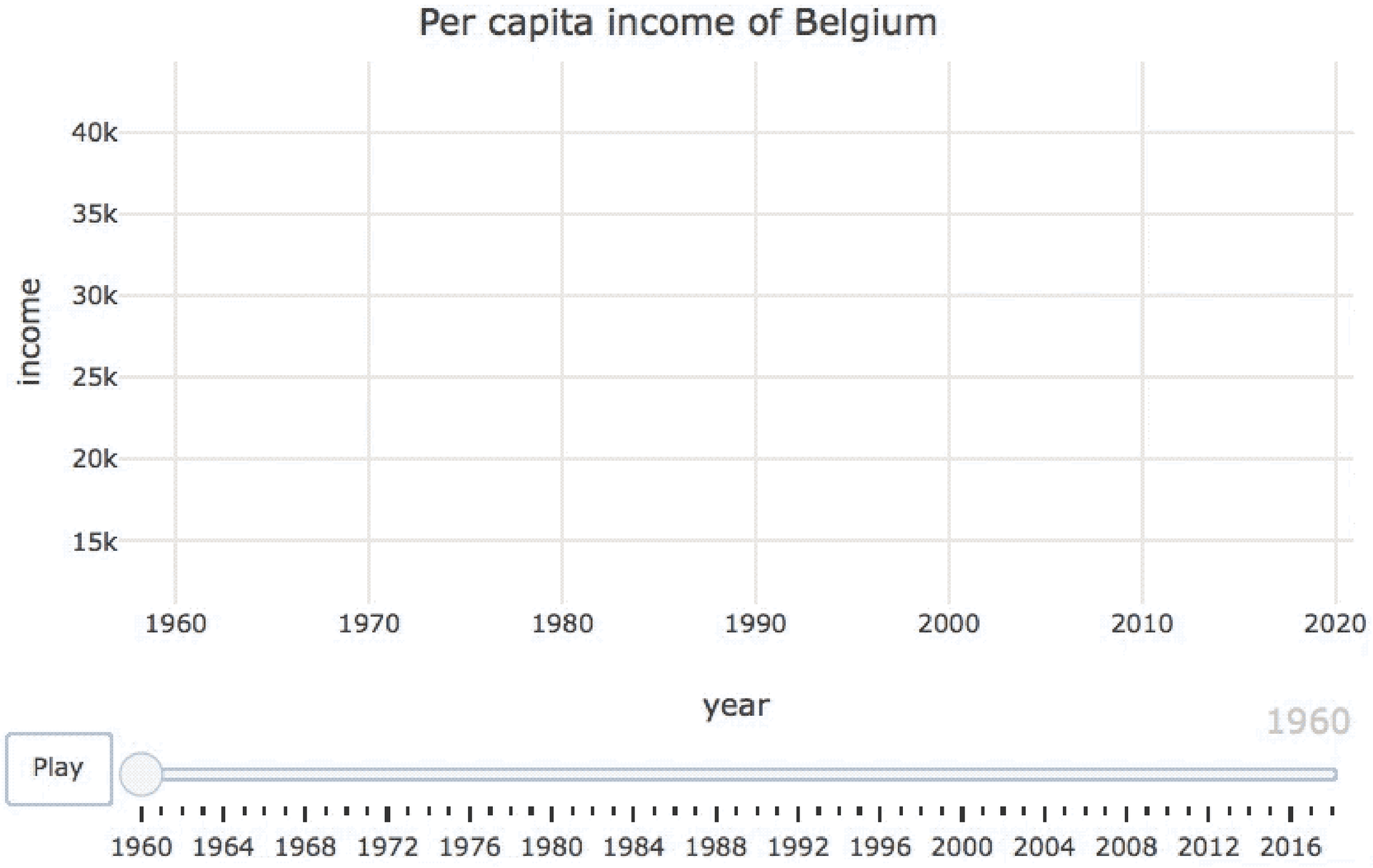
INTERMEDIATE INTERACTIVE DATA VISUALIZATION WITH PLOTLY IN R

Adam Loy
Statistician, Carleton College

Keyframe animation



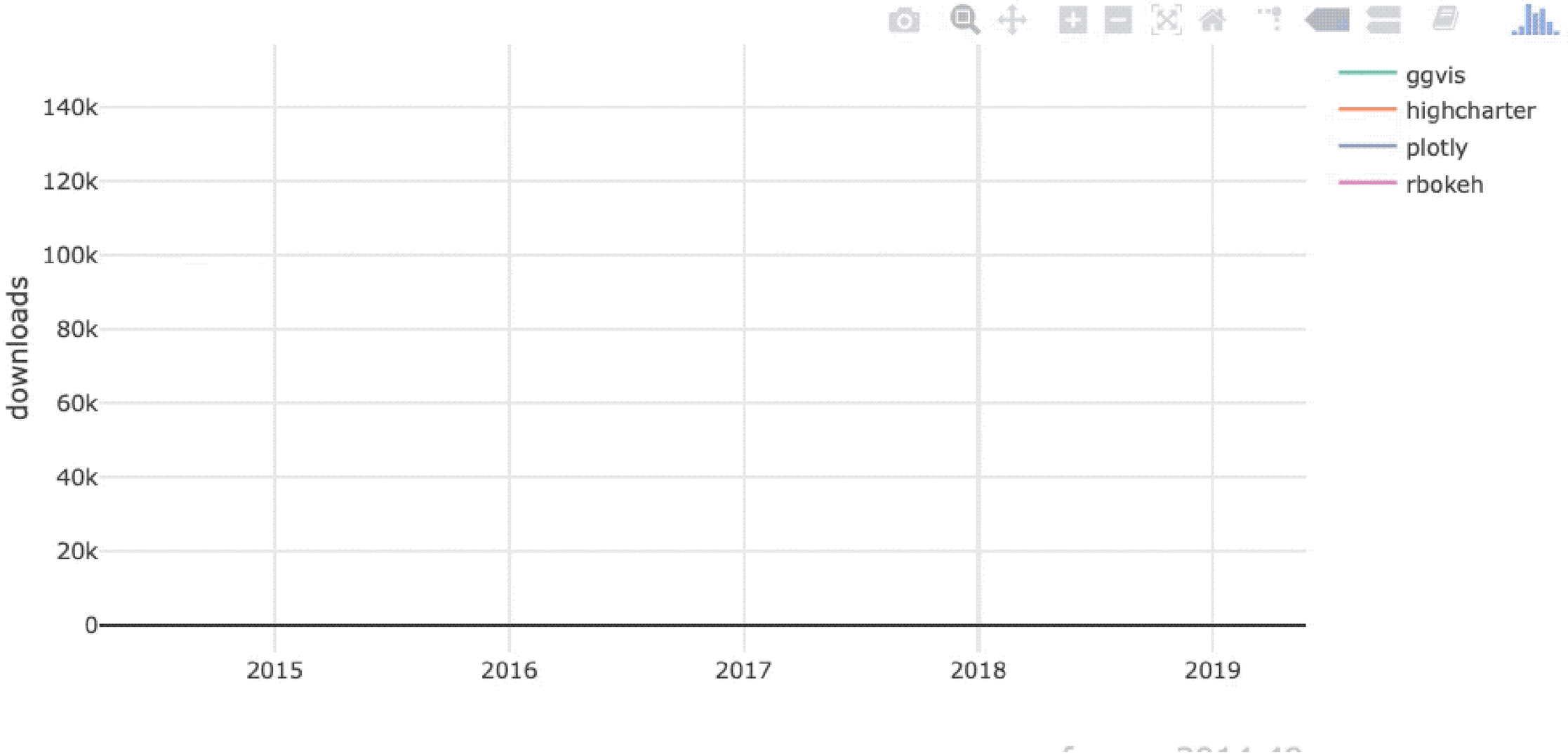

```
world_indicators %>%
  plot_ly(x = ~income, y = ~co2) %>%
  add_markers(frame = ~year, ids = ~country, showlegend = FALSE)
```


Per capita income of Belgium
income
40k
35k
30k
25k
20k
15k
1960
1970
1980
1990
2000
2010
2020
year
1960
Play
1960 1964 1968 1972 1976 1980 1984 1988 1992 1996 2000 2004 2008 2012 2016

Cumulative animations

```
library(dplyr)
library(purrr)
belgium %>%
  split(.$year) %>%
  accumulate(~bind_rows(.x, .y)) %>%
  set_names(1960:2018) %>%
  bind_rows(.id = "frame")
```

Coping with staggered starting points



CRAN download data

```
glimpse(monthly_logs)
```

```
Observations: 171
Variables: 4
$ package   <chr> "ggvis", "ggvis", "ggvis", "ggvis", "ggvis", "ggvis",…
$ date      <date> 2014-06-30, 2014-07-31, 2014-08-31, 2014-09-30, 2014…
$ dec_date  <dbl> 2014.49, 2014.58, 2014.66, 2014.75, 2014.83, 2014.91,…
$ downloads <dbl> 1344, 2120, 2035, 1702, 3590, 2899, 2427, 2227, 2708,…
```

What if we ignore the baseline issue?

```
monthly_logs %>%
  split(f = .$dec_date) %>%
  accumulate(., ~bind_rows(.x, .y)) %>%
  bind_rows(.id = "frame") %>%
  plot_ly(x = ~dec_date, y = ~downloads) %>%
  add_lines(color = ~package, frame = ~frame, ids = ~package)
```

```
Warning message:
In p$x$data[firstFrame] <- p$x$frames[[1]]$data :
  number of items to replace is not a multiple of replacement length
```

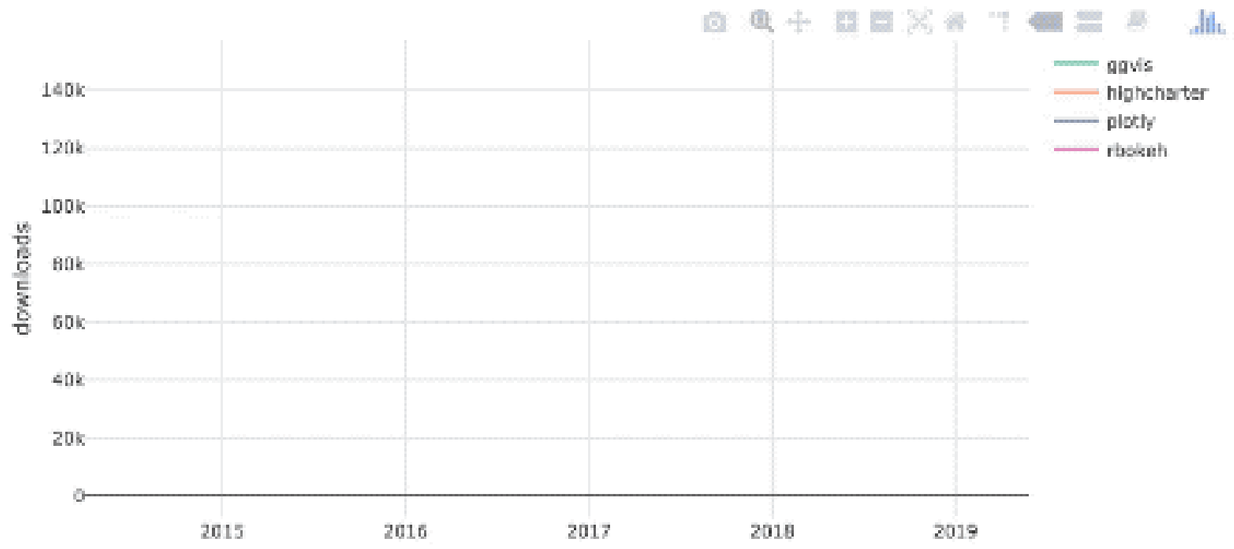
Completing the data set

```
library(tidyr)
complete_logs <- monthly_logs %>% complete(package, dec_date, fill = list(downloads = 0))
arrange(complete_logs, dec_date)
```

```
## A tibble: 228 x 4
  package     dec_date date       downloads
  <chr>          <dbl> <date>         <dbl>
1 ggvis          2014. 2014-06-30      1344
2 highcharter    2014. NA                 0
3 plotly         2014. NA                 0
4 rbokeh         2014. NA                 0
5 ggvis          2015. 2014-07-31      2120
# … with 223 more rows
```

Animating the completed data

```
complete_logs %>%
  split(f = .$dec_date) %>%
  accumulate(., ~bind_rows(.x, .y)) %>%
  bind_rows(.id = "frame") %>%
  plot_ly(x = ~dec_date, y = ~downloads) %>%
  add_lines(color = ~package, frame = ~frame, ids = ~package)
```



Let's practice!

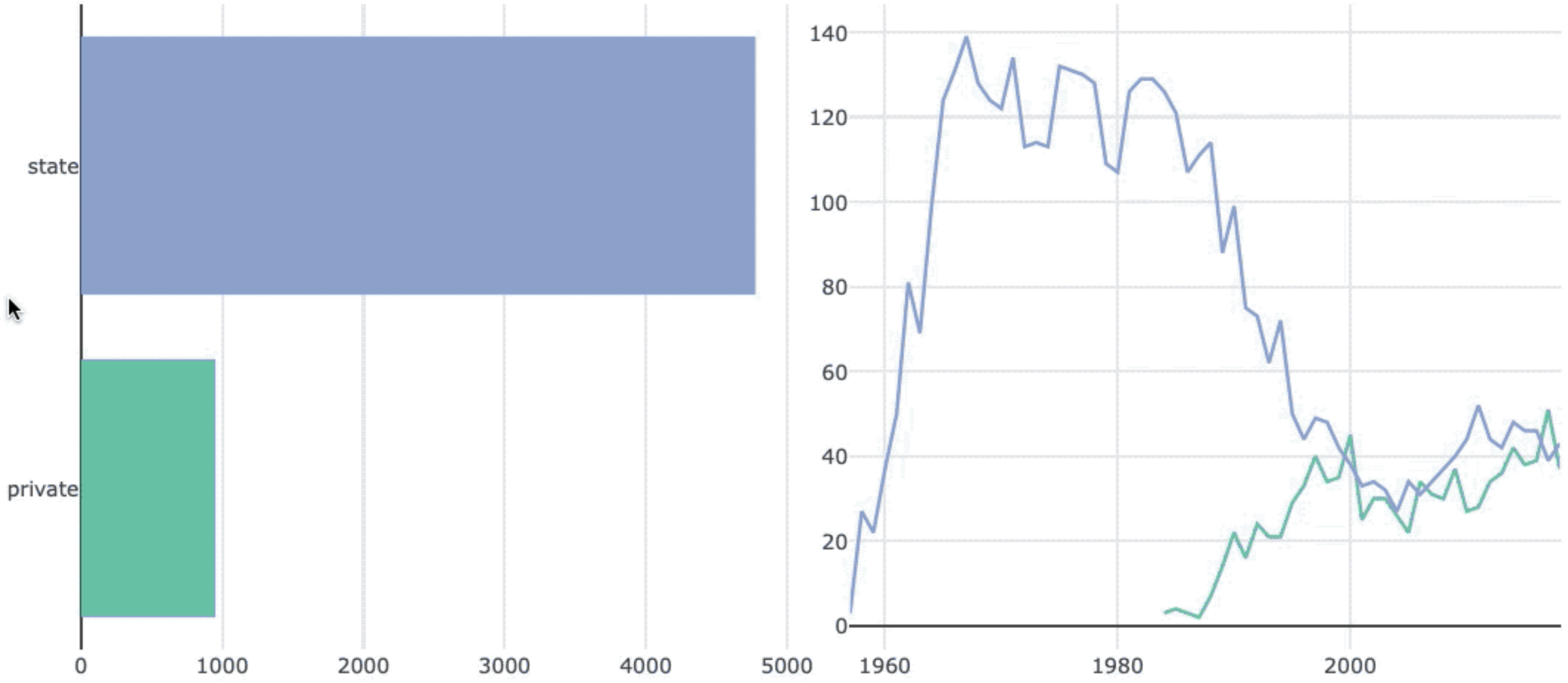
INTERMEDIATE INTERACTIVE DATA VISUALIZATION WITH PLOTLY IN R

Recap: linked views and selector widgets

INTERMEDIATE INTERACTIVE DATA VISUALIZATION WITH PLOTLY IN R

Adam Loy
Statistician, Carleton College

Linked views

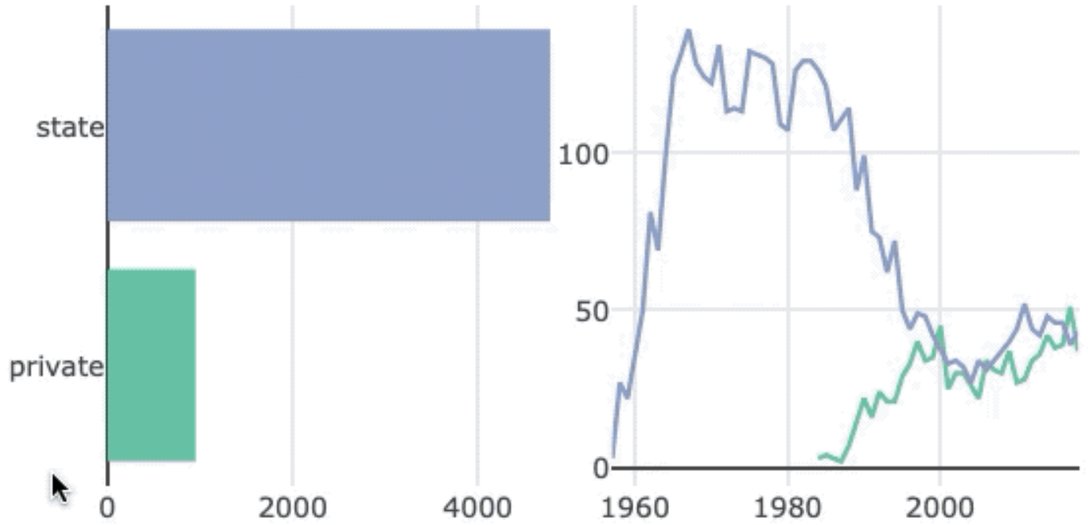


SharedData

```
shared_launches <- SharedData$new(launches, key = ~agency_type)
line_chart <- shared_launches %>%
  plot_ly(x = ~launch_year, y = ~n, color = ~agency_type) %>%
  count(launch_year, agency_type) %>%
  add_lines()  %>%
  hide_legend()
bar_chart <- shared_launches %>%
  plot_ly(y = ~fct_reorder(agency_type, n), x = ~n, color = ~agency_type) %>%
  count(agency_type) %>%
  add_bars() %>%
  layout(barmode = "overlay", yaxis = list(title = "")) %>%
  hide_legend()
```

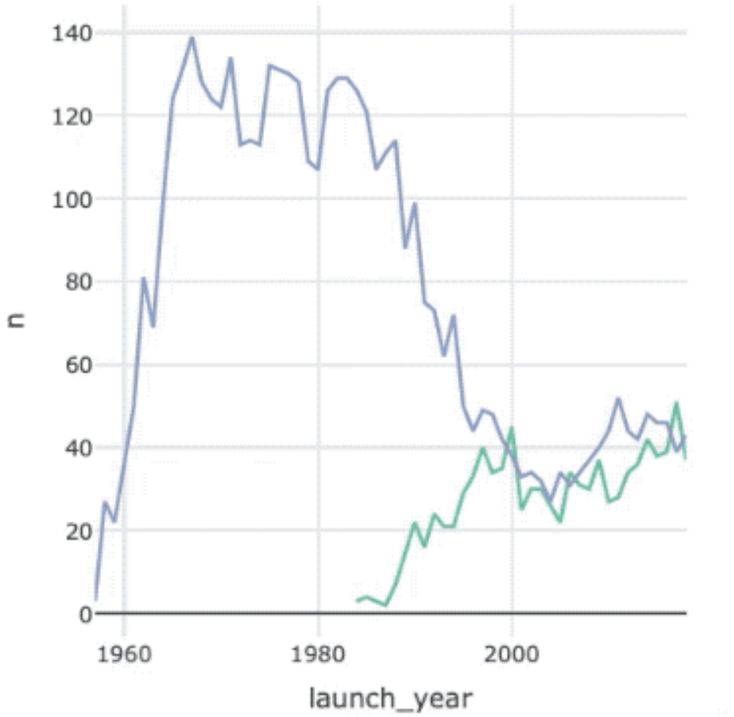
Linking views with subplot()

```
subplot(bar_chart, line_chart) %>%
  hide_legend() %>%
  highlight()
```



Linking views with bscols()

```
bscols(
  widths = c(4, NA),
  launch_state %>% highlight(),
  launch_ts %>% highlight()
)
```

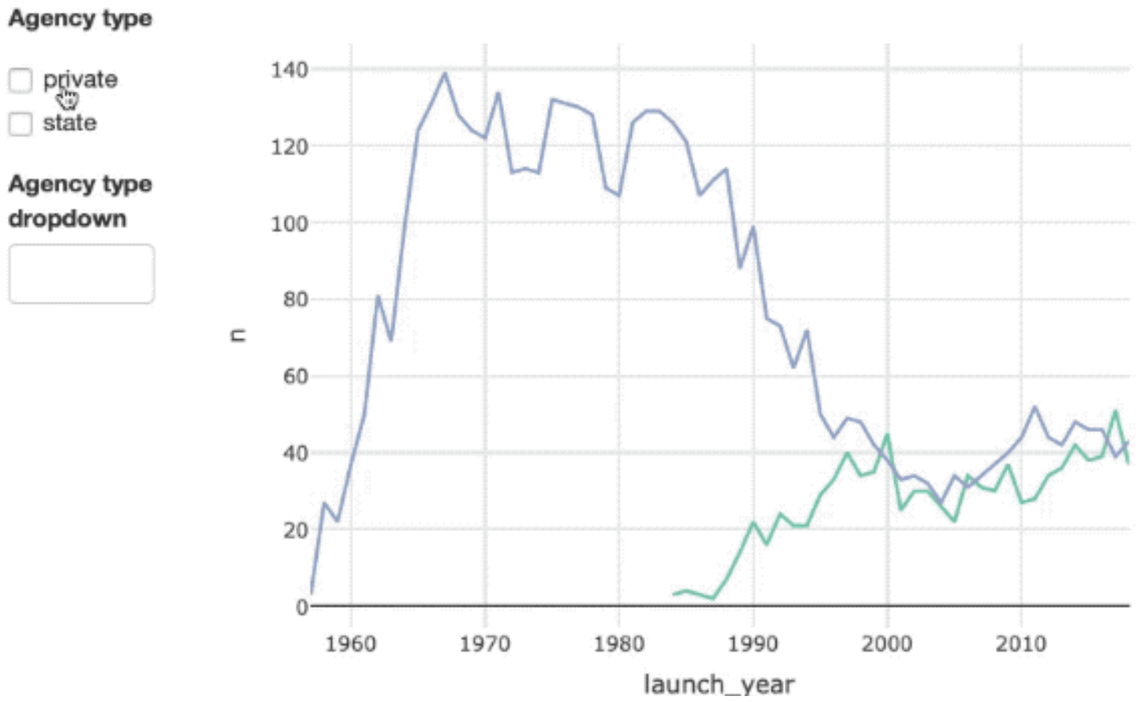


highlight() options

Argument	Description
`on`	selection event: `'plotly_click'`, `'plotly_hover'` or `'plotly_selected'`
`off`	event to turn off selection: `'plotly_doubleclick'`, `'plotly_deselect'`, or `'plotly_relayout'`
`persistent`	Should selections be persisent? `TRUE` or `FALSE`
`dynamic`	Add a widget to change colors? `TRUE` or `FALSE`
`color`	string of color(s) to use for highlighting selections
`selectize`	Add a selectize.js widget for selecting keys? `TRUE` or `FALSE`

Selector widgets

```
bscols(widths = c(2, NA),
  list(filter_checkbox(id = "agency", label = "Agency type", shared_launches, ~agency_type),
       filter_select(id = "agency2", label = "Agency type dropdown", shared_launches, ~agency_type)),
  line_chart %>% highlight(on = "plotly_selected", off = "plotly_deselect")
)
```



Launch vehicles

```
glimpse(lv)
```

```
# A tibble: 1,578 x 17
  name  family sfamily manufacturer variant alias min_stage max_stage length diameter
  <chr> <chr>  <chr>   <chr>        <chr>   <chr>      <dbl>     <dbl>  <dbl>    <dbl>
1 ?     Unkno… Unknown NA           NA      NA             1         1    NA       NA
2 Unkn… Unkno… Unknown NA           NA      NA             1         1    NA       NA
3 N-1 … N-1    N-1     OKB1         NA      NA             1         3   105.      14
4 Satu… Satur… SaturnV MSFC         NA      NA             1         4   111.      10.1
5 Satu… Satur… SaturnV MSFC         2       NA             1         3   105.      10.1
# … with 1,573 more rows, and 7 more variables: launch_mass <dbl>, leo_capacity <dbl>,
#   gto_capacity <dbl>, to_thrust <dbl>, class <chr>, apogee <dbl>, range <dbl>
```

Let's practice!

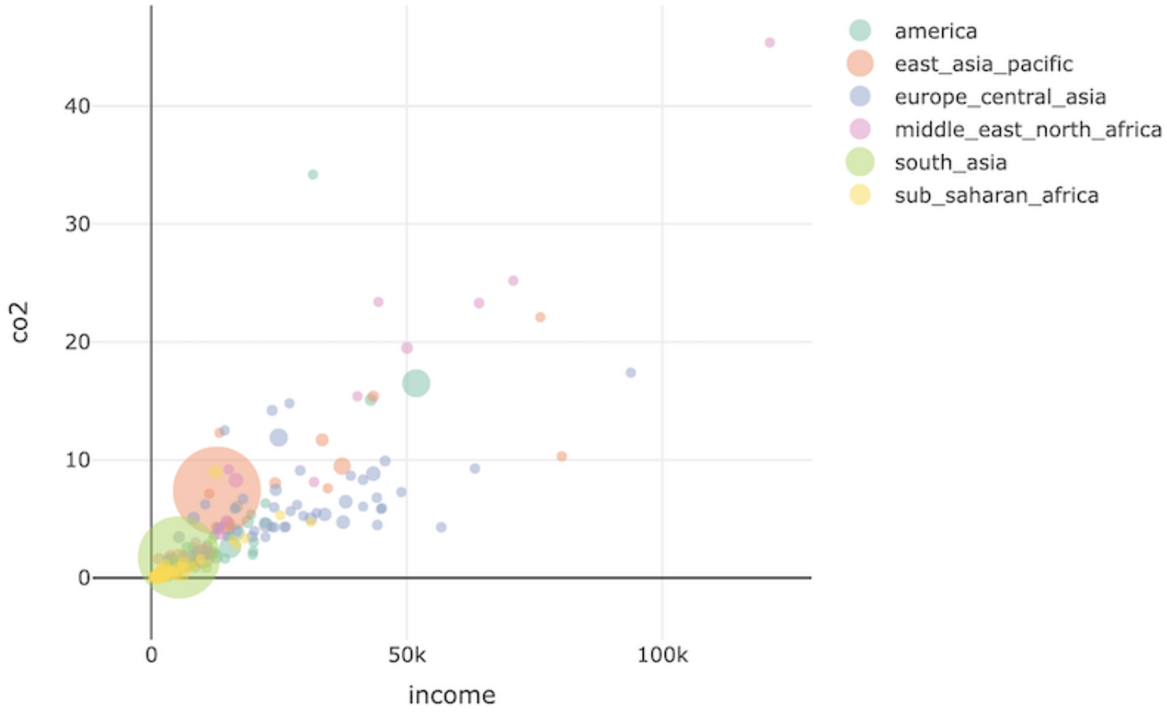
INTERMEDIATE INTERACTIVE DATA VISUALIZATION WITH PLOTLY IN R

Wrap-up

INTERMEDIATE INTERACTIVE DATA VISUALIZATION WITH PLOTLY IN R

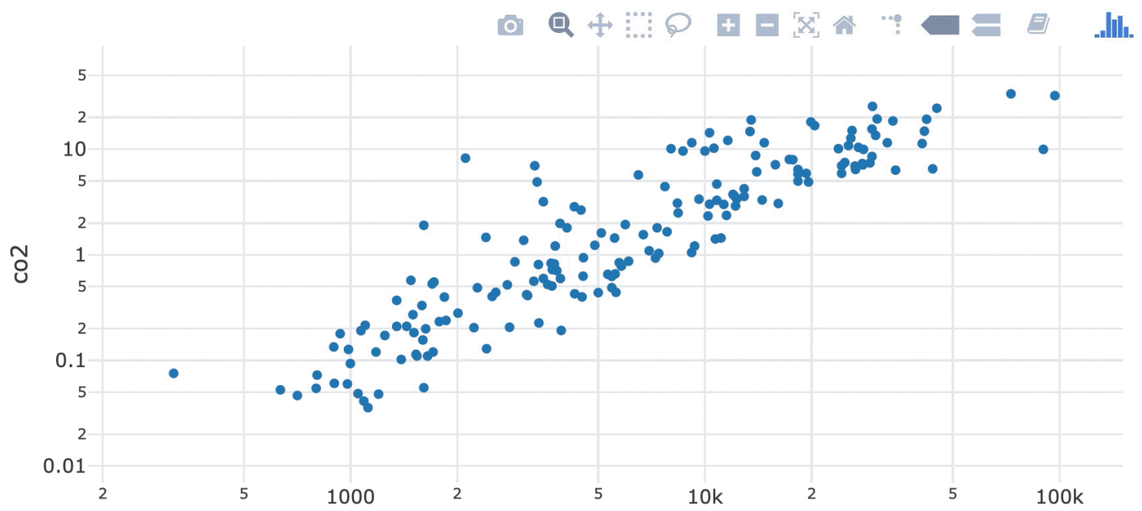
Adam Loy
Statistician, Carleton College

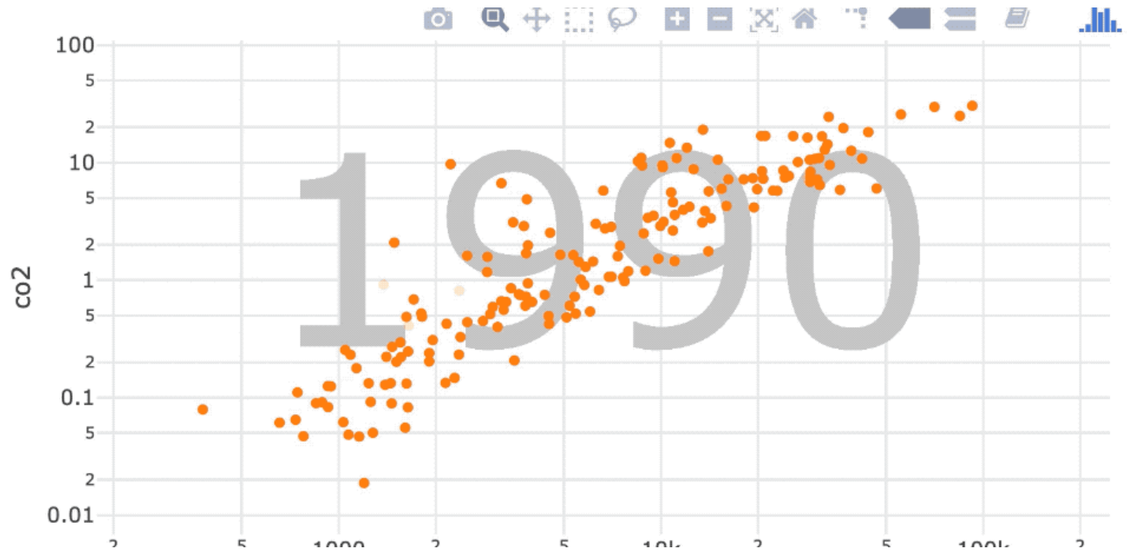
Chapter 1: plotly review

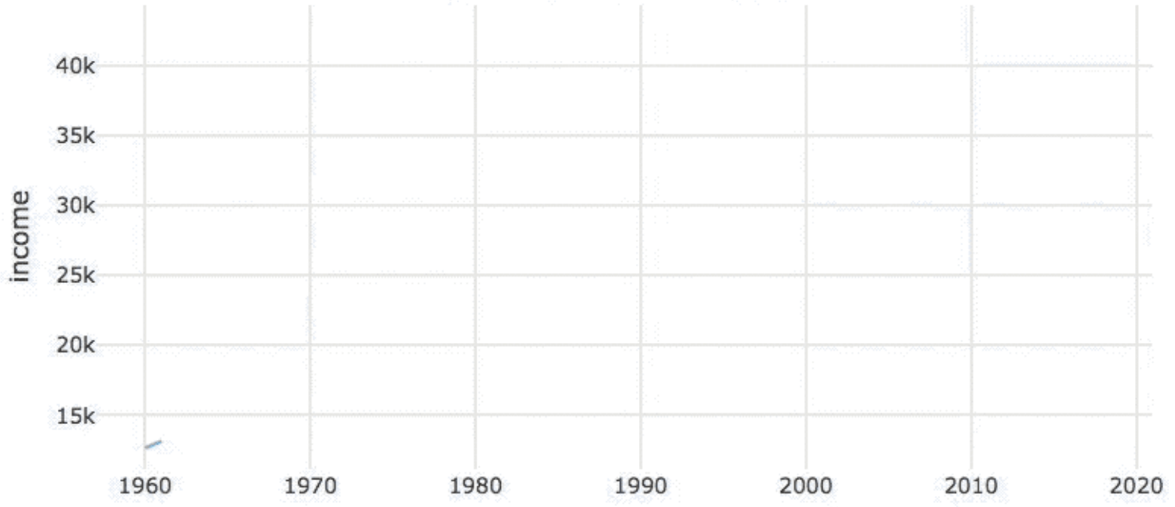


- Changing marker color, size, and symbols
- Polishing the `layout()` to provide context

Chapter 2: Animating your charts

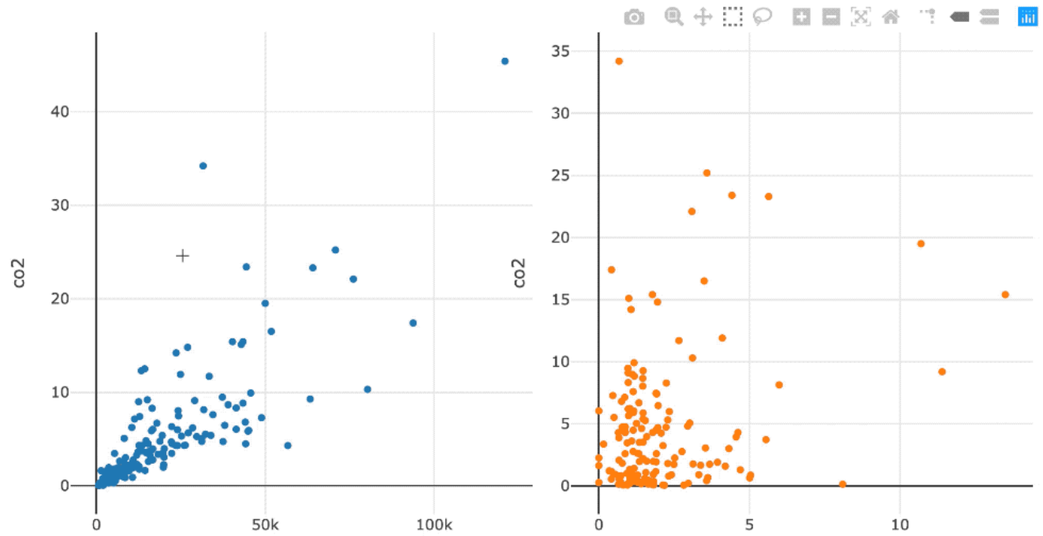






Chapter 3: Linked views and shinier charts

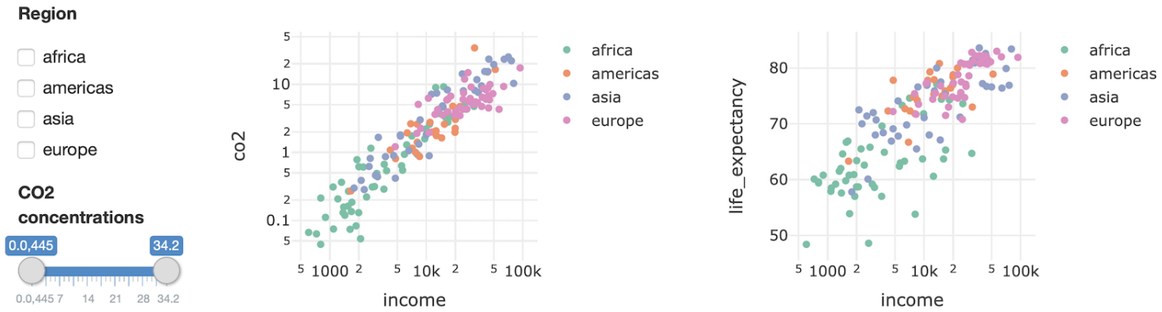
`crosstalk::SharedData$new()`



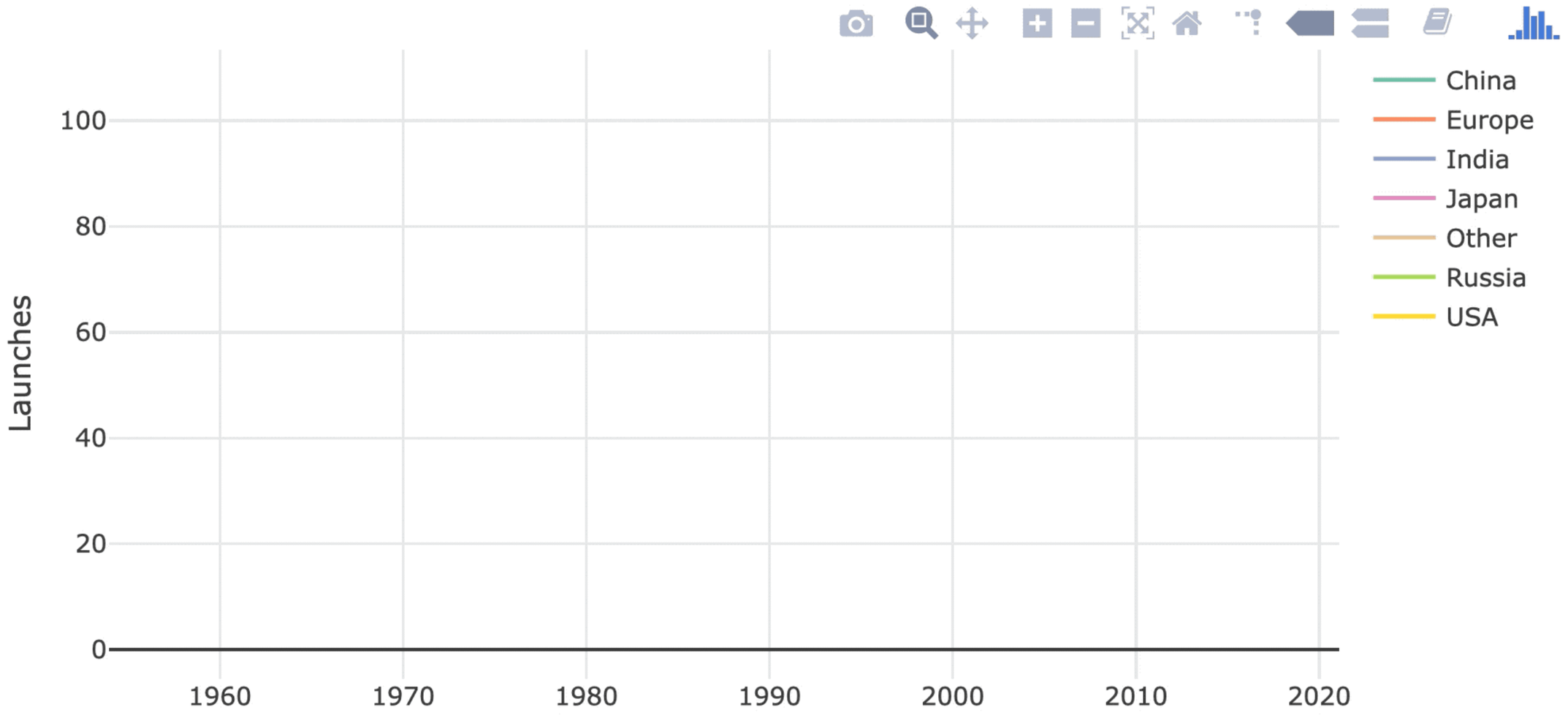
Selection

- Persistent vs. transient
- Direct vs. indirect

`crosstalk::bscols()`



Chapter 4: Space launches case study



Where to go from here

Explore more interactive plotting libraries

- leaflet
- highcharter
- trelliscope
- rbokeh

Learn Shiny

- plotly for R, by Carson Sievert
- DataCamp courses on Shiny

Thank you!

INTERMEDIATE INTERACTIVE DATA VISUALIZATION WITH PLOTLY IN R