

Data Visualization. 9 -

Case Studies

Kieran Healy

Code Horizons

December 10, 2023

Some Case Studies

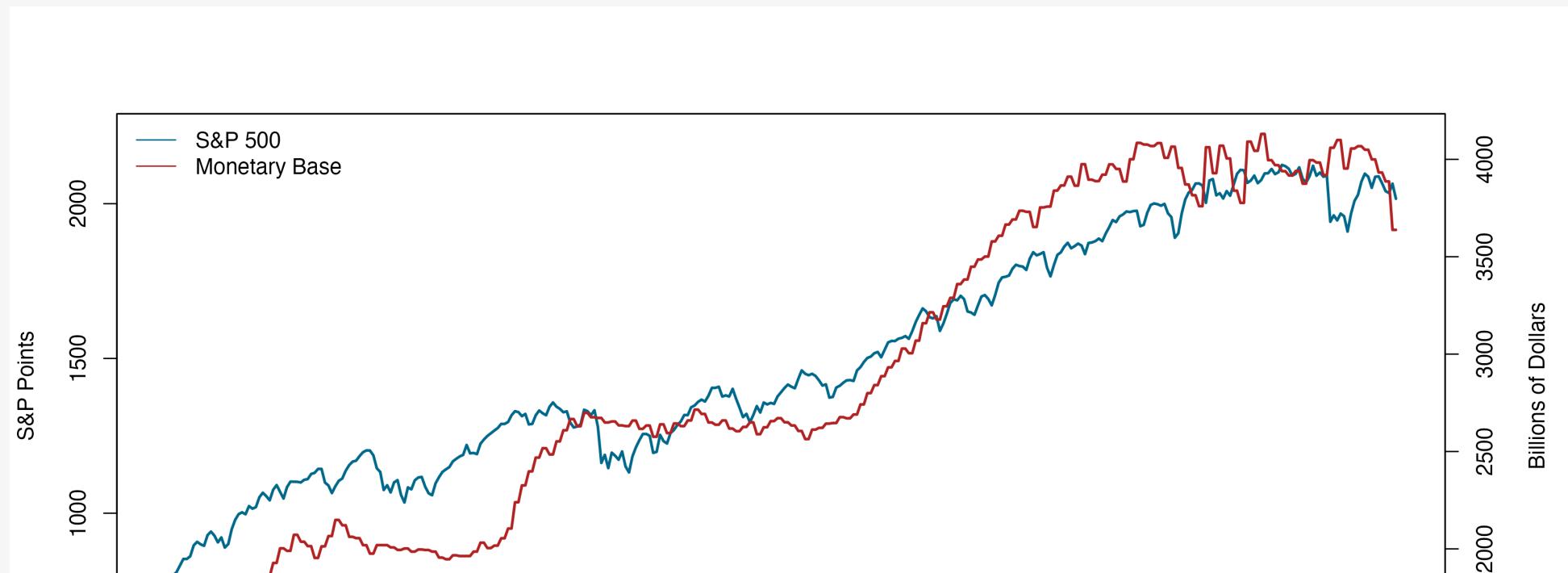
Load our packages

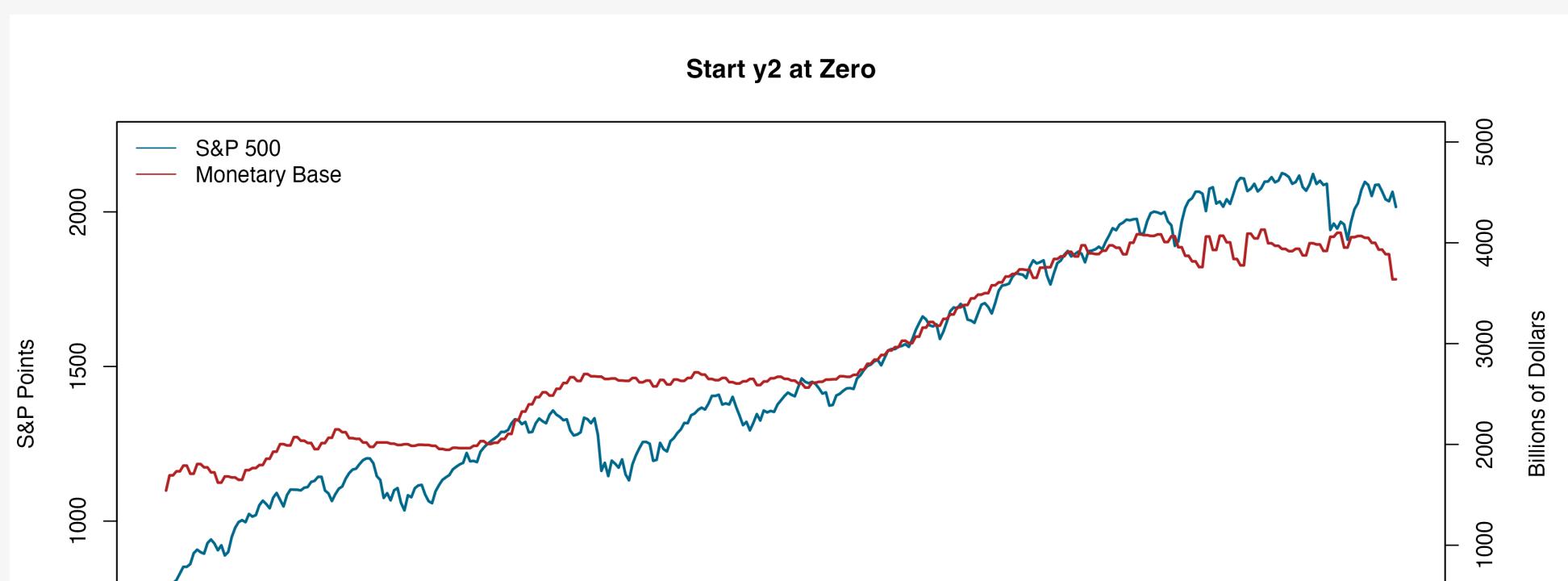
```
library(here)      # manage file paths
library(tidyverse) # your friend and mine
library(cavax)     # california vaccination exemption data
library(colorspace) # luminance-balanced palettes
library(demog)      # demographic data for a graph
library(ggforce)    # useful enhancements to ggplot
library(ggrepel)    # Text and labels
library(gssr)       # the gss packaged for r
library(patchwork)  # compose multiple plots

library(scales)     # scale adjustments and enhancements
library(socviz)     # data and some useful functions
```

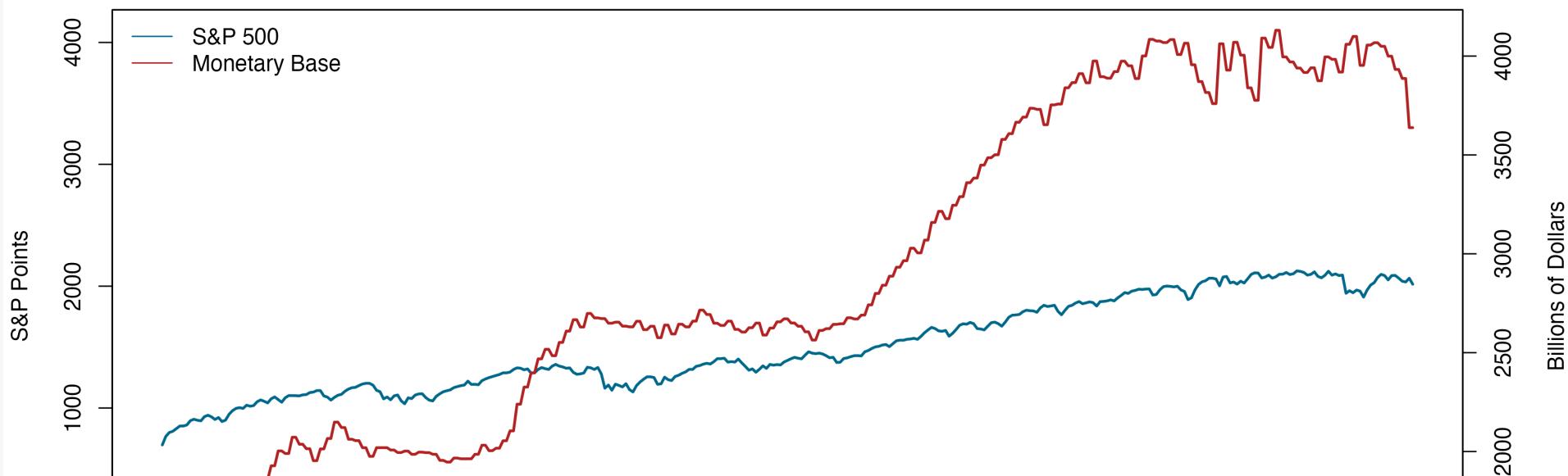
Two y-axes

≡





Start y1 at Zero; Max both at Max y2



What to do instead?

```
fredts ← as_tibble(fredts)
fredts

# A tibble: 357 × 5
  date      sp500 monbase sp500_i monbase_i
  <date>    <dbl>   <int>    <dbl>     <dbl>
1 2009-03-11  697. 1542228    100      100
2 2009-03-18  767. 1693133    110.     110.
3 2009-03-25  799. 1693133    115.     110.
4 2009-04-01  809. 1733017    116.     112.
5 2009-04-08  831. 1733017    119.     112.
6 2009-04-15  852. 1789878    122.     116.
7 2009-04-22  852. 1789878    122.     116.
8 2009-04-29  861. 1709369    124.     111.
9 2009-05-06  896. 1709369    129.     111.
10 2009-05-13 908. 1805373    130.     117.
# i 347 more rows
```

Pivot the data

```
1 fredts
```

```
# A tibble: 357 × 5
  date      sp500 monbase sp500_i monbase_i
  <date>    <dbl>   <int>    <dbl>     <dbl>
1 2009-03-11  697. 1542228    100      100
2 2009-03-18  767. 1693133    110.     110.
3 2009-03-25  799. 1693133    115.     110.
4 2009-04-01  809. 1733017    116.     112.
5 2009-04-08  831. 1733017    119.     112.
6 2009-04-15  852. 1789878    122.     116.
7 2009-04-22  852. 1789878    122.     116.
8 2009-04-29  861. 1709369    124.     111.
9 2009-05-06  896. 1709369    129.     111.
10 2009-05-13 908. 1805373   130.     117.
# i 347 more rows
```

Pivot the data

```
1 fredts %>%  
2   select(date, sp500_i, monbase_i)
```

```
# A tibble: 357 x 3  
  date       sp500_i  monbase_i  
  <date>     <dbl>    <dbl>  
1 2009-03-11    100     100  
2 2009-03-18    110.    110.  
3 2009-03-25    115.    110.  
4 2009-04-01    116.    112.  
5 2009-04-08    119.    112.  
6 2009-04-15    122.    116.  
7 2009-04-22    122.    116.  
8 2009-04-29    124.    111.  
9 2009-05-06    129.    111.  
10 2009-05-13   130.    117.  
# i 347 more rows
```

Pivot the data

```
1 fredts %>
2   select(date, sp500_i, monbase_i) %>
3   pivot_longer(sp500_i:monbase_i,
4                 names_to = "series",
5                 values_to = "score")
```

```
# A tibble: 714 x 3
  date      series    score
  <date>     <chr>     <dbl>
1 2009-03-11 sp500_i    100
2 2009-03-11 monbase_i  100
3 2009-03-18 sp500_i    110.
4 2009-03-18 monbase_i  110.
5 2009-03-25 sp500_i    115.
6 2009-03-25 monbase_i  110.
7 2009-04-01 sp500_i    116.
8 2009-04-01 monbase_i  112.
9 2009-04-08 sp500_i    119.
10 2009-04-08 monbase_i  112.
# i 704 more rows
```

Pivot the data

```
1 fredts >
2   select(date, sp500_i, monbase_i)  >
3   pivot_longer(sp500_i:monbase_i,
4                 names_to = "series",
5                 values_to = "score") ->
6   fredts_m
```

Pivot the data

```
1 fredts >
2   select(date, sp500_i, monbase_i)  >
3     pivot_longer(sp500_i:monbase_i,
4                   names_to = "series",
5                   values_to = "score") ->
6   fredts_m
```

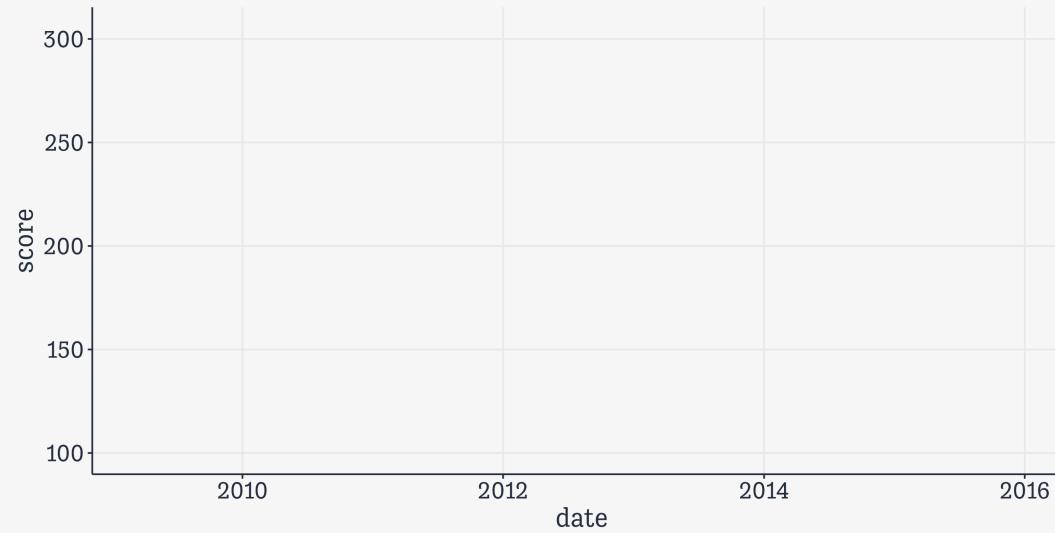
Make two plots

```
1 fredts_m
```

```
# A tibble: 714 × 3
  date      series    score
  <date>    <chr>     <dbl>
1 2009-03-11 sp500_i    100
2 2009-03-11 monbase_i  100
3 2009-03-18 sp500_i    110.
4 2009-03-18 monbase_i  110.
5 2009-03-25 sp500_i    115.
6 2009-03-25 monbase_i  110.
7 2009-04-01 sp500_i    116.
8 2009-04-01 monbase_i  112.
9 2009-04-08 sp500_i    119.
10 2009-04-08 monbase_i 112.
# i 704 more rows
```

Make two plots

```
1 fredts_m %>%  
2   ggplot(mapping =  
3           aes(x = date,  
4                     y = score,  
5                     color = series))
```



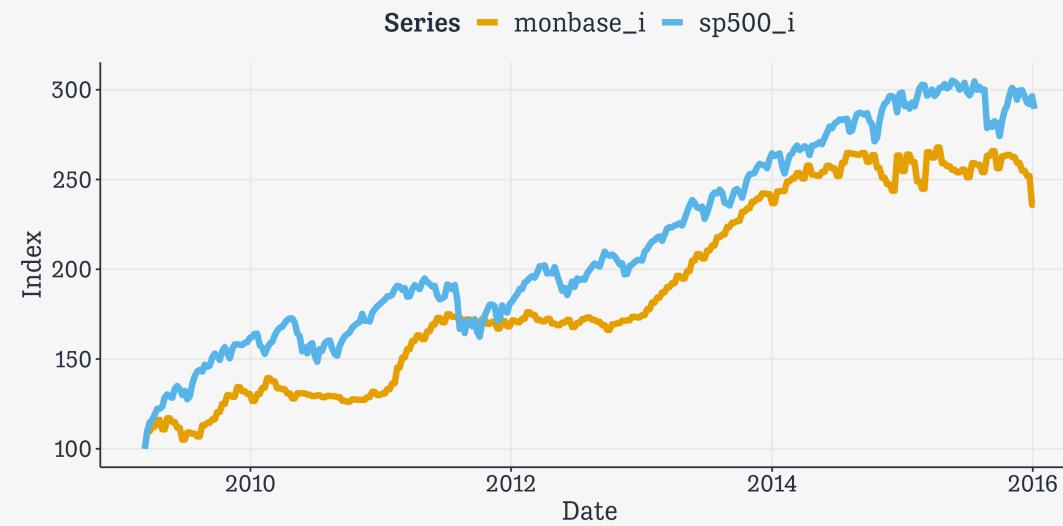
Make two plots

```
1 fredts_m >
2   ggplot(mapping =
3     aes(x = date,
4           y = score,
5           color = series)) +
6   geom_line(linewidth = 2)
```



Make two plots

```
1 fredts_m %>%
2   ggplot(mapping =
3     aes(x = date,
4          y = score,
5          color = series)) +
6   geom_line(linewidth = 2) +
7   labs(x = "Date", y = "Index",
8        color = "Series")
```



Make two plots

```
1 fredts_m >
2   ggplot(mapping =
3     aes(x = date,
4           y = score,
5           color = series)) +
6   geom_line(linewidth = 2) +
7   labs(x = "Date", y = "Index",
8         color = "Series") +
9   theme(axis.title.x = element_blank(),
10      axis.text.x = element_blank(),
11      axis.ticks.x = element_blank())
```



Make two plots

```
1 fredts_m >
2   ggplot(mapping =
3     aes(x = date,
4           y = score,
5           color = series)) +
6   geom_line(linewidth = 2) +
7   labs(x = "Date", y = "Index",
8         color = "Series") +
9   theme(axis.title.x = element_blank(),
10         axis.text.x = element_blank(),
11         axis.ticks.x = element_blank()) ->
12   p1
```

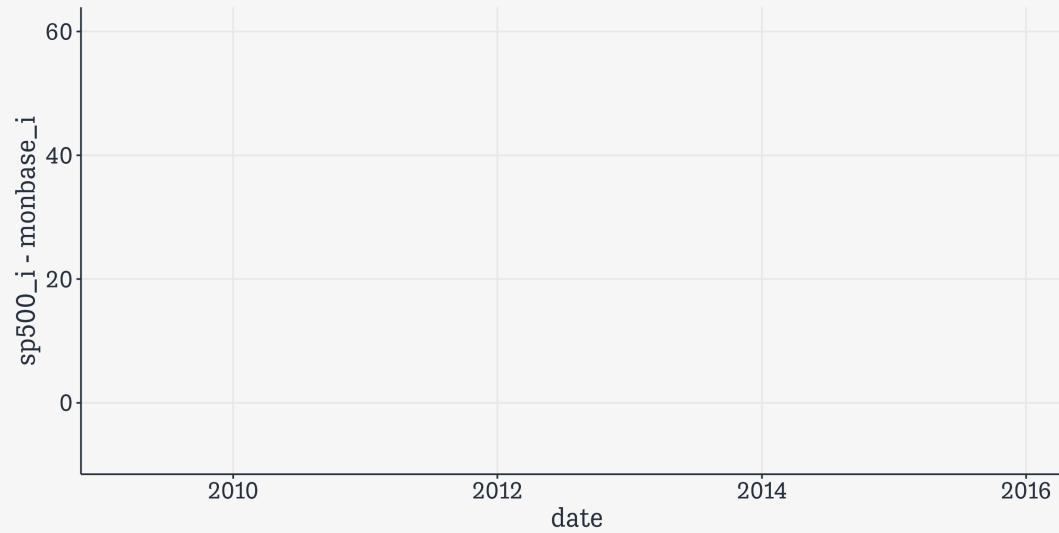
Make two plots

```
1 fredts_m >
2   ggplot(mapping =
3     aes(x = date,
4           y = score,
5           color = series)) +
6   geom_line(linewidth = 2) +
7   labs(x = "Date", y = "Index",
8         color = "Series") +
9   theme(axis.title.x = element_blank(),
10         axis.text.x = element_blank(),
11         axis.ticks.x = element_blank()) →
12   p1
13
14 # The original df
15 fredts
```

```
# A tibble: 357 × 5
  date      sp500 monbase sp500_i monbase_i
  <date>    <dbl>  <int>    <dbl>    <dbl>
1 2009-03-11  697. 1542228    100     100
2 2009-03-18  767. 1693133    110.    110.
3 2009-03-25  799. 1693133    115.    110.
4 2009-04-01  809. 1733017    116.    112.
5 2009-04-08  831. 1733017    119.    112.
6 2009-04-15  852. 1789878    122.    116.
7 2009-04-22  852. 1789878    122.    116.
8 2009-04-29  861. 1709369    124.    111.
9 2009-05-06  896. 1709369    129.    111.
10 2009-05-13 908. 1805373    130.    117.
# i 347 more rows
```

Make two plots

```
1 fredts_m >
2   ggplot(mapping =
3     aes(x = date,
4          y = score,
5          color = series)) +
6   geom_line(linewidth = 2) +
7   labs(x = "Date", y = "Index",
8        color = "Series") +
9   theme(axis.title.x = element_blank(),
10      axis.text.x = element_blank(),
11      axis.ticks.x = element_blank()) →
12 p1
13
14 # The original df
15 fredts >
16   ggplot(mapping =
17     aes(x = date,
18          y = sp500_i - monbase_i))
```



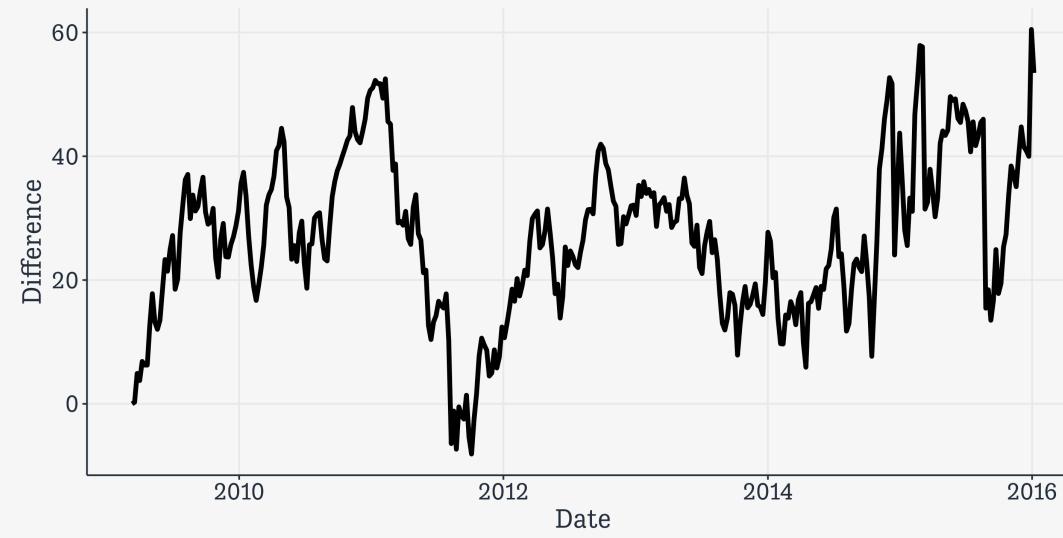
Make two plots

```
1 fredts_m >
2   ggplot(mapping =
3     aes(x = date,
4          y = score,
5          color = series)) +
6   geom_line(linewidth = 2) +
7   labs(x = "Date", y = "Index",
8        color = "Series") +
9   theme(axis.title.x = element_blank(),
10      axis.text.x = element_blank(),
11      axis.ticks.x = element_blank()) →
12 p1
13
14 # The original df
15 fredts >
16   ggplot(mapping =
17     aes(x = date,
18          y = sp500_i - monbase_i)) +
19   geom_line(linewidth = 1.5)
```



Make two plots

```
1 fredts_m >
2   ggplot(mapping =
3     aes(x = date,
4          y = score,
5          color = series)) +
6   geom_line(linewidth = 2) +
7   labs(x = "Date", y = "Index",
8        color = "Series") +
9   theme(axis.title.x = element_blank(),
10      axis.text.x = element_blank(),
11      axis.ticks.x = element_blank()) →
12 p1
13
14 # The original df
15 fredts >
16   ggplot(mapping =
17     aes(x = date,
18          y = sp500_i - monbase_i)) +
19   geom_line(linewidth = 1.5) +
20   labs(x = "Date", y = "Difference")
```



Make two plots

```
1 fredts_m >
2   ggplot(mapping =
3     aes(x = date,
4          y = score,
5          color = series)) +
6   geom_line(linewidth = 2) +
7   labs(x = "Date", y = "Index",
8        color = "Series") +
9   theme(axis.title.x = element_blank(),
10      axis.text.x = element_blank(),
11      axis.ticks.x = element_blank()) →
12 p1
13
14 # The original df
15 fredts >
16   ggplot(mapping =
17     aes(x = date,
18          y = sp500_i - monbase_i)) +
19   geom_line(linewidth = 1.5) +
20   labs(x = "Date", y = "Difference") →
21 p2
```

Make two plots

```
1 fredts_m >
2   ggplot(mapping =
3     aes(x = date,
4          y = score,
5          color = series)) +
6   geom_line(linewidth = 2) +
7   labs(x = "Date", y = "Index",
8        color = "Series") +
9   theme(axis.title.x = element_blank(),
10      axis.text.x = element_blank(),
11      axis.ticks.x = element_blank()) →
12 p1
13
14 # The original df
15 fredts >
16   ggplot(mapping =
17     aes(x = date,
18          y = sp500_i - monbase_i)) +
19   geom_line(linewidth = 1.5) +
20   labs(x = "Date", y = "Difference") →
21 p2
```

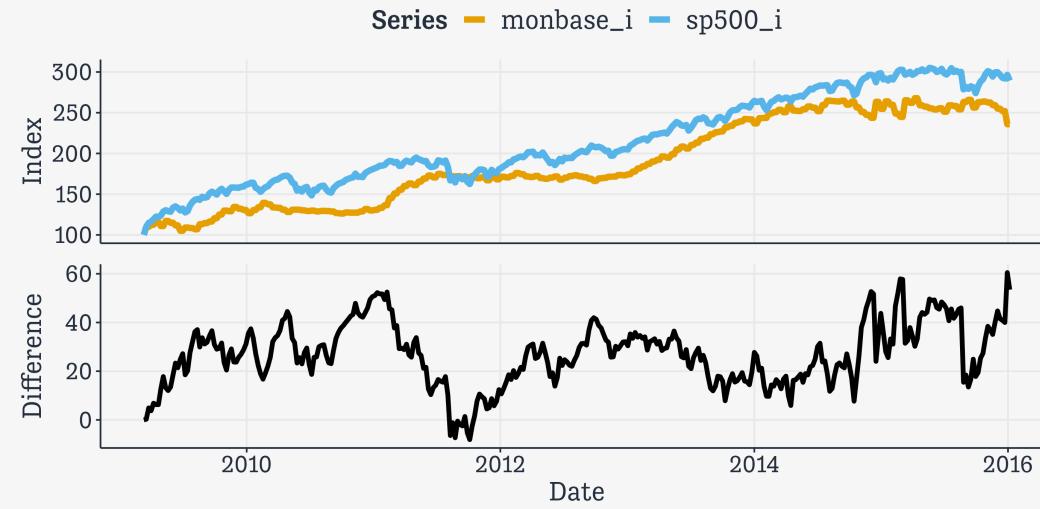
Combine with patchwork

```
1 library(patchwork)
```



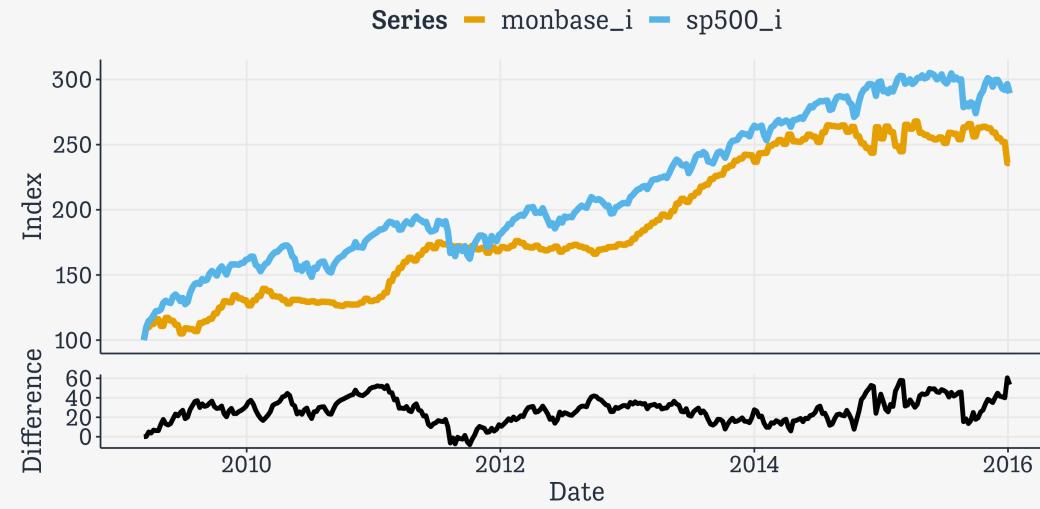
Combine with patchwork

```
1 library(patchwork)
2
3 (p1 / p2)
```



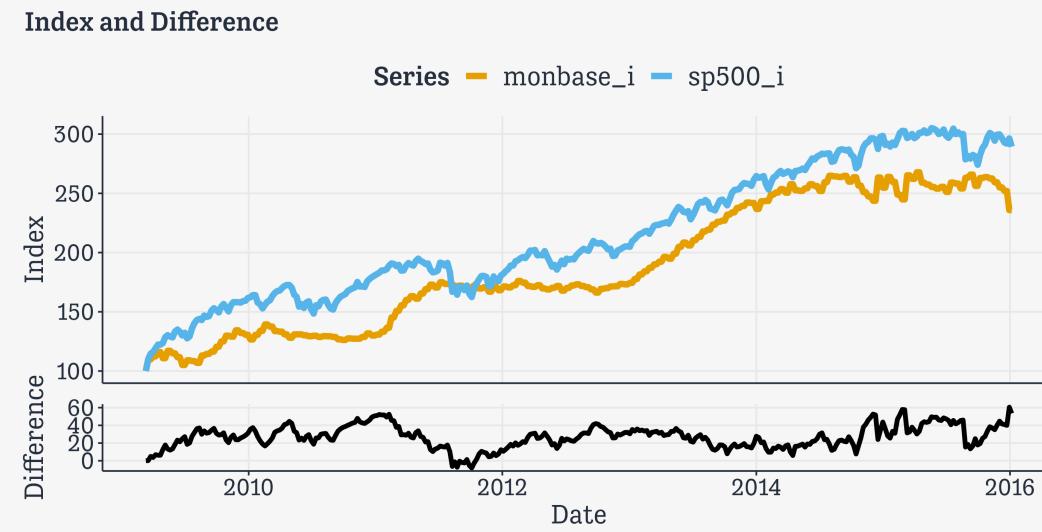
Combine with patchwork

```
1 library(patchwork)
2
3 (p1 / p2) +
4   plot_layout(heights = c(4, 1))
```



Combine with patchwork

```
1 library(patchwork)
2
3 (p1 / p2) +
4   plot_layout(heights = c(4, 1)) +
5   plot_annotation(title = "Index and Difference")
```

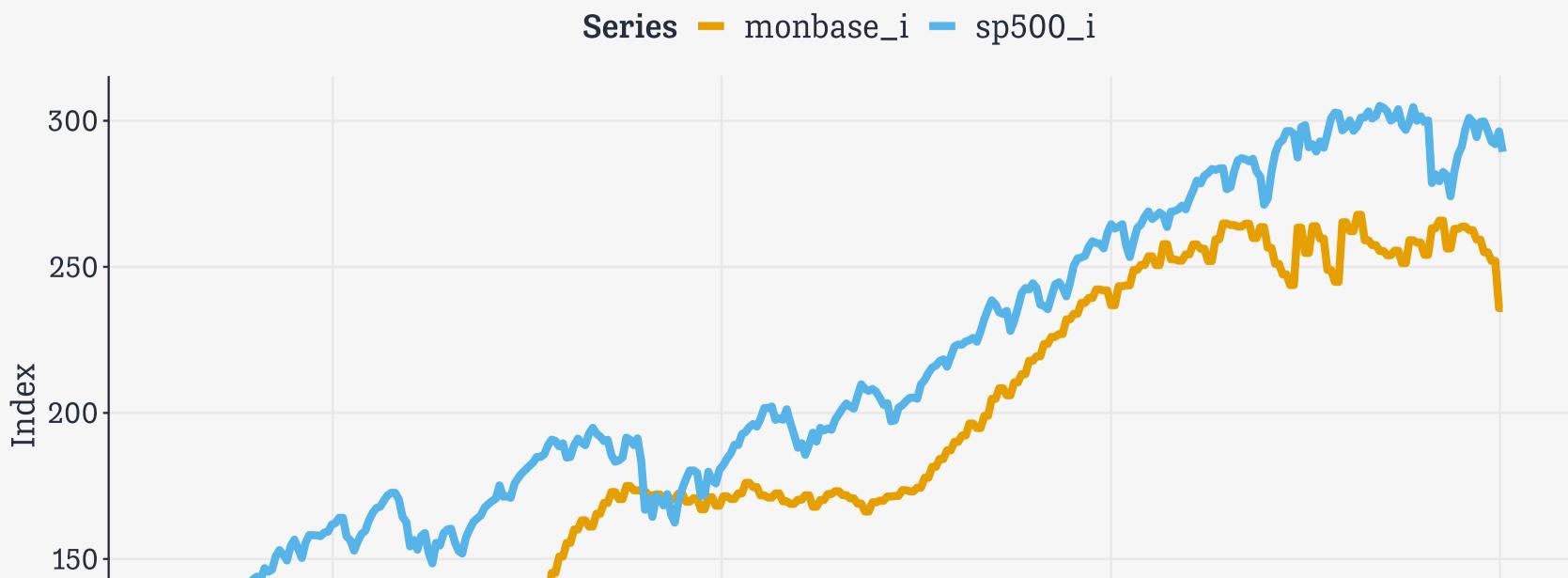


Combine with patchwork

```
1 library(patchwork)
2
3 (p1 / p2) +
4   plot_layout(heights = c(4, 1)) +
5   plot_annotation(title = "Index and Difference"
6   p_patch
```



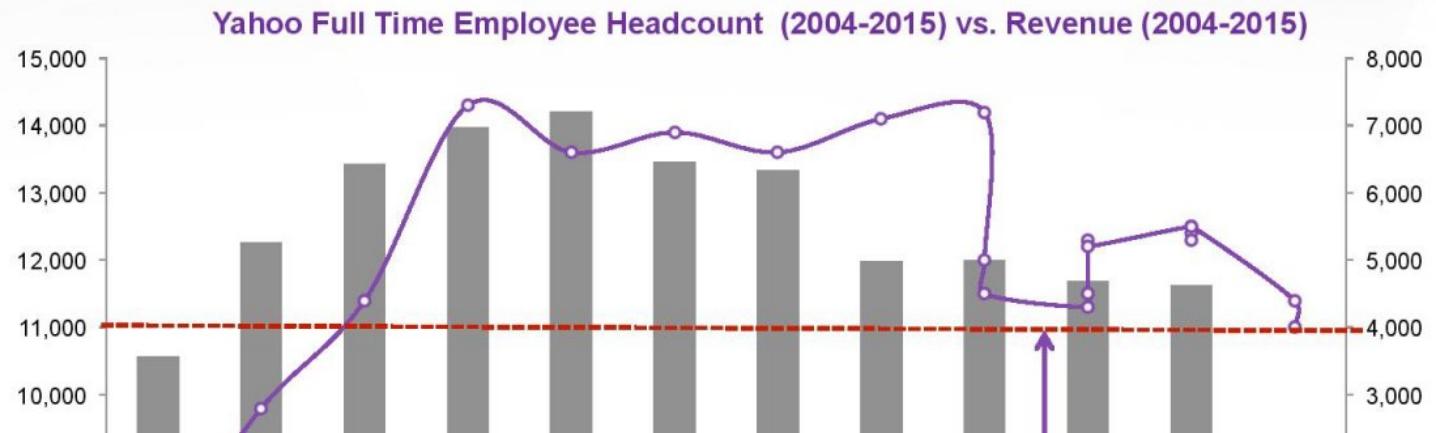
Index and Difference



≡

Redrawing a bad slide

Yahoo's Headcount Still Excessively High Given Revenues:



The data

```
yahoo
```

```
# A tibble: 12 × 4
  Year Revenue Employees Mayer
  <dbl>    <dbl>      <dbl> <chr>
1 2004     3574       7600 No
2 2005     5257       9800 No
3 2006     6425      11400 No
4 2007     6969      14300 No
5 2008     7208      13600 No
6 2009     6460      13900 No
7 2010     6324      13600 No
8 2011     4984      14100 No
9 2012     4986      12000 No
10 2012    4986      11500 Yes
11 2013     4680      12200 Yes
12 2014     4618      12500 Yes
```

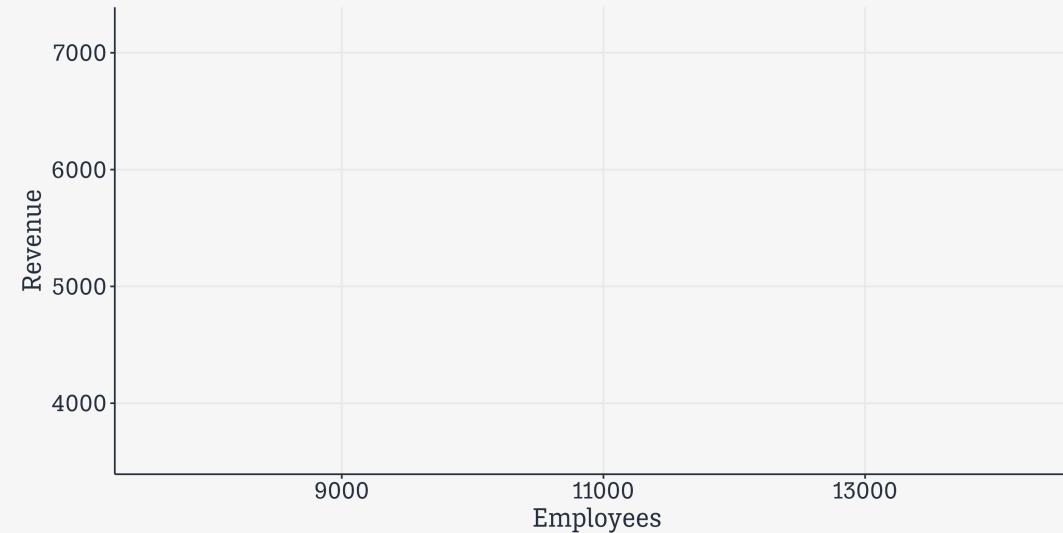
Option 1

```
1 yahoo
```

```
# A tibble: 12 × 4
  Year Revenue Employees Mayer
  <dbl>    <dbl>      <dbl> <chr>
1 2004     3574      7600 No
2 2005     5257      9800 No
3 2006     6425     11400 No
4 2007     6969     14300 No
5 2008     7208     13600 No
6 2009     6460     13900 No
7 2010     6324     13600 No
8 2011     4984     14100 No
9 2012     4986     12000 No
10 2012    4986     11500 Yes
11 2013    4680     12200 Yes
12 2014    4618     12500 Yes
```

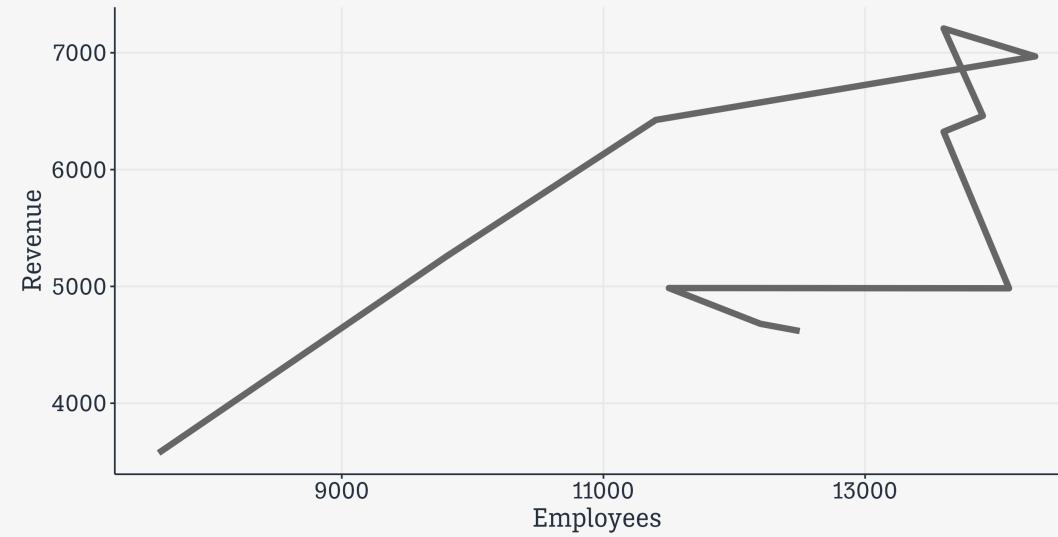
Option 1

```
1 yahoo %>
2   ggplot(mapping =
3     aes(x = Employees,
4          y = Revenue))
```



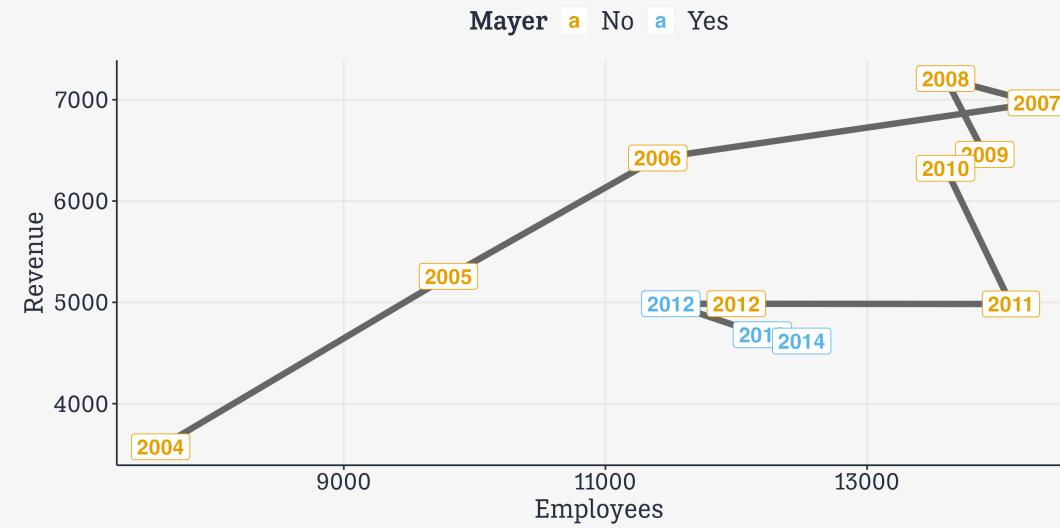
Option 1

```
1 yahoo %>
2   ggplot(mapping =
3     aes(x = Employees,
4          y = Revenue)) +
5   geom_path(color = "gray40",
6             linewidth = rel(2))
```



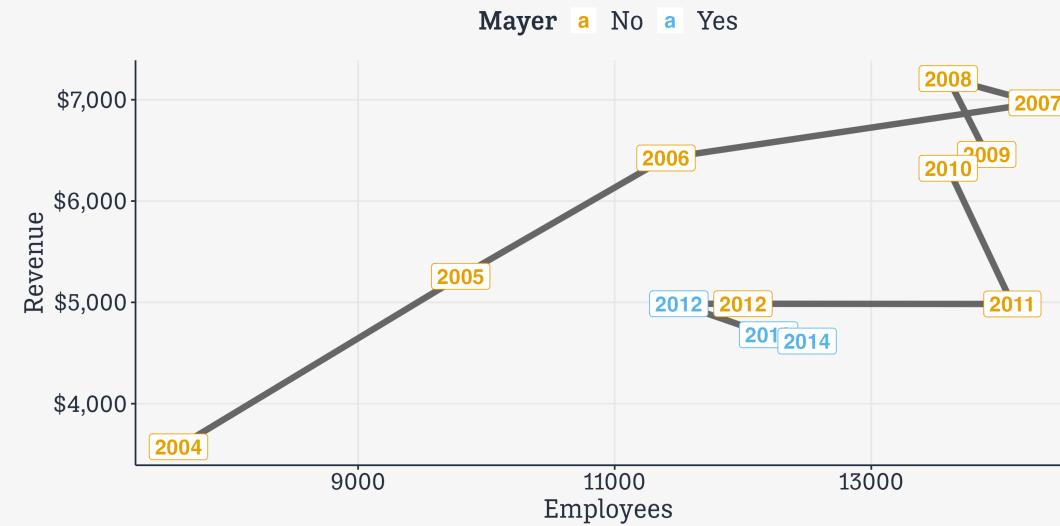
Option 1

```
1 yahoo >
2   ggplot(mapping =
3     aes(x = Employees,
4           y = Revenue)) +
5   geom_path(color = "gray40",
6             linewidth = rel(2)) +
7   geom_label(aes(color = Mayer,
8                 label = Year),
9             size = rel(5),
10            fontface = "bold")
```



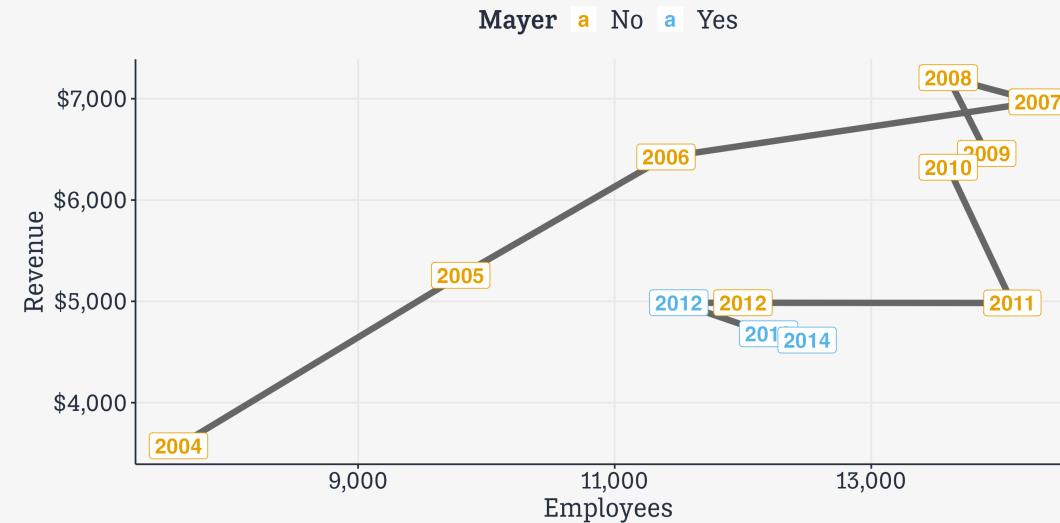
Option 1

```
1 yahoo %>
2   ggplot(mapping =
3     aes(x = Employees,
4          y = Revenue)) +
5   geom_path(color = "gray40",
6             linewidth = rel(2)) +
7   geom_label(aes(color = Mayer,
8                label = Year),
9              size = rel(5),
10             fontface = "bold") +
11   scale_y_continuous(labels = label_dollar())
```



Option 1

```
1 yahoo >
2   ggplot(mapping =
3     aes(x = Employees,
4           y = Revenue)) +
5   geom_path(color = "gray40",
6             linewidth = rel(2)) +
7   geom_label(aes(color = Mayer,
8                 label = Year),
9             size = rel(5),
10            fontface = "bold") +
11   scale_y_continuous(labels = label_dollar()) +
12   scale_x_continuous(labels = label_comma())
```



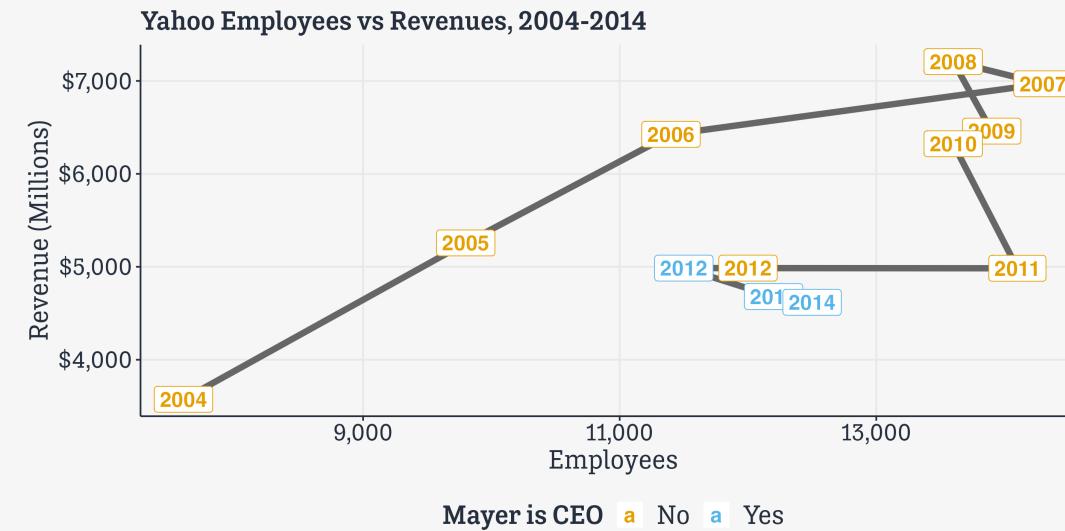
Option 1

```
1 yahoo >
2   ggplot(mapping =
3     aes(x = Employees,
4           y = Revenue)) +
5   geom_path(color = "gray40",
6             linewidth = rel(2)) +
7   geom_label(aes(color = Mayer,
8                 label = Year),
9             size = rel(5),
10            fontface = "bold") +
11   scale_y_continuous(labels = label_dollar()) +
12   scale_x_continuous(labels = label_comma()) +
13   theme(legend.position = "bottom")
```



Option 1

```
1 yahoo %>
2   ggplot(mapping =
3     aes(x = Employees,
4          y = Revenue)) +
5   geom_path(color = "gray40",
6             linewidth = rel(2)) +
7   geom_label(aes(color = Mayer,
8              label = Year),
9             size = rel(5),
10            fontface = "bold") +
11   scale_y_continuous(labels = label_dollar()) +
12   scale_x_continuous(labels = label_comma()) +
13   theme(legend.position = "bottom") +
14   labs(color = "Mayer is CEO",
15        x = "Employees", y = "Revenue (Millions)"
16        title = "Yahoo Employees vs Revenues, 2004-2014")
```

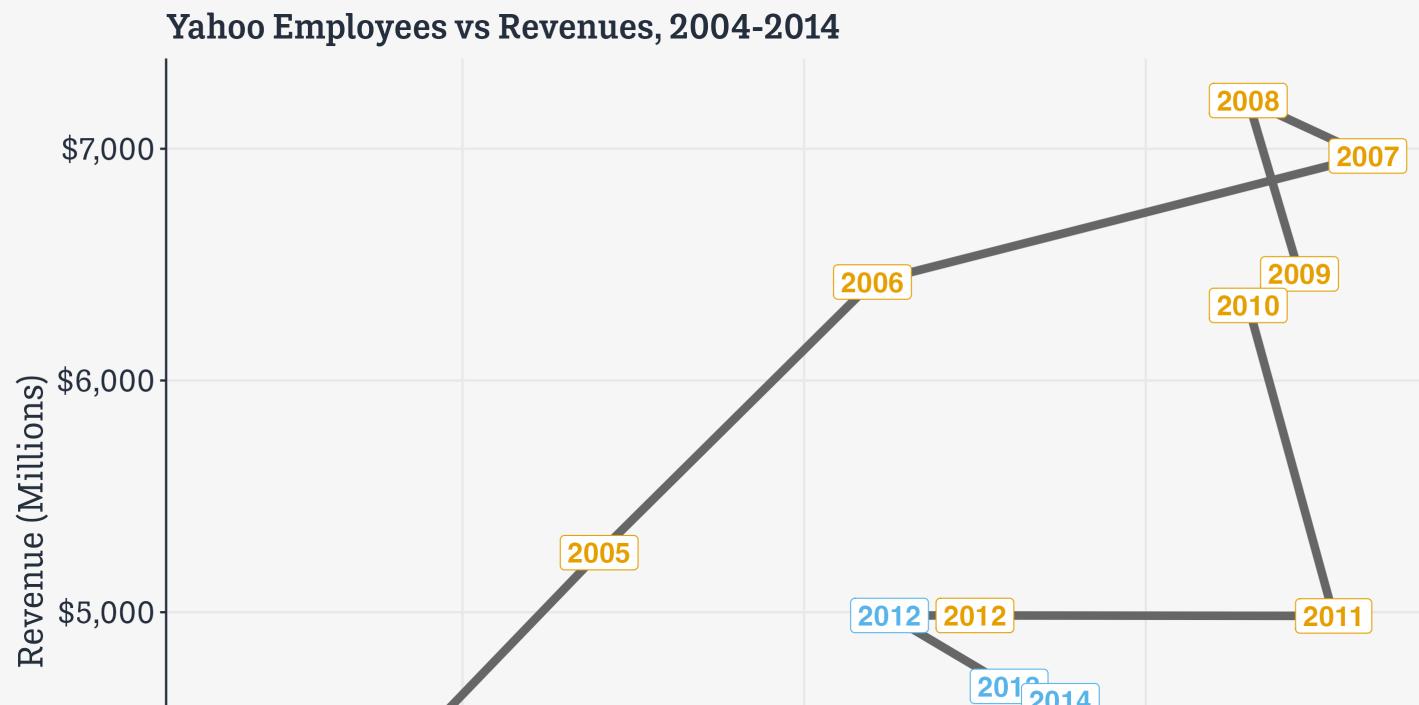


Option 1

```
1 yahoo >
2   ggplot(mapping =
3     aes(x = Employees,
4          y = Revenue)) +
5   geom_path(color = "gray40",
6             linewidth = rel(2)) +
7   geom_label(aes(color = Mayer,
8                label = Year),
9              size = rel(5),
10             fontface = "bold") +
11   scale_y_continuous(labels = label_dollar()) +
12   scale_x_continuous(labels = label_comma()) +
13   theme(legend.position = "bottom") +
14   labs(color = "Mayer is CEO",
15        x = "Employees", y = "Revenue (Millions)")
16   title = "Yahoo Employees vs Revenues, 200"
17 yahoo1
```

Option 1

```
1 yahoo >
2   ggplot(mapping =
3     aes(x = Employees,
4          y = Revenue)) +
5   geom_path(color = "gray40",
6             linewidth = rel(2)) +
7   geom_label(aes(color = Mayer,
8                 label = Year),
9              size = rel(5),
10             fontface = "bold") +
11   scale_y_continuous(labels = label_dollar()) +
12   scale_x_continuous(labels = label_comma()) +
13   theme(legend.position = "bottom") +
14   labs(color = "Mayer is CEO",
15        x = "Employees", y = "Revenue (Millions)")
16   title = "Yahoo Employees vs Revenues, 200"
17 yahoo1
```



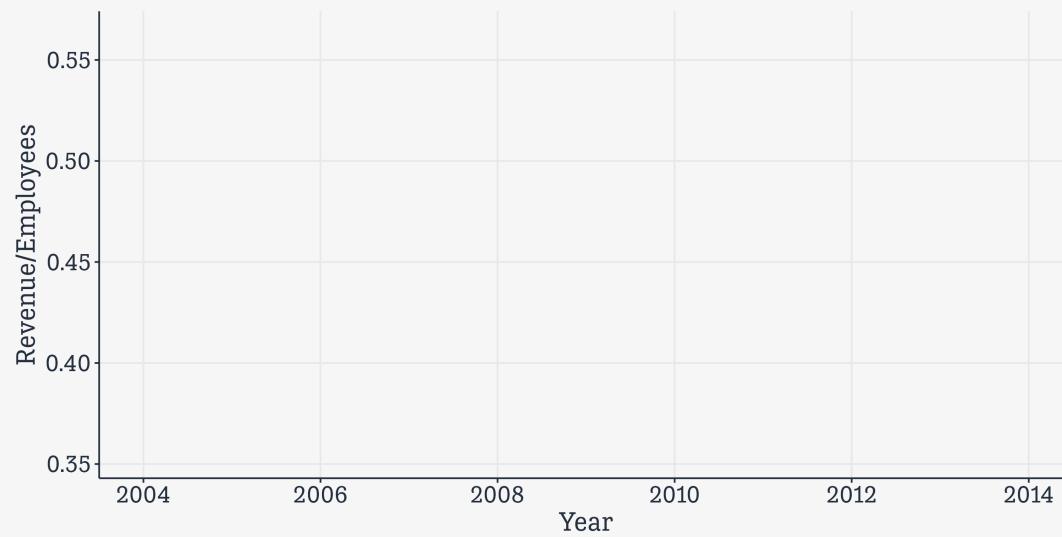
Alternatively ...

```
1 yahoo
```

```
# A tibble: 12 × 4
  Year Revenue Employees Mayer
  <dbl>    <dbl>      <dbl> <chr>
1 2004     3574      7600 No
2 2005     5257      9800 No
3 2006     6425     11400 No
4 2007     6969     14300 No
5 2008     7208     13600 No
6 2009     6460     13900 No
7 2010     6324     13600 No
8 2011     4984     14100 No
9 2012     4986     12000 No
10 2012    4986     11500 Yes
11 2013    4680     12200 Yes
12 2014    4618     12500 Yes
```

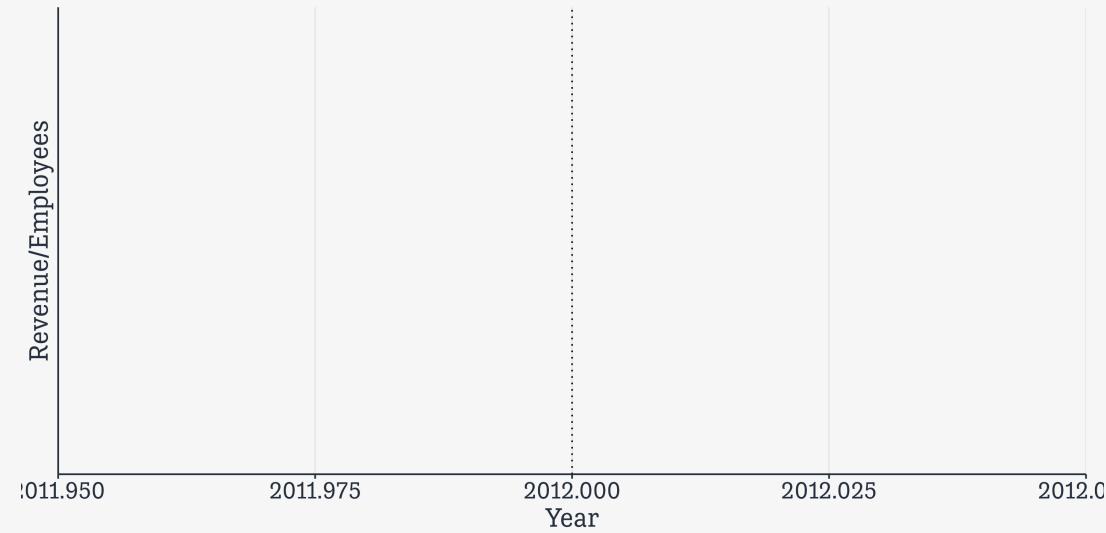
Alternatively ...

```
1 yahoo %>
2   ggplot(mapping =
3     aes(x = Year,
4          y = Revenue/Employees))
```



Alternatively ...

```
1 yahoo %>
2   ggplot(mapping =
3     aes(x = Year,
4          y = Revenue/Employees)) +
5   geom_vline(xintercept = 2012,
6              linewidth = rel(0.5),
7              linetype = "dotted")
```



Alternatively ...

```
1 yahoo %>
2   ggplot(mapping =
3     aes(x = Year,
4           y = Revenue/Employees)) +
5   geom_vline(xintercept = 2012,
6             linewidth = rel(0.5),
7             linetype = "dotted") +
8   geom_line(color = "royalblue", linewidth = rel
```



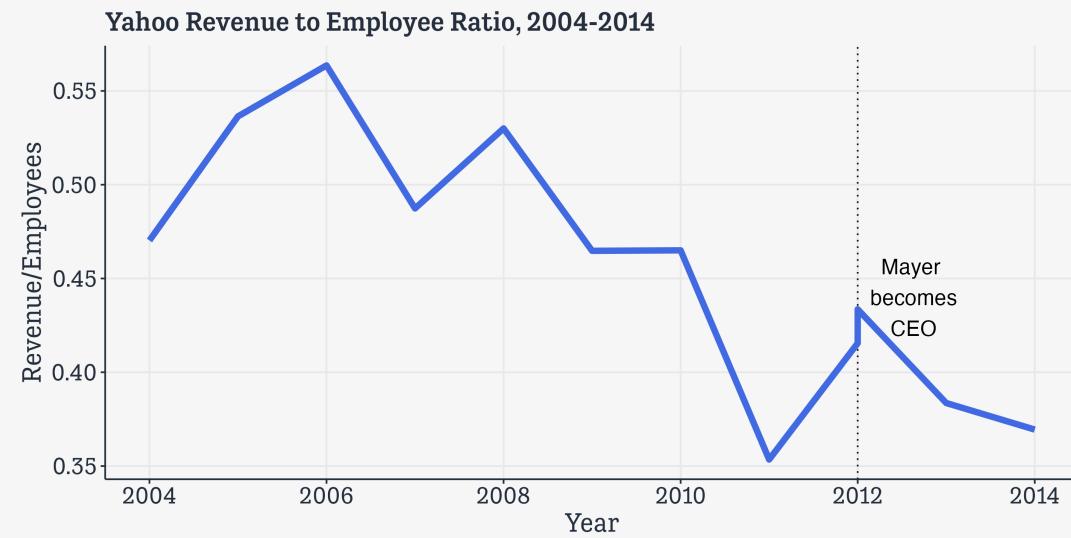
Alternatively ...

```
1 yahoo %>
2   ggplot(mapping =
3     aes(x = Year,
4           y = Revenue/Employees)) +
5   geom_vline(xintercept = 2012,
6             linewidth = rel(0.5),
7             linetype = "dotted") +
8   geom_line(color = "royalblue", linewidth = rel
9   annotate("text", x = 2012.6, y = 0.44,
10         label = "Mayer\n becomes\n CEO", size
```



Alternatively ...

```
1 yahoo %>
2   ggplot(mapping =
3     aes(x = Year,
4           y = Revenue/Employees)) +
5   geom_vline(xintercept = 2012,
6             linewidth = rel(0.5),
7             linetype = "dotted") +
8   geom_line(color = "royalblue", linewidth = rel
9             annotate("text", x = 2012.6, y = 0.44,
10                       label = "Mayer\n becomes\n CEO", size
11             labs(title = "Yahoo Revenue to Employee Ratio",
12                   x = "Year",
13                   y = "Revenue/Employees")
```



Alternatively ...

```
1 yahoo >
2   ggplot(mapping =
3     aes(x = Year,
4           y = Revenue/Employees)) +
5   geom_vline(xintercept = 2012,
6               linewidth = rel(0.5),
7               linetype = "dotted") +
8   geom_line(color = "royalblue", linewidth = rel
9   annotate("text", x = 2012.6, y = 0.44,
10         label = "Mayer\n becomes\n CEO", size
11   labs(title = "Yahoo Revenue to Employee Ratio,
12       x = "Year",
13       y = "Revenue/Employees") →
14   yahoo2
```

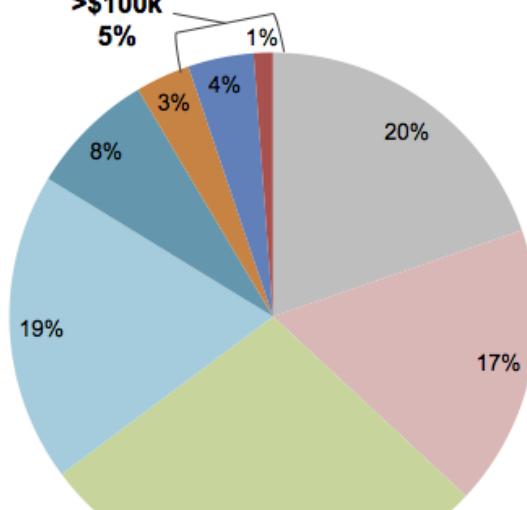
Yahoo Revenue to Employee Ratio, 2004-2014



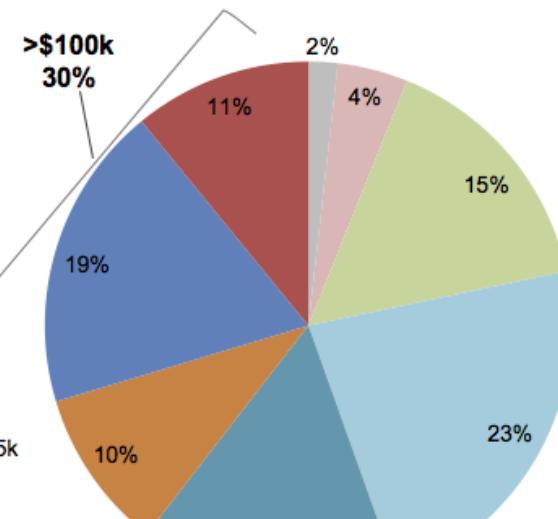
Say no to pie

Pie charts are easy to mess up

Borrower Distribution by Outstanding Balance
out of 44 million borrowers in 2016



Debt Distribution by Outstanding Balance
out of \$1.3 trillion in 2016



The data

```
studebt
```

```
# A tibble: 16 × 4
  Debt      type      pct Debtrc
  <ord>     <fct>    <int> <ord>
1 Under $5 Borrowers 20 Under $5
2 $5-$10  Borrowers 17 $5-$10
3 $10-$25 Borrowers 28 $10-$25
4 $25-$50 Borrowers 19 $25-$50
5 $50-$75 Borrowers  8 $50-$75
6 $75-$100 Borrowers  3 $75-$100
7 $100-$200 Borrowers  4 $100-$200
8 Over $200 Borrowers  1 Over $200
9 Under $5 Balances  2 Under $5
10 $5-$10 Balances  4 $5-$10
11 $10-$25 Balances 15 $10-$25
12 $25-$50 Balances 23 $25-$50
13 $50-$75 Balances 16 $50-$75
14 $75-$100 Balances 10 $75-$100
15 $100-$200 Balances 19 $100-$200
16 Over $200 Balances 11 Over $200
```

Debt and Debtrc are both ordered factors.



A little prep work

```
p_ylab ← "Amount Owed, in thousands of Dollars"
p_title ← "Outstanding Student Loans"
p_subtitle ← "44 million borrowers owe a total of $1.3 trillion"
p_caption ← "Source: FRB NY"

studebt ← studebt ▷
  mutate(type_label = recode(type, "Borrowers" = "Percent of all Borrowers",
                            "Balances" = "Percent of all Balances"))
```

```
studebt
```

#	Debt	type	pct	Debtrc	type_label
	<ord>	<fct>	<int>	<ord>	<fct>
1	Under \$5	Borrowers	20	Under \$5	Percent of all Borrowers
2	\$5-\$10	Borrowers	17	\$5-\$10	Percent of all Borrowers
3	\$10-\$25	Borrowers	28	\$10-\$25	Percent of all Borrowers
4	\$25-\$50	Borrowers	19	\$25-\$50	Percent of all Borrowers
5	\$50-\$75	Borrowers	8	\$50-\$75	Percent of all Borrowers
6	\$75-\$100	Borrowers	3	\$75-\$100	Percent of all Borrowers
7	\$100-\$200	Borrowers	4	\$100-\$200	Percent of all Borrowers
8	Over \$200	Borrowers	1	Over \$200	Percent of all Borrowers
9	Under \$5	Balances	2	Under \$5	Percent of all Balances
10	\$5-\$10	Balances	4	\$5-\$10	Percent of all Balances
11	\$10-\$25	Balances	15	\$10-\$25	Percent of all Balances
12	\$25-\$50	Balances	23	\$25-\$50	Percent of all Balances

Debt Plot 1

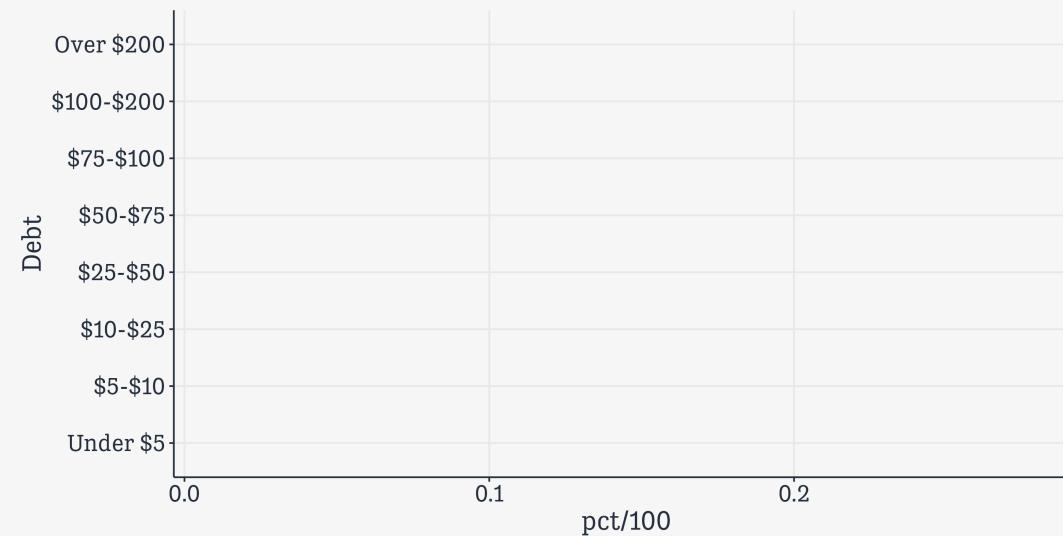
```
1 studebt
```

```
# A tibble: 16 × 5
  Debt      type     pct Debtrc    type_label
  <ord>     <fct>   <int> <ord>    <fct>
1 Under $5 Borrowers 20 Under $5 Percent of all Borrowers
2 $5-$10  Borrowers 17 $5-$10  Percent of all Borrowers
3 $10-$25 Borrowers 28 $10-$25 Percent of all Borrowers
4 $25-$50 Borrowers 19 $25-$50 Percent of all Borrowers
5 $50-$75 Borrowers  8 $50-$75 Percent of all Borrowers
6 $75-$100 Borrowers  3 $75-$100 Percent of all Borrowers
7 $100-$200 Borrowers  4 $100-$200 Percent of all Borrowers
8 Over $200 Borrowers  1 Over $200 Percent of all Borrowers
9 Under $5 Balances  2 Under $5 Percent of all Balances
10 $5-$10 Balances   4 $5-$10  Percent of all Balances
11 $10-$25 Balances  15 $10-$25 Percent of all Balances
12 $25-$50 Balances  23 $25-$50 Percent of all Balances
13 $50-$75 Balances  16 $50-$75 Percent of all Balances
14 $75-$100 Balances 10 $75-$100 Percent of all Balances
15 $100-$200 Balances 19 $100-$200 Percent of all Balances
16 Over $200 Balances 11 Over $200 Percent of all Balances
```



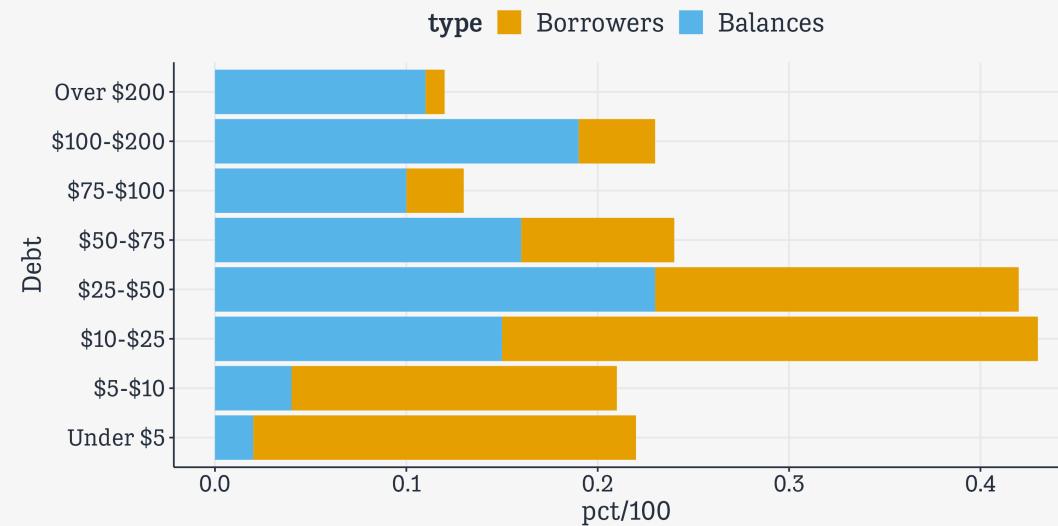
Debt Plot 1

```
1 studebt %>%
2   ggplot(mapping =
3     aes(x = pct/100,
4           y = Debt,
5           fill = type))
```



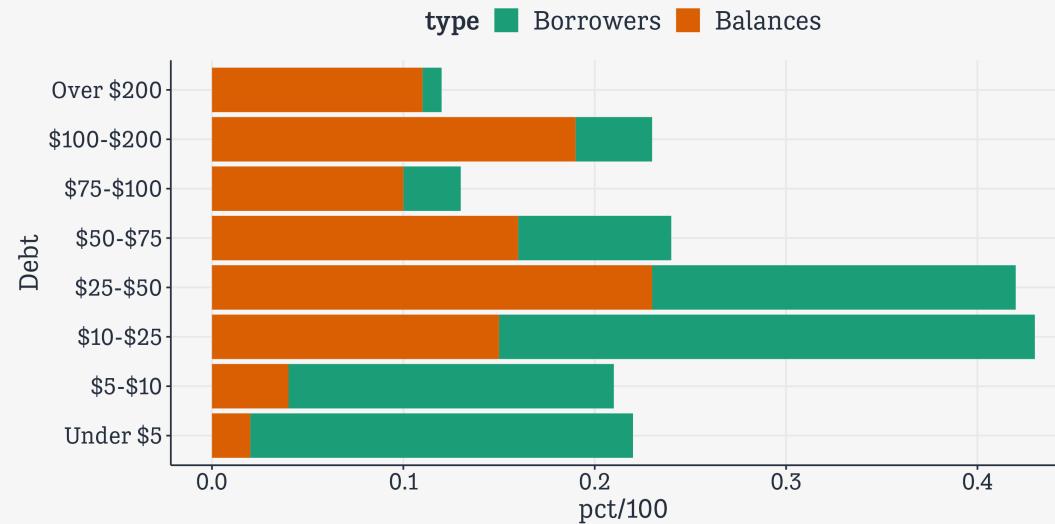
Debt Plot 1

```
1 studebt %>
2   ggplot(mapping =
3     aes(x = pct/100,
4           y = Debt,
5           fill = type)) +
6   geom_col()
```



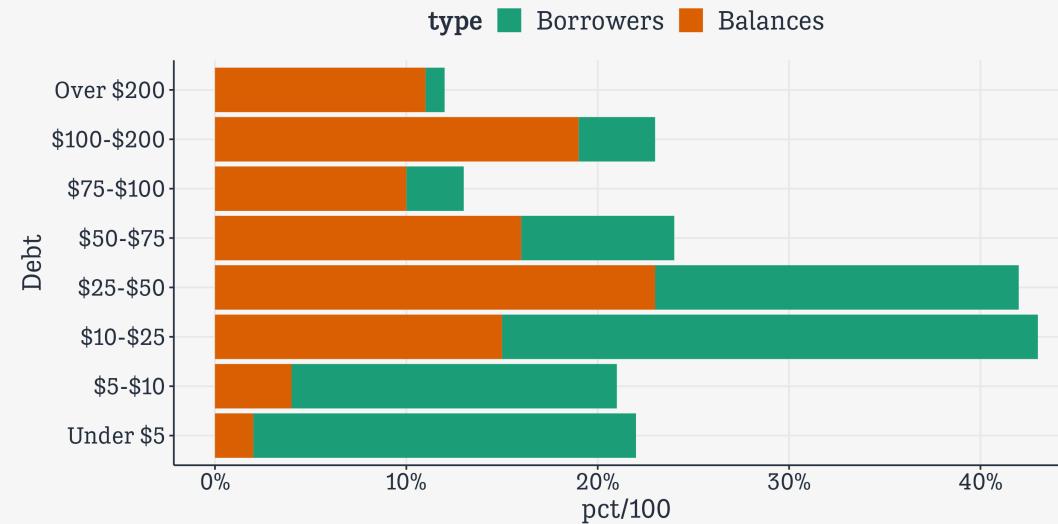
Debt Plot 1

```
1 studebt >
2   ggplot(mapping =
3     aes(x = pct/100,
4           y = Debt,
5           fill = type)) +
6   geom_col() +
7   scale_fill_brewer(type = "qual",
8                     palette = "Dark2")
```



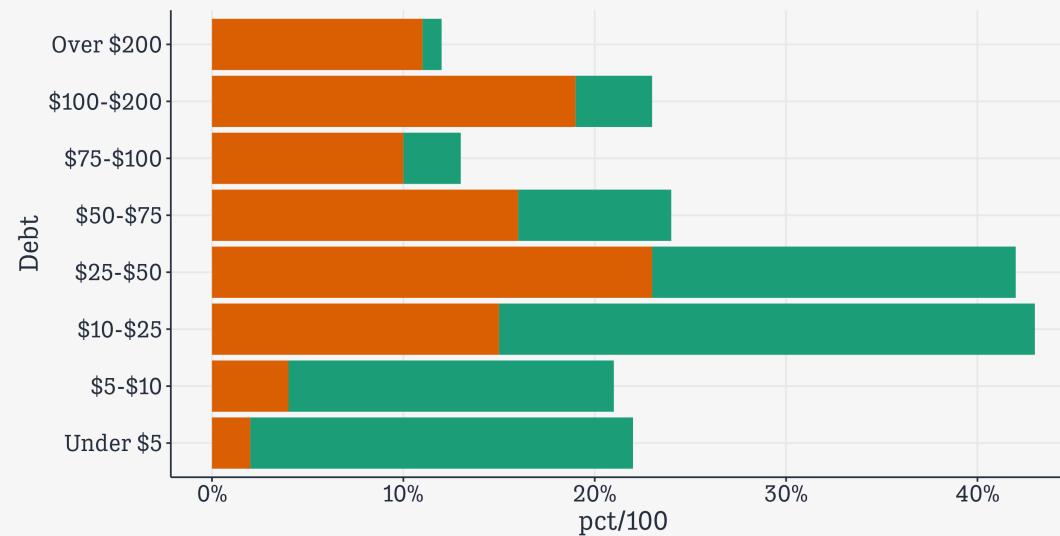
Debt Plot 1

```
1 studebt >
2   ggplot(mapping =
3     aes(x = pct/100,
4           y = Debt,
5           fill = type)) +
6   geom_col() +
7   scale_fill_brewer(type = "qual",
8                     palette = "Dark2") +
9   scale_x_continuous(labels = label_percent())
```



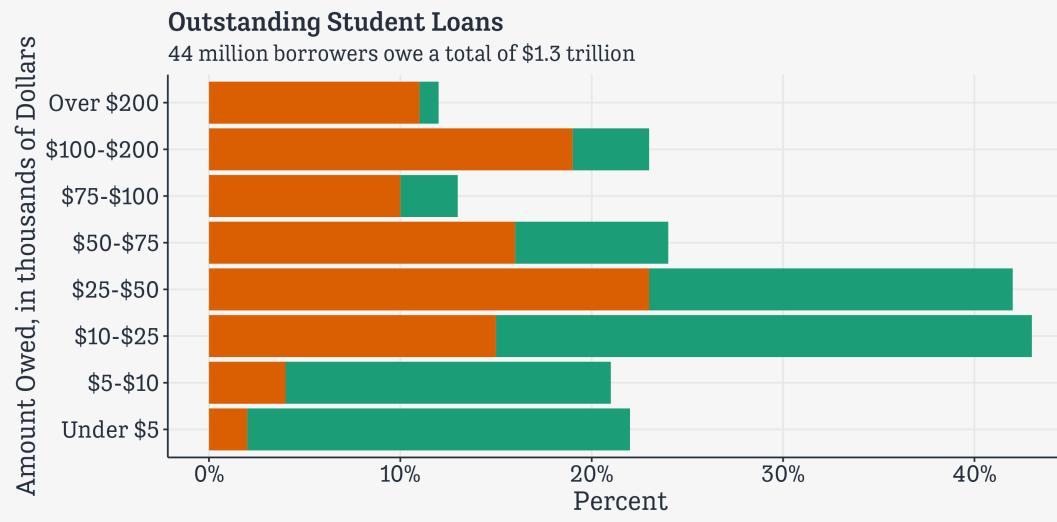
Debt Plot 1

```
1 studebt >
2   ggplot(mapping =
3     aes(x = pct/100,
4           y = Debt,
5           fill = type)) +
6   geom_col() +
7   scale_fill_brewer(type = "qual",
8                     palette = "Dark2") +
9   scale_x_continuous(labels = label_percent()) +
10  guides(fill = "none")
```



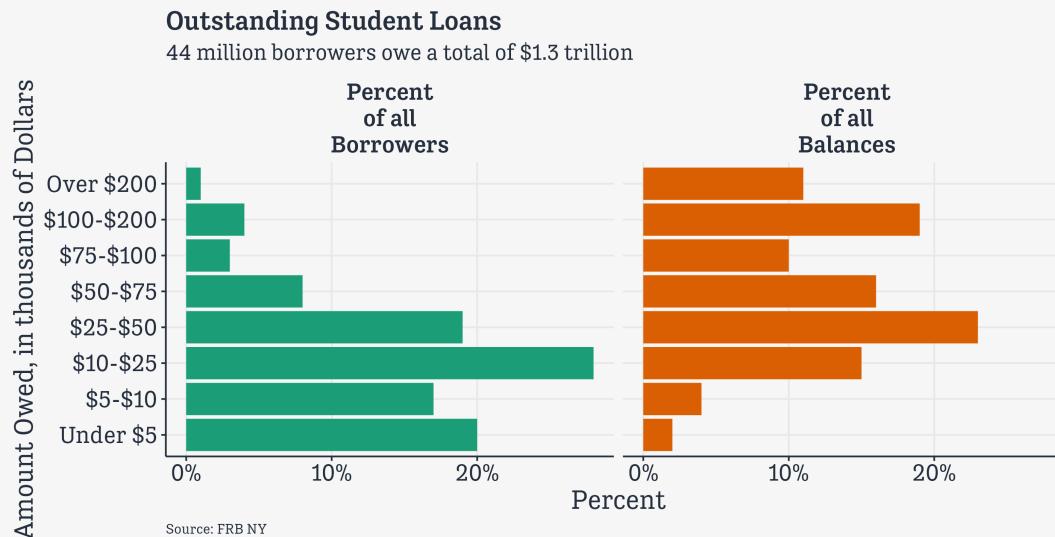
Debt Plot 1

```
1 studebt >
2   ggplot(mapping =
3     aes(x = pct/100,
4           y = Debt,
5           fill = type)) +
6   geom_col() +
7   scale_fill_brewer(type = "qual",
8                     palette = "Dark2") +
9   scale_x_continuous(labels = label_percent()) +
10  guides(fill = "none") +
11  labs(x = "Percent",
12        y = p_ylab,
13        caption = p_caption,
14        title = p_title,
15        subtitle = p_subtitle)
```



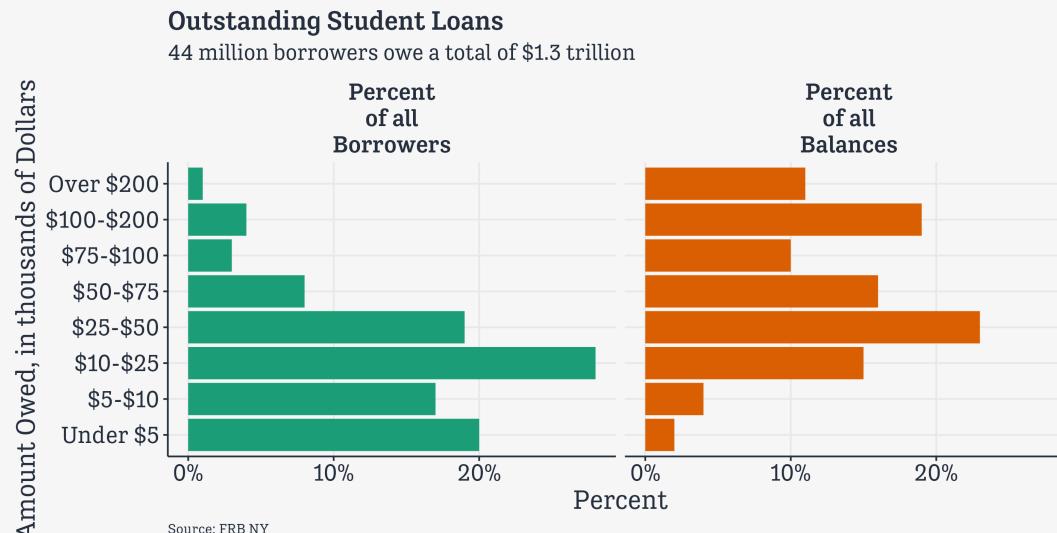
Debt Plot 1

```
1 studebt >
2   ggplot(mapping =
3     aes(x = pct/100,
4           y = Debt,
5           fill = type)) +
6   geom_col() +
7   scale_fill_brewer(type = "qual",
8                     palette = "Dark2") +
9   scale_x_continuous(labels = label_percent()) +
10  guides(fill = "none") +
11  labs(x = "Percent",
12        y = p_ylab,
13        caption = p_caption,
14        title = p_title,
15        subtitle = p_subtitle) +
16  facet_wrap(~ type_label,
17             labeller =
18               label_wrap_gen(width=10))
```



Debt Plot 1

```
1 studebt >
2   ggplot(mapping =
3     aes(x = pct/100,
4           y = Debt,
5           fill = type)) +
6   geom_col() +
7   scale_fill_brewer(type = "qual",
8                     palette = "Dark2") +
9   scale_x_continuous(labels = label_percent()) +
10  guides(fill = "none") +
11  labs(x = "Percent",
12        y = p_ylab,
13        caption = p_caption,
14        title = p_title,
15        subtitle = p_subtitle) +
16  facet_wrap(~ type_label,
17             labeller =
18               label_wrap_gen(width=10)) +
19  theme(strip.text.x =
20        element_text(face = "bold"))
```

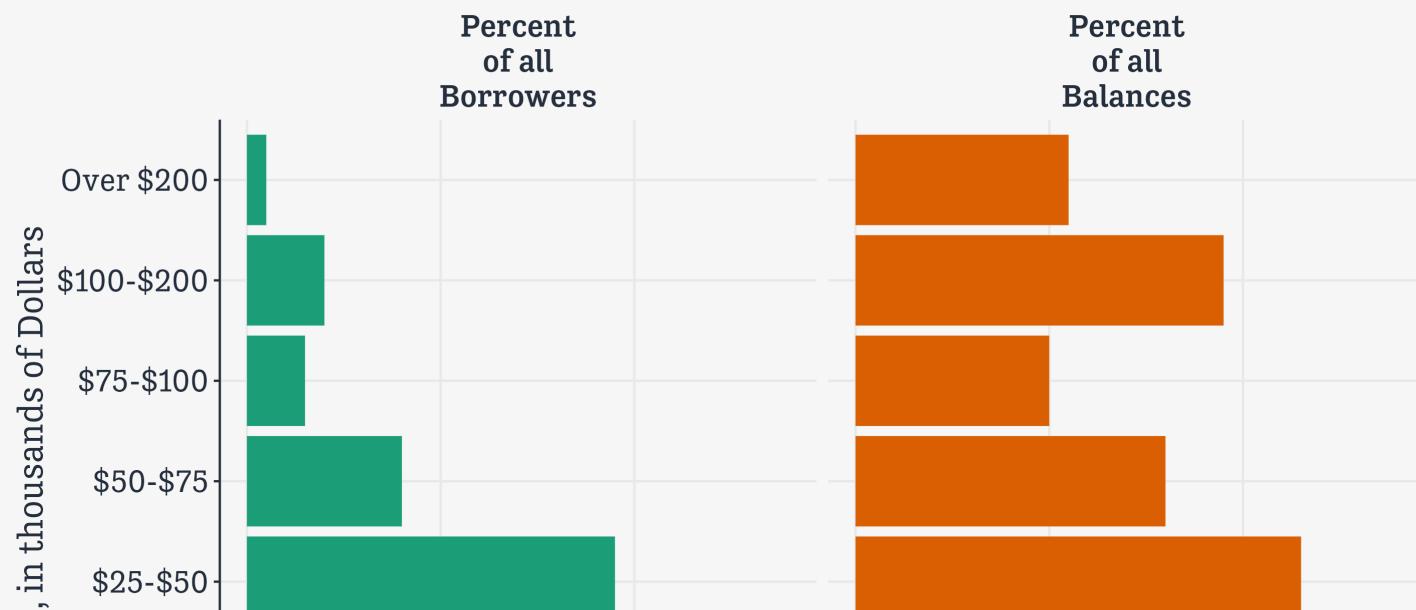


Debt Plot 1

```
1 studebt %>
2   ggplot(mapping =
3     aes(x = pct/100,
4          y = Debt,
5          fill = type)) +
6   geom_col() +
7   scale_fill_brewer(type = "qual",
8                      palette = "Dark2") +
9   scale_x_continuous(labels = label_percent()) +
10  guides(fill = "none") +
11  labs(x = "Percent",
12        y = p_ylab,
13        caption = p_caption,
14        title = p_title,
15        subtitle = p_subtitle) +
16  facet_wrap(~ type_label,
17             labeller =
18               label_wrap_gen(width=10)) +
19  theme(strip.text.x =
20        element_text(face = "bold")) →
21  p1_debt
```

Outstanding Student Loans

44 million borrowers owe a total of \$1.3 trillion



Alternatively, a kind of stacked bar chart

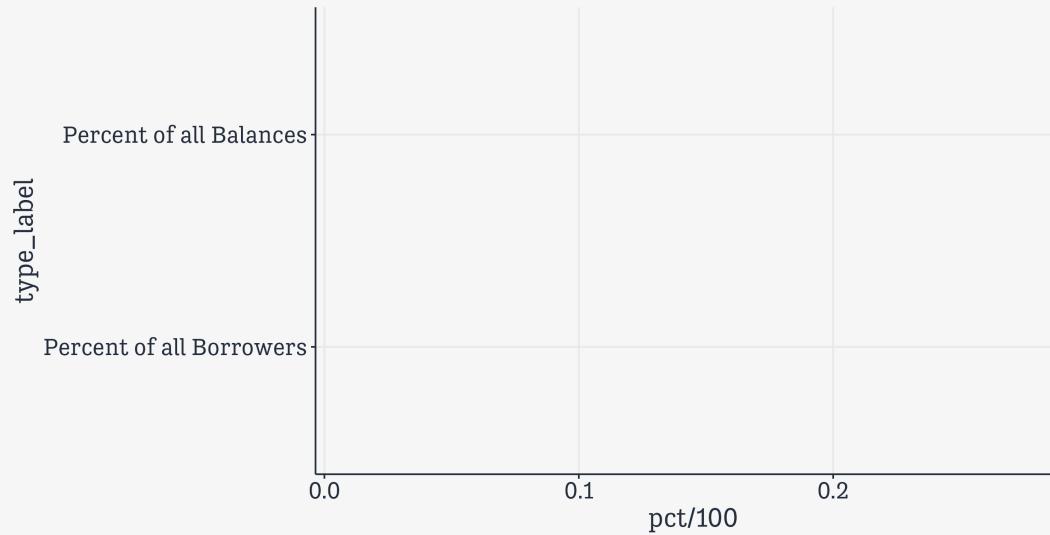
```
1 studebt
```

```
# A tibble: 16 × 5
  Debt      type     pct Debtrc type_label
  <ord>    <fct>   <int> <ord> <fct>
1 Under $5 Borrowers 20 Under $5 Percent of all Borrowers
2 $5-$10 Borrowers 17 $5-$10 Percent of all Borrowers
3 $10-$25 Borrowers 28 $10-$25 Percent of all Borrowers
4 $25-$50 Borrowers 19 $25-$50 Percent of all Borrowers
5 $50-$75 Borrowers  8 $50-$75 Percent of all Borrowers
6 $75-$100 Borrowers  3 $75-$100 Percent of all Borrowers
7 $100-$200 Borrowers  4 $100-$200 Percent of all Borrowers
8 Over $200 Borrowers  1 Over $200 Percent of all Borrowers
9 Under $5 Balances  2 Under $5 Percent of all Balances
10 $5-$10 Balances  4 $5-$10 Percent of all Balances
11 $10-$25 Balances 15 $10-$25 Percent of all Balances
12 $25-$50 Balances 23 $25-$50 Percent of all Balances
13 $50-$75 Balances 16 $50-$75 Percent of all Balances
14 $75-$100 Balances 10 $75-$100 Percent of all Balances
15 $100-$200 Balances 19 $100-$200 Percent of all Balances
16 Over $200 Balances 11 Over $200 Percent of all Balances
```



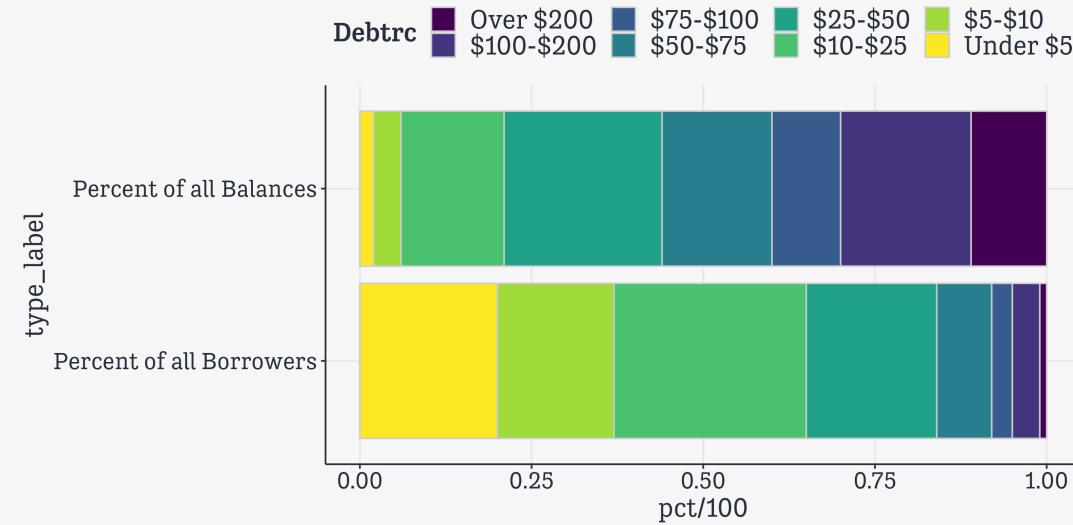
Alternatively, a kind of stacked bar chart

```
1 studebt %>%
2   ggplot(mapping = aes(x = pct/100,
3                         y = type_label,
4                         fill = Debtrc))
```



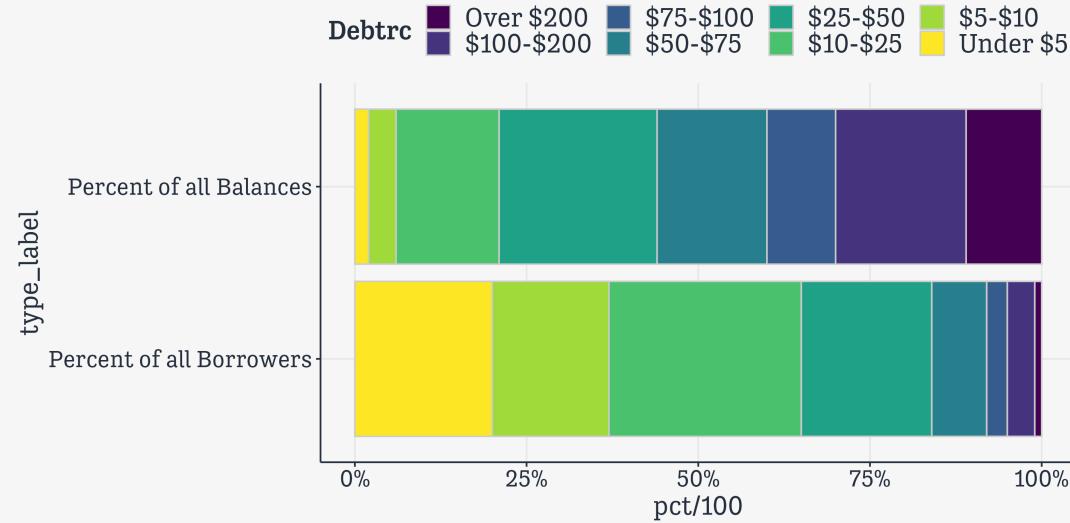
Alternatively, a kind of stacked bar chart

```
1 studebt %>
2   ggplot(mapping = aes(x = pct/100,
3                         y = type_label,
4                         fill = Debtrc)) +
5   geom_col(color = "gray80")
```



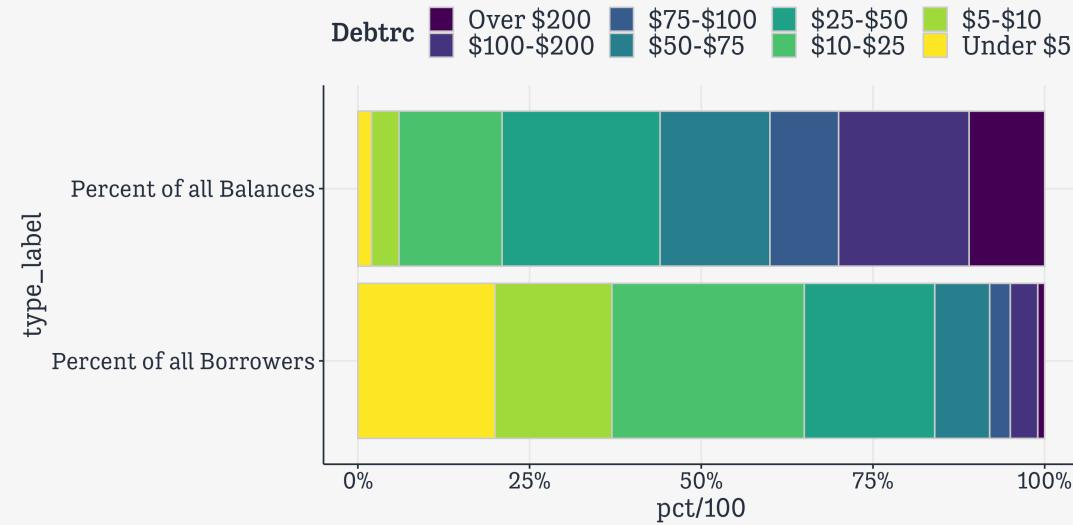
Alternatively, a kind of stacked bar chart

```
1 studebt %>
2   ggplot(mapping = aes(x = pct/100,
3                         y = type_label,
4                         fill = Debtrc)) +
5   geom_col(color = "gray80") +
6   scale_x_continuous(labels =
7     label_percent())
```



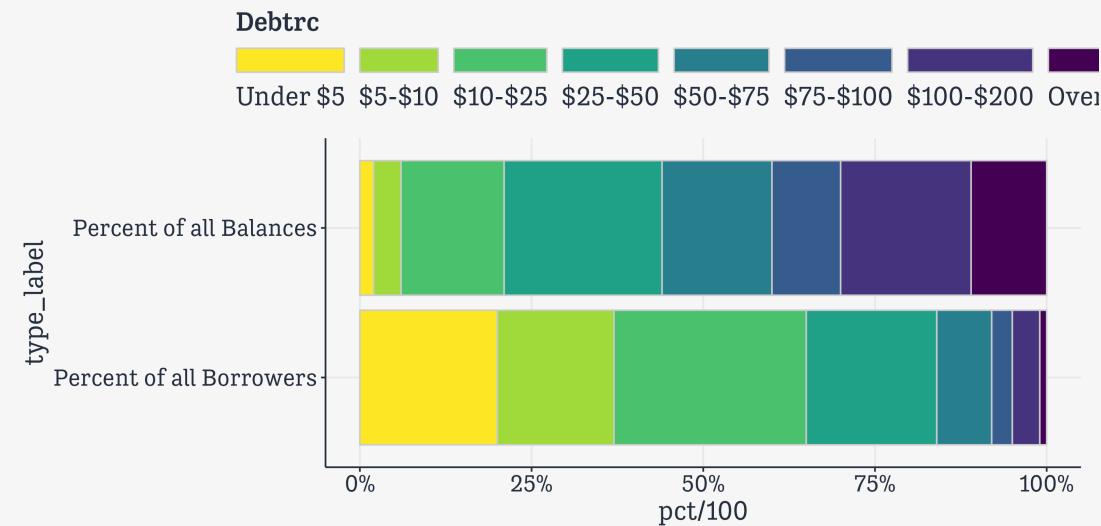
Alternatively, a kind of stacked bar chart

```
1 studebt %>
2   ggplot(mapping = aes(x = pct/100,
3                         y = type_label,
4                         fill = Debtrc)) +
5   geom_col(color = "gray80") +
6   scale_x_continuous(labels =
7     label_percent()) +
8   scale_fill_viridis_d()
```



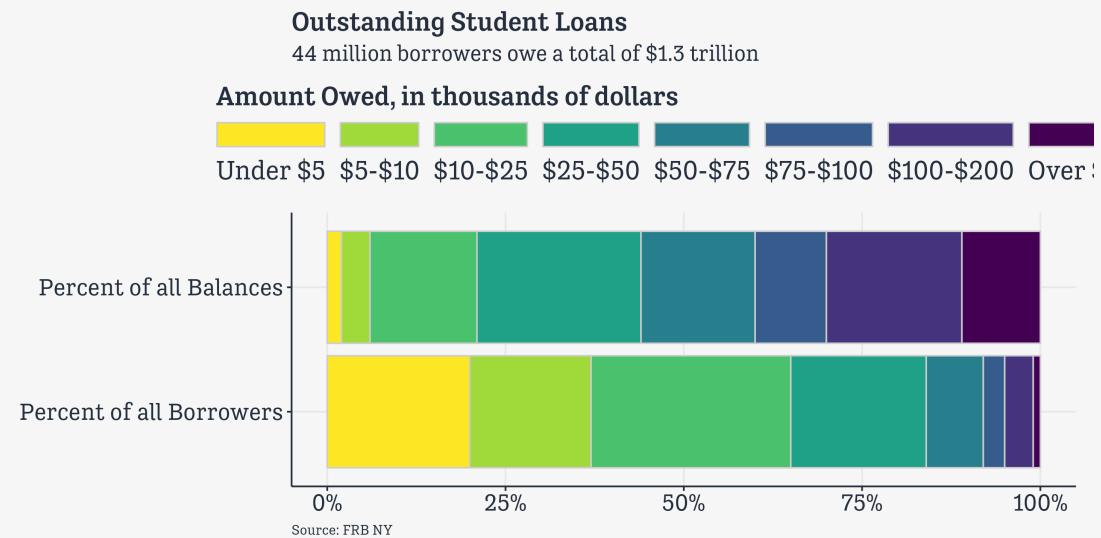
Alternatively, a kind of stacked bar chart

```
1 studebt >
2   ggplot(mapping = aes(x = pct/100,
3                         y = type_label,
4                         fill = Debtrc)) +
5   geom_col(color = "gray80") +
6   scale_x_continuous(labels =
7     label_percent()) +
8   scale_fill_viridis_d() +
9   guides(fill =
10      guide_legend(reverse = TRUE,
11                    title.position = "top",
12                    label.position = "bottom",
13                    keywidth = 3,
14                    nrow = 1))
```



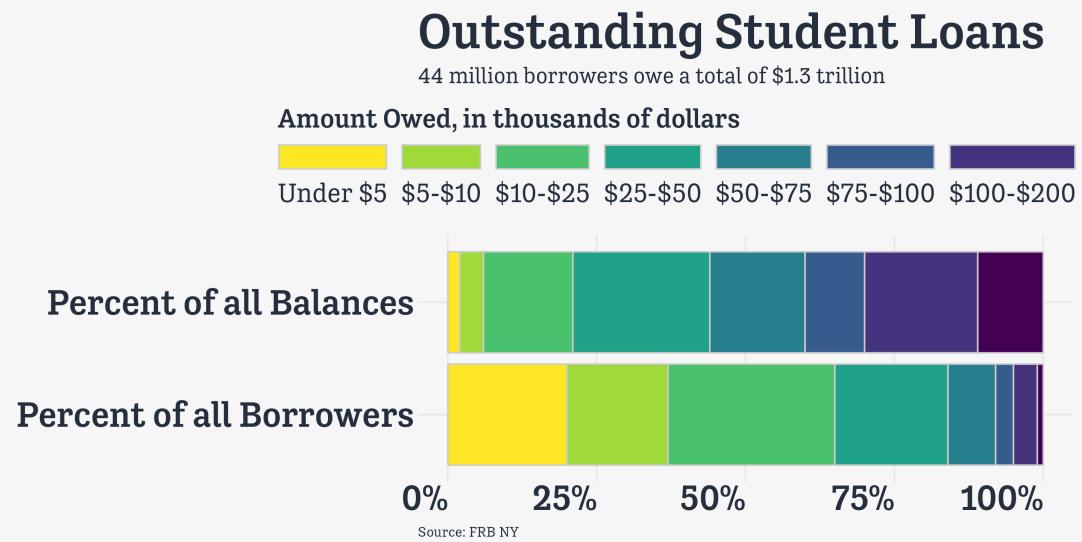
Alternatively, a kind of stacked bar chart

```
1 studebt >
2   ggplot(mapping = aes(x = pct/100,
3                         y = type_label,
4                         fill = Debtrc)) +
5   geom_col(color = "gray80") +
6   scale_x_continuous(labels =
7     label_percent()) +
8   scale_fill_viridis_d() +
9   guides(fill =
10      guide_legend(reverse = TRUE,
11                    title.position = "top",
12                    label.position = "bottom",
13                    keywidth = 3,
14                    nrow = 1)) +
15   labs(x = NULL, y = NULL,
16         fill = "Amount Owed, in thousands of dollars",
17         caption = p_caption, title = p_title,
18         subtitle = p_subtitle)
```



Alternatively, a kind of stacked bar chart

```
1 studebt >
2   ggplot(mapping = aes(x = pct/100,
3                         y = type_label,
4                         fill = Debtrc)) +
5   geom_col(color = "gray80") +
6   scale_x_continuous(labels =
7     label_percent()) +
8   scale_fill_viridis_d() +
9   guides(fill =
10     guide_legend(reverse = TRUE,
11                   title.position = "top",
12                   label.position = "bottom",
13                   keywidth = 3,
14                   nrow = 1)) +
15   labs(x = NULL, y = NULL,
16         fill = "Amount Owed, in thousands of dollars",
17         caption = p_caption, title = p_title,
18         subtitle = p_subtitle) +
19   theme(legend.position = "top",
20         plot.title = element_text(size = rel(2.8),
21                                   axis.text = element_text(face = "bold",
22                                               hjust = 1,
23                                               size = rel(2)),
```



Alternatively, a kind of stacked bar chart

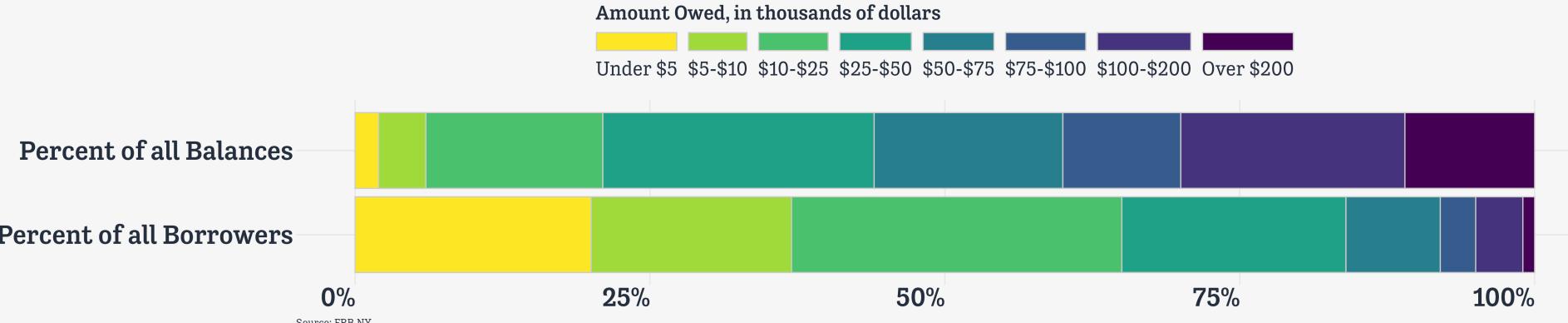
```
1 studebt >
2   ggplot(mapping = aes(x = pct/100,
3                         y = type_label,
4                         fill = Debtrc)) +
5   geom_col(color = "gray80") +
6   scale_x_continuous(labels =
7     label_percent()) +
8   scale_fill_viridis_d() +
9   guides(fill =
10      guide_legend(reverse = TRUE,
11                    title.position = "top",
12                    label.position = "bottom",
13                    keywidth = 3,
14                    nrow = 1)) +
15   labs(x = NULL, y = NULL,
16        fill = "Amount Owed, in thousands of dollars",
17        caption = p_caption, title = p_title,
18        subtitle = p_subtitle) +
19   theme(legend.position = "top",
20         plot.title = element_text(size = rel(2.8),
21                                     face = "bold",
22                                     hjust = 1,
23                                     size = rel(2)),
```

Alternatively, a kind of stacked bar chart

```
1 studebt >
2   ggplot(mapping = aes(x = pct/100,
3                         y = type_label,
4                         fill = Debtrc)) +
5   geom_col(color = "gray80") +
6   scale_x_continuous(labels =
7     label_percent()) +
8   scale_fill_viridis_d() +
9   guides(fill =
10      guide_legend(reverse = TRUE,
11                    title.position = "top",
12                    label.position = "bottom",
13                    keywidth = 3,
14                    nrow = 1)) +
15   labs(x = NULL, y = NULL,
16        fill = "Amount Owed, in thousands of dollars",
17        caption = p_caption, title = p_title,
18        subtitle = p_subtitle) +
19   theme(legend.position = "top",
20         plot.title = element_text(size = rel(2.8),
21                                     face = "bold",
22                                     hjust = 1,
23                                     size = rel(2)),
```

Outstanding Student Loans

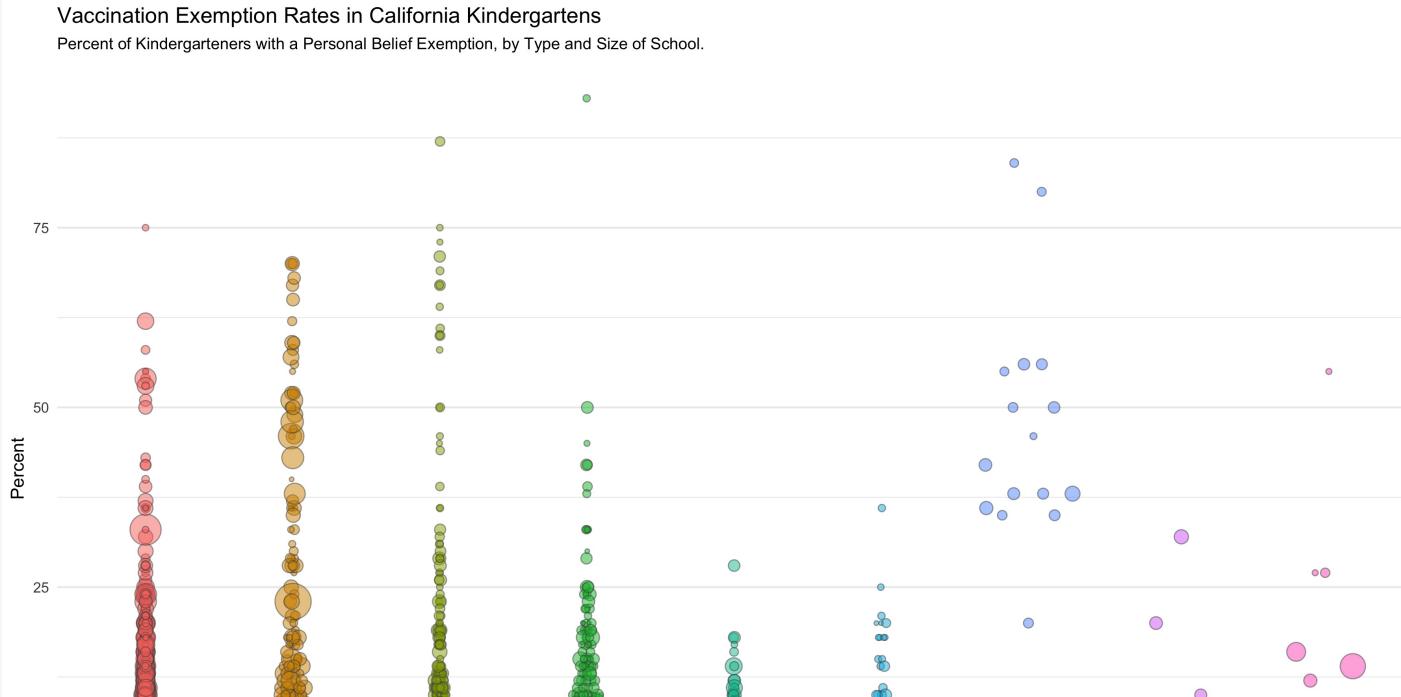
44 million borrowers owe a total of \$1.3 trillion



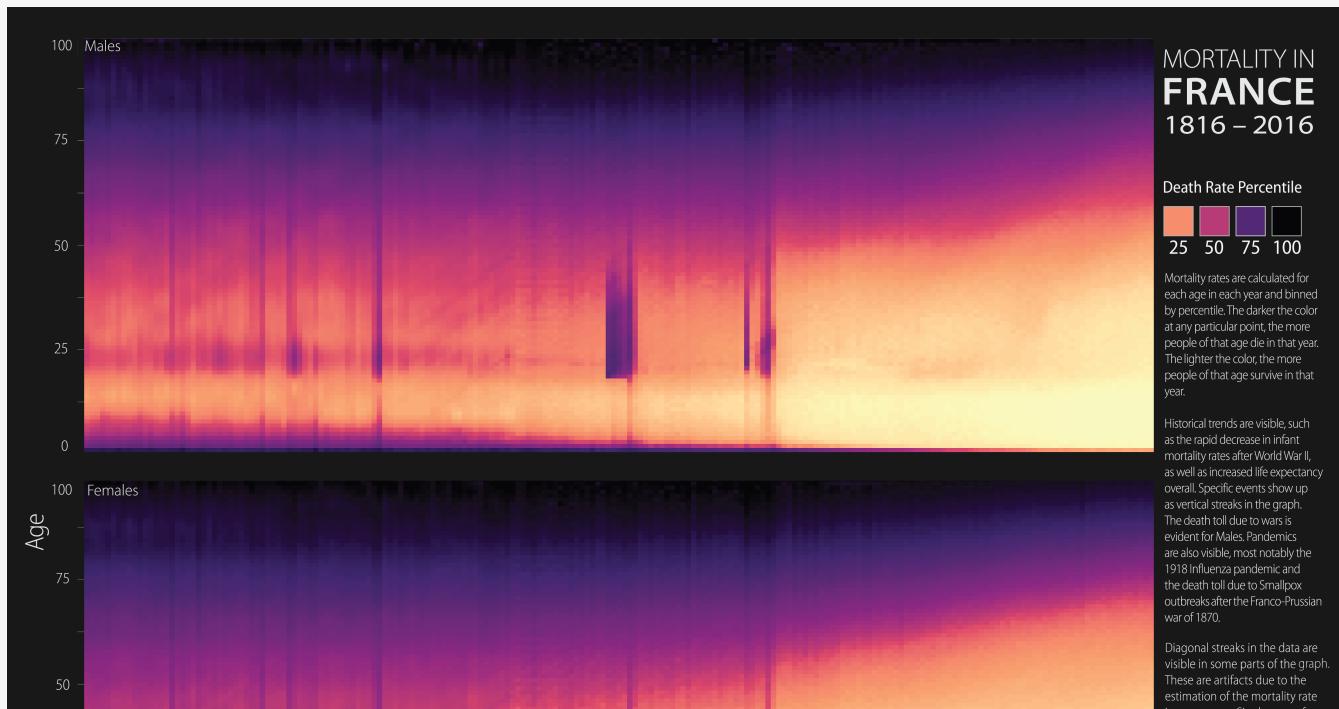
Pies redrawn as sideways-stacked columns

But I want
a pony

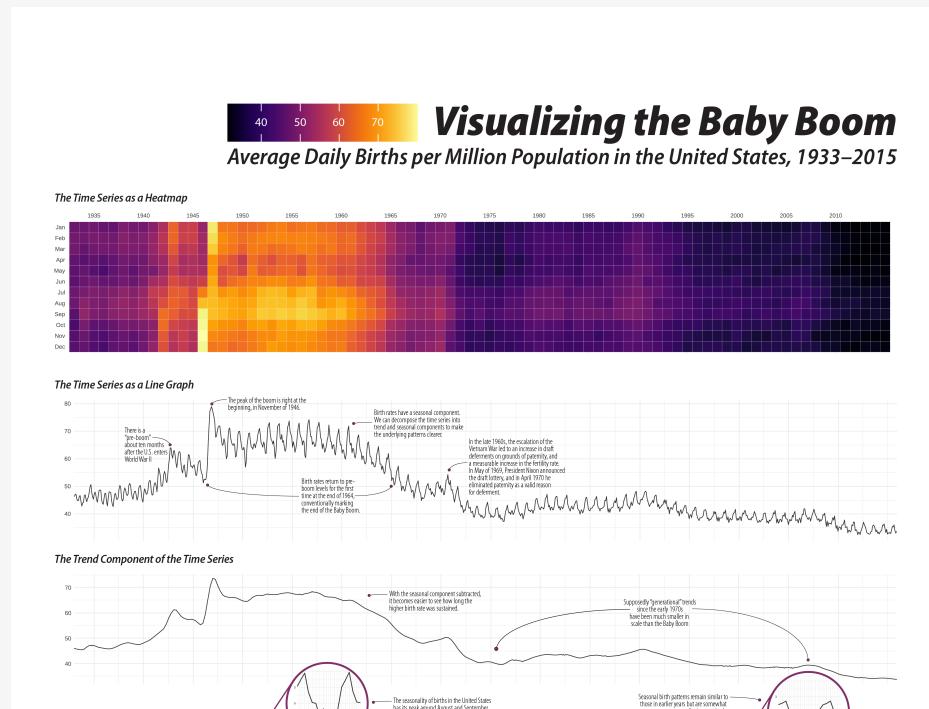
Show ponies



Show ponies



Show ponies



OK boomer

The **demog** package

```
# remotes::install_github("kjhealy/demog")
# library(demog)

okboomer

# A tibble: 1,644 × 12
  year month n_days births total_pop births_pct births_pct_day date
  <dbl> <dbl> <dbl>   <dbl>    <dbl>      <dbl>          <dbl> <date>
1 1938     1     31  51820  41215000  0.00126        40.6 1938-01-01
2 1938     2     28  47421  41215000  0.00115        41.1 1938-02-01
3 1938     3     31  54887  41215000  0.00133        43.0 1938-03-01
4 1938     4     30  54623  41215000  0.00133        44.2 1938-04-01
5 1938     5     31  56853  41215000  0.00138        44.5 1938-05-01
6 1938     6     30  53145  41215000  0.00129        43.0 1938-06-01
7 1938     7     31  53214  41215000  0.00129        41.6 1938-07-01
8 1938     8     31  50444  41215000  0.00122        39.5 1938-08-01
9 1938     9     30  50545  41215000  0.00123        40.9 1938-09-01
10 1938    10     31  50079  41215000  0.00122       39.2 1938-10-01
# i 1,634 more rows
# i 4 more variables: seasonal <dbl>, trend <dbl>, remainder <dbl>,
#   country <chr>
```

Boomer Line Graph

```
1 okboomer
```

```
# A tibble: 1,644 × 12
  year month n_days births total_pop births_pct births_pct_day date
  <dbl> <dbl> <dbl>   <dbl>    <dbl>      <dbl>          <dbl> <date>
1 1938     1     31  51820  41215000  0.00126        40.6 1938-01-01
2 1938     2     28  47421  41215000  0.00115        41.1 1938-02-01
3 1938     3     31  54887  41215000  0.00133        43.0 1938-03-01
4 1938     4     30  54623  41215000  0.00133        44.2 1938-04-01
5 1938     5     31  56853  41215000  0.00138        44.5 1938-05-01
6 1938     6     30  53145  41215000  0.00129        43.0 1938-06-01
7 1938     7     31  53214  41215000  0.00129        41.6 1938-07-01
8 1938     8     31  50444  41215000  0.00122        39.5 1938-08-01
9 1938     9     30  50545  41215000  0.00123        40.9 1938-09-01
10 1938    10     31  50079  41215000  0.00122       39.2 1938-10-01
# i 1,634 more rows
# i 4 more variables: seasonal <dbl>, trend <dbl>, remainder <dbl>,
#   country <chr>
```

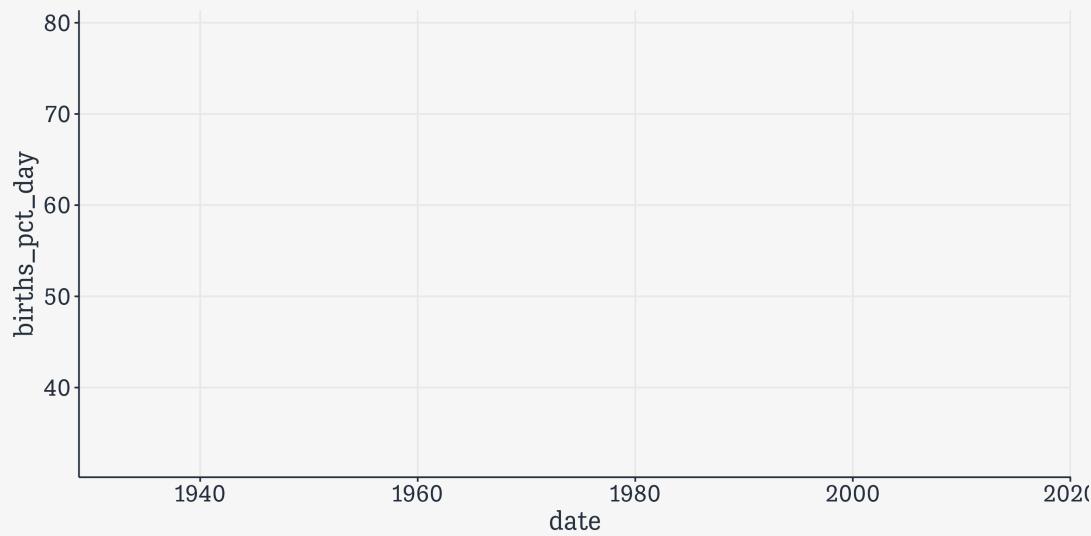
Boomer Line Graph

```
1 okboomer >
2     filter(country == "United States")
```

```
# A tibble: 996 x 12
   year month n_days births total_pop births_pct births_pct_day date
   <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <date>
 1 1933     1     31 180545 125579000 0.00144    46.4 1933-01-01
 2 1933     2     28 165986 125579000 0.00132    47.2 1933-02-01
 3 1933     3     31 183762 125579000 0.00146    47.2 1933-03-01
 4 1933     4     30 171354 125579000 0.00136    45.5 1933-04-01
 5 1933     5     31 174811 125579000 0.00139    44.9 1933-05-01
 6 1933     6     30 169255 125579000 0.00135    44.9 1933-06-01
 7 1933     7     31 180880 125579000 0.00144    46.5 1933-07-01
 8 1933     8     31 181856 125579000 0.00145    46.7 1933-08-01
 9 1933     9     30 167637 125579000 0.00133    44.5 1933-09-01
10 1933    10     31 167055 125579000 0.00133   42.9 1933-10-01
# i 986 more rows
# i 4 more variables: seasonal <dbl>, trend <dbl>, remainder <dbl>,
#   country <chr>
```

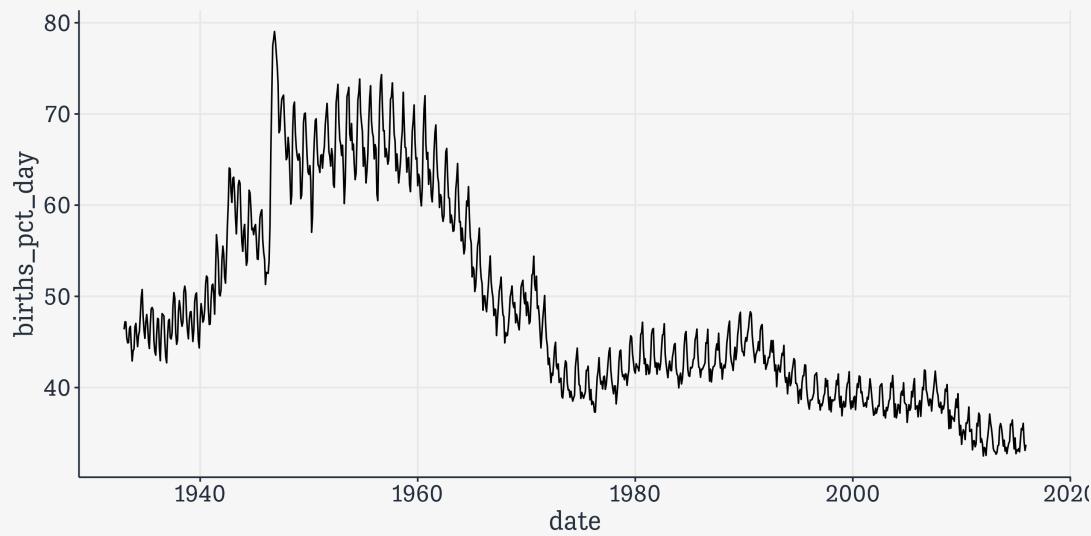
Boomer Line Graph

```
1 okboomer >
2   filter(country = "United States") >
3   ggplot(aes(x = date, y = births_pct_day))
```



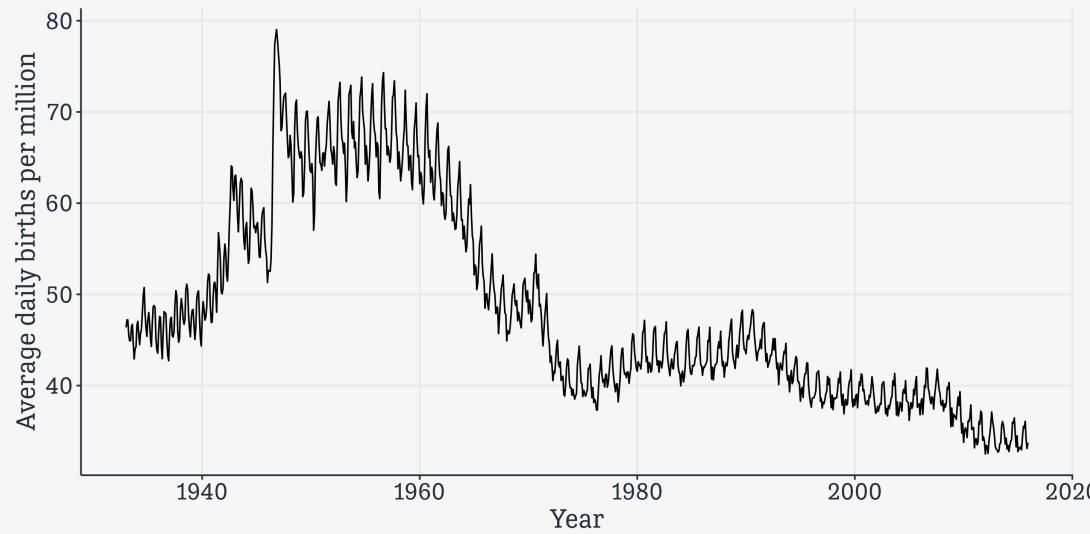
Boomer Line Graph

```
1 okboomer >
2   filter(country = "United States") >
3   ggplot(aes(x = date, y = births_pct_day)) +
4   geom_line(linewidth = 0.5)
```



Boomer Line Graph

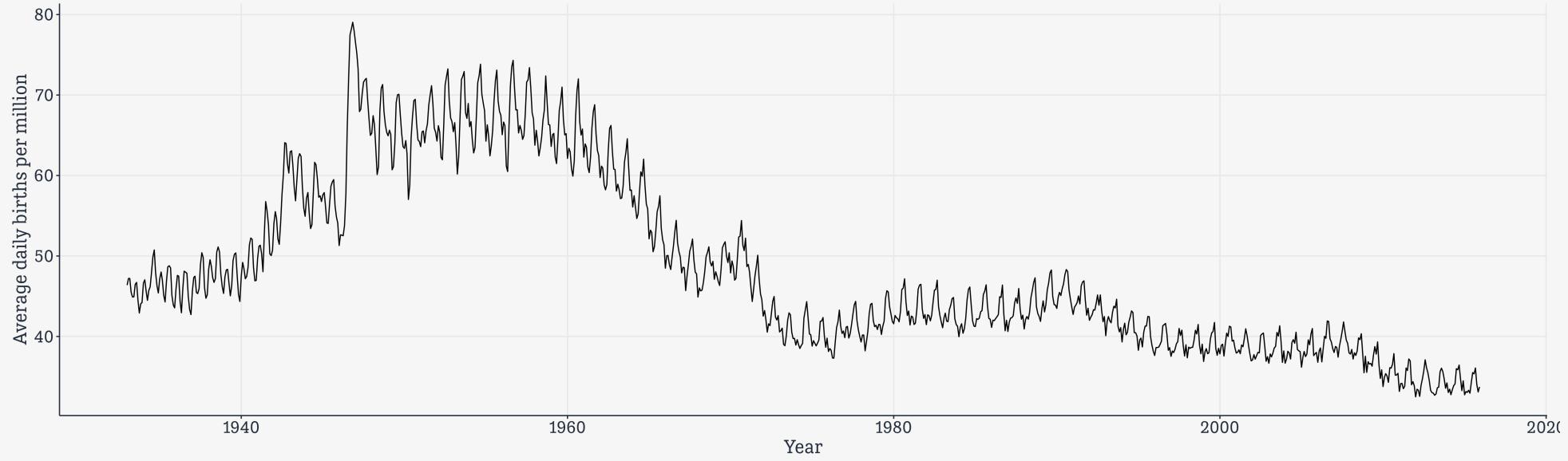
```
1 okboomer %>
2   filter(country = "United States") %>
3   ggplot(aes(x = date, y = births_pct_day)) +
4   geom_line(linewidth = 0.5) +
5   labs(x = "Year",
6        y = "Average daily births per million")
```



Boomer Line Graph

```
1 okboomer >
2   filter(country == "United States") >
3   ggplot(aes(x = date, y = births_pct_day)) +
4   geom_line(linewidth = 0.5) +
5   labs(x = "Year",
6        y = "Average daily births per million")
7 p_lineboom
```





The Baby Boom.

Tiled Heatmap

```
1 okboomer
```

```
# A tibble: 1,644 × 12
  year month n_days births total_pop births_pct
  <dbl> <dbl> <dbl>   <dbl>    <dbl>      <dbl>
  1 1938     1     31  51820  41215000  0.00126
  2 1938     2     28  47421  41215000  0.00115
  3 1938     3     31  54887  41215000  0.00133
  4 1938     4     30  54623  41215000  0.00133
  5 1938     5     31  56853  41215000  0.00138
  6 1938     6     30  53145  41215000  0.00129
  7 1938     7     31  53214  41215000  0.00129
  8 1938     8     31  50444  41215000  0.00122
  9 1938     9     30  50545  41215000  0.00123
 10 1938    10     31  50079  41215000  0.00122
 11 1938    11     30  49900  41215000  0.00122
 12 1938    12     31  49900  41215000  0.00122
 13 1939     1     31  51820  41215000  0.00126
 14 1939     2     28  47421  41215000  0.00115
 15 1939     3     31  54887  41215000  0.00133
 16 1939     4     30  54623  41215000  0.00133
 17 1939     5     31  56853  41215000  0.00138
 18 1939     6     30  53145  41215000  0.00129
 19 1939     7     31  53214  41215000  0.00129
 20 1939     8     31  50444  41215000  0.00122
 21 1939     9     30  50545  41215000  0.00123
 22 1939    10     31  50079  41215000  0.00122
 23 1939    11     30  49900  41215000  0.00122
 24 1939    12     31  49900  41215000  0.00122
 25 1940     1     31  51820  41215000  0.00126
 26 1940     2     28  47421  41215000  0.00115
 27 1940     3     31  54887  41215000  0.00133
 28 1940     4     30  54623  41215000  0.00133
 29 1940     5     31  56853  41215000  0.00138
 30 1940     6     30  53145  41215000  0.00129
 31 1940     7     31  53214  41215000  0.00129
 32 1940     8     31  50444  41215000  0.00122
 33 1940     9     30  50545  41215000  0.00123
 34 1940    10     31  50079  41215000  0.00122
 35 1940    11     30  49900  41215000  0.00122
 36 1940    12     31  49900  41215000  0.00122
 37 1941     1     31  51820  41215000  0.00126
 38 1941     2     28  47421  41215000  0.00115
 39 1941     3     31  54887  41215000  0.00133
 40 1941     4     30  54623  41215000  0.00133
 41 1941     5     31  56853  41215000  0.00138
 42 1941     6     30  53145  41215000  0.00129
 43 1941     7     31  53214  41215000  0.00129
 44 1941     8     31  50444  41215000  0.00122
 45 1941     9     30  50545  41215000  0.00123
 46 1941    10     31  50079  41215000  0.00122
 47 1941    11     30  49900  41215000  0.00122
 48 1941    12     31  49900  41215000  0.00122
 49 1942     1     31  51820  41215000  0.00126
 50 1942     2     28  47421  41215000  0.00115
 51 1942     3     31  54887  41215000  0.00133
 52 1942     4     30  54623  41215000  0.00133
 53 1942     5     31  56853  41215000  0.00138
 54 1942     6     30  53145  41215000  0.00129
 55 1942     7     31  53214  41215000  0.00129
 56 1942     8     31  50444  41215000  0.00122
 57 1942     9     30  50545  41215000  0.00123
 58 1942    10     31  50079  41215000  0.00122
 59 1942    11     30  49900  41215000  0.00122
 60 1942    12     31  49900  41215000  0.00122
 61 1943     1     31  51820  41215000  0.00126
 62 1943     2     28  47421  41215000  0.00115
 63 1943     3     31  54887  41215000  0.00133
 64 1943     4     30  54623  41215000  0.00133
 65 1943     5     31  56853  41215000  0.00138
 66 1943     6     30  53145  41215000  0.00129
 67 1943     7     31  53214  41215000  0.00129
 68 1943     8     31  50444  41215000  0.00122
 69 1943     9     30  50545  41215000  0.00123
 70 1943    10     31  50079  41215000  0.00122
 71 1943    11     30  49900  41215000  0.00122
 72 1943    12     31  49900  41215000  0.00122
 73 1944     1     31  51820  41215000  0.00126
 74 1944     2     28  47421  41215000  0.00115
 75 1944     3     31  54887  41215000  0.00133
 76 1944     4     30  54623  41215000  0.00133
 77 1944     5     31  56853  41215000  0.00138
 78 1944     6     30  53145  41215000  0.00129
 79 1944     7     31  53214  41215000  0.00129
 80 1944     8     31  50444  41215000  0.00122
 81 1944     9     30  50545  41215000  0.00123
 82 1944    10     31  50079  41215000  0.00122
 83 1944    11     30  49900  41215000  0.00122
 84 1944    12     31  49900  41215000  0.00122
 85 1945     1     31  51820  41215000  0.00126
 86 1945     2     28  47421  41215000  0.00115
 87 1945     3     31  54887  41215000  0.00133
 88 1945     4     30  54623  41215000  0.00133
 89 1945     5     31  56853  41215000  0.00138
 90 1945     6     30  53145  41215000  0.00129
 91 1945     7     31  53214  41215000  0.00129
 92 1945     8     31  50444  41215000  0.00122
 93 1945     9     30  50545  41215000  0.00123
 94 1945    10     31  50079  41215000  0.00122
 95 1945    11     30  49900  41215000  0.00122
 96 1945    12     31  49900  41215000  0.00122
 97 1946     1     31  51820  41215000  0.00126
 98 1946     2     28  47421  41215000  0.00115
 99 1946     3     31  54887  41215000  0.00133
100 1946     4     30  54623  41215000  0.00133
101 1946     5     31  56853  41215000  0.00138
102 1946     6     30  53145  41215000  0.00129
103 1946     7     31  53214  41215000  0.00129
104 1946     8     31  50444  41215000  0.00122
105 1946     9     30  50545  41215000  0.00123
106 1946    10     31  50079  41215000  0.00122
107 1946    11     30  49900  41215000  0.00122
108 1946    12     31  49900  41215000  0.00122
109 1947     1     31  51820  41215000  0.00126
110 1947     2     28  47421  41215000  0.00115
111 1947     3     31  54887  41215000  0.00133
112 1947     4     30  54623  41215000  0.00133
113 1947     5     31  56853  41215000  0.00138
114 1947     6     30  53145  41215000  0.00129
115 1947     7     31  53214  41215000  0.00129
116 1947     8     31  50444  41215000  0.00122
117 1947     9     30  50545  41215000  0.00123
118 1947    10     31  50079  41215000  0.00122
119 1947    11     30  49900  41215000  0.00122
120 1947    12     31  49900  41215000  0.00122
121 1948     1     31  51820  41215000  0.00126
122 1948     2     28  47421  41215000  0.00115
123 1948     3     31  54887  41215000  0.00133
124 1948     4     30  54623  41215000  0.00133
125 1948     5     31  56853  41215000  0.00138
126 1948     6     30  53145  41215000  0.00129
127 1948     7     31  53214  41215000  0.00129
128 1948     8     31  50444  41215000  0.00122
129 1948     9     30  50545  41215000  0.00123
130 1948    10     31  50079  41215000  0.00122
131 1948    11     30  49900  41215000  0.00122
132 1948    12     31  49900  41215000  0.00122
133 1949     1     31  51820  41215000  0.00126
134 1949     2     28  47421  41215000  0.00115
135 1949     3     31  54887  41215000  0.00133
136 1949     4     30  54623  41215000  0.00133
137 1949     5     31  56853  41215000  0.00138
138 1949     6     30  53145  41215000  0.00129
139 1949     7     31  53214  41215000  0.00129
140 1949     8     31  50444  41215000  0.00122
141 1949     9     30  50545  41215000  0.00123
142 1949    10     31  50079  41215000  0.00122
143 1949    11     30  49900  41215000  0.00122
144 1949    12     31  49900  41215000  0.00122
145 1950     1     31  51820  41215000  0.00126
146 1950     2     28  47421  41215000  0.00115
147 1950     3     31  54887  41215000  0.00133
148 1950     4     30  54623  41215000  0.00133
149 1950     5     31  56853  41215000  0.00138
150 1950     6     30  53145  41215000  0.00129
151 1950     7     31  53214  41215000  0.00129
152 1950     8     31  50444  41215000  0.00122
153 1950     9     30  50545  41215000  0.00123
154 1950    10     31  50079  41215000  0.00122
155 1950    11     30  49900  41215000  0.00122
156 1950    12     31  49900  41215000  0.00122
157 1951     1     31  51820  41215000  0.00126
158 1951     2     28  47421  41215000  0.00115
159 1951     3     31  54887  41215000  0.00133
160 1951     4     30  54623  41215000  0.00133
161 1951     5     31  56853  41215000  0.00138
162 1951     6     30  53145  41215000  0.00129
163 1951     7     31  53214  41215000  0.00129
164 1951     8     31  50444  41215000  0.00122
165 1951     9     30  50545  41215000  0.00123
166 1951    10     31  50079  41215000  0.00122
167 1951    11     30  49900  41215000  0.00122
168 1951    12     31  49900  41215000  0.00122
169 1952     1     31  51820  41215000  0.00126
170 1952     2     28  47421  41215000  0.00115
171 1952     3     31  54887  41215000  0.00133
172 1952     4     30  54623  41215000  0.00133
173 1952     5     31  56853  41215000  0.00138
174 1952     6     30  53145  41215000  0.00129
175 1952     7     31  53214  41215000  0.00129
176 1952     8     31  50444  41215000  0.00122
177 1952     9     30  50545  41215000  0.00123
178 1952    10     31  50079  41215000  0.00122
179 1952    11     30  49900  41215000  0.00122
180 1952    12     31  49900  41215000  0.00122
181 1953     1     31  51820  41215000  0.00126
182 1953     2     28  47421  41215000  0.00115
183 1953     3     31  54887  41215000  0.00133
184 1953     4     30  54623  41215000  0.00133
185 1953     5     31  56853  41215000  0.00138
186 1953     6     30  53145  41215000  0.00129
187 1953     7     31  53214  41215000  0.00129
188 1953     8     31  50444  41215000  0.00122
189 1953     9     30  50545  41215000  0.00123
190 1953    10     31  50079  41215000  0.00122
191 1953    11     30  49900  41215000  0.00122
192 1953    12     31  49900  41215000  0.00122
193 1954     1     31  51820  41215000  0.00126
194 1954     2     28  47421  41215000  0.00115
195 1954     3     31  54887  41215000  0.00133
196 1954     4     30  54623  41215000  0.00133
197 1954     5     31  56853  41215000  0.00138
198 1954     6     30  53145  41215000  0.00129
199 1954     7     31  53214  41215000  0.00129
200 1954     8     31  50444  41215000  0.00122
201 1954     9     30  50545  41215000  0.00123
202 1954    10     31  50079  41215000  0.00122
203 1954    11     30  49900  41215000  0.00122
204 1954    12     31  49900  41215000  0.00122
205 1955     1     31  51820  41215000  0.00126
206 1955     2     28  47421  41215000  0.00115
207 1955     3     31  54887  41215000  0.00133
208 1955     4     30  54623  41215000  0.00133
209 1955     5     31  56853  41215000  0.00138
210 1955     6     30  53145  41215000  0.00129
211 1955     7     31  53214  41215000  0.00129
212 1955     8     31  50444  41215000  0.00122
213 1955     9     30  50545  41215000  0.00123
214 1955    10     31  50079  41215000  0.00122
215 1955    11     30  49900  41215000  0.00122
216 1955    12     31  49900  41215000  0.00122
217 1956     1     31  51820  41215000  0.00126
218 1956     2     28  47421  41215000  0.00115
219 1956     3     31  54887  41215000  0.00133
220 1956     4     30  54623  41215000  0.00133
221 1956     5     31  56853  41215000  0.00138
222 1956     6     30  53145  41215000  0.00129
223 1956     7     31  53214  41215000  0.00129
224 1956     8     31  50444  41215000  0.00122
225 1956     9     30  50545  41215000  0.00123
226 1956    10     31  50079  41215000  0.00122
227 1956    11     30  49900  41215000  0.00122
228 1956    12     31  49900  41215000  0.00122
229 1957     1     31  51820  41215000  0.00126
230 1957     2     28  47421  41215000  0.00115
231 1957     3     31  54887  41215000  0.00133
232 1957     4     30  54623  41215000  0.00133
233 1957     5     31  56853  41215000  0.00138
234 1957     6     30  53145  41215000  0.00129
235 1957     7     31  53214  41215000  0.00129
236 1957     8     31  50444  41215000  0.00122
237 1957     9     30  50545  41215000  0.00123
238 1957    10     31  50079  41215000  0.00122
239 1957    11     30  49900  41215000  0.00122
240 1957    12     31  49900  41215000  0.00122
241 1958     1     31  51820  41215000  0.00126
242 1958     2     28  47421  41215000  0.00115
243 1958     3     31  54887  41215000  0.00133
244 1958     4     30  54623  41215000  0.00133
245 1958     5     31  56853  41215000  0.00138
246 1958     6     30  53145  41215000  0.00129
247 1958     7     31  53214  41215000  0.00129
248 1958     8     31  50444  41215000  0.00122
249 1958     9     30  50545  41215000  0.00123
250 1958    10     31  50079  41215000  0.00122
251 1958    11     30  49900  41215000  0.00122
252 1958    12     31  49900  41215000  0.00122
253 1959     1     31  51820  41215000  0.00126
254 1959     2     28  47421  41215000  0.00115
255 1959     3     31  54887  41215000  0.00133
256 1959     4     30  54623  41215000  0.00133
257 1959     5     31  56853  41215000  0.00138
258 1959     6     30  53145  41215000  0.00129
259 1959     7     31  53214  4121500
```

Tiled Heatmap

```
1 okboomer >
2     mutate(year_fct =
3         factor(year,
4                 levels = unique(year),
5                 ordered = TRUE),
6         month_fct = factor(month,
7                 levels = rev(c(1:12)),
8                 labels = rev(c("Jan", "Feb", "Mar", "Apr",
9                               "May", "Jun", "Jul", "Aug",
10                              "Sep", "Oct", "Nov", "Dec")),
11                 ordered = TRUE))
```

```
# A tibble: 1,644 × 14
  year month n_days births total_pop births_pct
  <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>
  1 1938 1 31 51820 41215000 0.00126
  2 1938 2 28 47421 41215000 0.00115
  3 1938 3 31 54887 41215000 0.00133
  4 1938 4 30 54623 41215000 0.00133
  5 1938 5 31 56853 41215000 0.00138
  6 1938 6 30 53145 41215000 0.00129
  7 1938 7 31 53214 41215000 0.00129
  8 1938 8 31 50444 41215000 0.00122
  9 1938 9 30 50545 41215000 0.00123
 10 1938 10 31 50079 41215000 0.00122
 11 1938 11 31 49850 41215000 0.00122
 12 1938 12 31 49750 41215000 0.00122
```

Tiled Heatmap

```
1 okboomer >
2     mutate(year_fct =
3         factor(year,
4                 levels = unique(year),
5                 ordered = TRUE),
6         month_fct = factor(month,
7                 levels = rev(c(1:12)),
8                 labels = rev(c("Jan", "Feb", "Mar", "Apr",
9                               "May", "Jun", "Jul", "Aug",
10                              "Sep", "Oct", "Nov", "Dec")),
11                 ordered = TRUE)) >
12     select(year, month, year_fct, month_fct, everything())
```

```
# A tibble: 1,644 × 14
  year month year_fct month_fct n_days births
  <dbl> <dbl> <ord>    <ord>      <dbl> <dbl>
  <dbl>   <dbl>
  1 1938   1 1938 Jan       31  51820
  41215000 0.00126
  2 1938   2 1938 Feb       28  47421
  41215000 0.00115
  3 1938   3 1938 Mar       31  54887
  41215000 0.00133
  4 1938   4 1938 Apr       30  54623
  41215000 0.00133
  5 1938   5 1938 May       31  56853
  41215000 0.00138
  6 1938   6 1938 Jun       30  53145
  41215000 0.00129
  7 1938   7 1938 Jul       31  53214
  41215000 0.00129
  8 1938   8 1938 Aug       31  50444
  41215000 0.00122
  9 1938   9 1938 Sep       30  50545
  41215000 0.00123
  10 1938  10 1938 Oct       31  50079
  41215000 0.00122
```

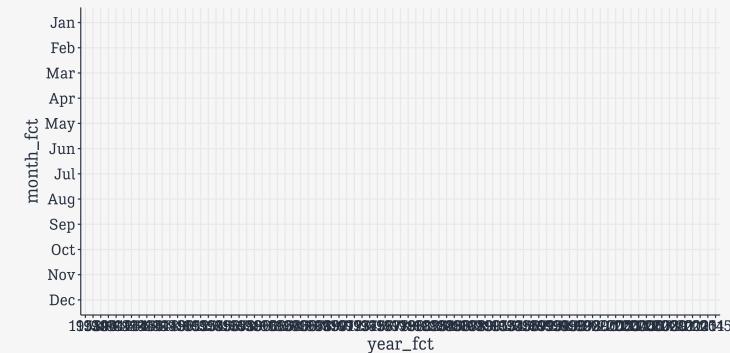
Tiled Heatmap

```
1 okboomer >
2     mutate(year_fct =
3         factor(year,
4                 levels = unique(year),
5                 ordered = TRUE),
6         month_fct = factor(month,
7                 levels = rev(c(1:12)),
8                 labels = rev(c("Jan", "Feb", "Mar", "Apr",
9                               "May", "Jun", "Jul", "Aug",
10                              "Sep", "Oct", "Nov", "Dec")),
11                 ordered = TRUE)) >
12     select(year, month, year_fct, month_fct, everything()) >
13     filter(country = "United States")
```

```
# A tibble: 996 x 14
   year month year_fct month_fct n_days births
   <dbl> <dbl> <ord>    <ord>      <dbl> <dbl>
1 1933  1 1933 Jan       31 180545
2 125579000 0.00144
3 1933  2 1933 Feb       28 165986
4 125579000 0.00132
5 1933  3 1933 Mar       31 183762
6 125579000 0.00146
7 1933  4 1933 Apr       30 171354
8 125579000 0.00136
9 1933  5 1933 May       31 174811
10 125579000 0.00139
11 1933  6 1933 Jun       30 169255
12 125579000 0.00135
13 1933  7 1933 Jul       31 180880
14 125579000 0.00144
15 1933  8 1933 Aug       31 181856
16 125579000 0.00145
17 1933  9 1933 Sep       30 167637
18 125579000 0.00133
19 1933 10 1933 Oct       31 167055
```

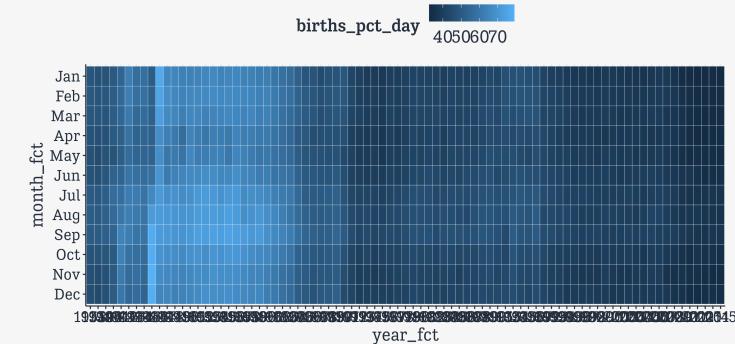
Tiled Heatmap

```
1 okboomer >
2   mutate(year_fct =
3     factor(year,
4       levels = unique(year),
5       ordered = TRUE),
6     month_fct = factor(month,
7       levels = rev(c(1:12)),
8       labels = rev(c("Jan", "Feb", "Mar", "Apr",
9                     "May", "Jun", "Jul", "Aug",
10                    "Sep", "Oct", "Nov", "Dec")),
11       ordered = TRUE)) >
12   select(year, month, year_fct, month_fct, everything()) >
13   filter(country = "United States") >
14   ggplot(aes(x = year_fct, y = month_fct))
```



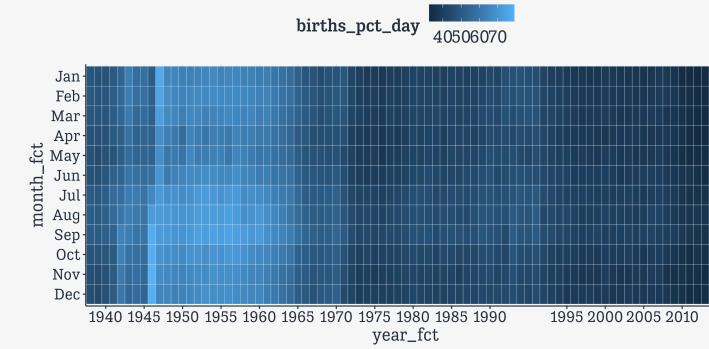
Tiled Heatmap

```
1 okboomer >
2   mutate(year_fct =
3     factor(year,
4       levels = unique(year),
5       ordered = TRUE),
6     month_fct = factor(month,
7       levels = rev(c(1:12)),
8       labels = rev(c("Jan", "Feb", "Mar", "Apr",
9                     "May", "Jun", "Jul", "Aug",
10                    "Sep", "Oct", "Nov", "Dec")),
11       ordered = TRUE)) >
12   select(year, month, year_fct, month_fct, everything()) >
13   filter(country = "United States") >
14   ggplot(aes(x = year_fct, y = month_fct)) +
15   geom_tile(mapping = aes(fill = births_pct_day),
16             color = "white")
```



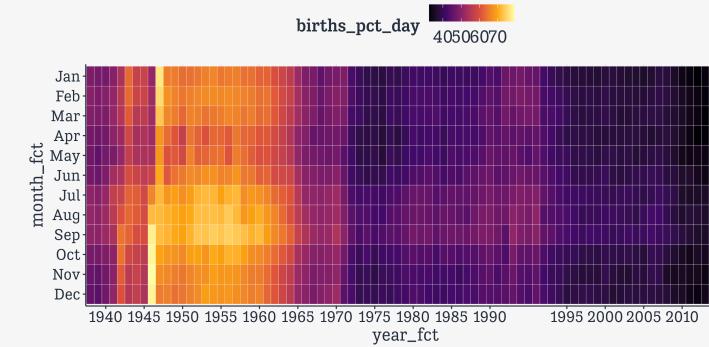
Tiled Heatmap

```
1 okboomer >
2   mutate(year_fct =
3     factor(year,
4       levels = unique(year),
5       ordered = TRUE),
6     month_fct = factor(month,
7       levels = rev(c(1:12)),
8       labels = rev(c("Jan", "Feb", "Mar", "Apr",
9                     "May", "Jun", "Jul", "Aug",
10                    "Sep", "Oct", "Nov", "Dec")),
11       ordered = TRUE)) >
12   select(year, month, year_fct, month_fct, everything()) >
13   filter(country = "United States") >
14   ggplot(aes(x = year_fct, y = month_fct)) +
15   geom_tile(mapping = aes(fill = births_pct_day),
16             color = "white") +
17   scale_x_discrete(breaks = seq(1940, 2010, 5))
```



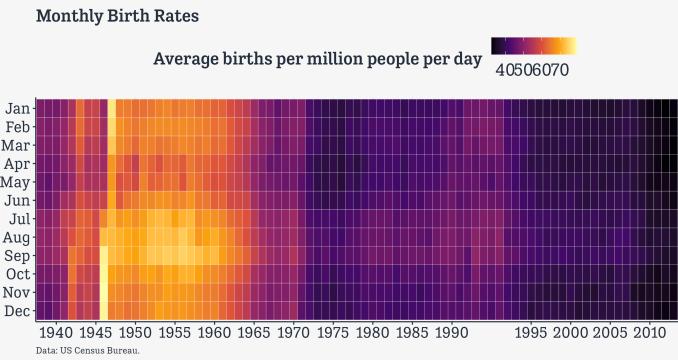
Tiled Heatmap

```
1 okboomer >
2   mutate(year_fct =
3     factor(year,
4       levels = unique(year),
5       ordered = TRUE),
6     month_fct = factor(month,
7       levels = rev(c(1:12)),
8       labels = rev(c("Jan", "Feb", "Mar", "Apr",
9         "May", "Jun", "Jul", "Aug",
10        "Sep", "Oct", "Nov", "Dec")),
11       ordered = TRUE)) >
12   select(year, month, year_fct, month_fct, everything()) >
13   filter(country == "United States") >
14   ggplot(aes(x = year_fct, y = month_fct)) +
15   geom_tile(mapping = aes(fill = births_pct_day),
16             color = "white") +
17   scale_x_discrete(breaks = seq(1940, 2010, 5)) +
18   scale_fill_viridis_c(option = "B")
```



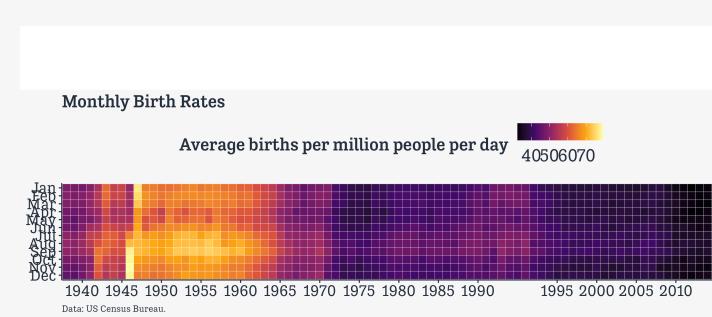
Tiled Heatmap

```
1 okboomer >
2   mutate(year_fct =
3     factor(year,
4       levels = unique(year),
5       ordered = TRUE),
6     month_fct = factor(month,
7       levels = rev(c(1:12)),
8       labels = rev(c("Jan", "Feb", "Mar", "Apr",
9                     "May", "Jun", "Jul", "Aug",
10                    "Sep", "Oct", "Nov", "Dec")),
11       ordered = TRUE)) >
12   select(year, month, year_fct, month_fct, everything()) >
13   filter(country == "United States") >
14   ggplot(aes(x = year_fct, y = month_fct)) +
15   geom_tile(mapping = aes(fill = births_pct_day),
16             color = "white") +
17   scale_x_discrete(breaks = seq(1940, 2010, 5)) +
18   scale_fill_viridis_c(option = "B") +
19   labs(x = NULL, y = NULL,
20        title = "Monthly Birth Rates",
21        fill = "Average births per million people per day",
22        caption = "Data: US Census Bureau.")
```



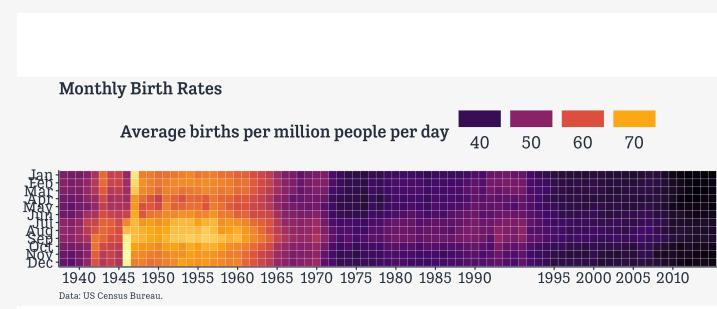
Tiled Heatmap

```
1 okboomer >
2   mutate(year_fct =
3     factor(year,
4       levels = unique(year),
5       ordered = TRUE),
6     month_fct = factor(month,
7       levels = rev(c(1:12)),
8       labels = rev(c("Jan", "Feb", "Mar", "Apr",
9         "May", "Jun", "Jul", "Aug",
10        "Sep", "Oct", "Nov", "Dec"))),
11       ordered = TRUE)) >
12   select(year, month, year_fct, month_fct, everything()) >
13   filter(country == "United States") >
14   ggplot(aes(x = year_fct, y = month_fct)) +
15   geom_tile(mapping = aes(fill = births_pct_day),
16             color = "white") +
17   scale_x_discrete(breaks = seq(1940, 2010, 5)) +
18   scale_fill_viridis_c(option = "B") +
19   labs(x = NULL, y = NULL,
20     title = "Monthly Birth Rates",
21     fill = "Average births per million people per day",
22     caption = "Data: US Census Bureau.") +
23   coord_fixed()
```



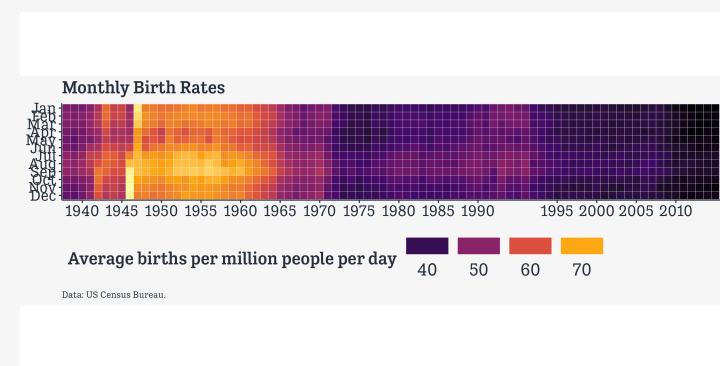
Tiled Heatmap

```
1 okboomer >
2   mutate(year_fct =
3     factor(year,
4       levels = unique(year),
5       ordered = TRUE),
6     month_fct = factor(month,
7       levels = rev(c(1:12)),
8       labels = rev(c("Jan", "Feb", "Mar", "Apr",
9         "May", "Jun", "Jul", "Aug",
10        "Sep", "Oct", "Nov", "Dec")),
11       ordered = TRUE)) >
12   select(year, month, year_fct, month_fct, everything()) >
13   filter(country == "United States") >
14   ggplot(aes(x = year_fct, y = month_fct)) +
15   geom_tile(mapping = aes(fill = births_pct_day),
16             color = "white") +
17   scale_x_discrete(breaks = seq(1940, 2010, 5)) +
18   scale_fill_viridis_c(option = "B") +
19   labs(x = NULL, y = NULL,
20     title = "Monthly Birth Rates",
21     fill = "Average births per million people per day",
22     caption = "Data: US Census Bureau.") +
23   coord_fixed() +
```



Tiled Heatmap

```
1 okboomer >
2   mutate(year_fct =
3     factor(year,
4       levels = unique(year),
5       ordered = TRUE),
6     month_fct = factor(month,
7       levels = rev(c(1:12)),
8       labels = rev(c("Jan", "Feb", "Mar", "Apr",
9         "May", "Jun", "Jul", "Aug",
10        "Sep", "Oct", "Nov", "Dec")),
11       ordered = TRUE)) >
12   select(year, month, year_fct, month_fct, everything()) >
13   filter(country == "United States") >
14   ggplot(aes(x = year_fct, y = month_fct)) +
15   geom_tile(mapping = aes(fill = births_pct_day),
16             color = "white") +
17   scale_x_discrete(breaks = seq(1940, 2010, 5)) +
18   scale_fill_viridis_c(option = "B") +
19   labs(x = NULL, y = NULL,
20     title = "Monthly Birth Rates",
21     fill = "Average births per million people per day",
22     caption = "Data: US Census Bureau.") +
23   coord_fixed() +
```



Tiled Heatmap

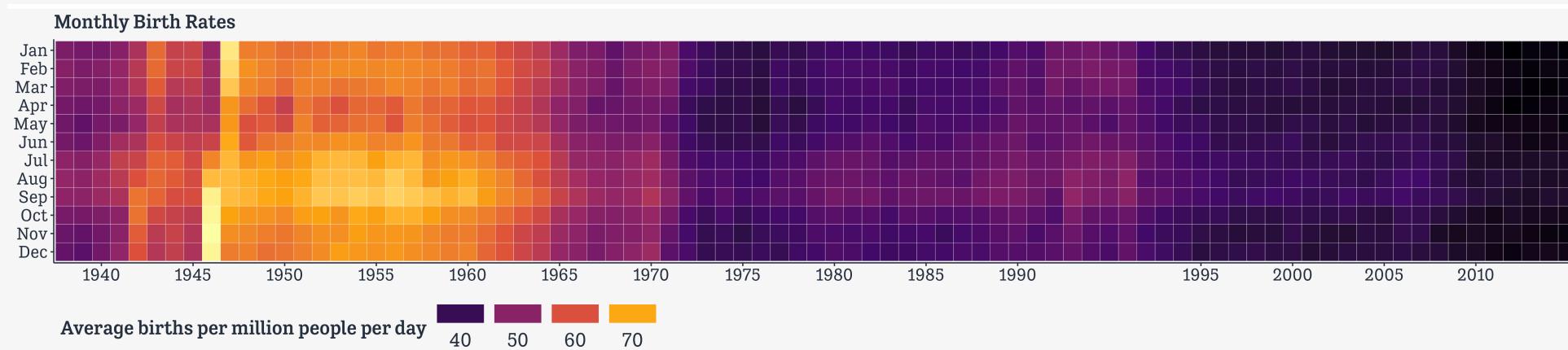
```
1 okboomer >
2     mutate(year_fct =
3         factor(year,
4                 levels = unique(year),
5                 ordered = TRUE),
6         month_fct = factor(month,
7                 levels = rev(c(1:12)),
8                 labels = rev(c("Jan", "Feb", "Mar", "Apr",
9                               "May", "Jun", "Jul", "Aug",
10                              "Sep", "Oct", "Nov", "Dec")),
11                 ordered = TRUE)) >
12     select(year, month, year_fct, month_fct, everything()) >
13     filter(country == "United States") >
14     ggplot(aes(x = year_fct, y = month_fct)) +
15     geom_tile(mapping = aes(fill = births_pct_day),
16               color = "white") +
17     scale_x_discrete(breaks = seq(1940, 2010, 5)) +
18     scale_fill_viridis_c(option = "B") +
19     labs(x = NULL, y = NULL,
20           title = "Monthly Birth Rates",
21           fill = "Average births per million people per day",
22           caption = "Data: US Census Bureau.") +
23     coord_fixed() +
```



Tiled Heatmap

```
1 okboomer >
2     mutate(year_fct =
3         factor(year,
4                 levels = unique(year),
5                 ordered = TRUE),
6         month_fct = factor(month,
7             levels = rev(c(1:12)),
8             labels = rev(c("Jan", "Feb", "Mar", "Apr",
9                         "May", "Jun", "Jul", "Aug",
10                        "Sep", "Oct", "Nov", "Dec")),
11             ordered = TRUE)) >
12     select(year, month, year_fct, month_fct, everything()) >
13     filter(country == "United States") >
14     ggplot(aes(x = year_fct, y = month_fct)) +
15     geom_tile(mapping = aes(fill = births_pct_day),
16               color = "white") +
17     scale_x_discrete(breaks = seq(1940, 2010, 5)) +
18     scale_fill_viridis_c(option = "B") +
19     labs(x = NULL, y = NULL,
20           title = "Monthly Birth Rates",
21           fill = "Average births per million people per day",
22           caption = "Data: US Census Bureau.") +
23     coord_fixed() +
```





Data: US Census Bureau.

The Baby Boom as a tiled temporal heatmap

Beeswarms and bespoke labels

The **cavax** package

```
# remotes::install_github("kjhealy/cavax")
library(cavax)

cavax

# A tibble: 7,032 × 13
  code county name type district city enrollment pbe_pct exempt med_exempt
  <dbl> <chr>  <chr> <chr>   <chr>    <dbl>    <dbl>    <dbl>    <dbl>
1 1.10e5 ALAME... FAME... PUBL... ALAMEDA... NEWA...     109      13  12.8      0
2 6.00e6 ALAME... COX ... PUBL... ALAMEDA... OAKL...     115       1  0.87  0.87
3 6.00e6 ALAME... LAZE... PUBL... ALAMEDA... OAKL...      40       0   0      0
4 1.24e5 ALAME... YU M... PUBL... ALAMEDA... OAKL...      52       10  9.62      0
5 6.10e6 ALAME... AMEL... PUBL... ALAMEDA... ALAM...     128       2  1.56      0
6 6.11e6 ALAME... BAY ... PUBL... ALAMEDA... ALAM...      70       1  1.43      0
7 6.09e6 ALAME... DONA... PUBL... ALAMEDA... ALAM...     100       3   3      0
8 6.09e6 ALAME... EDIS... PUBL... ALAMEDA... ALAM...      70       1  1.43      0
9 6.09e6 ALAME... FRAN... PUBL... ALAMEDA... ALAM...      95       1  1.05  1.05
10 6.09e6 ALAME... FRAN... PUBL... ALAMEDA... ALAM...      50       2   2      0
# i 7,022 more rows
# i 3 more variables: rel_exempt <dbl>, mwc <fct>, kind <fct>
```

Here we will do some custom manual labeling.

Aux Info Panel

```
1 library(ggbeeswarm)
```



Aux Info Panel

```
1 library(ggbeeswarm)
2 make_comma ← scales::label_comma()
```



Aux Info Panel

```
1 library(ggbeeswarm)
2 make_comma ← scales::label_comma()
3
4 cavax
```

```
# A tibble: 7,032 × 13
  code county name   type district city enrollment pbe_pct exempt
  <dbl> <chr>  <chr>  <chr>  <chr>    <dbl>    <dbl>    <dbl>
med_exempt
<dbl>
  1 1.10e5 ALAME... FAME... PUBL... ALAMEDA... NEWA...     109      13  12.8
  0
  2 6.00e6 ALAME... COX ... PUBL... ALAMEDA... OAKL...     115      1  0.87
  0.87
  3 6.00e6 ALAME... LAZE... PUBL... ALAMEDA... OAKL...     40       0  0
  0
  4 1.24e5 ALAME... YU M... PUBL... ALAMEDA... OAKL...     52      10  9.62
  0
  5 6.10e6 ALAME... AMEL... PUBL... ALAMEDA... ALAM...     128      2  1.56
  0
  6 6.11e6 ALAME... BAY ... PUBL... ALAMEDA... ALAM...     70       1  1.43
  0
  7 6.09e6 ALAME... DONA... PUBL... ALAMEDA... ALAM...     100      3  3
  0
  8 6.09e6 ALAME... EDIS... PUBL... ALAMEDA... ALAM...     70       1  1.43
  0
  9 6.09e6 ALAME... FRAN... PUBL... ALAMEDA... ALAM... 1.05
  95      1  1.05
 10 6.09e6 ALAME... FRAN... PUBL... ALAMEDA... ALAM...     50       2  2
  0
```

Aux Info Panel

```
1 library(ggbeeswarm)
2 make_comma ← scales::label_comma()
3
4 cavax ▷
5   group_by(mwc)
```

```
# A tibble: 7,032 × 13
# Groups:   mwc [11]
  code county name  type district city enrollment pbe_pct exempt
  <dbl> <chr>  <chr> <chr>  <chr>    <dbl>    <dbl>    <dbl>
  1 1.10e5 ALAME... FAME... PUBL... ALAMEDA... NEWA...     109      13  12.8
  2 6.00e6 ALAME... COX ... PUBL... ALAMEDA... OAKL...     115      1   0.87
  3 6.00e6 ALAME... LAZE... PUBL... ALAMEDA... OAKL...      40       0   0
  4 1.24e5 ALAME... YU M... PUBL... ALAMEDA... OAKL...      52      10  9.62
  5 6.10e6 ALAME... AMEL... PUBL... ALAMEDA... ALAM...     128      2   1.56
  6 6.11e6 ALAME... BAY ... PUBL... ALAMEDA... ALAM...      70      1   1.43
  7 6.09e6 ALAME... DONA... PUBL... ALAMEDA... ALAM...     100      3     3
  8 6.09e6 ALAME... EDIS... PUBL... ALAMEDA... ALAM...      70      1   1.43
  9 6.09e6 ALAME... FRAN... PUBL... ALAMEDA... ALAM...     95      1   1.05
 10 6.09e6 ALAME... ERAN... PUBL... ALAMEDA... ALAM...     50      2     2
```



Aux Info Panel

```
1 library(ggbeeswarm)
2 make_comma ← scales::label_comma()
3
4 cavax ▷
5   group_by(mwc) ▷
6   summarize(n_schools=n(),
7             n_students = sum(enrollment, na.rm=TRUE))
```

	# A tibble: 11 × 3	n_schools	n_students
mwc	<fct>	<int>	<dbl>
1	Public	5314	472802
2	Charter	314	19863
3	Private Non-Specific	591	16697
4	Private Christian	336	8836
5	Private Catholic	334	9869
6	Private Montessori	99	2112
7	Private Waldorf	16	513
8	Charter Montessori	5	227
9	Public Montessori	11	706
10	Private Christian Montessori	4	78
11	Private Jewish/Islamic	8	237

Aux Info Panel

```
1 library(ggbeeswarm)
2 make_comma ← scales::label_comma()
3
4 cavax ▷
5   group_by(mwc) ▷
6   summarize(n_schools=n(),
7             n_students = sum(enrollment, na.rm=TRUE)) ▷
8   drop_na()
```

	# A tibble: 11 × 3	n_schools	n_students
mwc	<fct>	<int>	<dbl>
1	Public	5314	472802
2	Charter	314	19863
3	Private Non-Specific	591	16697
4	Private Christian	336	8836
5	Private Catholic	334	9869
6	Private Montessori	99	2112
7	Private Waldorf	16	513
8	Charter Montessori	5	227
9	Public Montessori	11	706
10	Private Christian Montessori	4	78
11	Private Jewish/Islamic	8	237

Aux Info Panel

```
1 library(ggbeeswarm)
2 make_comma ← scales::label_comma()
3
4 cavax ▷
5   group_by(mwc) ▷
6   summarize(n_schools=n(),
7             n_students = sum(enrollment, na.rm=TRUE)) ▷
8   drop_na() ▷
9   mutate(n_schools_fmt = make_comma(n_schools),
10         n_students_fmt = make_comma(n_students),
11         info_schools = paste(n_schools_fmt, "Schools E",
12         info_students = paste(n_students_fmt, "Kinderga
```

# A tibble: 11 × 7	mwc	n_schools	n_students	n_schools_fmt	n_students_fmt	
info_schools	<fct>	<int>	<dbl>	<chr>	<chr>	<chr>
Schoo...	1 Public	5314	472802	5,314	472,802	5,314
Schools...	2 Charter	314	19863	314	19,863	314
Schools...	3 Private Non-S...	591	16697	591	16,697	591
Schools...	4 Private Chris...	336	8836	336	8,836	336
Schools...	5 Private Catho...	334	9869	334	9,869	334
Schools ...	6 Private Monte...	99	2112	99	2,112	99
Schools ...	7 Private Waldo...	16	513	16	513	16
Schools ...	8 Charter Monte...	5	227	5	227	5
Schools ...	9 Public Montes...	11	706	11	706	11
Schools ...	10 Private Chris...	4	78	4	78	4
Schools ...	11 Private Lewis...	8	237	8	237	8



Aux Info Panel

```
1 library(ggbeeswarm)
2 make_comma ← scales::label_comma()
3
4 cavax ▷
5   group_by(mwc) ▷
6   summarize(n_schools=n(),
7             n_students = sum(enrollment, na.rm=TRUE)) ▷
8   drop_na() ▷
9   mutate(n_schools_fmt = make_comma(n_schools),
10         n_students_fmt = make_comma(n_students),
11         info_schools = paste(n_schools_fmt, "Schools Enrolled"),
12         info_students = paste(n_students_fmt, "Kindergarten Students"))
13 aux_info
```



A little kludge

```
1 ## This is not an efficient way to do this  
2 aux_info
```

```
# A tibble: 11 × 7  
#>   mwc          n_schools n_students n_schools_fmt  
#>   <fct>        <int>     <dbl> <chr>      <chr>  
#>   <chr>  
#>   1 Public       5314     472802  5,314     472,802  
#>   5,314 Schoo...  
#>   2 Charter      314      19863   314      19,863  
#>   314 Schools...  
#>   3 Private Non-S...  591      16697   591      16,697  
#>   591 Schools...  
#>   4 Private Chris...  336      8836    336      8,836  
#>   336 Schools...  
#>   5 Private Catho...  334      9869    334      9,869  
#>   334 Schools...  
#>   6 Private Monte...  99       2112    99      2,112  
#>   99 Schools ...  
#>   7 Private Waldo...  16       513     16      513  
#>   16 Schools ...  
#>   8 Charter Monte...  5        227     5       227  
#>   5 Schools E...  
#>   9 Public Montes...  11      706     11      706  
#>   11 Schools ...  
#>   10 Private Chris...  4        78      4       78  
#>   4 Schools E...
```

A little kludge

```
1 ## This is not an efficient way to do this
2 aux_info %>
3   select(mwc, info_schools, info_students)
```

```
# A tibble: 11 × 3
#>   mwc          info_schools
#>   <fct>        <chr>           <chr>
#> 1 Public       5,314 Schools Enrolling 472,802
#> 2 Charter      314 Schools Enrolling  19,863
#> 3 Private Non-Specific 591 Schools Enrolling 16,697
#> 4 Private Christian 336 Schools Enrolling  8,836
#> 5 Private Catholic 334 Schools Enrolling  9,869
#> 6 Private Montessori 99 Schools Enrolling  2,112
#> 7 Private Waldorf  16 Schools Enrolling  513
#> 8 Charter Montessori 5 Schools Enrolling  227
#> 9 Public Montessori 11 Schools Enrolling  706
#> 10 Private Christian Montessori 4 Schools Enrolling  78
#> 11 Private Jewish/Islamic 8 Schools Enrolling  237
```



A little kludge

```
1 ## This is not an efficient way to do this
2 aux_info %>
3   select(mwc, info_schools, info_students) %>
4   mutate(across(everything(), as.character))
```

```
# A tibble: 11 × 3
#>   mwc          info_schools
#>   <chr>        <chr>           <chr>
#> 1 Public       5,314 Schools Enrolling 472,802
#> 2 Charter      314 Schools Enrolling  19,863
#> 3 Kindergarteners
#> 4 Private Non-Specific
#> 5 Private Christian
#> 6 Private Catholic
#> 7 Private Montessori
#> 8 Private Waldorf
#> 9 Public Montessori
#> 10 Private Christian Montessori
#> 11 Private Jewish/Islamic
```

```
          Enrolling    16,697
          Enrolling    8,836
          Enrolling    9,869
          Enrolling    2,112
          Enrolling    513
          Enrolling    227
          Enrolling    706
          Enrolling    78
          Enrolling    237
```

A little kludge

```
1 ## This is not an efficient way to do this
2 aux_info %>
3   select(mwc, info_schools, info_students) %>
4   mutate(across(everything(), as.character())) %>
5   group_by(mwc)
```

```
# A tibble: 11 × 3
# Groups:   mwc [11]
#   mwc          info_schools
#   <chr>        <chr>      <chr>
# 1 Public       5,314 Schools Enrolling 472,802
# 2 Charter      314 Schools Enrolling  19,863
# 3 Kindergarteners
# 4 Private Non-Specific  591 Schools Enrolling  16,697
# 5 Kindergarteners
# 6 Private Christian  336 Schools Enrolling  8,836
# 7 Kindergarteners
# 8 Private Catholic  334 Schools Enrolling  9,869
# 9 Kindergarteners
# 10 Private Montessori 99 Schools Enrolling  2,112
# 11 Kindergarteners
# 12 Private Waldorf  16 Schools Enrolling  513
# 13 Kindergarteners
# 14 Charter Montessori 5 Schools Enrolling  227
# 15 Kindergarteners
# 16 Public Montessori 11 Schools Enrolling  706
# 17 Kindergarteners
# 18 Private Christian Montessori 4 Schools Enrolling  78
# 19 Kindergarteners
```

A little kludge

```
1 ## This is not an efficient way to do this
2 aux_info %>
3   select(mwc, info_schools, info_students) %>
4   mutate(across(everything(), as.character)) %>
5   group_by(mwc) %>
6   group_keys()
```

```
# A tibble: 11 × 1
  mwc
  <chr>
  1 Charter
  2 Charter Montessori
  3 Private Catholic
  4 Private Christian
  5 Private Christian Montessori
  6 Private Jewish/Islamic
  7 Private Montessori
  8 Private Non-Specific
  9 Private Waldorf
 10 Public
 11 Public Montessori
```

A little kludge

```
1 ## This is not an efficient way to do this
2 aux_info >
3   select(mwc, info_schools, info_students) >
4   mutate(across(everything(), as.character)) >
5   group_by(mwc) >
6   group_keys() >
7   pull()
```

```
[1] "Charter"                  "Charter Montessori"
[3] "Private Catholic"          "Private Christian"
[5] "Private Christian Montessori" "Private Jewish/Islamic"
[7] "Private Montessori"         "Private Non-Specific"
[9] "Private Waldorf"           "Public"
[11] "Public Montessori"
```

A little kludge

```
1 ## This is not an efficient way to do this
2 aux_info >
3   select(mwc, info_schools, info_students) >
4   mutate(across(everything(), as.character)) >
5   group_by(mwc) >
6   group_keys() >
7   pull() >
8   as.character()
```

```
[1] "Charter"                  "Charter Montessori"
[3] "Private Catholic"          "Private Christian"
[5] "Private Christian Montessori" "Private Jewish/Islamic"
[7] "Private Montessori"         "Private Non-Specific"
[9] "Private Waldorf"           "Public"
[11] "Public Montessori"
```

A little kludge

```
1 ## This is not an efficient way to do this
2 aux_info >
3   select(mwc, info_schools, info_students) >
4   mutate(across(everything(), as.character)) >
5   group_by(mwc) >
6   group_keys() >
7   pull() >
8   as.character() ->
9   keys
```

A little kludge

```
1 ## This is not an efficient way to do this
2 aux_info >
3   select(mwc, info_schools, info_students) >
4   mutate(across(everything(), as.character())) >
5   group_by(mwc) >
6   group_keys() >
7   pull() >
8   as.character() →
9   keys
10
11 aux_info
```

```
# A tibble: 11 × 7
#>   mwc          n_schools n_students n_schools_fmt
#>   <fct>        <int>     <dbl>    <chr>      <chr>
#>   <chr>
#>   1 Public       5314     472802  5,314      472,802
#>   5,314 Schoo...
#>   2 Charter      314      19863   314      19,863
#>   314 Schools...
#>   3 Private Non-S... 591      16697   591      16,697
#>   591 Schools...
#>   4 Private Chris... 336      8836    336      8,836
#>   336 Schools...
#>   5 Private Catho... 334      9869    334      9,869
#>   334 Schools...
#>   6 Private Monte... 99       2112    99      2,112
#>   99 Schools ...
#>   7 Private Waldo... 16       513     16      513
#>   16 Schools ...
#>   8 Charter Monte... 5        227     5       227
#>   5 Schools E...
#>   9 Public Montes... 11      706     11      706
#>   11 Schools ...
#>   10 Private Chris... 4        78      4       78
#>   4 Schools E...
```

A little kludge

```
1 ## This is not an efficient way to do this
2 aux_info %>
3   select(mwc, info_schools, info_students) %>
4   mutate(across(everything(), as.character())) %>
5   group_by(mwc) %>
6   group_keys() %>
7   pull() %>
8   as.character() %>
9   keys
10
11 aux_info %>
12   select(mwc, info_schools, info_students)
```

```
# A tibble: 11 × 3
# ... with 11 variables:
#   mwc      <fct>    info_schools <chr>    <chr>
#   info_students <fct>    <chr>      <chr>    <chr>
#   1 Public     5,314 Schools Enrolling 472,802
#   Kindergarteners 314 Schools Enrolling 19,863
#   2 Charter    591 Schools Enrolling 16,697
#   Kindergarteners 336 Schools Enrolling 8,836
#   3 Private Non-Specific 334 Schools Enrolling 9,869
#   Kindergarteners 99 Schools Enrolling 2,112
#   4 Private Christian 16 Schools Enrolling 513
#   Kindergarteners 5 Schools Enrolling 227
#   5 Private Catholic 11 Schools Enrolling 706
#   Kindergarteners 4 Schools Enrolling 78
#   6 Private Montessori 8 Charter Montessori 11 Schools Enrolling 706
#   Kindergarteners 4 Schools Enrolling 78
#   7 Private Waldorf 9 Public Montessori 11 Schools Enrolling 706
#   Kindergarteners 4 Schools Enrolling 78
#   8 Charter Montessori 10 Private Christian Montessori 4 Schools Enrolling 78
#   Kindergarteners 8 Schools Enrolling 237
#   9 Public Montessori 11 Private Jewish/Islamic 8 Schools Enrolling 237
```



A little kludge

```
1 ## This is not an efficient way to do this
2 aux_info %>
3   select(mwc, info_schools, info_students) %>
4   mutate(across(everything(), as.character)) %>
5   group_by(mwc) %>
6   group_keys() %>
7   pull() %>
8   as.character() %>
9   keys
10
11 aux_info %>
12   select(mwc, info_schools, info_students) %>
13   mutate(across(everything(), as.character))
```

```
# A tibble: 11 × 3
#>   mwc          info_schools
#>   <chr>        <chr>           <chr>
#> 1 Public       5,314 Schools Enrolling 472,802
#> 2 Charter      314 Schools Enrolling  19,863
#> 3 Private Non-Specific 591 Schools Enrolling 16,697
#> 4 Private Christian 336 Schools Enrolling  8,836
#> 5 Private Catholic 334 Schools Enrolling  9,869
#> 6 Private Montessori 99 Schools Enrolling  2,112
#> 7 Private Waldorf  16 Schools Enrolling  513
#> 8 Charter Montessori 5 Schools Enrolling  227
#> 9 Public Montessori 11 Schools Enrolling  706
#> 10 Private Christian Montessori 4 Schools Enrolling  78
#> 11 Private Jewish/Islamic 8 Schools Enrolling  237
```



A little kludge

```
1 ## This is not an efficient way to do this
2 aux_info >
3   select(mwc, info_schools, info_students) >
4   mutate(across(everything(), as.character)) >
5   group_by(mwc) >
6   group_keys() >
7   pull() >
8   as.character() →
9   keys
10
11 aux_info >
12   select(mwc, info_schools, info_students) >
13   mutate(across(everything(), as.character)) >
14   group_split(mwc)
```

```
<list_of<
tbl_df<
  mwc           : character
  info_schools : character
  info_students: character
>
>[1]>
[[1]]
# A tibble: 1 × 3
  mwc      info_schools       info_students
  <chr>    <chr>            <chr>
1 Charter 314 Schools Enrolling 19,863 Kindergarteners

[[2]]
# A tibble: 1 × 3
  mwc           info_schools       info_students
  <chr>         <chr>            <chr>
1 Charter Montessori 5 Schools Enrolling 227 Kindergarteners

[[3]]
# A tibble: 1 × 3
  mwc           info_schools       info_students
  <chr>         <chr>            <chr>
1 Private Catholic 334 Schools Enrolling 9,869 Kindergarteners
```

A little kludge

```
1 ## This is not an efficient way to do this
2 aux_info >
3   select(mwc, info_schools, info_students) >
4   mutate(across(everything(), as.character)) >
5   group_by(mwc) >
6   group_keys() >
7   pull() >
8   as.character() →
9   keys
10
11 aux_info >
12   select(mwc, info_schools, info_students) >
13   mutate(across(everything(), as.character)) >
14   group_split(mwc) >
15   set_names(keys) # There's a better way ...
```

```
<list_of<
tbl_df<
  mwc           : character
  info_schools : character
  info_students: character
>
>[11]>
$Charter
# A tibble: 1 × 3
  mwc      info_schools      info_students
  <chr>    <chr>            <chr>
1 Charter  314 Schools Enrolling 19,863 Kindergarteners

$`Charter Montessori`
# A tibble: 1 × 3
  mwc      info_schools      info_students
  <chr>    <chr>            <chr>
1 Charter Montessori 5 Schools Enrolling 227 Kindergarteners

$`Private Catholic`
# A tibble: 1 × 3
  mwc      info_schools      info_students
  <chr>    <chr>            <chr>
1 Private Catholic 334 Schools Enrolling 9,869 Kindergarteners
```

A little kludge

```
1 ## This is not an efficient way to do this
2 aux_info %>
3   select(mwc, info_schools, info_students) %>
4   mutate(across(everything(), as.character)) %>
5   group_by(mwc) %>
6   group_keys() %>
7   pull() %>
8   as.character() %>
9   keys
10
11 aux_info %>
12   select(mwc, info_schools, info_students) %>
13   mutate(across(everything(), as.character)) %>
14   group_split(mwc) %>
15   set_names(keys) %># There's a better way ...
16   map_chr(.f = paste, sep = "", collapse = "\n")
```

```
Charter          "Charter\n314 Schools Enrolling\n19,863
Kindergarteners"
                           Charter
Montessori        "Charter Montessori\n5 Schools Enrolling\n227
Kindergarteners"
                           Private
Catholic          "Private Catholic\n334 Schools Enrolling\n9,869
Kindergarteners"
                           Private
Christian         "Private Christian\n336 Schools Enrolling\n8,836
Kindergarteners"
                           Private Christian
Montessori        "Private Christian Montessori\n4 Schools Enrolling\n78
Kindergarteners"
                           Private
Jewish/Islamic    "Private Jewish/Islamic\n8 Schools Enrolling\n237
Kindergarteners"
                           Private
```



A little kludge

```
1 ## This is not an efficient way to do this
2 aux_info >
3   select(mwc, info_schools, info_students) >
4   mutate(across(everything(), as.character)) >
5   group_by(mwc) >
6   group_keys() >
7   pull() >
8   as.character() →
9   keys
10
11 aux_info >
12   select(mwc, info_schools, info_students) >
13   mutate(across(everything(), as.character)) >
14   group_split(mwc) >
15   set_names(keys) > # There's a better way ...
16   map_chr(.f = paste, sep = "", collapse = "\n") →
17   special_x_labs
```



A little kludge

```
1 ## This is not an efficient way to do this
2 aux_info >
3   select(mwc, info_schools, info_students) >
4   mutate(across(everything(), as.character)) >
5   group_by(mwc) >
6   group_keys() >
7   pull() >
8   as.character() →
9   keys
10
11 aux_info >
12   select(mwc, info_schools, info_students) >
13   mutate(across(everything(), as.character)) >
14   group_split(mwc) >
15   set_names(keys) > # There's a better way ...
16   map_chr(.f = paste, sep = "", collapse = "\n") →
17   special_x_labs
```



At last, the Beeplot

```
1 cavax
```

```
# A tibble: 7,032 x 13
  code county name  type district city enrollment pbe_pct
  <dbl> <chr>  <chr> <chr>   <chr>    <dbl>    <dbl>
exempt med_exempt
  <dbl> <dbl>   <dbl> <dbl>   <dbl>    <dbl>    <dbl>
  1 1.10e5 ALAME... FAME... PUBL... ALAMEDA... NEWA...
12.8      0          109      13
  2 6.00e6 ALAME... COX ... PUBL... ALAMEDA... OAKL...
0.87      0.87        115      1
  3 6.00e6 ALAME... LAZE... PUBL... ALAMEDA... OAKL...
0          40         0       0  0
  4 1.24e5 ALAME... YU M... PUBL... ALAMEDA... OAKL...
9.62      0          52       10
  5 6.10e6 ALAME... AMEL... PUBL... ALAMEDA... ALAM...
1.56      0          128      2
  6 6.11e6 ALAME... BAY ... PUBL... ALAMEDA... ALAM...
1.43      0          70       1
  7 6.09e6 ALAME... DONA... PUBL... ALAMEDA... ALAM...
0          100      3     3
  8 6.09e6 ALAME... EDIS... PUBL... ALAMEDA... ALAM...
1.43      0          70       1
  9 6.09e6 ALAME... FRAN... PUBL... ALAMEDA... ALAM...
1.05      1.05        95       1
 10 6.09e6 ALAME... FRAN... PUBL... ALAMEDA... ALAM...
0          50       2     2
```



At last, the Beeplot

```
1 cavax %>%
2   filter(mwc %nin% c("Private Christian Montessori",
3   "Charter Montessori",
4   "Private Jewish/Islamic"))
```

```
# A tibble: 7,015 × 13
  code county name type district city enrollment pbe_pct
  <dbl> <chr>  <chr> <chr>  <chr>    <dbl>      <dbl>
1 1.10e5 ALAME... FAME... PUBL... ALAMEDA... NEWA... 109     13
2 6.00e6 ALAME... COX ... PUBL... ALAMEDA... OAKL... 115      1
3 6.00e6 ALAME... LAZE... PUBL... ALAMEDA... OAKL... 40      0   0
4 1.24e5 ALAME... YU M... PUBL... ALAMEDA... OAKL... 52      10
5 6.10e6 ALAME... AMEL... PUBL... ALAMEDA... ALAM... 128      2
6 6.11e6 ALAME... BAY ... PUBL... ALAMEDA... ALAM... 70      1
7 6.09e6 ALAME... DONA... PUBL... ALAMEDA... ALAM... 100     3   3
8 6.09e6 ALAME... EDIS... PUBL... ALAMEDA... ALAM... 70      1
9 6.09e6 ALAME... FRAN... PUBL... ALAMEDA... ALAM... 95      1
10 6.09e6 ALAME... FRAN... PUBL... ALAMEDA... ALAM... 50      2   2
```

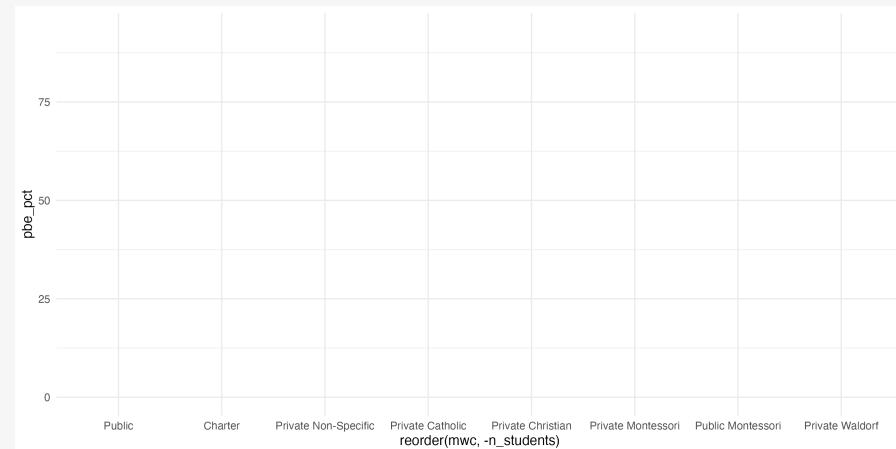
At last, the Beeplot

```
1 cavax %>
2   filter(mwc %in% c("Private Christian Montessori",
3                     "Charter Montessori",
4                     "Private Jewish/Islamic")) %>
5   left_join(aux_info, by = "mwc")
```

```
# A tibble: 7,015 x 19
  code county name type district city enrollment pbe_pct
  <dbl> <chr>  <chr> <chr>  <chr>    <dbl>      <dbl>
exempt med_exempt
<dbl> <dbl>  <dbl> <dbl>  <dbl>    <dbl>      <dbl>
<dbl>  <dbl>
1 1.10e5 ALAME... FAME... PUBL... ALAMEDA... NEWA...
12.8     0          109       13
2 6.00e6 ALAME... COX ... PUBL... ALAMEDA... OAKL...
0.87     0.87        115       1
3 6.00e6 ALAME... LAZE... PUBL... ALAMEDA... OAKL...
0          40         0         0
4 1.24e5 ALAME... YU M... PUBL... ALAMEDA... OAKL...
9.62     0          52        10
5 6.10e6 ALAME... AMEL... PUBL... ALAMEDA... ALAM...
1.56     0          128       2
6 6.11e6 ALAME... BAY ... PUBL... ALAMEDA... ALAM...
1.43     0          70        1
7 6.09e6 ALAME... DONA... PUBL... ALAMEDA... ALAM...
0          100       3        3
8 6.09e6 ALAME... EDIS... PUBL... ALAMEDA... ALAM...
1.43     0          70        1
9 6.09e6 ALAME... FRAN... PUBL... ALAMEDA... ALAM...
1.05     1.05        95        1
10 6.09e6 ALAME... FRAN... PUBL... ALAMEDA... ALAM...
0          50        2        2
```

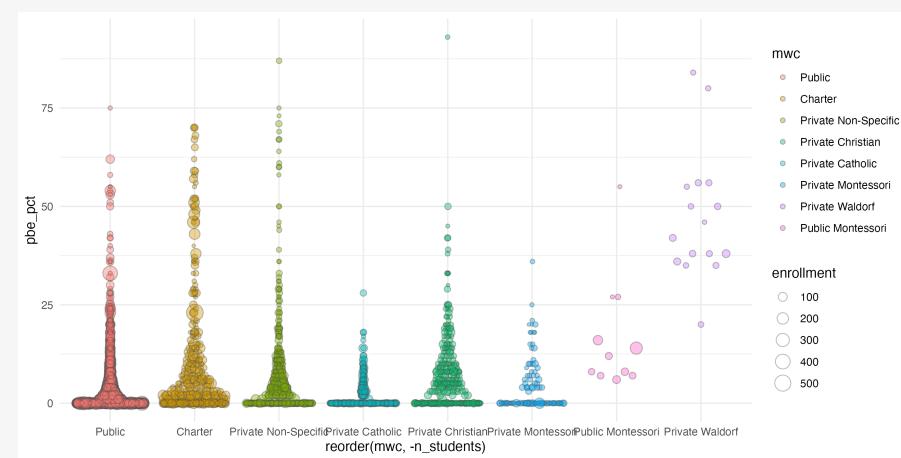
At last, the Beeplot

```
1 cavax %>
2   filter(mwc %in% c("Private Christian Montessori",
3                     "Charter Montessori",
4                     "Private Jewish/Islamic")) %>
5   left_join(aux_info, by = "mwc") %>
6   ggplot(mapping =
7     aes(y = pbe_pct,
8         x = reorder(mwc, -n_students),
9         size = enrollment,
10        fill = mwc))
```



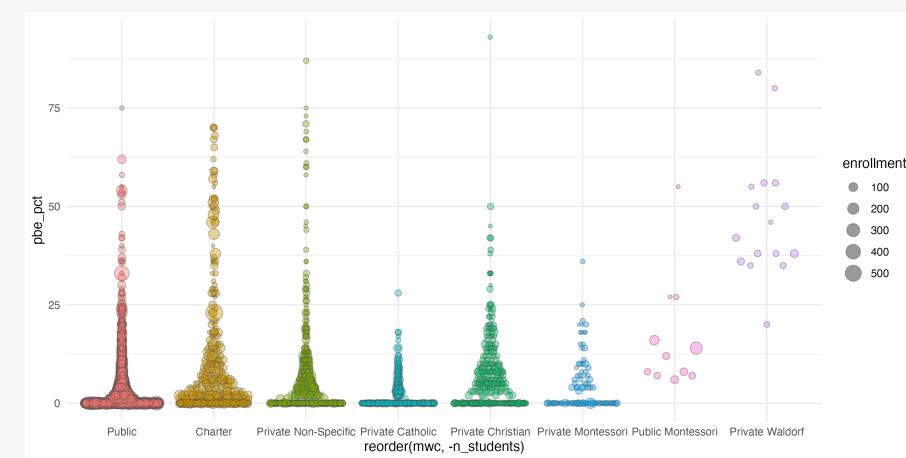
At last, the Beeplot

```
1 cavax %>
2   filter(mwc %in% c("Private Christian Montessori",
3   "Charter Montessori",
4   "Private Jewish/Islamic")) %>
5   left_join(aux_info, by = "mwc") %>
6   ggplot(mapping =
7     aes(y = pbe_pct,
8         x = reorder(mwc, -n_students),
9         size = enrollment,
10        fill = mwc)) +
11   geom_quasirandom(shape=21,
12     alpha = 0.4,color="gray30",
13     method = "quasirandom",
14     varwidth = FALSE,
15     bandwidth = 0.9)
```



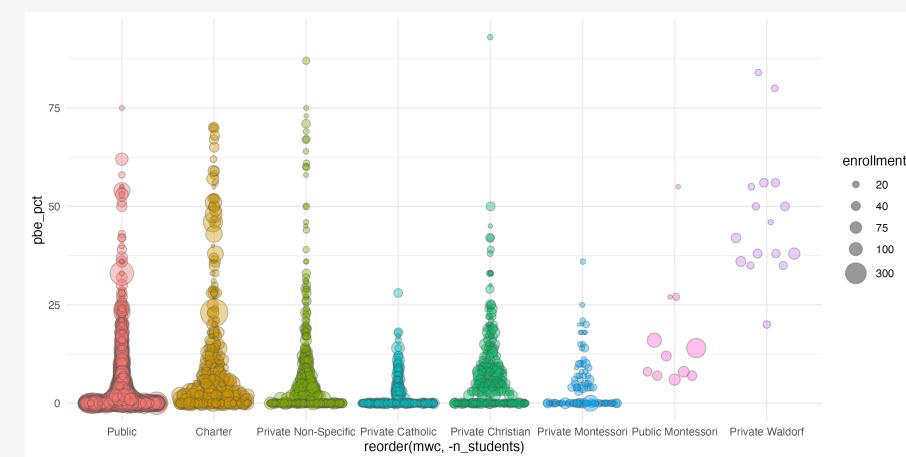
At last, the Beeplot

```
1 cavax >
2   filter(mwc %nin% c("Private Christian Montessori",
3   "Charter Montessori",
4   "Private Jewish/Islamic")) >
5   left_join(aux_info, by = "mwc") >
6   ggplot(mapping =
7     aes(y = pbe_pct,
8       x = reorder(mwc, -n_students),
9       size = enrollment,
10      fill = mwc)) +
11   geom_quasirandom(shape=21,
12     alpha = 0.4,color="gray30",
13     method = "quasirandom",
14     varwidth = FALSE,
15     bandwidth = 0.9) +
16   guides(color = "none",
17     shape= "none",
18     fill= "none",
19     size = guide_legend(override.aes =
20       list(fill = "black")))
```



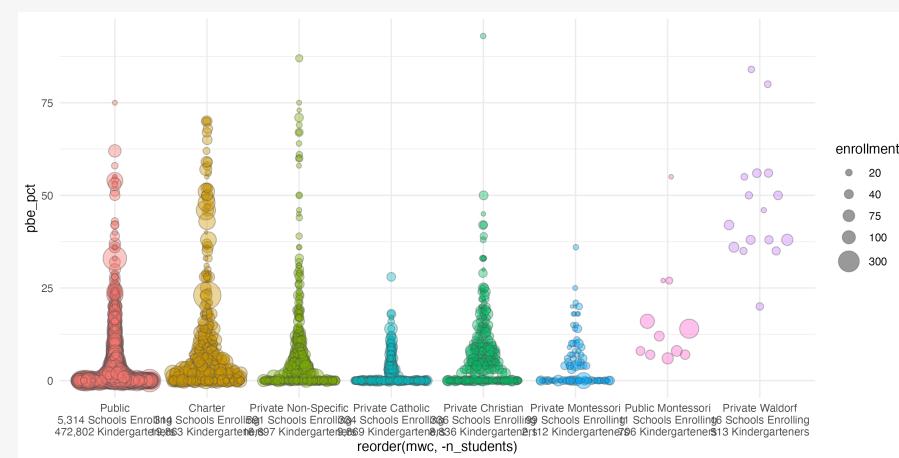
At last, the Beeplot

```
1 cavax >
2   filter(mwc %nin% c("Private Christian Montessori",
3   "Charter Montessori",
4   "Private Jewish/Islamic")) >
5   left_join(aux_info, by = "mwc") >
6   ggplot(mapping =
7     aes(y = pbe_pct,
8       x = reorder(mwc, -n_students),
9       size = enrollment,
10      fill = mwc)) +
11   geom_quasirandom(shape=21,
12     alpha = 0.4,color="gray30",
13     method = "quasirandom",
14     varwidth = FALSE,
15     bandwidth = 0.9) +
16   guides(color = "none",
17     shape= "none",
18     fill= "none",
19     size = guide_legend(override.aes =
20       list(fill = "black")) ) +
21   scale_size(breaks=c(20, 40, 75, 100, 300),
22             range=c(1,10))
```



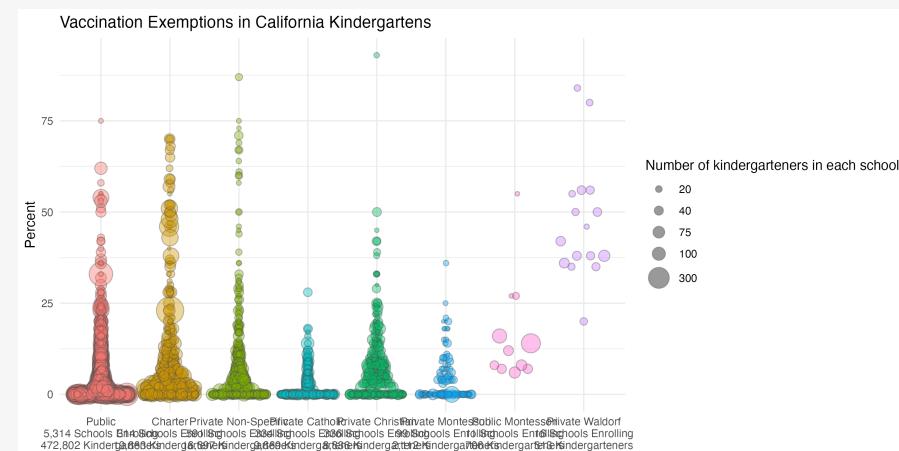
At last, the Beeplot

```
1 cavax %>
2   filter(mwc %nin% c("Private Christian Montessori",
3   "Charter Montessori",
4   "Private Jewish/Islamic")) %>
5   left_join(aux_info, by = "mwc") %>
6   ggplot(mapping =
7     aes(y = pbe_pct,
8       x = reorder(mwc, -n_students),
9       size = enrollment,
10      fill = mwc)) +
11   geom_quasirandom(shape=21,
12     alpha = 0.4,color="gray30",
13     method = "quasirandom",
14     varwidth = FALSE,
15     bandwidth = 0.9) +
16   guides(color = "none",
17     shape= "none",
18     fill= "none",
19     size = guide_legend(override.aes =
20       list(fill = "black")))) +
21   scale_size(breaks=c(20, 40, 75, 100, 300),
22             range=c(1,10)) +
23   scale_x_discrete(labels = special_x_labs)
```



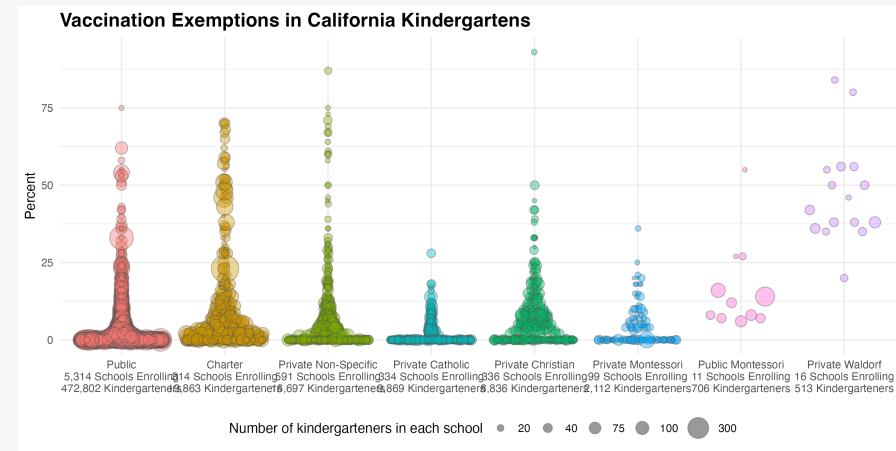
At last, the Beeplot

```
1 cavax >
2   filter(mwc %nin% c("Private Christian Montessori",
3                     "Charter Montessori",
4                     "Private Jewish/Islamic")) >
5   left_join(aux_info, by = "mwc") >
6   ggplot(mapping =
7     aes(y = pbe_pct,
8         x = reorder(mwc, -n_students),
9         size = enrollment,
10        fill = mwc)) +
11   geom_quasirandom(shape=21,
12                     alpha = 0.4,color="gray30",
13                     method = "quasirandom",
14                     varwidth = FALSE,
15                     bandwidth = 0.9) +
16   guides(color = "none",
17          shape= "none",
18          fill= "none",
19          size = guide_legend(override.aes =
20                            list(fill = "black"))) +
21   scale_size(breaks=c(20, 40, 75, 100, 300),
22              range=c(1,10)) +
23   scale_x_discrete(labels = special_x_labs) +
```



At last, the Beeplot

```
1 cavax >
2   filter(mwc %nin% c("Private Christian Montessori",
3   "Charter Montessori",
4   "Private Jewish/Islamic")) >
5   left_join(aux_info, by = "mwc") >
6   ggplot(mapping =
7     aes(y = pbe_pct,
8       x = reorder(mwc, -n_students),
9       size = enrollment,
10      fill = mwc)) +
11   geom_quasirandom(shape=21,
12     alpha = 0.4,color="gray30",
13     method = "quasirandom",
14     varwidth = FALSE,
15     bandwidth = 0.9) +
16   guides(color = "none",
17     shape= "none",
18     fill= "none",
19     size = guide_legend(override.aes =
20       list(fill = "black")) ) +
21   scale_size(breaks=c(20, 40, 75, 100, 300),
22             range=c(1,10)) +
23   scale_x_discrete(labels = special_x_labs) +
```



At last, the Beeplot

```
1 cavax %>
2   filter(mwc %in% c("Private Christian Montessori",
3                     "Charter Montessori",
4                     "Private Jewish/Islamic")) %>
5   left_join(aux_info, by = "mwc") %>
6   ggplot(mapping =
7     aes(y = pbe_pct,
8         x = reorder(mwc, -n_students),
9         size = enrollment,
10        fill = mwc)) +
11   geom_quasirandom(shape=21,
12     alpha = 0.4,color="gray30",
13     method = "quasirandom",
14     varwidth = FALSE,
15     bandwidth = 0.9) +
16   guides(color = "none",
17     shape= "none",
18     fill= "none",
19     size = guide_legend(override.aes =
20       list(fill = "black")))) +
21   scale_size(breaks=c(20, 40, 75, 100, 300),
22             range=c(1,10)) +
23   scale_x_discrete(labels = special_x_labs) +
```

At last, the Beeplot

```
1 cavax %>
2   filter(mwc %in% c("Private Christian Montessori",
3   "Charter Montessori",
4   "Private Jewish/Islamic")) %>
5   left_join(aux_info, by = "mwc") %>
6   ggplot(mapping =
7     aes(y = pbe_pct,
8         x = reorder(mwc, -n_students),
9         size = enrollment,
10        fill = mwc)) +
11   geom_quasirandom(shape=21,
12     alpha = 0.4,color="gray30",
13     method = "quasirandom",
14     varwidth = FALSE,
15     bandwidth = 0.9) +
16   guides(color = "none",
17     shape= "none",
18     fill= "none",
19     size = guide_legend(override.aes =
20       list(fill = "black")))) +
21   scale_size(breaks=c(20, 40, 75, 100, 300),
22             range=c(1,10)) +
23   scale_x_discrete(labels = special_x_labs) +
```

Vaccination Exemptions in California Kindergartens

