

# STAT 231: R Q & A Discussion Forum

Sept. 3, 2020

## Original code

Below is the code, slightly modified, from the extended example on historical baby names in section 3.3.1. You can work off the section(s) of this code relevant to your assigned question as you explore answers.

**QX.** [Insert the question you're answering here]

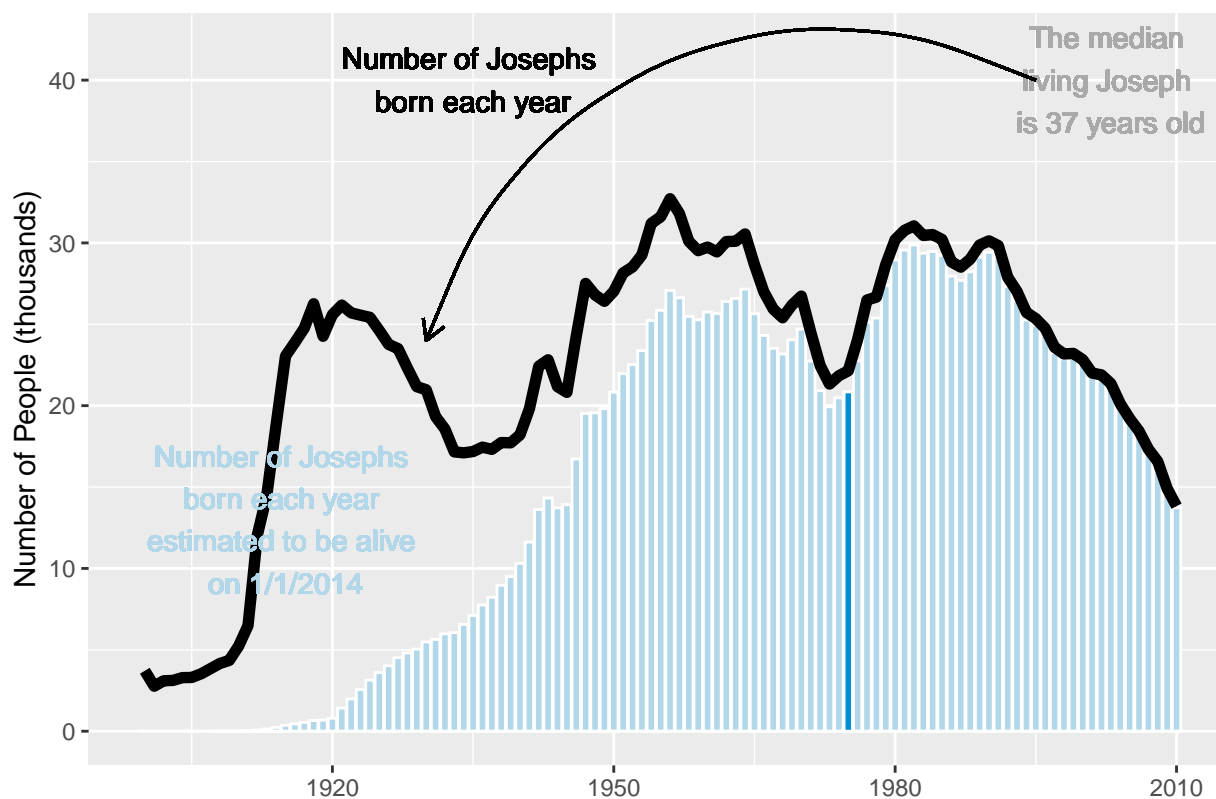
Include your answer and explanation here. You can intersperse text with code in your explanation.

How does `geom_curve()` work? How was the shape and position of the arrow to median year of birth set? What does “curvature = 0.5” do? What is the purpose of the arrow argument?

The `geom_curve()` function creates a curved line that can be added on to ggplots. In the case of the babies named Joseph plot, the curve points to the median age of Josephs. The position of the curve is defined by x and y ranges, which are decided by the x/xend and y/yend values. The 0.5 curvature value corresponds to a 90 degree counter-clockwise turn; a 1 value creates a semicircle, while a -0.5 curvature value moves in the clockwise direction. The arrow argument is a style addition placed at the end of the curve, which effectively points to the median value of the plot. See nameplots x, y, and z below for visualizations of three of these `geom_curve()` inputs.

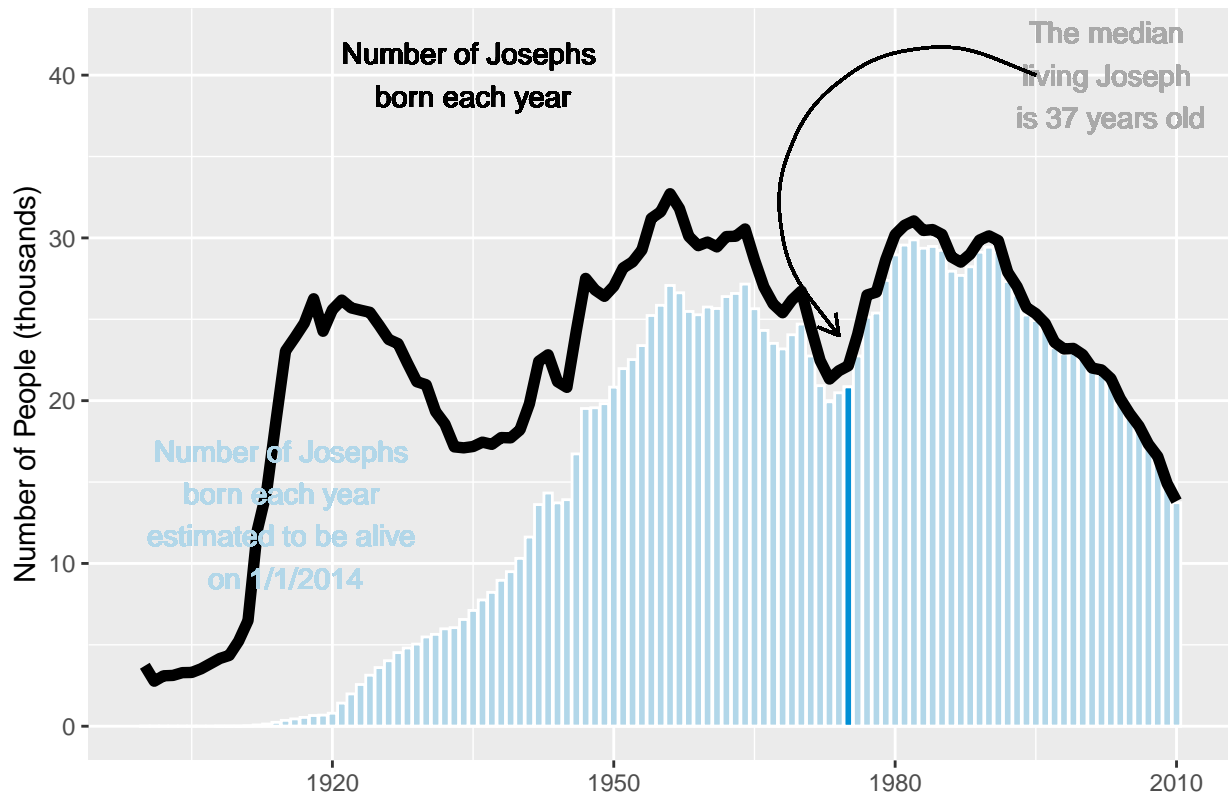
```
#Position Adjustment: xend adjusted to 1930 to lengthen the curve.
name_plotx <- name_plot +
  geom_bar(stat = "identity", color = "white", fill = "#008fd5"
    , aes(y = ifelse(year == median_yob, est_alive_today/1000, 0))) +
  ggtitle("Age Distribution of American Boys Named Joseph") +
  geom_text(x = 1935, y = 40, label = "Number of Josephs \nborn each year") +
  geom_text(x = 1915, y = 13
    , label = "Number of Josephs \nborn each year \nestimated to be alive \non 1/1/2014"
    , color = "#b2d7e9") +
  geom_text(x = 2003, y = 40, label = "The median \nliving Joseph \nis 37 years old"
    , color = "darkgray") +
  geom_curve(x = 1995, xend = 1930, y = 40, yend = 24
    , arrow = arrow(length = unit(0.3,"cm")), curvature = 0.5) +
  ylim(0,42)
name_plotx
```

## Age Distribution of American Boys Named Joseph



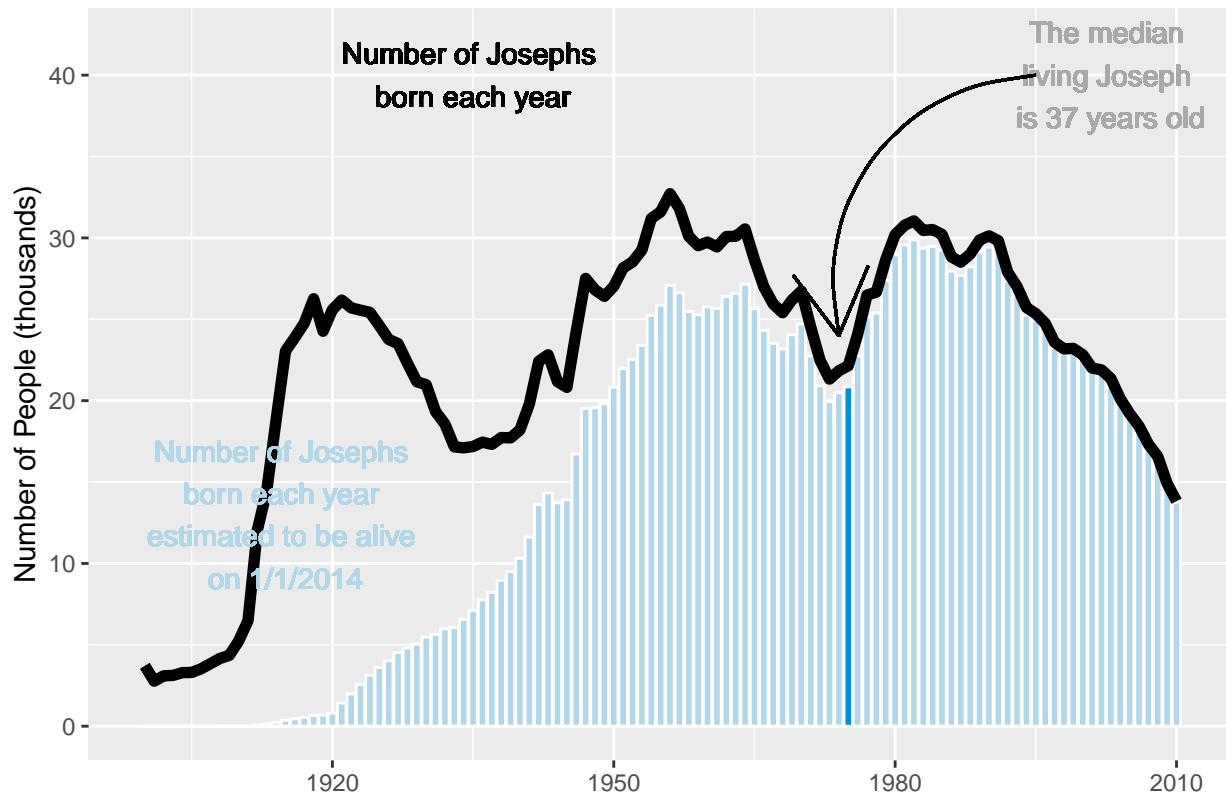
```
#Curvature Adjustment: curvature changed to 1 to change the angle of the curve.
name_ploty <- name_plot +
  geom_bar(stat = "identity", color = "white", fill = "#008fd5"
    , aes(y = ifelse(year == median_yob, est_alive_today/1000, 0))) +
  ggtitle("Age Distribution of American Boys Named Joseph") +
  geom_text(x = 1935, y = 40, label = "Number of Josephs \nborn each year") +
  geom_text(x = 1915, y = 13
    , label = "Number of Josephs \nborn each year \nestimated to be alive \non 1/1/2014"
    , color = "#b2d7e9") +
  geom_text(x = 2003, y = 40, label = "The median \nliving Joseph \nis 37 years old"
    , color = "darkgray") +
  geom_curve(x = 1995, xend = 1974, y = 40, yend = 24
    , arrow = arrow(length = unit(0.3,"cm")), curvature = 1) +
  ylim(0,42)
name_ploty
```

## Age Distribution of American Boys Named Joseph



```
#Arrow Adjustment: arrow length changed to 1 cm.
name_plotz <- name_plot +
  geom_bar(stat = "identity", color = "white", fill = "#008fd5"
    , aes(y = ifelse(year == median_yob, est_alive_today/1000, 0))) +
  ggtitle("Age Distribution of American Boys Named Joseph") +
  geom_text(x = 1935, y = 40, label = "Number of Josephs \nborn each year") +
  geom_text(x = 1915, y = 13
    , label = "Number of Josephs \nborn each year \nestimated to be alive \non 1/1/2014"
    , color = "#b2d7e9") +
  geom_text(x = 2003, y = 40, label = "The median \nliving Joseph \nis 37 years old"
    , color = "darkgray") +
  geom_curve(x = 1995, xend = 1974, y = 40, yend = 24
    , arrow = arrow(length = unit(1,"cm")), curvature = 0.5) +
  ylim(0,42)
name_plotz
```

## Age Distribution of American Boys Named Joseph



Once you're satisfied with your explanation:

- set the r code chunk option “include=FALSE” for the code chunk above (for the original code); we don't want to see all the original code in your answer and that option will suppress that code and output from being printed (but will still evaluate the code)
- delete the text and section header for the “# Original Code” section
- knit the document to pdf or html
- go to the “R Questions and Answers” forum on Moodle, and click on the discussion topic “Problem Set 1A Questions”. Click “Reply”, indicate the question you're answering, and attach your knit file with the response