Q3
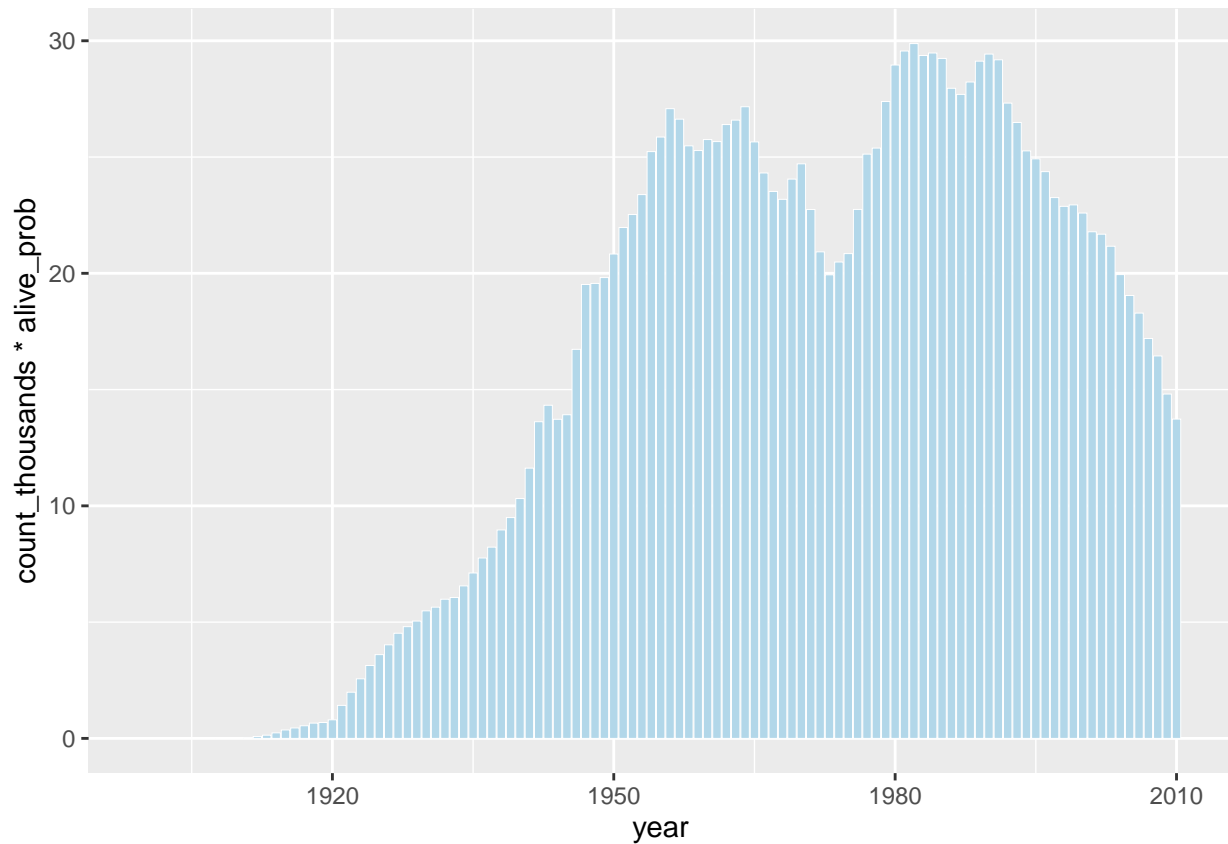
R Markdown

```
library(babynames)
```

Including Plots

```
BabynamesDist <- make_babynames_dist()

joseph <- BabynamesDist %>%
  filter(name == "Joseph" & sex == "M")
name_plot <- ggplot(data = joseph, aes(x = year)) +
  geom_bar(stat = "identity", aes(y = count_thousands*alive_prob)
          , fill = "#b2d7e9", color = "white", size = 0.1)
print(name_plot)
```
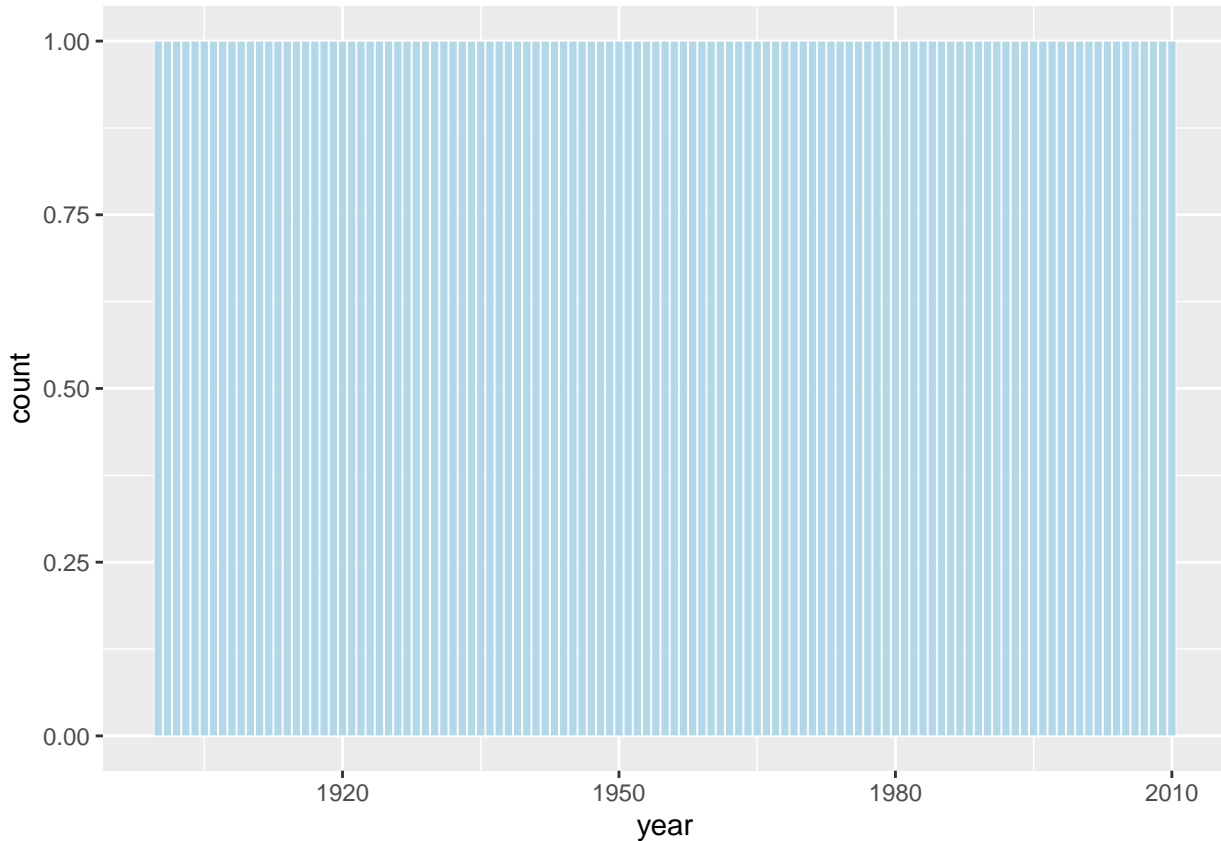

Question 1: What does stat = "identity" do?

```
BabynamesDist <- make_babynames_dist()
```

```
joseph <- BabynamesDist %>%
  filter(name == "Joseph" & sex == "M")
name_plot <- ggplot(data = joseph, aes(x = year)) +
  geom_bar(fill = "#b2d7e9", color = "white", size = 0.1)
print(name_plot)
```
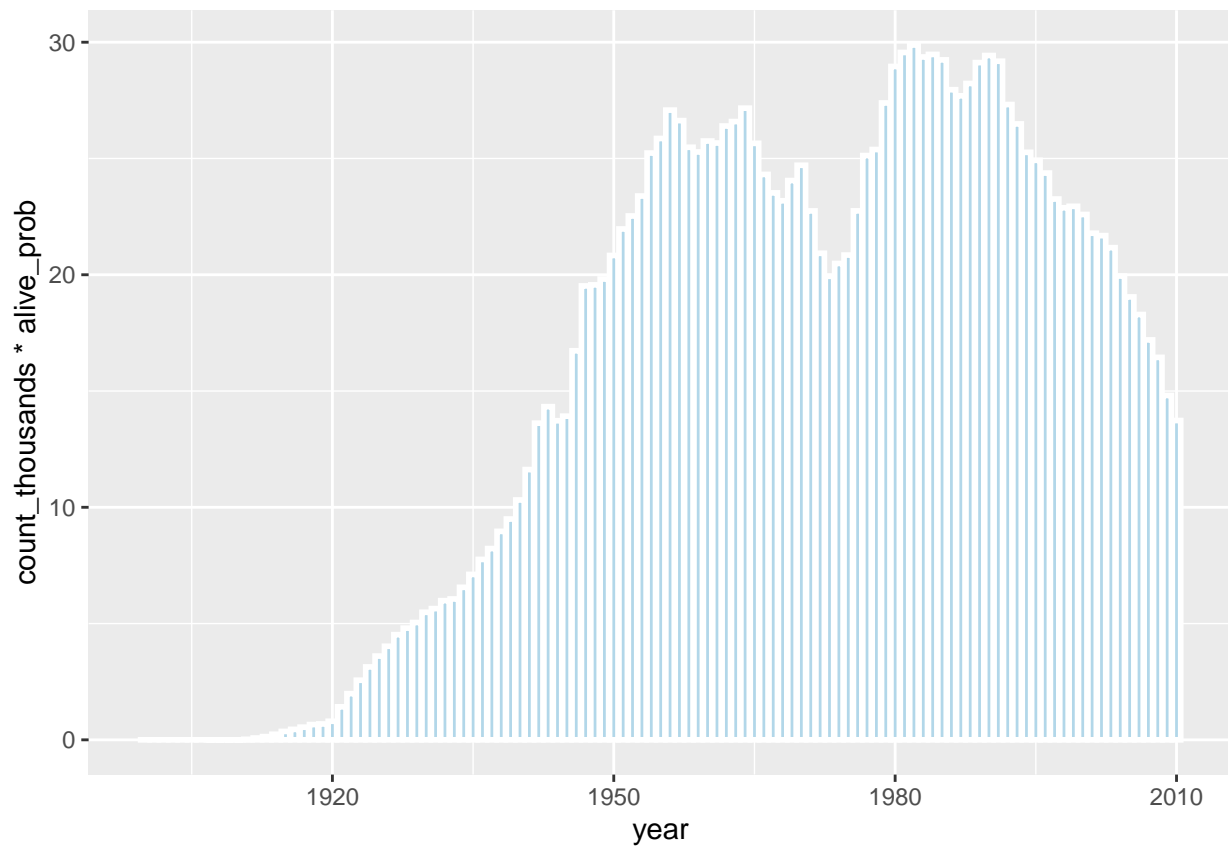

What stat = "identity" does: By default, geom_bar takes in one variable and uses stat = "count" on one variable. However, if you use stat = "identity", it allows you to specify the second variable to plot against. Thus, geom_bar can now take in two variables.

Question 2: What does size do?

```
BabynamesDist <- make_babynames_dist()

joseph <- BabynamesDist %>%
  filter(name == "Joseph" & sex == "M")
name_plot <- ggplot(data = joseph, aes(x = year)) +
  geom_bar(stat = "identity", aes(y = count_thousands*alive_prob)
           , fill = "#b2d7e9", color = "white", size = 1)
print(name_plot)
```

Size in geom_bar determines the thickness of the outline of each bar. In order for it to do anything, the color of the outline must be specified, in this case color = "white". Otherwise, nothing happens.