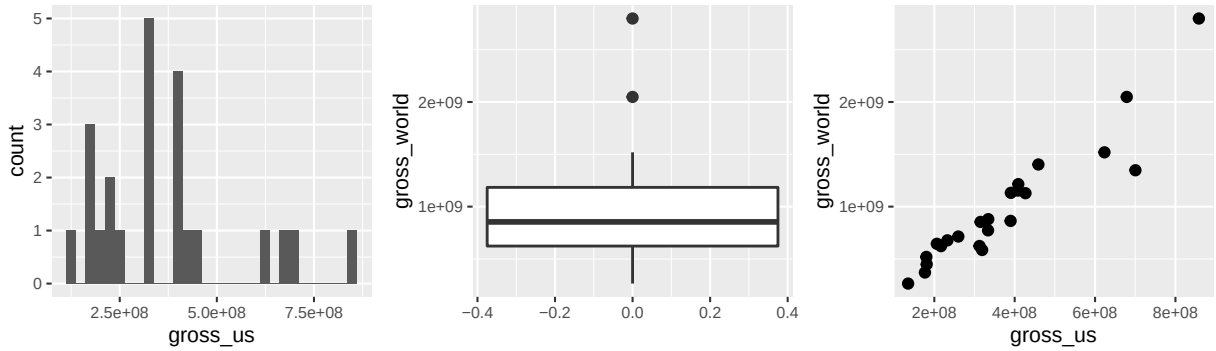
Lab 2B: Grammar of Graphics

Andrew Bray

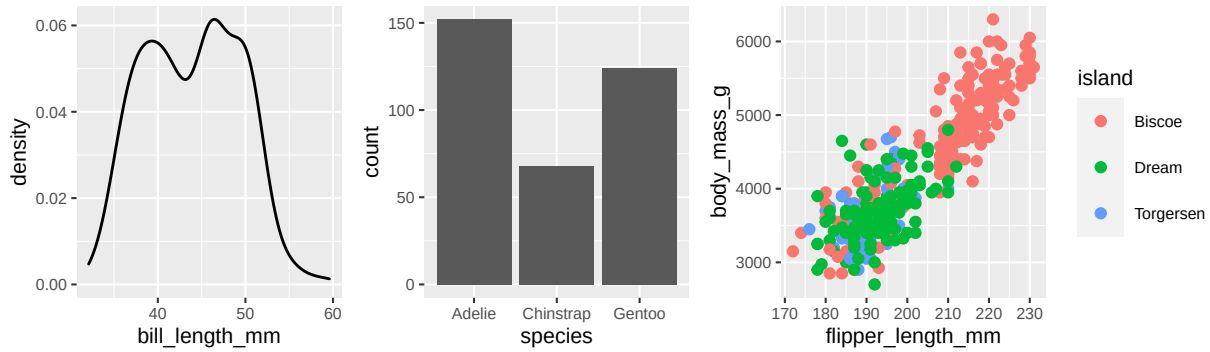
9/1/2021

```
ggplot2::theme_set(ggplot2::theme_gray(base_size = 8))
```

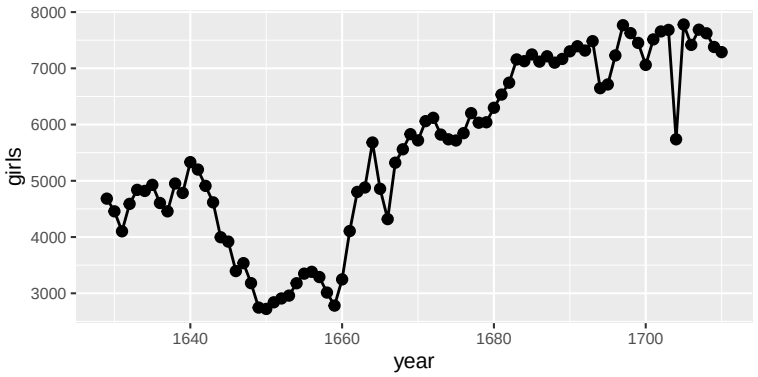
1. The following three plots come from a data set called `mcu_films` inside the `openintro` package. Please write out the `ggplot2` code that will produce each one.



2. The following three plots come from a data set called `penguins` inside the `palmerpenguins` package. Please write out the `ggplot2` code that will produce each one.



3. Improve the scatter plot in question 2 by doing two things.
   a. Capitalize the labels on the axes and the color legend, replace underscores with spaces, and add an appropriate title at the top.
   b. Add transparency to the points so that overplotting is less of an issue.
4. Recreate this modified version of a line plot from Lab 2A using the `arbuthnot` data frame from the `stat20data` package. Note that it is possible to apply more than one geometry to a ggplot.



5*. Fivethirtyeight.com is a news website that focuses on data journalism: nearly all of their stories are inspired by a data set. The editors of the website make that data available online for all to explore, and several fine members of R community put it into an R package so that we can have easy access to it. You can view the names of the data sets by typing the following into the console.

```
data(package = "fivethirtyeight")
```

You can learn more about the listed data sets by pulling up with helpfile by running a command like `?bob_ross` at the console. If that seems a data set of interest to you, you can load it into R and view it with:

```
data("bob_ross")
View(bob_ross)
```

Select a data set (not `bob_ross`) and produce two plots. There are countless plots that you *could* generate, but aim to make plots that reveal interesting structure in the data and are thoughtfully constructed. Be sure to add useful axis labels and a title plot. Provide a sentence or two for each plot describing what it reveals about the subject matter. Hint: usually data sets with a mix of categorical and numerical data are easier to visualize.

Only after you have produced these two plots, go ahead and read the story that this data set inspired. Those can be found by clicking into the appropriate folder at this website and following the link.

Do the two plots that you created (or near likenesses), appear in the new story?