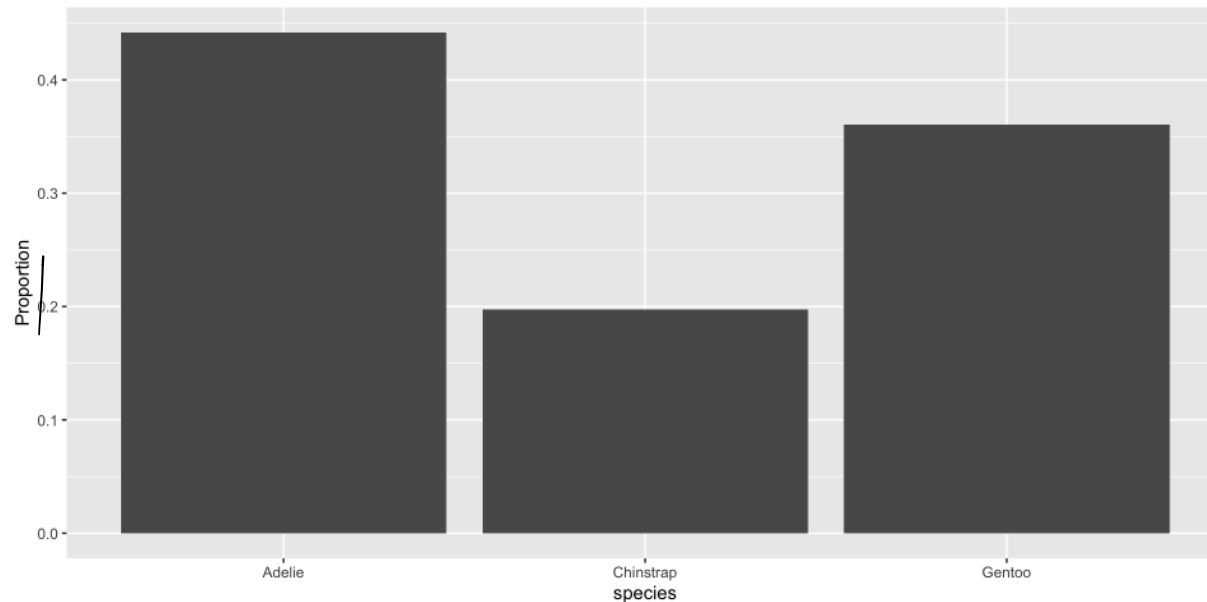


# Bar charts with proportions

- `after_stat()` indicates the aesthetic mapping is performed after statistical transformation
- Use `after_stat(count)` to access the `stat_count()` called by `geom_bar()`

```
1 penguins |>
2   ggplot(aes(x = species)) +
3   geom_bar(aes(y = after_stat(count) / sum(after_stat(count)))) +
4   labs(y = "Proportion")
```

*Handwritten note: `nrow(penguins)` is written above the denominator `sum(after_stat(count))` in the code, indicating that the denominator represents the total number of rows in the data frame.*



# Compute and display the proportions directly

- Use `group_by()`, `summarize()`, and `mutate()` in a pipeline to compute then display the proportions directly
- Need to indicate we are displaying the `y` axis as given, i.e., the identity function

```
1 penguins |>
2   group_by(species) |>
3   summarize(count = n(), .groups = "drop") |>
4   mutate(total = sum(count),
5           prop = count / total) |>
6   ggplot(aes(x = species)) +
7   geom_bar(aes(y = prop), stat = "identity")
```

# Useful to order categories by frequency with forcats

```
1 penguins |>
2   group_by(species) |>
3   summarize(count = n(), .groups = "drop") |>
4   mutate(total = sum(count),
5           prop = count / total,
6           se = sqrt(prop * (1 - prop) / total),
7           lower = prop - 2 * se,
8           upper = prop + 2 * se,
9           species = fct_reorder(species, prop)) |>
10  ggplot(aes(x = species)) +
11  geom_bar(aes(y = prop), stat = "identity") +
12  geom_errorbar(aes(ymin = lower, ymax = upper),
13              color = "red")
```

# Useful to order categories by frequency with forcats

