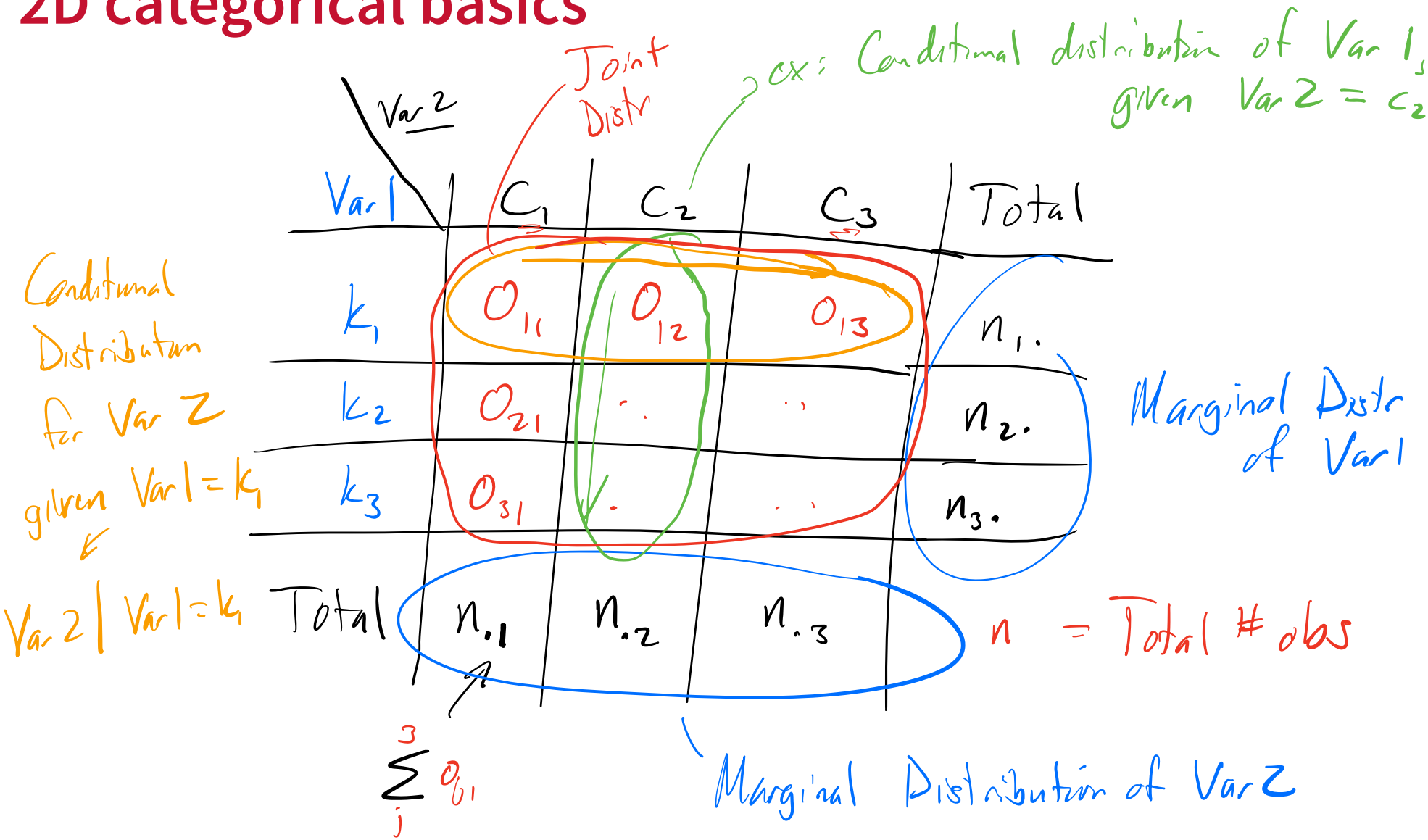
2D categorical basics



2D categorical basics

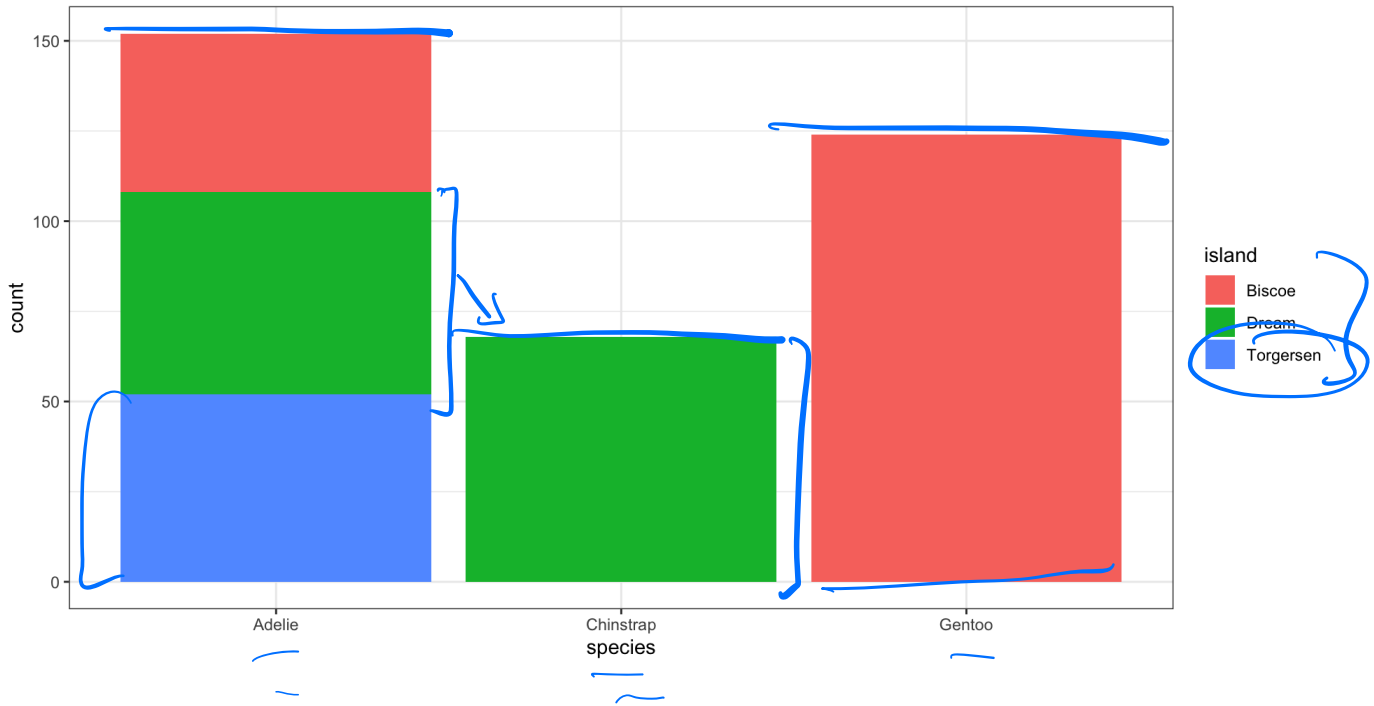
```
1 addmargins(table("Species" = penguins$species, "Island" = penguins$island))
```

```
           Island
Species     Biscoe Dream Torgersen Sum
  Adelie        44    56        52 152
  Chinstrap      0    68         0  68
  Gentoo       124     0         0 124
  Sum          168   124        52 344
```

- Column and row sums: marginal distributions
- Values within rows: conditional distribution for `Island` given `Species`
- Values within columns: conditional distribution for `Species` given `Island`
- Bottom right: total number of observations

Stacked bar charts - a bar chart of spine charts

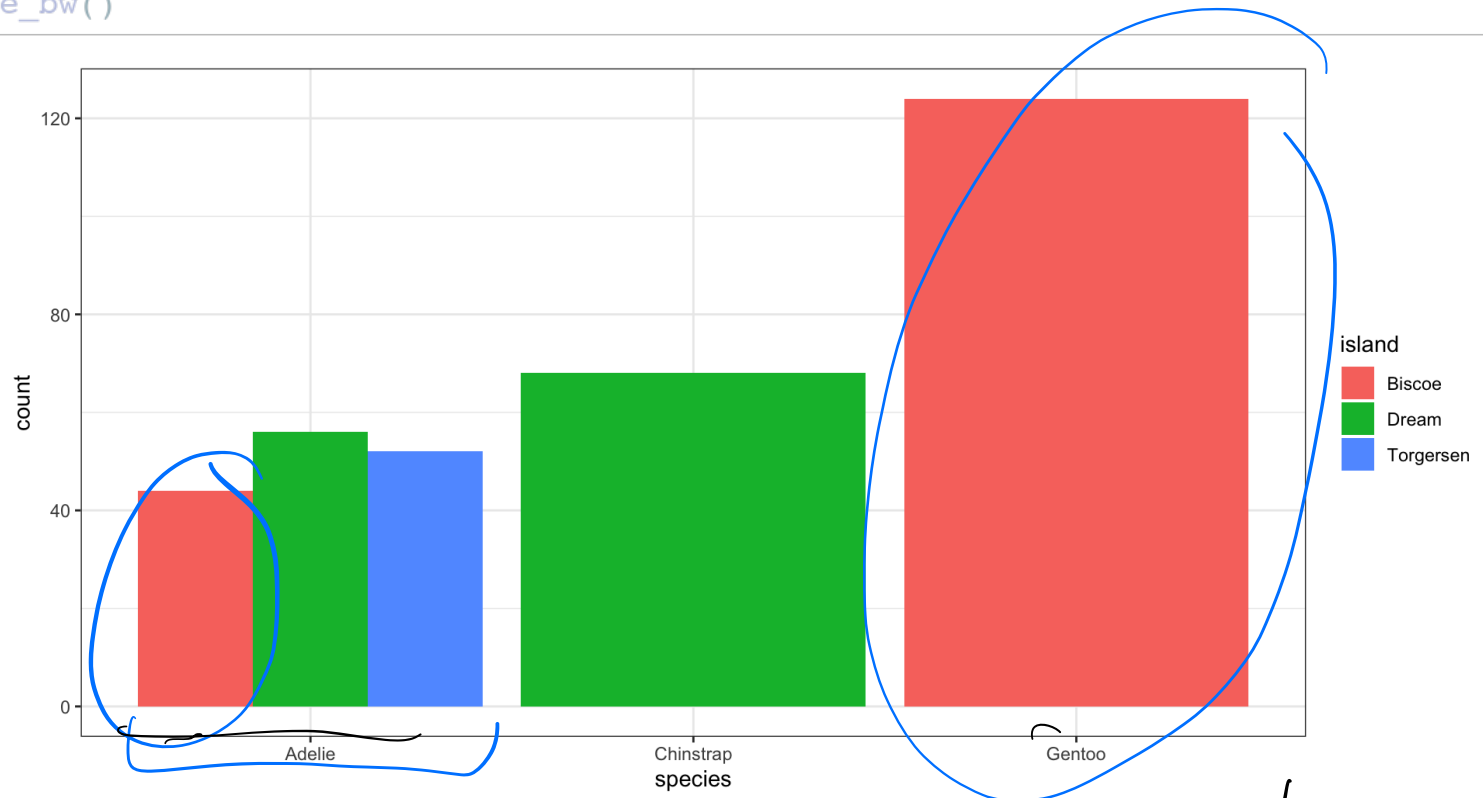
```
penguins |>
  ggplot(aes(x = species, fill = island)) +
  geom_bar() +
  theme_bw()
```



- Easy to see marginal of `species`, i.e., $P($ `x` $)$
- Can see conditional of `island` | `species`, i.e., $P($ `fill` $|$ `x` $)$
- Harder to see conditional of `species` | `island`, i.e., $P($ `x` $|$ `fill` $)$

Side-by-side bar charts

```
penguins |>
  ggplot(aes(x = species, fill = island)) +
  geom_bar(position = "dodge") +
  theme_bw()
```



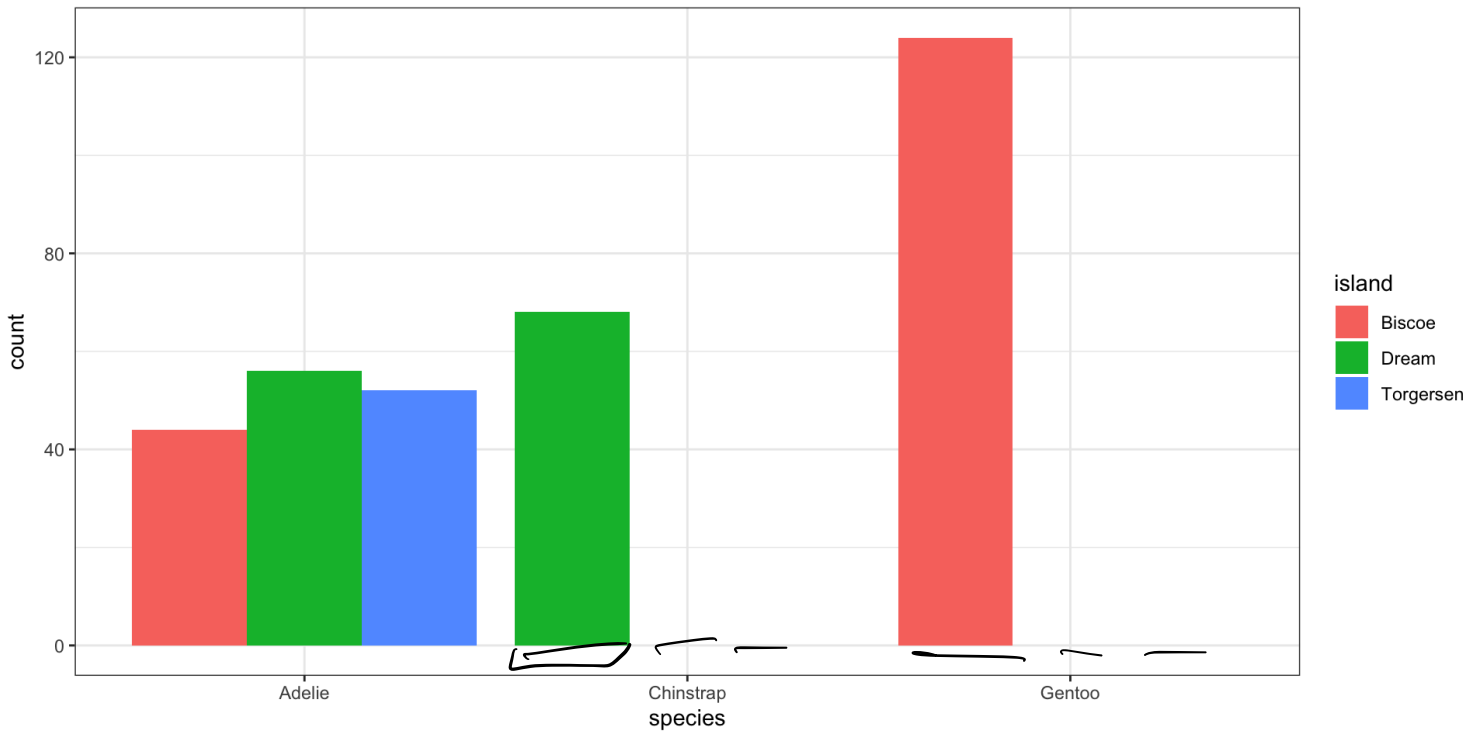
- Easy to see conditional of `island` | `species`, i.e., $P($ `fill` $|$ `x` $)$
- Can see conditional of `species` | `island`, i.e., $P($ `x` $|$ `fill` $)$

We can also see: joint

Hard to see: marginal

Side-by-side bar charts

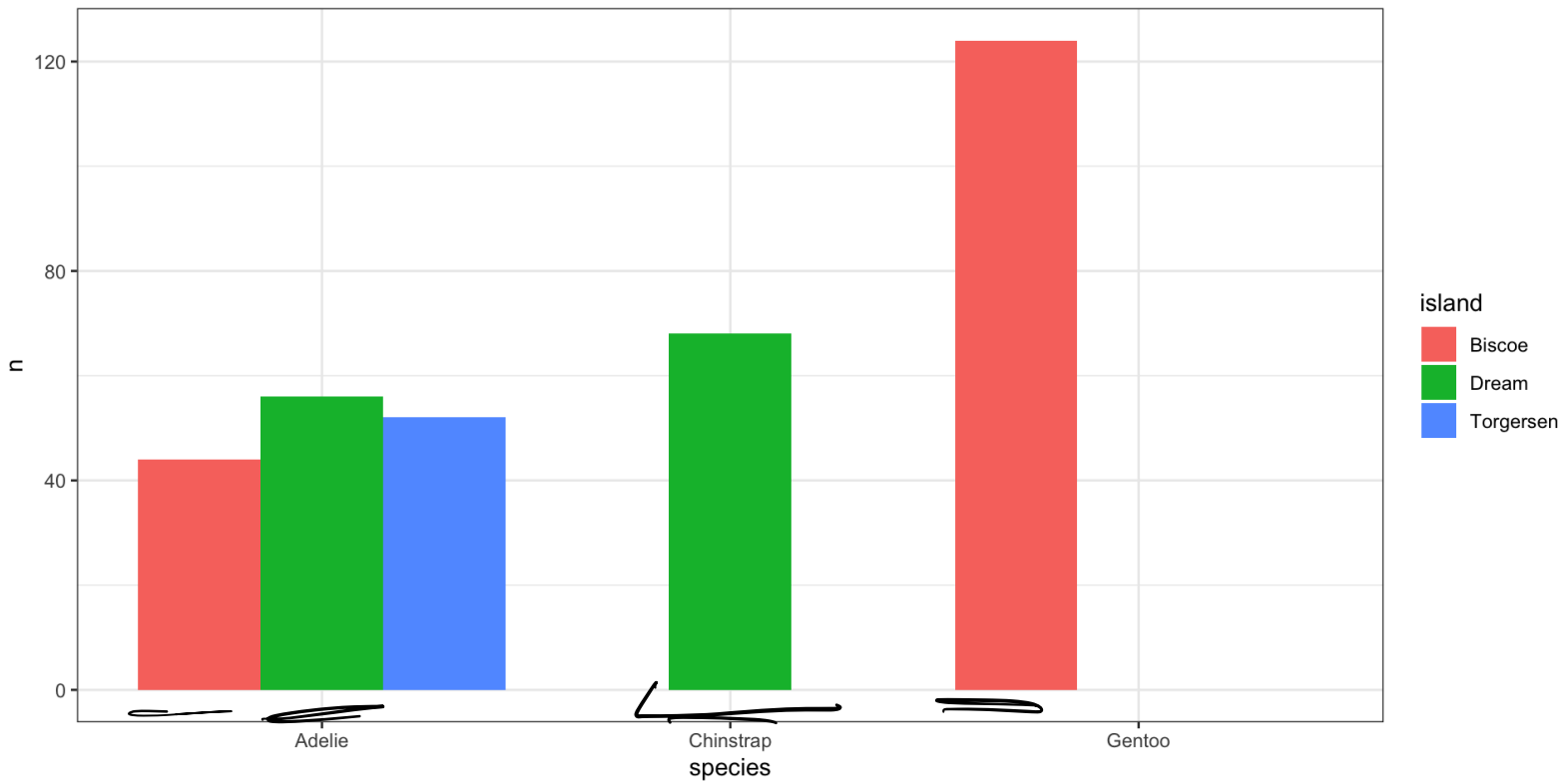
```
penguins |>
  ggplot(aes(x = species, fill = island)) +
  geom_bar(position = position_dodge(preserve = "single")) +
  theme_bw()
```



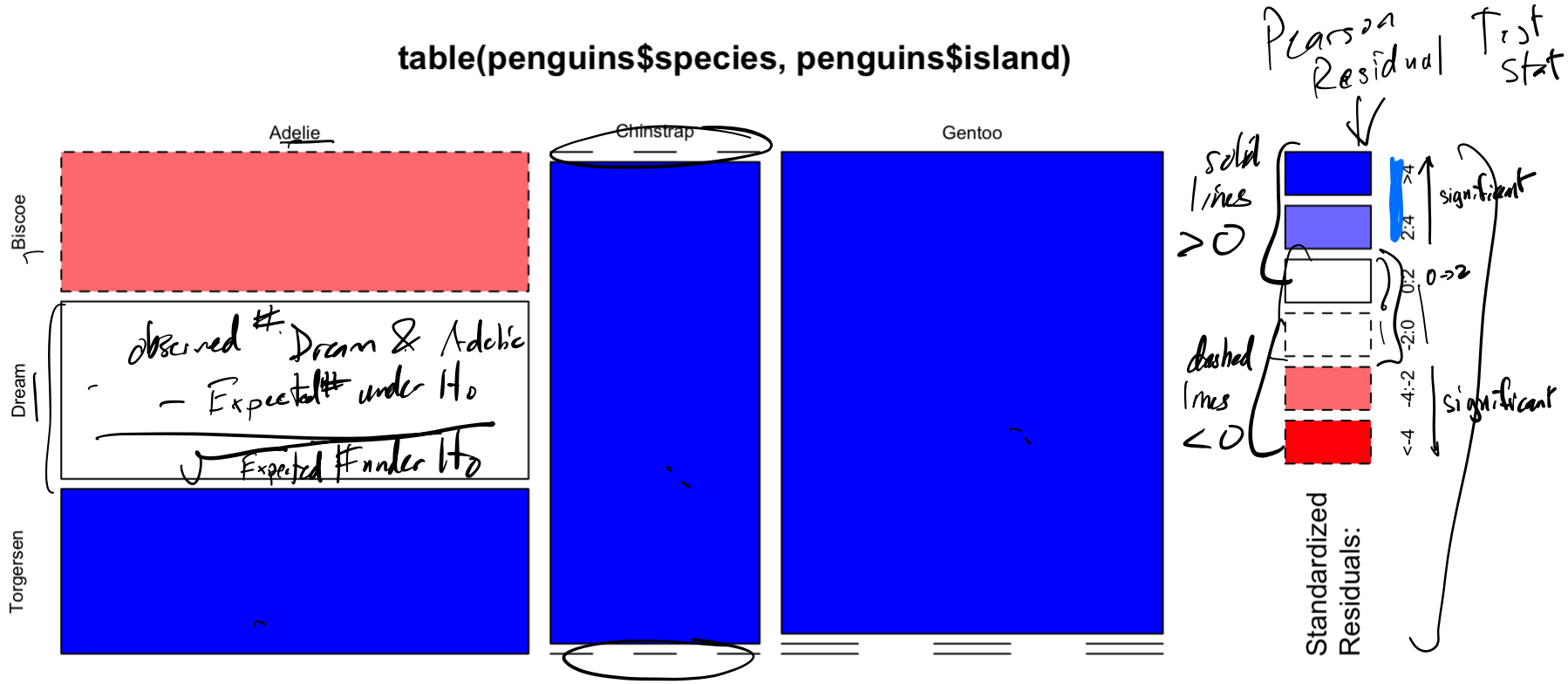
- Easy to see conditional of `island` | `species`, i.e., $P($ `fill` $|$ `x` $)$
- Can see conditional of `species` | `island`, i.e., $P($ `x` $|$ `fill` $)$

Complete missing values to preserve location

```
penguins |>
  count(species, island) |>
  complete(species = unique(species), island = unique(island),
           fill = list(n = 0)) |>
  ggplot(aes(x = species, y = n, fill = island)) +
  geom_bar(stat = "identity", position = "dodge") +
  theme_bw()
```



```r
mosaicplot(table(penguins$species, penguins$island), shade = TRUE)
```



```
mosaicplot(table(penguins$island, penguins$sex), shade = TRUE,
           main = "Distribution of penguins' sex does not vary across islands")
```