Lizard perch heights

An artificial data set

---

28 marks

See the lizards (artificial) background file for context.

The data are available as `lizards` from the `loon.data` package:

```
data("lizards", package = "loon.data")
```

Note again that these data are **not real**.

a. **(7 marks)** On `sex` and `species`.

   i. *(2 marks)* Produce a histogram of the sex counts faceted by species.

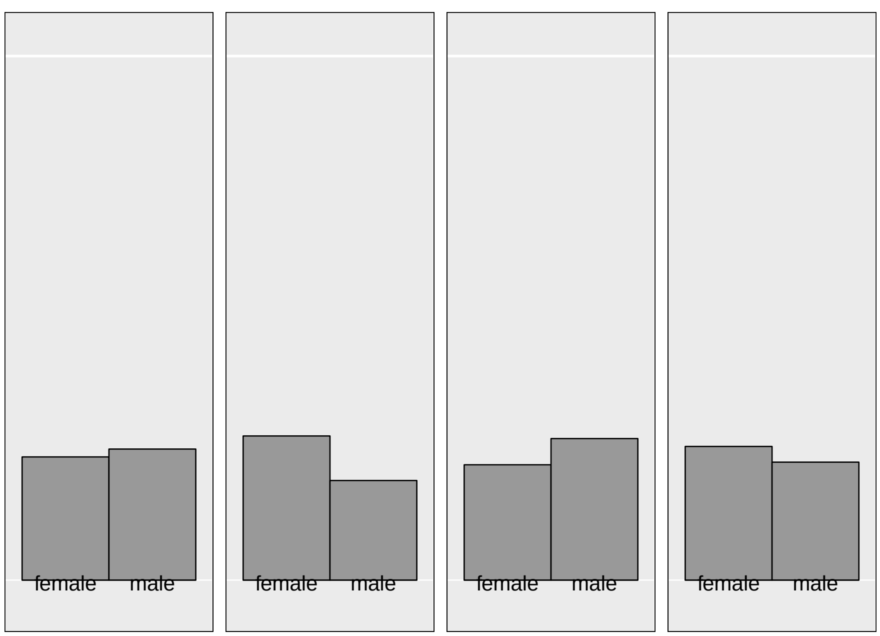
      The resulting faceted display should

      - have as many facets as there are species
      - show the corresponding histogram of sex in each facet
      - have linking group `"lizards"`
      - be assigned to the variable `h_sex_by_species`

      Show your code.

      Show the resulting faceted plot.

```
p <- with(lizards,
          l_hist(lizards[ , "sex"], xlabel = "sex", title = "sex", linkingGroup = "lizards"))
h_sex_by_species <- l_facet(p, by = lizards$species, linkingGroup = "lizards")
plot(h_sex_by_species)
```



i. *(2 marks)* What differences, if any, do you see in the sex distributions?

From the faceted display from part i, we can see that for species A, the two bars of male and female are almost at the same level. This indicates that the sex and species for species A are almost evenly distributed. However, for species B, C, and D, each of the two bars are clearly not at the same level, this suggest that the sex is not evenly distributed.

ii. *(3 marks)* Formally test the hypothesis that `sex` and `species` are independently distributed. What conclusions do you draw?

Show your code and its results.

```r
table_sex_species <- paid_type <- with(lizards, table(sex, species))
table_sex_species
```

```
##         species
## sex       A  B  C  D
##   female 47 55 44 51
##   male   50 38 54 45
```

```r
chisq.test(table_sex_species)
```

```
##
##  Pearson's Chi-squared test
##
## data:  table_sex_species
## X-squared = 4.3382, df = 3, p-value = 0.2272
```

- A p-value of 0.2272 is pretty large, suggesting that, by this test, there is little to no evidence against the hypothesis of independence.

b. **(10 marks)** On perch preferences.

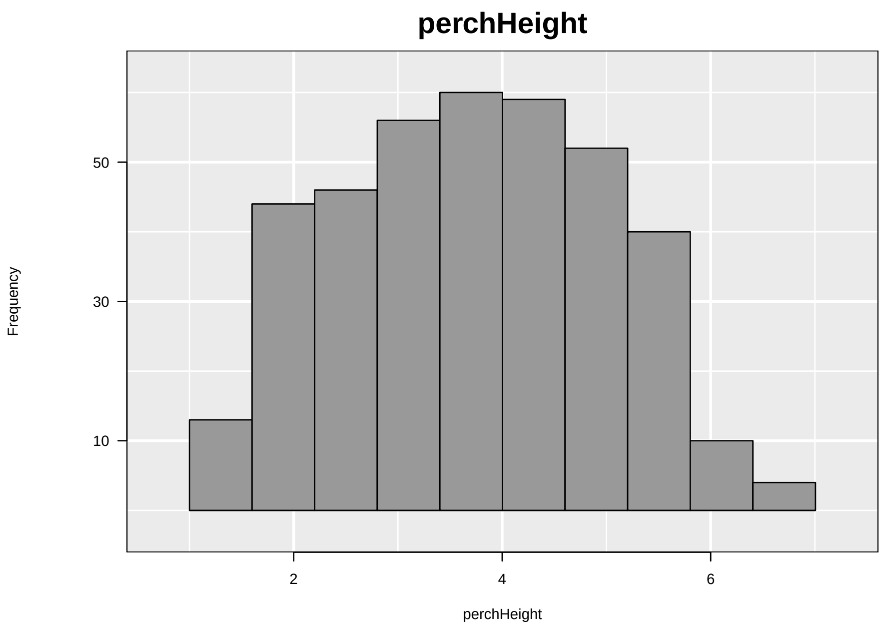
   i. *(2 marks)* Construct an interactive histogram of the `perchHeight`.

   Show your code and the resulting plot.

   The histogram must

   - have `linkingGroup = "lizards"`
   - be assigned to the variable `h_height`

```r
h_height <- with(lizards,
          l_hist(lizards[ , "perchHeight"], xlabel = "perchHeight",showScales = TRUE, title = "perchHei
plot(h_height)
```



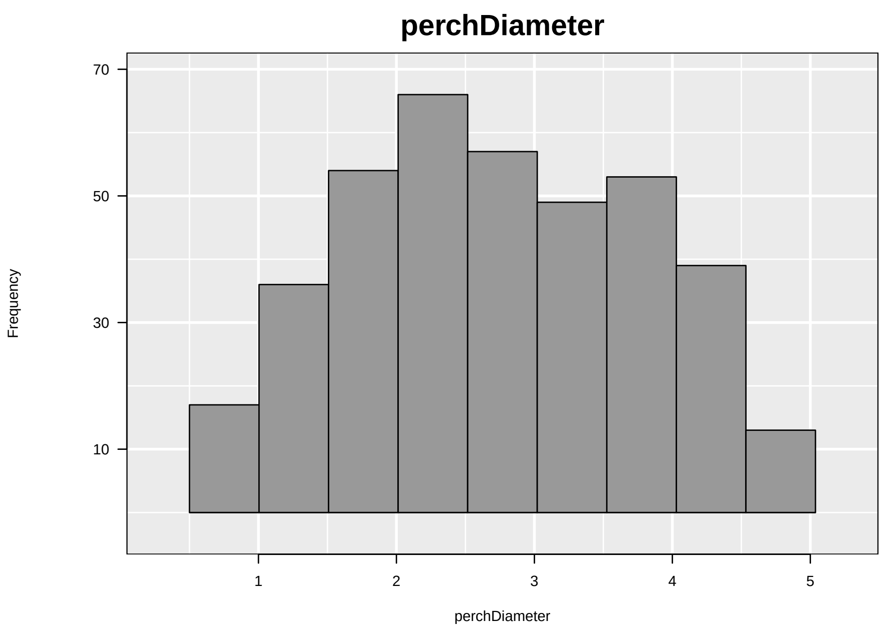
i. *(2 marks)* Construct an interactive histogram of the `perchDiameter`.

Show your code and the resulting plot.

The histogram must

- have `linkingGroup = "lizards"`
- be assigned to the variable `h_diameter`

```
h_diameter <- with(lizards,
          l_hist(lizards[ , "perchDiameter"], xlabel = "perchDiameter",showScales = TRUE, title = "perc
plot(h_diameter)
```



i. *(2 marks)* Brush the `perchHeight` histogram `h_height` and observe the effect on the `perchDiameter` histogram `h_diameter`.

Describe the relationship you observe, if any, between `perchHeight` and `perchDiameter`.

Does this make sense?

I observed that when when we brush the left of the perchHeight histogram, the perchDiameter mostly appears to be the larger diameters. This means that when the perchHeight is small, the perchDiameter is likely to be big. When we brush towards the high perchHeight, the perchDiameter seems to decrease. This indicates that when the perchHeight increases, we are more likely to observe a smaller perchDiameter. This makes senses because in natural, it is more likely to have greater diameter when perch height is low.

ii. *(1 mark)* Brush the `perchHeight` histogram `h_height` and observe the effect on the faceted display of `sex` by `species`.

What does this suggest?

When we brush the perchHeight histogram, we observed that the species that appears in the lowest perchHeight is A. When we brush towards the higher perchHeight, the species appears the most is B. And when we brush further to the right of the histogram, the majority species is C. Last, the species that appears in the highest perchHeight is D. This suggests us that there are relationship between the perchHeight and the species.

iii. *(2 marks)* Facet the `perchHeight` histogram `h_height` by `sex`.

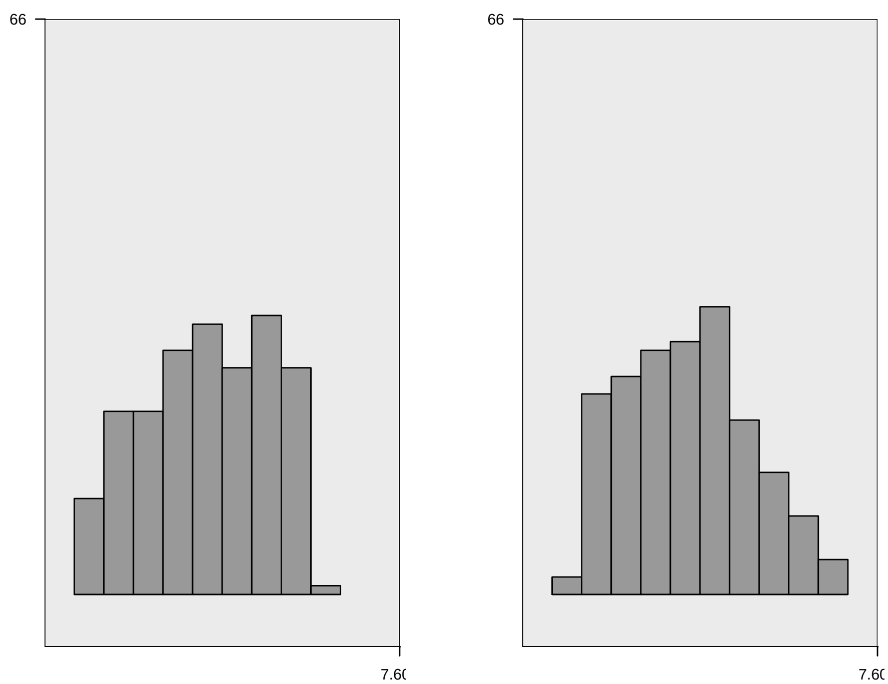
The faceting should

- created with `layout = "wrap"`
- be by `sex`
- be arranged in a single column
- assigned to `h_height_by_sex`

Show your code.

Show the resulting plot.

```
h_height_by_sex <- l_facet(h_height,by = lizards$sex, layout = "wrap")
plot(h_height_by_sex)
```



i. *(1 mark)* Which sex appears most likely to be found at the highest perch heights?

It appears that the sex that is most likely to be found ast the highest perch heights is male.

ii. **(11 marks)** On the relationship between perch diameter and perch height.

   i. *(3 marks)* Fit a straight line model `perchDiameter ~ perchHeight` to the lizards data.

      Write down the equation of the fitted line.

      Does this make sense? Explain.

      Do **not** show your code.

      perchDiameter = 4.637412 - 0.510926*perchHeight

      Yes, I think that this equation makes sense because it is reasonable that the perchDiameter is 4.63 when the height is 0. Also, the slope is negative, this means when the perchHeight increases, the perch diameter will decrease. This results is same as what we observed previously, therefore the equation makes sense.

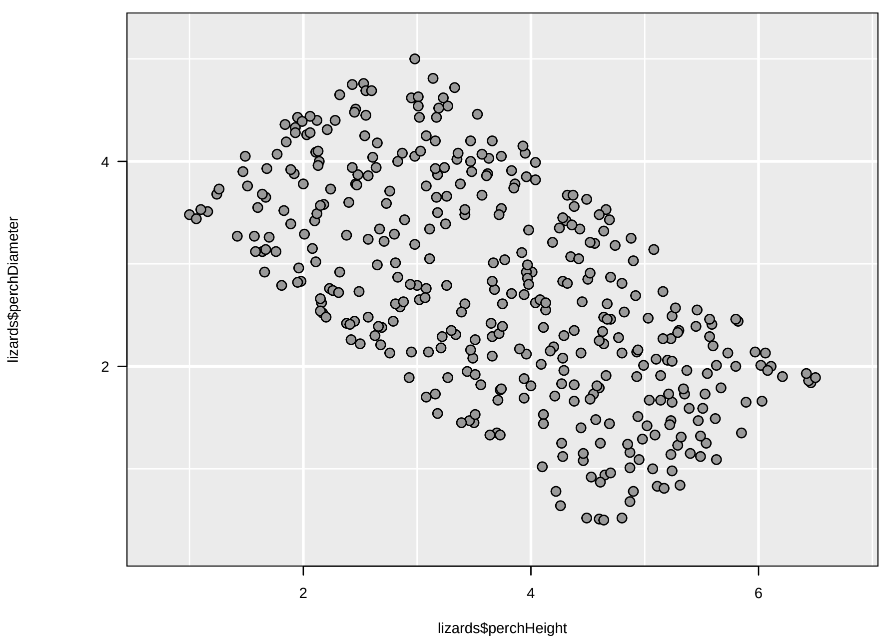
   ii. *(2 marks)* Draw an interactive scatterplot with `x = perchHeight` and `y = perchDiameter`.

      The plot should

      - have linking group "'lizards"
      - be assigned to the variable `p_perch`

      Show your code and the resulting plot.

```
p_perch <- with(lizards,
                l_plot(lizards$perchHeight, lizards$perchDiameter,showScales = TRUE,
                       linkingGroup = "lizards"))
plot(p_perch)
```



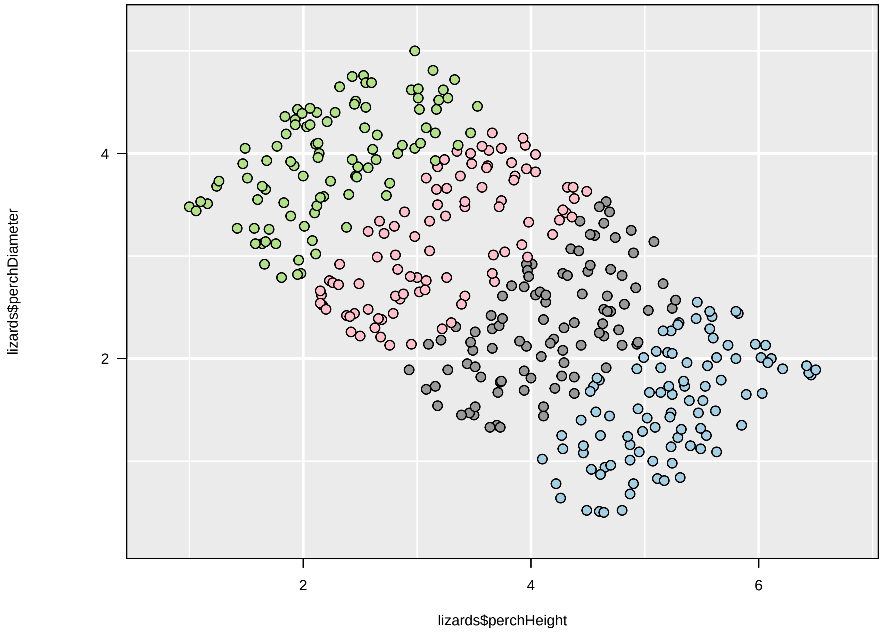
i. *(1 mark)* On the basis of the above plot **alone**, write a one sentence summary of the relationship between the perch height and the perch diameter, at least as preferred by these (fictitious) lizards.

   According to the plot above, we can see that the lizards preferred the habitat equally (because the dots looks evenly distributed) when perchHeight is low/perchDiameter is high, medium perch-Height/medium perchDiameter, and perchHeight is high/perchDiameter is low.

ii. *(1 mark)* Colour the `p_perch` scatterplot by `species`.

Show your code and the resulting plot.

```
p_perch['color'] <- lizards$species
plot(p_perch)
```



i. *(2 marks)* On the basis of the above coloured `p_perch` scatterplot, what do you conclude about the perch preferences of these (fictitious) lizards?

   From the coloured scatterplot, it is obvious that there is no overlap between different species. Each species is found in a different combination of perchHeight and perchDiameter, this means that the preferences of the habitat is different in every species.

ii. *(2 marks)* By interacting with the plots as you see fit, investigate the relationship between perch diameter and height **only** for female lizards from species "D".

   Describe how these lizards prefer their perch as well as the relationship between perch diameter and height for these lizards.

   The female lizards from species "D" prefered the highest perch height among all other three species female lizards. And they prefered the lowest perch diameter among all other three species female lizards.