Differences in petal width between three Iris species.
Yaniv Brandvain. Applied Biostats. Project 1. Example

Introduction

Plant species can vary tremendously or can be quite similar. Edgar Anderson was interested to know if different *Iris* species could be clearly defined by morphological differences in flower size – focussing on four traits: Petal Length, Petal Width, Sepal Length and Sepal Width. Here we are interested in knowing the extent to which species and petal length predict petal width among three *Iris* species.

Details of the sampling scheme are unclear, so I cannot comment on the extent of bias and non-independence. If plants represented random samples from independent populations, the answers below would be most trustworthy and reliable. However, if plant were picked as "exemplars" of the species we may be biased toward seeing more differences on phenotypes (or alternatively, if plants that were the most similar to the other species were chosen, we would be biased towards underestimating the differences between species)

Methods

We summarized observations of one hundred and fifty *Iris* flowers – fifty from each of three species. All data were analyzed in R with the tidyverse packages – ggplot2, dplyr, readr, and janitor. We focus on petal size, as both petal width and petal length, throughout the analyses below.

Results.

Both petal length and width are unimodal within species, and their distribution is roughly symmetric in all species (Figure 1). Petal size differs substantially between species (Figures 1A). *Iris setosa* has by far the narrowest petals. While both *Iris versicolor* and *Iris setosa* have large petals, *Iris setosa*'s petals are consistently the widest. *Iris setosa* has quite narrow petals (mean = 0.246), approximately one fifth as wide as petals of *Iris versicolor* (mean = 1.33). *Iris virginica*'s petals are wider still (mean = 2.03). Relative to its modest mean, *Iris setosa* has quite variable petal widths (sd = 0.105, coefficient of variation = 42.8), as compared to *Iris versicolor* (sd = 0.198, coefficient of variation = 14.9) and *Iris virginica* (sd = 0.275, coefficient of variation = 0.275).

Across the data set as a whole, Petal Length and Petal Width are strongly and positively correlated (r = 0.96), with Petal width increasing by 0.415 cm for each 1 cm increase in Petal length (dotted black line in Figure 3). Some of this relationship is likely attributable to the fact that species differ in both Petal Length and Petal width (Figure 1B). Again, *Iris setosa* has exceptionally short petals (mean = 1.46), which are approximately three times shorter than petals of *Iris versicolor* (mean = 4.26). *Iris virginica*'s petals are longer still (mean = 5.55). After accounting for differences in mean petal length, all species have similar variability in petal length, with standard deviations of 0.174, 0.470, and 0.552, and coefficients of variation of 11.9, 11.0, and 9.94, for *Iris setos*a, *Iris versicolor*, and *Iris virginica*, respectively. In fact, including species in our model further increased the proportion of variance explained by the model (r^2 increases from 0.928 without species in the model to 0.946 with species in the model).

Nonetheless, some of this increase in petal width with petal length is also found within species - the positive correlation within species varies from modest (0.33 and 0.32 in *Iris setosa* and *Iris virginica*, respectively) to quite strong (0.79 in versicolor). Similarly, the slope of this line was modest in *Iris setosa* (b = 0.20) and *Iris virginica* (b = 0.16), and strongest in *Iris versicolor* (b = 0.33) (See colored lines in Figure 2).

Conclusions

On the whole, petal length and width are highly correlated in these three *Iris species* (Figure 2). While some of this relationship is due to differences between species, there are still meaningful correlations between petal length and width within species. Intriguingly, this correlation is strongest in the species with most intermediate values of petal length and width, *Iris versicolor*. As such both species and petal length likely causally contribute to differences in petal width.

Implications for species classification: Additionally, Because observed sepal widths and lengths do not overlap between *Iris setosa* and the other two species (Figure 1), petal size could be used to differentiate this species from the others. Although there is some overlap in petal length and width between Iris virginica and Iris versicolor, this overlap is minimal and it seems likely that most plants could be reliably assigned to one species or the other.

Consideration and caveats: Although the observed differences in species mean petal size are unlikely to arise by sampling error, we cannot rule out a contribution of non-independence and sampling bias to these results without additional information about the sampling process. If species were randomly placed and grown in a "common garden" the differences observed could be attributable to species, however if these flowers represented natural collections from differing locations, we would not be able to rule out an effect of the environment in generating these differences.

Figures

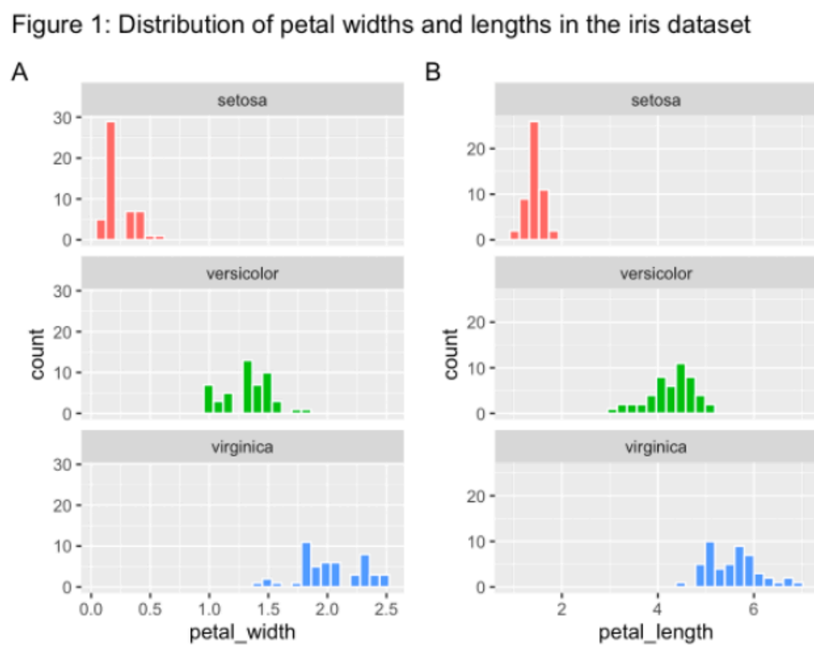

Figure 1) The distribution of petal width (A) and length (B) in fifty flowers of each of three *Iris* species (shown separately by facet, and highlighted in color)

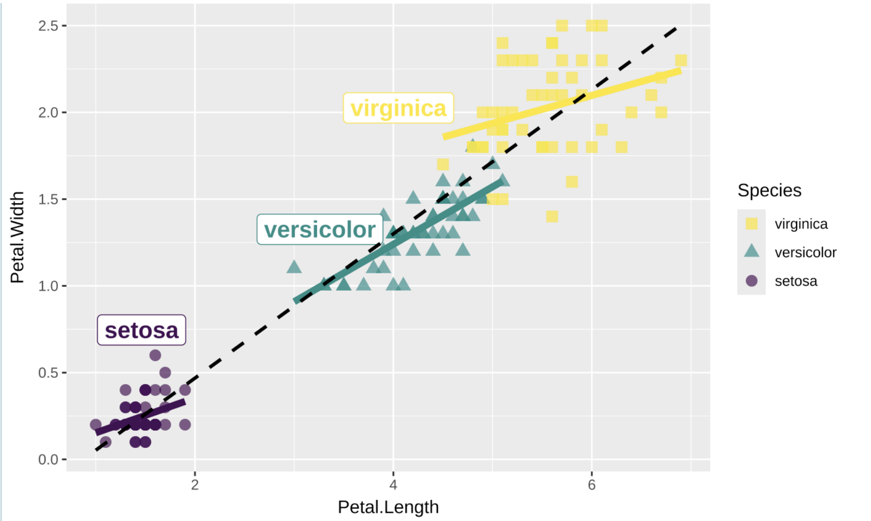

Figure 2) The relationship between petal length and width for fifty flowers in each of three *Iris* species (shown by color). Lines represent estimates of a linear regression. The black dotted line shows the slope in the data set as a whole, while the colored straight lines show the relationship within each species.