

Assignment 1 Question 2

Sheen Thusoo

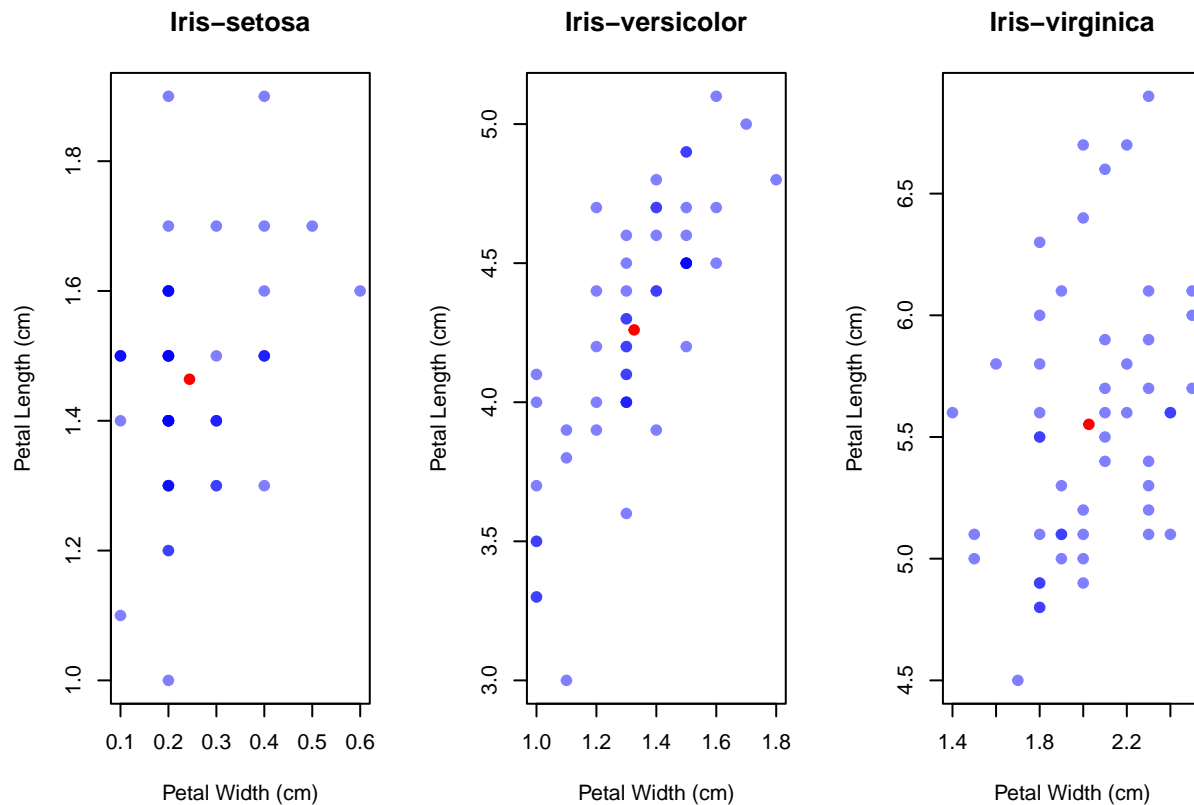
g. Using R, create a figure with 3 panels (1×3 plots). In each panel, generate the scatterplot of the petal length (y-axis) vs. the petal width (x-axis) of one species of iris. The title of each plot should be the iris species of that plot. Add to each plot a red point indicating the mean of the data in that plot. In particular, the x-coordinate and y-coordinate of the red dot should be, respectively, the mean petal length and the mean petal width of the particular iris species in that panel. Comment on any trends you observe in this plot.

```
iris_data <- read.csv("Iris.csv")

par(mfrow=c(1,3))
setosa.PetalWidth <- iris_data[iris_data$Species == "Iris-setosa", "PetalWidth"]
setosa.PetalLength <- iris_data[iris_data$Species == "Iris-setosa", "PetalLength"]
setosa.PetalWidth.mean <- mean(setosa.PetalWidth)
setosa.PetalLength.mean <- mean(setosa.PetalLength)
plot( setosa.PetalWidth, setosa.PetalLength , main="Iris-setosa", xlab="Petal Width (cm)",
      ylab="Petal Length (cm)", pch = 19, col=adjustcolor("blue", alpha = 0.5 ) )
points(x = setosa.PetalWidth.mean, y = setosa.PetalLength.mean, pch = 19, col="red",
      cex = 1)

versicolor.PetalWidth <- iris_data[iris_data$Species == "Iris-versicolor", "PetalWidth"]
versicolor.PetalLength <- iris_data[iris_data$Species == "Iris-versicolor", "PetalLength"]
versicolor.PetalWidth.mean <- mean(versicolor.PetalWidth)
versicolor.PetalLength.mean <- mean(versicolor.PetalLength)
plot( versicolor.PetalWidth, versicolor.PetalLength, main="Iris-versicolor",
      xlab="Petal Width (cm)", ylab="Petal Length (cm)", pch = 19,
      col=adjustcolor("blue", alpha = 0.5 ) )
points(x = versicolor.PetalWidth.mean, y = versicolor.PetalLength.mean, pch = 19, col="red",
      cex = 1)

virginica.PetalWidth <- iris_data[iris_data$Species == "Iris-virginica", "PetalWidth"]
virginica.PetalLength <- iris_data[iris_data$Species == "Iris-virginica", "PetalLength"]
virginica.PetalWidth.mean <- mean(virginica.PetalWidth)
virginica.PetalLength.mean <- mean(virginica.PetalLength)
plot( virginica.PetalWidth, virginica.PetalLength , main="Iris-virginica",
      xlab="Petal Width (cm)", ylab="Petal Length (cm)", pch = 19,
      col=adjustcolor("blue", alpha = 0.5 ) )
points(x = virginica.PetalWidth.mean, y = virginica.PetalLength.mean, pch = 19, col="red",
      cex = 1)
```



Trends Observed: For the iris-setosa plot, there seems to be many data points that are in the same location as many points are a darker shade of blue (indicates overlapping points). The mean (red point) seems to be located towards the center-left of the plot. Except some outliers, the petal width seems to be between 0.1 to 0.4 cm and the petal length seems to be 1.2 to 1.7 cm.

For the iris-versicolor plot, there seems to be an almost linear relationship between the petal width and the petal length.

For the iris-virginica plot, there seems to be less overlapping points then the previous two plots. The mean of this plot is almost at the center of the plot. It has a larger range of petal lengths, from 4.5 cm to 7 cm.