A1_HYDE_20261425, Mari Hyde

2025-01-07

Repository: https://github.com/marihyde/Biol432_A1

The first section of code opens the updated measurements file, which includes calculated limb volumes. head function is used to display a sample of the table.

```
library(dplyr)
```

```
##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##      filter, lag

## The following objects are masked from 'package:base':
##
##      intersect, setdiff, setequal, union
```

```
library(ggplot2)
library(tidyr)

Measurements<-read.csv("Measurements")
head(Measurements)
```

```
##             Organism Limb_width Limb_height Observer Limb_volume
## 1            Beagle       1.14        9.34    Laura    38.13348
## 2 Golden Retriever       1.03        7.04     Paul    23.46373
## 3   German Shepard       1.30        5.72      Jim    30.36915
## 4            Boxer       1.24        5.66      Jim    27.34070
## 5           Poodle       1.39        8.29    Laura    50.31923
## 6           Beagle       1.45        8.12      Jim    53.63421
```

Dplyr is used to order the rows in Measurements by Organism, observer, then limb volume, using the arrange function.

```
Sort_Species<- Measurements %>% arrange(Organism, Observer, Limb_volume)
head(Sort_Species)
```

```
##   Organism Limb_width Limb_height Observer Limb_volume
## 1   Beagle       1.11        6.72      Jim    26.01148
## 2   Beagle       1.19        9.44      Jim    41.99676
## 3   Beagle       1.45        8.12      Jim    53.63421
## 4   Beagle       1.74        8.21      Jim    78.08930
## 5   Beagle       1.94        7.06      Jim    83.47531
## 6   Beagle       1.06        5.39    Laura    19.02613
```

This section groups the data by each species using the group_by function. Then each volume is calculated using the mean function, and printed to display the average limb volume for each species.

```
Average_Volume<- Sort_Species %>% group_by(Organism) %>% summarize(Avg_Volume= mean(Limb_volume, na.rm=
print(Average_Volume)
```

```
## # A tibble: 5 x 2
##   Organism        Avg_Volume
##   <chr>                <dbl>
## 1 Beagle                57.6
## 2 Boxer                 55.3
## 3 German Shepard        56.0
## 4 Golden Retriever      54.2
## 5 Poodle                50.2
```

This chunk sorts the data by observer and species using the group_by function. the Summarise function is counting the number of organisms viewed by each observer for each species, and displaying them as "ObservationCount" in the table.

```
number_observations<- Sort_Species %>% group_by(Observer, Organism) %>% summarise(ObservationCount=n(),
print(number_observations)
```

```
## # A tibble: 15 x 3
##    Observer Organism         ObservationCount
##    <chr>    <chr>                       <int>
##  1 Jim      Beagle                          5
##  2 Jim      Boxer                           9
##  3 Jim      German Shepard                  7
##  4 Jim      Golden Retriever                4
##  5 Jim      Poodle                          5
##  6 Laura    Beagle                          9
##  7 Laura    Boxer                           5
##  8 Laura    German Shepard                  8
##  9 Laura    Golden Retriever                7
## 10 Laura    Poodle                          8
## 11 Paul     Beagle                          6
## 12 Paul     Boxer                           6
## 13 Paul     German Shepard                  5
## 14 Paul     Golden Retriever                9
## 15 Paul     Poodle                          7
```

This chunk displays a boxplot showing the distribution of limb volumes across different species.

```
ggplot(Measurements, aes(x= Organism, y=Limb_volume))+
  geom_boxplot()+
  theme_light()+
  labs(title= "Limb Volume as a Function of Species",
       x="Species",
       y="Limb Volume")
```

Limb Volume as a Function of Species
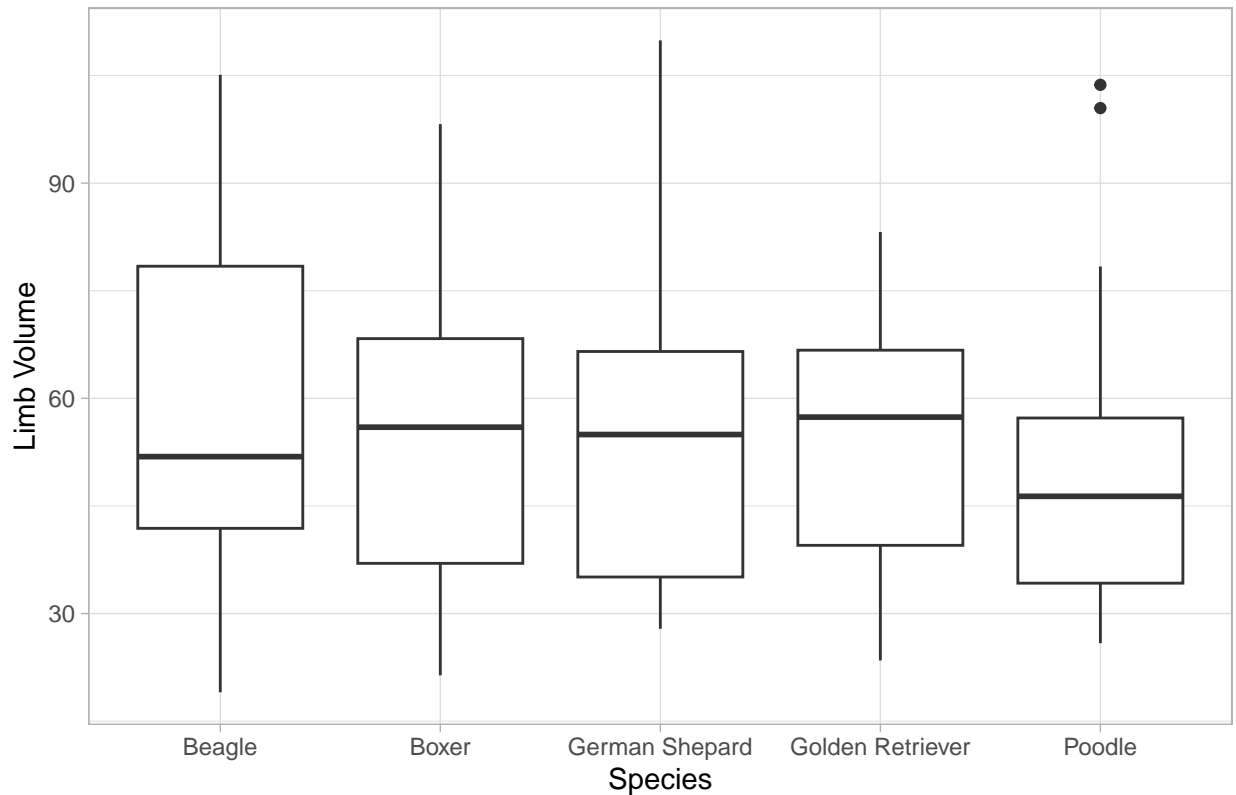
This chunk of code converted the limb volume to elongated measurements, making it easier to see the distribution in the next section.

```
elongated_measurements<- Measurements %>% gather(key="Measurement_Type", value="limbvolume", Limb_volume
head(elongated_measurements)
```

```
##             Organism Limb_width Limb_height Observer Measurement_Type limbvolume
## 1           Beagle          1.14        9.34    Laura      Limb_volume   38.13348
## 2  Golden Retriever        1.03        7.04     Paul      Limb_volume   23.46373
## 3     German Shepard       1.30        5.72      Jim      Limb_volume   30.36915
## 4            Boxer          1.24        5.66      Jim      Limb_volume   27.34070
## 5           Poodle          1.39        8.29    Laura      Limb_volume   50.31923
## 6           Beagle          1.45        8.12      Jim      Limb_volume   53.63421
```

This chunk is displaying the distribution of the elongated limb volume measurements for each species on a histogram. The function facet_wrap splits the data between species, for multiple histograms.

```
ggplot(elongated_measurements, aes(x=limbvolume))+
  geom_histogram(bins=20, fill= "lightpink", color="black", alpha=1)+
  facet_wrap(~Organism)+
  theme_light()+
  labs(title="Distribution of Limb Volume across Dog Species", x="Limb Volume", y="")
```

Distribution of Limb Volume across Dog Species