

Universidad de Comillas
Escuela Técnica Superior de Ingeniería
Master Big Data Technologies and Advanced Analytics

Alumno: Boris Cabrera Barandica (Código 201710808)

Asignatura: Mathematical Foundations of Data Analysis

Práctica 2:

Cargue los datos de AnimalData y responda:

1- ¿Cuántas variables hay en el set de datos?

2- De los primeros 10 animales del dataset, ¿cuántos fueron adoptados?

3- ¿Cuántos días permanece un animal en el centro antes de ser adoptado?

Apoye las respuestas con un gráfico y obtenga tres conclusiones de los datos a partir de los gráficos

Respuestas

1- Hay 24 variables, de 473 entradas.

Código usado:

```
> View(AnimalData)
```

2- hay 4 animales adoptados en los 10 primeros registros

Código:

```
> firsts<-AnimalData[1:10,]
```

```
> firsts<-firsts[,23]
```

```
> y<-(firsts=='Adoption')
```

```
> sum(y)
```

```
[1] 4
```

3- En promedio 20.8 días están en el centro,
código

```
> days<-AnimalData[,24]
```

```
> dim(days)
```

```
[1] 473 1
```

```
> sum(days)/dim(days)[1]
```

```
[1] 20.79915
```

No sé porque no pude hacer funcionar a mean()

Algunas conclusiones:

Es un set de 473 entradas con 24 variables cada una (algunas pueden ser NA), donde se describen animales que están en un refugio, se describen datos que los describen (peso, edad, raza, si es inteligente), y que detallen su ingreso y permanencia en el centro (condiciones de ingreso y salida). Se analizan los datos para encontrar patrones de interés.

Con la función de `summary(AnimalData)` encontramos varios elementos interesantes:

- La edad de ingreso en promedio es de 2.3 años, con mediana de 1, siendo la mayor 17 años, esto nos habla de una población joven de animales.
- El peso de los animales tiene por promedio 10.25 lbs (libras, supongo), con mediana de 20.9 lbs y 3er cuartil de 35.50 lbs, con máximo de 131 lbs; claramente la mayoría son animales ligeros, menos de 30 lbs y unos pocos muy pesados que modifican mucho el promedio.

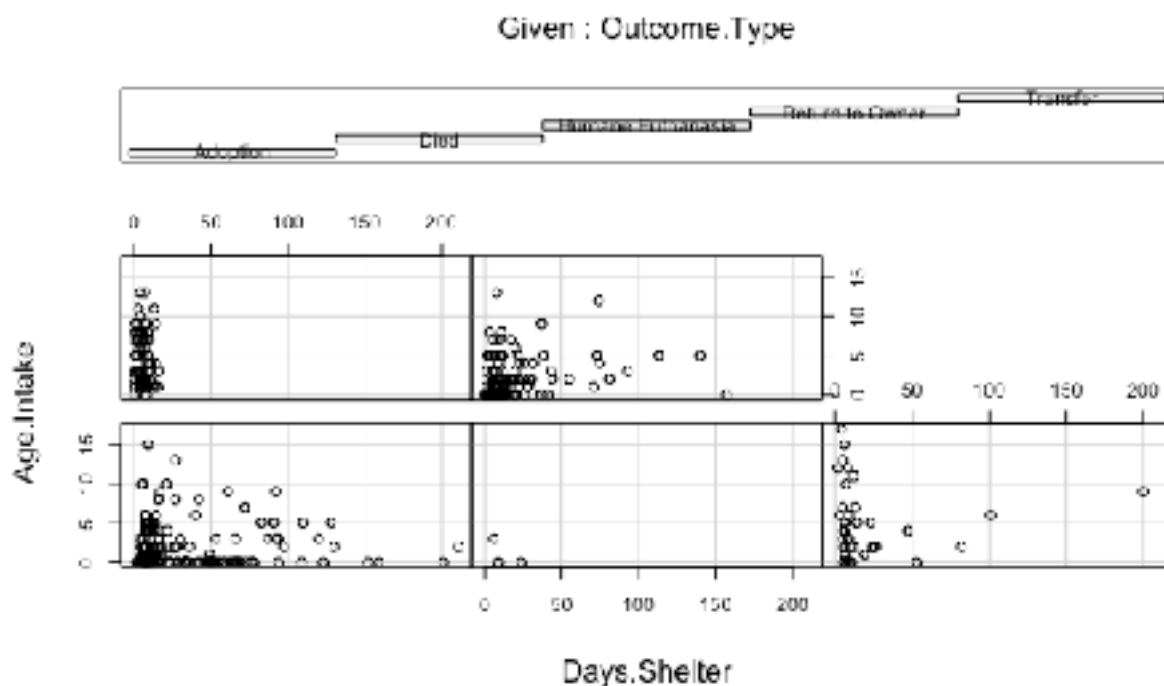
- Igual que con el peso para con los días en el refugio, mediana de 9 días, 3er cuartil de 20 y promedio de 20.8, algunos pocos animales llevan mucho tiempo en el albergue.

Gráficos

```
> coplot(Age.Intake~Days.Shelter|Outcome.Type,data=AnimalData)
```

Se pueden notar varias cosas interesantes,

- Los retornos a los dueños ocurren en pocos días (podemos suponer que si se pierde la mascota, de las primeras cosas que se hace es llamar al refugio y recuperarla)
- Las eutanasias ocurren muchas veces en poco tiempo de la llegada de los animales, es posible que se sacrifiquen animales que ya estén heridos (debo considerar esta hipótesis posteriormente)
- Es un lugar bastante seguro, solo murieron 3 animales, no vale la pena estudiar esta dirección porque son pocos animales.
- Las adopciones ocurren en cualquier momento, hay hasta casos de animales adoptados después de 200 días en el lugar. También se muestra una preferencia por animales con 0 años de vida (ver la línea negra horizontal). Si la mascota tiene menos de 5 años su tiempo en el albergue será mucho menor.

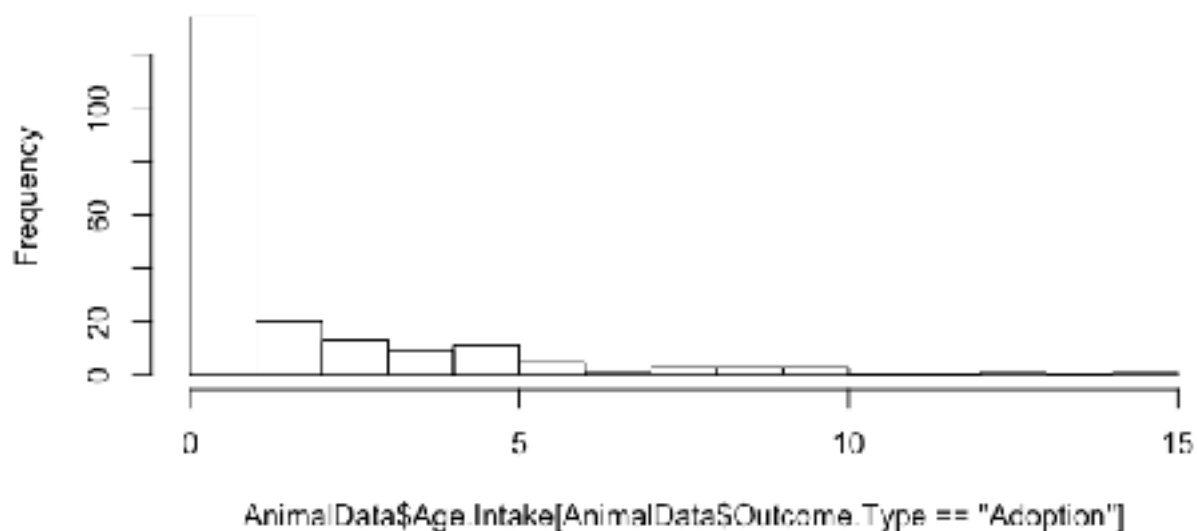


Edad de Adopción

Dibujé el histograma de la edad de los animales cuyo Outcome.Type=='Adoption'. Efectivamente se nota como hay una fuerte preferencia por los animales más jóvenes; los animales de 0 años son los preferidos por las personas que adoptan.

```
> hist(AnimalData$Age.Intake[AnimalData$Outcome.Type=='Adoption'],breaks=15)
```

Histogram of AnimalData\$Age.Intake[AnimalData\$Outcome.Type == "Adopti



Eutanasia

¿Porqué sacrifican a los animales?. Hago el subconjunto de animales sacrificados

```
Euta<-AnimalData[AnimalData$Outcome.Type=="Humane Euthanasia",]
```

Se tienen 39 casos de eutanasia, donde 23 son animales que estan "heridos o enfermos", o son "agresivos".

```
> dim(Euta[Euta$Aggressive=="Y" | Euta$Condition=="Injured or Sick",])  
[1] 23 24
```

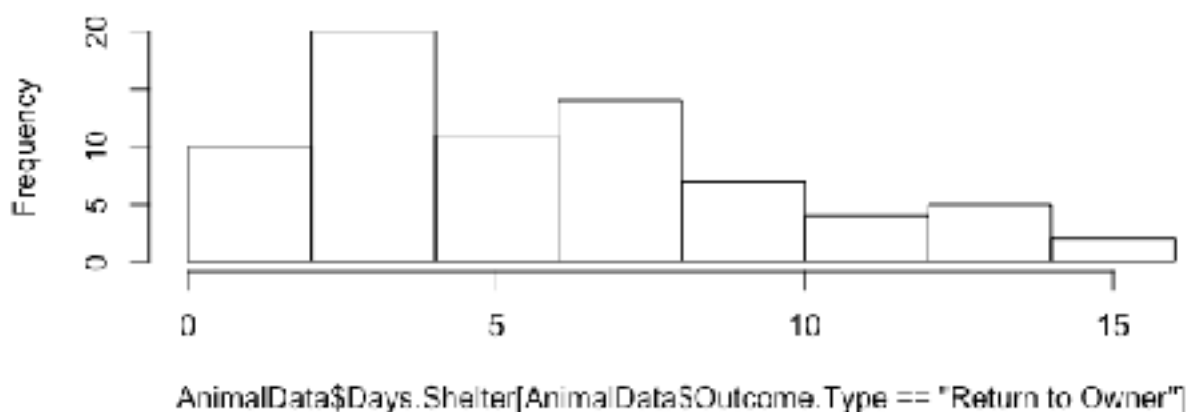
Retorno a Dueños

Los animales se retornan pronto a sus dueños, el máximo demoró 16 días. usando el siguiente código

```
> hist(AnimalData$Days.Shelter[AnimalData$Outcome.Type=='Return to Owner'])  
> max(AnimalData$Days.Shelter[AnimalData$Outcome.Type=='Return to Owner'])  
[1] 16
```

Se dibuja el histograma

Histogram of AnimalData\$Days.Shelter[AnimalData\$Outcome.Type == "Return to



Conclusiones:

- Los animales son adoptados en cualquier momento, pero existe una clara preferencia por animales cachorros, de menos de 1 año de edad.
- Los animales sacrificados están heridos o son agresivos en su mayoría.
- Los animales que son recuperados por sus dueños demoran poco en la instalación.