

Big Data Examen Final

Daniel Vallejo

9 de mayo de 2018

Use la biblioteca dplyr

Use ggplot

Lea el archivo msleep

Explique el conjunto de datos: Cual es el numero de columnas, cual es el numero de filas, nombre de columnas, tipo de cada columna

De un resumen de cada columna.

Muestre las columnas name, sleep_rem

Muestre todas las columnas excepto awake y bodywt

Muestre todas las columnas cuyo nombre inicie con “sl”

Muestre todas las columnas cuyo nombre inicie con “b”

Muestre todas las columnas cuyo nombre sea “name” o contengan “wt”

Muestre todos los animales que duerman en total menos de 16

Muestre los animales que duerman en total al menos 16 y pesen al menos 1

Muestre todos los primates o Perissodactyla

Que son estos ultimos en lenguaje coloquial?

Use el operador pipe en todo lo que sigue.

Muestre las columnas name y sleep_total, solo las primeras 5 filas

Ordene por la columna “order” (muestre solo los primeros 5 elementos).

Ordene por columna sleep_total

Cual es el animal que duerme menos?

Cual es el animal que duerme mas?

Muestre las columnas name, order, sleep_total ordenadas en forma descendente de acuerdo a la columna order, cuando duerman por lo menos 10 (use el operador pipe).

Agregar una columna al data frame con la proporcion de sleep_rem del total sleep_total.

Agregar una columna con el peso en gramos

Cual es el promedio del total de sueno?

Que animales se acercan mas a ese promedio?

Halle el promedio del total de sueno asi como el minimo y el maximo de los animales agrupados por orden.

Que orden es el que duerme mas en promedio? El que duerme menos?

Grafique sleep_total contra cada una de las demas variables, muestre cada uno de los ordenes en diferente color (use ggplot).

Encuentre que tan fuerte es la relacion lineal entre sleep_total y cada una de las demas variable. Ordene las variable de acuerdo a este valor.

Que variable es la que mas determina el sueno total?

La variable que menos determina el sueno total?

A entregar: un script en R con cada una de las preguntas anteriores como un comentario seguida del o los comandos en R que dan el resultado deseado.