

9,0

Tarea 01

Nombre del estudiante: Alejandro Sebiani

####1. Llena las palabras faltantes:

El grupo entero de individuos sobre el cual la información es requerida es llamada Los individuos son llamados La es la parte que es examinada para extraer información.

- a. población, variables explicativas, subgrupo
- b. población, unidades de muestreo, muestra**
- c. muestra, unidades de muestreo, población meta
- d. conjunto, ítem de interés, estrato



####2. En un estudio de la ecología de una especie de pez se colectaron las siguientes variables para cada individuo:

Sexo, peso inicial, temperatura corporal, peso ganado.

Escoge la respuesta correcta respecto al tipo de variables que se colectaron

- a. Nominal, racional, racional, racional
- b. Nominal, racional, intervalo, intervalo
- c. Ordinal, racional, intervalo, racional
- d. Nominal, racional, intervalo, racional**
- e. Ordinal, intervalo, racional, intervalo



####3. A partir de una muestra, el peso (kg) de mapaches del Parque Nacional Manuel Antonio resulta en los siguientes valores:

6.7, 2.7, 2.5, 3.6, 3.4, 4.1, 4.8, 5.9, 8.3

Ayuda: Crea el siguiente objeto “a” en R con los pesos de los mapaches:

```
a <- c(6.7, 2.7, 2.5, 3.6, 3.4, 4.1, 4.8, 5.9, 8.3)
```

¿Cuáles son el peso medio y el desvío estándar?

Peso medio

```
mean(a)
```

```
## [1] 4.666667
```

Desvío estándar

```
sd(a)
```

```
## [1] 1.952562
```

####4. Estos son los largos totales (cm) de 5 caimanes:

165, 175, 176, 159, 170

Ayuda: crea un objeto “b” en R con las medidas de los caimanes:

```
b <- c(165, 175, 176, 159, 170)
```

¿Cuáles son la mediana muestral y la media muestral?

Media muestral

```
mean(b)
```

```
## [1] 169
```

Mediana muestral

```
median(b)
```

```
## [1] 170
```

####5. Si la mayoría de los valores de un conjunto de datos son de aproximadamente de la misma magnitud excepto por unas pocas medidas que son bastante más grandes, ¿cómo serán la media y la mediana del conjunto de datos y que forma tendría el histograma?

- a. La media sería más pequeña que la mediana y el histograma sería asimétrico con una larga cola izquierda.
- b. La media será más grande que la mediana y el histograma sería asimétrico con una larga cola derecha. ✓
- c. La media sería más larga que la mediana y el histograma sería asimétrico con una larga cola izquierda.
- d. La media sería más pequeña que la mediana y el histograma sería asimétrico con una larga cola derecha.** ✗
- e. La media sería igual a la mediana y el histograma sería simétrico.

####6. ¿Cuál de los siguientes enunciados es VERDADERO?

- a. La media muestral no es sensible a valores extremos.
- b. El rango muestral no es sensible a valores extremos.
- c. El desvío estándar es una medida de dispersión alrededor de la media.** ✓
- d. El desvío estándar es una medida tendencia central alrededor de la media.
- e. Si la distribución es simétrica, entonces la media no es igual a la mediana.

####7. Se te permite escoger 4 números del 1 al 10 (sin reemplazo). ¿Cuál de los siguientes enunciados es VERDADERO?

- a. Los números 4,5,6,7, tienen el mayor desvío estándar.
- b. Los números 1,2,3,4, tienen el mayor desvío estándar.
- c. Los números 1,5,6,10, tienen el mayor desvío estándar.
- d. Los números 1,2,9,10, tienen el mayor desvío estándar.** ✓
- e. Los números 7,8,9,10, tienen el menor desvío estándar.

####8. ¿Cuál de los siguientes enunciados es VERDADERO?:

- a. Los números 3,3,3 tienen un desvío estándar de 3.
- b. Los números 3,4,5 tienen un desvío estándar menor al desvío estándar de los números 1003, 1004, 1005.
- c. La varianza es el desvío estándar al cuadrado.** ✓
- d. La moda es el valor que menos se repite.
- e. La media es una medida de dispersión de los datos.

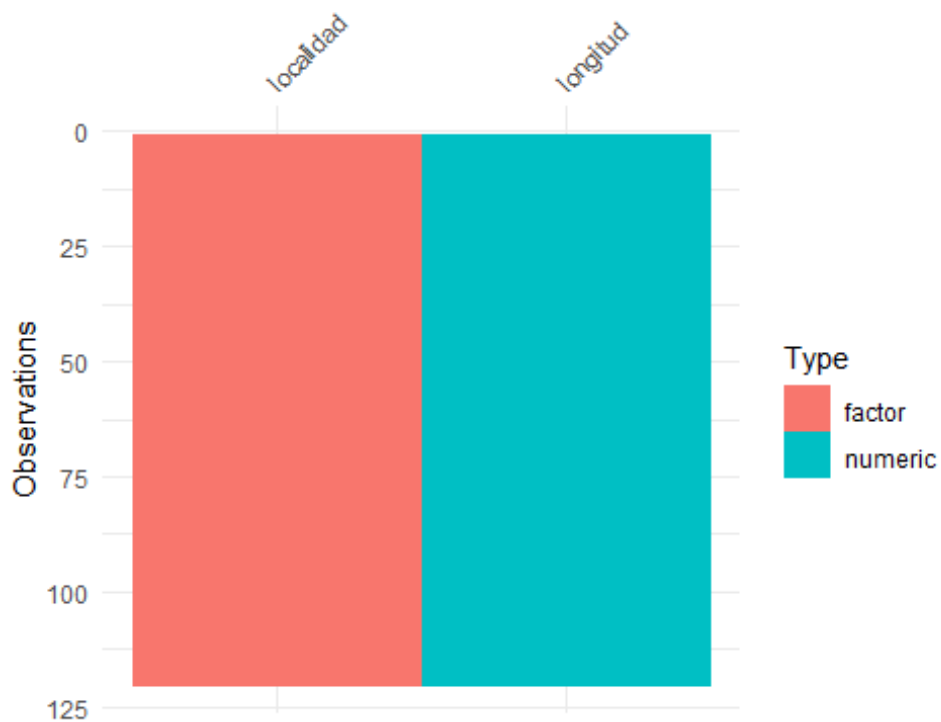
####9. Con los datos “alas.txt” (carpeta Datos) que contiene las longitudes de alas de una especie de aves en 2 localidades diferentes, calcula la media para cada localidad. Inserta un R chunk y escribe el código

```
library(rio)
library(visdat)
library(tidyverse)
```

```
alas <- import("datos/alas.txt") %>%
  na.omit(alas)
```

```
alas$localidad<-factor(alas$localidad)
```

```
vis_dat(alas)
```



```
medias<- alas %>%
  group_by(localidad) %>%
  summarise(promedio=mean(longitud))
```

```
medias
## # A tibble: 2 x 2
##   localidad promedio
##   <fct>         <dbl>
## 1 A             24.1
## 2 B             25.2
```



####10. Con los datos “alas.txt” que contiene las longitudes de alas de una especie de aves en 2 localidades diferentes, realiza un gráfico de caja (boxplot) y un gráfico de violín para cada localidad. Inserta un R chunk y escribe el código

```
ggplot(alas, aes(x=localidad, y=longitud))+
  geom_boxplot()+
  geom_violin(alpha=0.2)+
  geom_jitter(width=0.2)+
  geom_point(data=medias, aes(x=localidad, y=promedio), color="red", width=1)
```

