

Nombre del estudiante: **Carlos Rafael Arita Portillo**

Señala con **letra** (en este documento) la respuesta correcta. Devuelve el documento en formato PDF. Asegúrate de incluir tu nombre.

####1. Llena las palabras faltantes:

El grupo entero de individuos sobre el cual la información es requerida es llamada **población**. Los individuos son llamados **unidades de muestreo**. La **muestra** 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

b. **población, unidades de muestreo, muestra**

####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
- d. **Nominal, racional, intervalo, racional**

```
library(tidyverse)
library(rio)
library(fabricein)
library(visdat)
library(vtree)
library(collapse)
library(DescTools)
library(emmeans)
```

```
library(modelbased)
library(see)
library(parameters)
```

####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)
```

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

```
a
```

```
## [1] 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?

```
media = mean(a)
```

```
media
```

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

```
desviacion = sd(a)
```

```
desviacion
```

```
## [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)
```

```
b = c(165, 175, 176, 159, 170)
```

```
b
```

```
## [1] 165 175 176 159 170
```

¿Cuáles son la mediana muestral y la media muestral?

```
mediana = median(b)
```

```
mediana
```

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

```
media = mean(b)
```

```
media
```

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

####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.

b. La media será más grande que la mediana y el histograma sería asimétrico con una larga cola derecha.

####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.
- c. El desvío estándar es una medida de dispersión alrededor de la media.**

####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.

d. Los números 1,2,9,10, tienen el mayor 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.

c. La varianza es el desvío estándar al cuadrado.

####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

```
aves = import("alas.txt")
```

```
aves
```

```
##      longitud localidad
## 1  29.31140         A
## 2  25.96292         A
## 3  24.47853         A
## 4  22.94904         A
## 5  22.25191         A
## 6  25.20959         A
## 7  19.33376         A
## 8  25.78789         A
## 9  22.56246         A
## 10 25.52483         A
## 11 22.73126         A
## 12 23.48272         A
## 13 25.81436         A
## 14 23.61457         A
## 15 24.72731         A
## 16 28.36228         A
## 17 24.51804         A
## 18 16.98068         A
## 19 28.37925         A
```

## 20	24.20590	A
## 21	26.11796	A
## 22	29.06467	A
## 23	22.06382	A
## 24	22.88138	A
## 25	23.44479	A
## 26	21.18421	A
## 27	20.47674	A
## 28	23.77729	A
## 29	20.26934	A
## 30	19.18757	A
## 31	32.91597	A
## 32	24.23003	A
## 33	23.11119	A
## 34	19.48177	A
## 35	29.05983	A
## 36	22.77284	A
## 37	23.14654	A
## 38	26.00667	A
## 39	22.45082	A
## 40	22.07384	A
## 41	27.90548	A
## 42	22.82924	A
## 43	24.01340	A
## 44	19.48367	A
## 45	22.19061	A
## 46	23.63771	A
## 47	26.16500	A
## 48	28.88330	A
## 49	28.86669	A
## 50	21.99278	A
## 51	29.81984	B
## 52	25.58656	B
## 53	25.08861	B
## 54	29.47527	B
## 55	29.49170	B
## 56	23.38806	B
## 57	24.27135	B
## 58	20.95520	B
## 59	29.63321	B
## 60	25.35347	B
## 61	29.38996	B
## 62	17.82669	B
## 63	27.87755	B
## 64	23.40487	B
## 65	32.02936	B
## 66	28.57462	B
## 67	32.44551	B
## 68	25.92111	B
## 69	31.36365	B

## 70	27.43410	B
## 71	16.89100	B
## 72	28.58179	B
## 73	24.75652	B
## 74	27.93270	B
## 75	28.75290	B
## 76	25.62206	B
## 77	23.56552	B
## 78	23.08194	B
## 79	22.88308	B
## 80	17.25769	B
## 81	29.86668	B
## 82	30.05387	B
## 83	23.35198	B
## 84	21.43363	B
## 85	22.10865	B
## 86	28.06292	B
## 87	22.74691	B
## 88	21.86922	B
## 89	20.94718	B
## 90	26.03530	B
## 91	16.06792	B
## 92	30.20290	B
## 93	26.30017	B
## 94	30.51081	B
## 95	15.18989	B
## 96	27.29669	B
## 97	30.14747	B
## 98	23.99693	B
## 99	23.05739	B
## 100	28.78991	B
## 101	29.36690	B
## 102	26.98399	B
## 103	24.55058	B
## 104	24.88922	B
## 105	18.89952	B
## 106	26.95935	B
## 107	34.19588	B
## 108	24.66259	B
## 109	27.19062	B
## 110	23.20390	B
## 111	19.50692	B
## 112	17.15765	B
## 113	27.53538	B
## 114	29.07388	B
## 115	30.81347	B
## 116	21.89080	B
## 117	14.76077	B
## 118	21.59737	B

```
## 119 20.59379      B
## 120 25.92797      B

medias <- aves %>%
  group_by(localidad) %>%
  summarize(media = mean(longitud))

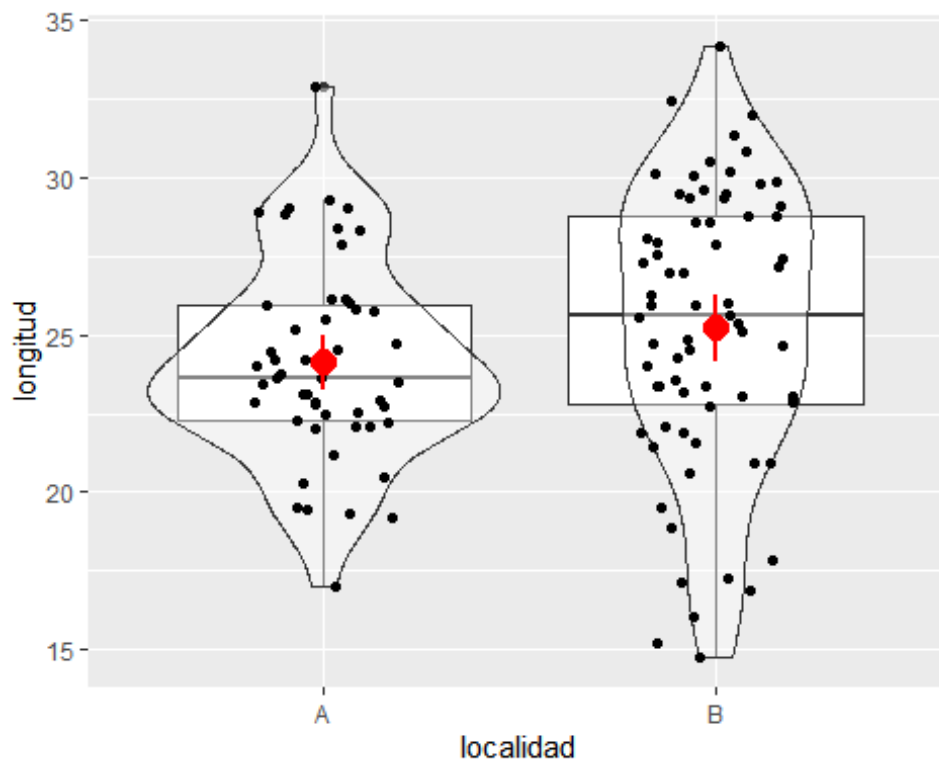
## `summarise()` ungrouping output (override with `.groups` argument)

medias

## # A tibble: 2 x 2
##   localidad media
##   <chr>     <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(aves, aes(localidad, longitud)) +
  geom_boxplot() +
  geom_violin(alpha = 0.4) +
  geom_jitter(width = 0.2) +
  stat_summary(fun.data = "mean_cl_normal", colour= "red", size =1)
```



03 Octubre del 2020