



Universidad Simón Bolívar

Departamento de Cómputo científico y
Estadística

Estadística para Ingenieros – CO3321

Enero-Marzo 2019

Reporte Laboratorio 1

Antonella Requena. Carnet: 15-11196

1. ¿Cuántos hombres y mujeres reporta?

Según los datos proporcionados, se reportan 24 hombres y 16 mujeres, para una muestra total de 40 personas.

Este cálculo se obtuvo utilizando el comando `length()` de R que indica la longitud de un vector, primero accediendo al vector columna de sexo mediante `table$sexo` (table corresponde a la variable donde se almacenó la tabla de datos), y se filtra mediante la sintaxis de acceso de datos con índices lógicos.

```
> length(table$sexo  
[table$sexo=='mujer'])
```

```
[1] 16
```

```
> length(table$sexo  
[table$sexo=='hombre'])
```

```
[1] 24
```

2. ¿Cuánto es el gasto promedio? y ¿segregado por género?

Gasto promedio

Con el comando `table$consumo` se extrajo el vector columna del consumo de toda la muestra, posteriormente con la función `mean()` de R (media aritmética), se calculó el promedio de estos datos. Así, se obtuvo lo siguiente:

```
> mean(table$consumo)
```

```
[1] 100375
```

Gasto promedio por factor género

Para filtrar el consumo por género se utilizó la sintaxis de acceso a elementos con índices lógicos

De la siguiente manera:

```
> cons_h = table$consumo  
[table$sexo == 'hombre' ]
```

Y esto retorno el vector columna del consumo de hombres únicamente, luego, se repitió el mismo procedimiento que en el apartado anterior

```
> mean(cons_h)
```

Arrojando:

```
[1] 106250
```

Análogamente para el consumo por mujeres:

```
> cons_m = table$consumo  
[table$sexo == 'mujer' ]
```

```
> mean(cons_m)
```

```
[1] 91562.5
```

3. ¿Se puede decir que en promedio los hombre gastaron más que las mujeres?

Si, porque la media aritmética calculada con R nos dice que para los hombres es 106250 y para las mujeres 91562.5, siendo el de los hombres mayor.

¿Cuáles son las respectivas desviaciones estándar?

Con la función `sd()` calculamos la desviación estándar para consumo de hombres y consumo de mujeres

(Recordando que `cons_m` corresponde a la variable del consumo de mujeres, y `cons_h` el consumo de hombres)

```
> sd(cons_m, na.rm=FALSE)
[1] 37404.04
```

```
> sd(cons_h, na.rm=FALSE)
[1] 40626.88
```

4. Indique el género que consume más café

Para realizar este cálculo, al igual que en el apartado 1., se utilizó la función `length()` de R

```
> length(table$Bebida
[table$sexo=='mujer' &
table$Bebida == 'cafe'])
[1] 12
```

```
> length(table$Bebida
[table$sexo=='hombre' &
table$Bebida == 'cafe'])
[1] 12
```

Dado que 12 hombres y 12 mujeres consumen café, ambos consumen café en cantidades iguales.

5. Indique el género que consume más arepa y café

Al igual que en el apartado anterior, accedemos a las columnas que queremos agregando mas condiciones lógicas, en este caso, según la tabla, es fácil saber quienes comen arepa ya que comen arepa cuando la una fila de la columna de empanada es `FALSE`, así que añadimos esa condición y hacemos los mismos cálculos

```
> length(table$sexo
[table$sexo=='mujer' &
table$Bebida == 'cafe' &
table$Consume_empanada ==
FALSE])
[1] 6
```

```
> length(table$sexo
[table$sexo=='hombre' &
table$Bebida == 'cafe' &
table$Consume_empanada ==
FALSE])
[1] 6
```

Nuevamente se arrojó que ambos géneros consumen la misma cantidad de café y arepa.

Para corroborar todos los cálculos utilizados, véase en este mismo repositorio el archivo `calculo.r`, allí podrá encontrar el script en R.