

Estadística Descriptiva

Paúl Arévalo

2024-07-03

Reporte de la muestra de 500 combos de Sándwiches, Salsas, Papas y Refrescos

Para comenzar el desarrollo del reporte es necesario cargar inicialmente el archivo .csv con la muestra

```
combos <- read.csv("combo.csv")
```

1 Comportamiento de la variable carne

1.1 Tendencia

Calculamos las medidas de tendencia central

1.1.1 Media

```
mean(combos$Carne)
```

```
## [1] 90.02181
```

1.1.2 Mediana

```
median(combos$Carne)
```

```
## [1] 90.02615
```

1.1.3 Cuartiles

```
quantile(combos$Carne)
```

```
##      0%      25%      50%      75%     100%  
## 83.16160 88.59600 90.02615 91.36860 95.79860
```

1.2 Variabilidad

1.2.1 Varianza

```
var(combos$Carne)
```

```
## [1] 4.170799
```

1.2.2 Desviación estandar

```
sd(combos$Carne)
```

```
## [1] 2.042254
```

1.2.3 Amplitud

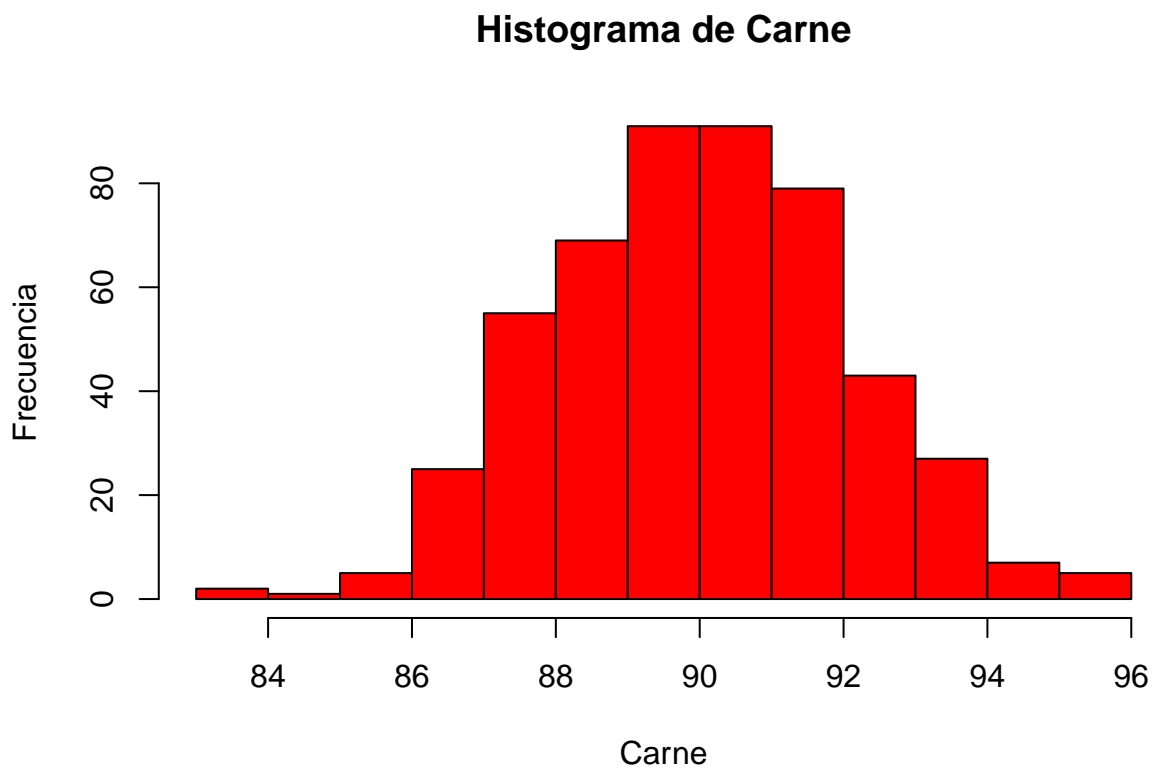
```
max(combos$Carne) - min(combos$Carne)
```

```
## [1] 12.637
```

1.3 Distribución

1.3.1 Histograma

```
hist(combos$Carne, col = "red", main = "Histograma de Carne", xlab = "Carne", ylab =  
  ↪ "Frecuencia")
```



1.3.2 Curtosis

```
library(moments)  
kurtosis(combos$Carne)
```

```
## [1] 2.911783
```

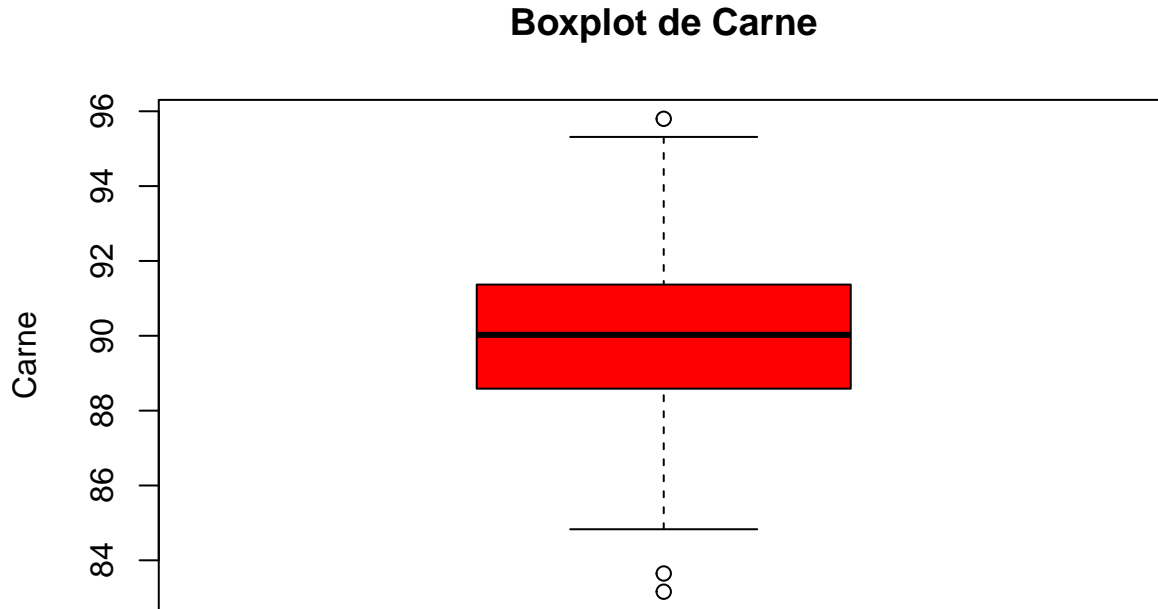
1.3.3 Asimetría

```
skewness(combos$Carne)
```

```
## [1] 0.008819527
```

1.4 Valores Atípicos

```
boxplot(combos$Carne, main = "Boxplot de Carne", ylab = "Carne", col = "red")
```



2 Comportamiento de la variable salsa

2.1 Tendencia

2.1.1 Media

```
mean(combos$Salsa)
```

```
## [1] 5.60811
```

2.1.2 Mediana

```
median(combos$Salsa)
```

```
## [1] 5.60025
```

2.1.3 Cuartiles

```
quantile(combos$Salsa)
```

```
##      0%      25%      50%      75%     100%  
## 4.92999 5.46569 5.60025 5.76792 6.36586
```

2.2 Variabilidad

2.2.1 Varianza

```
var(combos$Salsa)
```

```
## [1] 0.05488086
```

2.2.2 Desviación estandar

```
sd(combos$Salsa)
```

```
## [1] 0.2342666
```

2.2.3 Amplitud

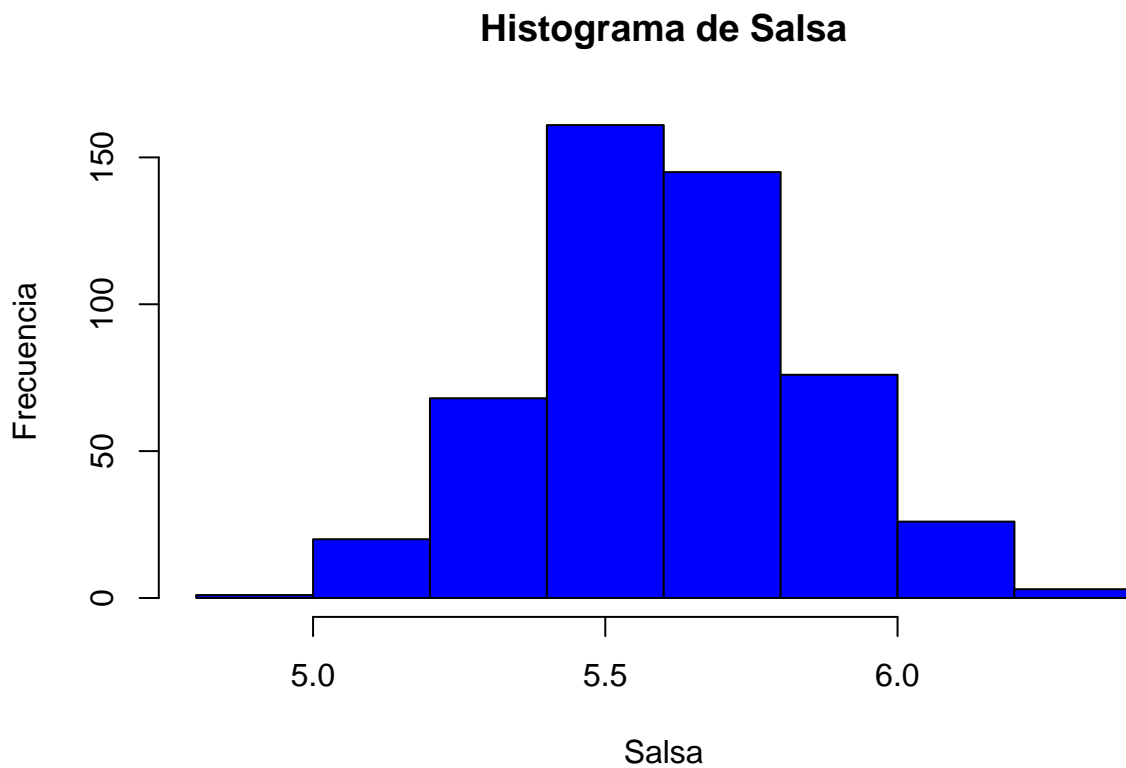
```
max(combos$Salsa) - min(combos$Salsa)
```

```
## [1] 1.43587
```

2.3 Distribución

2.3.1 Histograma

```
hist(combos$Salsa, col = "blue", main = "Histograma de Salsa", xlab = "Salsa", ylab =  
  ↪ "Frecuencia")
```



2.3.2 Curtosis

```
library(moments)
kurtosis(combos$Salsa)
```

```
## [1] 2.972769
```

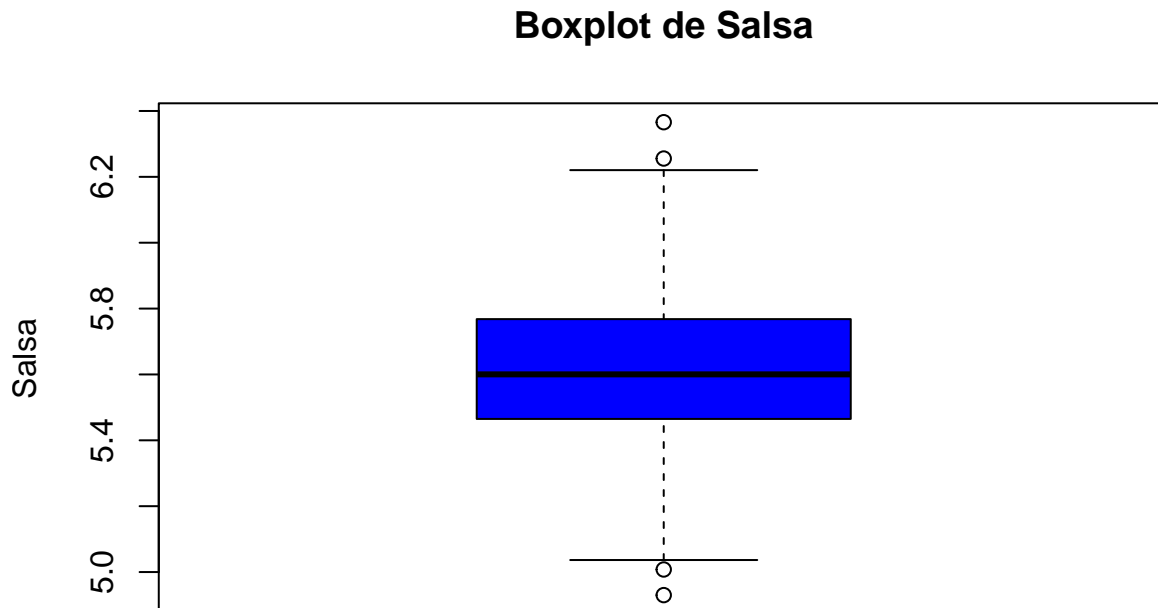
2.3.3 Asimetría

```
skewness(combos$Salsa)
```

```
## [1] 0.09954532
```

2.4 Valores Atípicos

```
boxplot(combos$Salsa, main = "Boxplot de Salsa", ylab = "Salsa", col = "blue")
```



3 Comportamiento de la variable papas

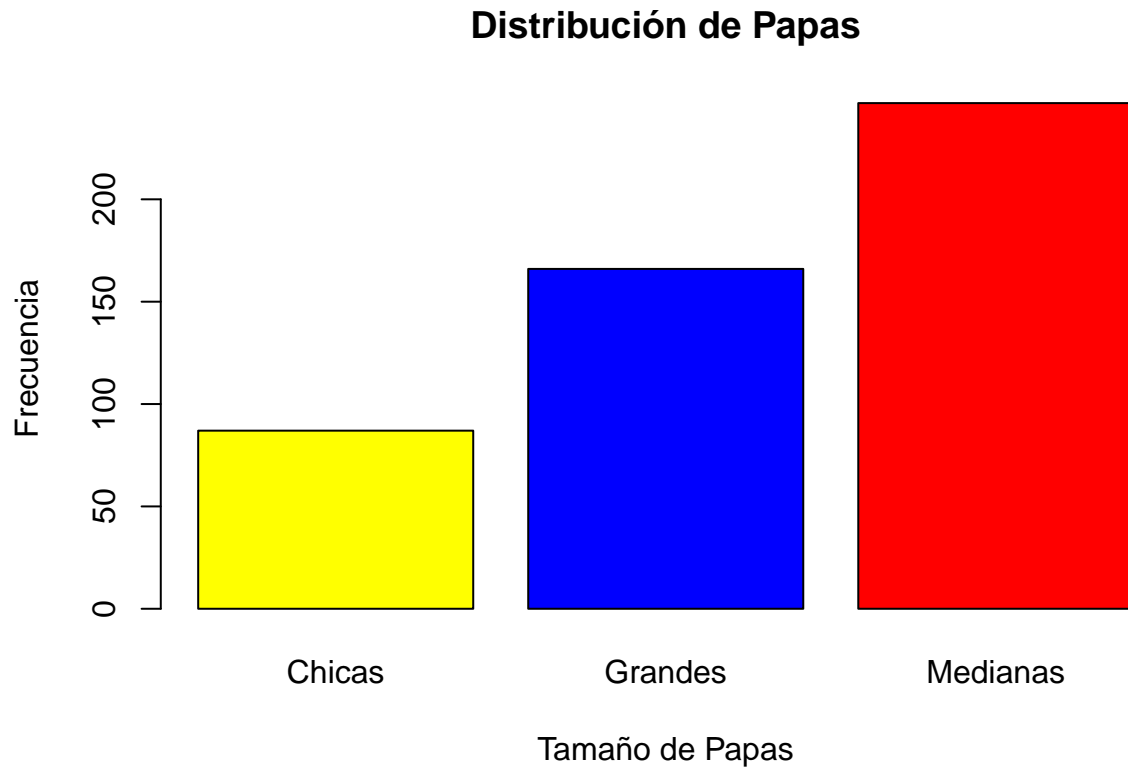
3.1 Tabla de frecuencias

```
tabla_papas <- table(combos$Papas)
tabla_papas
```

```
##
## Chicas  Grandes Medianas
##      87      166      247
```

3.2 Gráfico de barras

```
barplot(tabla_papas, main = "Distribución de Papas", xlab = "Tamaño de Papas", ylab =
  ↪ "Frecuencia", col = c("yellow", "blue", "red"))
```



4 Comportamiento bivariado de las variables carne y salsa

4.1 Cálculo del coeficiente de correlación entre las variables

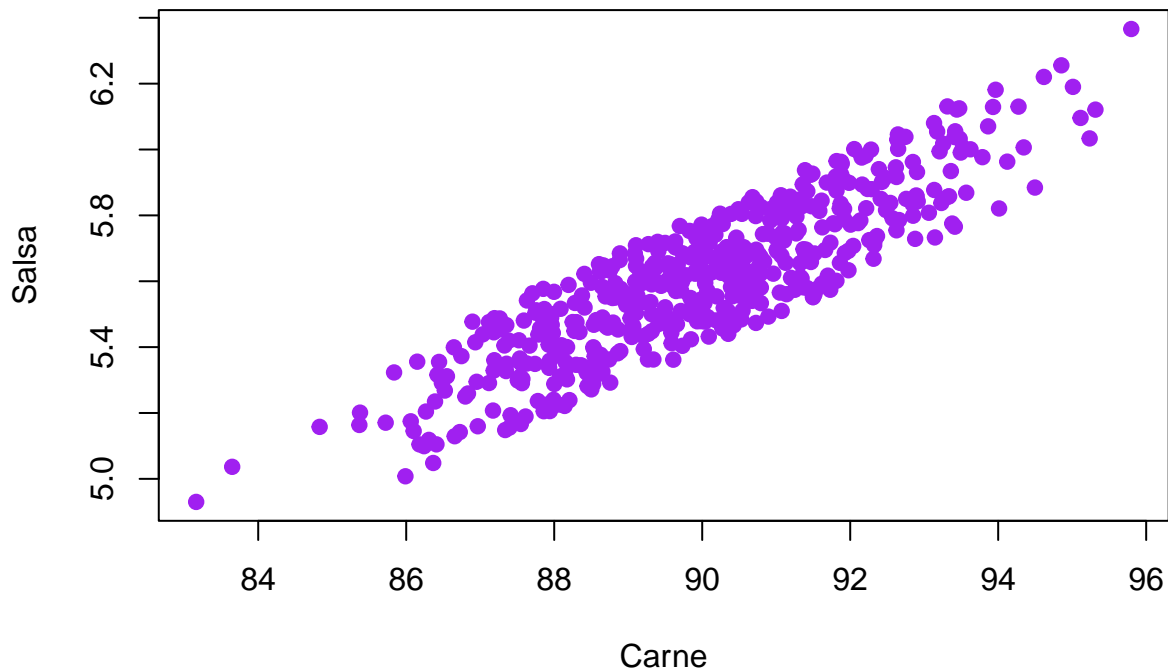
```
cor(combos$Carne, combos$Salsa)
```

```
## [1] 0.8816783
```

4.2 Análisis gráfico mediante un diagrama de dispersión

```
plot(combos$Carne, combos$Salsa,  
     main = "Diagrama de dispersión de Carne vs Salsa",  
     xlab = "Carne",  
     ylab = "Salsa",  
     pch = 19,  
     col = "purple"  
)
```

Diagrama de dispersión de Carne vs Salsa



4.3 Interpretación del análisis de los resultados

Mediante el análisis gráfico se puede inferir que las variables están correlacionadas, esto se demuestra con el resultado del coeficiente de correlación que es más cercano a uno, lo que significa que existe una relación positiva de las variables.

5 Combinación más frecuente de papas y refrescos

5.1 Tabla de frecuencias cruzada entre papas y refrescos

```
tabla_frecuencias <- table(combos$Papas, combos$Refresco)
tabla_frecuencias
```

```
##
##           Chico Grande Mediano
## Chicas       35     16      36
## Grandes      54     32     80
## Medianas     78     35    134
```

5.2 Identificar la combinación más frecuente

```
tabla_frecuencias[which.max(tabla_frecuencias)]
```

```
## [1] 134
```

A partir del valor obtenido podemos decir que la combinación de papas medianas y refresco mediano es más frecuente

6 Combinación menos frecuente de papas y refrescos

6.1 Indentificar la combinación más frecuente

```
tabla_frecuencias[which.min(tabla_frecuencias)]
```

```
## [1] 16
```

A partir del valor obtenido podemos decir que la combinación de papas chicas y refresco grande es más frecuente

7 Probabilidades de pedidos de un cliente

Calculamos el total de combos

n = Total de combos

```
n_combos <- nrow(combos)
```

7.1 Probabilidad de papas medianas

Eventos:

A = Número de papas medianas

$$P(A) = \frac{A}{n}$$

```
p_a <- dim(combos[which(combos$Papas == "Medianas"), ])/n_combos
p_a
```

```
## [1] 0.494 0.008
```

7.2 Probabilidad de papas medianas o refresco chico

Eventos:

A = Número de papas medianas

B = Número de refresco chico

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

```
p_b <- dim(combos[which(combos$Refresco == "Chico"), ])/n_combos
p_a_and_b <- dim(combos[which(combos$Papas == "Medianas" & combos$Refresco == "Chico"),
  ↪ ])/n_combos
p_a[1] + p_b[1] - p_a_and_b[1]
```

```
## [1] 0.672
```


7.3 Probabilidad de papas grandes y refresco chico

Eventos:

A = Número de papas grandes

B = Número de refresco chico

$$P(A \cap B)$$

```
(dim(combos[which(combos$Papas == "Grandes" & combos$Refresco == "Chico"), ]) /  
  ↪ n_combos)[1]
```

```
## [1] 0.108
```

7.4 Probabilidad de refresco chico si ya pidió papas grandes

Eventos:

A = Número de papas grandes

B = Número de refresco chico

$$P(B|A) = \frac{P(A \cap B)}{P(A)}$$

```
p_a_and_b_ <- dim(combos[which(combos$Papas == "Grandes" & combos$Refresco == "Chico"),  
  ↪ ]) / n_combos  
p_a_ <- dim(combos[which(combos$Papas == "Grandes"), ]) / n_combos  
p_a_and_b_[1] / p_a_[1]
```

```
## [1] 0.3253012
```

8 Verificación de independencia de los eventos papas grandes y refresco grande

8.1 Eventos

A = Número de papas grandes

B = Número de refresco grande

8.2 Regla de independencia

$$P(A \cap B) = P(A) \times P(B)$$

Debe cumplirse la igualdad para que los eventos A y B sean independientes

8.2.1 Cálculo de $P(A \cap B)$

```
p_izquierda <- dim(combos[which(combos$Papas == "Grandes" & combos$Refresco == "Grande"),  
  ↪ ]) / n_combos  
p_izquierda[1]
```

```
## [1] 0.064
```

8.2.2 Cálculo de $P(A) \times P(B)$

Usando la variable `p_a_` del ejercicio anterior

```
p_derecha <- p_a_ * dim(combos[which(combos$Refresco == "Grande"), ]) / n_combos
p_derecha[1]
```

```
## [1] 0.055112
```

8.3 Interpretación

```
p_izquierda[1] == p_derecha[1]
```

```
## [1] FALSE
```