



FACULTAD DE MATEMÁTICAS  
PONTIFICIA UNIVERSIDAD  
CATÓLICA DE CHILE

# EYP1113 - Probabilidad y Estadística

## Taller R - 02: Estadística Descriptiva

Ricardo Aravena C. - Cristian Capetillo C. - Ingrid Guevara R.  
Bladimir Morales T. - Ricardo Olea O. - Daniel Saavedra M.

Facultad de Matemáticas  
Departamento de Estadística  
Pontificia Universidad Católica de Chile

Segundo Semestre 2024

# Contenido I

Manipulación de Data en R

Medidas Descriptivas

Aplicación



# Manipulación de Data en R

- ▶ `import()` librería "rio" (importa data en distintos formatos)
- ▶ `filter()` librería "dplyr" (Filtra bases de datos)
- ▶ `select()` librería "dplyr" (Selecciona variables en una bases de datos)
- ▶ `spread()` librería "tidyr" (data frame vertical → horizontal)
- ▶ `aggregate()` (Resume bases de datos)
- ▶ `merge()` (Empalma bases de datos)
- ▶ `as.factor()` (Convierte una variable cuantitativa en cualitativa)
- ▶ `as.numeric()` (Convierte una variable cualitativa en cuantitativa)
- ▶ `data.frame()` (Construye una base de datos)
- ▶ `as.Date()` (Asigna atributo de fecha a un vector de fechas)

# Medidas Descriptivas

## Medidas Centrales

- ▶ Media: `mean()`
- ▶ Mediana: `median()`
- ▶ Moda: `mlv()` librería "modeest"

## Medidas Posición

- ▶ Cuantiles: `quantile()`
- ▶ Mínimo: `min()`
- ▶ Máximo: `max()`

# Medidas Descriptivas

## Medidas Dispersión

- ▶ Varianza: `var()`
- ▶ Desviación estándar: `sd()`
- ▶ Coeficiente de variación: `sd()/mean()`
- ▶ Rango: `max()-min()`
- ▶ Rango intercuartil: `IQR()`

## Medidas Asociación

- ▶ Covarianza: `cov()`
- ▶ Correlación: `cor()`

# Medidas Descriptivas

## Asimetría

- ▶ skewness: `skewness()` librería "moments"

## Kurtosis

- ▶ skewness: `kurtosis()` librería "moments"

## Resumen Estadístico

- ▶ summary: `summary()`
- ▶ describe: `describe()` librería "psych"

# Aplicación

El abalón es un molusco marino de la familia Haliotidae, conocido por su concha en forma de oreja, que es iridiscente y muy valorada tanto en joyería como en artesanía. En términos biológicos, el abalón es un tipo de caracol marino, y su carne es considerada un manjar en muchas culturas, especialmente en Asia.

La base de datos `ABALON.xlsx` contiene información correspondiente a la recolección de abalones en 4 centros distintos: Caldera, Coquimbo, Puerto Montt y Chiloé. En este taller se busca realizar un análisis descriptivo utilizando funciones y librerías de R.



# Aplicación

En particular, se quiere estudiar la relación existente entre las variables cuantitativas (largo, diámetro, alto, peso total, peso del cuerpo, peso de la concha, anillos) y si existen diferencias entre los abalones capturados en cada centro. Para ello, responda las siguientes interrogantes.

- a) Transforme la variable correspondiente al centro de recolección a una tipo factor.
- b) Utilizando la correlación, ¿podríamos decir que existe relación entre las dimensiones de los abalones, su peso y su edad?
- c) Compare, utilizando el promedio, las variables cuantitativas en los distintos centros. ¿Dónde se encuentran los abalones más grandes?, ¿dónde los más pequeños?, ¿dónde los más pesados?, ¿dónde los más livianos?, ¿dónde los más jóvenes y dónde los más viejos?



# Aplicación

- d) Realice la comparación anterior pero a través de la desviación estándar. Si tuviera que elegir el centro más disperso a un nivel general, ¿cuál sería este?
- e) Realice la comparación de (c) pero utilizando el cuantil 80, es decir, observar que ocurre con los abalones pesados, grandes y viejos. ¿Se mantienen las interpretaciones emitidas?, ¿qué puede decir al respecto?
- f) Por último, si tuviese que invertir en alguno de estos centros, ¿a priori cuál sería éste? y ¿qué centro le parecería más riesgoso de invertir?

