

# Estadística I

## Guía de Ejercicios 1

### *Vectores, data frames y tibbles*

**Profesor:** Víctor Macías E.

#### Pregunta 1

- (a) Genere un vector que contenga la siguiente secuencia de números de 1 al 10: 1, 2, 3, ..., 10
- (b) Genere un vector del 1 al 5 que contenga 10 elementos
- (c) Genere un vector que contenga los números: 10, 30, 50 y luego calcule su suma, suma acumulada y suma de los cuadrados de cada uno de los números.

#### Pregunta 2

El vector  $x$  contiene los años de vida útil restante de 5 máquinas, excepto para las máquinas  $B$  y  $D$ :

```
x = c("A"=6, "B"=NA, "C"=10, "D"=NA, "E"=20)
```

- (a) Defina un vector  $y$  que excluya los missing values (NA).
- (b) Calcule el promedio aritmético de la vida útil de las 5 máquinas
- (c) Calcule la proporción de máquinas que tienen una vida útil restante de 10 años o menos
- (d) ¿Cuántas máquinas tienen una vida útil menor o igual a 10 años?

#### Pregunta 3

La siguiente tabla presenta el número de unidades en stock de los productos 1 y 2 en cuatro zonas:

zona	stock_1	stock_2
A	800	700
B	200	300
C	700	600
D	300	400

- (a) Construya dicha tabla como un data frame/tibble y asigne el dataframe a `stock_total_df` y el tibble a `stock_total_tb`
- (b) Escriba en el script `summary(stock_total_tb)` ¿Qué obtiene?
- (c) Escriba en el script `str(stock_total_tb)` y `dplyr::glimpse(stock_total_tb)` ¿Qué obtiene?
- (d) Presente las primeras tres y últimas dos observaciones de `stock_total_tb`

- (e) Construya un tibble con las observaciones de las variables `zona` y `stock_2`
- (f) Construya un tibble con los stocks de las zonas *A* y *C*
- (g) Construya una variable `stock_total` que sea igual a la suma de `stock_1` y `stock_2`
- (h) Genere una variable que toma el valor “alto” si el stock total de una zona determinada es mayor a 1000 e igual a “bajo”, si no son mayores.
- (i) Guarde `stock_total_tb` con extensión *.csv*