

# Laboratorio 1

Juan Diego Sique Martínez

5/8/2020 AD

## Primer inciso

Ha sido contratado para trabajar en una consultoría a una embotelladora nacional. La embotelladora se encarga de distribuir su producto a distintos clientes, utilizando diferentes equipos de transporte y pilotos.

Se requiere

- Unificar todos los archivos en una tabla única.
- Agregar una columna adicional que identifique al mes y año de ese archivo, por ejemplo: Fecha: 01-2018.
- Exportar ese archivo en formato csv o Excel.
- Adjuntar el link de su Git Rmarkdown de R con lo que realizó lo anterior.
- Adjuntar archivo csv o Excel unificado que genera el archivo de R.

**Primera parte** Se cargarán los archivos. El nombre del archivo se guardará en la variable fecha.

```
# Cargando librerías
library(readr)
library(readxl)
library(dplyr)

# Obteniendo una lista de archivos
archivos = list.files(path = "Lab1", pattern = ".xlsx")

# Función para leer
leer_y_fecha <- function(x) {
  excel <- read_excel(paste("Lab1/", x, sep = ""))
  excel$fecha <- substr(x, 1, 7)
  return(excel)
}

# Retorna una lista de dataframes (tibble)
guardar = lapply(archivos, FUN = leer_y_fecha)

# Uniendo todos los dataframes
gran_dataframe = bind_rows(guardar)
```

**Segunda parte** Ahora se guardarán en un CSV todos juntos.

```
write.csv(gran_dataframe, file = "UnificadoInciso1.csv")
```

## Segundo inciso

Utilizando la función lapply, encuentre la moda de cada vector de una lista de por lo menos 3 vectores.

```

# Calcular la moda de un sólo vector
get_mode <- function(v) {
  uniqv <- unique(v)
  uniqv[which.max(tabulate(match(v, uniqv)))]
}

# Calcular la moda de varios vectores
vector_mode <- function(x) {
  return(get_mode(x))
}

# Haciendo vectores de ejemplo.
v <- c(2,1,2,3,1,2,3,4,1,5,5,3,2,3)
v2 <- c(2,1,1,1,1,1,1,1,2,3,4,1,2,3)
v3 <- c(0,0,0,0,0,0,0,0,0,0,0,0,2)

# Calculando la moda de los tres vectores
result <- lapply(list(v, v2, v3), vector_mode)

# Devuelve una lista con las modas de cada uno de los vectores correspondientes
result

## [[1]]
## [1] 2
##
## [[2]]
## [1] 1
##
## [[3]]
## [1] 0

```

Idea base tomada de [TutorialsPoint](#)<sup>1</sup> y [StackOverflow](#)<sup>2</sup>

### Tercer inciso

- Descargue de la página web de la SAT el archivo<sup>3</sup> de Parque Vehicular de Enero 2019.
- Leer el archivo en R. (Nota: usar `read_delim()` del paquete `readr`)

```

csv_leido_delimitador <- read_delim("INE_PARQUE_VEHICULAR_080219.txt", delim = "|")

# Mostrando la primera parte de la tabla
head(csv_leido_delimitador)

## # A tibble: 6 x 11
##   ANIO_ALZA MES  NOMBRE_DEPARTAM~ NOMBRE_MUNICIPIO MODELO_VEHICULO
##   <dbl> <chr> <chr>          <chr>          <chr>
## 1     2007 05  HUEHUETENANGO  "HUEHUETENANGO" 2007
## 2     2007 05  EL PROGRESO    "EL JICARO"      2007
## 3     2007 05  SAN MARCOS     "OCOS"           2007
## 4     2007 05  ESCUINTLA     "SAN JOS\x9c9"   2006
## 5     2007 05  JUTIAPA       "MOYUTA"         2007
## 6     2007 05  GUATEMALA     "FRAIJANES"      1997

```

<sup>1</sup>TutorialsPoint. (Desconocido). R - Mean, Median and Mode

<sup>2</sup>StackOverflow. (2014). Is there a built-in function for finding the mode?

<sup>3</sup>Corregido error de escritura en esta palabra

```
## # ... with 6 more variables: LINEA_VEHICULO <chr>, TIPO_VEHICULO <chr>,  
## #   USO_VEHICULO <chr>, MARCA_VEHICULO <chr>, CANTIDAD <dbl>, X11 <chr>
```

Archivo para lectura obtenido de Portal SAT.<sup>4</sup>

---

<sup>4</sup>SAT. (2020). Análisis Estadístico del Parque Vehicular