

Lab 1

Hans Walter

8/6/2020

Cargar Librerías

```
library(dplyr)
library(readr)
library(tidyverse)
library(readxl)
```

Ejercicio 1 Parte 1

Primero usamos la función *list.files* para que nos retorne un vector con los nombres de todos los archivos de excel del directorio. Hacemos un lapply para que lea de uno solo todos los archivos y los almacene en una lista de dataframes.

```
file_list <- list.files(pattern = '.xlsx')
df_list <- lapply(file_list, read_excel)
```

Ejercicio 1 Parte 2

Para aplicarle el nombre correcto a cada dataframe, eliminamos la extensión “.xlsx” de cada nombre del vector original **file_list** y almacenamos el resultado en **nombres_dataframes**. Aplicamos los nombres a la lista de dataframes.

```
nombres_dataframes <- str_remove_all(file_list, ".xlsx")
names(df_list) <- nombres_dataframes
```

Ejercicio 1 Parte 3

Usamos la función *Map* que va a agregar una nueva columna en cada dataframe de la lista, y va a asignarle la fecha que le corresponde por posición. Es decir, en la posición uno del vector de **nombres_dataframes** tiene “01-2018”, este lo va a asignar a cada fila del primer data frame, y así consecutivamente con los demás dataframes.

```
df_nueva <- Map(cbind, df_list, Fecha = nombres_dataframes)
df_nueva2 <- bind_rows(df_nueva)
df_final <- subset(df_nueva2, select = -c(TIPO, ...10))
head(df_final)
```

##	COD_VIAJE	CLIENTE	UBICACION	CANTIDAD
## 1	10000001	EL PINCHE OBELISCO / Despacho a cliente	76002	1200
## 2	10000002	TAQUERIA EL CHINITO Faltante	76002	1433
## 3	10000003	TIENDA LA BENDICION / Despacho a cliente	76002	1857
## 4	10000004	TAQUERIA EL CHINITO	76002	339
## 5	10000005	CHICHARRONERIA EL RICO COLESTEROL Faltante	76001	1644
## 6	10000006	UBIQUO LABS FALTANTE	76001	1827

##	PILOTO	Q CREDITO	UNIDAD	Fecha
## 1	Fernando Mariano Berrio	300.00	30 Camion Grande	01-2018
## 2	Hector Aragonés Frutos	358.25	90 Camion Grande	01-2018
## 3	Pedro Alvarez Parejo	464.25	60 Camion Grande	01-2018
## 4	Angel Valdez Alegria	84.75	30 Panel	01-2018
## 5	Juan Francisco Portillo Gomez	411.00	30 Camion Grande	01-2018
## 6	Luis Jaime Urbano	456.75	30 Camion Grande	01-2018

Ejercicio 1 Parte 4

Exportamos el archivo Excel con todas las correcciones hechas.

```
archivo_nuevo <- writexl::write_xlsx(df_final, "embotelladora.xlsx")
```

Ejercicio 2

Creamos 3 vectores, con un tamaño de 20 números y que escoga aleatoriamente entre un rango de 1 a 20. Dejamos replace TRUE para que puedan repetirse números dentro del vector. Finalmente almacenamos los 3 vectores en una lista

```
vector1 <- sample(1:20, size = 20, replace = TRUE)
vector2 <- sample(1:20, size = 20, replace = TRUE)
vector3 <- sample(1:20, size = 20, replace = TRUE)
lista_vectores <- list(vector1, vector2, vector3)
```

Función de Moda

```
moda <- function(x){
  y <- table(x)
  z <- data.frame(y)
  colnames(z) <- c("Número", "Frecuencia")
  resp <- z[which.max(z$Frecuencia),]
  return(resp)
}
```

Resultado

```
resultado <- lapply(lista_vectores, moda)
resultado
```

```
## [[1]]
##   Número Frecuencia
## 2      6          4
##
## [[2]]
##   Número Frecuencia
## 3      4          3
##
## [[3]]
##   Número Frecuencia
## 10     19          4
```

Ejercicio 3

Descargamos el dataset de la página de la SAT, previamente a realizar el código se analizó el texto para ver que delimitador se tenía que utilizar.

```
parque <- readr::read_delim("pvj2020.txt", delim = "|")
head(parque)
```

```
## # A tibble: 6 x 11
##   ANIO_ALZA MES  NOMBRE_DEPARTAM~ NOMBRE_MUNICIPIO MODELO_VEHICULO
##   <dbl> <chr> <chr>              <chr>              <chr>
## 1    2007 05   EL PROGRESO        "EL JICARO"        2007
## 2    2007 05   ESCUINTLA         "SAN JOS\xc9"      2006
## 3    2007 05   JUTIAPA           "MOYUTA"           2007
## 4    2007 05   GUATEMALA         "FRAIJANES"        1997
## 5    2007 05   QUETZALTENANGO    "QUETZALTENANGO"   2007
## 6    2007 05   HUEHUETENANGO     "CUILCO"           1999
## # ... with 6 more variables: LINEA_VEHICULO <chr>, TIPO_VEHICULO <chr>,
## #   USO_VEHICULO <chr>, MARCA_VEHICULO <chr>, CANTIDAD <dbl>, X11 <chr>
```