

UNIVERSIDAD DE EL SALVADOR
FACULTAD MULTIDISCIPLINARIA DE
OCCIDENTE
DEPARTAMENTO DE MATEMÁTICA
LICENCIATURA EN ESTADÍSTICA



PRACTICAS REALIZADAS EN EL
SOFTWARE PYTHON

DOCENTE:

LICENCIADO. Luis Ernesto Aquino Salazar

PRESENTADO POR:

Jenniffer Beatriz Alemán Castillo

Tabla de contenidos

Práctica de Restaurante	2
Exploración de la base de datos	2

Práctica de Restaurante

Exploración de la base de datos

El objetivo de esta práctica es descubrir el porqué de las disminuciones de las ventas del restaurante y así poder ayudar al dueño a encontrar las soluciones adecuadas a dicho problema, para esto se debe comenzar con la revisión de los datos que tiene la base, las columnas y el número de filas de esta. Para esto se seguirán los siguientes pasos:

1. Carga de librerías que se utilizaran
2. Carga de la base
3. Exploración básica de la base de datos
4. Creación de 5 gráficos

```
# Carga de las librerías que se utilizaran
```

```
library(ggplot2)
library(readxl)
library(dplyr)
```

Attaching package: 'dplyr'

The following objects are masked from 'package:stats':

```
filter, lag
```

The following objects are masked from 'package:base':

```
intersect, setdiff, setequal, union
```

```
Restaurante <- read_excel("D:/Beatriz/Ciclo II 2023/Análisis Estadístico con el Paquete R/Práctica 1/Restaurante.xlsx")
```

```
# Nombre de las columnas/variables
colnames(Restaurante)
```

```
[1] "Orden"      "Fecha"      "Hora de Cobro" "Mesa"
[5] "Atendió"    "Tipo"       "Producto"      "Categoría"
[9] "Precio"     "Costo"      "Propina"       "Tipo de Cliente"
```

```
#Número de filas y columnas
dim(Restaurante)
```

```
[1] 30132      12
```

```
#Visualización de los primeros 10 datos de la base
head(Restaurante)
```

```
# A tibble: 6 x 12
```

```
Orden Fecha      `Hora de Cobro` Mesa Atendió Tipo Producto
```

```

      <dbl> <dtm>          <dtm>          <dbl> <chr>   <chr>   <chr>
1      1 2014-02-01 00:00:00 1899-12-31 18:43:12      1 Rodolfo Bebida Tecate
2      1 2014-02-01 00:00:00 1899-12-31 18:43:12      1 Rodolfo Bebida Tisana
3      1 2014-02-01 00:00:00 1899-12-31 18:43:12      1 Rodolfo Bebida Café ameri~
4      1 2014-02-01 00:00:00 1899-12-31 18:43:12      1 Rodolfo Bebida Te helado
5      1 2014-02-01 00:00:00 1899-12-31 18:43:12      1 Rodolfo Comida Pera Rosti~
6      1 2014-02-01 00:00:00 1899-12-31 18:43:12      1 Rodolfo Comida Pan de Elo~
# i 5 more variables: Categoria <chr>, Precio <dbl>, Costo <dbl>,
#   Propina <dbl>, `Tipo de Cliente` <chr>

```

Ahora se procede a obtener gráficos estadístico de la base de datos que ayudaran a entender el comportamiento de los datos:

Margen de beneficio = (Precio de venta - Costo del producto) / Precio de venta x 100

```

# Agregar una columna de margen de beneficio a la base de datos Restaurante
Restaurante <- Restaurante %>% mutate(margen_beneficio = (Precio - Costo) / Precio * 100)

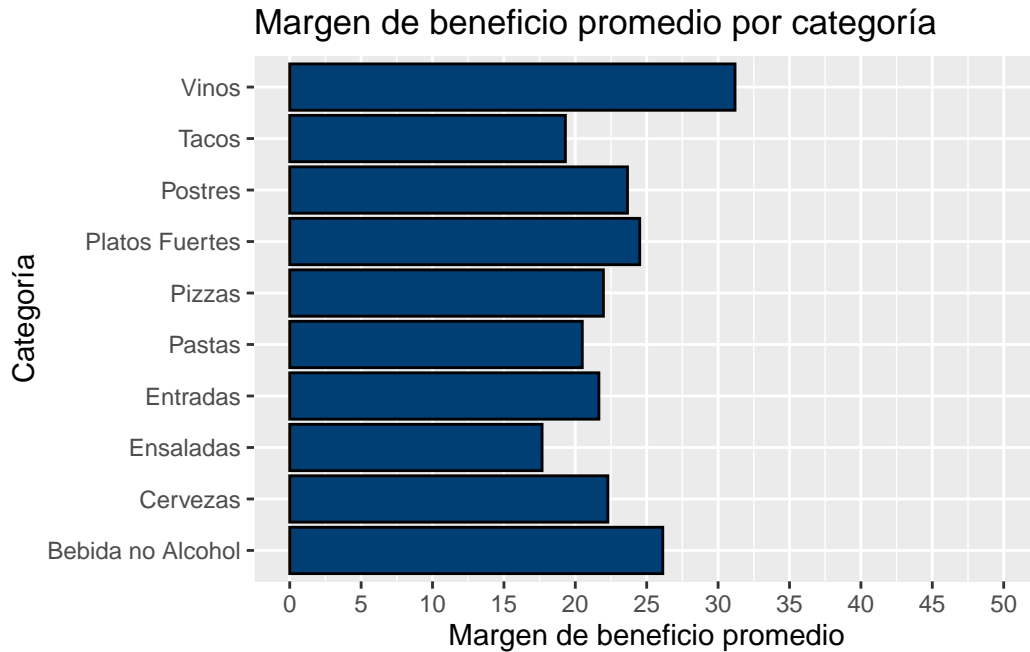
# Crear un nuevo dataframe con el margen de beneficio promedio por categoría
margen_promedio <- Restaurante %>% group_by(Categoria) %>% summarize(margen_beneficio_promedio =

# Ordenar el dataframe por margen de beneficio promedio
margen_promedio <- margen_promedio[order(margen_promedio$margen_beneficio_promedio),]

# Crear el gráfico de barras con el eje y contado de 5 en 5
graf_margen <- ggplot(data = margen_promedio, aes(x = Categoria, y = margen_beneficio_promedio))
  geom_bar(stat = "identity", fill = "#003E74", col="black") +
  coord_flip() +
  labs(title = "Margen de beneficio promedio por categoría", x = "Categoría", y = "Margen de bene
  scale_y_continuous(limits = c(0, 50), breaks = seq(0, 50, 5))

# Mostrar el gráfico
graf_margen

```



El gráfico anterior los márgenes de beneficios que se obtienen en promedio por cada categoría, se puede apreciar que las entradas son la categoría que menos márgenes de beneficio obtiene y que los vinos son los que obtiene un mayor beneficio.

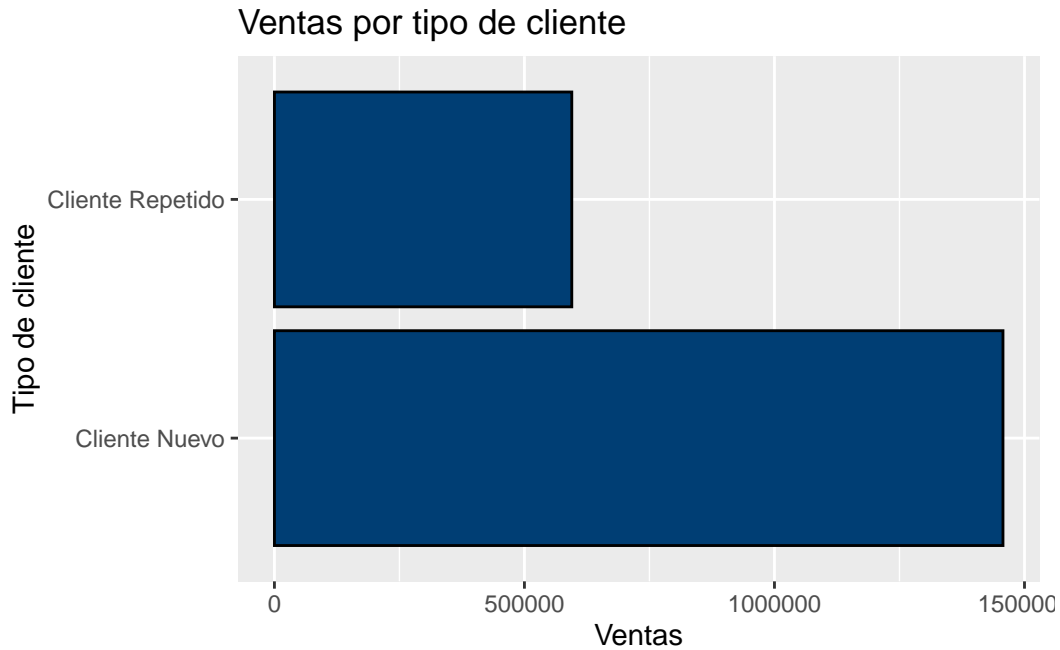
Ahora se procederá a realizar un gráfico para conocer cuál es la situación del restaurante con respecto a la clientela de este:

```
# Crear un nuevo dataframe con las ventas por tipo de cliente
ventas_tipo_cliente <- Restaurante %>% group_by(`Tipo de Cliente`) %>% summarize(ventas = sum(Pre

# Ordenar el dataframe por ventas
ventas_tipo_cliente <- ventas_tipo_cliente[order(ventas_tipo_cliente$ventas),]

# Crear el gráfico de barras
graf_ventas <- ggplot(data = ventas_tipo_cliente, aes(x = `Tipo de Cliente`, y = ventas)) +
  geom_bar(stat = "identity", fill = "#003E74", col="black") +
  coord_flip() +
  labs(title = "Ventas por tipo de cliente", x = "Tipo de cliente", y = "Ventas")

# Mostrar el gráfico
graf_ventas
```

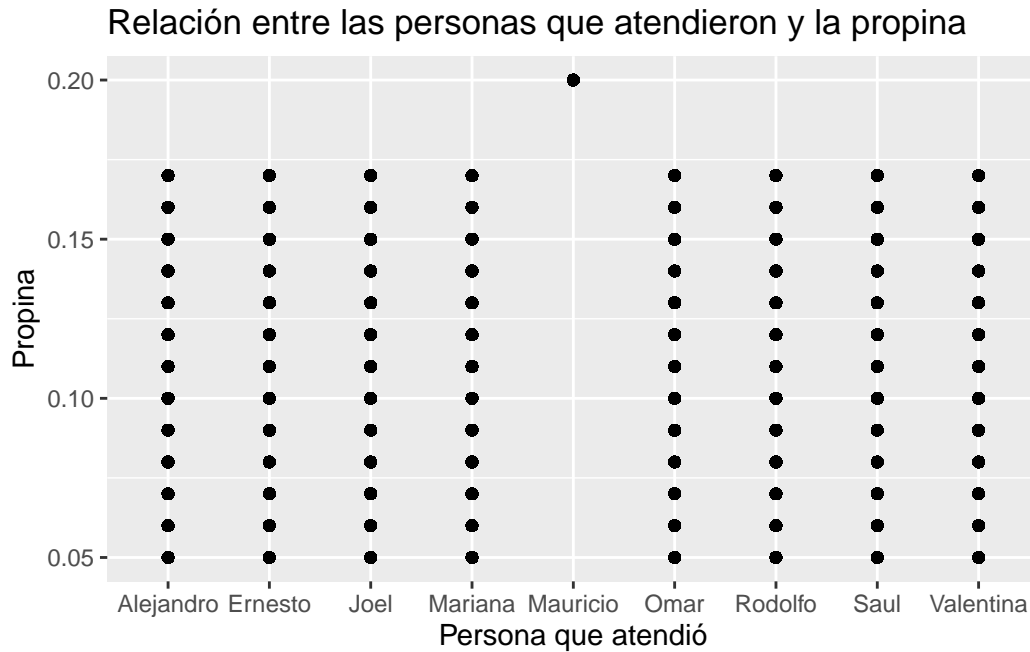


En el gráfico anterior se puede apreciar que hay muchos más clientes nuevos que clientes repetidos, esto puede llegar a resultar en que los clientes nuevos que van llegando no disfrutaron del servicio por lo que deciden no regresar o por alguna situación similar deciden no volver, esto puede ser una de las razones donde el dueño del restaurante debe enfocarse para poder resolver el problema de disminución en las ventas del restaurante

Ahora se procederá a realizar un gráfico para conocer las propinas que los clientes dejan con respecto a la persona que les atendió, si bien esto no es una razón para asegurar que la persona que atendió es buena o mala en su trabajo esto puede ayudar a saber si podría ser un factor para la disminución de ventas.

```
# Crear el gráfico de dispersión
graf_propina <- ggplot(data = Restaurante, aes(x = Atendió, y = Propina)) +
  geom_point() +
  labs(title = "Relación entre las personas que atendieron y la propina", x = "Persona que atendió")

# Mostrar el gráfico
graf_propina
```



El gráfico anterior muestra que, si todos han recibido propinas de diferente porcentaje, no se observa nada en específico, excepto en Mauricio que todas sus propinas fueron del 20

Además de estos análisis también se debe conocer que comportamiento tienen las ventas del restaurante por mes, para esto se tomó en cuenta el precio que los clientes pagaron, sin tomar en cuenta las propinas:

```
# Crear un nuevo dataframe con las ventas por mes
ventas_mes <- Restaurante %>% mutate(Mes = format(as.Date(Fecha), "%m-%Y"), Año = format(as.Date(Fecha), "%Y"))

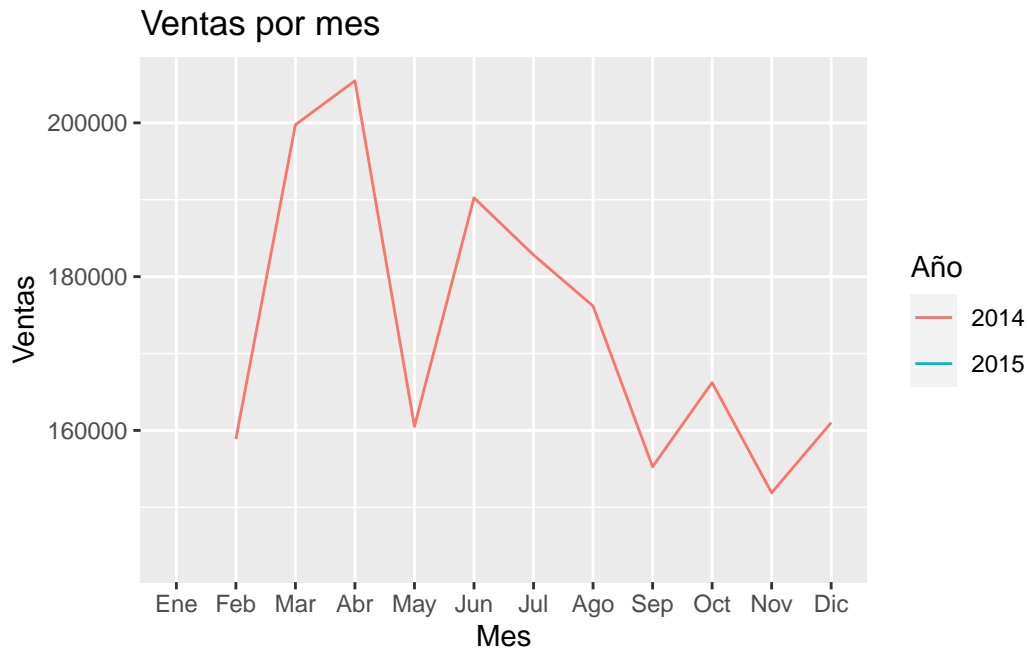
`summarise()` has grouped output by 'Mes'. You can override using the ``.groups`
argument.

# Filtrar los datos por los años 2014 y 2015
ventas_mes <- ventas_mes[ventas_mes$Año %in% c("2014", "2015"),]

# Crear el gráfico de líneas
graf_ventas <- ggplot(data = ventas_mes, aes(x = Mes, y = ventas, color = Año, group = Año)) +
  geom_line() +
  labs(title = "Ventas por mes", x = "Mes", y = "Ventas") +
  scale_x_discrete(labels = c("Ene", "Feb", "Mar", "Abr", "May", "Jun", "Jul", "Ago", "Sep", "Oct", "Nov", "Dic"))

# Mostrar el gráfico
```

graf_ventas



En el gráfico anterior se puede observar como se tuvieron mayores ventas en el mes de abril y que a partir de ese mes las ventas comenzaron a disminuir. El año de 2015 no se observa dado que solo se tienen datos del primer mes del año. También es importante tener en cuenta el margen de beneficios que se han tenido por mes:

```
# Crear un nuevo dataframe con el margen de beneficio por mes
beneficio_mes <- Restaurante %>% mutate(Mes = format(as.Date(Fecha), "%m-%Y"), Año = format(as.Date(Fecha), "%Y"))

`summarise()` has grouped output by 'Mes'. You can override using the ``.groups`
argument.
```

```
# Filtrar los datos por los años 2014 y 2015
beneficio_mes <- beneficio_mes[beneficio_mes$Año %in% c("2014", "2015"),]

# Crear el gráfico de líneas
graf_beneficio <- ggplot(data = beneficio_mes, aes(x = Mes, y = beneficio, color = Año, group = Año)) +
  geom_line() +
  labs(title = "Margen de beneficio por mes", x = "Mes", y = "Margen de beneficio") +
  scale_x_discrete(labels = c("Ene", "Feb", "Mar", "Abr", "May", "Jun", "Jul", "Ago", "Sep", "Oct", "Nov", "Dic"))
```



```
# Mostrar el gráfico  
graf_beneficio
```

