**Proyecto**

**Caso de negocio:**

Construimos una base de datos referente a una venta de automóviles, en donde tenemos datos distintos de vehículos, precio, fecha de venta, país de venta, información de los clientes, características del vehículo etc. Con la finalidad de ejecutar diferentes análisis aplicando los conocimientos adquiridos en el curso de arquitectura de datos.

**Base de datos:**



Primeramente vamos a importar la base de datos como dataframe

#data frame

Proyecto = as.data.frame(Projecto\_DATA)

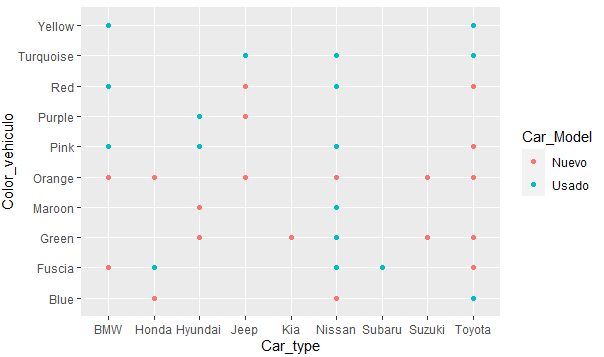
View(Proyecto)

**Caso 1 Chile**

1-Queremos conocer que tipos de vehículos se vendieron en el país de Chile, esto no solamente va a incluir el tipo sino también el color de vehículo y si es nuevo o usado

filter1\_Chile = filter(Proyecto, Country == "Chile")

ggplot(data=filter1\_Chile, aes(Car\_type, Color\_vehiculo, color=Car\_Model)) + geom\_point()



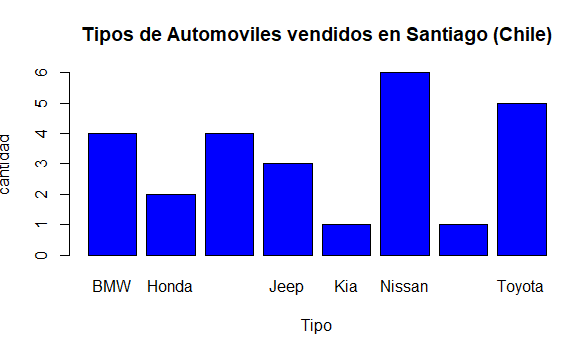
En el grafico anterior podemos identificar los diferentes Vehiculos, su color y su estado que se vendieron en Chile.

2- Especificamente en Santiago de Chile que tipo de Vehiculos se vendieron

plot(as.factor(Proyecto[Proyecto$City == 'Santiago',]$Car\_type),

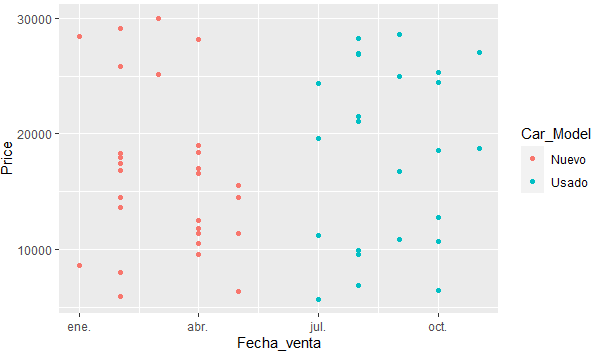
xlab = 'Tipo', ylab = 'cantidad', col = 'blue',

main='Tipos de Automoviles vendidos en Santiago (Chile)')



3- adicionalmente queremos visualizar los datos que incluya los precios y las fechas en que fueron vendidos los vehículos así como su estado (Nuevo o usado) esto con el fin para la toma de decisiones y el próximo plan de marketing de la compañía.

ggplot(data=filter1\_Chile, aes(Fecha\_venta, Price, color=Car\_Model)) + geom\_point()

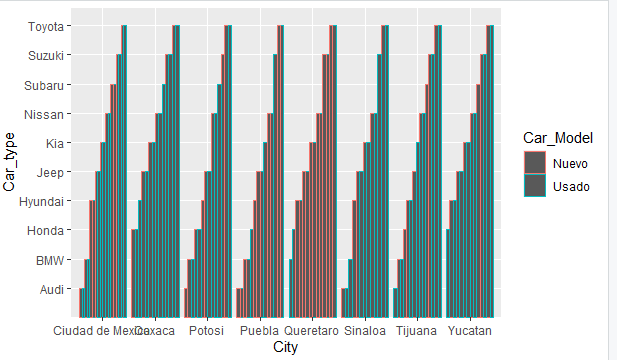


Si visualizamos por las diferentes nubes que se forman, podemos notar que los primeros meses de enero se vendieron mas autos nuevos y en mayor cantidad que la segunda parte del año. Notamos una fuerte tendencia en el mes de abril al comprar automóviles nuevos entre 10 y 20 mil dólares

Caso 2 México

1-Queremos conocer que tipos de vehículos se vendieron en el país de México y en cuales regiones o provincias, esto no solamente va a incluir el tipo sino si es nuevo o usado

ggplot(data=filter2\_mexico, aes(City, Car\_type, color=Car\_Model)) + geom\_bar(stat="identity", position="dodge")



A pesar de tener la información detallada sobre vehículo y provincia no es muy clara al ser muchos datos en el gráfico, vamos a proceder a desmenuzar un poco más la información

2- Vamos a filtrar por todas las provincias disponibles menos ciudad de México y efectuar el conteo para verificar cuales tipos de vehículos se vendieron en las provincias

filter3\_mexico\_city = filter(filter2\_mexico, City != "Ciudad de Mexico")

Group\_CarType\_mexico = group\_by(filter3\_mexico\_city, Car\_type, City)

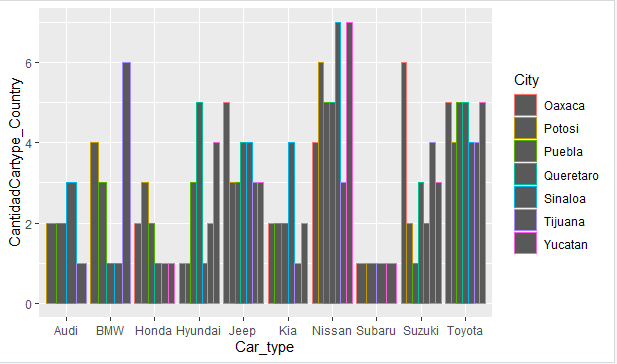
data\_Agrupada\_MX = summarise(Group\_CarType\_mexico,

CantidadCartype\_Country = n())

Ordenar\_mexico\_city = arrange(data\_Agrupada\_MX, City)

View(Ordenar\_mexico\_city)

ggplot(data = Ordenar\_mexico\_city, aes(Car\_type, CantidadCartype\_Country, color=City)) + geom\_bar(stat="identity", position="dodge")



En el anterior grafico podemos observar una mejor distribución de los tipos de automóviles, cuantos se vendieron en cada provincia, podemos observar como Nissan fue la marca más vendida en Sinaloa y Yucatán

3- Así mismo queremos averiguar solamente la marca Toyota, la cantidad de autos vendidos en cada provincia

filter5\_Toyota = filter(filter2\_mexico, Car\_type == "Toyota")

Group\_CarType\_Toyota = group\_by(filter5\_Toyota, Car\_type, City)

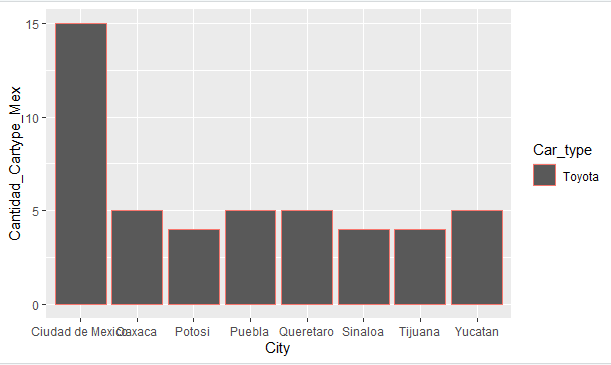
data\_Agrupada\_MX\_3 = summarise(Group\_CarType\_Toyota,

Cantidad\_Cartype\_Mex = n())

Ordenar\_Toyota = arrange(data\_Agrupada\_MX\_3, Cantidad\_Cartype\_Mex)

View(Ordenar\_Toyota)

ggplot(data = Ordenar\_Toyota, aes(City, Cantidad\_Cartype\_Mex, color=Car\_type)) + geom\_bar(stat="identity", position="dodge")



Podemos apreciar la marca Toyota que en la ciudad de México fue donde más se vendió

4- se creo una función para determinar la fecha de venta y el precio de los diferentes automóviles en diferentes ciudades

select (Ordenar\_Toyota, Car\_type, City)

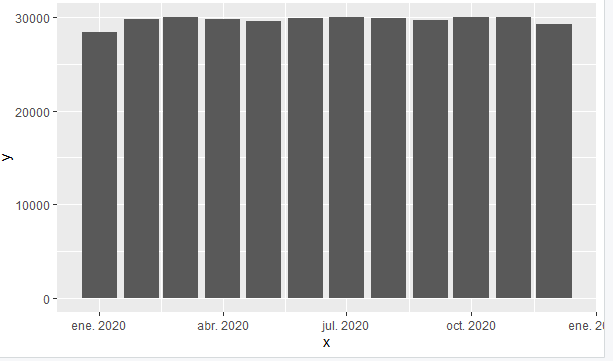
distinct (Ordenar\_Toyota, Cantidad\_Cartype\_Mex)

mifuncion <- function(Car\_type,City, x,y) {

ggplot(data = Proyecto, aes(x, y))+

geom\_bar(stat="identity", position="dodge")}

mifuncion("Toyota","Ciudad de Mexico",Proyecto$Fecha\_venta,Proyecto$Price)



### Link del proyecto [Repositorio\_Proyecto\_Arquitectura\_curso5](https://github.com/David66300/Repositorio_Proyecto_Arquitectura_curso5) Private