



dcc

CIENCIAS DE LA COMPUTACIÓN  
UNIVERSIDAD DE CHILE

# Hito 3 : Analisis clientes

## ***Block***

María Hernández  
Lung Pang  
Cristóbal Saldías  
Víctor Vidal

# Block

## NUESTROS DESTACADOS



ZAPATILLA VANS OLD SKOOL  
MUJER NEGRO

Vans  
\$49.990



ZAPATILLA VANS OLD SKOOL  
HOMBRE NEGRO

Vans  
\$49.990



ZAPATILLA CONVERSE  
CHUCK TAYLOR  
PLATAFORMA MUJER  
BLANCO



ZAPATILLA VANS SK8-LOW  
UNISEX NEGRO

Vans  
\$49.990



ZAPATILLA VANS ERA MUJER  
AZUL

Vans  
\$39.990

## LAS MEJORES MARCAS



**Despachamos en RM  
al día hábil siguiente**  
Regiones desde 2  
días hábiles

VER COMUNAS



**¿No conoces tu talla?**  
¡Revisa nuestra guía  
y encuéntrala!

VER GUÍA

# Motivación

Ventas Online crecen exponencialmente durante pandemia

Las empresas cambian su estrategia de negocios a un medio digital



# Objetivos principales

Estudiar el comportamiento del cliente chileno mediante Técnicas de minería de datos.

Estudiar qué tan útiles son las técnicas de minerías de datos para generar futuras recomendaciones



# Preguntas y problemas

1. ¿Qué marcas se compran juntas? (En un mismo pedido)
2. ¿Se puede saber de qué comuna se hizo una compra de acuerdo a su historial de compras?
3. ¿Qué tipo de producto puede que compre alguien que ya adquirió x tipo de producto?
4. Comportamiento de clientes con respecto a la cantidad de compras y su monto total.

# Datos

Se obtuvieron 2 tablas desde Shopify

Luego de una limpieza las tablas quedaron de la forma :

- **Orders**(ID, Email, Precio Pedido, Numero Pedido, ID Cliente, Fecha Compra, Año Compra, Mes Compra, Cantidad Pedidos Cliente, Total Gastado Cliente, RUT, Direccion 1, Direccion 2, Comuna, Nombre Cliente, REGION CON CODIGO, Tipo despacho)
- **Order\_detail**(id, Numero Pedido, Fecha Pedido, Nombre Producto, SKU, Cantidad, Precio Producto, Precio Total Productos, Marca)

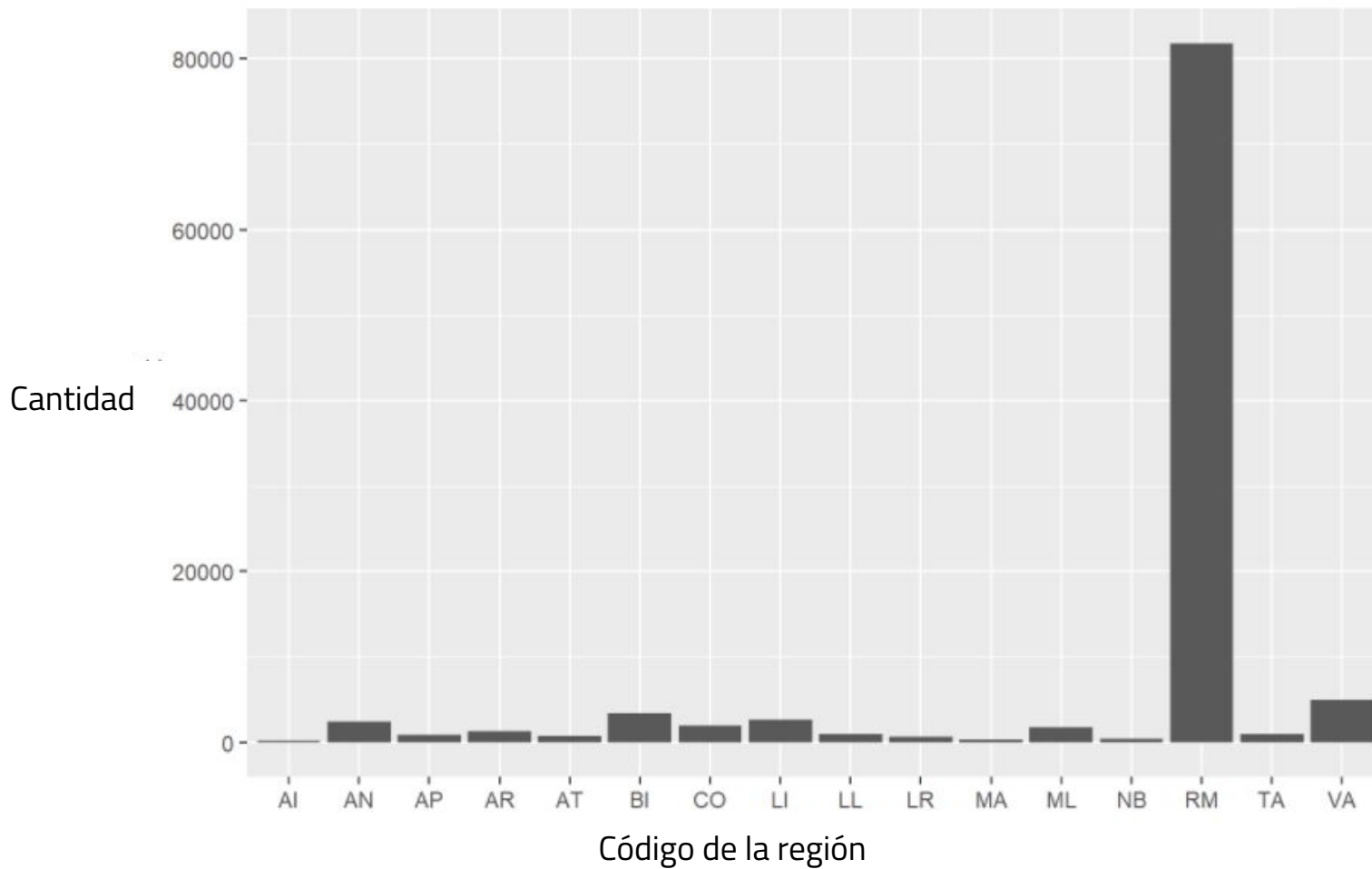
**-Se anonimizaron los datos de las personas**

**-Se filtraron nuevamente para hacer los experimentos**

# EXPLORACIÓN DE DATOS

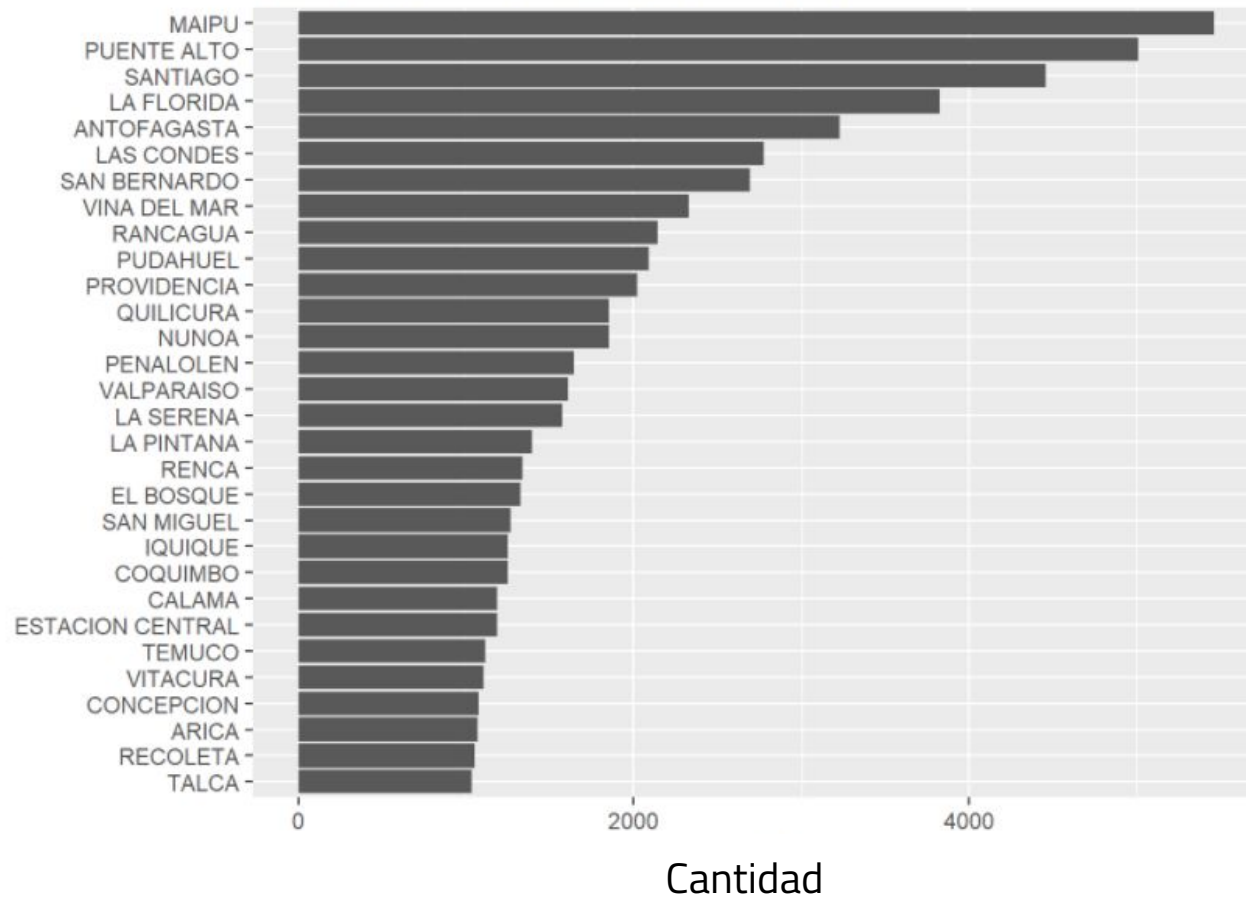


## Cantidad de pedidos por región



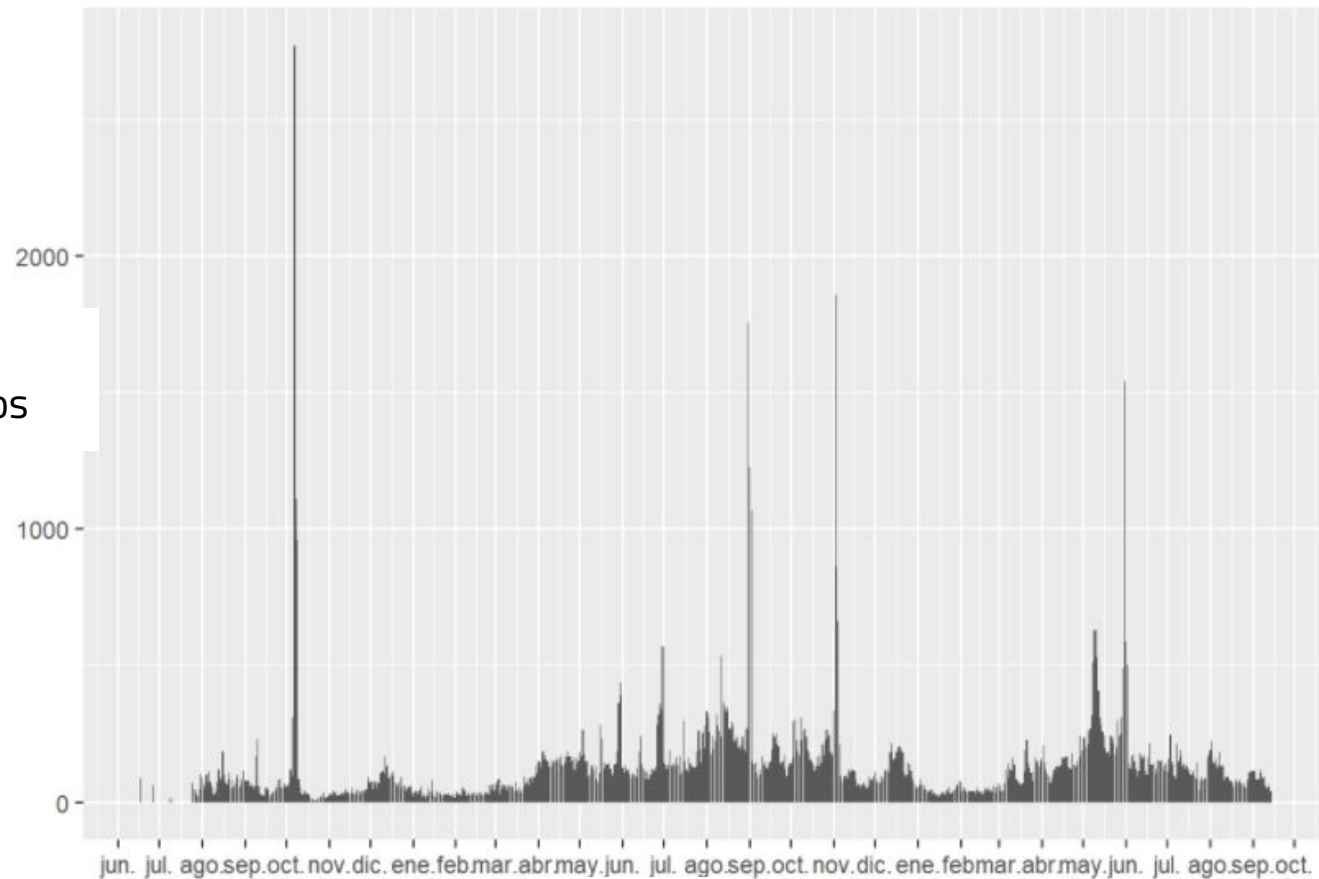


## Las 30 comunas con más pedidos del país



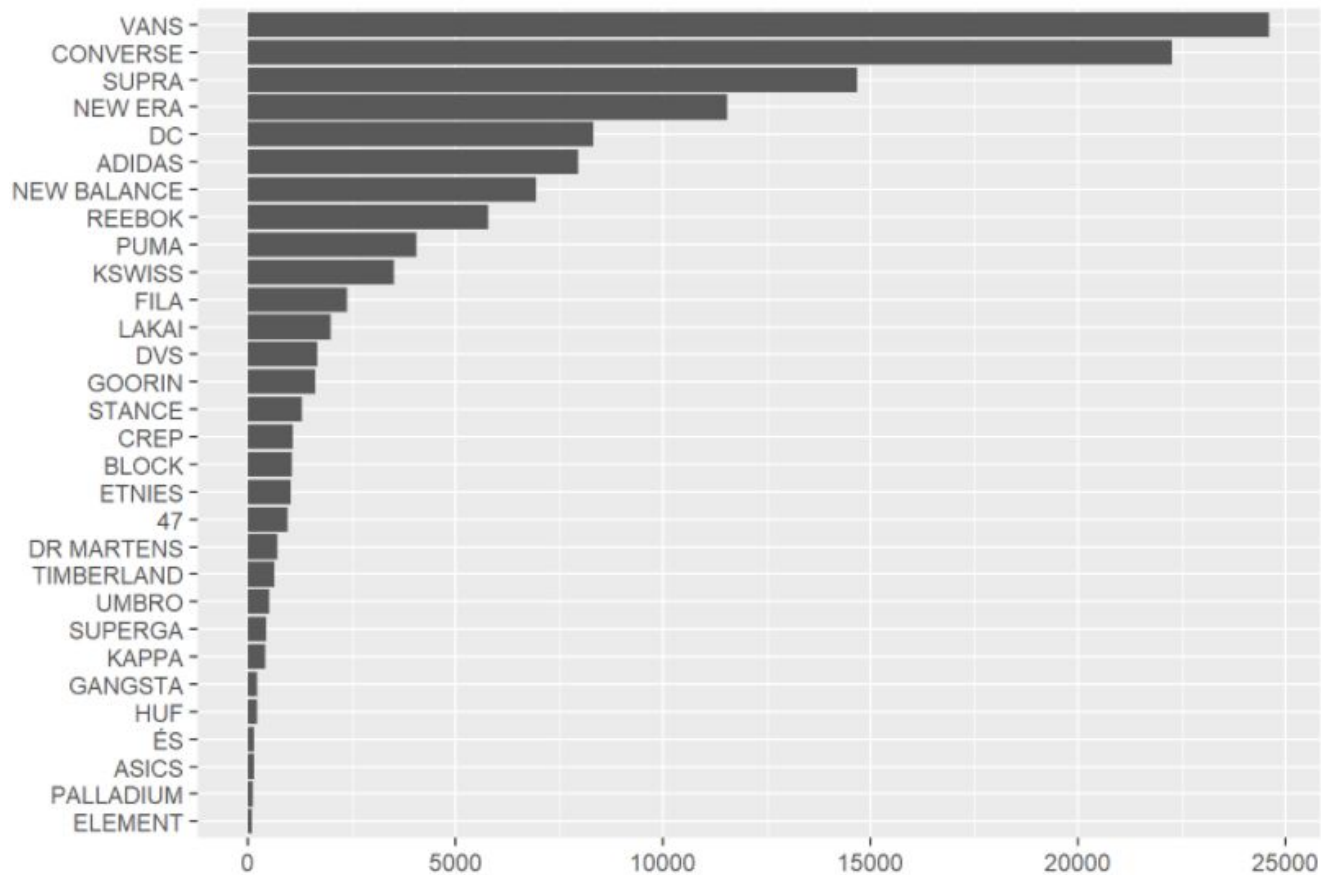
Nombre de  
la comuna

## Cantidad de compras entre Junio 2019 y Septiembre 2021



Mes

## Las 30 marcas con más pedidos del país

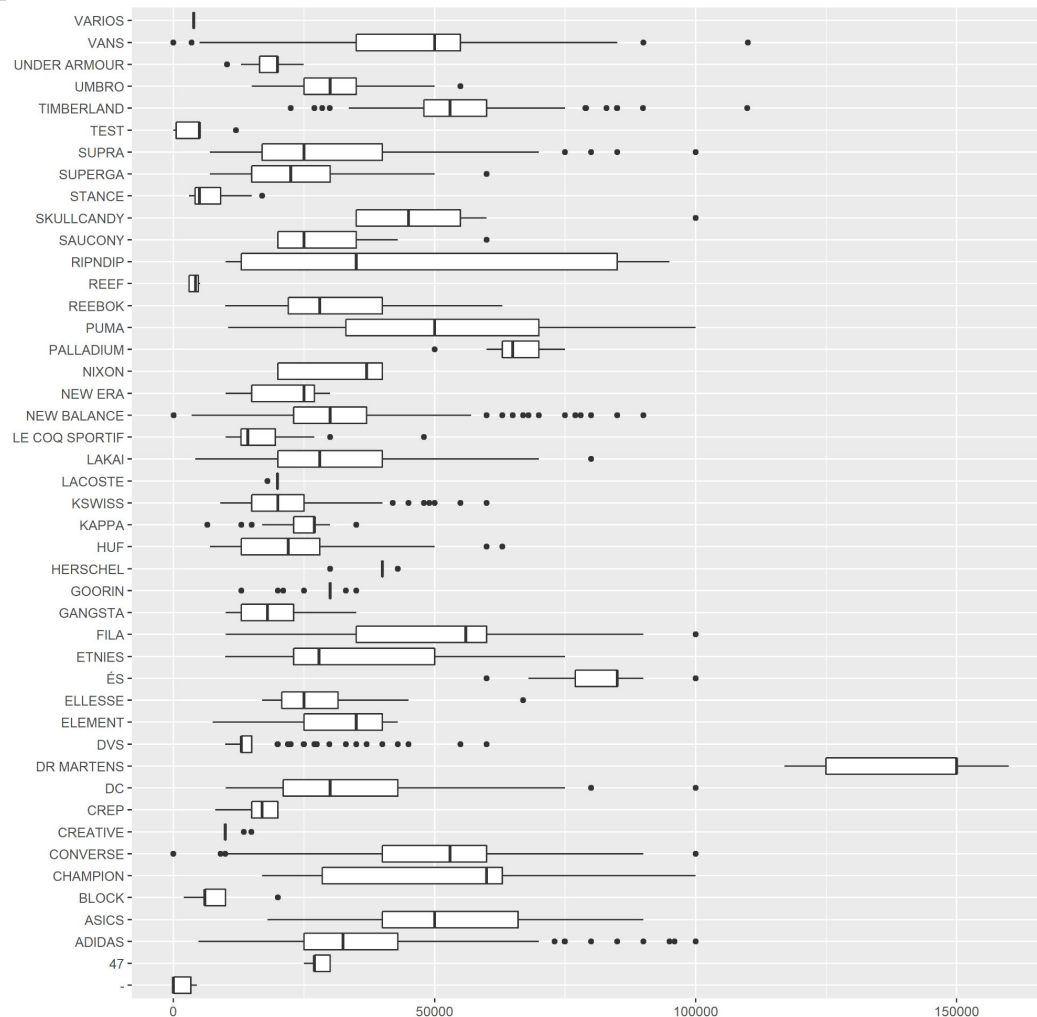


Marca

Cantidad de pedidos

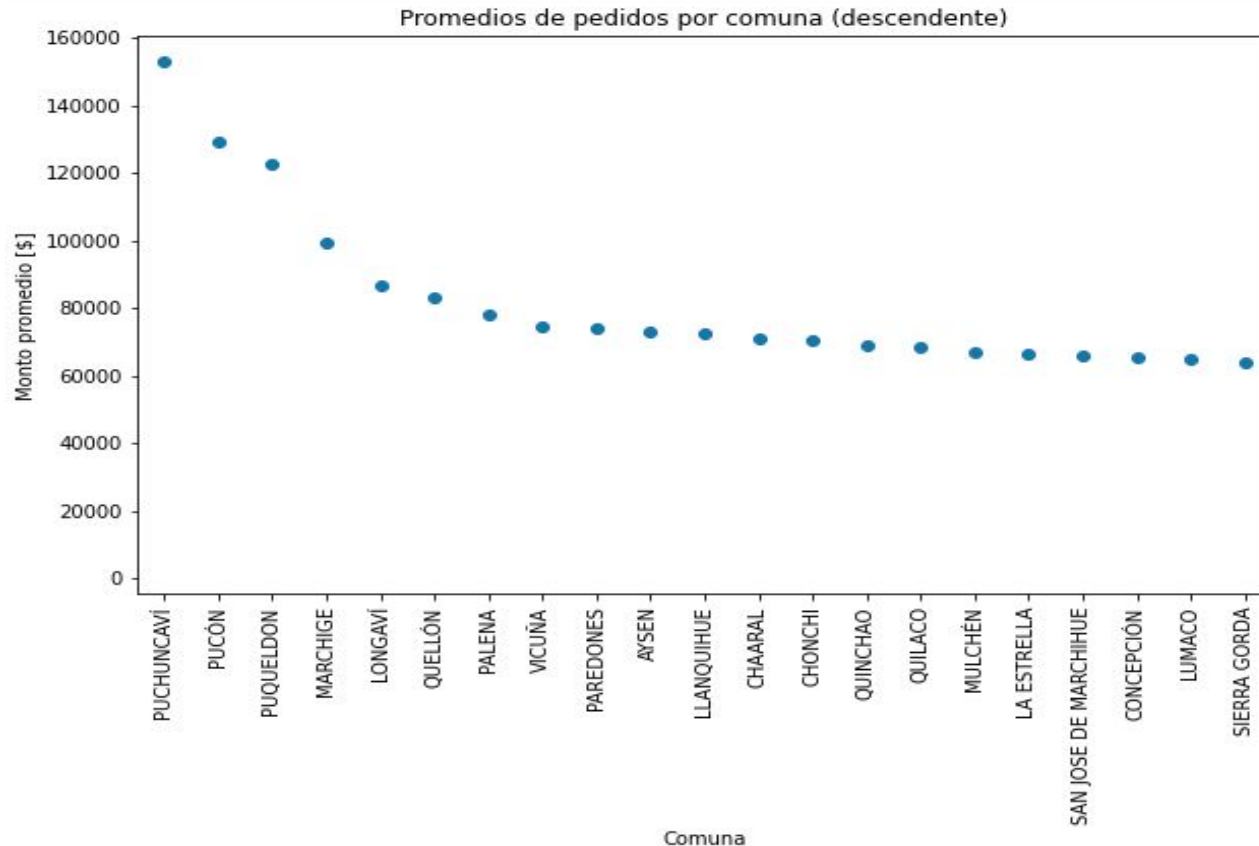
# Variación de precios de cada marca

Marca



Precio

# ¿Cómo se relaciona el precio de la compra con el lugar desde donde se pidió?

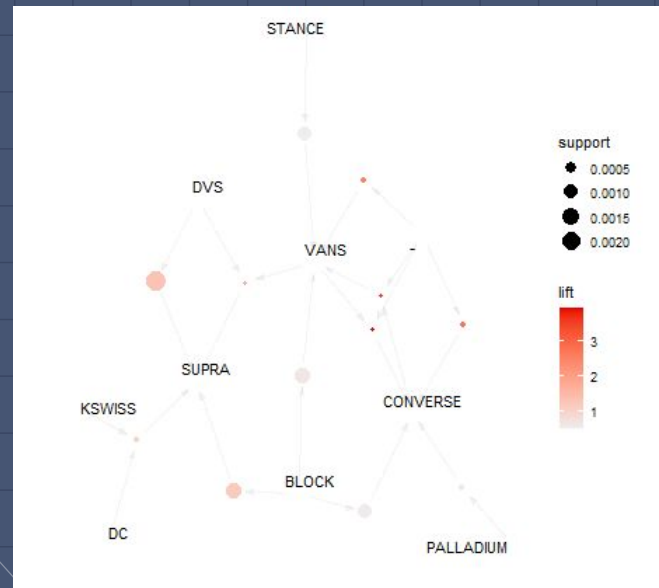
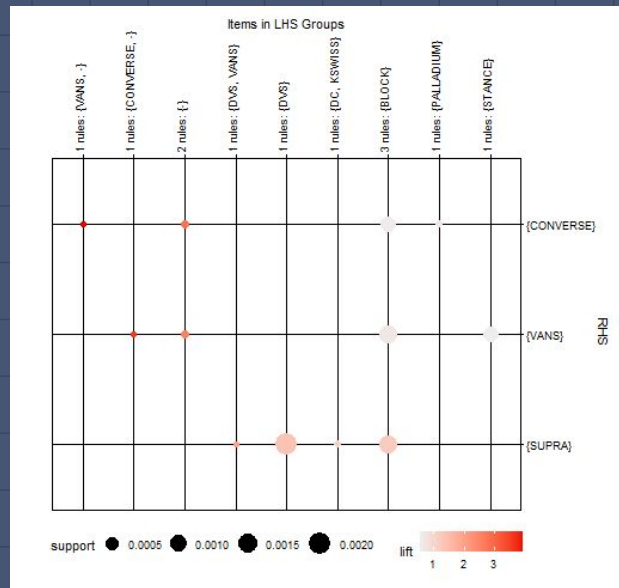


# Resultados Relevantes



# P1: ¿Qué marcas se compran juntas?

lhs	rhs
{-, VANS}	=> {CONVERSE}
{-, CONVERSE}	=> {VANS}
{-}	=> {CONVERSE}
{-}	=> {VANS}
{DVS, VANS}	=> {SUPRA}
{DVS}	=> {SUPRA}
{BLOCK}	=> {SUPRA}
{DC, KSWISS}	=> {SUPRA}
{PALLADIUM}	=> {CONVERSE}
{BLOCK}	=> {VANS}
{BLOCK}	=> {CONVERSE}
{STANCE}	=> {VANS}



## P2: ¿Se puede saber de qué comuna se hizo una compra de acuerdo a su historial?

Método: Clasificación mediante árbol de decisión

33.2% de los datos para testear

66.7% de los datos de entrenamiento

Accuracy en test set: 0.014340344168260038





**P3: ¿Qué tipo de producto puede que compre alguien que ya adquirió x tipo de producto?**



lhs	rhs	support	count	confidence
{Zapatilla SUPRA 1,Zapatilla SUPRA 3}	{Zapatilla SUPRA 4}	0.00104209	89	0.30795848
{Zapatilla SUPRA 2,Zapatilla SUPRA 3}	{Zapatilla SUPRA 4}	0.00101868	87	0.29491525
{Beanie 2}	{Beanie 3}	0.00100697	86	0.20283019
{Zapatilla SUPRA 3}	{Zapatilla SUPRA 4}	0.00134653	115	0.22727273
{Jockey 1,Jockey 2}	{Jockey 3}	0.00207248	177	0.21096544
{Jockey 2}	{Jockey 3}	0.0027633	236	0.1712627
{Zapatilla SUPRA 1,Zapatilla SUPRA 2}	{Zapatilla SUPRA 3}	0.00252913	216	0.20571429
{Zapatilla SUPRA 2}	{Zapatilla SUPRA 3}	0.00345413	295	0.1610262
{Zapatilla CONVERSE 1,Zapatilla CONVERSE 2}	{Zapatilla CONVERSE 3}	0.00160412	137	0.12309075
{Zapatilla VANS 1,Zapatilla VANS 2}	{Zapatilla VANS 3}	0.00177976	152	0.13309982
{Zapatilla KSWISS 1}	{Zapatilla KSWISS 2}	0.00231837	198	0.10301769
{Jockey 1}	{Jockey 2}	0.00982378	839	0.10930172
{Zapatilla SUPRA 1}	{Zapatilla SUPRA 2}	0.01229436	1050	0.10918166

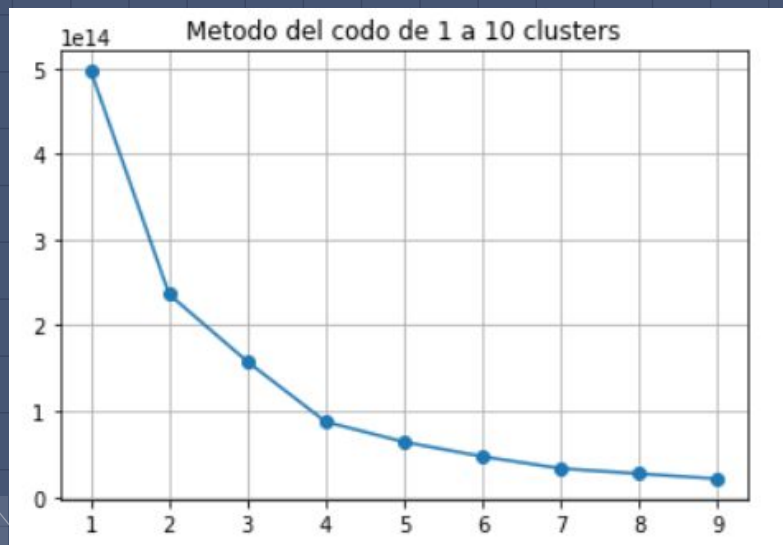
## P4: Comportamiento de clientes respecto a la cantidad de compras y su monto total

Cantidad de pedidos y total gastado por cliente

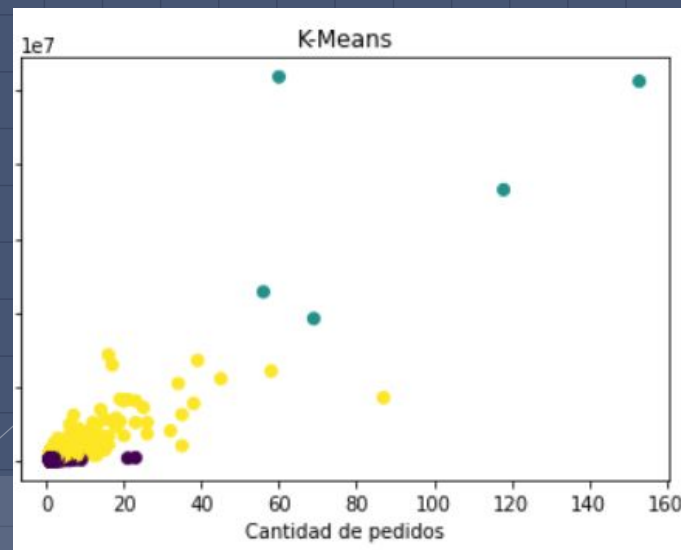
Cantidad Pedidos Cliente	Total Gastado Cliente
13.0	444337.0
1.0	37419.0
1.0	40419.0
1.0	12329.0
14.0	656793.0

# P4: Comportamiento de clientes respecto a la cantidad de compras y su monto total

Método del codo



Clusters



# Futuras direcciones

Hacer una exhaustiva limpieza de datos al inicio del proyecto

Mejorar la organización en el trabajo

Plantear mejores preguntas de investigación dado los conocimientos adquiridos sobre minería de datos durante el semestre



# Gracias!



# ***Block***

