

Casos de Negocio

Antonio Pita Lozano

Máster en Data Science

- **Business Intelligence:** Introducción a las Métricas Robustas.
- **Supply Chain Value:** Forecast de demanda para planificar la producción.
- **Marketing Digital:** Aplicación de diseño de Experimentos.
- **Customer Analytics:** Segmentación de Clientes mediante el modelo RFM.



Introducción a las métricas robustas



*Del Dato
al Conocimiento*

Finalidad

Construir estadísticos robustos

Objetivo

Resumir los datos mediante estadísticos no afectados por la falta de normalidad de los datos o la presencia de valores atípicos u outliers.

Descripción

En las presentaciones de resultados pasados suelen utilizarse estadísticos basados en la media para resumir la realidad.

Estos estadísticos pierden representatividad en las variables económicas que no presentan una distribución normal.

La estadística robusta define estadístico que tratan de resolver este problema.

Datos

- Datos de viviendas como precio, metros cuadrados, etc...

Técnicas Analíticas

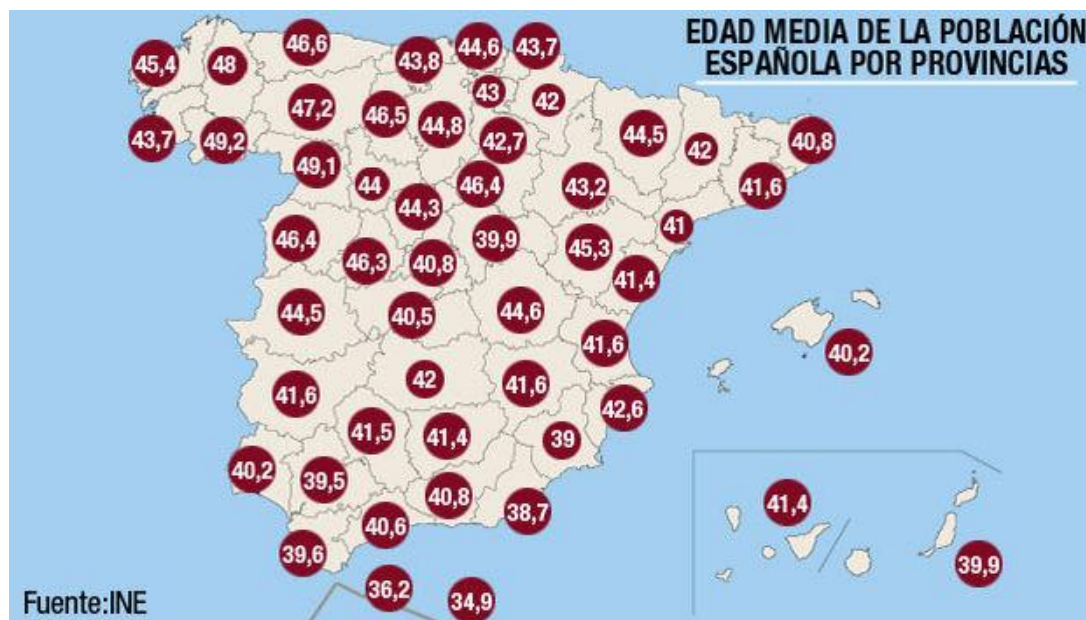
- Mediana
- Media alfa-windsorizada
- Media alfa-recortada
- Regresión robusta
- ...



Si tenemos que presentar datos en un informe

¿Qué estadísticos presentaríamos que representasen a la distribución?

¿Cuál es la edad media de Madrid?



¿Qué significa esto?



Introducción a las métricas robustas



*Del Dato
al Conocimiento*

Los **estadísticos robustos** ayudan a una mejor comprensión ante la presencia de outliers

Estadísticos Clásicos		Estadísticos Robustos
Media Aritmética	Posición	Mediana Media recortada Media winsorizada
Desviación típica	Desviación	Desviación absoluta mediana (MAD) Recorrido intercuartílico

$$IQR = X_{(3n/4)} - X_{(n/4)}$$
$$MAD = \operatorname{median}_i \{ |Y_i - \operatorname{median}_j(Y_j)| \}$$



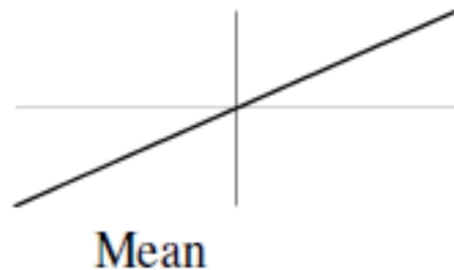
Introducción a las métricas robustas



*Del Dato
al Conocimiento*

El **punto de ruptura** mide la capacidad de robustez de un estadístico como el porcentaje de datos extremos que su variación no afecta al valor del estadístico

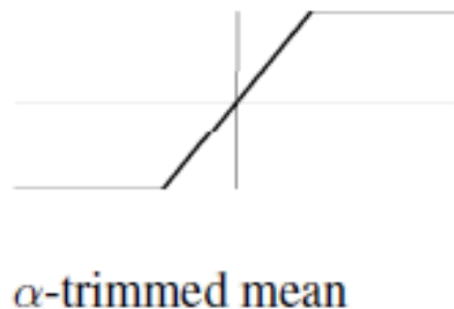
0



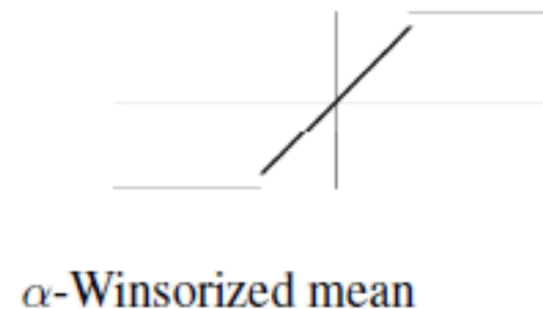
0,5



$0 < \alpha < 0,5$



$0 < \alpha < 0,5$





Introducción a las métricas robustas



*Del Dato
al Conocimiento*





Finalidad

Predicción a futuro

Objetivo

Estimar un modelo que permita realizar una previsión a futuro de una métrica

Descripción

Considerando la información histórica de una métrica determinar un modelo que aprenda las componentes de tendencia, cíclicas y estacionales.

Estos modelos se enriquecen con información histórica de series correlacionadas, principalmente macroeconómicas.

Datos

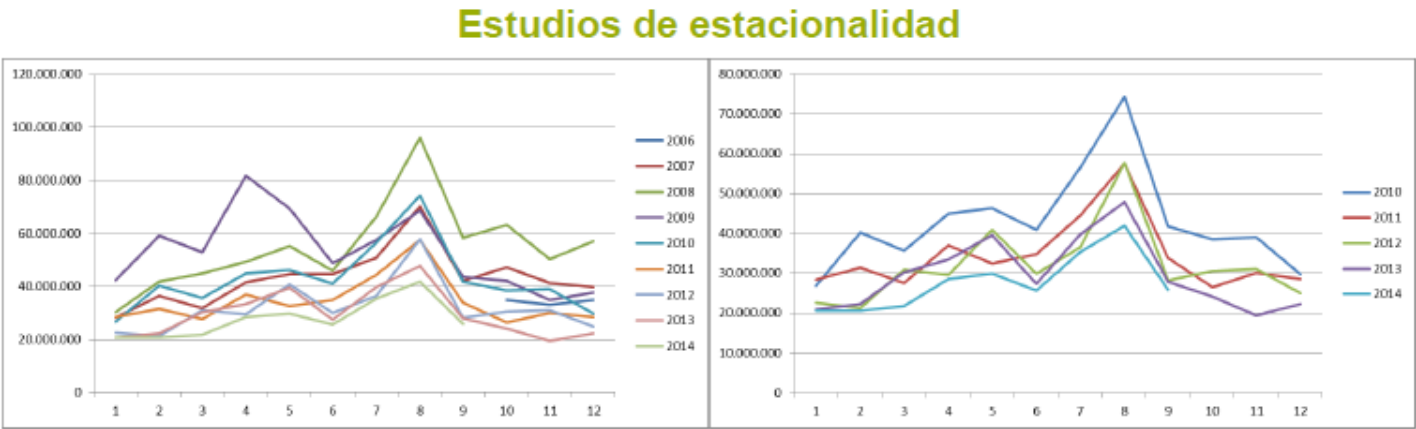
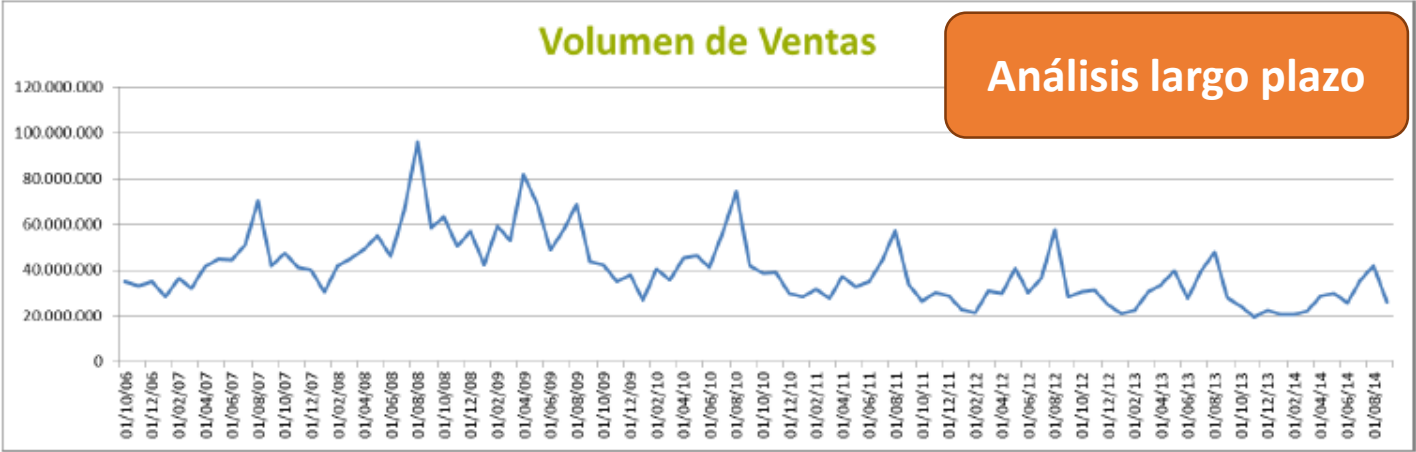
- Serie temporal a estimar
- Series macroeconómicas
- ...

Técnicas Analíticas

- Medias Móviles
- Suavizado Exponencial (Holt Winters)
- Modelos Regresivos (SARIMA)
- ...



Forecast de demanda



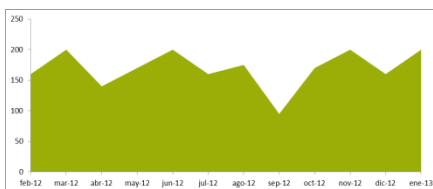
Análisis corto plazo

Franja Horaria	Lunes	Martes	Miercoles	Jueves	Viernes	Total Franja Horaria
09:00 - 09:30	1,51%	1,63%	1,80%	1,54%	1,26%	7,74%
09:30 - 10:00	1,01%	0,92%	1,80%	1,05%	0,99%	5,76%
10:00 - 10:30	0,77%	0,96%	2,05%	0,91%	0,92%	5,61%
10:30 - 11:00	0,76%	0,72%	1,71%	0,81%	0,88%	4,88%
11:00 - 11:30	0,67%	0,95%	1,69%	1,06%	1,00%	5,37%
11:30 - 12:00	0,54%	0,67%	1,29%	0,84%	0,77%	4,11%
12:00 - 12:30	0,72%	0,64%	1,40%	0,94%	0,74%	4,44%
12:30 - 13:00	0,50%	0,74%	1,13%	0,72%	0,63%	3,72%
13:00 - 13:30	0,49%	0,63%	0,97%	0,66%	0,80%	3,55%
13:30 - 14:00	0,53%	0,53%	0,93%	0,76%	0,70%	3,46%
14:00 - 14:30	0,56%	0,67%	0,99%	0,76%	0,69%	3,67%
14:30 - 15:00	0,50%	0,74%	1,03%	0,84%	0,80%	3,91%
15:00 - 15:30	0,69%	0,76%	1,27%	0,94%	0,74%	4,38%
15:30 - 16:00	1,10%	0,86%	1,48%	1,36%	1,10%	5,90%
16:00 - 16:30	1,14%	1,19%	1,78%	1,39%	1,05%	6,55%
16:30 - 17:00	1,21%	1,12%	1,85%	1,49%	1,25%	6,91%
17:00 - 17:30	1,90%	2,00%	2,49%	2,01%	1,78%	10,19%
17:30 - 18:00	1,90%	1,95%	2,17%	1,56%	2,24%	9,82%
Total Diario	16,48%	17,70%	27,82%	19,65%	18,35%	100,00%

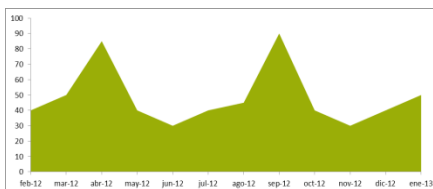
1

Datos de entrada

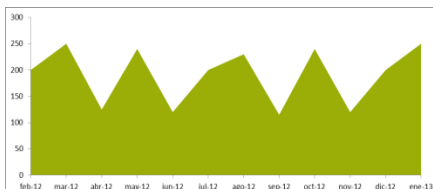
Ventas



Ventas



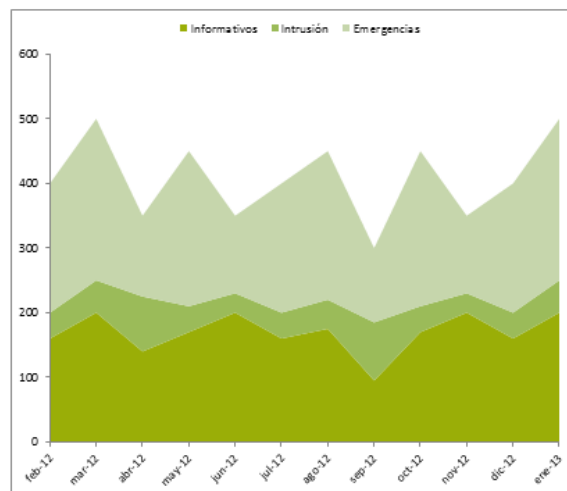
Ventas



2

Modelo estimado

Modelo estimado

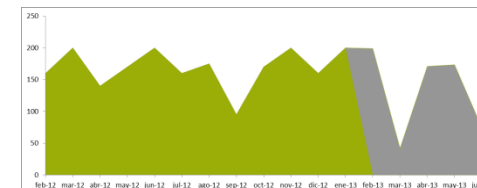


$$z_t = 1.2z_{t-2} - 0.32z_{t-1} + a_t$$

3

Datos de salida

Predicciones a futuro





Forecast de demanda



*Del Dato
al Conocimiento*





Finalidad

Metodología de estudios observacionales para la medición de efectos

Objetivo

Estimar el efecto de una acción realizada mediante un experimento que permita aislar el resultado del resto de efectos, principalmente los no medibles.

Descripción

El efecto de una acción en una variable objetivo puede estar contaminado por factores externos o internos que a su vez pueden ser medibles o no medibles.

Con el objetivo de medir el efecto de la acción, se diseñan experimentos que permitan aislar el resto de factores mediante la comparación de grupos homogéneos cuya única diferencia sea si han sido o no impactados por la acción. Esta comparación se realiza a través de contrastes de hipótesis.

En el caso de grupos no homogéneos, existen experimentos y contrastes más sofisticados.

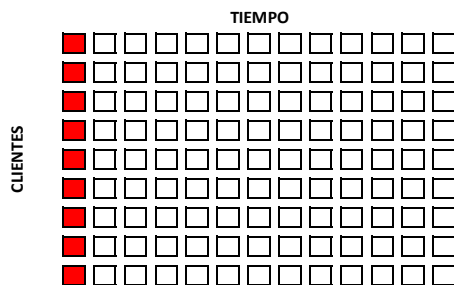
Datos

-

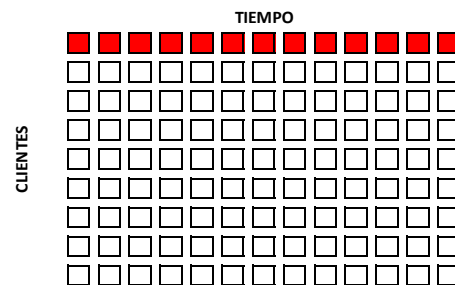
Técnicas Analíticas

- Contrastes de Hipótesis
- Diseño de Experimentos.
- ...

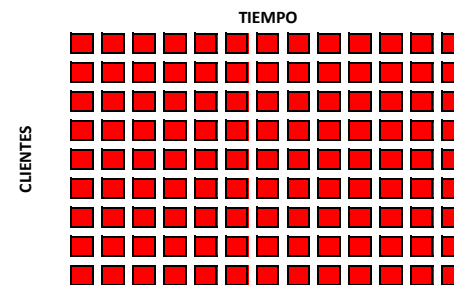
Sección Cruzada



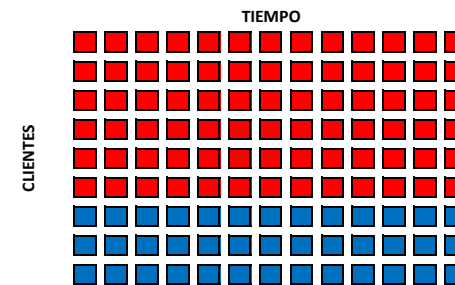
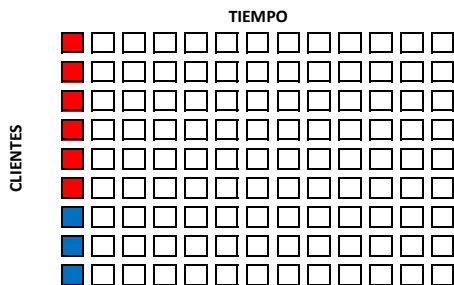
Longitudinales



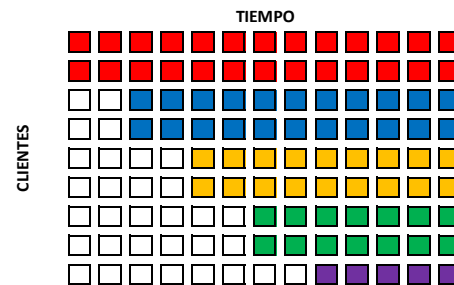
Datos de Panel



Caso - Control



Cohortes

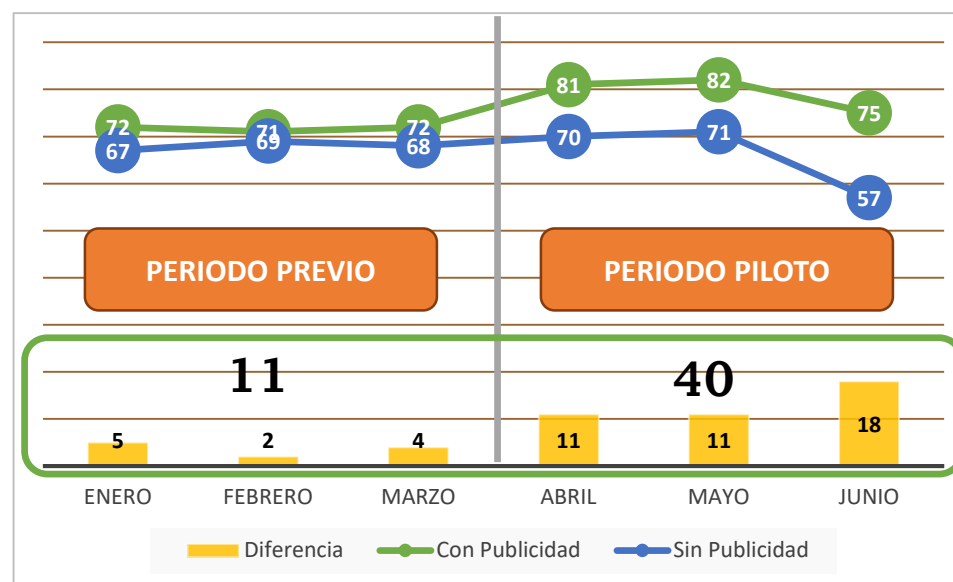




Ejemplos:

- ❖ Efectividad de una campaña.
- ❖ Selección de canal de contacto.
- ❖ Beneficio de una actividad formativa.
- ❖ Efecto de nuevas políticas.
- ❖ Evaluación de modelos predictivos en producción.
- ❖ Satisfacción de los usuarios.

Ejemplo Ilustrativo





Diseño de Experimentos



*Del Dato
al Conocimiento*





Segmentación comportamental RFM



*Del Dato
al Conocimiento*

Finalidad

Seguimiento del cliente

Objetivo

Construir una agrupación de clientes en función al comportamiento que tienen con la empresa

Descripción

El objetivo es clasificar a los clientes mediante su comportamiento utilizando técnicas de aprendizaje no supervisado puesto que los patrones se desconocen.

La información a utilizar debe ser exclusivamente la información relacionada con el comportamiento del cliente con la empresa.

El modelo RFM permite describir el comportamiento del cliente de forma sencilla.

Datos

- Datos transaccionales
- Modelo RFM
- ...

Técnicas Analíticas

- Jerárquico
- K-medias
- K-medianas
- DBSCAN
- ...



4

Segmentación comportamental RFM



*Del Dato
al Conocimiento*

El **modelo RFM**, permite representar los **comportamientos** de los clientes:

Recencia

Días desde la última operación

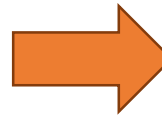
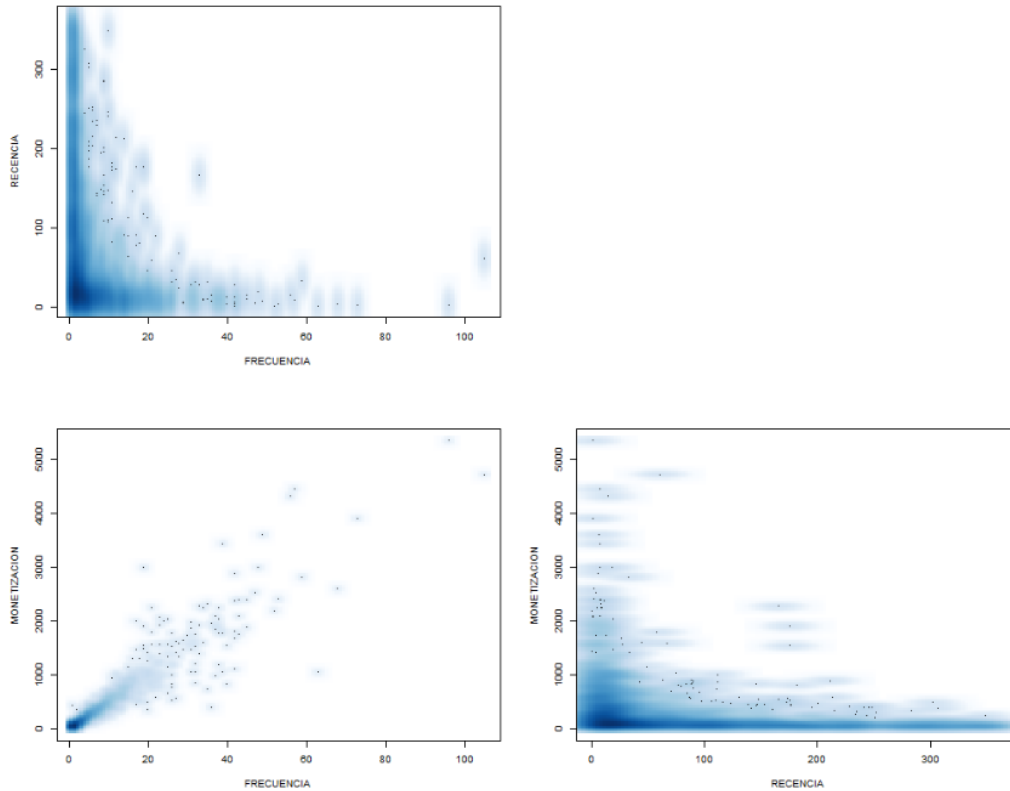
Frecuencia

Número de transacciones en el periodo

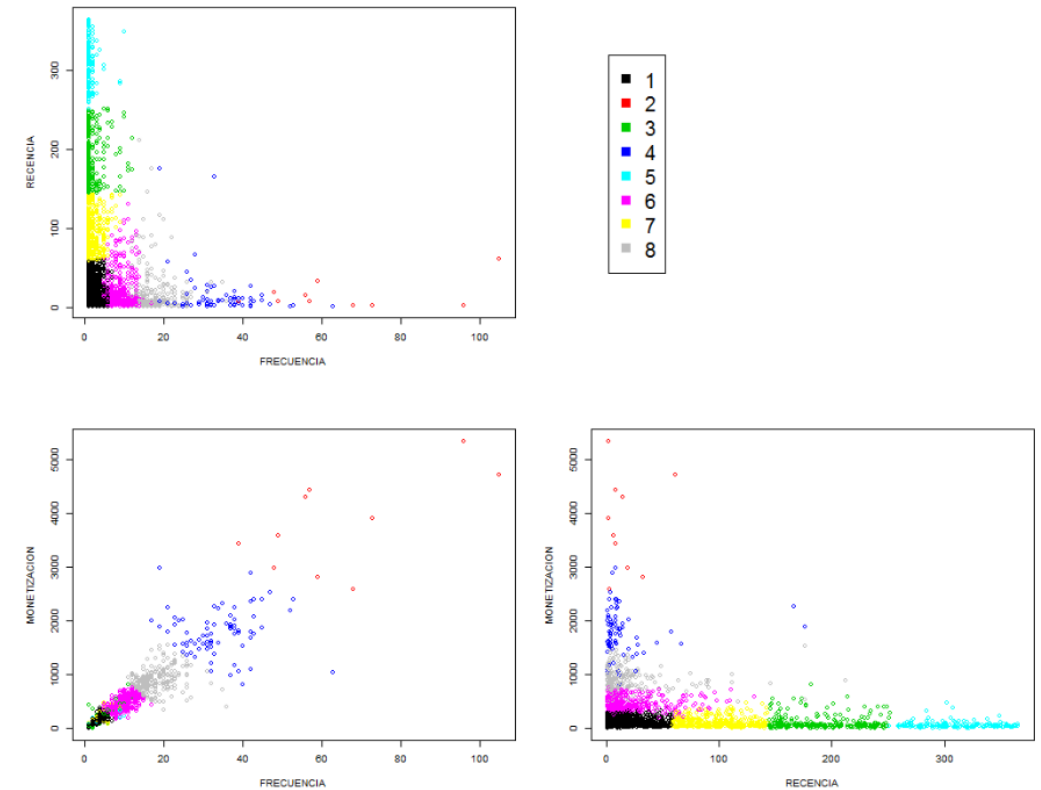
Monetización

Importe de las transacciones en el periodo

Densidad de clientes mediante Modelo RFM 12 meses



Clusterización kmeans de clientes mediante Modelo RFM 12 meses

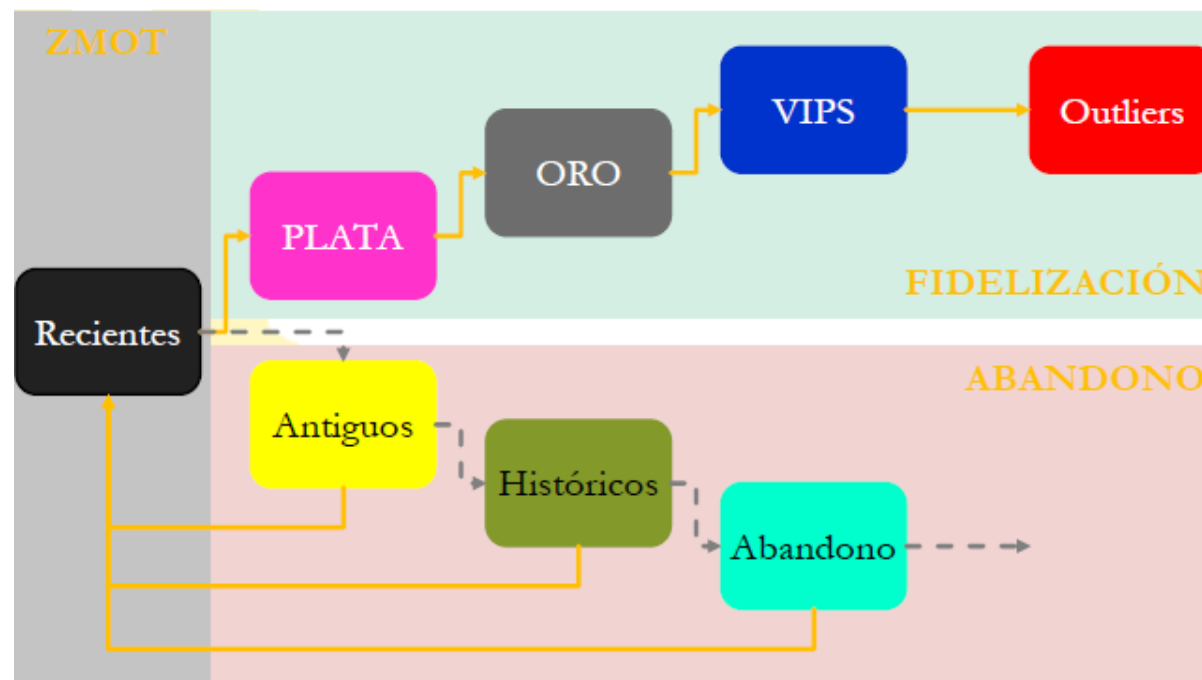




Segmentación comportamental RFM

Representación de los clusters

Segmento	Recencia	Frecuencia	Monetizacion	Contador	Segmento
1	21,76	2,60	127,21	743	NUEVOS CLIENTES
2	15,80	65,00	3811,26	10	OUTLIERS
3	191,81	2,18	109,99	274	HISTORICOS
4	15,00	33,65	1776,75	74	VIPS
5	309,11	1,43	67,66	159	ABANDONO
6	22,87	9,23	453,36	372	PLATA
7	96,68	2,28	113,51	390	ANTIGUOS
8	21,87	18,18	897,63	208	ORO





Segmentación comportamental RFM



*Del Dato
al Conocimiento*



Casos de Negocio

Antonio Pita Lozano

Máster en Data Science

Suerte con el proyecto