

## Introducción

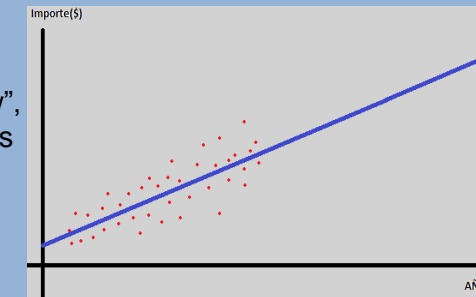
- ¿Donde? → CABA
- ¿Qué? → PUBLICIDAD
- ¿Cuando? → 2016 a 2019
- Objetivos →
- Cuanto dinero se destina a publicidad
  - Como se distribuyen los importes
  - Predicciones para años futuros

## Machine Learning

Aplicamos 3 modelos de Regresión lineal

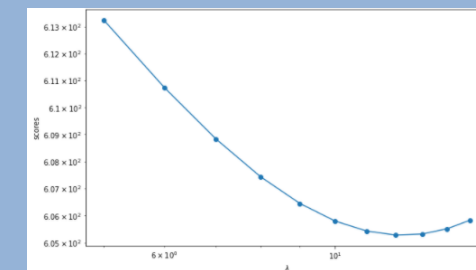
### Regresión Lineal

Utiliza como inputs los meses (X) y los parámetros “w”, y a través de entrenamiento de los datos utilizando los importes, nos dará como salida una función lineal tratando de asemejarse lo máximo posible a los valores reales, en nuestro caso tendremos una recta de pendiente positiva, ya que al avance de los años, el importe es mayor.



### Ridge Regresión

La diferencia de este modelo con el de regresión lineal, es que este modelo utiliza un hiper parámetro (λ) L2, el cual es elegido por el usuario según el que sea más conveniente, penalizando los parámetros “w”.



### SVR

El modelo de Support Vector Regression es el tercer modelo que utilizamos. Se busca el hiper plano que maximice el margen, en el cual se tiene una función de costo (c) en la que se penalizan las muestras que caen fuera del margen.

## Resultados

Luego de aplicar los 3 modelos lineales, observamos los resultados que obtuvo cada modelo con sus errores y el coeficiente de determinación. Estos nos permitieron averiguar cual es el modelo más conveniente para este data set.

Error cuadrático medio (MSE)

$$MSE = \frac{\sum (\hat{y}_i - y_i)^2}{n}$$

Media del error (MAE)

$$MAE = \frac{|\sum (\hat{y}_i - y_i)|}{n}$$

Coeficiente de determinación

$$R^2 = \frac{TSS - RSS}{TSS}$$

Resultado obtenido

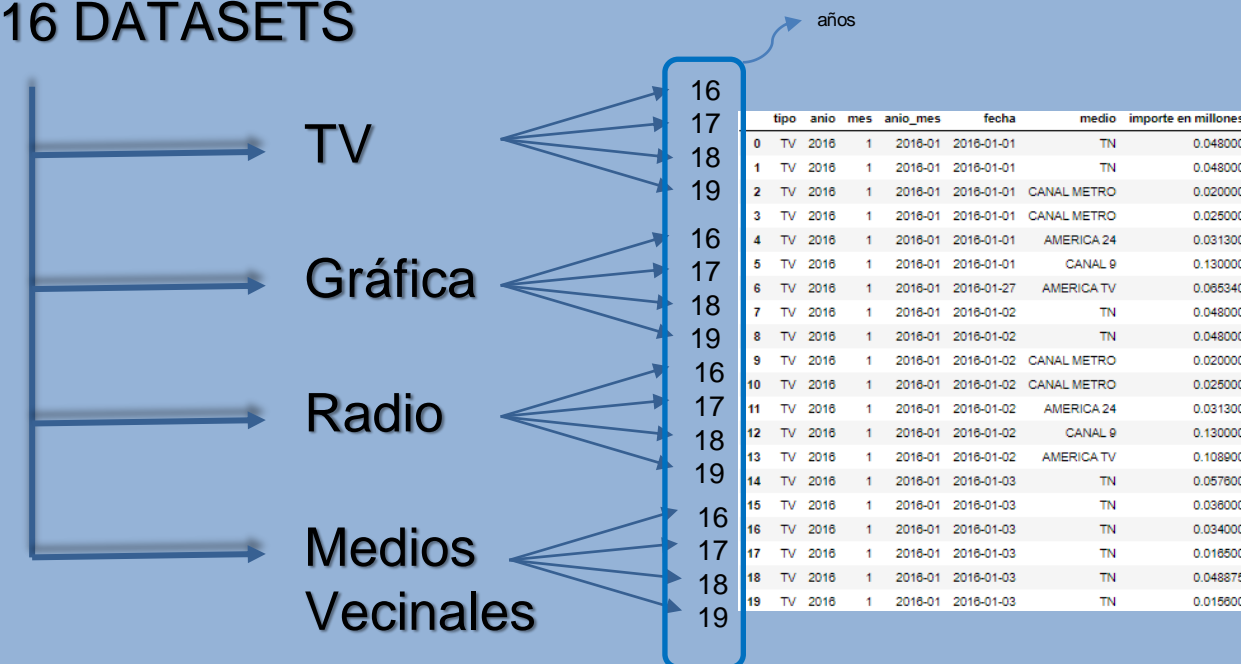
Model	Features	R2	MSE	MAE
Linear	Lineal	0.086	848.999	23.757
SVR	Lineal	0.183	759.165	21.520
Ridge	Lineal	0.085	850.431	23.623

Predicciones a futuro

DATOS REALES	PREDICCIONES
> 2016: 272.3 Millones	> 2020: 814.5 Millones
> 2017: 408.4 Millones	> 2021: 954.5 Millones
> 2018: 470.6 Millones	
> 2019: 686.1 Millones	

## Datasets

16 DATASETS



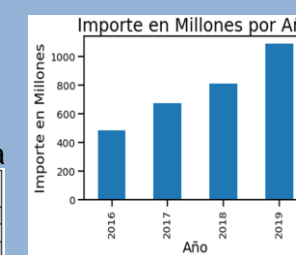
## Análisis Exploratorio de datos

El objetivo de nuestro análisis de datos fue poder hallar mediante gráficos cuanto se invierte según cada tipo de publicidad y detectar variaciones obteniendo posibles hipótesis que las justifiquen.

### Análisis anual

Podemos ver un incremento del importe en millones de publicidad total por año. Sin embargo viéndolo en términos reales hubo caída con respecto a la inflación en el año 2018 como 2019

Año	Incremento de Importe	Inflación*
2017	38%	24.80%
2018	20%	47.62%
2019	34%	53.83%



### Análisis mensual

Vemos que existen picos importantes en mayo de 2019. Haciendo un análisis de esta situación podemos sacar posible hipótesis de que esto se debe a que al mes siguiente fueron las elecciones PASO. Por lo que en ese mes el gobierno de la ciudad gastó casi \$200 millones. Lo mismo podemos notar en el pico de septiembre de 2019, también mes previo a las elecciones



### Análisis por tipo de publicidad

En este caso podemos ver lo que era de esperarse, el tipo de publicidad que más importe conlleva es la televisión ya que es el medio donde mayor audiencia hay teniendo un Costo mayor de publicidad.



## Conclusiones

A la hora de analizar las publicidades actuales ya existentes, podemos notar que mientras transcurren ciertos fenómenos en la sociedad se puede observar un aumento o disminución del importe en la publicidad.

Por otro lado, estamos seguros que este predecir el año 2020 es una tarea complicada ya que ha sido un año atípico, y este nuevo escenario global es impredecible por cualquier modelo de Machine Learning.

Por otro lado, teniendo en cuenta las predicciones realizadas con los modelos lineales podemos llegar a la conclusión de que la alta variabilidad de los datos mes a mes, generan un gran error al utilizar modelos sub-ajustados no permitiendo que nuestra variable independiente explique con efectividad la variación de la variable dependiente.

De todas formas, al analizar las predicciones a futuro podemos notar que son números que a simple vista son razonables, y nos da una leve noción de que el modelo sigue un criterio coherente a la hora de predecir.