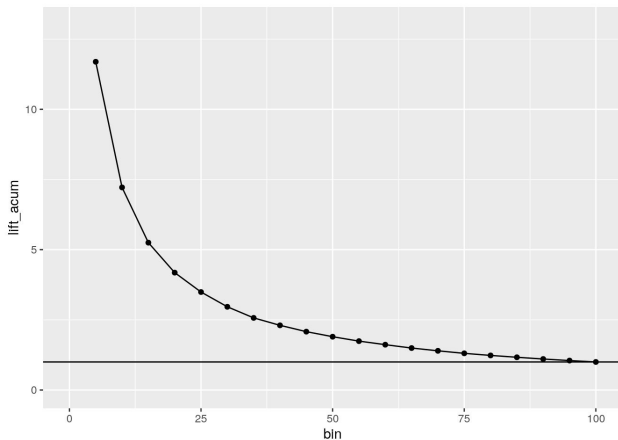


Conectando con el Negocio

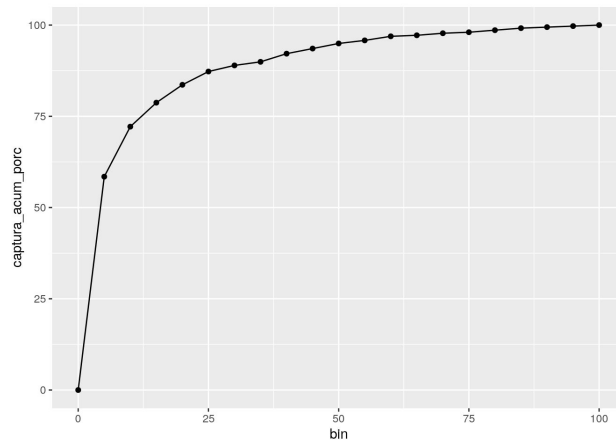
Data Mining Economía y Finanzas - 2021 - Comisión 1

Alejandro Bolaños

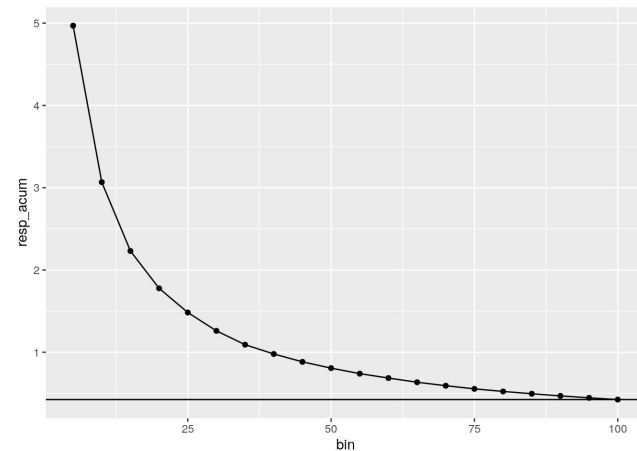
Métricas → Usos y Expectativas



Lift: Cuánto mejora la probabilidad de encontrar una **baja** en los **x** mejores del modelo, con respecto al tomar una muestra al azar



Captura: Qué porcentaje de **bajas** se captura si se toman el mejor **x** porcentaje según el modelo.



Respuesta: Qué porcentaje de **bajas** se espera entre los **x** mejores según el modelo.

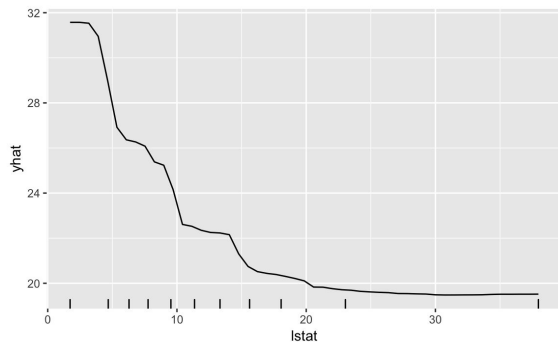
Interpretabilidad → Confianza + Insights

Globales

Variable Importance

Influencia de la variable en el modelo

Partial Dependence Plot

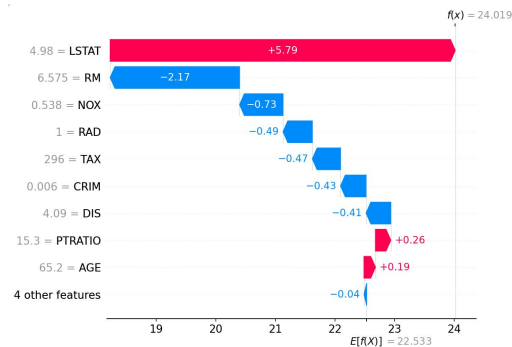


Locales

ICE

Similar al PDP, pero para un caso

LIME / (*) SHAP



Además

- EDA
- Visualizaciones
- Segmentos / Profiling
- Prototipos

Comunicando al Negocio ... y al Mundo

Recolección de recomendaciones para presentar

No es lo mismo...

Presentar en una conferencia / TED

Dar un demo

Elevator pitch

Presentar online / presencial

Presentar a persona con tu misma cultura / extranjero

....

En cada caso existen guías y recomendaciones, que es importante tener en cuenta.

Una historia de 2 cerebros

No se trata sólo de dar un mensaje. Es importante cómo darlo entendiendo que hay barreras biológicas y culturales entre los oradores y oyentes que pueden reducir la calidad de lo que se trata transmitir:

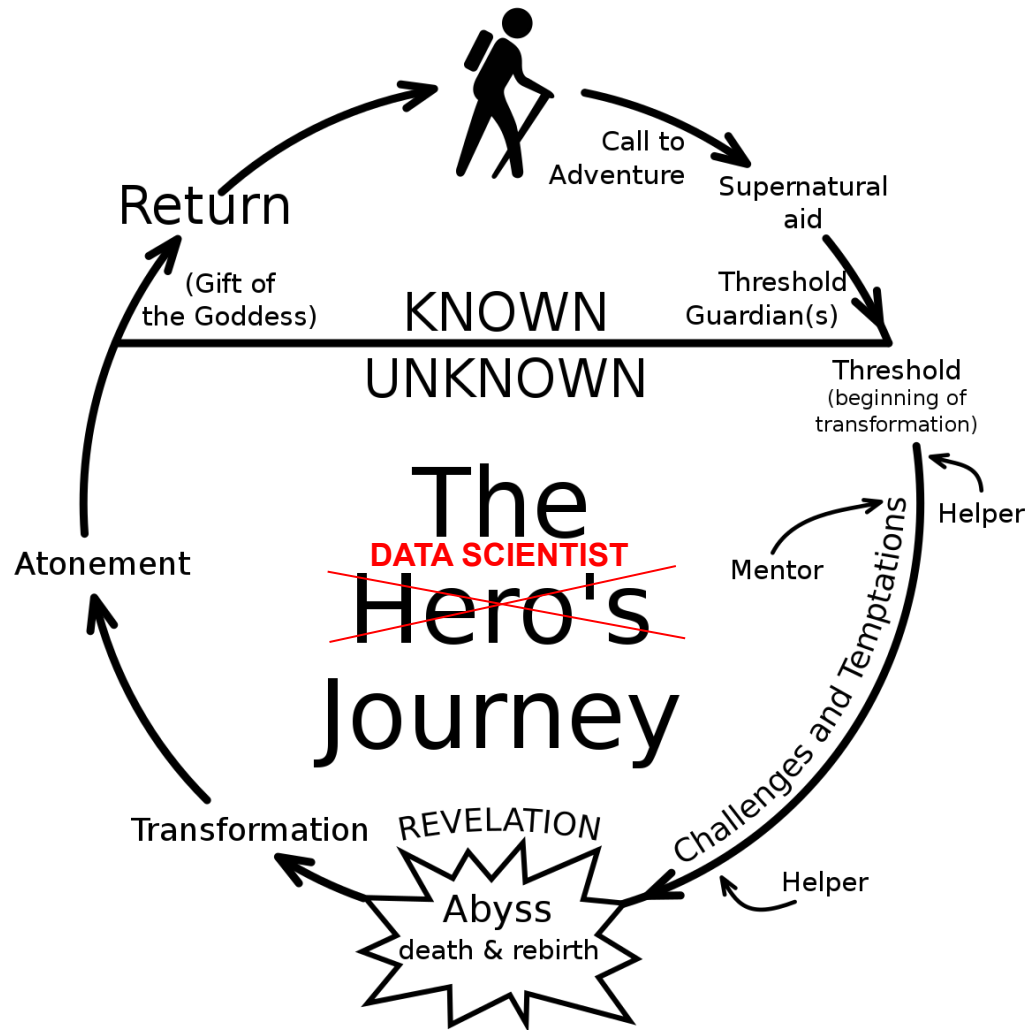
- Sesgos previos residentes en los oyentes
- Falta de atención
- Saturar de información
- Dificultad en la expresión
 - Frases muy complejas
 - Latiguillos
 - Lenguaje corporal inadecuado
- Capacidad de absorción de conceptos

Entre otros...

Por eso es muy importante...

- Armar la presentación PENSANDO en el oyente. Conocer al oyente lo más posible y usar ese conocimiento para conectar.
- Hacer foco en POCOS puntos. Se recomiendan no más de 3, y que toda la presentación esté siempre alrededor de estos 3 puntos. Se DIRECTOS.
- Simpleza ante todo.
 - Hablar lo justo, frases precisas
 - No mucho texto en las diapositivas (haz lo que digo...)
 - Gráficos claros: Data-Ink Ratio de Tufte
- Ser agradables
- STORYTELLING es la forma que ancestral se transmitía información, lo tenemos incorporada.

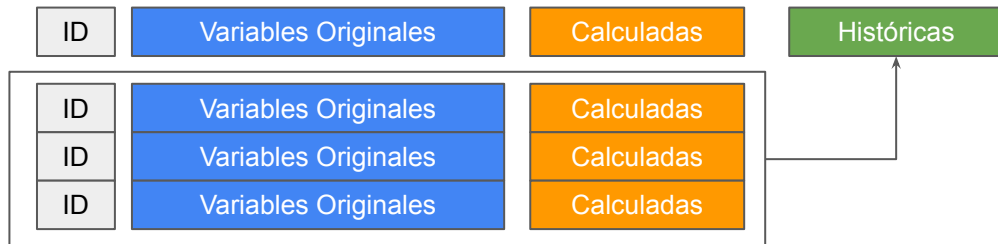
<https://forbes.co/2020/08/21/actualidad/jeff-bezos-prohibe-usar-powerpoint-en-reuniones-de-amazon/>



Sobre los errores y el tiempo

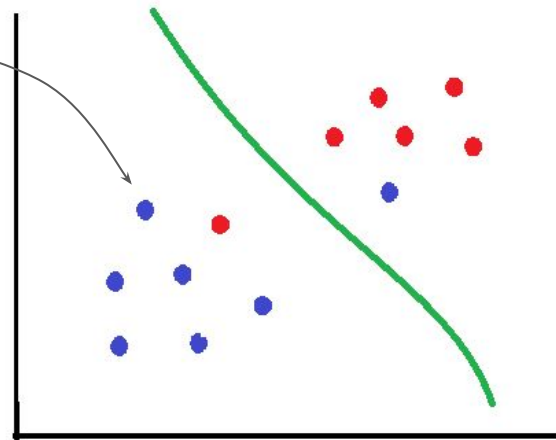
El tiempo da más información, nos ayuda a probar mejor... y es un bien escaso.

El tiempo da más información



Objetivo: Sumar info consolidada de los casos a través del tiempo, ejemplos:

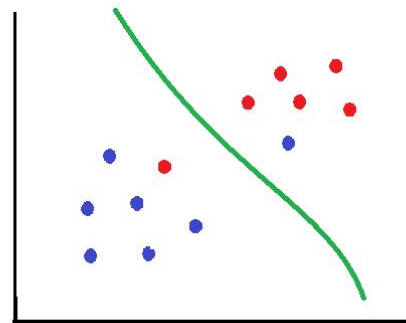
- Lags
- Deltas
- Max
- Min
- Medias
- Tendencias
 - Beta
 - Ratios
- Propuestas?
- Modelos históricos?



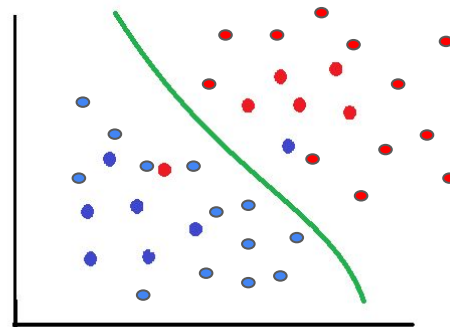
https://github.com/dmecoynfin/dmeyf/blob/master/src/bftf/951_dataset_epic.r

El tiempo da más información

202009	Variables
--------	-----------



202011	Variables
202010	Variables
202009	Variables
202008	Variables
202007	Variables
202006	Variables



El tiempo da más información... trae sus problemas

Crece rápidamente el tamaño de los datos

- Las variables históricas multiplican por N horizontalmente
 - Las variables importantes se repiten más
- Sumar meses multiplican verticalmente

Debatamos las posibles **soluciones**

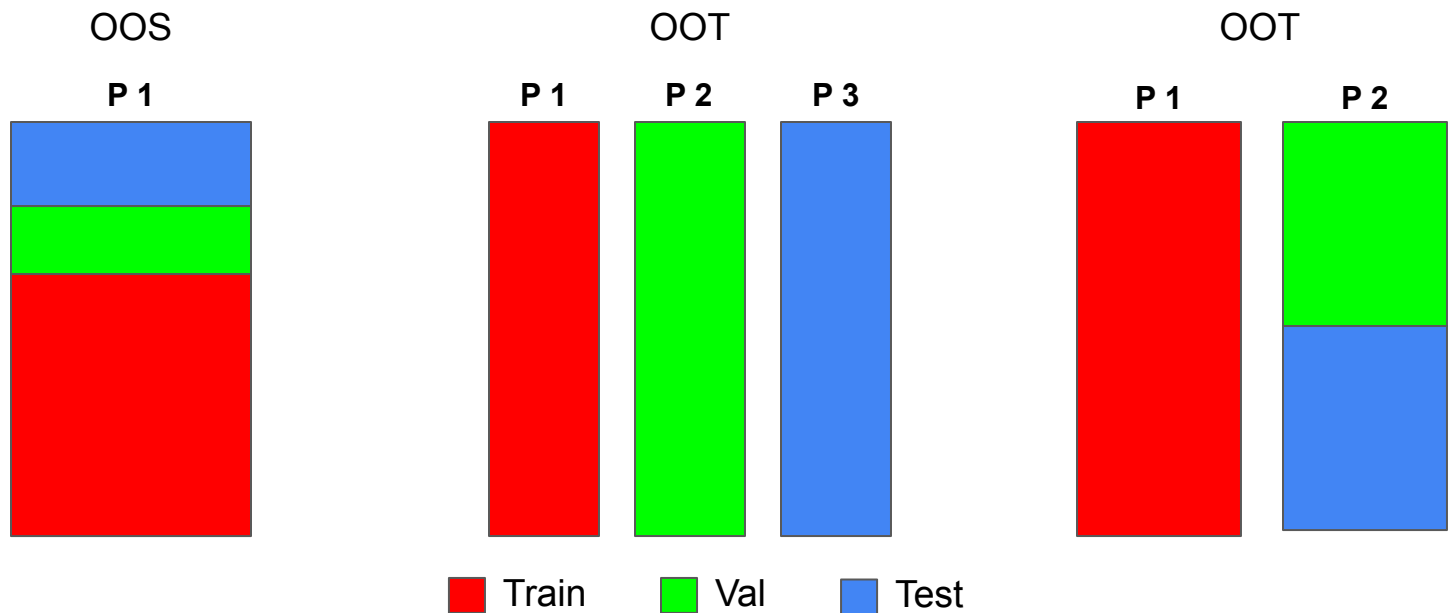
El tiempo da más información... trae sus problemas

Sumar meses mezclan con distintas características estacionales y con efectos temporales como la temible **INFLACIÓN**

Debatamos las posibles **soluciones**

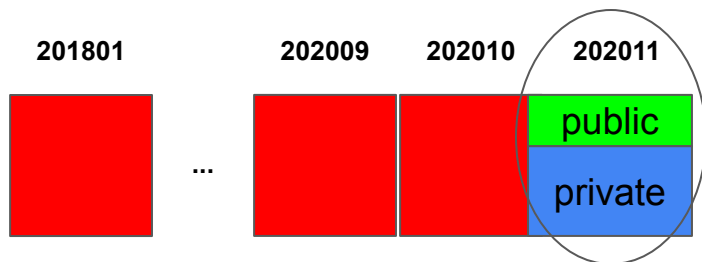
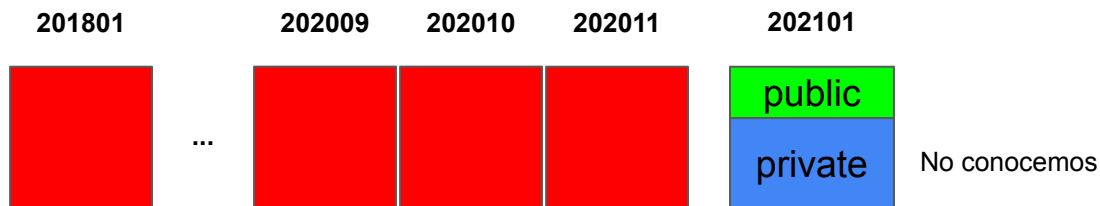
El tiempo nos ayuda a probar mejor...

Pasamos de medir y ajustar de un modelo **Out of Sample** a un **Out of Time**



El tiempo nos ayuda a probar mejor...

Nos permite en otro espacio de tiempo, repetir las condiciones y probar las estrategias de la competencia



- Medir efectos de las diferencias estrategias.
- Medir la variabilidad de los modelos
- Generar confianza y validez de estrategias más sofisticadas y difíciles de trabajar como el stacking

El tiempo es un bien escaso...

Tienen que utilizarlo sabiamente.

- Los modelos tardan mucho más tiempo.
- Medir correctamente es más importante y esto eleva los tiempos

