

Analítica Predictiva y el Teorema de Bayes

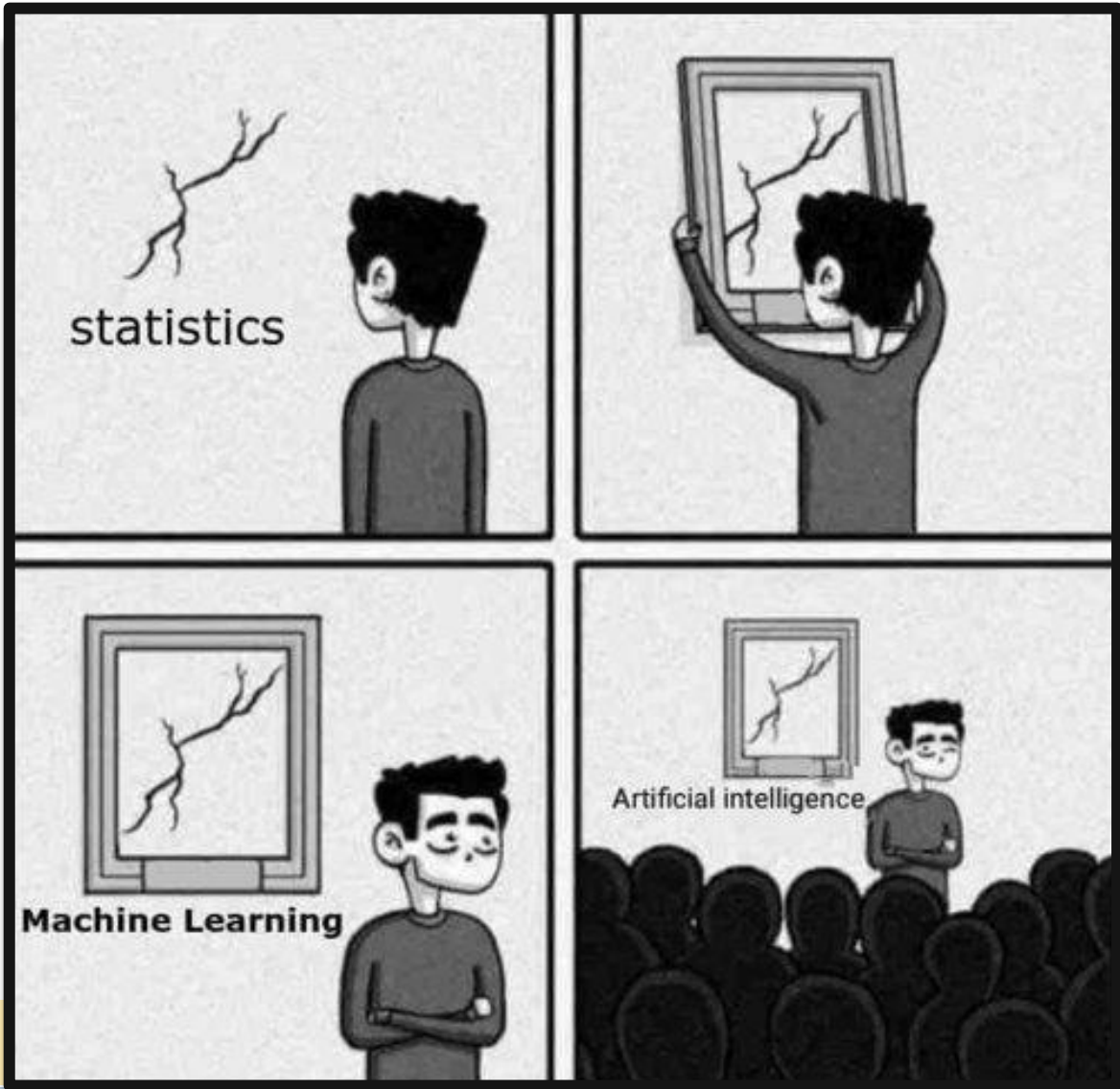
31 de mayo y 8 de junio 2019

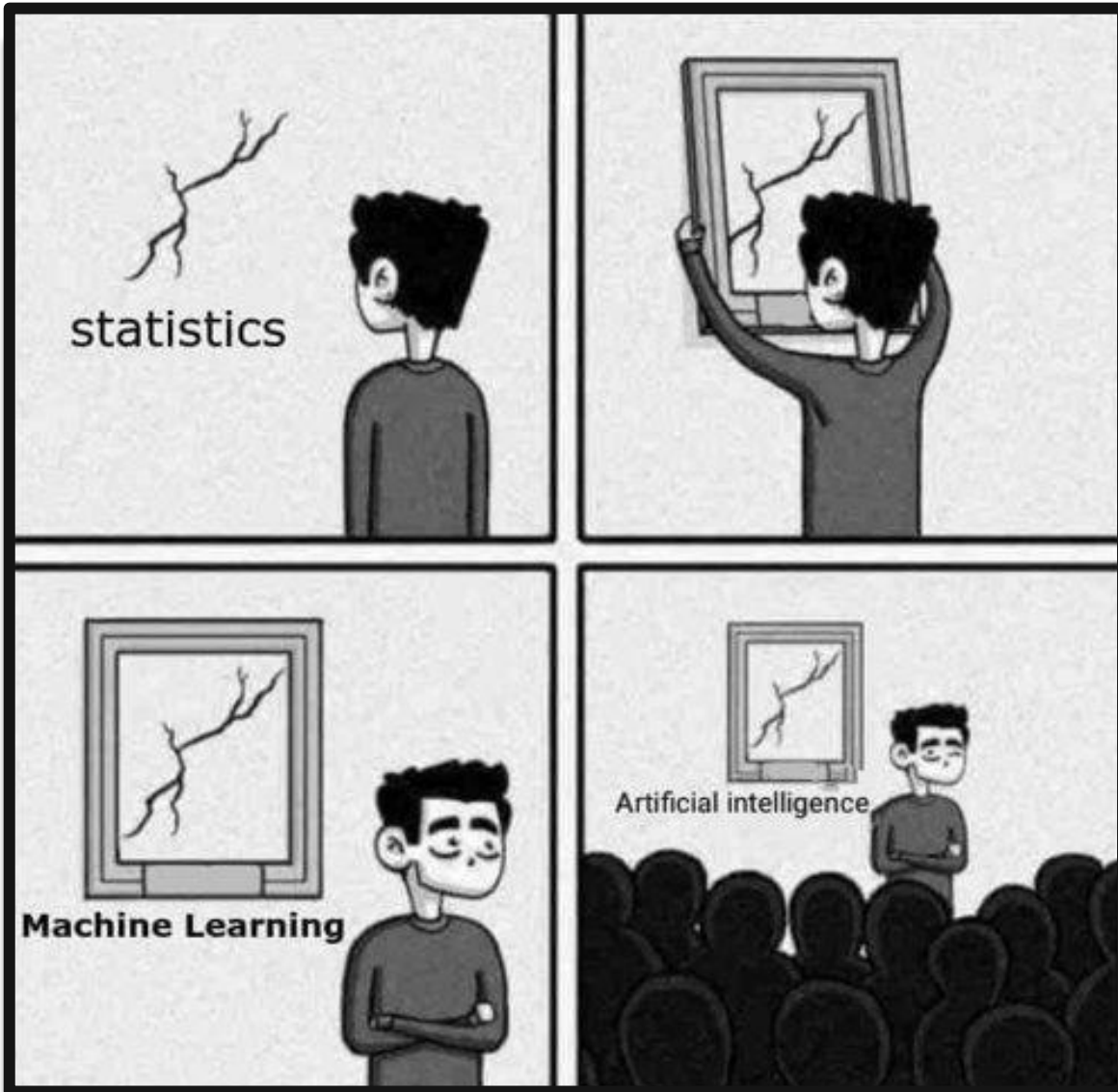
URJC - MASTER EN DATA SCIENCE



Romy Rodriguez-Ravines

romy.rodriguez@innova-tsn.com





David Smith

@revodavid

Follow



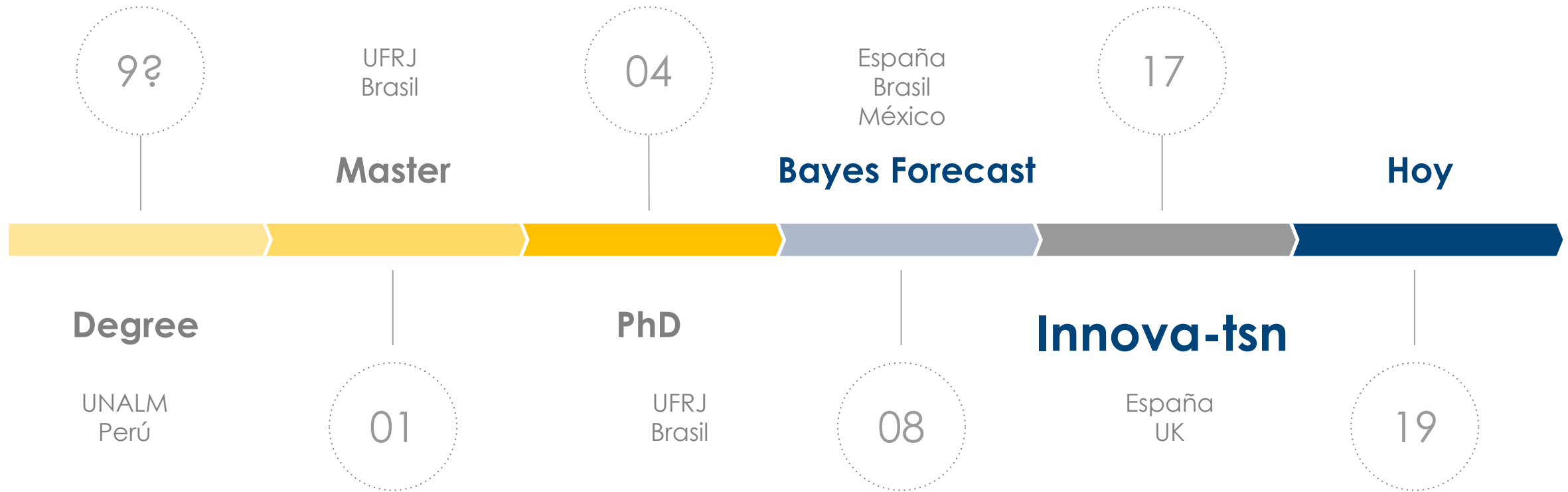
That does it. I am officially switching back to "Statistician".



7:36 PM - 28 May 2018

← **LEARNING** STATISTICS →

EXPERIENCING STATISTICS →





inova-tsn

15 años

3 ciudades

+220 consultores

Soluciones a medida

+200 proyectos/año

5 áreas

- Solutions Architecture
- Client Intelligence & Digital Marketing
- Business Discovery
- **Advanced Analytics & Big Data**
- In Labs



Analítica Avanzada y Seguro: perspectivas desde Dirección

innova-tsn



Regional
Partner of the Year
SE EMEA
2018



Premios Tecnología Siglo XXI 2019
INNOVA-TSN
Business Intelligence



Halifax, Nova Scotia - Canada
August 13-17, 2017

1º EMEA & 7º GLOBALLY

Highway Tollgates
Traffic Flow Prediction in China





DETECCIÓN DE FRAUDE

- Classification
- Gradient Boosting Machine
- SAS, R, Spark



RESERVA DE BILLETES

- Demand Forecasting
- GBM, k-NN, Quantile regression
- SAS, R



DEMANDA DE ENERGÍA

- Demand Forecasting
- ARIMA, Neural Networks, Clustering
- SAS, Python, Spark, AWS



AUDIENCIA DE PUBLICIDAD

- Demand Forecasting
- GBM,
- SAS, R, H2O



PRECIO ÓPTIMO

- Demand Forecasting, Optimization
- ARIMA, GBM, Hill Climbing
- SAS, R



MANTENIMIENTO PREDICTIVO

- Classification
- Gradient Boosting Machine
- R, Qlik Sense



EXPERIENCIA DE CLIENTE

- Classification
- Gradient Boosting Machine
- SAS, R



VENTA CRUZADA

- Topic Modelling
- Text Mining
- R



RECOMENDACIÓN DE CUPONES

- Clustering
- k-NN
- Python, Spark



CLASIFICACIÓN DE SOLICITUDES

- Documents Classification
- Text Mining, Deep Learning
- Python

DANDO FORMA A LAS IDEAS

“

Ayudamos a nuestros clientes a tomar las mejores decisiones basadas en la información proporcionada por nuestras soluciones.





www.innova-tsn.com



<https://twitter.com/innovatsn>



<https://www.linkedin.com/company/innovatsn>



[Canal YouTube Innova-tsn](#)

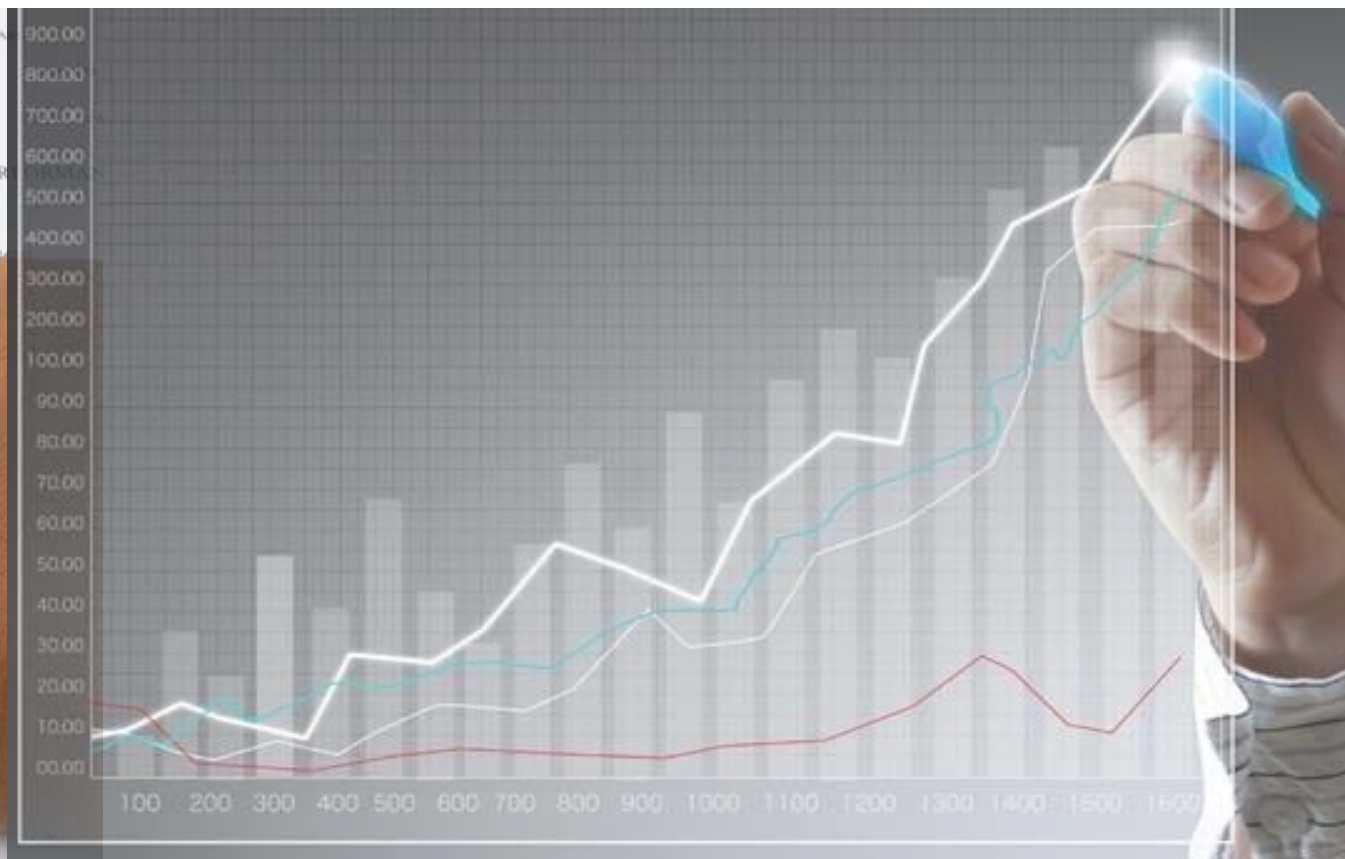
Si quieres formar parte de nuestro equipo, **envíanos tu CV**
recruiting@innova-tsn.com



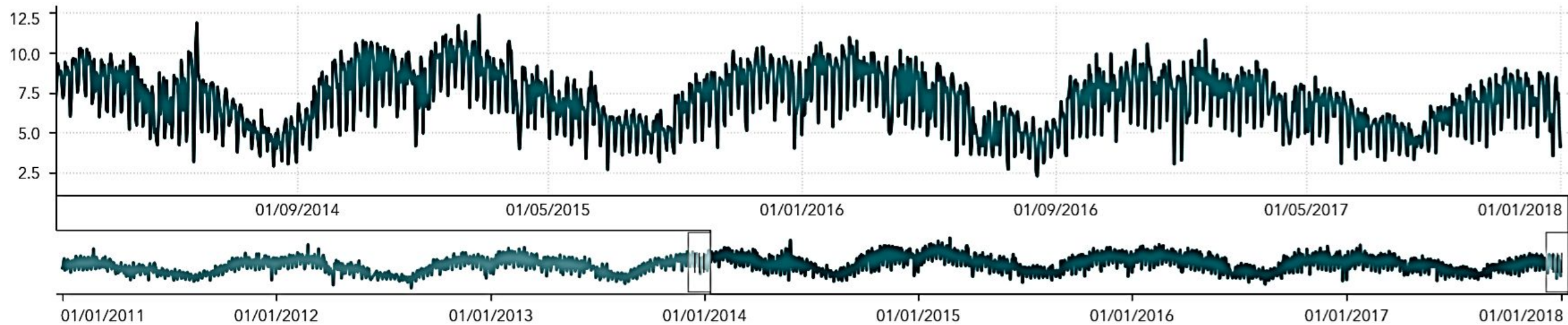


¿Cómo predecir la demanda y actualizar nuestro conocimiento?

Objetivo: Demanda



Datos: Series Temporales

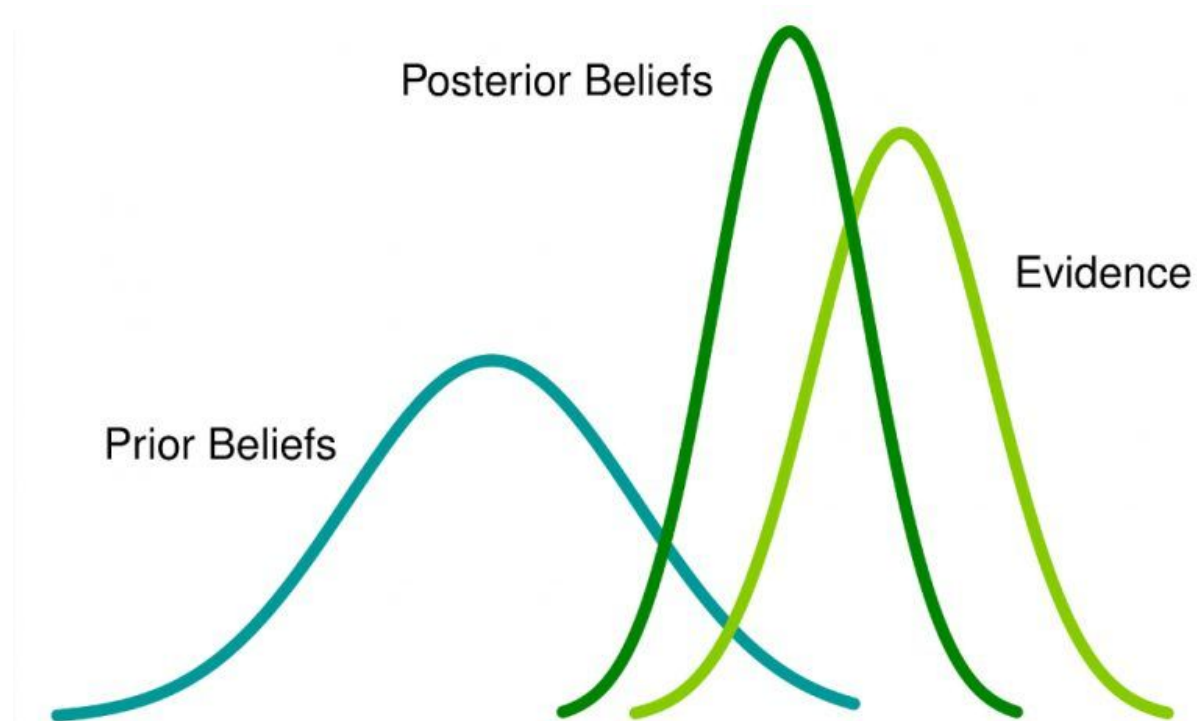


Modelo: Variables que actúan



Inferencia: **Enfoque Bayesiano**

$$P(A|B) = \frac{P(A, B)}{P(B)}$$

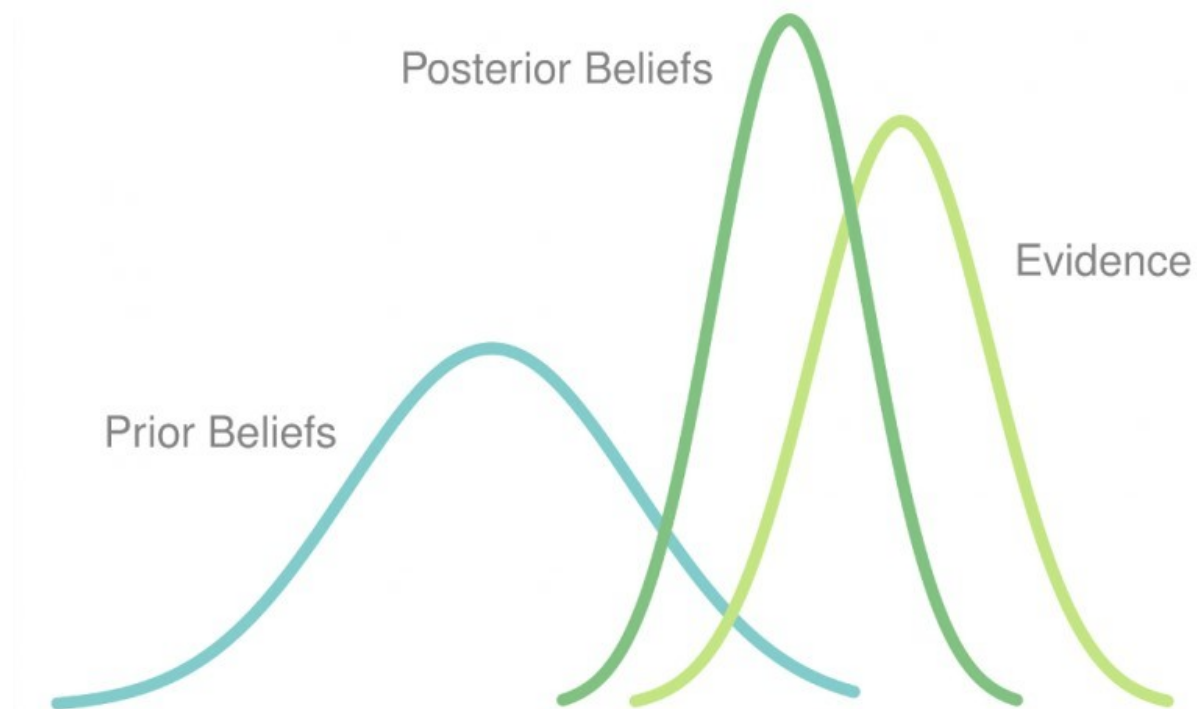
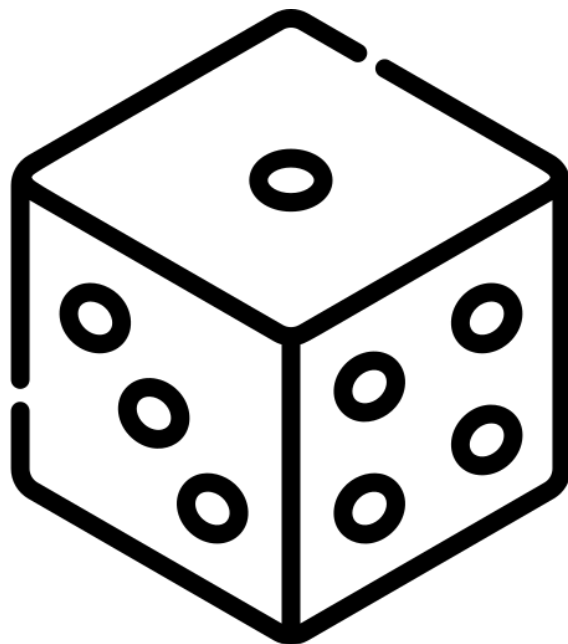


<https://empresas.blogthinkbig.com/como-predecir-el-futuro-con-inferencia-bayesiana/>

Conocimiento + Datos = Nuevo Conocimiento

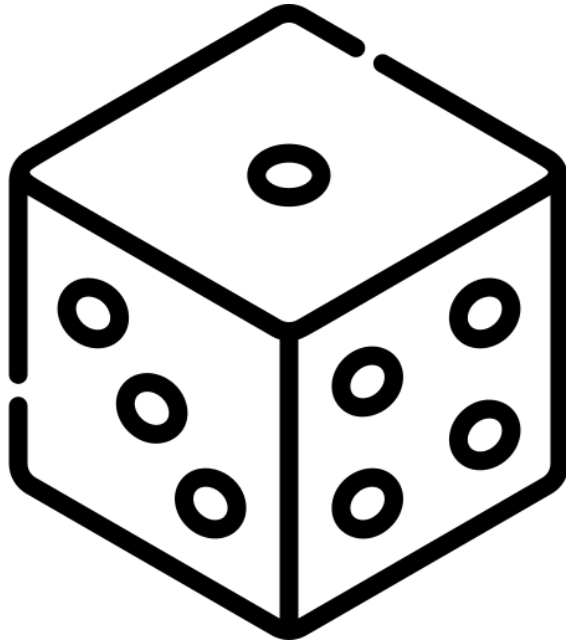
Creencias iniciales x **Nuevos datos objetivos** = **Creencia nueva y mejorada**

Inferencia: **Enfoque Bayesiano**



<https://empresas.blogthinkbig.com/como-predecir-el-futuro-con-inferencia-bayesiana/>

Creencias iniciales x **Nuevos datos objetivos** = **Creencia nueva y mejorada**



Acabamos de terminar varias partidas a un juego de mesa con dados. Mientras guardamos el material en la caja, lanzo un dado y lo cubro con la mano. “¿Qué probabilidad hay de que haya sacado un 6?”, te pregunto. “Es fácil”, respondes, “la probabilidad es de $1/6$ ”.

Miro debajo de mi mano y te revelo: «Es un número par. ¿Cuál es la probabilidad de que siga siendo un 6?». Ahora actualizarás tu vieja hipótesis gracias a la nueva información, de manera que responderás que la probabilidad pasa a ser $1/3$. Ha aumentado.

A continuación, aún te revelo más: «Y no es un 4». ¿Cuál será ahora la probabilidad de un 6? Una vez más, necesitas actualizar tu última hipótesis con la nueva información y llegarás a la conclusión de que la nueva probabilidad es $1/2$, Ha vuelto a aumentar.

¡Enhorabuena! ¡Acabas de realizar un análisis de inferencia bayesiana! Cada nuevo dato objetivo te ha obligado a revisar tu probabilidad original.

<https://empresas.blogthinkbig.com/como-predcir-el-futuro-con-inferencia-bayesiana/>

Estadística & Machine Learning

“

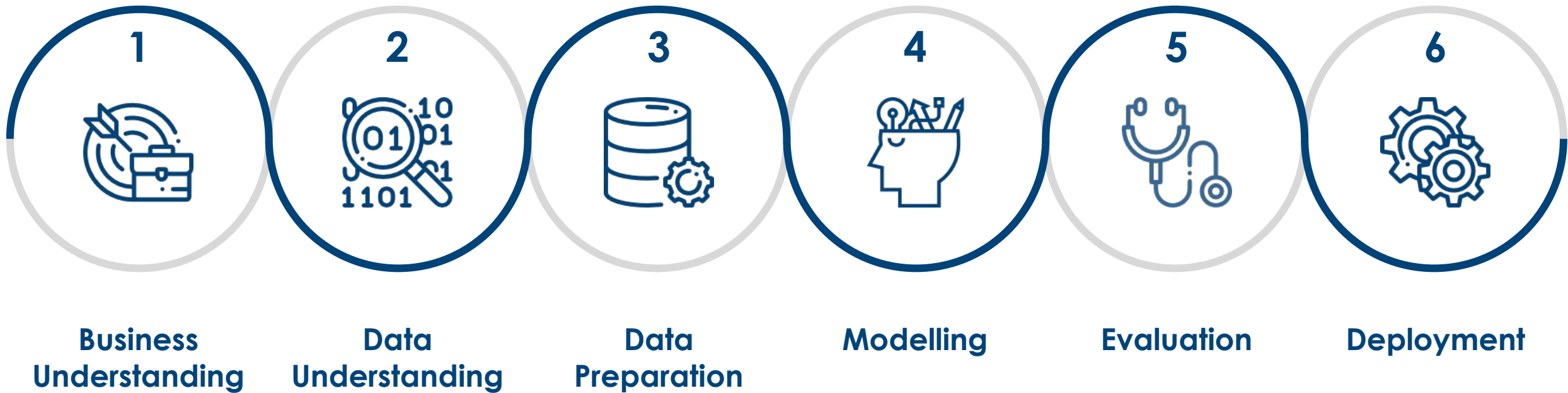
Machine learning requires no prior assumptions about the underlying relationships between the variables. You just have to throw in all the data you have, and the algorithm processes the data and discovers patterns, using which you can make predictions on the new data set. Machine learning treats an algorithm like a black box, as long it works.

In contrast, **statisticians** must understand how the data was collected, statistical properties of the estimator, the underlying distribution of the population they are studying and the kinds of properties you would expect if you did the experiment many times. You need to know precisely what you are doing and come up with parameters that will provide the predictive power.



<http://www.kdnuggets.com/2016/11/machine-learning-vs-statistics.html>

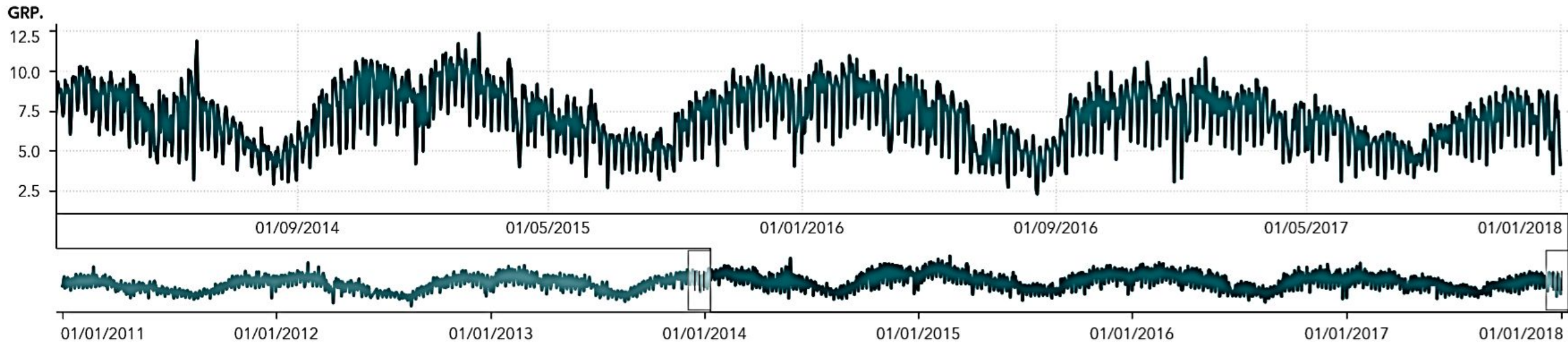
Fases de la Modelización



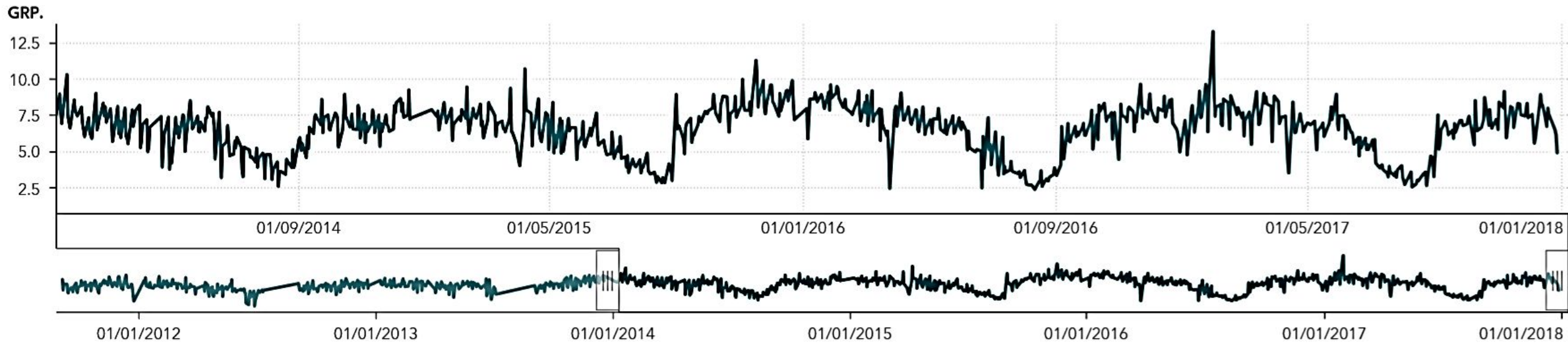


Demanda

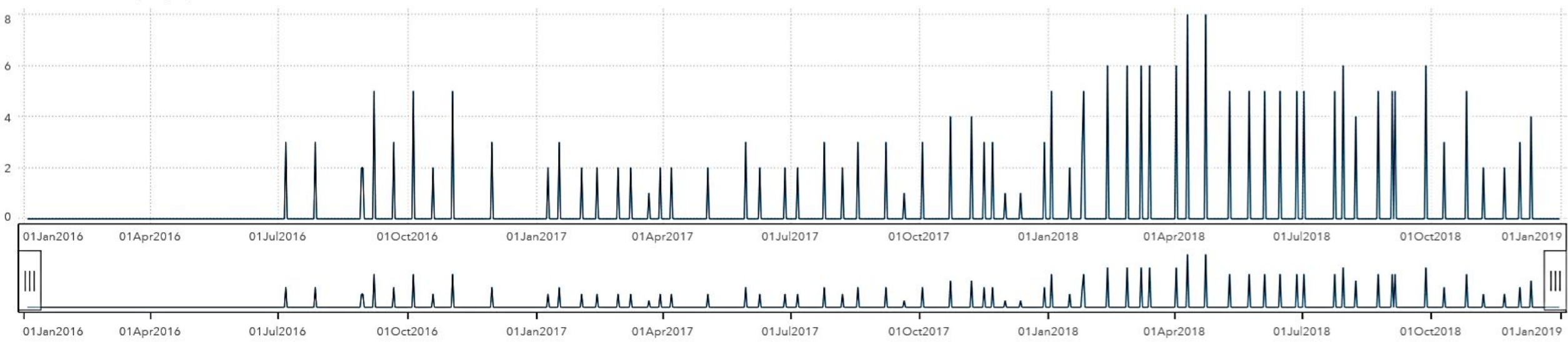
Audiencia TV (1)



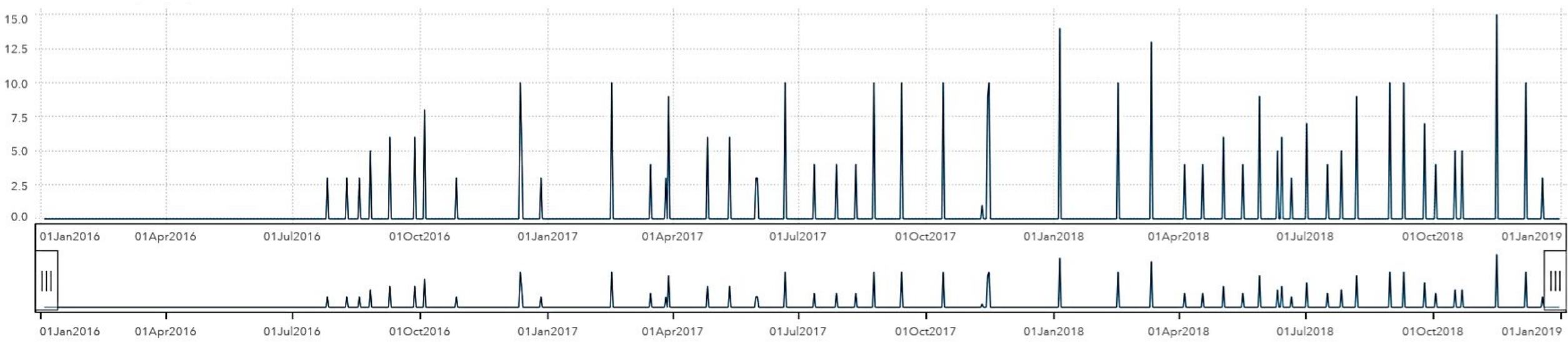
Audiencia TV (2)



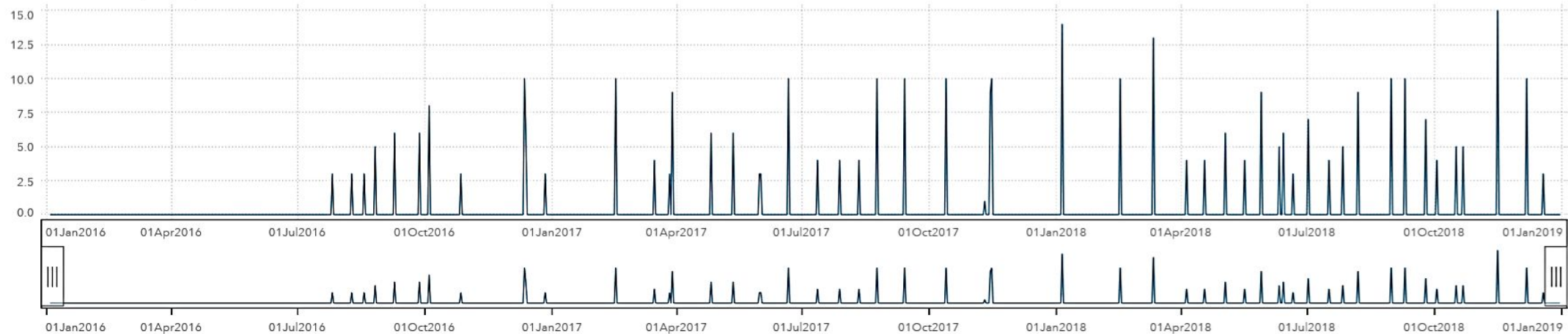
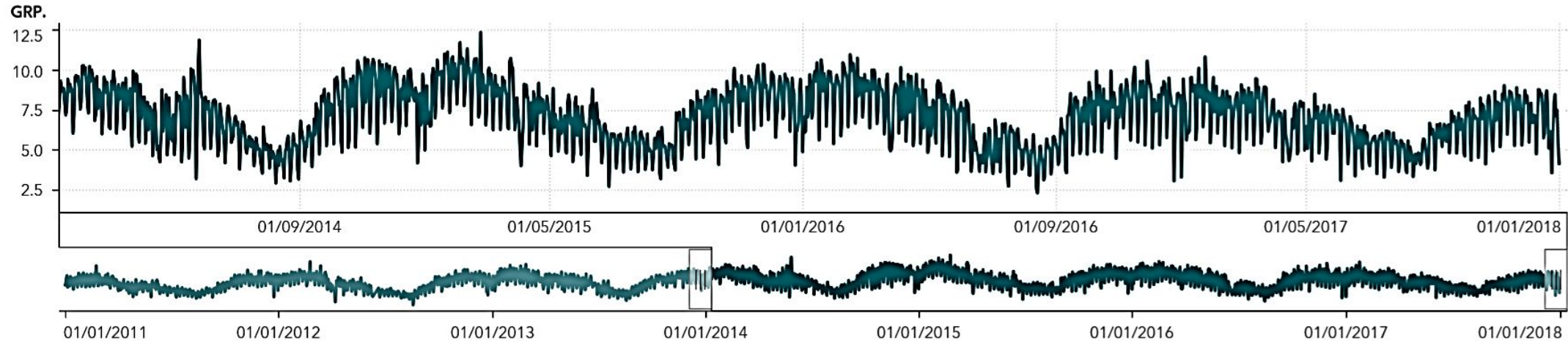
Medicina (1)



Medicina (2)



¿Cómo hacer la previsión?



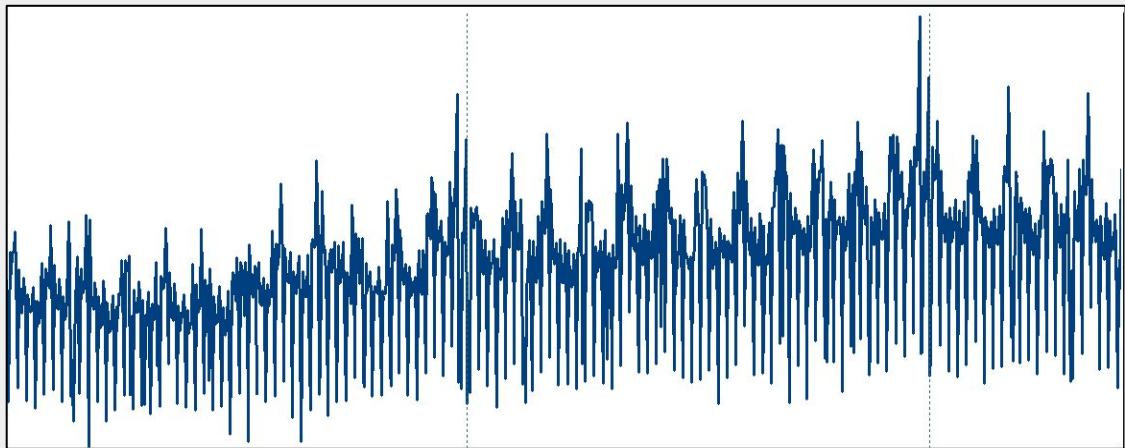
¿Cómo se mide la Demanda en Teleco?

- Planta / Portfolio / Parque
- Altas / Bajas
- Tráfico
- Recargas
- Canales de TV
- ...

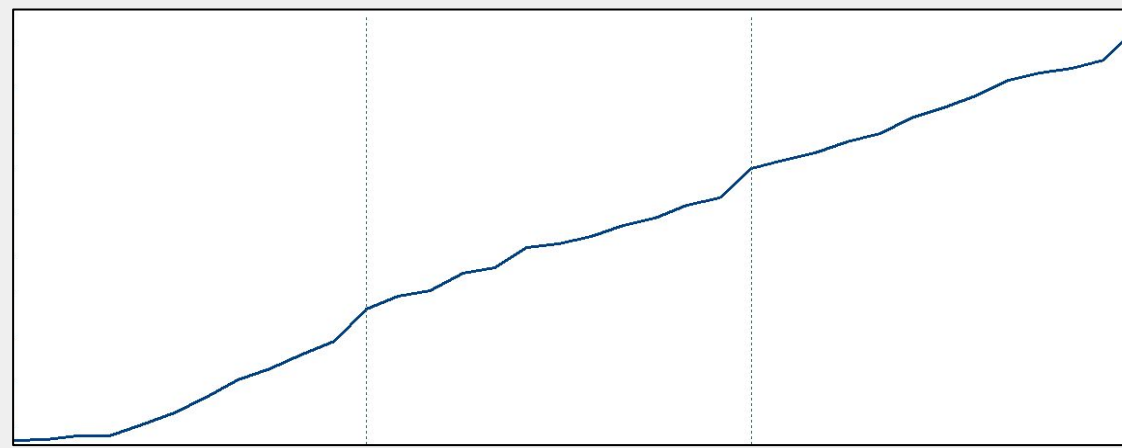


Frecuentemente, los productos pueden medirse en **unidades físicas** y esto es lo que recomendamos. En algunos casos, no obstante, podemos usar **unidades monetarias** para representar un producto de dos unidades físicas heterogéneas.

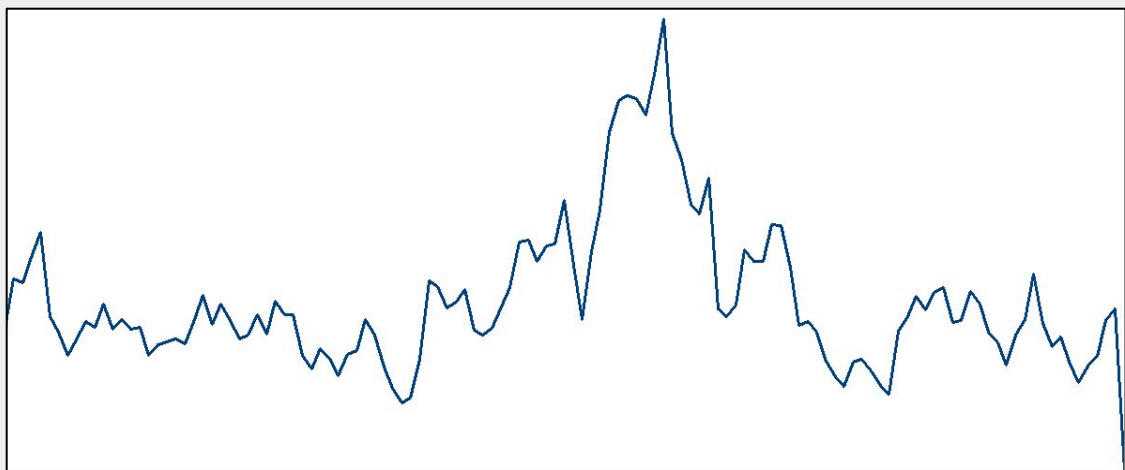
Recargas Diarias Pre-Pago



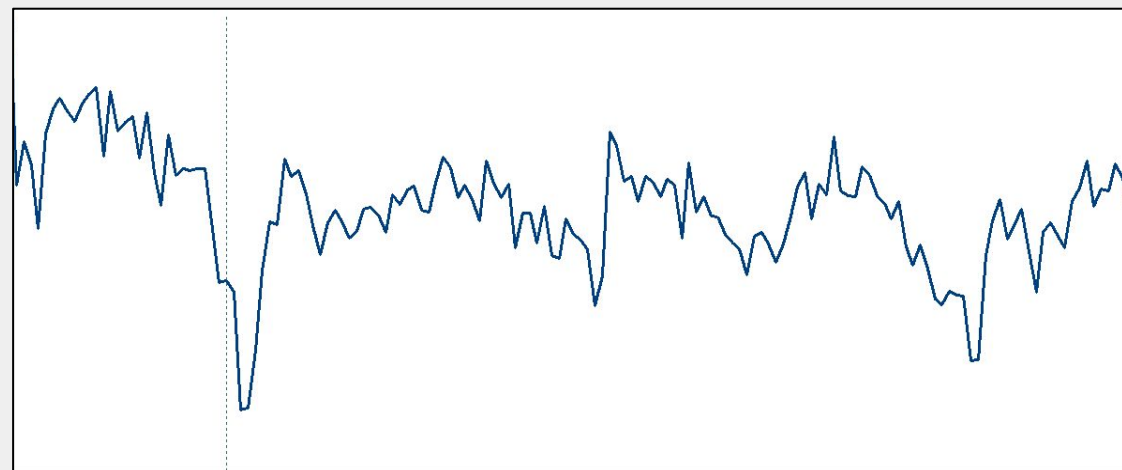
Portfolio Mensual



Llamadas Semanales Atendidas



Televentas Totales Semanales



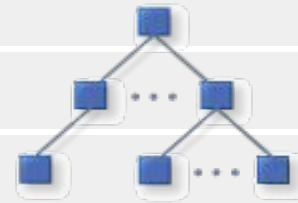
¿Qué Output(s) considerar?

¿Cuántos Outputs?

PRODUCTO

SEGMENTO

CANAL



De lo general a lo particular, de forma integrada.

¿Qué variable(s)?

TOTAL CLIENTES

ALTAS

BAJAS

Los drivers afectan de forma diferente a cada variable. P.e: Precio, Competencia, etc.

¿Qué frecuencia?

DIARIO

SEMANTAL

MENSUAL

Los efectos de los drivers pueden analizarse mejor en diferentes frecuencias. P.e.: Calendario, Economía, etc.

¿Qué transformación?

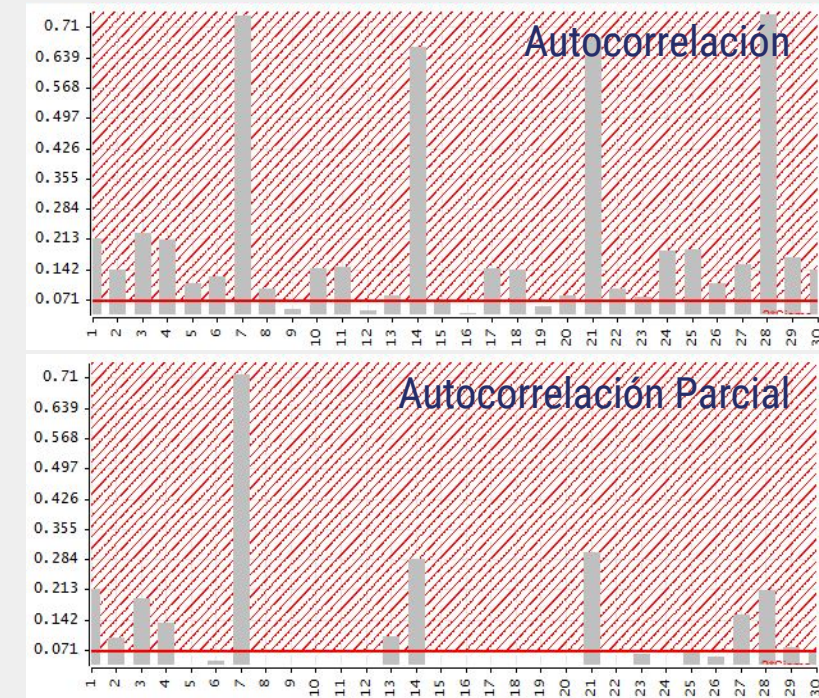
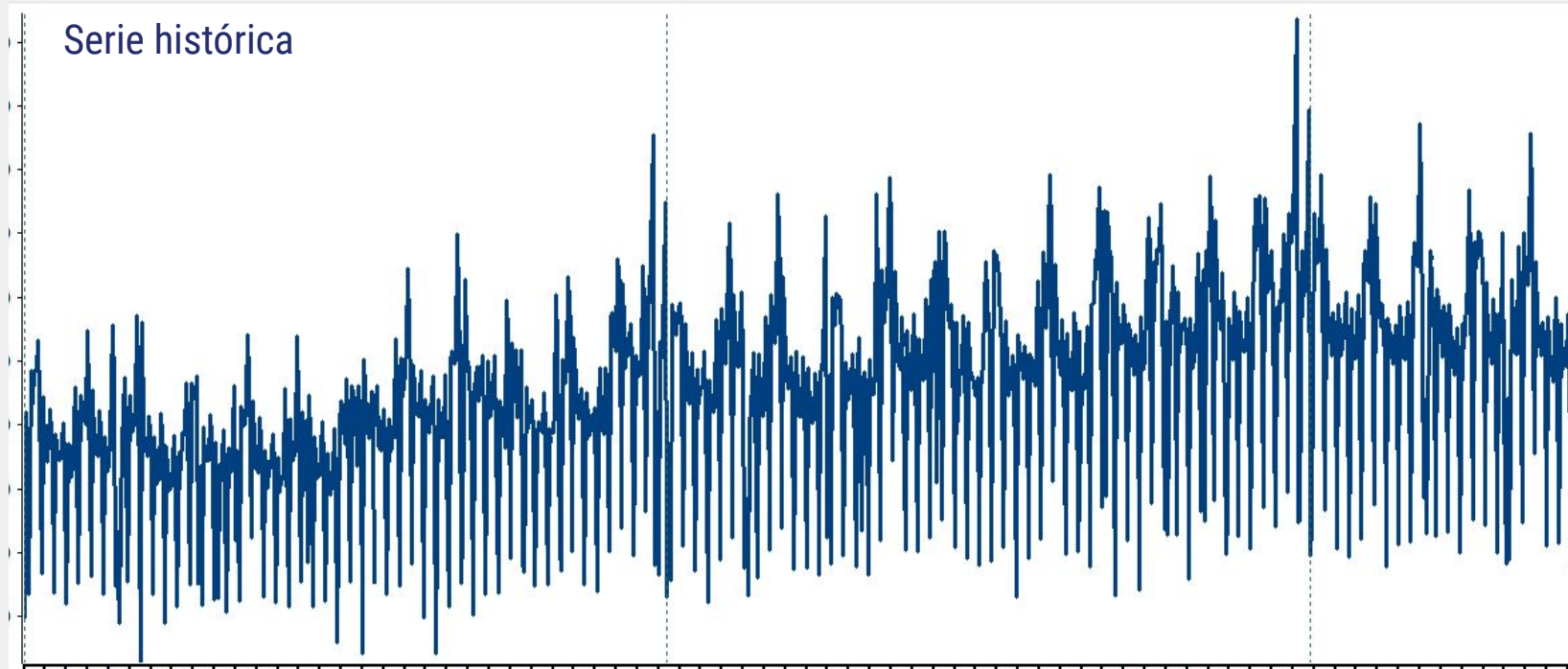
NINGUNA

LOGARITMO

VARIACION

Los modelos multiplicativos representan la interacción de manera natural.

Ejemplo: Recargas Pre-pago



- Frecuencia: Diaria.
- Histórico: Ene, 2008 - May, 2010
- Unidad: Moneda local.

1. ¿Cómo son las recargas por semana? ¿por mes?
2. ¿Cómo es la tasa de crecimiento mensual?
3. ¿Cómo es la serie en logaritmos?

Time Series Machine Learning Regression Framework

Building a time series forecasting pipeline to predict weekly sales transaction



Pourya [Follow](#)

Apr 6 · 15 min read ★

How (not) to use Machine Learning for time series forecasting: Avoiding the pitfalls



Vegard Flovik [Follow](#)

Jun 7, 2018 · 9 min read ★

An End-to-End Project on Time Series Analysis and Forecasting with Python



Susan Li [Follow](#)

Jul 9, 2018 · 9 min read

Time series Forecasting— ARIMA models



Sangarshanan [Follow](#)

Oct 3, 2018 · 5 min read

Facing the ARIMA Model against Neural Networks

Time series Forecasting for Beijing Contamination



Miguel Torres [Follow](#)

Apr 16 · 10 min read ★

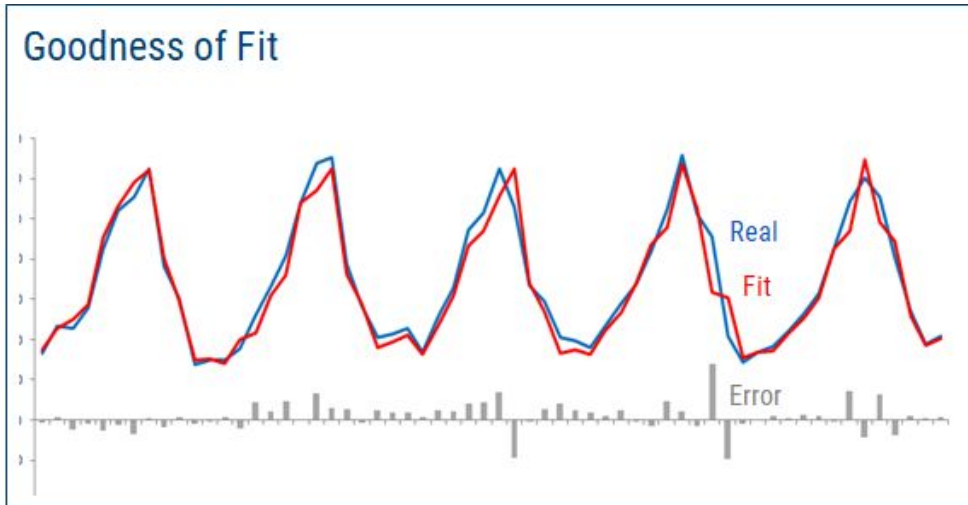
The best Forecast Techniques or how to Predict from Time Series Data



Edwin Lisowski [Follow](#)

May 23 · 5 min read ★

Evaluación de la Calidad Predictiva



Coeficiente de Determinación

$$R^2 = \frac{SSR}{SST} = 1 - \frac{SSE}{SST},$$

AIC/BIC

$$AIC = 2k - 2\ln(\hat{L})$$
$$BIC = \ln(n)k - 2\ln(\hat{L})$$

Análisis de Residuos

$$MAPE = \frac{1}{n} \sum_{i=1}^n \frac{100 \cdot |\hat{y}_i - y_i|}{y_i}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{n}}$$

$$RSE = \sqrt{\frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{\sum_{i=1}^n (\bar{y} - y_i)^2}}$$

$$MAE = \frac{\sum_{i=1}^n |\hat{y}_i - y_i|}{n}$$

- **MAPE** (mean absolute percentage Error)
- **RMSE** (root mean squared error) o error cuadrado medio
- **MAE** (mean absolute error) o error absoluto medio
- **RAE** (relative absolute error) o error relativo absoluto
- **RSE** (relative squared error) o error relativo cuadrado
- **Coeficiente de determinación**,
- **AIC** [Akaike information criterion](#)
- **BIC**. [Bayesian information criterion](#)
- **DIC**. [Deviance information criterion](#)



Variables que actúan

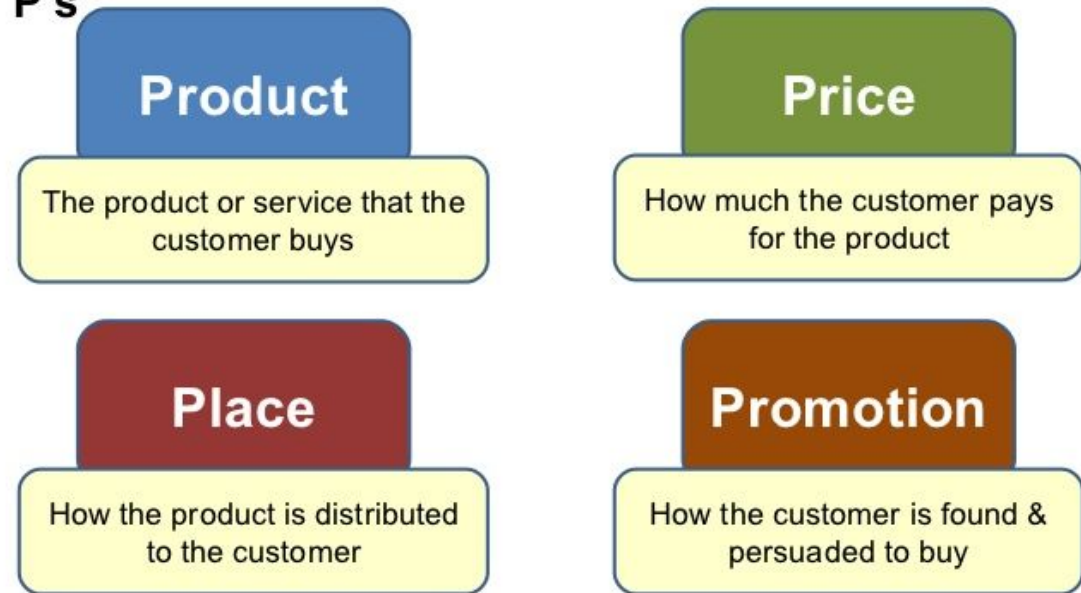
Las Cuatro P



Se trata de los cuatro elementos de los que dispone la empresa para conseguir que sus **estrategias de marketing** resulten efectivas y alcancen los fines previamente propuestos. Factores que ofrecen un gran margen de **maniobra** para ser modificadas, por lo que el resultado y que la venta del producto sea o no exitosa dependerá mucho de las **decisiones** que se vayan tomando al respecto.

The Four P's

The marketing mix is often referred to as the Four P's



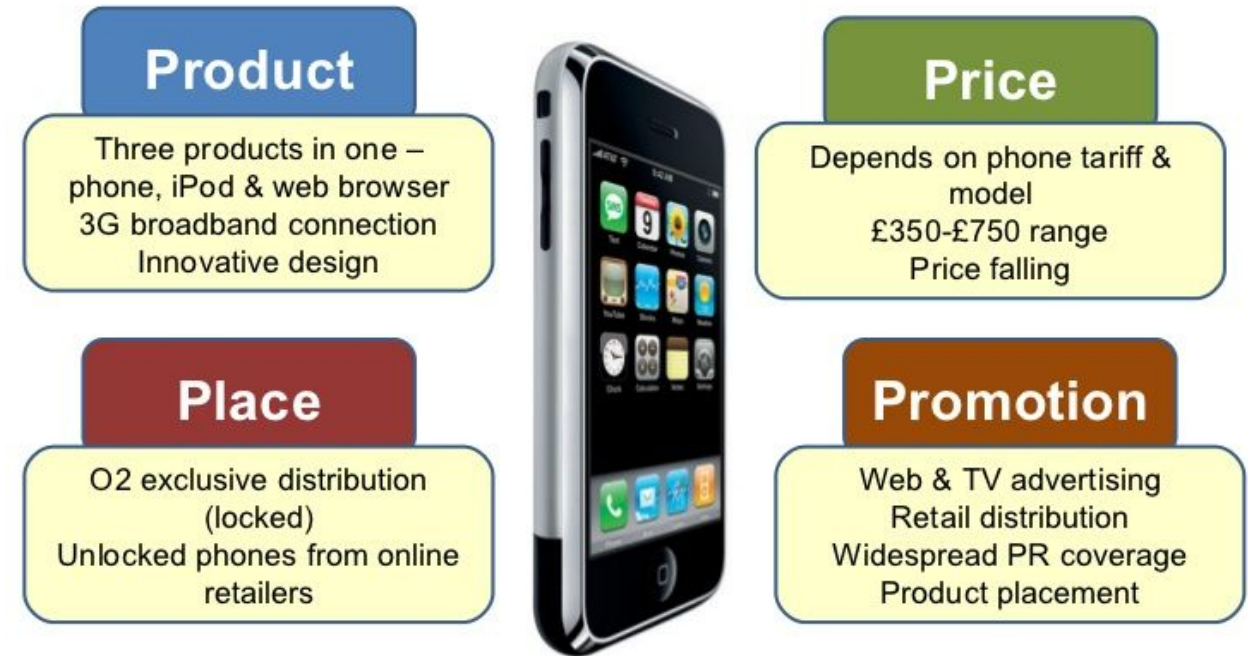
Cómo hacer un marketing mix
<https://www.solomarketing.es/como-hacer-un-marketing-mix/>

Las Cuatro P



Se trata de los cuatro elementos de los que dispone la empresa para conseguir que sus **estrategias de marketing** resulten efectivas y alcancen los fines previamente propuestos. Factores que ofrecen un gran margen de **maniobra** para ser modificadas, por lo que el resultado y que la venta del producto sea o no exitosa dependerá mucho de las **decisiones** que se vayan tomando al respecto.

Example – launch of iPhone



Cómo hacer un marketing mix

<https://www.solomarketing.es/como-hacer-un-marketing-mix/>

No se puede entender la publicidad aisladamente

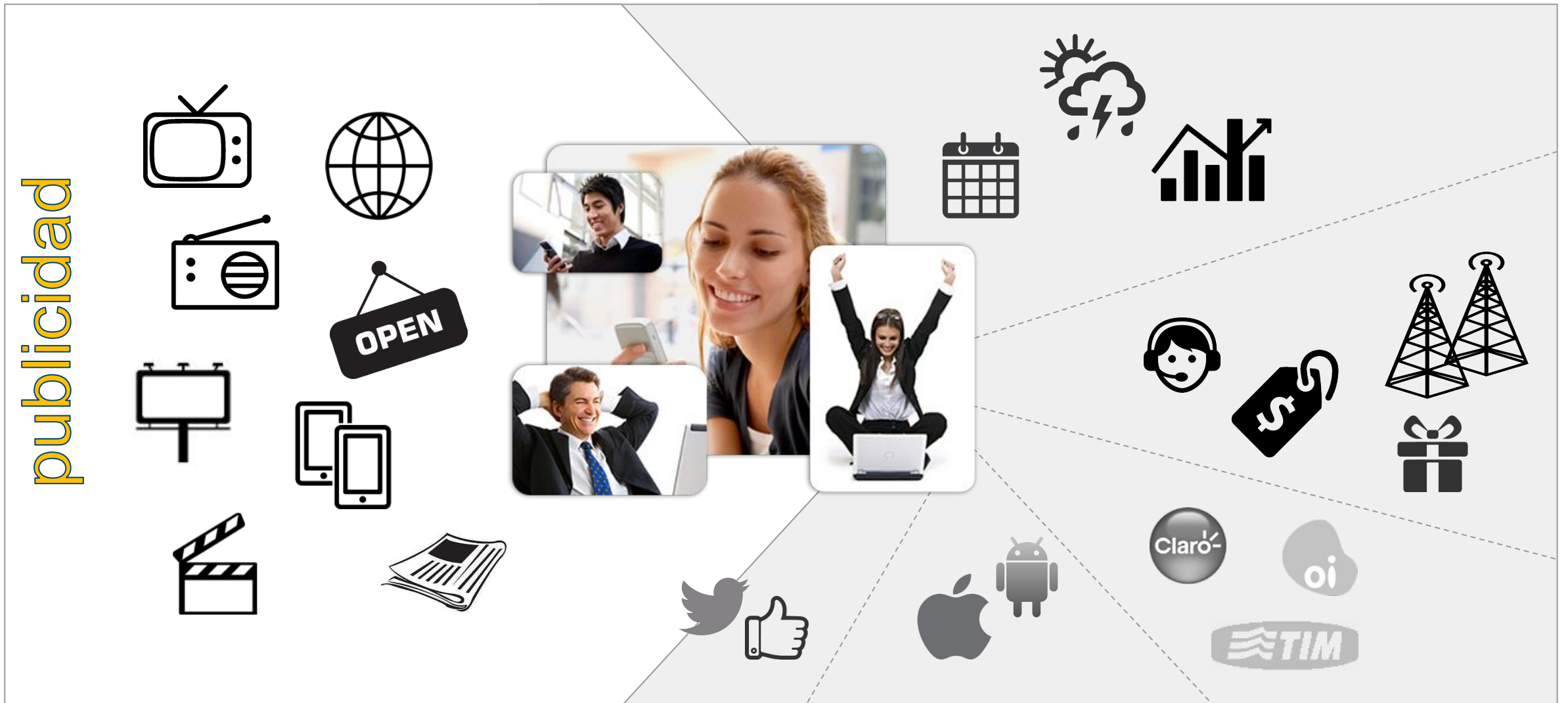
Para medir adecuadamente el efecto del marketing, que es relativamente pequeño y muy distribuido en el tiempo, debemos medir los otros efectos, que son mayores y que impiden su visión directa.

Ello nos lleva a la necesidad de entender y medir el mercado en su conjunto.

La señal de que nuestro entendimiento es apropiado es nuestra capacidad predictiva.



Muchas variables actúan sobre la Demanda

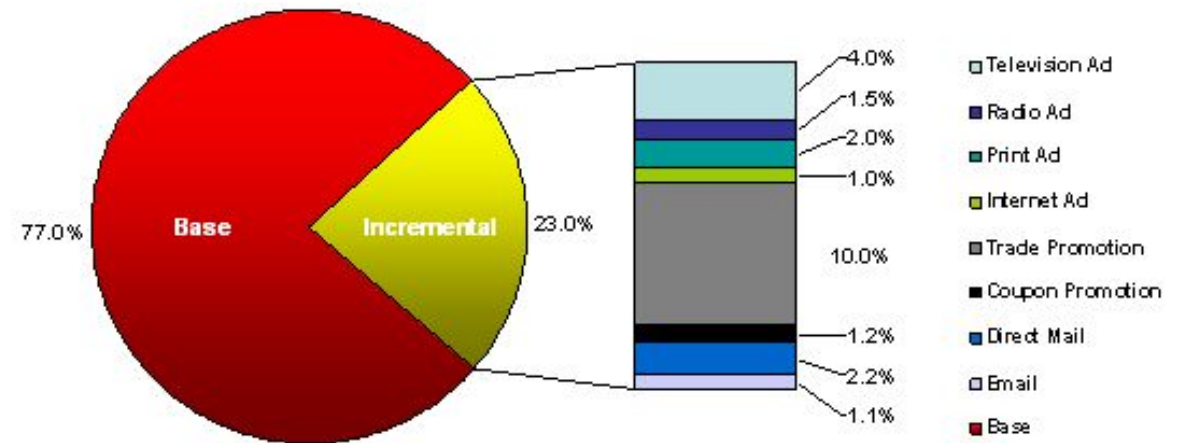


Modelos de Marketing Mix



Marketing mix modeling (MMM) is **statistical analysis** such as **multivariate regressions** on sales and marketing **time series** data to estimate the **impact** of various **marketing** tactics (**marketing mix**) on sales and then **forecast** the impact of future sets of tactics. It is often used to **optimize** advertising mix and promotional tactics with respect to sales revenue or profit.

Product Sources of Volume: Year Ago



Marketing mix modeling

https://en.wikipedia.org/wiki/Marketing_mix_modeling

<https://www.slideshare.net/indhudprincy/market-mix-modelling>

<https://www.bayesforecast.com/bayes-drivers-of-growth-due-to-analytics/>

¿Qué son los modelos?

Input (X)

$$f(X) = F$$
$$Y = F + N$$

Output (Y)



Variables sobre las que se tiene **control**



Variables que están bajo el control de la **competencia**



Variables de **entorno**



¿Qué son los modelos?

Nivel de Incertidumbre

Input (X)

$$f(X) = F$$
$$Y = F + N$$

Output (Y)



Variables sobre las que se tiene **control**

Precio
Publicidad
Promoción
Distribución

Conocidas:
Optimizables



Variables que están bajo el control de la **competencia**

Precio
Publicidad
Promoción

Desconocidas:
Su futuro se representa en forma de distribuciones de probabilidad



Variables de **entorno**

Economía
Calendario

Deterministas

¿Cómo?

MODELO CONCEPTUAL

¿Cómo es la relación DRIVER-DEMANDA, desde el punto de vista de negocio?

DATOS

¿Qué indicador(es) o métrica(s) representa(n) el DRIVER? ¿En qué unidades? ¿Se perciben problemas de medida?

INPUT

¿Qué función de transferencia representa bien la relación DRIVER-DEMANDA?

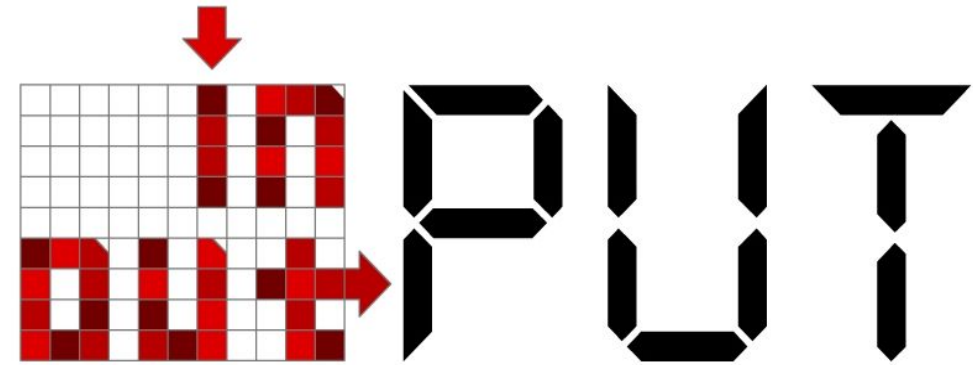
EFECTO

¿Cómo representar el impacto sobre la Demanda? - Respuesta a Preguntas de Negocio -

ACCIÓN

¿Cómo se usan los resultados? - Planificación y Optimización -

¿Cómo es la relación ?

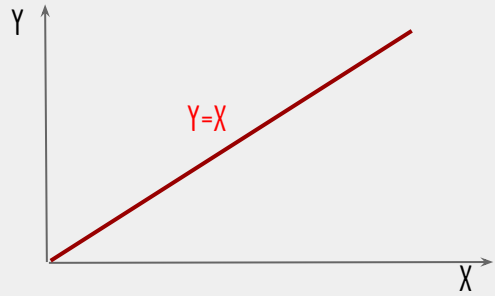


- ✓ ¿Relación positiva o negativa?
- ✓ ¿Relación lineal o no lineal?
- ✓ ¿Existe memoria? ¿Cuánto dura?
- ✓ ¿Existen rendimientos decrecientes?
- ✓ ¿Existe saturación?
- ✓ ¿Efecto multiplicativo o aditivo?
- ✓ ¿Interacción con otros inputs? ¿Cómo?

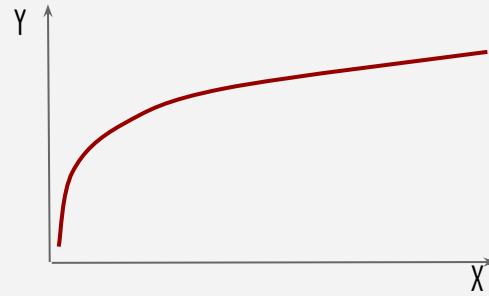
Funciones de Transferencia

Transformación Instantánea del input

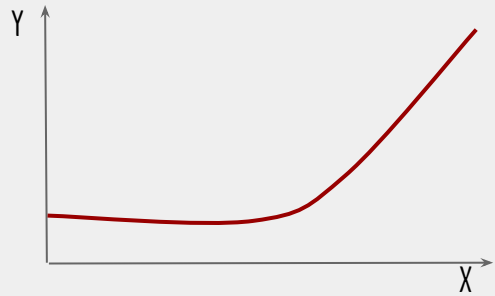
Identidad



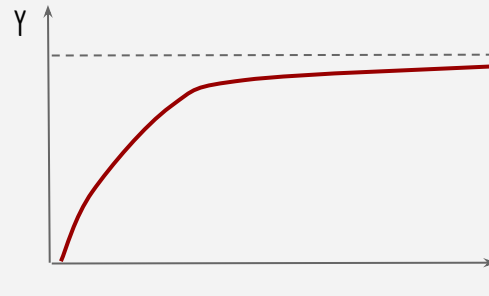
Logaritmo



Piecewise

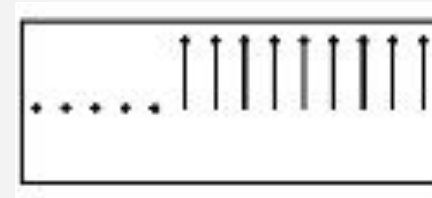


RD-Saturación

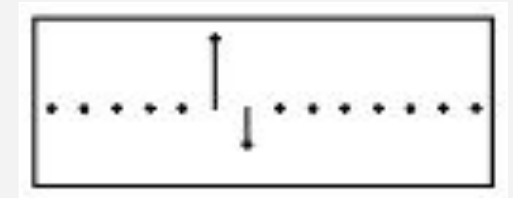


Estructura temporal de la respuesta

Nuevo Nivel



Compensación



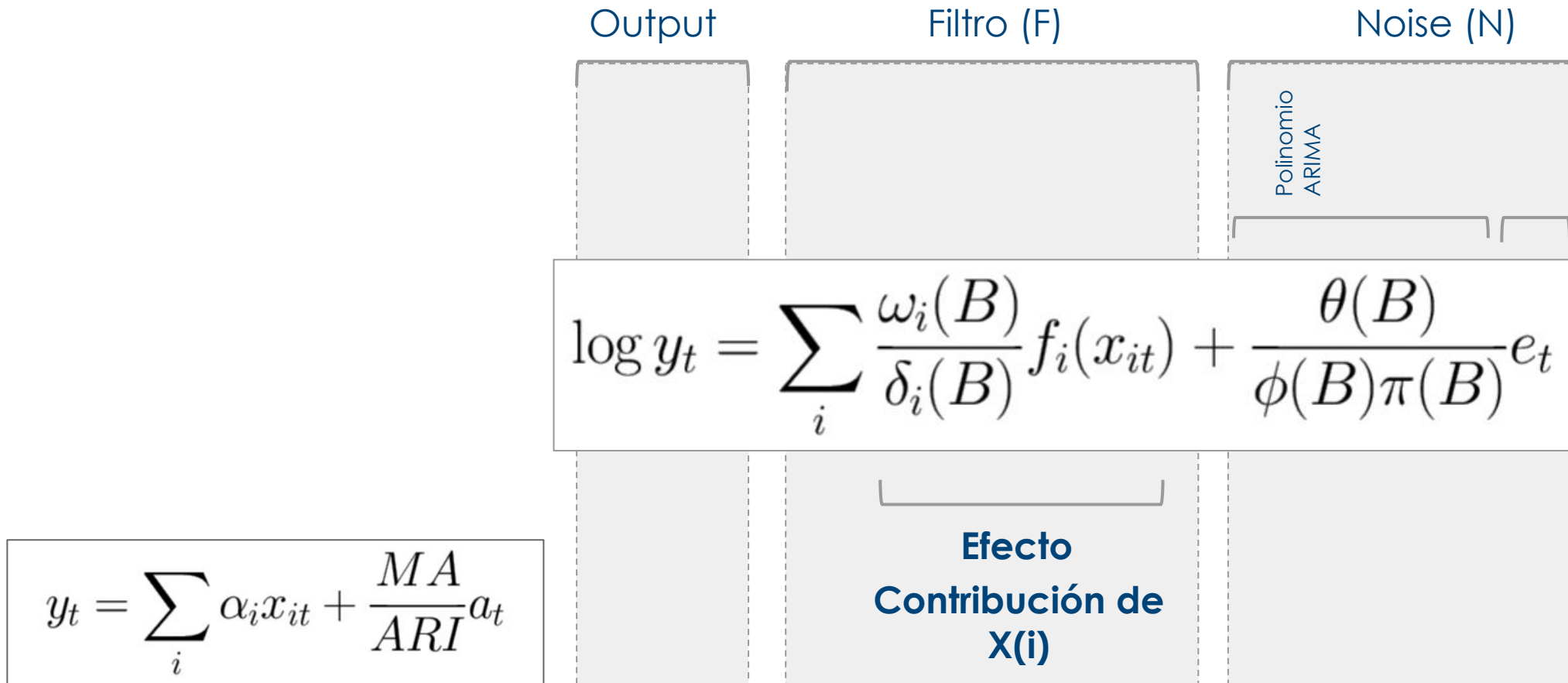
Decreciente



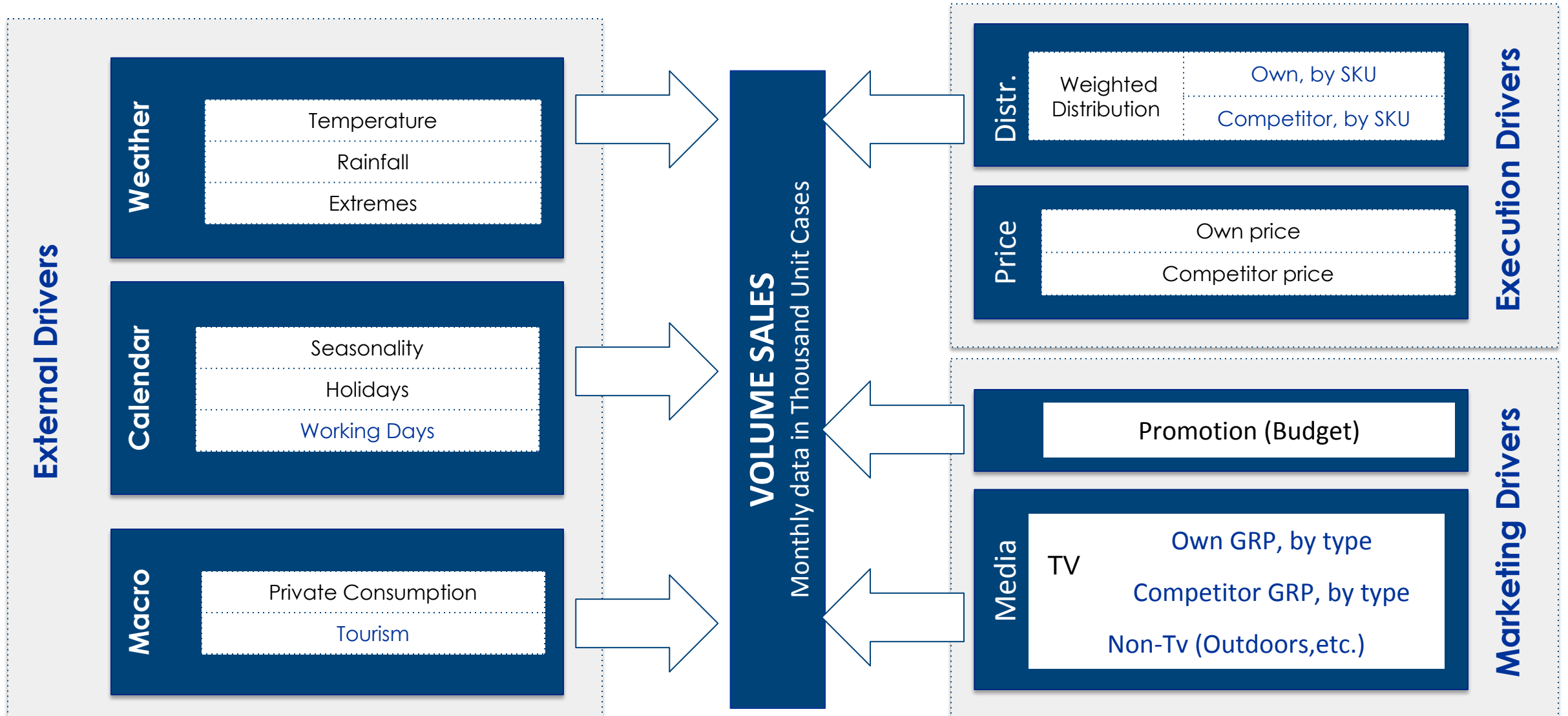
Creciente



Un ejemplo de modelo



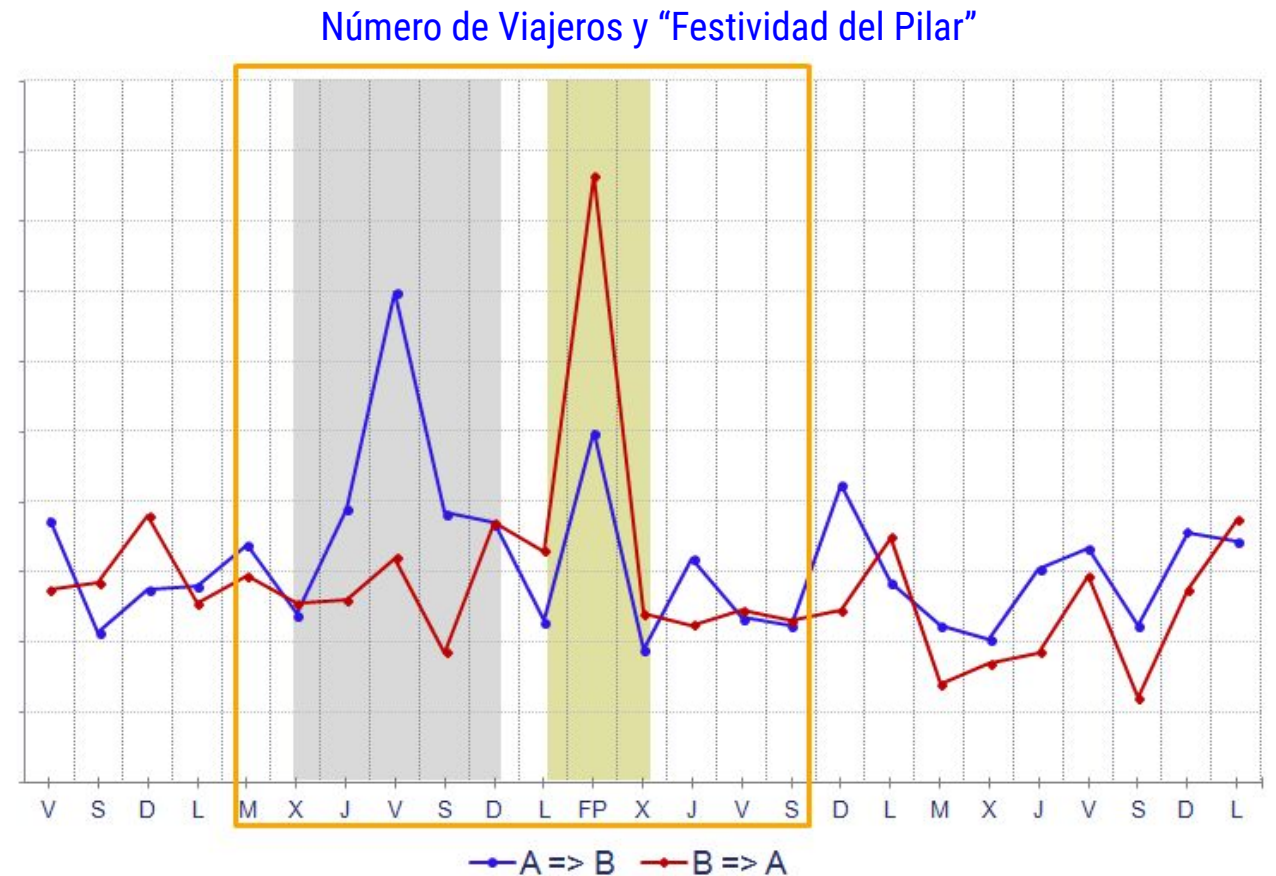
Ejemplo de MMM



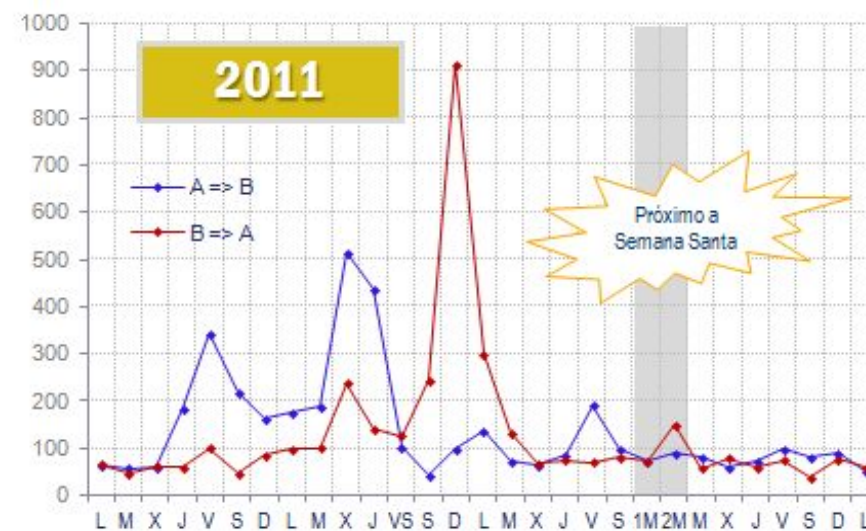
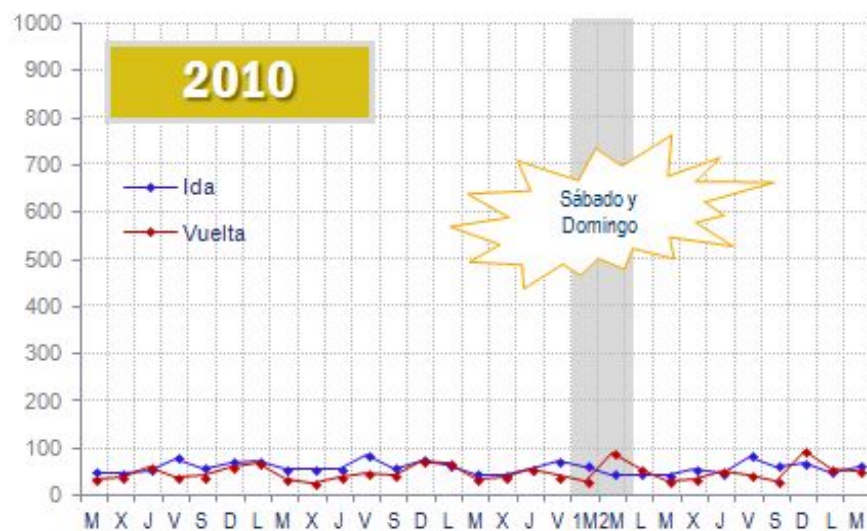
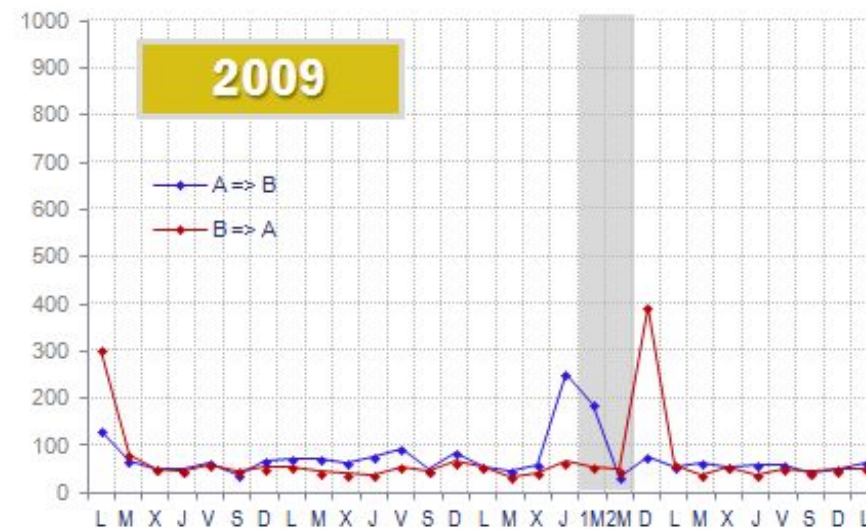
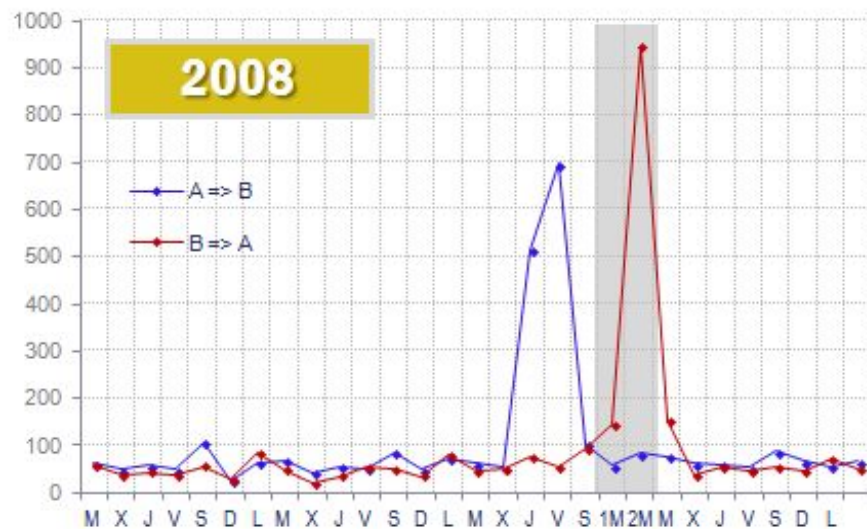
Calendario

- Efecto en el tiempo
- Compensación

- ¿Cómo afecta el calendario social a mi demanda?
 - ¿Efecto positivo/negativo?
 - ¿Efecto traslado?
- ¿El efecto cambia según el día de la semana en que ocurre el feriado? ¿y si hay “puente”?
- ¿Hay festivos locales? ¿Cómo representar el efecto de un festivo no nacional?
- ¿Hay diferencia entre el efecto de Semana Santa y el efecto de la Navidad?
- ¿Hay otros efectos previsibles, dentro de un mismo mes?



Ejemplo: Puente de Mayo



Economía

Consumo Privado



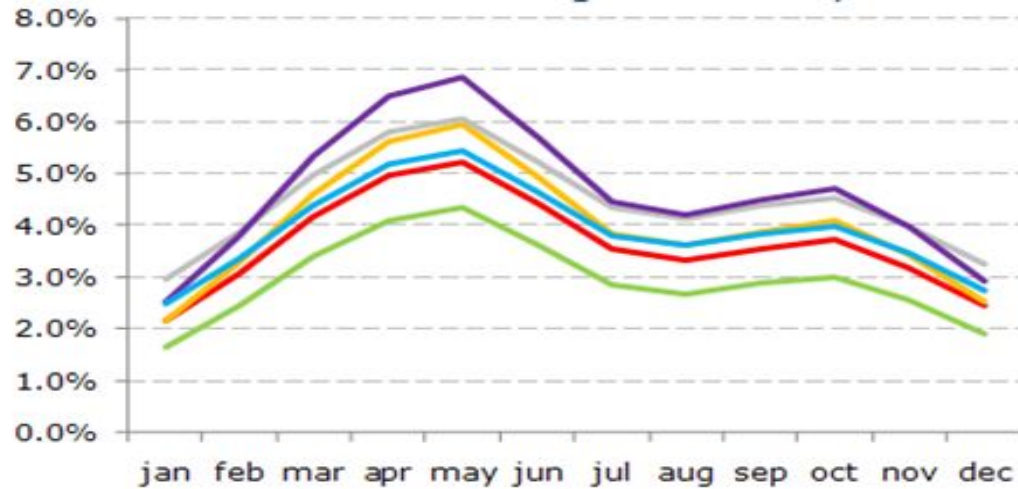
Tasa de Desempleo



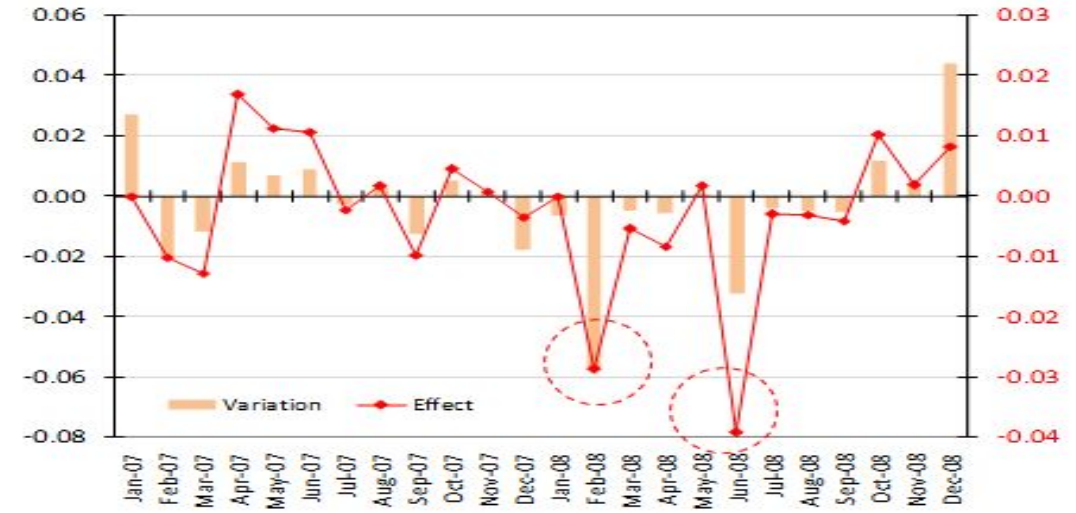
- ¿Cómo afecta el contexto económico a mi demanda?
- ¿Qué indicador representa mejor el contexto económico relacionado con mi negocio?
- ¿Cómo representar la relación entre la economía del país?
- ¿El efecto sobre mis ventas es igual antes / durante / después de 2008?
- ¿Qué dificultades pueden presentarse para hacer uso de las variables económica?

Temperatura

Efecto de 1°C extra



Impacto en las Ventas



- ¿Qué ocurre si la temperatura media se eleva 1°C?
- ¿Cómo representamos el efecto de la temperatura?
- ¿Qué otras variables pueden representar el clima? ¿Qué es más importante en España: temperatura máxima o precipitación?
- ¿Qué dificultades pueden presentarse para hacer uso de las variables climatológicas?



La Publicidad

Comunicación y Venta

mensaje



grp

memoria

acumulación

adstock

saturación

rendimientos
decrecientes

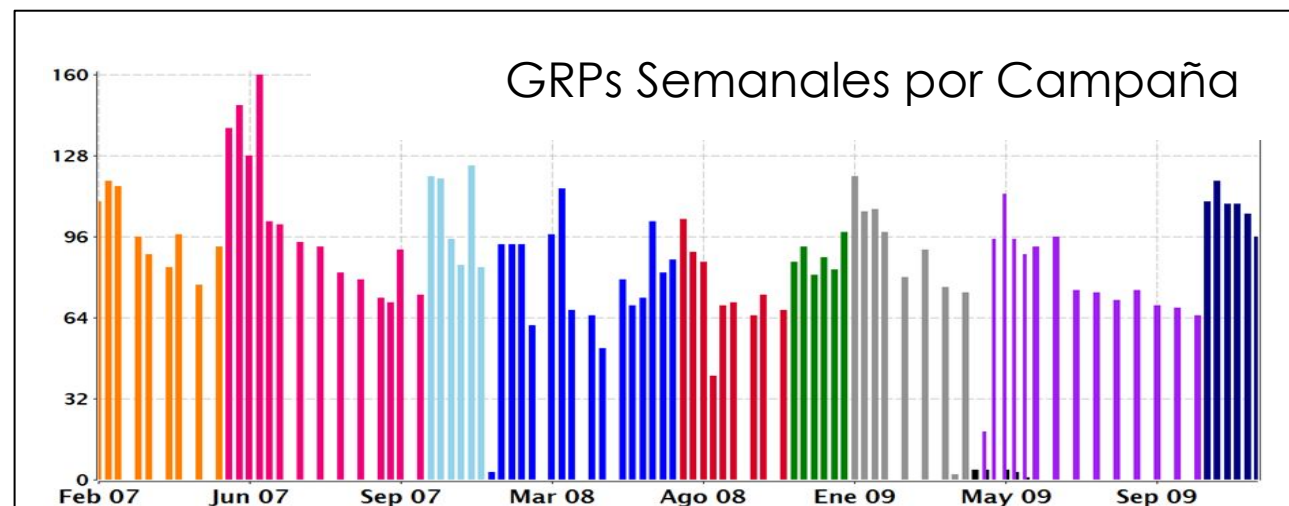
Gross Rating Point (GRP)

Gross Rating Point, es una unidad de medida de la audiencia: es el producto entre esfuerzo en los medios y el porcentaje del mercado objetivo alcanzado.

$$\text{GRP} = \text{Cobertura} \times \text{Frecuencia}$$

Cobertura: porcentaje de individuos objetivo, que han visto la publicidad, durante un periodo de tiempo determinado

Frecuencia: número de exposiciones del público objetivo a la publicidad en un período de tiempo determinado



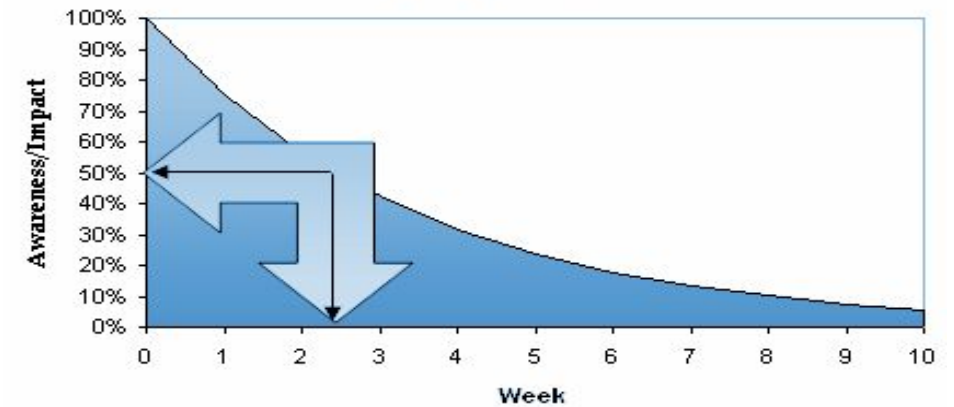
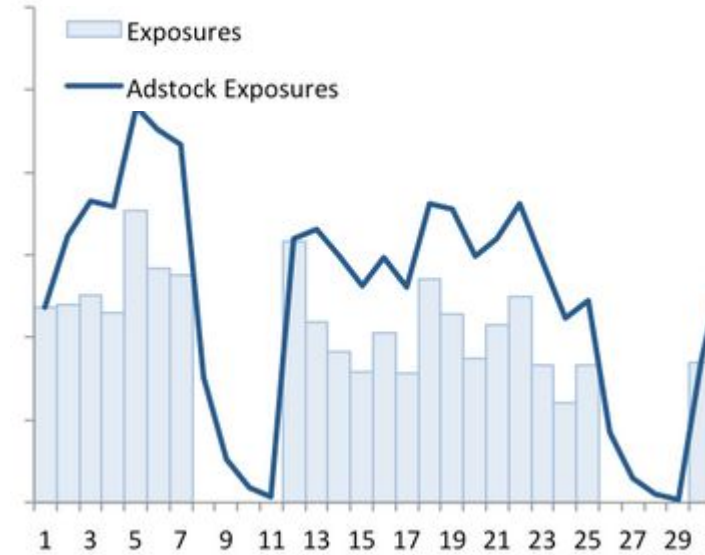
- ✓ GRPs para ser comparables deben estar expresados en el mismo target para todas las marcas objeto del estudio.
- ✓ GRPs no deben estar normalizados a 20" (un 40" no tiene doble efecto en ventas que un 20").
- ✓ Una alta cantidad de GRPs en una cadena con muy poca audiencia (por ej: canales de pago musicales) debe tratarse de modo diferenciado

Adstock

El **Adstock** es la medida de la persistencia del mensaje publicitario en la mente del público, recoge cómo se distribuye el impacto de la publicidad a lo largo del tiempo.

$$AS_t = (1 - \delta)GRP_t + \delta AS_{t-1}$$

Un concepto asociado al término de Adstock, es la **Vida Media** que se define como el número de periodos que son necesarios para que se haya consumido el impacto de la mitad de los GRPs.



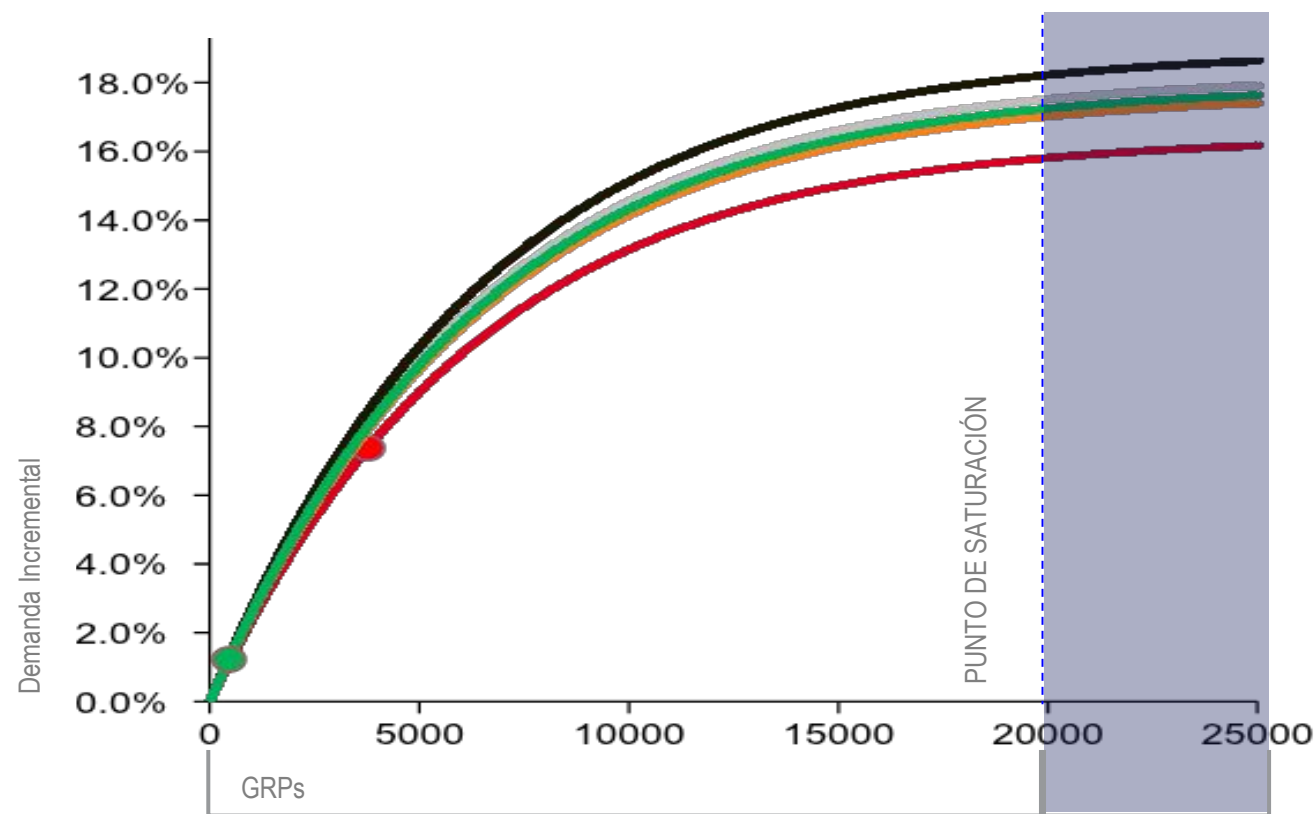
Rendimientos Decrecientes y Saturación

Al aumentar el número de GRP's aumenta la cobertura (% personas del target contactadas) pero también la frecuencia (número de veces que se contacta en promedio a cada individuo del target).

Por tanto, el número de individuos alcanzados no aumenta de la misma manera y llega un momento que no hay nuevos contactos – sólo frecuencia:

Punto de Saturación

Se puede llegar a un punto de saturación que por más GRPs que añada no llega a más gente.



El aumento de GRPs implica aumento de Demanda. El aumento de la Demanda no es lineal (idéntico) al aumento de GRPs

El aumento de GRPs no aumenta la Demanda.

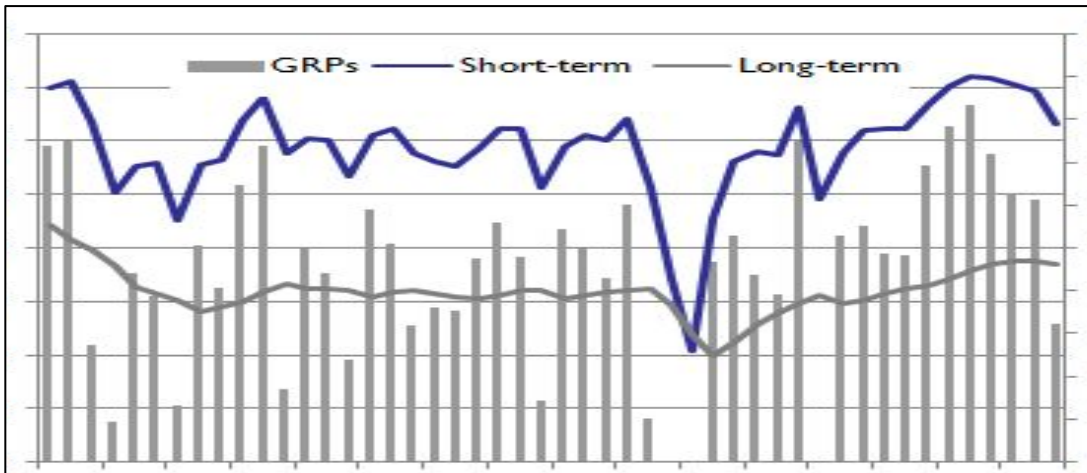
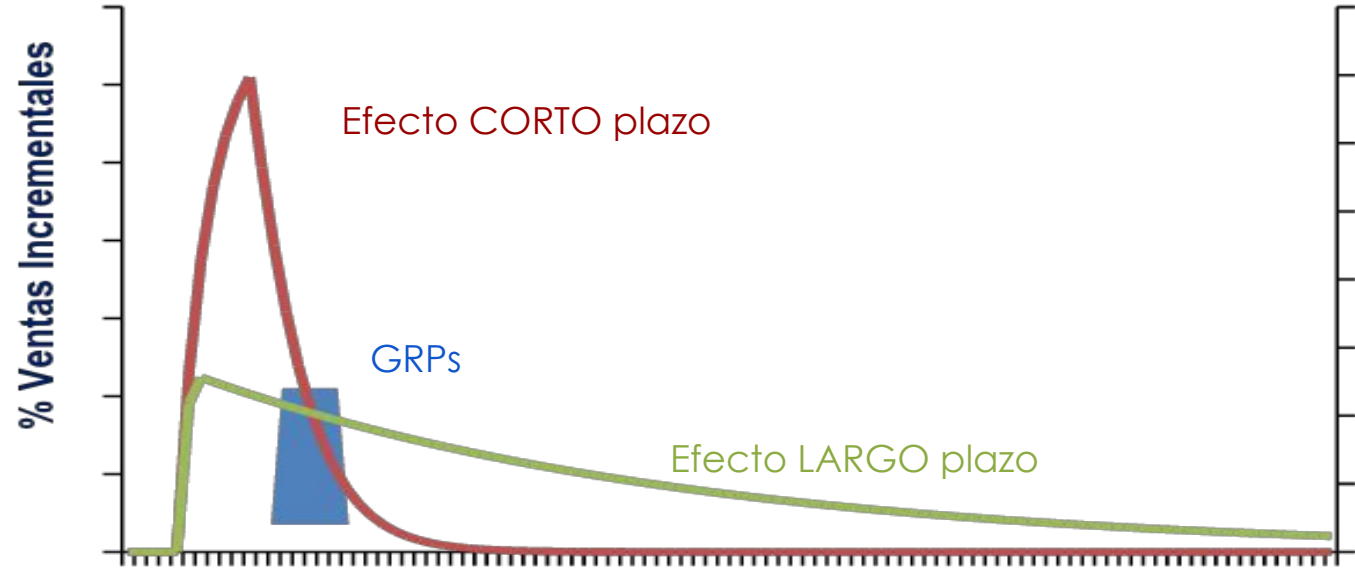
Efecto de Corto y Largo Plazo

Inmediato / Corto plazo

Incremento de las ventas a corto plazo, con una vida media no superior a cinco semanas.

Largo plazo

La publicidad permite afianzar a largo plazo una marca en el mercado.



El Input

La idea básica es aceptar que los GRPs son almacenados en la memoria de los consumidores, creando una cantidad de recuerdos denominados Adstock

$$AS_t = (1 - \delta)GRP_t + \delta AS_{t-1}$$

$$AS_t = \frac{(1 - \delta)GRP_t}{1 - \delta}$$

Una transformación no lineal del GRP se introduce para obtener los efectos de los rendimientos decrecientes de los anuncios y en algunos casos el nivel de saturación de los efectos.

$$\log(V_t) = \beta(1 - \exp(-\gamma AS_t)) + R_t$$

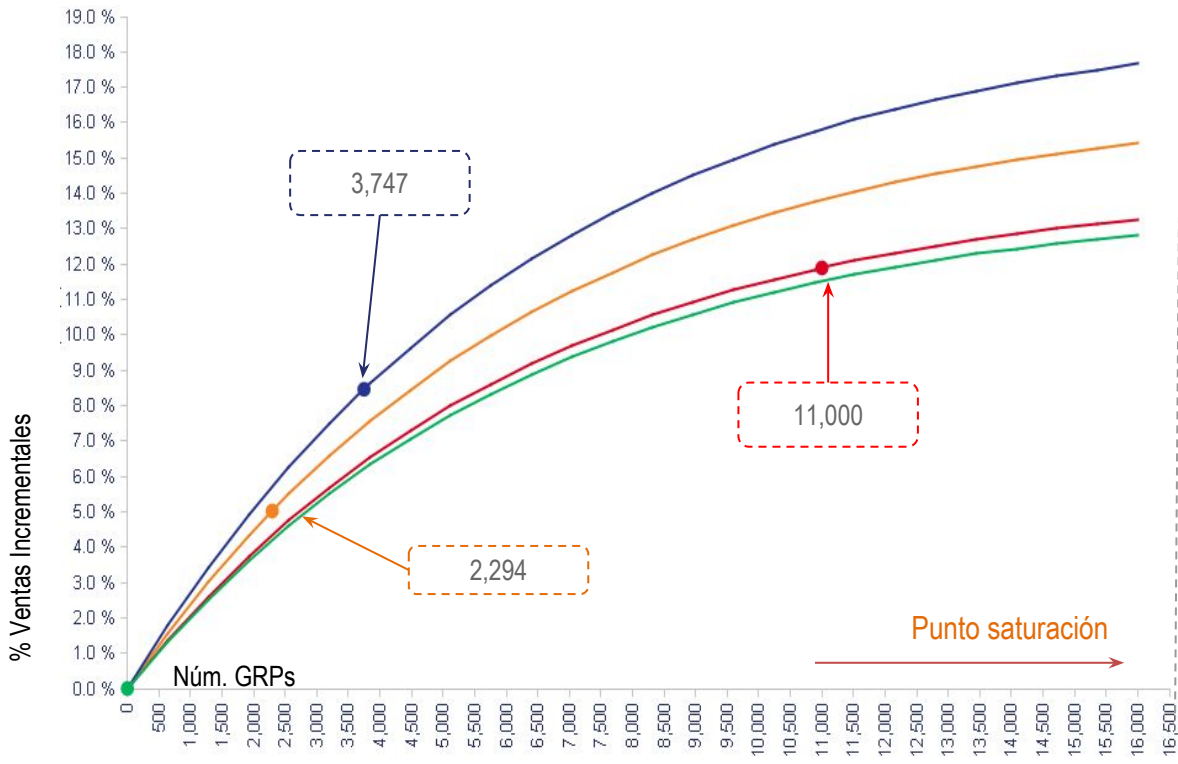
$$\log(V_t) = \beta(1 - r^{AS_t}) + R_t,$$

También se puede utilizar una parametrización alternativa, más simple, pero que no es capaz de expresar un grado de saturación.

$$\log(V_t) = \alpha \log AS_t + R_t.$$

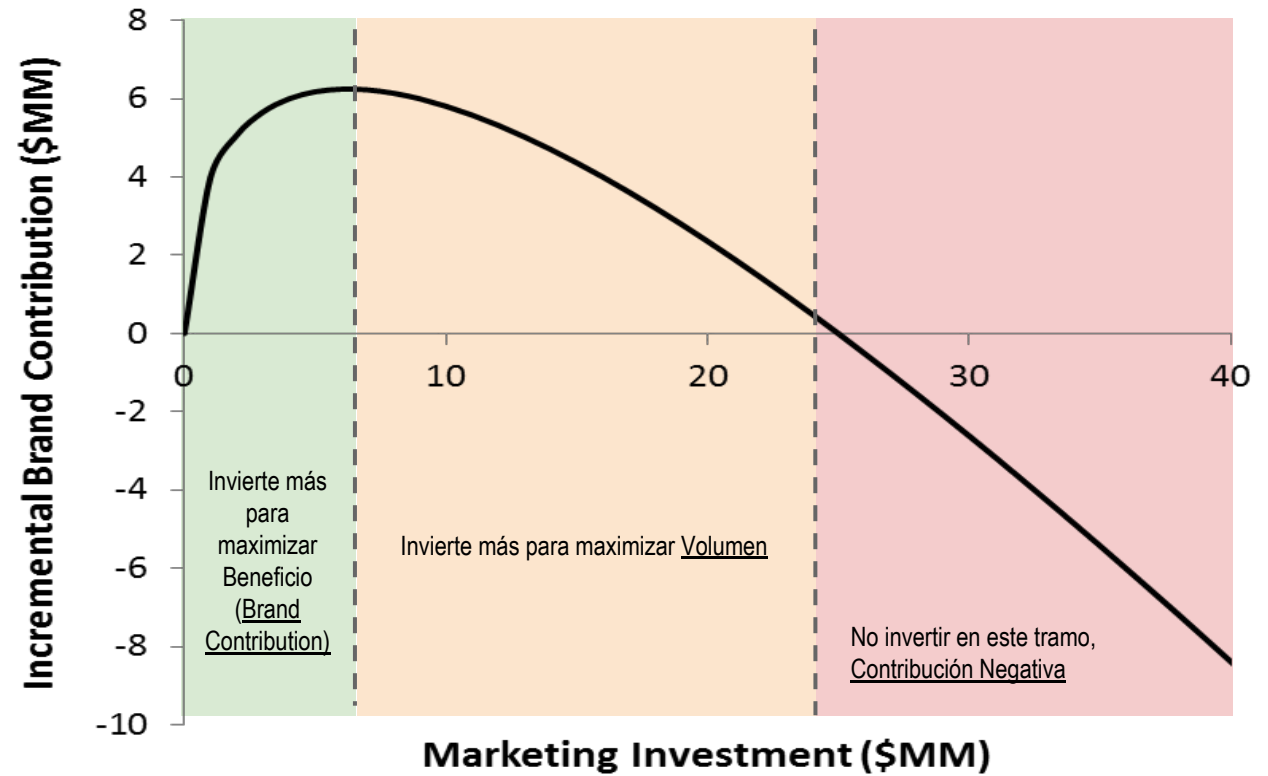
Curva de Respuesta

¿Cuánto de mis ventas se debe a mi inversión en publicidad en TV?



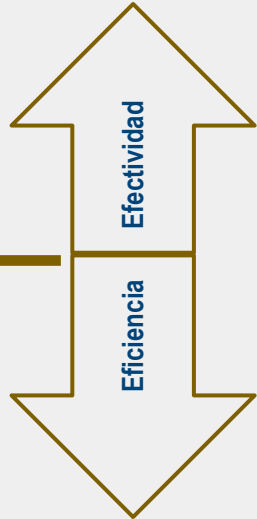
RESPONSE CURVE

¿Cuánto de mi beneficio incremental se debe a mi inversión en publicidad en TV?



PROFIT CURVE

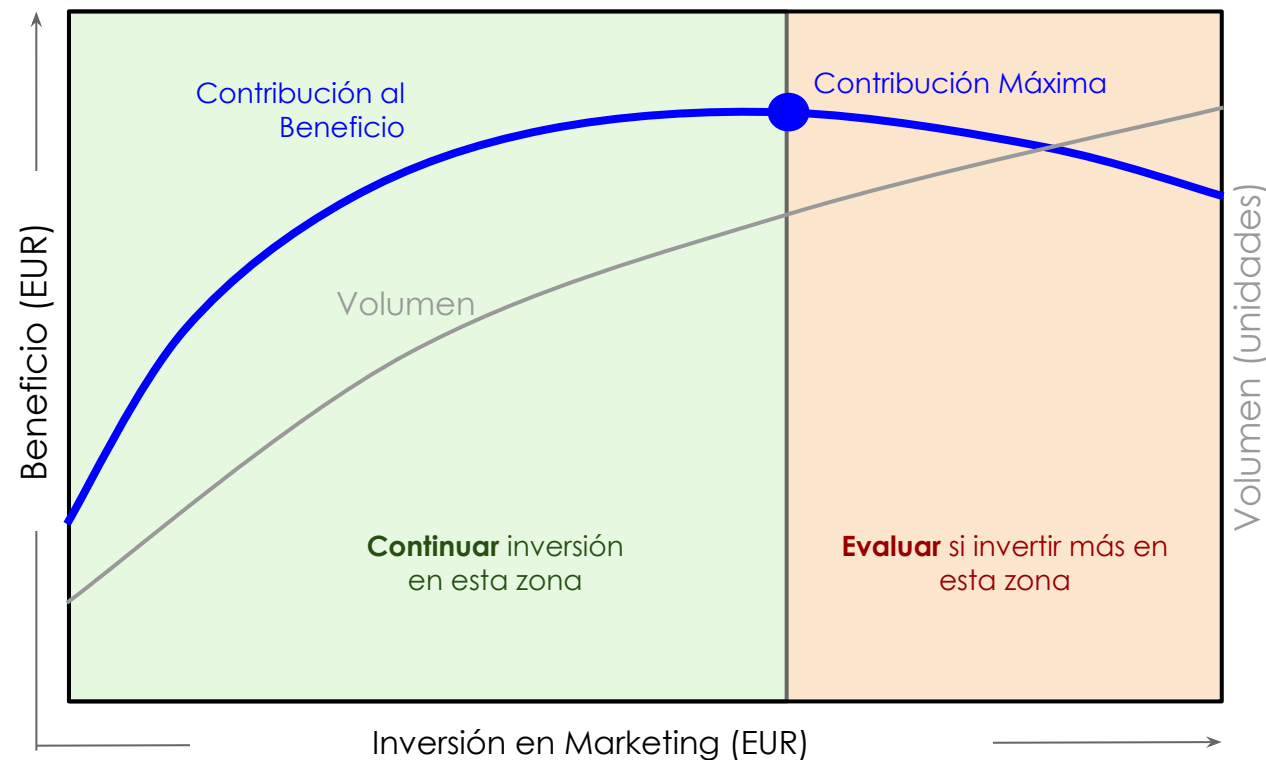
Return on Investment (ROI)

$$\text{ROI} = \frac{\text{Aumento en el Beneficio Bruto (Contribución)}}{\text{Gasto (Inversión)}}$$


El aumento en la Demanda obtenido del MMM, debe ser expresado en unidades monetarias (numerador).

El coste de la publicidad depende del tipo de acuerdo con la agencia y la configuración acordada (denominador)

En general, cuánto más se invierte en publicidad, más volumen e ingresos se obtienen. Sin embargo, el **beneficio no aumenta indefinidamente**, sino que después de alcanzar la contribución máxima empieza a decaer ya que el incremento de los ingresos no cubre el incremento de gasto necesario

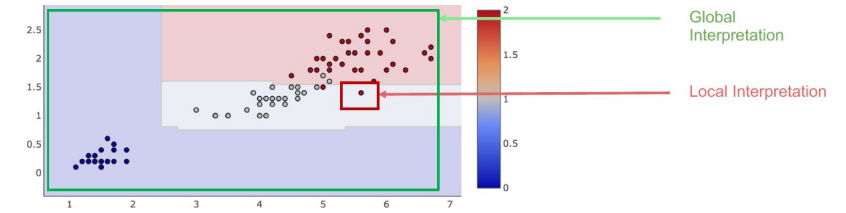


Interpretabilidad de Modelos

Interpretación de modelos: habilidad para explicar y presentar un modelo en una forma comprensible para los seres humanos (Doshi and Been, 2017)

Los modelos de *Machine learning* son conocidos por su alto poder predictivo, pero no por la facilidad de interpretación de sus resultados. Una regresión logística es un ejemplo modelo interpretable. Un modelo *gbm* es un ejemplo donde la interpretación no es trivial. Dada la importancia que tiene, desde la perspectiva de la toma de decisiones de negocio, explicar porque se considera que un cliente va a darse de baja o va a presentar una reclamación, hay un creciente número de publicaciones donde se presentan alternativas para abordar este tema.

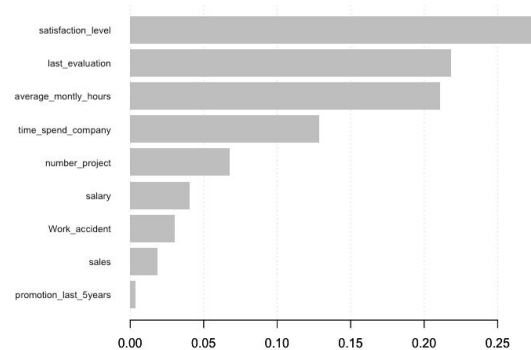
La interpretación de los modelos se puede realizar con dos perspectivas:



Interpretación GLOBAL

Permite explicar la interacción entre la variable respuesta (*target*) y las variables explicativas (*features*) utilizando el **conjunto de datos completo**.

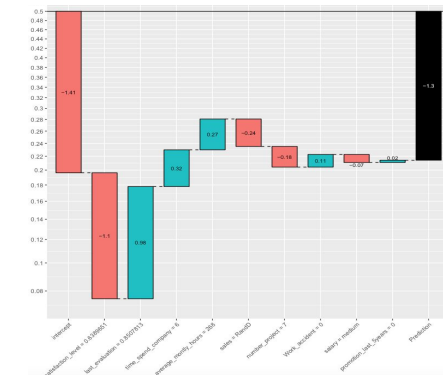
- Importancia de cada variable - **Feature importance:** ¿Cuáles son las variables (*features*) más importantes?
- Efecto de cada variable - **Feature effects:** ¿Cómo influye una característica en la predicción? - Efectos locales acumulados (ALE), gráficos de dependencia parcial (PDP) y curvas de expectativa condicional individuales (ICE) -



Interpretación LOCAL






Permite explicar la interacción entre la variable respuesta (*target*) y las variables explicativas (*features*) para una **única observación** o individuo.

- **Model-Agnostic Explanations:** ¿Cómo afectan los valores de cada variable a la predicción puntual? - Algunas opciones: LIME, Shapley value, BreakDown.
- Árboles sustitutos - **Surrogate trees:** ¿Podemos aproximar el modelo de caja negra subyacente con un árbol de decisión corto?





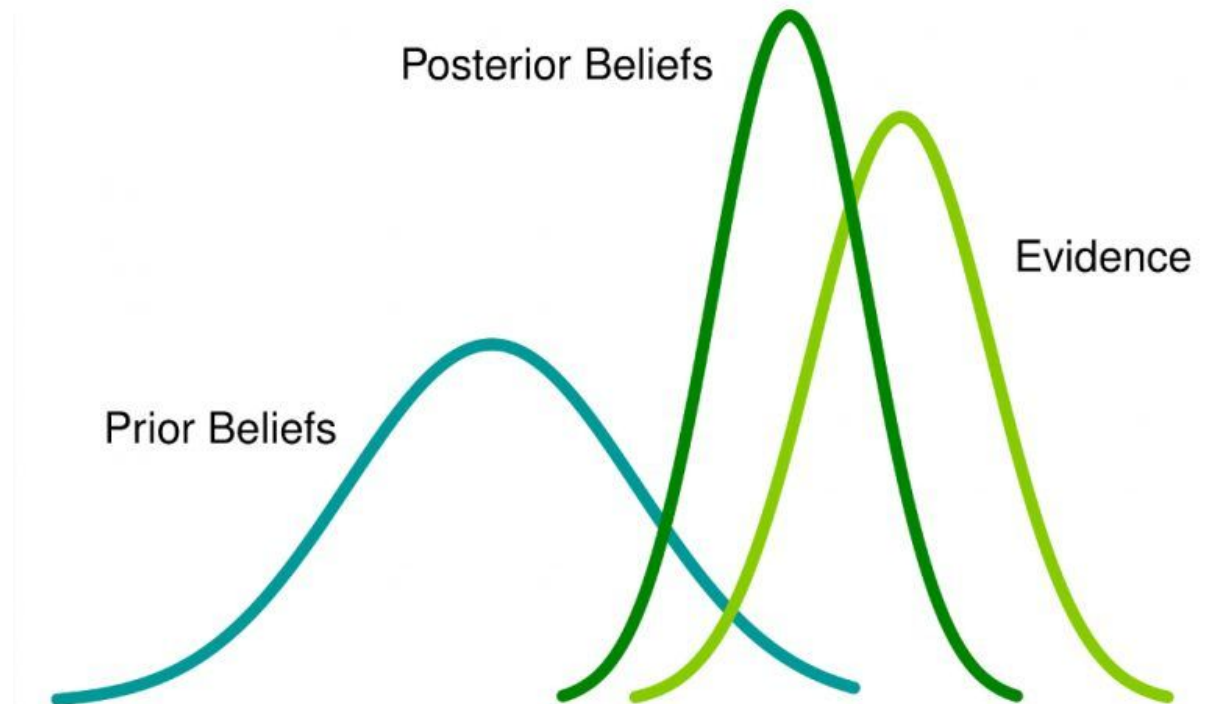
Inferencia Bayesiana

the theory  that would
 not die 
how bayes' rule cracked
 the enigma code,
hunted down russian
submarines & emerged
triumphant from two 
centuries of controversy
sharon bertsch mcgrayne



Inferencia Bayesiana

In which of my hypothesis should I believe in, and how strongly, given the collected data?



Regresión Lineal Simple

```
lm(formula = y ~ x)
```

Residuals:

Min	1Q	Median	3Q	Max
-0.8627	-0.5863	-0.4621	-0.2363	9.4099

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	-54.77802	7.47154	-7.332	4.47e-08	***
x	0.39172	0.03892	10.065	5.69e-11	***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

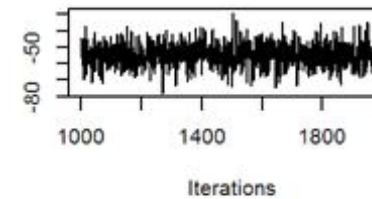
Residual standard error: 1.862 on 29 degrees of freedom

Multiple R-squared: 0.7775, Adjusted R-squared: 0.7698

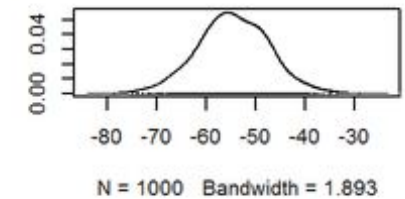
F-statistic: 101.3 on 1 and 29 DF, p-value: 5.692e-11

https://en.wikipedia.org/wiki/Bayesian_linear_regression

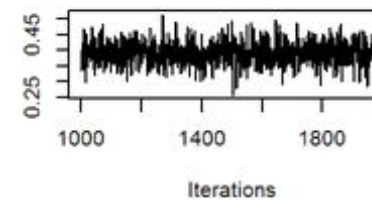
Trace of (Intercept)



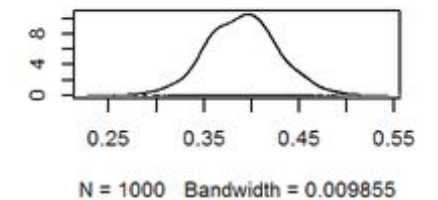
Density of (Intercept)



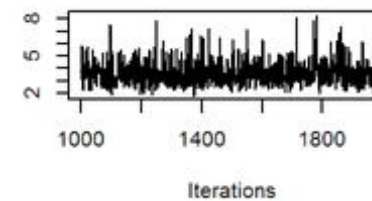
Trace of x



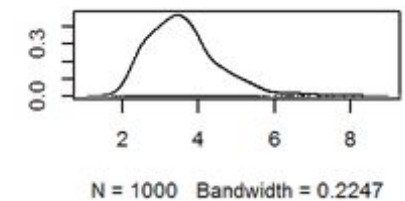
Density of x



Trace of sigma2



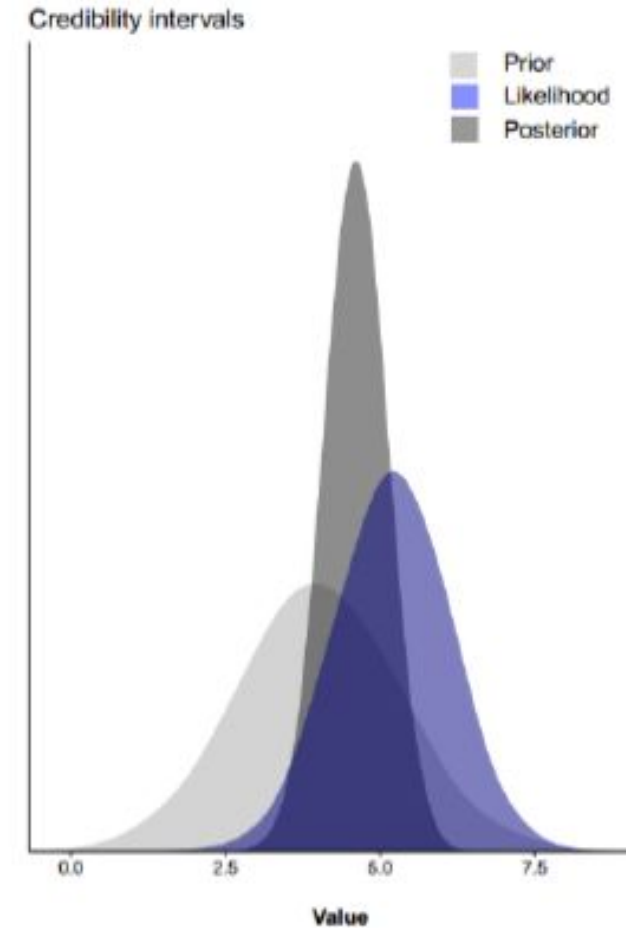
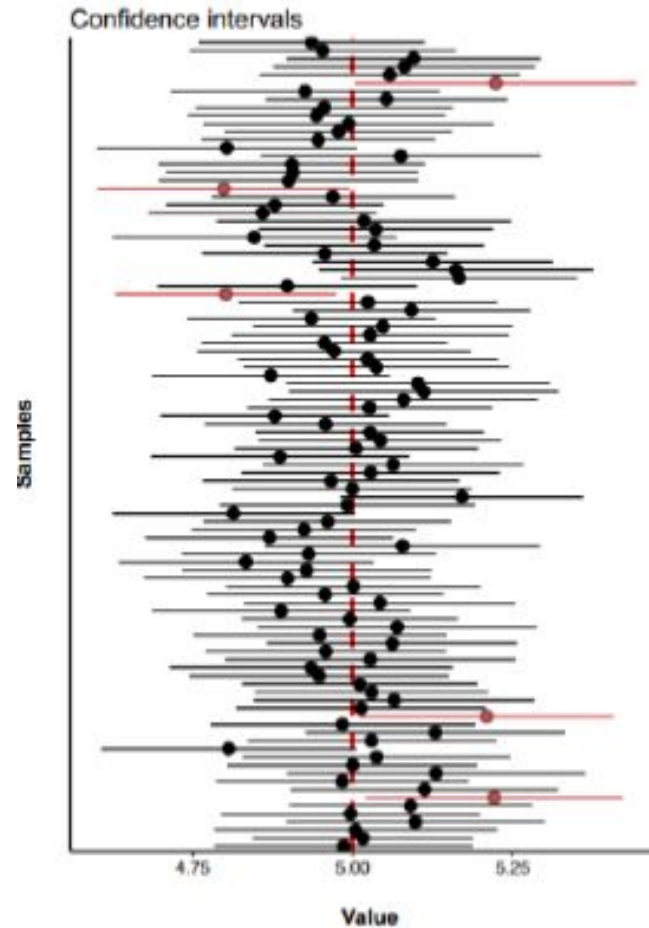
Density of sigma2



Incertidumbre

INTERVALOS DE CONFIANZA

Con un 95% de confianza, el intervalo contiene el verdadero valor del parámetro de la población?



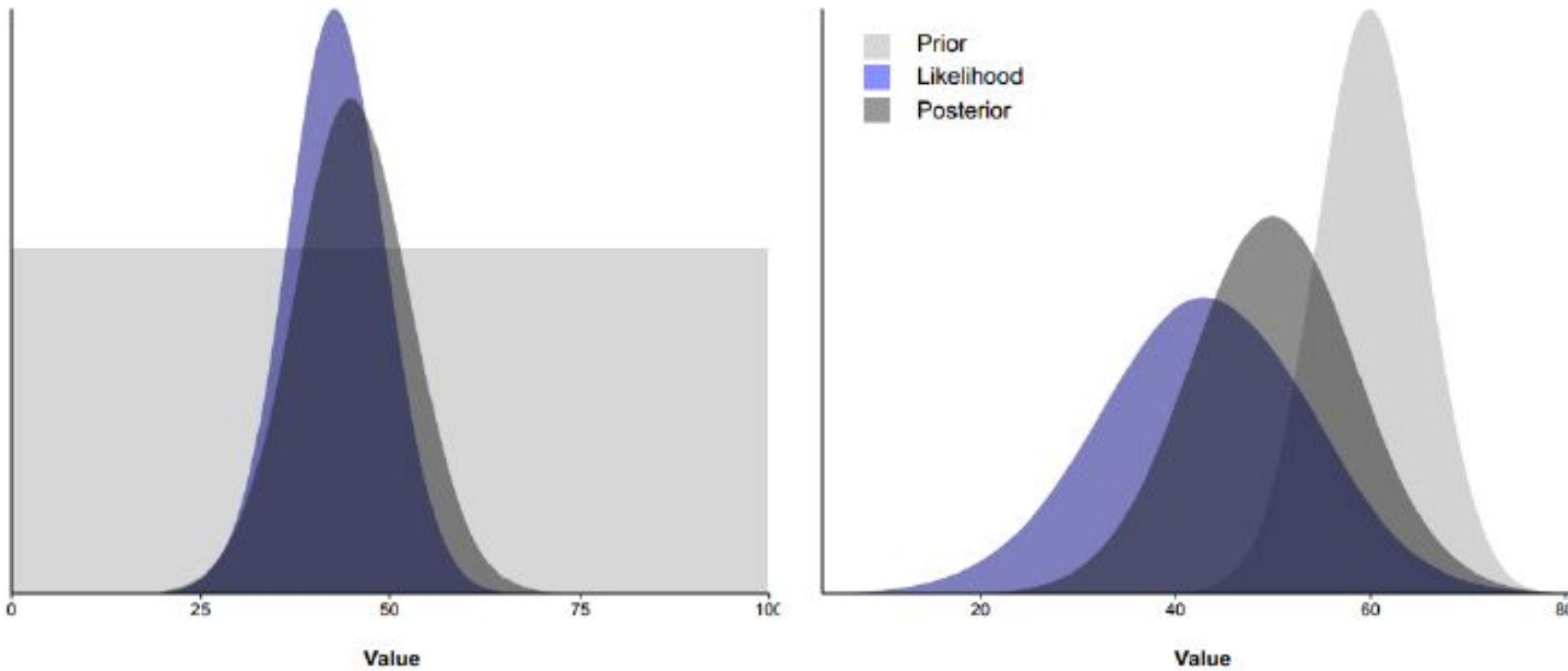
INTERVALOS DE CREDIBILIDAD

Existe una probabilidad de 0.95 de que el intervalo contenga el valor del parámetro de la población

Calidad del Prior

Less informative

More informative

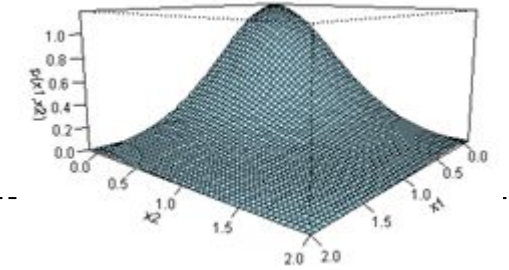


Formas de Priori

El parámetro es POSITIVO o NEGATIVO

Restricción de Dominio

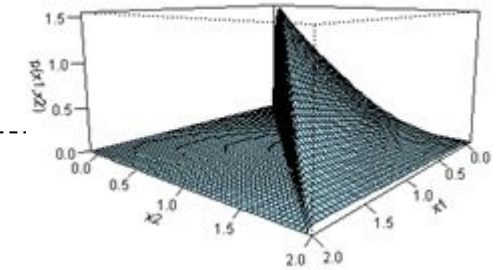
$$\beta_B > 0$$



Valor MEDIO del parámetro debe ser ...

Estudios previos y Teoría

$$\beta_i = m \pm v$$

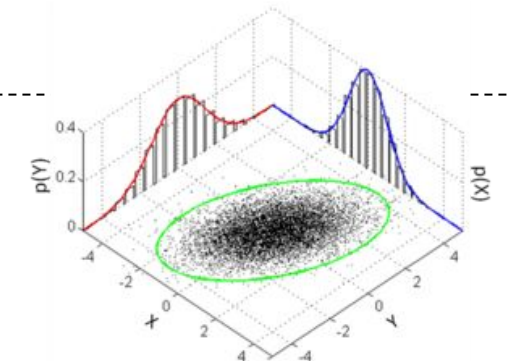


Estos parámetros son SIMILARES ...

Diversos grupos con diferentes respuestas

$$\beta_B \approx \beta_A + \xi$$

$$\beta_A \approx \beta_T + \psi$$



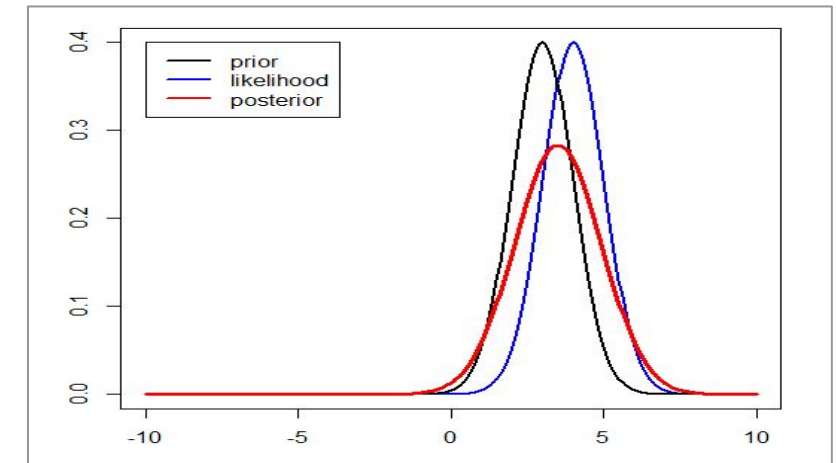
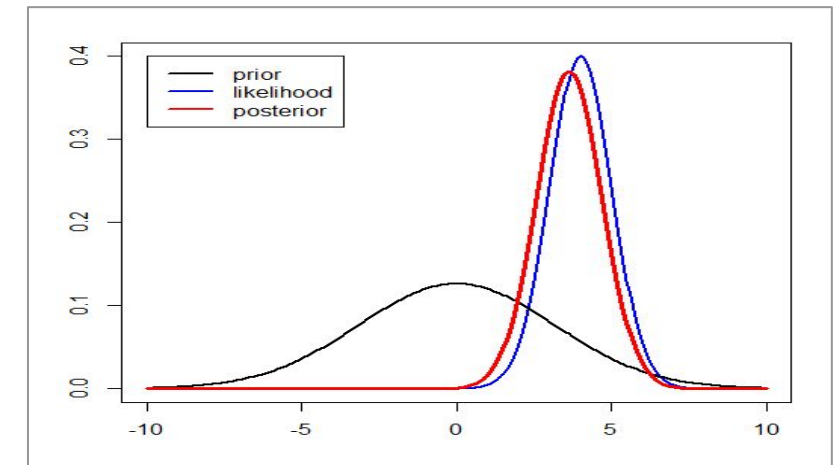
El ORDEN entre estos parámetros es ...

Un grupo tiene la mayor respuesta

$$\beta_1 > \beta_2$$

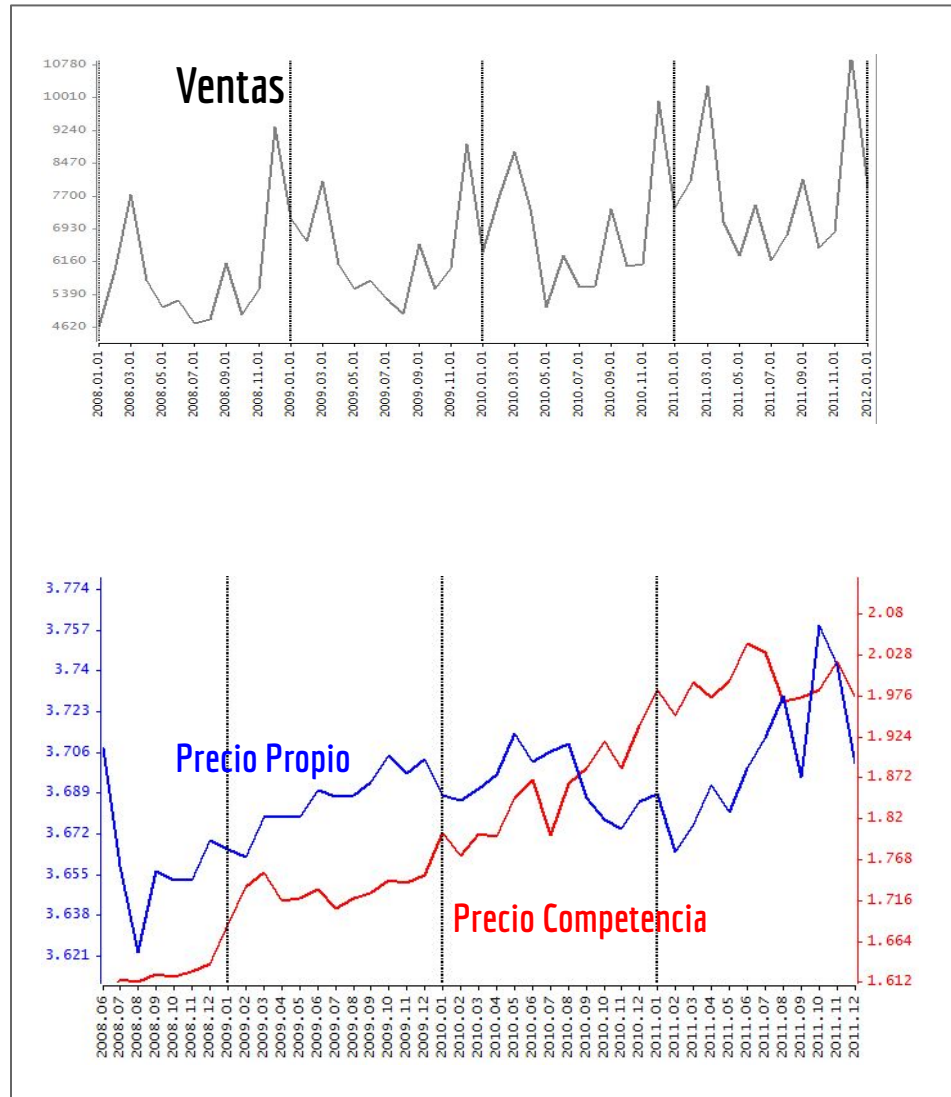
$$\beta_1 = \beta_2 + \beta_3$$

¿Dónde hay información a priori?



Teoría Económica, Experiencia & Conocimiento de Negocio, Investigaciones o Estudios Paralelos, Priors Revelados, etc.

Ejemplo: Elasticidad-Precio

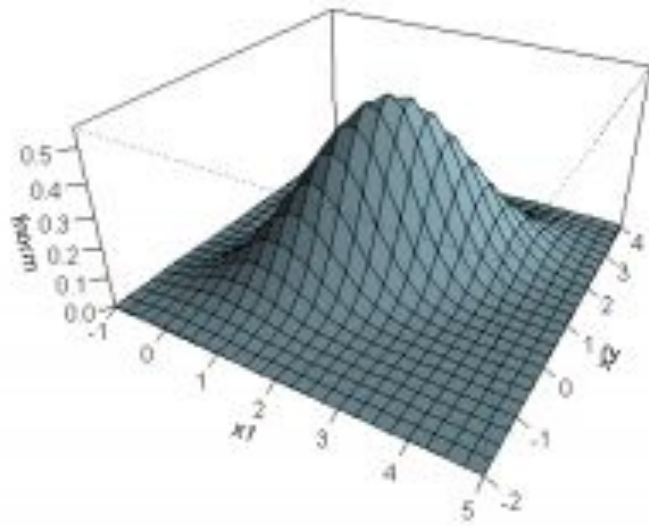


	PRIOR 01	PRIOR 02	PRIOR 03
Precio Propio	$N(-1,0.3) I[-2,0]$	$N(-2,1.5) I[-8,0]$	--
Precio Competencia	$N(0.2,0.1) I[0,1]$	$N(0.3,0.6) I[0,4]$	--

	PRIOR 01		PRIOR 02		PRIOR 03	
	mean	sigma	mean	sigma	mean	sigma
Precio Propio	-1.029	0.303	-2.280	1.285	-301.984	32.475
Precio Competencia	0.199	0.091	0.546	0.379	1.952	2.044
Temperatura P0	0.021	0.018	0.021	0.017	0.010	0.009
Temperatura P1	0.061	0.036	0.061	0.036	0.053	0.031
AR(1)	0.900	0.071	0.886	0.082	0.238	0.185
MA(12)	0.607	0.209	0.550	0.207	0.464	0.216

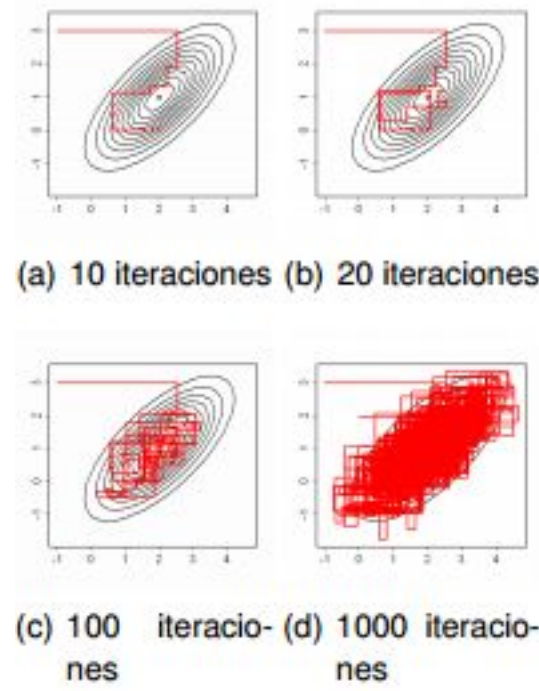
¿Cómo se estima?

MCMC - Gibbs Sampling

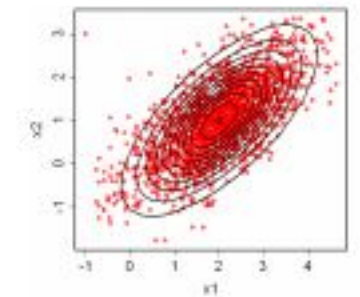


```
M <-1000
x <-c(-1,3)
xs <-x
for (i in 1:M){
  x[1]<-rnorm( 1,m[1]+S[1,2]/S[2,2]*(x[2]-m[2]),
              sqrt(S[1,1]-S[1,2]^2/S[2,2]) )
  x[2]<-rnorm( 1,m[2]+S[2,1]/S[1,1]*(x[1]-m[1]),
              sqrt(S[2,2]-S[2,1]^2/S[1,2]) )
  xs <-rbind(xs,x)
}
```

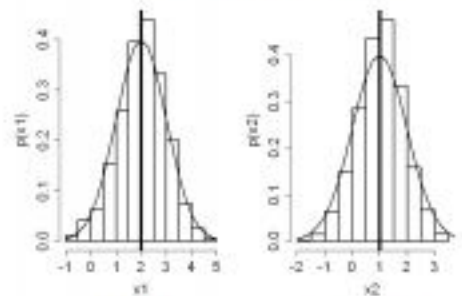
$$(x_1, x_2) \sim N(,) \text{ donde } = (m_1, m_2)' \text{ y } \begin{pmatrix} s_1 & s_{12} \\ s_{21} & s_2 \end{pmatrix}$$
$$x_1|x_2 \sim N(m_1 + (s_{12}/s_2) \times (x_2 - m_2); s_1 - s_{12}s_{21}/s_2)$$
$$x_2|x_1 \sim N(m_2 + (s_{21}/s_1) \times (x_1 - m_1); s_2 - s_{21}s_{12}/s_1)$$



(a) 10 iteraciones (b) 20 iteraciones
(c) 100 iteraciones (d) 1000 iteraciones



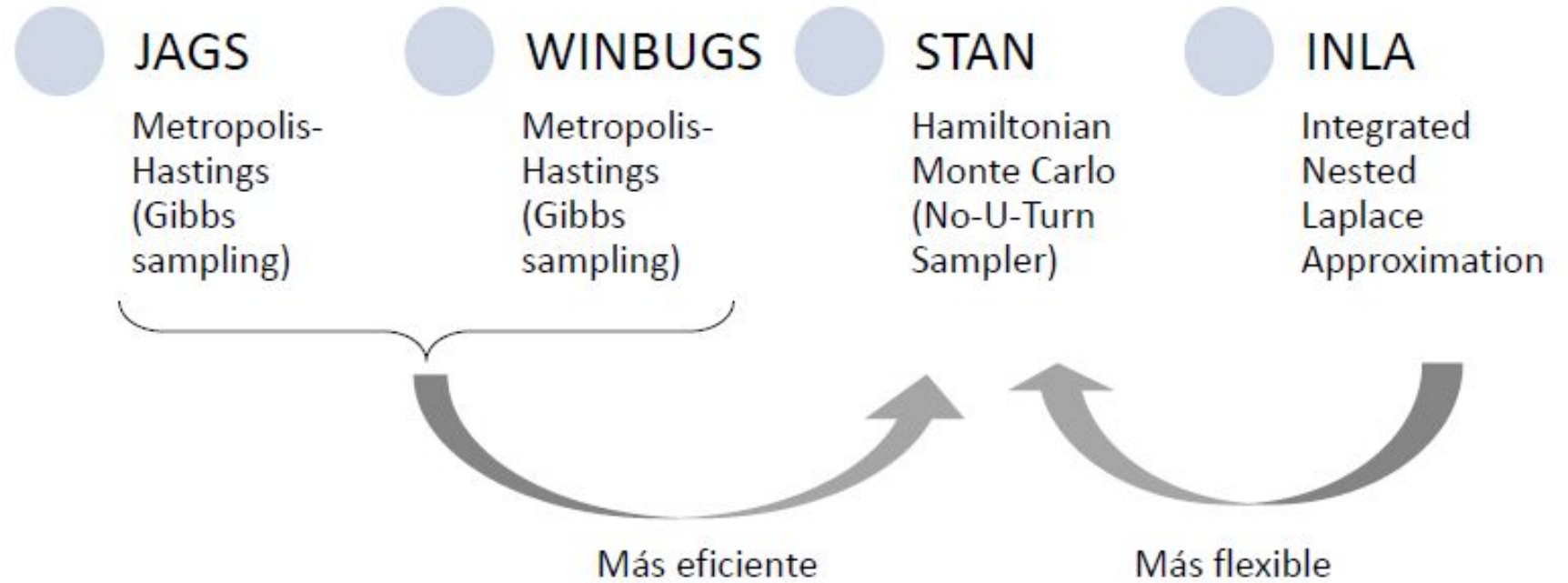
(a) Muestra conjunta



(b) Muestra marginal

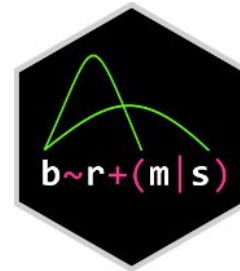
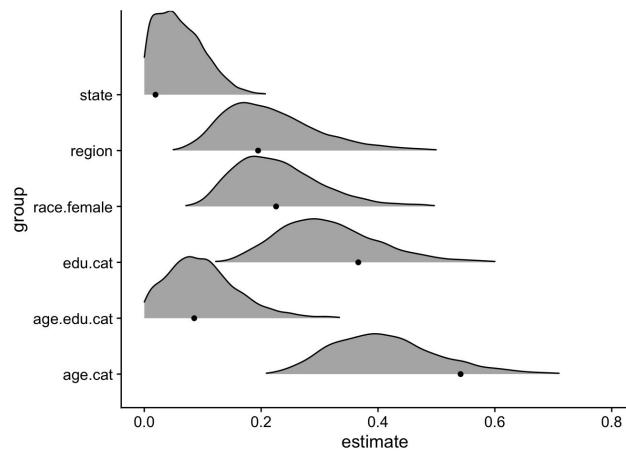
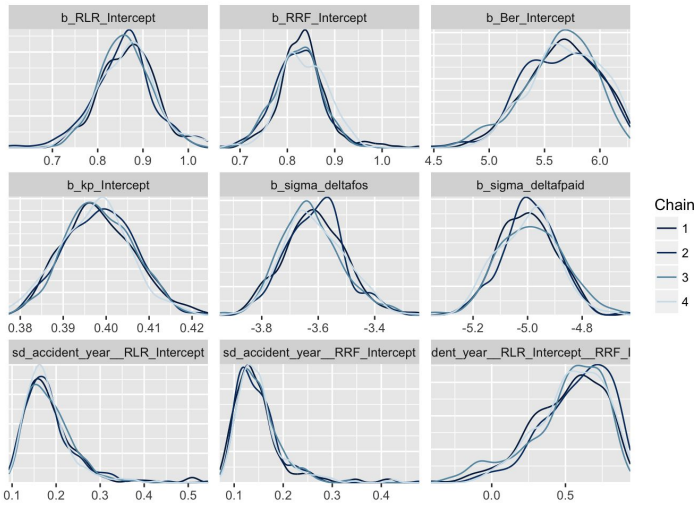
¿Dónde?

→ Script propio



→ Paquetes de R:
r2winbugs, rstanarm, mcmcpack

MCMC en R

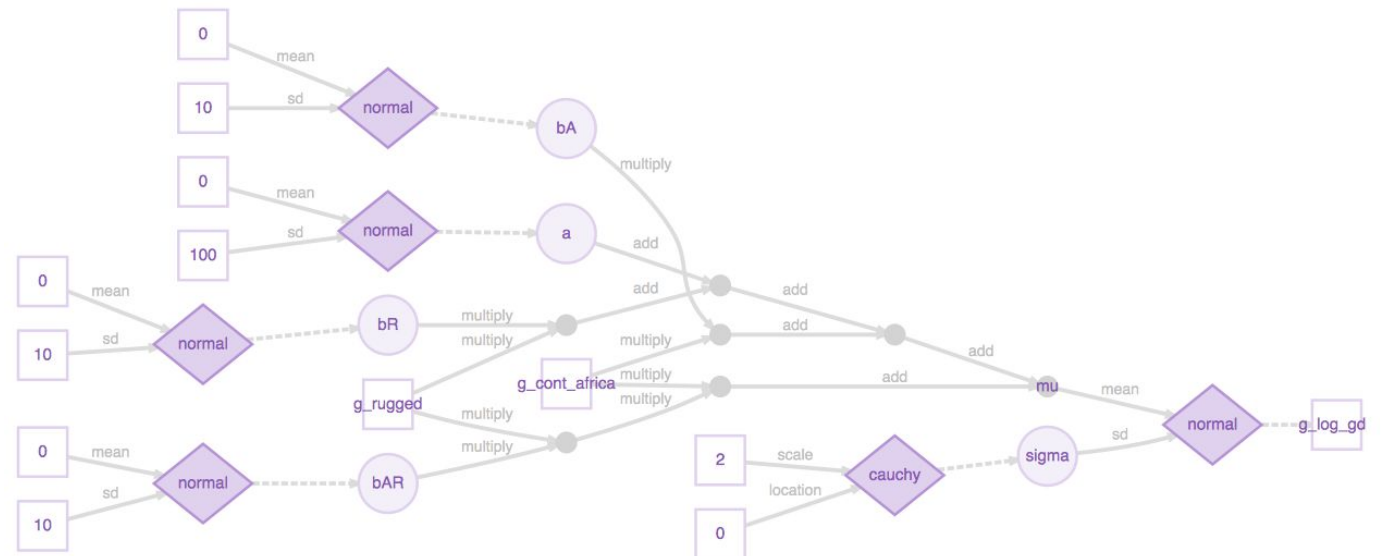


<https://bit.ly/2MxaCTB>

greta

<https://bit.ly/2MzCy9D>

<https://bit.ly/2tF7EFy>



Precios

Asumiendo que las elasticidades directas y cruzadas son constantes:

$$\log(V_t) = \alpha(B) \log\left(\frac{P_t}{CPI_t}\right) + \sum_i \beta_i(B) \log\left(\frac{CP_t}{P_t}\right) + R_t$$

ventas (pointing to $\log(V_t)$)
 precio propio (pointing to $\frac{P_t}{CPI_t}$)
 precio del i-ésimo competidor (pointing to $\frac{CP_t}{P_t}$)
 IPC (pointing to $\frac{P_t}{CPI_t}$)
 otros componentes (pointing to R_t)

- Los parámetros deben satisfacer:

un aumento en el precio de la competencia debe estimular las ventas

$$\begin{aligned} \alpha &< 0 \\ \beta_i &> 0 \\ \alpha - \sum_i \beta_i &< -1 \end{aligned}$$

un aumento en el precio propio mientras el IPC se mantiene estable debe producir un efecto negativo

elasticidad de la marca al precio propio

- Además...
 - los efectos absolutos derivados de los cambios en precios relativos entre productos que compiten también deberían estar restringidos
 - las respuestas al precio podrían ser no instantáneas, especialmente para canales de consumo inmediato y packs.

El papel del Enfoque Bayesiano



Permite la **acumulación** de conocimiento.



Permite el **diálogo** de los datos con el conocimiento de negocio y por ello, adecuada para modificar dicho conocimiento con el peso de la evidencia empírica.



El factor crítico de todo este proceso es la **diagnosis**, que compara los supuestos teóricos con la evidencia.



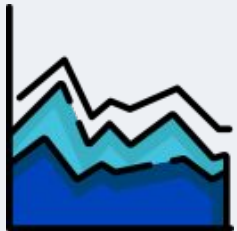
Resumiendo



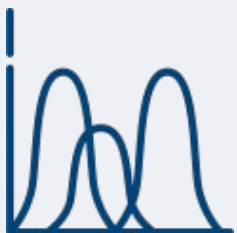
**Orientación
a Negocio**



**Naturaleza
de los Datos**



Explicatividad



**Teorema
de Bayes**



Q & A





Thank you

Living in Real Time

Data followers | Solutions Leaders