

La delgada línea entre predictivo y Análisis prescriptivo (con ejemplos)

Y cómo subir las escaleras hacia el estado del arte de Analytics



Víctor Almeida · Seguir

Publicado en Cultura Geek

Lectura de 8 minutos · 14 de abril de 2021



Escuchar



Compartir



Más

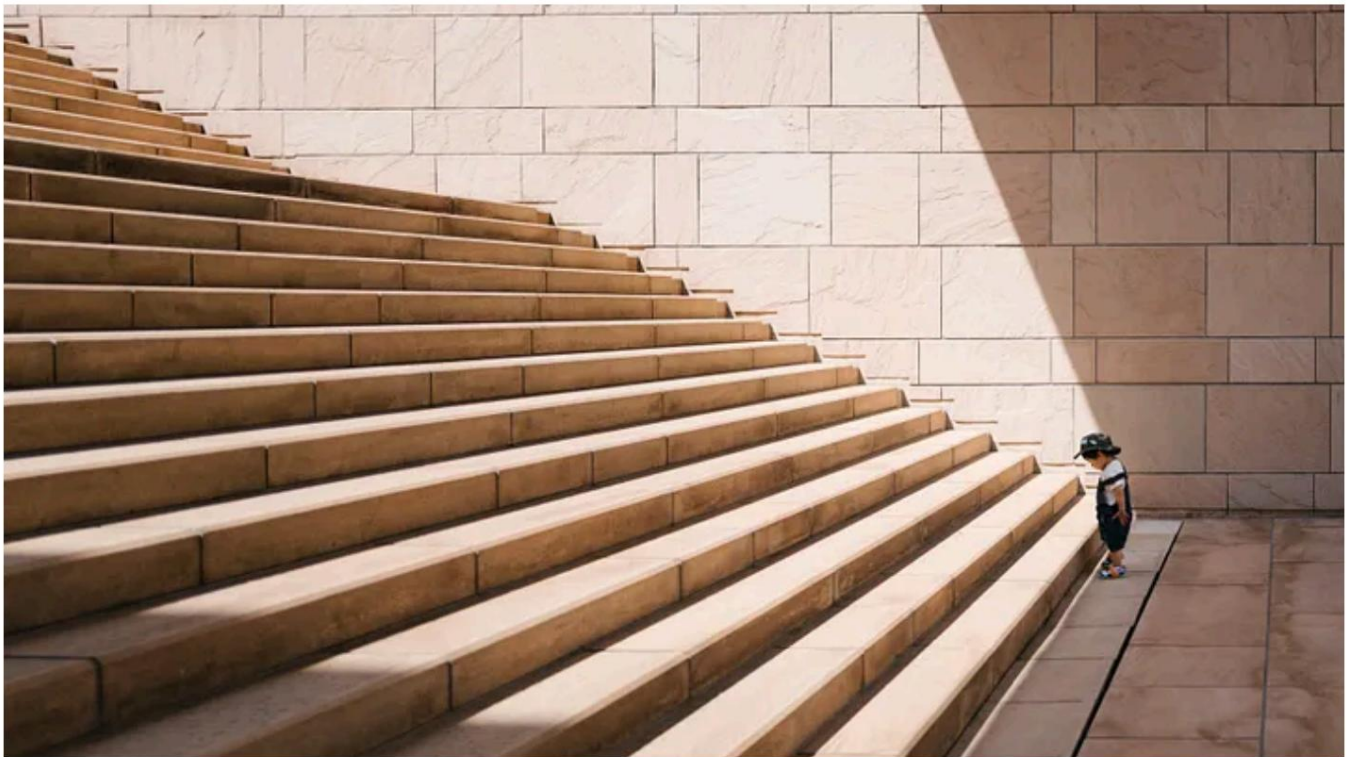
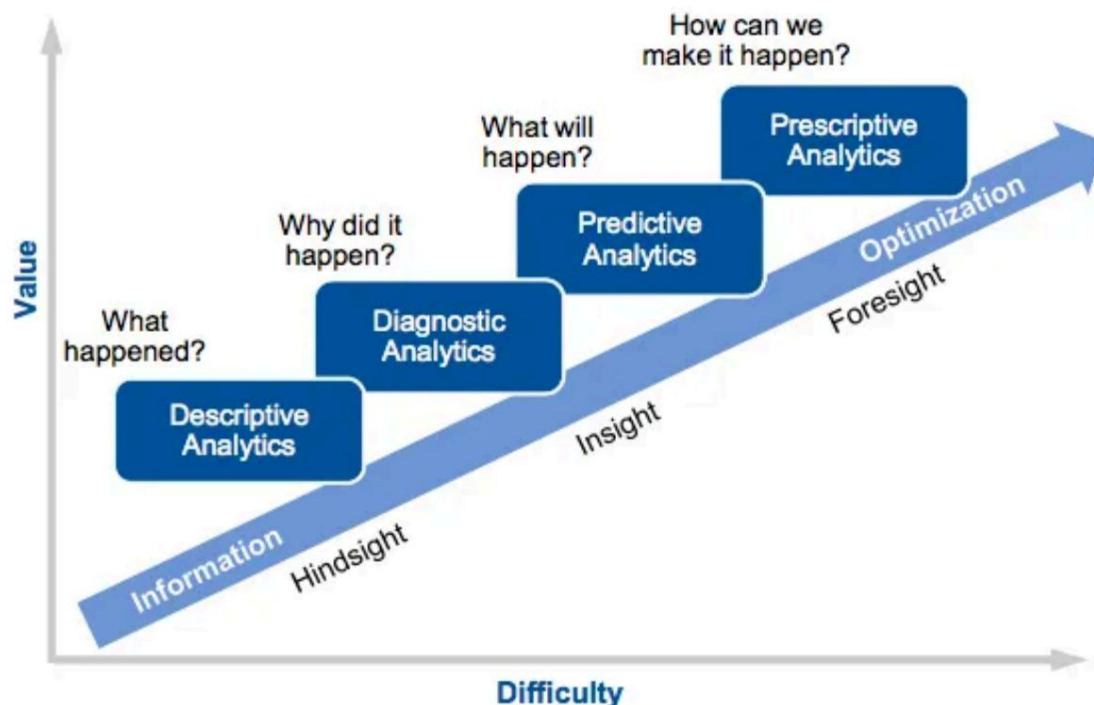


Foto de [Jukan Tateisi](#) en [Unsplash](#)

Desde 2012, hemos escuchado ese famoso dicho de Gartner de que la analítica, en un De manera general, tenemos cuatro niveles antes de llegar al estado del arte. Aunque esto artículo fue publicado hace 9 años, representa bien (incluso demasiado) la realidad de la mayoría de las empresas.

Los cuatro pasos, o los cuatro tipos principales de análisis, se pueden definir como:

- Descriptivo: ¿Qué ha sucedido?
- Diagnóstico: ¿Por qué sucedió?
- Predictivo: ¿Qué pasará?
- Prescriptivo: ¿Cómo podemos lograrlo?



Source: Gartner (March 2012)

El famoso modelo de ascendencia de Gartner

Análisis descriptivo Podemos

entender fácilmente los dos primeros ya que su idea se ha difundido entre las empresas.

Podemos decir que el análisis descriptivo llegó en la primera generación de BI e hizo famosas las hojas de cálculo de Excel. La idea principal es que podemos describir (significado más obvio) los números de nuestra organización:

- ¿Cuáles son los ingresos mensuales?
- ¿Cuánto gastamos en determinados gastos?
- ¿Cómo son los niveles de inventario?

Es el primer paso para comprender nuestros datos, trabajando sobre información del pasado.

Análisis de diagnóstico Una

vez que sabemos lo que sucedió en el pasado, el segundo paso es el análisis de diagnóstico, preguntando por qué sucedió. Esto se soluciona utilizando la segunda generación de BI, con la que podemos correlacionar diferentes datos, provenientes de diferentes fuentes. Algunas preguntas de ejemplo son:

- ¿Por qué nuestros ingresos son siempre mayores al final del año?
- ¿Por qué gastamos tanto en gastos de automoción?
- ¿Por qué nuestro inventario está disminuyendo tan rápidamente?

Este siguiente nivel requiere un poco de curiosidad por parte del usuario y voluntad de resolver problemas.

Aquí nos centramos en los problemas reales y por qué ocurren.

Si miramos más de cerca, podríamos decir que la analítica descriptiva mira hacia el pasado mientras que el diagnóstico tiende a observar el presente. Según esta lógica, pensaréis que el siguiente nivel debería ser mirar hacia el futuro. Y adivina qué, ¡tienes razón!

Lo que la mayoría de la gente no entiende es que sólo el análisis predictivo observa el futuro, pero no el prescriptivo.

Y luego, después de toda esta madurez analítica 101, volvemos al punto principal de este artículo.

¿Cuál es la diferencia entre análisis predictivo y prescriptivo?

El tercer y cuarto nivel se hicieron famosos con la llegada de la ciencia de datos y el aprendizaje automático. No confundamos estos términos con "prever" el futuro, esto sigue siendo imposible incluso con la tecnología que tenemos hoy en día. Lo que hacemos los científicos de datos es inferir (o predecir) el escenario más probable, basándonos en datos históricos. Entonces, las verdaderas preguntas no son como:

- ¿Cuáles serán mis ingresos durante los próximos 6 meses?
- ¿Cuánto gastaremos en gastos automotrices el próximo mes?
- ¿Cuánto inventario debemos tener para no tener atrasos en las próximas semanas?

Sino más bien:

- ¿Cuáles serán mis ingresos durante los próximos 6 meses en función de los últimos dos años?

- ¿Cuánto gastaremos en gastos automotrices el próximo mes según el histórico?
- ¿Uso de nuestra flota, marca del vehículo y número de pedidos de entrega abiertos?
- ¿Cuánto stock deberíamos tener para no tener retrasos en las próximas semanas según RFM?
- ¿Y número de campañas realizadas el mes pasado?

El Machine Learning siempre tiene que basarse en datos históricos, no podemos inferir algo surgido de la nada. Y además, utilizamos las preguntas respondidas en el diagnóstico. fase como etapa inicial del desarrollo del modelo.

Lo que es importante tener en cuenta es el resultado del análisis predictivo:

- ¿Cuáles serán mis ingresos durante los próximos 6 meses en función de los últimos dos años?

Años: \$14.242.924,52

- ¿Cuánto gastaremos en gastos automotrices el próximo mes según el histórico?
- ¿Uso de nuestra flota, marca del vehículo y número de pedidos de entrega abiertos?

Años: \$241,242.08

- ¿Cuánto inventario deberíamos tener para no tener retrasos en las próximas semanas según ¿RFM y número de campañas realizadas el mes pasado?

Respuesta: 6236 artículos

Y también obtenemos un modelo, que puede ser una ecuación matemática, que usamos para generar estos números, por ejemplo:

Ingresos = 1152,94 + 32,98 * <número de productos vendidos> + 790 * <número de campañas hecho>

Por ejemplo, digamos que vendimos 35.743 productos e hicimos 15 campañas. Obtenemos:

Ingresos = 1152,94 + 32,98 * 35.743 + 790 * 15

Ingresos = \$ 1.191.807,08

Por supuesto, utilizamos técnicas mucho más avanzadas para crear modelos, pero esa es la idea principal del aprendizaje automático: crear una 'ecuación' basada en datos históricos para predecir el escenario futuro más probable. Concéntrate en esa palabra, probable.

¿Y qué pasa con el análisis prescriptivo? Desperdiciamos nuestra línea de tiempo: pasado, presente y futuro. ¿Qué queda?

En análisis prescriptivo nos centramos en lograr estas cifras probables. Las preguntas ahora son:

- ¿Cómo logramos unos ingresos de \$14.242.924,52?
- ¿Cómo presupuestamos los gastos automotrices en \$241,242.08?
- ¿Cómo garantizamos que no tendremos retrasos en determinados artículos?

Y algunas de las respuestas podrían ser:

- ¿Cómo logramos unos ingresos de \$14.242.924,52?

Respuesta: Reducir el margen en un 3% y aumentar las ventas en un 10%

- ¿Cómo presupuestamos los gastos automotrices en \$241,242.08?

Respuesta: Utilice esta ruta para el vehículo 1, esta ruta para el vehículo 2...

- ¿Cómo garantizamos que no tendremos retrasos en determinados artículos?

Respuesta: Utilice los 6236 elementos como punto de referencia y agregue un 10% más como margen de seguridad. Además, compre estos artículos con estos proveedores: <lista>

En resumen, la analítica prescriptiva se centra en la decisión y/o la acción.

La principal diferencia entre el análisis predictivo y el prescriptivo es que, en el análisis predictivo, tenemos una máquina que nos ayuda a tomar decisiones, mientras que en el análisis prescriptivo tendremos la máquina que nos dirá qué hacer para lograr los números que obtuvimos en el análisis predictivo. Si usaremos o no la recomendación de la máquina será una decisión humana.

Pero ¿por qué necesitamos un modelo predictivo para construir uno prescriptivo?

Es hora de practicar

Pongamos un ejemplo. La fijación de precios es una práctica minorista de definir un precio que optimiza las ganancias (no los ingresos). En el ejemplo, usaremos una base de datos privada y no nos centraremos en el código, sino en el caso de negocio.

Ya sabemos por los cursos básicos de administración en la academia que:

Beneficio = Ingresos - Costo

Ingresos = Cantidad * Precio

Entonces,

Beneficio = Cantidad * Precio de coste

Podemos estimar un modelo de demanda simple, donde apuntamos a la cantidad en función de la precio solo. Y sabemos que cuanto mayor es el precio, menos productos se venden.

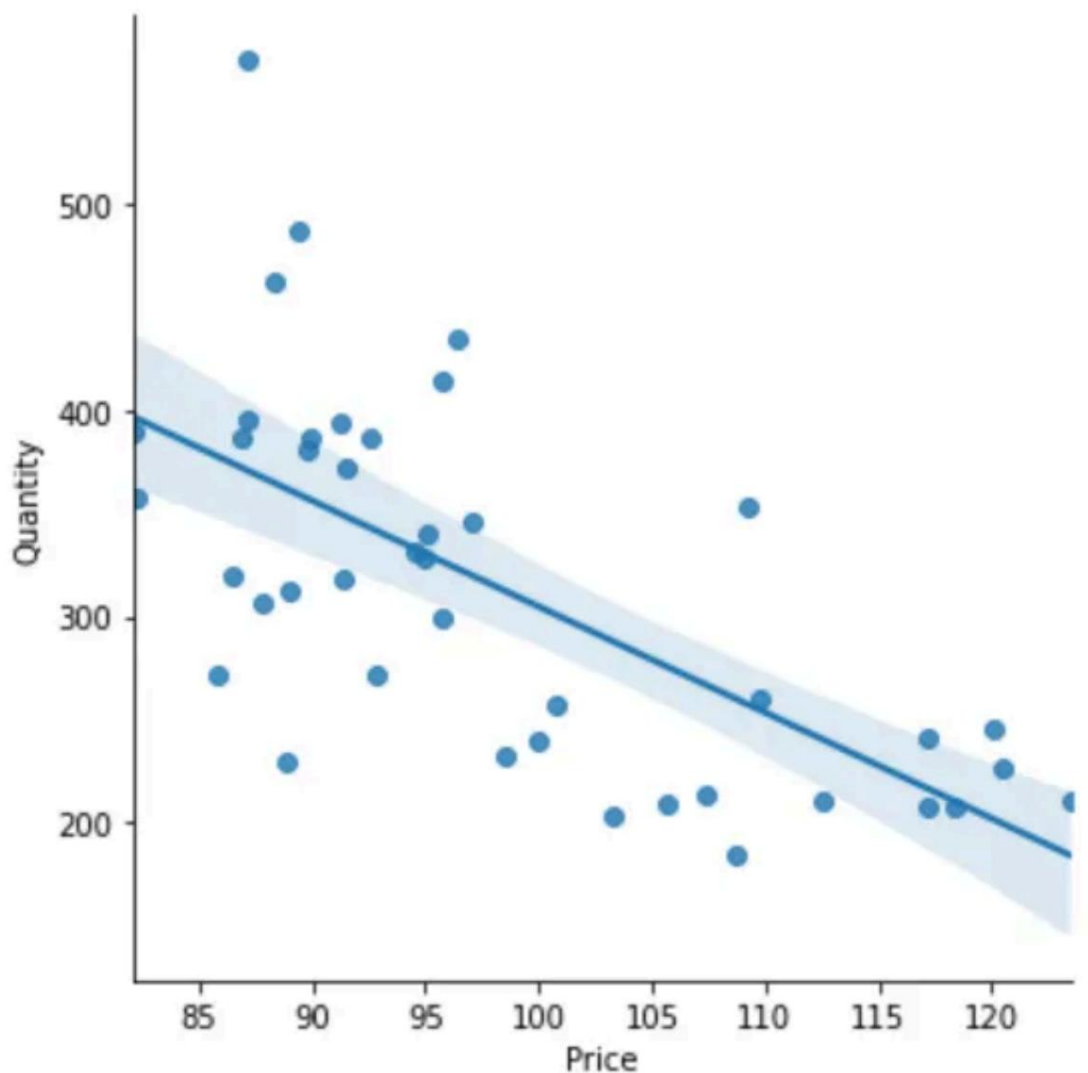
Usemos algunos datos sobre un determinado producto en el comercio minorista:

		Price	Quantity
Year	Month		
2016	01	103.31	203.0
	02	108.66	184.0
	03	112.57	211.0
	04	107.40	213.0
	05	117.23	241.0
Abrir en la aplicación ↗			

<

Ventas mensuales de un solo producto aleatorio.

Si trazamos estos datos, podemos ver un patrón a la baja. Como esperábamos, cuanto más
Cuanto mayor sea el precio, menos productos venderemos.



Modelo lineal para precio x cantidad (hecho con Seaborn)

De la teoría de la demanda, deberíamos esperar la ecuación lineal ($y = ax + b$) como:

$$\text{Cantidad} = -a \cdot \text{Precio} + b$$

Donde los coeficientes a y b estarán definidos por el modelo estadístico. el menos
firmar en a refleja la tendencia a la baja.

Usaremos el modelo más simple disponible para nosotros, una Regresión Lineal (usando Ordinario
Mínimos Cuadrados u MCO). Una vez que creamos el modelo, obtenemos los siguientes resultados:

OLS Regression Results

Dep. Variable:	Quantity	R-squared:	0.439			
Model:	OLS	Adj. R-squared:	0.425			
Method:	Least Squares	F-statistic:	31.28			
Date:	Tue, 13 Apr 2021	Prob (F-statistic):	1.76e-06			
Time:	14:22:35	Log-Likelihood:	-235.89			
No. Observations:	42	AIC:	475.8			
Df Residuals:	40	BIC:	479.3			
Df Model:	1					
Covariance Type:	nonrobust					
	coef	std err	t	P> t 	[0.025	0.975]
Intercept	818.7170	90.812	9.016	0.000	635.180	1002.254
Price	-5.1388	0.919	-5.593	0.000	-6.996	-3.282
Omnibus:	4.248	Durbin-Watson:	1.483			
Prob(Omnibus):	0.120	Jarque-Bera (JB):	3.031			
Skew:	0.602	Prob(JB):	0.220			
Kurtosis:	3.533	Cond. No.	853.			

Modelo creado con python lib statsmodels

Aunque hay mucha información aquí, nos centraremos en la columna coef, donde obtenemos la información para la Intercepción y el Precio. El resto de La información nos informa sobre el rendimiento del modelo, que no cubriremos aquí.

Nuestro modelo Cantidad x Precio ahora se convierte en:

$$\text{Cantidad} = 818,72 - 5,14 * \text{Precio}$$

Por supuesto, tenemos un error asociado con esta ecuación, podemos notarlo en la Cantidad x Precio traza un área azul más clara, ese es el error. Una gran parte de nuestra jobs es minimizar este error tanto como sea posible, para que obtengamos resultados más precisos.

De todos modos, la ecuación final que obtenemos es la ecuación que representa la más probable.

(aquí está esa palabra importante nuevamente) escenario que ocurrirá.

Una vez que tengamos este modelo (predictivo), podemos sustituir en nuestra ecuación de ganancias:

$$\text{Beneficio} = \text{Cantidad} * \text{Precio} - \text{Costo} = (818,72 - 5,14 * \text{Precio}) * \text{Precio de coste}$$

Notamos que tenemos el precio al cuadrado y nuevamente regresamos a nuestra universidad.

clases, sabemos que estas ecuaciones cuyos coeficientes están elevados a 2 (o

ecuaciones de segundo orden) tienen una curva, de la cual podemos estimar el punto máximo

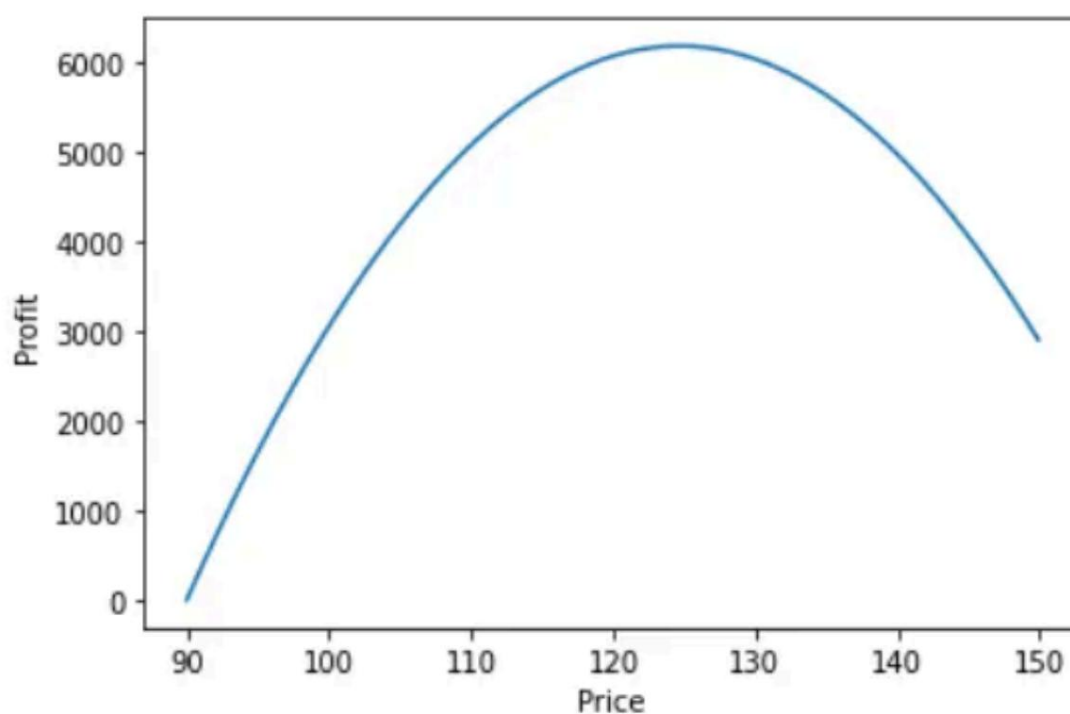
que representa el máximo beneficio que podemos obtener de un determinado producto.

Para el costo, asumiremos un valor fijo, pero como habrás notado, podríamos

También crear un modelo de costos basado también en la demanda. Pero mantengamos las cosas simples,

Supondremos que el costo es de \$90.00 por producto. Trazando esta ecuación obtendremos

siguiente cuadro:



Curva de beneficios

Donde podemos obtener fácilmente el máximo beneficio fijando el precio en 124,70 dólares y

eso, fin de nuestro análisis prescriptivo.

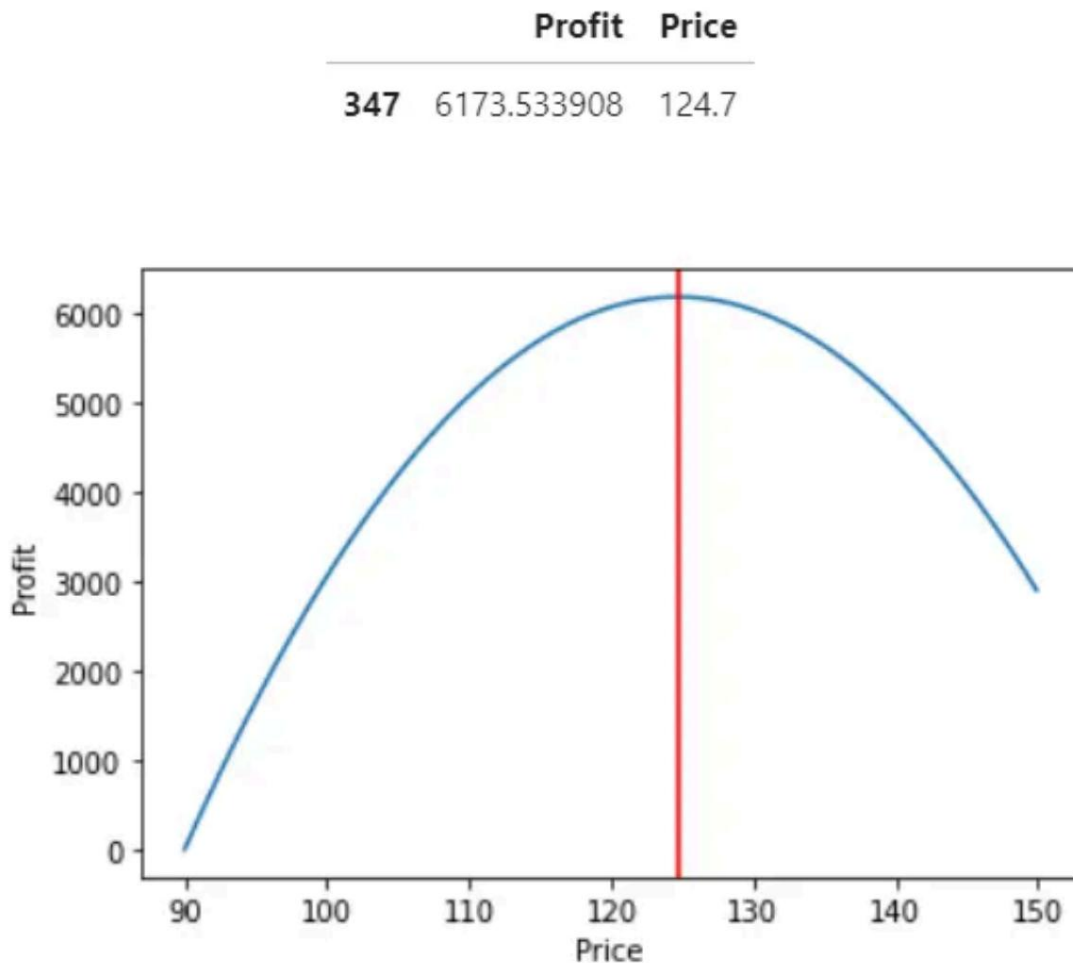


Tabla de valor máximo y valor.

Al final del día lo que tendremos es un robot recomendando a tu empresa un precio por cada producto que vendes.

Envolver

Espero que hayas disfrutado leyendo este artículo mientras lo escribo y espero que te ayude a comprender dónde se encuentra y dónde quiere estar en su viaje analítico.

Tenga en cuenta que cada nivel analítico es un paso hacia el siguiente, así que no se apresure a llegar al siguiente. última fase sin cruzar los niveles anteriores. Cada uno de los principales tipos de Los análisis le brindarán la información correcta que necesita para ascender al siguiente.

Ciencia de los datos

Hacia la ciencia de datos

Analítica

Aprendizaje automático

Análisis prescriptivo

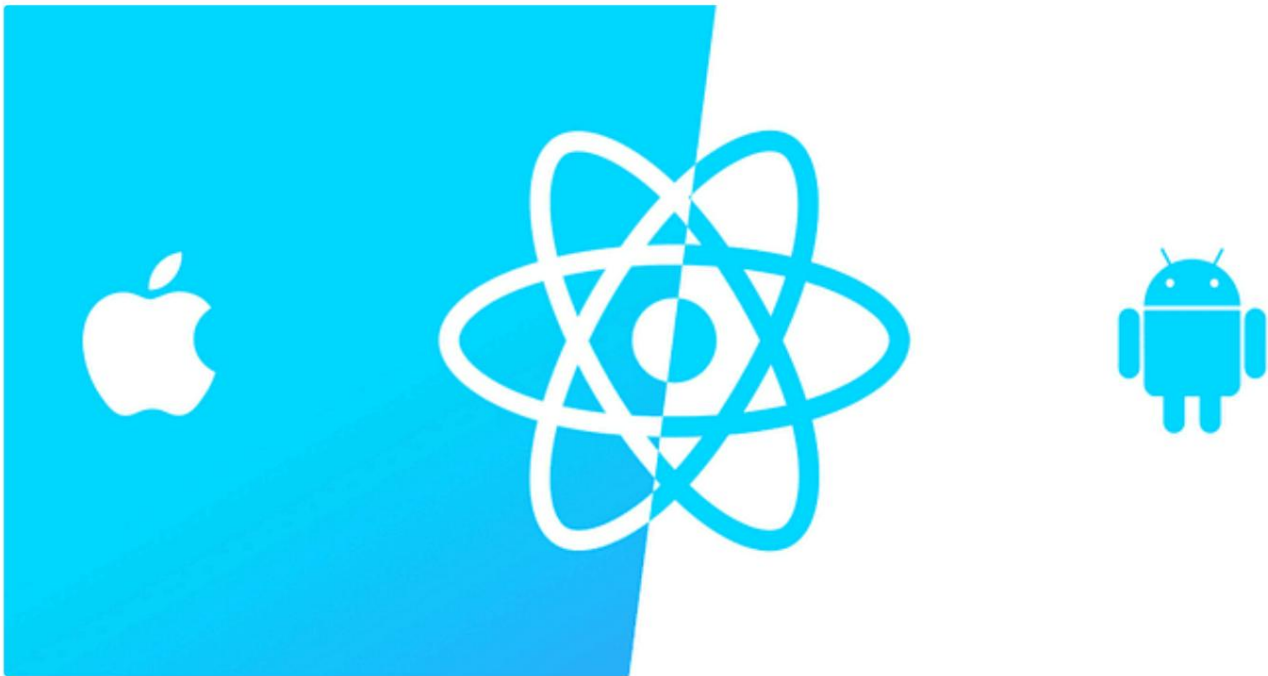
[Seguir](#)

Escrito por Víctor Almeida

9 seguidores · Escritor para Cultura Geek

Entusiasta de la tecnología, una persona tremendamente curiosa, que intenta hacer que la gente entienda cosas complejas. Además, un científico de datos por pasión.

Más de Víctor Almeida y Cultura Geek



Anshul BorawakeCultura geek

React Native Generate APK: depurar y publicar APK

Generar depuración y lanzamiento de APK en React Native; Windows, iOS y Linux

Lectura de 3 minutos · 3 de abril de 2021



1,7K



11





Cultura geek de Masud Afsar

¿Cómo instalar Node.js mediante NVM?

Instale y administre múltiples versiones de Node.js con nvm

Lectura de 3 minutos · 14 de septiembre de 2021



146



2



Cultura geek de Hasitha Subhashana

Patrón de disyuntor (patrones de diseño para microservicios)

En un sistema distribuido no tenemos idea de cómo fallarían otros componentes. Problemas de red podría ocurrir, los componentes podrían fallar o un enrutador o un...

Lectura de 10 minutos · 12 de junio de 2021



618



5



Cultura geek de Shu Ishida

Instalación de Linux (Ubuntu 20.04) en un SSD portátil externo y dificultades ser consciente de

Lleve su sistema operativo Ubuntu a cualquier parte sin estar limitado por el hardware

Lectura de 11 minutos · 22 de abril de 2022



263



2



Ver todo de Víctor Almeida

Ver todo de Cultura Geek

Recomendado desde Medio



agosto severn

Dominar la estrategia de datos: una guía en vídeo de tres partes

Publicado anteriormente: <https://capitoldataanalytics.com/mastering-data-strategy/>

Lectura de 2 minutos · 4 de marzo de 2024





Codificación para subir de nivel Somnath Singh

La era de los empleos tecnológicos bien remunerados ha terminado

La muerte de los empleos tecnológicos.



· Lectura de 14 minutos · 1 de abril de 2024



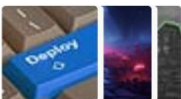
9.8K



253



Liza



Modelado predictivo con Python

20 historias · 1135 guardados



Guías prácticas para el aprendizaje automático

10 historias · 1365 guardados



Procesamiento natural del lenguaje

1411 historias · 908 guardados



ciencia de datos e inteligencia artificial

40 historias · 139 guardados

30/4/24, 15:33



Roland Nagy

Análisis de tráfico peatonal: análisis del comportamiento del cliente sin conexión y Método en línea

El análisis de datos basado en el tráfico peatonal puede ser clave para el éxito de cualquier negocio de física. Comprender el comportamiento del cliente puede ayudar a lograr...



· Lectura de 5 minutos · 8 de diciembre de 2023



3



Benoît Pimpaud

El analista de datos que todo CEO quiere

El analista de datos es probablemente el trabajo más subestimado en la industria de los datos.



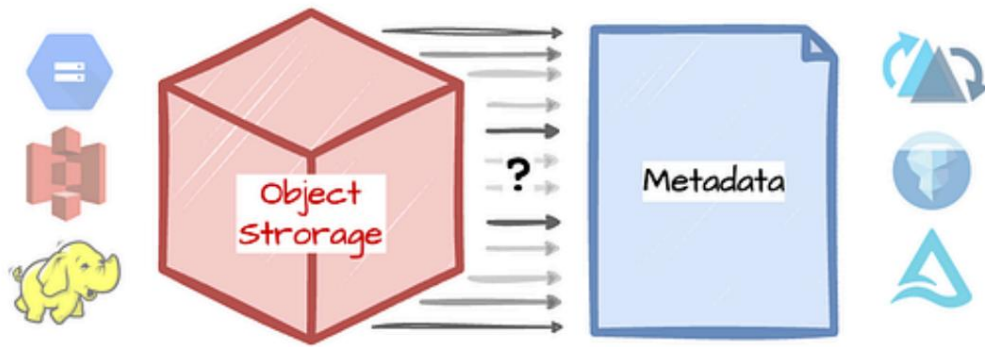
Lectura de 7 minutos · 2 de abril de 2024



77



2



Do we need the Lakehouse architecture?



Vu Trinh Ingeniero de datos Cosas

¿Necesitamos la arquitectura Lakehouse?

Cuando los lagos y almacenes de datos no son suficientes.

Lectura de 10 minutos · 20 de abril de 2024

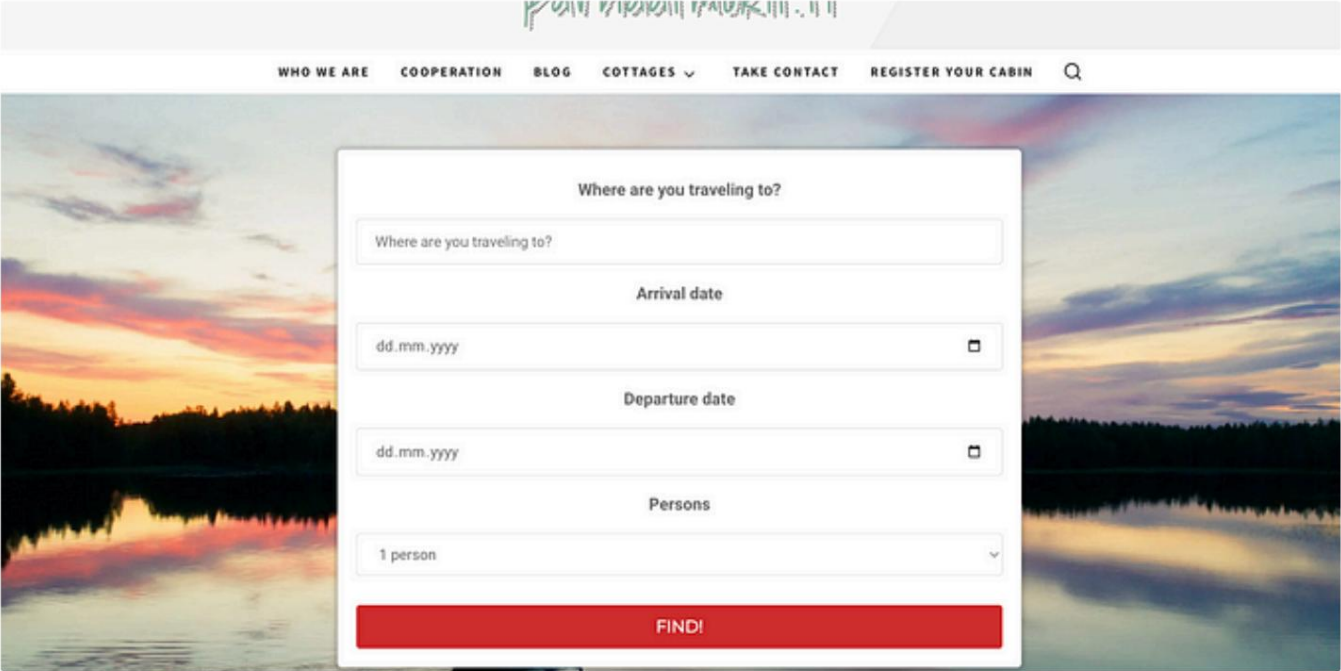


214



8







Artturi Jalli

Creé una aplicación en 6 horas que genera \$1500 al mes

¡Copia mi estrategia!



Lectura de 3 minutos · 23 de enero de 2024



17,5K



188



Ver más recomendaciones