

Introducción a la Ciencia de Datos (*Data Science*)



TECHIDEAS
ENGINEER YOUR DREAMS

© 2017 Engineered by TechIDEAS in Barcelona

Pablo Hernández
CEO & Co-Founder
@phermar



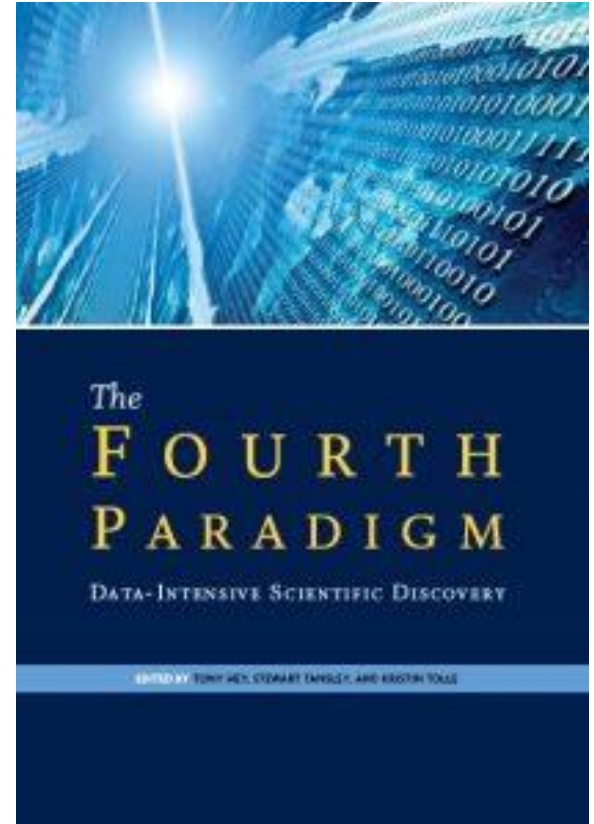
¿Qué es la Ciencia de Datos?

Definición

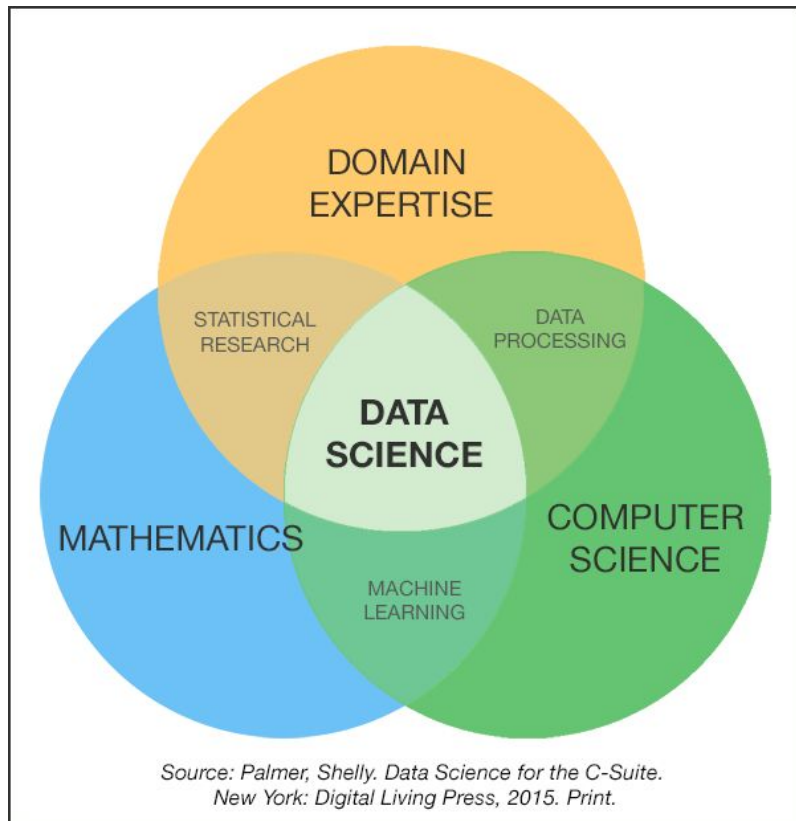
“La ciencia de datos es un campo interdisciplinario que involucra métodos científicos, procesos y sistemas para extraer conocimiento o un mejor entendimiento de datos en sus diferentes formas, ya sea estructurados o no estructurados”, Wikipedia

A new science paradigm

Jim Gray, ganador de un premio Turing, creía que la ciencia de datos es el cuarto paradigma para la exploración científica, que comenzó con la experimentación, continuó con la teorización y posteriormente con la computación. (“Living in The Fourth Paradigm”, January 15, 2010, Microsoft Research)



Data Science for the C-Suite - Shelly Palmer



El análisis de datos usando el método científico con el objetivo de transformar la información en acción.



¿Quién es un Científico de Datos?

MODERN DATA SCIENTIST

Data Scientist, the sexiest job of 21st century requires a mixture of multidisciplinary skills ranging from an intersection of mathematics, statistics, computer science, communication and business. Finding a data scientist is hard. Finding people who understand who a data scientist is, is equally hard. So here is a little cheat sheet on who the modern data scientist really is.

MATH & STATISTICS

- ☆ Machine learning
- ☆ Statistical modeling
- ☆ Experiment design
- ☆ Bayesian inference
- ☆ Supervised learning: decision trees, random forests, logistic regression
- ☆ Unsupervised learning: clustering, dimensionality reduction
- ☆ Optimization: gradient descent and variants

DOMAIN KNOWLEDGE & SOFT SKILLS

- ☆ Passionate about the business
- ☆ Curious about data
- ☆ Influence without authority
- ☆ Hacker mindset
- ☆ Problem solver
- ☆ Strategic, proactive, creative, innovative and collaborative



PROGRAMMING & DATABASE

- ☆ Computer science fundamentals
- ☆ Scripting language e.g. Python
- ☆ Statistical computing package e.g. R
- ☆ Databases SQL and NoSQL
- ☆ Relational algebra
- ☆ Parallel databases and parallel query processing
- ☆ MapReduce concepts
- ☆ Hadoop and Hive/Pig
- ☆ Custom reducers
- ☆ Experience with xaaS like AWS

COMMUNICATION & VISUALIZATION

- ☆ Able to engage with senior management
- ☆ Story telling skills
- ☆ Translate data-driven insights into decisions and actions
- ☆ Visual art design
- ☆ R packages like ggplot or lattice
- ☆ Knowledge of any of visualization tools e.g. Flare, D3.js, Tableau

Habilidades del científico de datos

Data science skill-set explained

By Krzysztof Zawadzki on August 30, 2014

Data scientists are in high demand. There is simply not enough talent to fill the jobs. Why? Because the sexiest job of 21st century requires a mixture of broad, multidisciplinary skills ranging from an intersection of mathematics, statistics, computer science, communication and business.

LinkedIn, October 2016



Employers need employees with cloud and distributed computing, statistical analysis and data mining skills to stay competitive.



¿Por qué ser un Científico de Datos?

Harvard Business Review, October 2012

DATA

Data Scientist: The Sexiest Job of the 21st Century

by Thomas H. Davenport and D.J. Patil

FROM THE OCTOBER 2012 ISSUE



The Economist, May 2017



Regulating the internet giants

The world's most valuable resource is no longer oil, but data

The data economy demands a new approach to antitrust rules



THE LARGEST COMPANIES BY MARKET CAP

The oil barons have been replaced by the whiz kids of Silicon Valley



Top 5 Publicly Traded Companies (by Market Cap)



Tech



Other





¿Cuántos Datos Generamos?

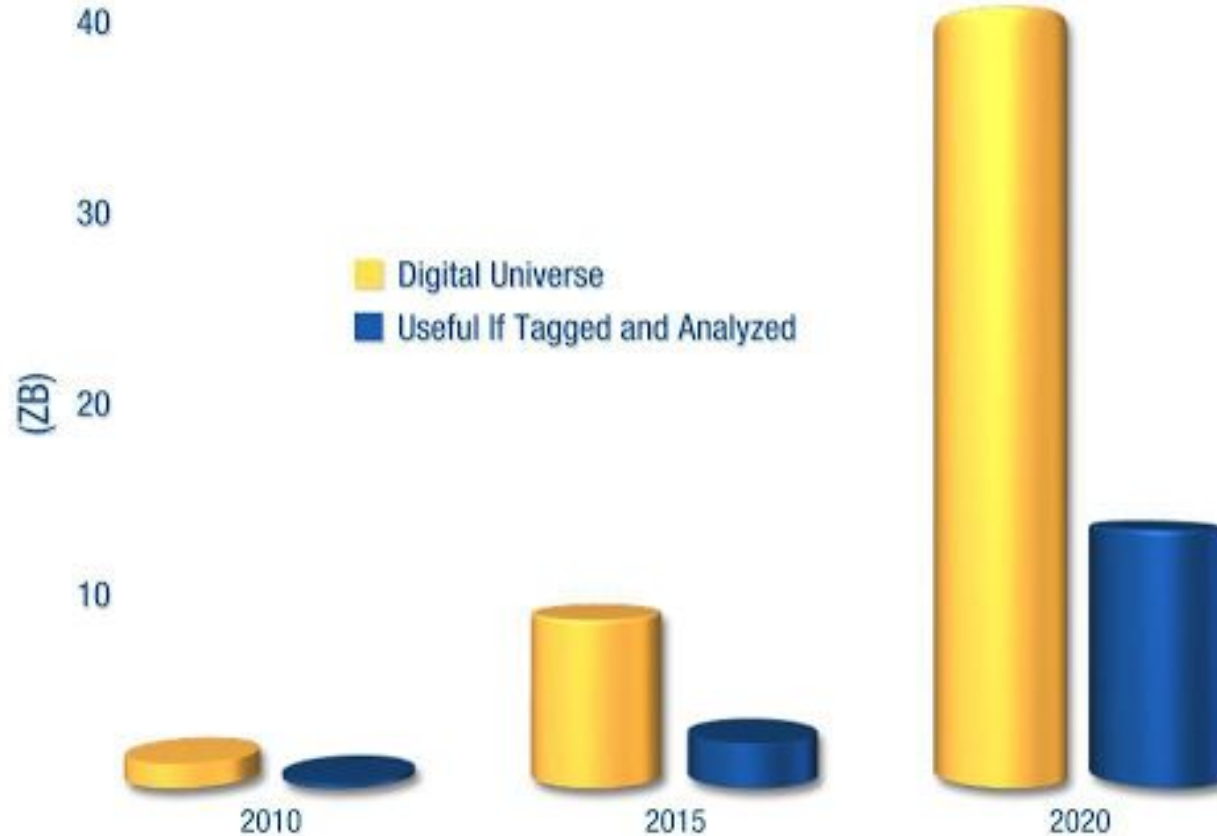
2017 This Is What Happens In An Internet Minute



Created By:
@LoriLewis
@OfficiallyChadd



Opportunity for Big Data



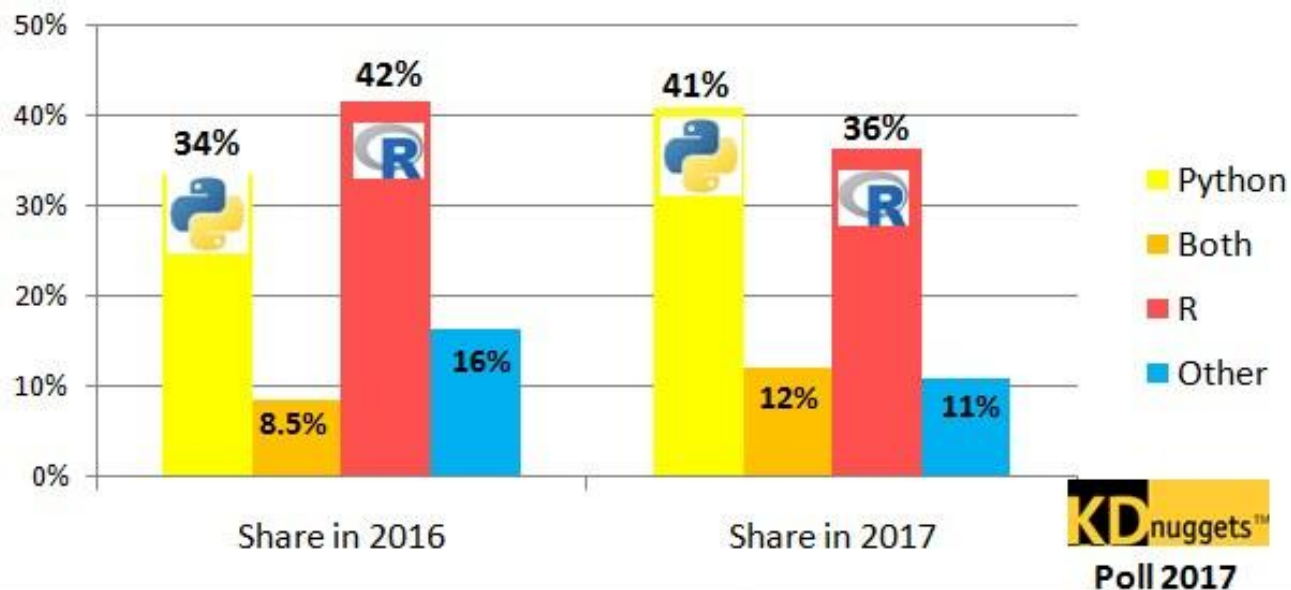
Source: IDC's Digital Universe Study, sponsored by EMC, December 2012



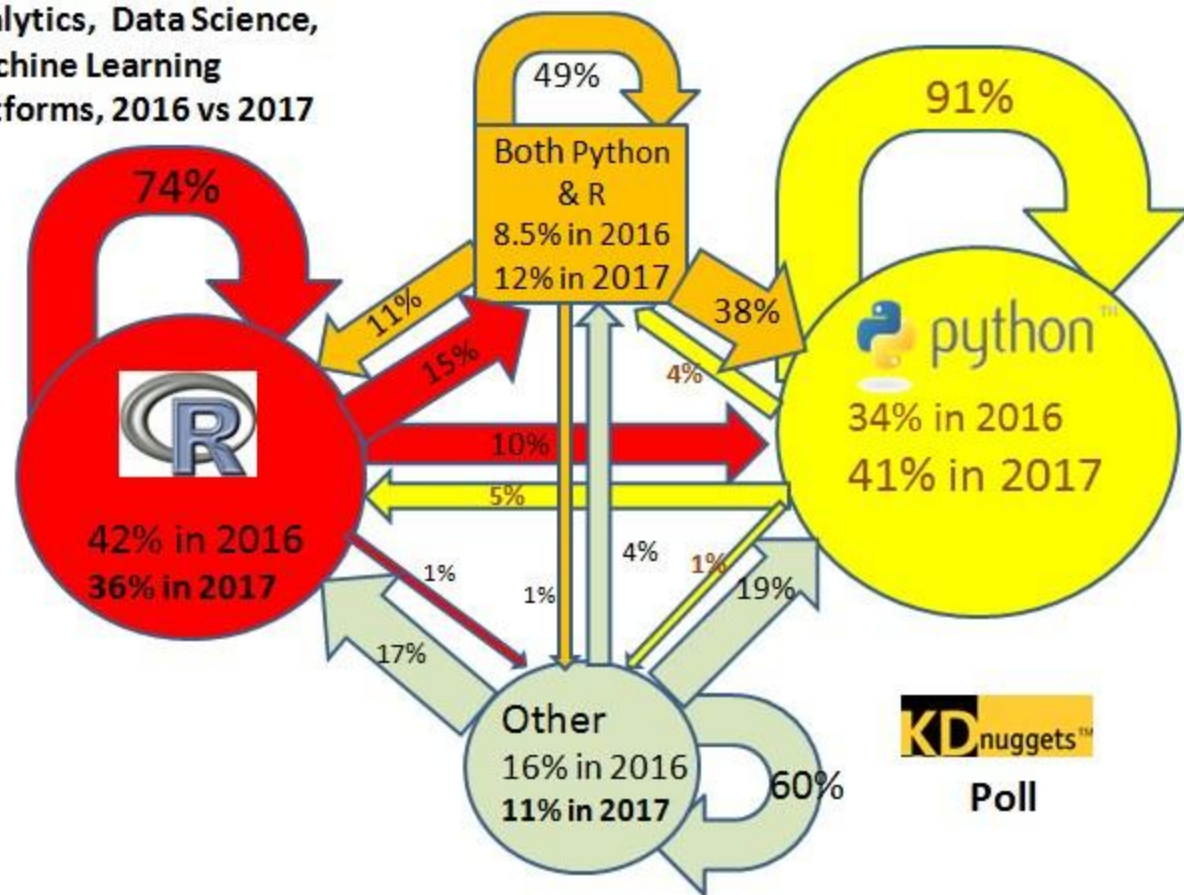
¿Qué Lenguaje de Programación?

Python o R

Python, R, Both, or Other platforms for Analytics, Data Science, Machine Learning



**Analytics, Data Science,
Machine Learning
Platforms, 2016 vs 2017**



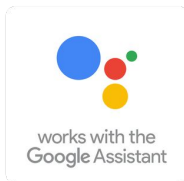


¿Cuáles son las aplicaciones de la
Ciencia de Datos?

Aplicaciones

- Asistentes virtuales por voz: Amazon, Google, Apple
- Predicciones del tráfico: Google (Waze), Uber
- Video-vigilancia
- Redes sociales: Facebook, Pinterest
- Filtrado de correo electrónico (Spam y Malware)
- Servicio de atención al cliente (Chatbot)
- Buscadores
- Recomendaciones de productos: Amazon
- Detección de pagos fraudulentos: PayPal

Asistentes Virtuales por Voz



Predicciones del Tráfico



Recomendadores

facebook Search Home

Find friends from different parts of your life
Use the tickboxes below to search for friends by hometown, school, employer and more.

Hometown
☐ Bhopalpur
☐ Delhi, India
☐ New Delhi, India
 Enter another city

Current location
☐ New Delhi, India
☐ Mumbai, Maharashtra
 Enter another location

Secondary school
☐ Marwari college Bhopalpur (SC/Mah)
☐ Anandam Dhandhaya Saraswati Vidya Mandir - Bhopalpur
☐ Up board
 Enter another secondary school

Mutual friend
☐ Chulbul Pradhan
☐ Ritika Joshi
☐ Shiv Bhaker
 Enter another name

College or university
☐ T.M.Bhopalpur University Bhopalpur, India
☐ University of Delhi
☐ Siddim Manipal University
 Enter another college

Facebook Friend Search Filter

Ritika Sharma
2 mutual friends
Add as Friend

Kostas Vardakas
Add as Friend

Syed Yaqoob Yaqoob
4 mutual friends
Add as Friend

Berlin Oda
Add as Friend

Abhineet Tiwari
1 mutual friend
Add as Friend

Jazz Malhotra
Add as Friend

Pankhuri Kumari
1 mutual friend
Add as Friend

Supra Sippu
Add as Friend

Yash Prakash
1 mutual friend
Add as Friend

Khan Zuban
2 mutual friends
Add as Friend

Sophie Lily
48 mutual friends
Add as Friend

Meenal Oberoi
2 mutual friends
Add as Friend

Sam Web Hosting
93 mutual friends
Add as Friend

NineNine Desk
73 mutual friends
Add as Friend

Neha Chaturvedi
24 mutual friends
Add as Friend

More

www.amazon.com/Lean-Startup-Entrepreneurs-Continuous-Innovation/dp/0307887898/ref=sr_1_1?ie=UTF8&qid=1455532634&sr=8-1&keywords=lean+startup

The Lean Startup: How Today's Entrepreneurs Use Continuous Innovation to Create Radically Successful Businesses Hardcover – September 13, 2011

by Eric Ries (Author)

★★★★★ 1,543 customer reviews

See all 4 formats and editions

Hardcover \$16.49	Paperback From \$13.23	Audiobook \$3.00
----------------------	---------------------------	---------------------

63 Used from \$9.92 13 Used from \$13.43 92 New from \$10.00 4 New from \$13.23

Free with your Audible trial

Most startups fail. But many of those failures are preventable. The Lean Startup is a new approach being adopted across the globe, changing the way companies are built and new products are launched.

Eric Ries delivers a startup as an organization dedicated to creating something new under conditions of extreme uncertainty. This is just as true for one person in a garage or a group of seasoned professionals in a Fortune 500 boardroom. What they have in common is a mission to penetrate that fog of uncertainty to discover a successful path to a sustainable business.

The Lean Startup approach fosters companies that are both more capital efficient and that leverage human creativity more effectively. Inspired by lessons from lean manufacturing, it relies on "validated learning" to discover a successful path to a sustainable business.

See the Best Books of the Month
Want to know our Editors' picks for the best books of the month? Browse Best Books of the Month, featuring our favorite new books in more than a dozen categories.

Frequently Bought Together

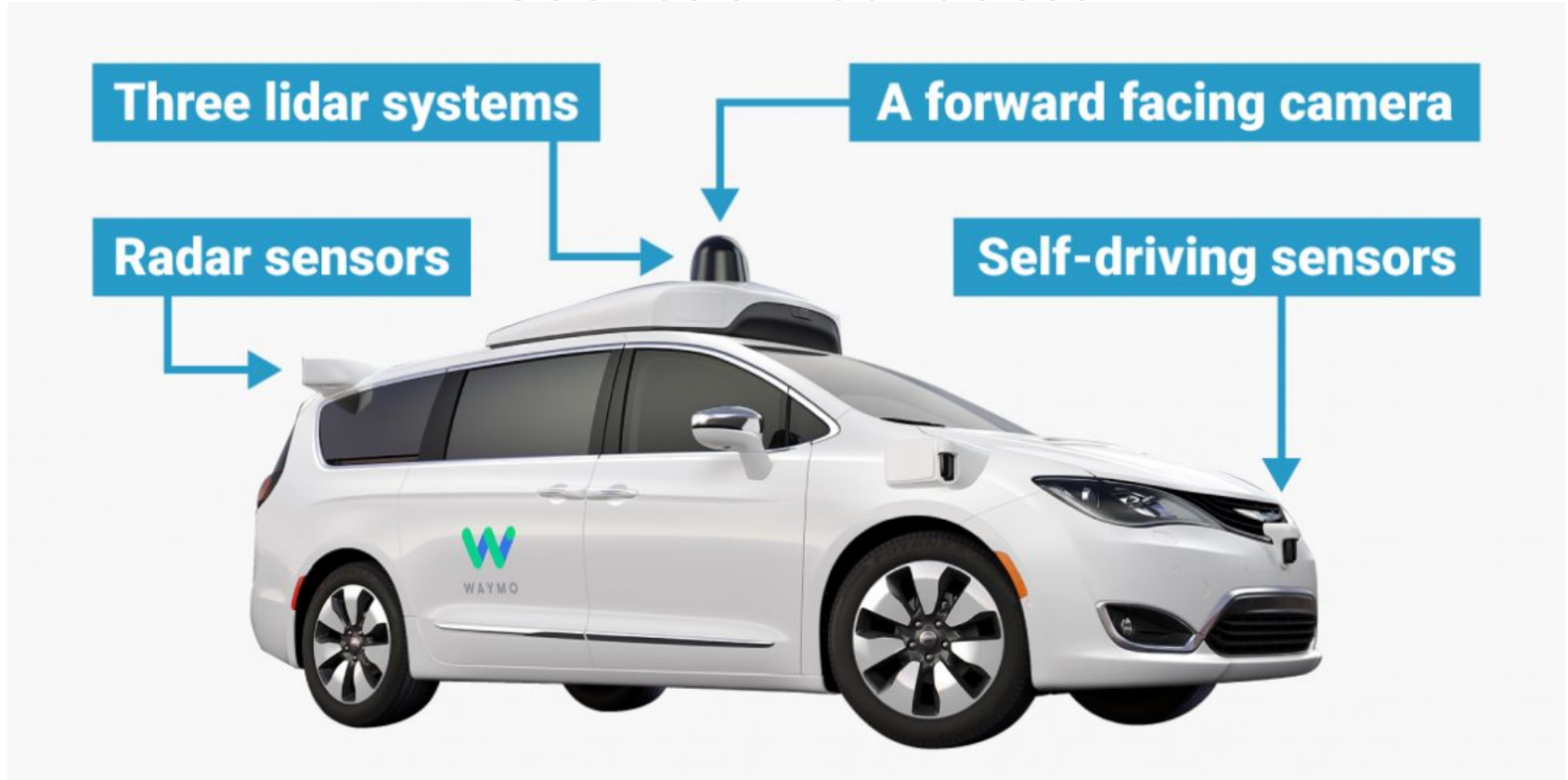
The Lean Startup + **Zero to One**
Total price: \$32.82
Add both to Cart
Add both to List

Customers Who Bought This Item Also Bought

Zero to One: Notes on Entrepreneurship by Peter Thiel
Business Model Generation: How the World's Most Innovative Companies Grow New Businesses by Alexander Osterwalder
The Startup Owner's Manual: How to Build a Successful Business from a Small Idea to a Scalable Venture by Reid Hoffman
How to Win Friends & Influence People by Dale Carnegie
The \$100 Startup: Discover How to Grow a Business from Your Home by Chris Pinney
The Innovator's Dilemma: The Revolutionary Disruption that Will Destroy Big Business by Clayton M. Christensen
Value Proposition Design: How to Create Competitive Advantage Using New Customer-Centric Models, Canvas, and Tools by Alexander Osterwalder
The Hard Thing About Hard Things: Building a Business That Lasts by Ben Horowitz
Traction: How Any Startup Can Increase Visibility, Understand Customers, and Secure Funding by Steve Kravis
The 4-Hour Workweek: Escape 9-5, Live Anywhere, and Join the New Rich by Timothy Ferriss



Coches sin conductor





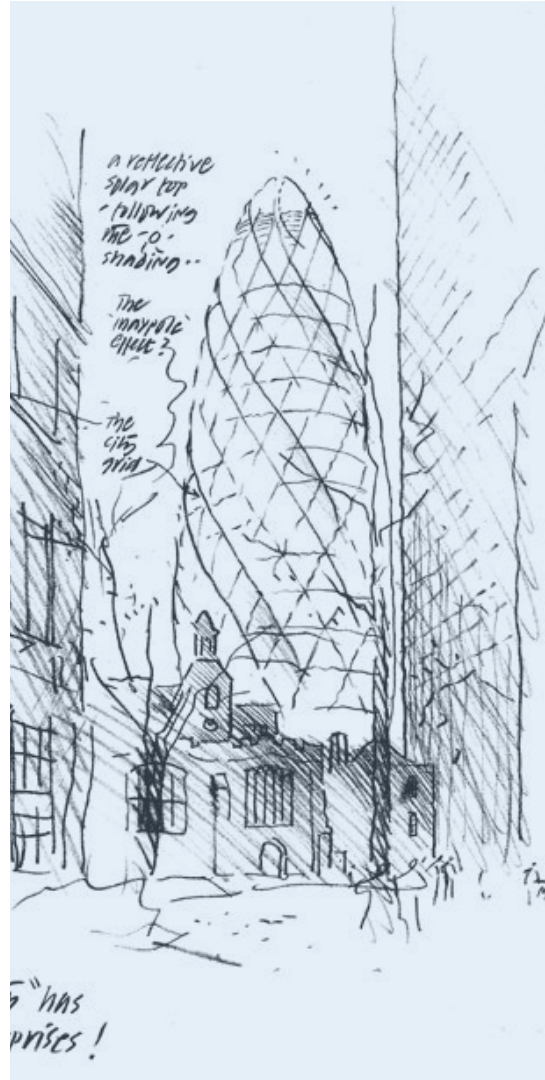
TechIDEAS's Profile

Section SubTitle

About us

“We transform IDEAS into compelling TECHNOLOGY”

TechIDEAS is a Software Engineering company specialized in disruptive innovation



MIT Collaboration



Massachusetts
Institute of
Technology

La pyme catalana TechIDEAS colaborará con el MIT de Boston

La firma tecnológica participará en los programas de formación

El prestigioso instituto de tecnología de Boston (MIT) alcanzará por un acuerdo con la firma TechIDEAS. La pyme barcelonesa participará en los programas de formación que la academia estadounidense ofrece a sus investigadores.

PERCECIBIO M. SANDEU

BARCELONA. — Desde Boston al País Vasco. La empresa de software TechIDEAS, ubicada en el centro empresarial de Barcelona Activa, firmó ayer un acuerdo con el Massachusetts Institute of Technology (MIT) mediante el cual esta pyme barcelonesa podrá formar parte de un selecto lista de empresas que colaboran con el programa educativo y profesional del instituto estadounidense líder del mundo.

Los ciudadanos e investigadores del MIT de Boston tendrán ahora la posibilidad de perfeccionar sus conocimientos y trabajar en proyectos avanzados en el sector de la tecnología en Barcelona durante una estancia TechIDEAS. Es la primera pyme catalana que se involucra en este programa de internacionalización del instituto norteamericano.

Así, a partir del próximo mes de junio, llegará a Barcelona el primer colabante estadounidense para participar en un proyecto de software aplicado al sector del transporte.



Pablo Hernández, consejero delegado y fundador de TechIDEAS

En el futuro la intención es ampliar este acuerdo de colaboración bilateral para que investigadores catalanes tengan la posibilidad de llevar a cabo sus proyectos en las instalaciones del MIT en Boston.

"Se trata de un reconocimiento muy importante", declara Pablo Hernández, consejero delegado y fundador de la firma. "Es la primera vez que en España se puede hacer investigación de calidad", agrega. El MIT ya tiene acuerdos similares con firmas como el IITPA o la construcción Acciona.

La elección de España no es casual. Según el instituto norteamericano, la Península "está avanzando

rápidamente y es un jugador dinámico en el paisaje económico europeo". "Los estadounidenses ven que la economía funciona, pero que el nivel de investigación en España es bajo. Se dan cuenta de que algo está pasando, pero al mismo tiempo saben que hay mucho potencial creativo", explica Hernández.

TechIDEAS nació en el 2000, justo después del colapso de la burbuja de Internet y facturó el año pasado 600.000 euros. La empresa destaca en torno al 30% de sus ingresos a proyectos de I+D en el desarrollo de software avanzado para la movilidad y los sistemas informáticos de código abierto.



Featured Apps



The European Paradox

“ **Why so few European SMEs grow to become global corporations?** ”



“One important explanation is that technologies and services (developed by small and highly innovative European firms) often can only grow by attracting US investments. It is flattering that the US investors recognize the quality and attractiveness of our research, but it is a signal that Europeans are failing to risk money in getting gains from our brains. This is really a tragic loss of assets: not only do we fail to exploit our potential, but we pump our knowledge capital offshore often accompanied by the innovators, our most innovative and entrepreneurial assets of all.”

Viviane Reding Speech at Biennial Conference 2008, Athens, 8 September 2008.

Thank you



TECHIDEAS
ENGINEER YOUR DREAMS

© 2017 Engineered by TechIDEAS in Barcelona

Pablo Hernández
CEO & Co-Founder
@phermar