

Científicos de datos (Data Scientist)

Dr. Gustavo Illescas



Dr. Arturo Mora-Soto



Dr. Guillermo Rodríguez



Unidad 1

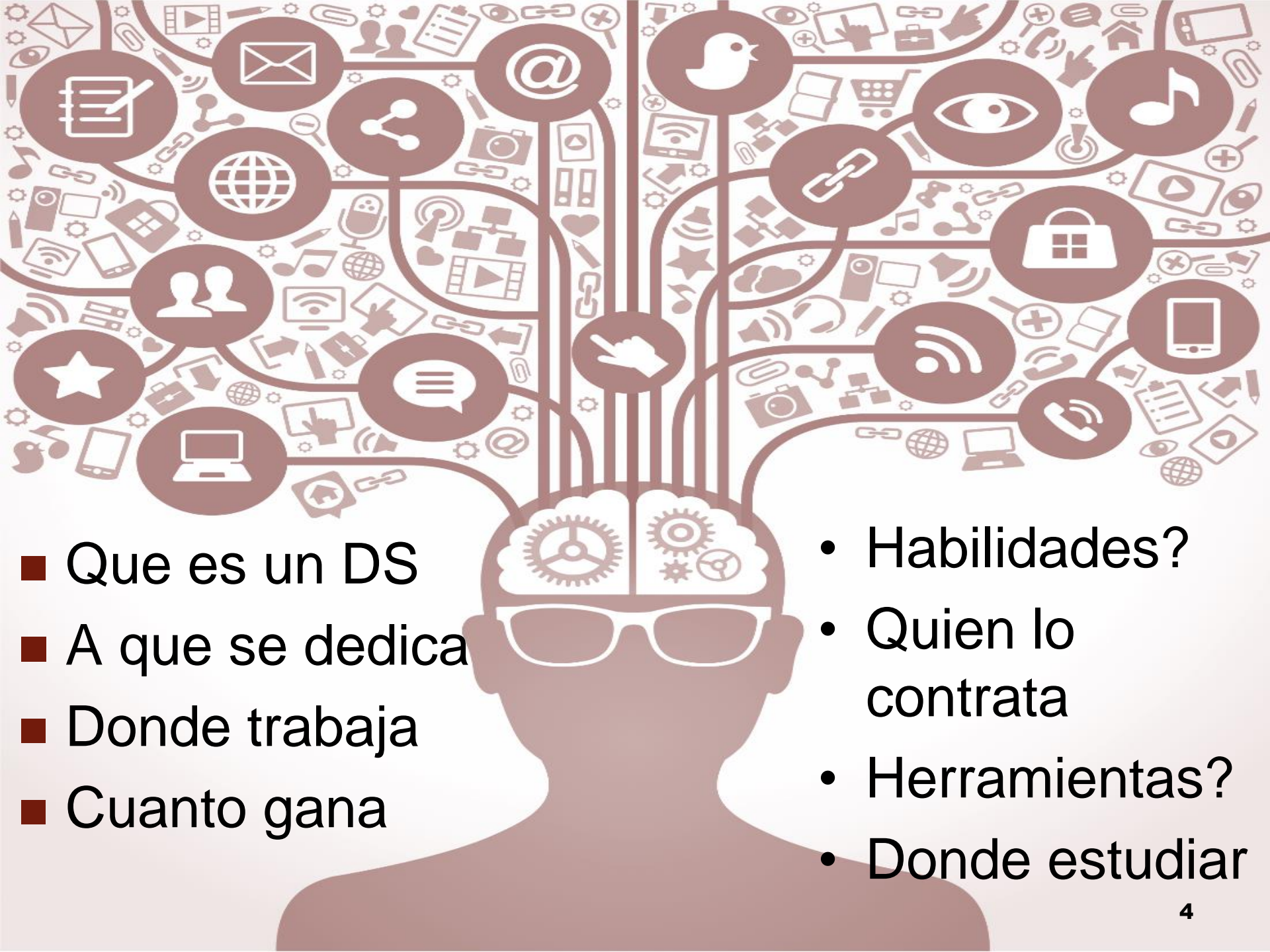
El científico de datos (Data Scientist-DS).
Funciones, habilidades, como pensar como un DS.
El DS en el contexto de la gestión del
conocimiento, la inteligencia de negocios (Business
Intelligence) y el Big Data.



DS desde la informática de gestión



La informática de gestión, también conocida como informática del negocio, es una disciplina de estudio que contempla la aplicación de la informática en el área gerencial. Como tal, presenta una confluencia interdisciplinaria entre disciplinas científicas, gestión empresarial e ingeniería.



- Que es un DS
- A que se dedica
- Donde trabaja
- Cuanto gana

- Habilidades?
- Quien lo contrata
- Herramientas?
- Donde estudiar

HBR.ORG

Harvard Business Review



OCTOBER 2012

Oct 2012

46 The Big Idea

The True Measures
Of Success

Michael J. Mauboussin

84 International Business

10 Rules for Managing
Global Innovation

Keeley Wilson and Yves L. Doz

93 Leadership

What Ever Happened
To Accountability?

Thomas E. Ricks

GETTING
CONTROL
OF

BIG DATA

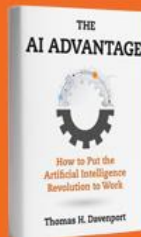


How vast new streams of
information are changing
the art of management
PAGE 59



Speaker. Author. Advisor.

Tom provokes, stimulates, and equips audiences with the



The AI Advantage How to Put the Artificial Intelligence Revolution to Work

"Once again, Tom Davenport has taken on the role of the business sherpa, guiding decision makers through the maze of technology to the business transformation we all want." "A must-read for executives who wish to truly become cognitive corporation leaders."

[Order Book](#)[Read More](#)

Speaking

Tom is an experienced speaker who has delivered keynotes and event speeches.

[View Topics ▸](#)

Writing

Tom has written more than fifteen books and over 250 articles for prominent publications.

[Read Articles ▸](#)

Research & Consulting

Tom is an experienced consultant who has worked with many of the world's best-known corporations.

[Contact ▸](#)



Data Scientist:

The Sexiest Job of the 21st Century

**Meet the people who
can coax treasure out of
messy, unstructured data.**

*by Thomas H. Davenport
and D.J. Patil*

When Jonathan Goldman arrived for work in June 2006 at LinkedIn, the business networking site, the place still felt like a start-up. The company had just under 8 million accounts, and the number was growing quickly as existing members invited their friends and colleagues to join. But users weren't seeking out connections with the people who were already on the site at the rate executives had expected. Something was apparently missing in the social experience. As one LinkedIn manager put it, "It was like arriving at a conference reception and realizing you don't know anyone. So you just stand in the corner sipping your drink—and you probably leave early."

70 Harvard Business Review October 2012

Caso Goldman: 2006, LinkedIn

Meet the people who
can coax treasure out of
messy, unstructured data.
by Thomas H. Davenport
and D.J. Patil

When Jonathan Goldman arrived for work in June 2005 at LinkedIn, the business networking site, the place still felt like a start-up. The company had just under 8 million accounts, and the number was growing quickly as existing members invited their friends and colleagues to join. But users weren't making real connections with the people who were already on the site at the rate executives had expected. Something was apparently missing in the social experience. As one LinkedIn manager put it, "It was like arriving at a conference reception and realizing you don't know anyone. So you just stand in the corner sipping your drink—and you probably leave early."

39 Harvard Business Review October 2009

- Dr en física
- estaba **intrigado** por la vinculación y por la riqueza de los perfiles de usuario
- personas que aún no se habían relacionados pero parecía que iban a ser “conocidas”
- no creían demasiado en lo que estaba haciendo
- alto grado de autonomía: comenzó con las pruebas apoyado por el CEO
- un buen ejemplo de un nuevo jugador clave en organizaciones: el "científico de datos"

¿Quiénes son esas personas?

Meet the people who
can coax treasure out of
messy, unstructured data
by Thomas H. Davenport
and D.J. Patil

When Jonathan Goldman arrived for work in June 2001 at LinkedIn, the business networking site, the place still felt like a start-up. The company had just under 8 million accounts, and the number was growing quickly as existing members invited their friends and colleagues to join. But users weren't making real connections with the people who were already on the site at the rate executives had expected. Something was apparently missing in the social experience. As one LinkedIn manager put it, "It was like arriving at a conference reception and realizing you don't know anyone. So you just stand in the corner sipping your drink—and you probably leave early."

36 Harvard Business Review October 2009

Dificultades

- cómo identificar el talento, atraerlo a una empresa y hacerlo productivo
- **comienzan a existir** programas universitarios
- hay poco consenso sobre el papel que cumple en una organización
- cómo pueden añadir más valor
- cómo se debe medir su rendimiento



¿Quiénes son esas personas?

- entender lo que hacen en las empresas
- ¿qué habilidades necesitan?
- ¿en qué campos se encuentran más fácilmente esas habilidades?

¿Quiénes son esas personas?

Una vez que tienen “datos” necesitan personas que puedan administrarlos y encontrar ideas en ellos.

Los DS son capaces de:

- “estructurar” grandes cantidades de datos sin forma y hacer posible el análisis.
- Identifican datos enriquecidos
- Unir fuentes potencialmente incompletas
- los datos nunca dejan de fluir, los DS cambian el análisis ad hoc: una conversación continua con datos.
- Enfrentan limitaciones técnicas pero no permiten que eso empañe su trabajo y busca soluciones novedosas.
- Mientras hacen descubrimientos, comunican lo que han aprendido y sugieren sus implicaciones para nuevas direcciones comerciales.
- Son creativos para mostrar información visualmente.
- Asesoran a ejecutivos y gerentes de producto sobre las implicaciones de los datos para productos, procesos y toma de decisiones.
- A menudo recae en los DS diseñar sus propias herramientas.

¿Qué clase de persona hace todo esto?

Meet the people who
can coax treasure out of
messy, unstructured data.
by Thomas H. Davenport
and D.J. Patil

When Jonathan Goldman arrived for work in June 2007 at LinkedIn, the business networking site, the place still felt like a start-up. The company had just under 8 million accounts, and the number was growing quickly as existing members invited their friends and colleagues to join. But users weren't working and communicating with the people who were already on the site at the rate executives had expected. Something was apparently missing in the social experience. As one LinkedIn manager put it, "It was like arriving at a conference reception and realizing you don't know anyone. So you just stand in the corner sipping your drink—and you probably leave early."

39 Harvard Business Review October 2009

- Pensar en él o ella como un híbrido de hacker, analista, comunicador, y asesor de confianza. La combinación es muy potente y rara.
- La habilidad más básica: escribir código.
- Necesidad de los DS para comunicarse

El rasgo dominante es:

- una curiosidad, un intenso deseo de ir por debajo de la superficie de un problema,
- encontrar las preguntas en su corazón
- y extraer un conjunto muy claro de las hipótesis que se pueden probar.

Data Scientist:

The Sexiest Job of the 21st Century

*Meet the people who
can coax treasure out of
messy, unstructured data.*
by Thomas H. Davenport
and D.J. Patil

When Jonathan Goldman arrived for work in June 2006 at LinkedIn, the business networking site, the place still didn't like a start-up. The company had just under 8 million accounts, and the number was growing quickly as existing members invited their friends and colleagues to join. But users weren't seeking out connections with the people who were already on the site as the site executives had expected. Something was apparently missing in the social experience. As one LinkedIn manager put it, "It was like arriving at a conference reception and realizing you don't know anyone. So you just stand in the corner sipping your drink—and you probably leave early."

78 Harvard Business Review October 2009




DataCamp
Learn R & Data Analysis
By Doing

THE DATA SCIENCE INDUSTRY WHO DOES WHAT

A LOOK AT THE KEY ROLES IN DATA SCIENCE

The vast range of careers in the data science industry is accompanied by an avalanche of job postings. DataCamp took a closer look at recent data science job postings to learn more about the differences between various roles in the industry and to unravel the meaning behind many cool-sounding job titles.



Roles en la industria

¿Quien hace que?

- El mercado de trabajo industria de la ciencia de datos es actual
- Visión detallada de todas las **ofertas de trabajo**, su descripción y sus requisitos de formación.
- Estos anuncios ayudan a responder a preguntas como
 - ¿Qué tan similares o diferentes son ciertas funciones?
 - ¿Qué conjunto de tecnología necesito dominar?
 - ¿Cuál es el modo de pensar que necesito al aceptar este trabajo?

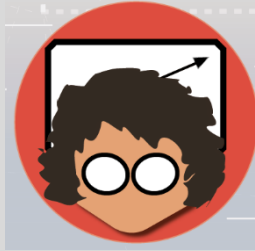
DATA ANALYST

'DATA DETECTIVE'



BUSINESS ANALYST

'CHANGE AGENT'



DATABASE ADMINISTRATOR

'DATABASE CARETAKER'



DATA ENGINEER

'SOFTWARE ENGINEERS BY TRADE'



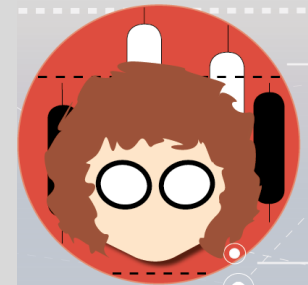
DATA ARCHITECT

'THE CONTEMPORARY DATA MODELLER'



DATA SCIENTIST

'AS RARE AS UNICORNS'



STATISTICIAN

'HISTORIC LEADERS OF DATA'

DATA AND ANALYTICS MANAGER

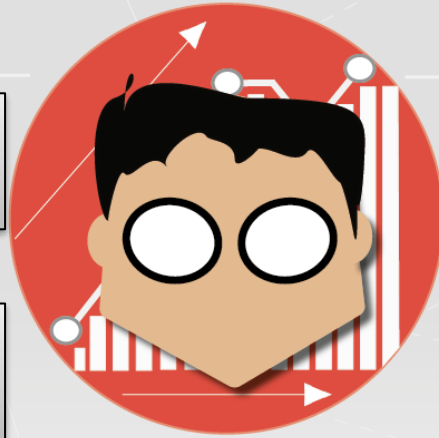
'DATA SCIENCE TEAM LEADER'

DATA ANALYST

'DATA DETECTIVE'

Role

Mindset



Languages

Skills & Talents

HIRED BY

DATA ENGINEER

SOFTWARE ENGINEERS BY TRADE

Role

Mindset



HIRED BY

Languages

Skills & Talents

Languages

Skills & Talents

DATA SCIENTIST
"AS RARE AS UNICORNS"

Role

Mindset

HIRED BY



DATA ARCHITECT

'THE CONTEMPORARY DATA MODELLER'

Languages

SQL, XML, Hive, Pig, Spark

Skills & Talents

- ✓ Data warehousing solutions
- ✓ In-depth knowledge of database architecture
- ✓ Extraction Transformation and Load (ETL), spreadsheet and BI tools
- ✓ Data modeling
- ✓ Systems development



Role:

Creates blueprints for data management systems to integrate, centralize, protect and maintain data sources

Mindset:

Inquiring ninja with a love for data architecture design patterns

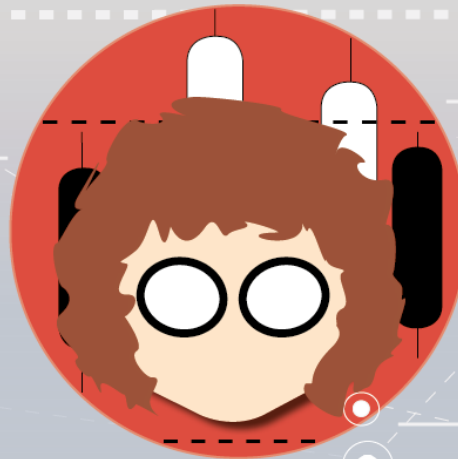
HIRED BY
VISA *Coca-Cola* logitech

Languages

R, SAS, SPSS, Matlab, Stata, Python,
Perl, Hive, Pig, Spark, SQL

Skills & Talents

- ✓ Statistical theories & methodology
- ✓ Data mining & machine learning
- ✓ Distributed Computing (Hadoop)
- ✓ Database systems (SQL and NO SQL based)
- ✓ Cloud tools



STATISTICIAN

"HISTORIC LEADERS OF DATA"

Role

*Collects, analyzes and interprets-
qualitative as well as quantitative
data with statistical theories and
methods*

Mindset

*Logical and enthusiastic stats
genius*

HIRED BY

Linked in *Johnson & Johnson*  **PEPSICO**

DATABASE ADMINISTRATOR

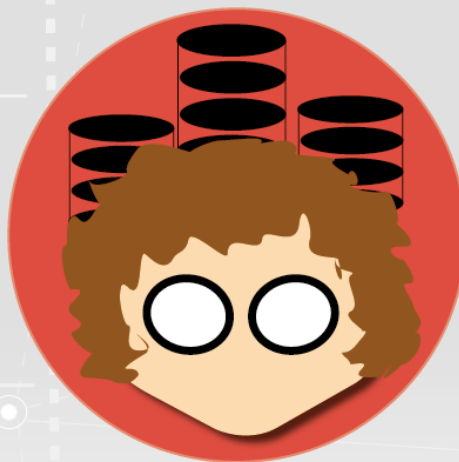
'DATABASE CARETAKER'

Role

Ensures that the database is available to all relevant users, is performing properly and is being kept safe

Mindset

Master of Disaster Prevention



HIRED BY



Languages

SQL, Java, Ruby on Rails, XML, C#, Python

Skills & Talents

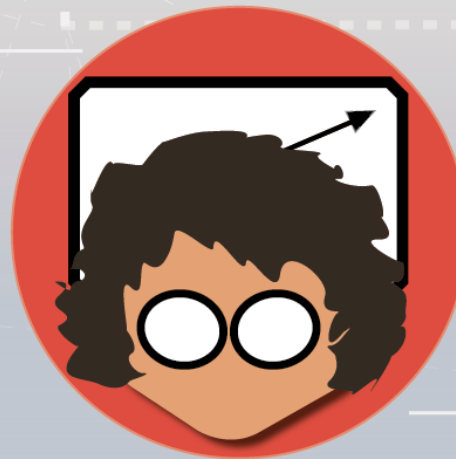
- ✓ Backup & recovery
- ✓ Data modeling and design
- ✓ Distributed Computing (Hadoop)
- ✓ Database systems (SQL and NO SQL based)
- ✓ Data security
- ✓ ERP & business knowledge

Languages

SQL

Skills & Talents

- ✓ Basic tools (e.g. MS Office)
- ✓ Data visualization tools (e.g. Tableau)
- ✓ Conscious listening and storytelling
- ✓ Business Intelligence understanding
- ✓ Data modeling



BUSINESS ANALYST

'CHANGE AGENT'

Role

Improves business processes as intermediary between business and IT

Mindset

Resilient project juggler

HIRED BY



DATA AND ANALYTICS MANAGER

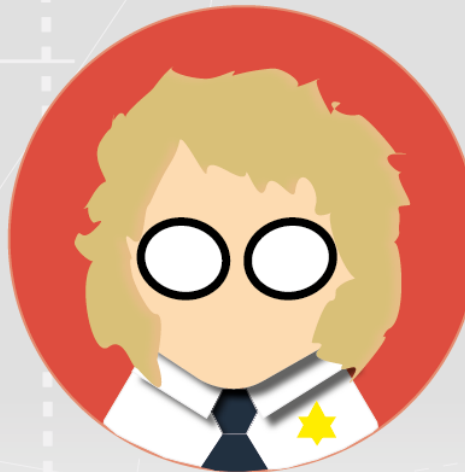
'DATA SCIENCE TEAM LEADER'

Role

Manages a team of analysts and data scientists

Mindset

Data Wizards' Cheerleader



Languages

SQL, R, SAS, Python, Matlab, Java

Skills & Talents

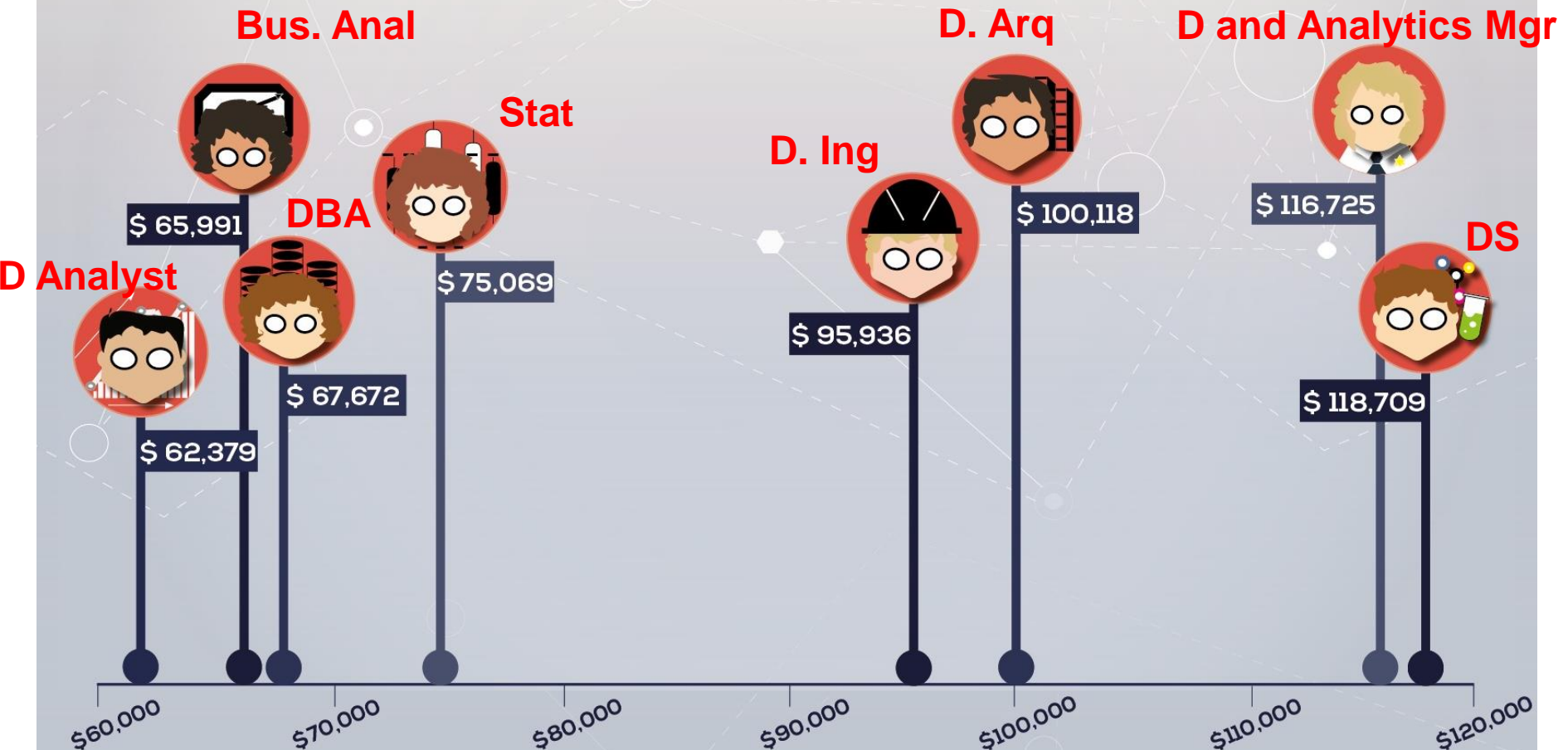
- ✓ *Database systems (SQL and NO SQL based)*
- ✓ *Leadership & project management*
- ✓ *Interpersonal communication*
- ✓ *Data mining & predictive modeling*

HIRED BY

coursera  **slack**  **MOTOROLA SOLUTIONS**

WHO EARNS WHAT SALARY

NATIONAL AVERAGE SALARIES (US)



Source: glassdoor.com

https://www.glassdoor.com.ar/Sueldos/buenos-aires-analista-de-sistemas-sueldo-SRCH_IL.0,12_IM963_KO13,33.htm?clickSource=searchBtn

glassdoor

analista de sistemas

Sueldos

Buenos Aires

Buscar

Empleos Empresas Sueldos Entrevistas

Para empleadores

Sueldos para Analista De Sistemas en Área Buenos Aires

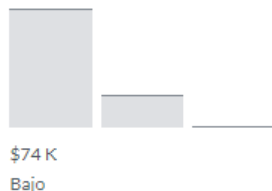
Actualizado el 10 de may de 2022

Confianza muy alta

\$ 130.638 /mes

Sueldo base promedio

675 sueldos



Remuneración adicional en efectivo ⓘ

Promedio: **\$ 148.291** Intervalo: \$ 2.379 - \$ 679.590

¿Cuánto gana un Analista De Sistemas en Buenos Aires?

El sueldo promedio de Analista De Sistemas es \$ 130.638 por mes en Área B para un Analista De Sistemas en Área Buenos Aires es de \$ 148.291, con un sueldo base basan en 675 sueldos enviados anónimamente a Glassdoor por empleado

Sueldos en Área Buenos Aires

Empresa

Sueldo base promedio en (ARS)



Accenture

Analista De Sistemas - Por mes

\$ 102.712 /mes

4,1 ★

27 sueldos [Ver 30 sueldos de todas las ubicaciones](#)



Telecom Argentina

Analista De Sistemas - Por mes

\$ 146.617 /mes

3,8 ★

18 sueldos [Ver 21 sueldos de todas las ubicaciones](#)



Telefónica

Analista De Sistemas - Por mes

\$ 146.654 /mes

4,0 ★

15 sueldos [Ver 19 sueldos de todas las ubicaciones](#)



Mercado Libre

Analista De Sistemas - Por mes

\$ 128.466 /mes

4,5 ★

8 sueldos [Ver 10 sueldos de todas las ubicaciones](#)



ExxonMobil

Analista De Sistemas - Por mes

\$ 197.506 /mes

3,1 ★

7 sueldos [Ver 7 sueldos de todas las ubicaciones](#)

https://www.glassdoor.com.ar

glassdoor

🔍 Data Scientist

Sueldos

Buenos Aires

Buscar



Empleos Empresas **Sueldos** Entrevistas

Para empleadores

Sueldos para Data Scientist en Área Buenos Aires

Actualizado el 10 de may de 2022

Confianza muy alta

\$ 200.000 /mes

Sueldo base promedio

208 sueldos



Remuneración adicional en efectivo ⓘ

Promedio: **\$ 199.642** Intervalo: \$ 2 - \$ 639.321

¿Cuánto gana un Data Scientist en Buenos Aires?

El sueldo promedio de Data Scientist es \$ 200.000 por mes en Área Buenos Aires. La rem Data Scientist en Área Buenos Aires es de \$ 199.642, con un rango de entre \$ 2 y \$ 639.321. 208 sueldos enviados anónimamente a Glassdoor por empleados con el cargo de Data Sc

Sueldos en Área Buenos Aires



Telefónica

Data Scientist - Por mes

\$ 224.299 /mes

4,0 ★

15 sueldos [Ver 17 sueldos de todas las ubicaciones](#)



Gobierno de la Ciudad de Buenos Aires

Data Scientist - Por mes

\$ 122.500 /mes

3,4 ★

10 sueldos [Ver 15 sueldos de todas las ubicaciones](#)



Mercado Libre

Data Scientist - Por mes

\$ 212.825 /mes

4,5 ★

10 sueldos [Ver 16 sueldos de todas las ubicaciones](#)



Equifax

Data Scientist - Por mes

\$ 186.128 /mes

3,9 ★

9 sueldos [Ver 10 sueldos de todas las ubicaciones](#)



Accenture

Data Scientist - Por mes

\$ 140.785 /mes

4,1 ★

6 sueldos [Ver 8 sueldos de todas las ubicaciones](#)



Banco Galicia

Data Scientist - Por mes

\$ 268.878 /mes



Nº 1 SAP



4,4



[Ver evaluaciones](#) | [Ver empleos](#)

“ ¡Excelente ambiente laboral, grandes posibilidades de crecimiento y de capacitaciones!

Nº 2 Tenaris



4,3



[Ver evaluaciones](#) | [Ver empleos](#)

“ Estabilidad laboral, sueldo promedio, beneficios gym y transporte

Nº 3 Mercado Libre



4,3



[Ver evaluaciones](#) | [Ver empleos](#)

“ Excelentes desafíos, alto nivel de competencia, buen ambiente laboral.

Nº 4 Accenture



4,2



[Ver evaluaciones](#) | [Ver empleos](#)

“ Ambiente laboral amigable, excelente edificio y beneficios, realmente se preocupan, son muy flexibles.

Nº 5 BBVA



4,2



[Ver evaluaciones](#) | [Ver empleos](#)

“ Muchas capacitaciones, beneficios en productos bancarios, ajuste mensual por inflación.

Nº 6 Telefónica



4,2



[Ver evaluaciones](#) | [Ver empleos](#)

“ Home office, descuentos en seguros, maternidad extendida, rotaciones internacionales, beneficios como empleado.

Nº 7 ypf



4,1



[Ver evaluaciones](#) | [Ver empleos](#)

“ Sueldo, prestaciones, horarios, cobertura salud, premios.

Nº 8 Banco Galicia



4,1



[Ver evaluaciones](#) | [Ver empleos](#)

“ Oportunidades de crecimiento, flexibilidad, capacitación constante, remuneración, beneficios en productos del Banco.

Nº 9 Despegar.com



4,1



[Ver evaluaciones](#) | [Ver empleos](#)

“ El ambiente laboral es muy ameno, se tienen muchos beneficios, los sueldos son bastante competitivos y los servicios por los que se...

Nº 10 Banco de la Nación



4,1



[Ver evaluaciones](#) | [Ver empleos](#)

“ Te sentís seguro, hay oportunidades de crecimiento, es un constante aprendizaje.



Principales empleos que te recomendamos



Data Analytics

Santander Argentina · Provincia de Buenos Aires, Argentina



1 antiguo empleado



Solicitud sencilla



Data Scientist

Galicia Seguros · Buenos Aires, Buenos Aires Province, Argentina



Jóvenes Profesionales

Axionlog · Provincia de Buenos Aires, Argentina



En busca de personal



Data Scientist

JPMorgan Chase & Co. · Buenos Aires, Buenos Aires Province, Argentina



1 contacto



Data Scientist

Walmart · Buenos Aires y alrededores



En busca de personal



Solicitud sencilla



Supervisora / Supervisor Regional de Nuevos Canales y Estrategia Point - Mercado Pago - NUEVO

Mercado Libre · Buenos Aires y alrededores



2 contactos



Analista Senior de Datos para Fintech

Mercado Libre · Buenos Aires y alrededores



2 contactos



Data Scientist

Ibm · Buenos Aires, Buenos Aires Province, Argentina



Sé de los primeros 5 solicitantes



Data Engineer

Santander Tecnología Argentina · Buenos Aires, Provincia de Buenos Aires, Argentina



2 contactos



Solicitud sencilla



Data Scientist - Applied AI

JPMorgan Chase & Co. · Buenos Aires, Buenos Aires Province, Argentina



Data Scientist

Progres S.A. · Buenos Aires, Buenos Aires Province, Argentina



En busca de personal



Growth Data Analyst

Santander Argentina · Buenos Aires, Provincia de Buenos Aires, Argentina



En busca de personal



Solicitud sencilla

Figure 1. Magic Quadrant for Data Science and Machine Learning Platforms



<https://www.dominodatalab.com/blog/domino-named-a-visionary-in-gartner-magic-quadrant-for-completeness-of-vision-and-ability-to-execute/>

[Free Courses Launched] Linear Programming for Data Science Professionals | Naive Bayes from Scratch

🔍 Pendientes x



➡ **Analytics Vidhya** <bhanu@analyticsvidhya.com> [Cancelar suscripción](#)
para mí ▾

28 mar. 2020 5:53



🌐 inglés ▾ > español ▾ [Traducir mensaje](#)

[Desactivar para: inglés](#) x

Two Brand New Free Courses | [Linear Programming](#) | [Naive Bayes from Scratch](#)

Hey there

The weekend is here and to make your homestay even more enriching, we are excited to bring you the two free data science courses of the week! The learning must continue, and we at Analytics Vidhya firmly believe in giving back to the community.

Below are the details of both the courses

[Linear Programming for Data](#)

[Scientists](#)

What'll you learn in this course:

- What is Linear Programming?
- Solve Linear Program using different platforms
- Methods of Solving Linear Programming
- Applications of Linear Programming

[Enroll For Free](#)

[Naive Bayes from Scratch](#)

What'll you learn in this course:

- What is Naive Bayes?
- What are the different types of Naive Bayes algorithms?
- Components of Naive Bayes and the role of probability
- How to implement Naive Bayes?

[Enroll For Free](#)

We have many other Free Courses in Data Science for you right here:

[View All Free Courses](#)

How to Become a Data Scientist



Dice Staff

November 20, 2019

42 min read

ARTIFICIAL INTELLIGENCE DATA DATA ANALYST DATA ENGINEER
DATA SCIENCE DATA SCIENTIST MACHINE LEARNING



What is a data scientist? If you ask the Harvard Business Review, it's the “[sexiest job of the](#)”

Trending Stories

[Are H-1B Visa Workers Overpaid or Underpaid?](#)

[Top Tech Skills Amazon is Hiring For](#)

[Airbnb Laid-Off Technologists Could Head to... JPMorgan?](#)

[24 Programming Languages That Pay Top Salaries: Scala and More](#)

Related Jobs

Search jobs in these related categories...

[Data Scientist](#)

[Data Analyst](#)

[Data Engineer](#)

Salary Predictor

Ciencia de Datos: qué es, dónde estudiar y por qué Argentina lidera en el mundo

Argentina destaca a nivel mundial por sus conocimientos en bases de datos y habilidades IT: cuáles son las claves para no quedar atrás.



<https://www.infotechnology.com/online/Ciencia-de-Datos-que-es-donde-estudiar-y-por-que-Argentina-lidera-en-el-mundo-20200717-0003.html>

7 TIPOS DE CIENTÍFICOS DE DATOS, LA PROFESIÓN QUE OPTIMIZARÁ TECNOLOGÍA Y NEGOCIOS

Por AETecno en Empresas 25 de Noviembre de 2015, 13:10



Conozca los diversos perfiles de analistas y expertos en Big Data que dan forma a la profesión de moda en el mundo y la “más sexy del siglo XXI”, clave para conducir proyectos TI en las empresas.

The End