Mi vida después del bootcamp

Data Science Bootcamp

The Bridge

Ya he terminado el bootcamp…

¿Y ahora qué?


I SHOULD START A JOB SEARCH
made on imgur


WE ARE READY
FOR THAT VACATION
imgflip.com

Tiempo de reset y pensar por dónde tirar

Vale, ¿y después qué?

¿Qué te gusta/se te da bien?

De todo lo que has visto, ¿con qué te quedas?

Ramp Up

1. Programación con Python
2. Programación orientada a objetos

Data Analysis

1. Acceso a datos: APIs, Bases de datos, SQL
2. Limpieza de datos: Numpy, Pandas
3. Automatización: de procesos, web scrapping
4. Visualización con Python
5. Dashboards de Streamlit
6. Analítica de los datos
7. Estadística
8. Business Intelligence
9. Mapas

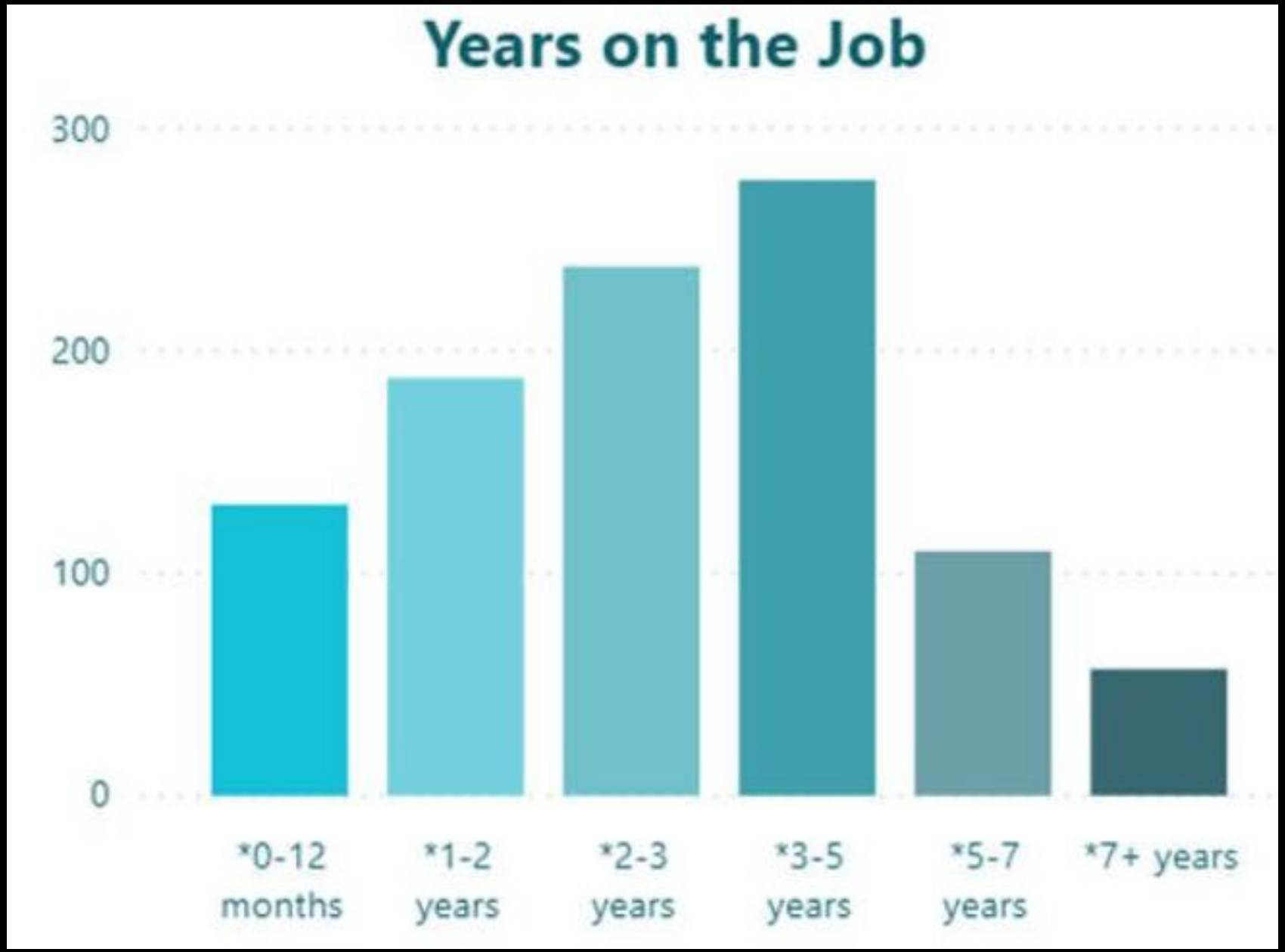
Machine Learning

1. Automatizar toma de decisiones
2. Algoritmo supervisados
3. Comprensión de los algoritmos y variables que impactan en la toma de decisión
4. Algoritmos no supervisados
5. NLP
6. Deep Learning
7. Time Series
8. Aprendizaje por refuerzo

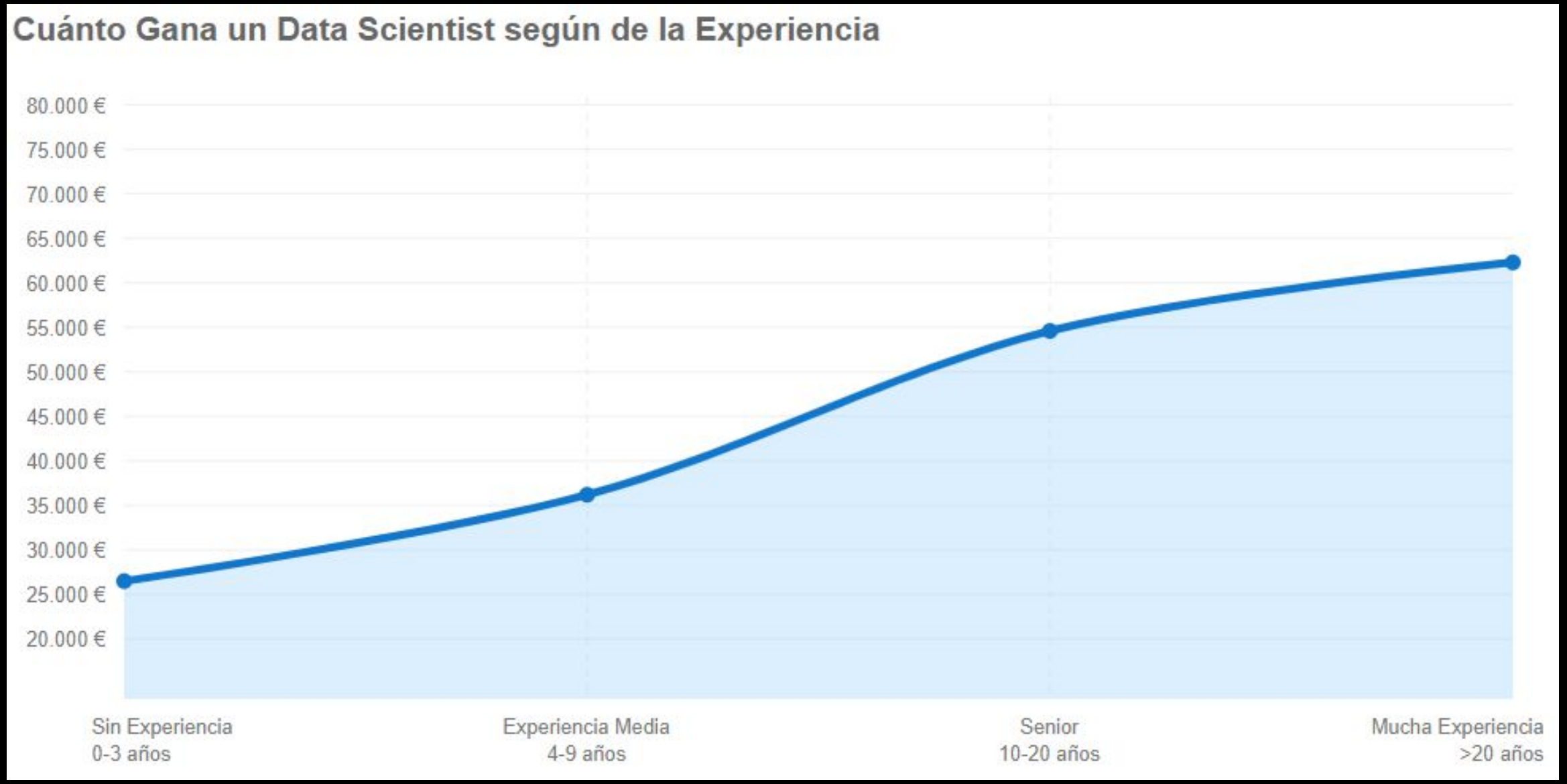
Data Engineer

1. APIs
2. Cloud, despliegue de recursos
3. Máquinas, servidores, redes
4. Big Data & Spark
5. Airflow
6. Docker

¿Cómo está el mercado?

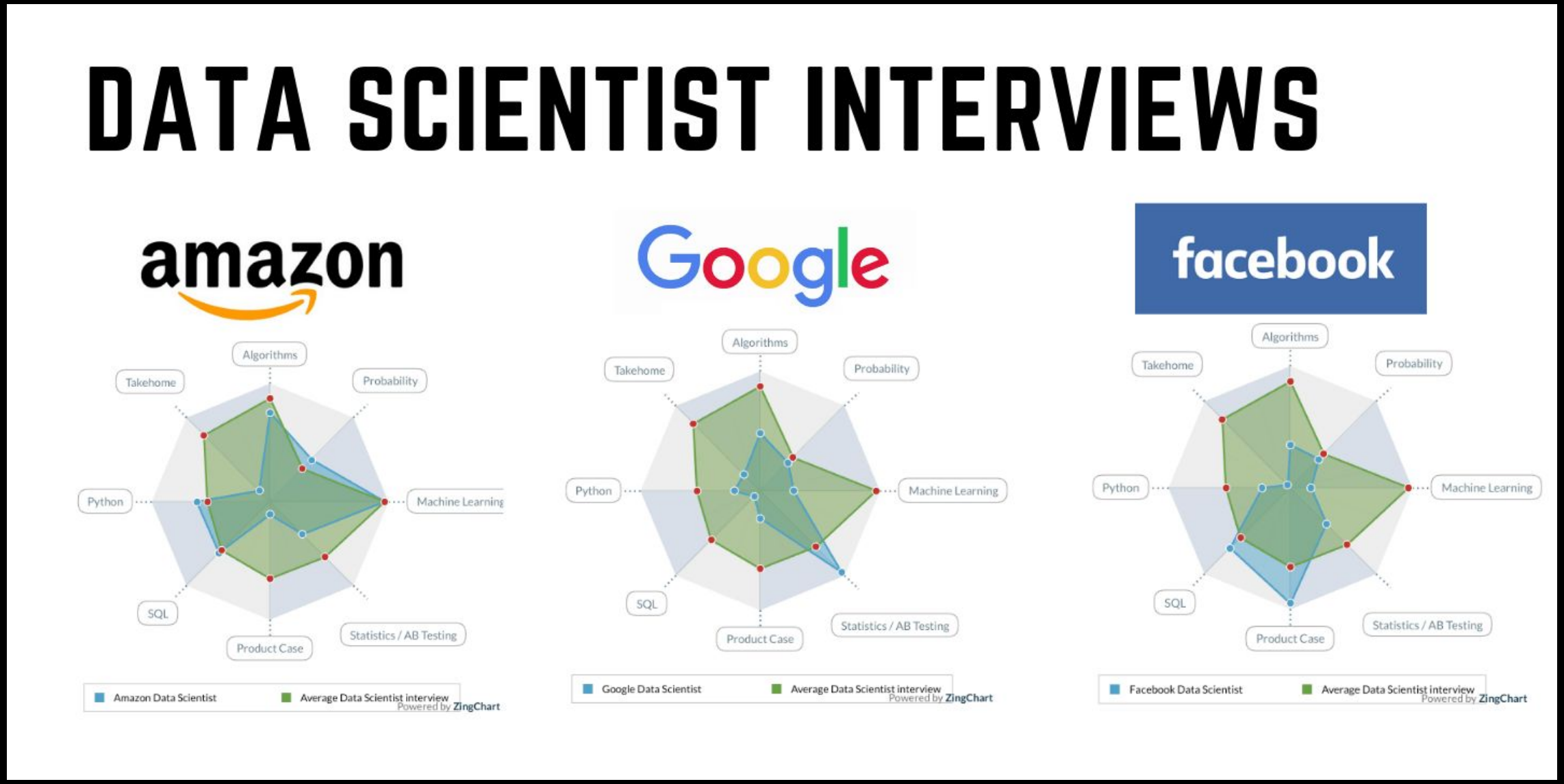

Years on the Job
300
200
100
0
*0-12 months
*1-2 years
*2-3 years
*3-5 years
*5-7 years
*7+ years

Fuente

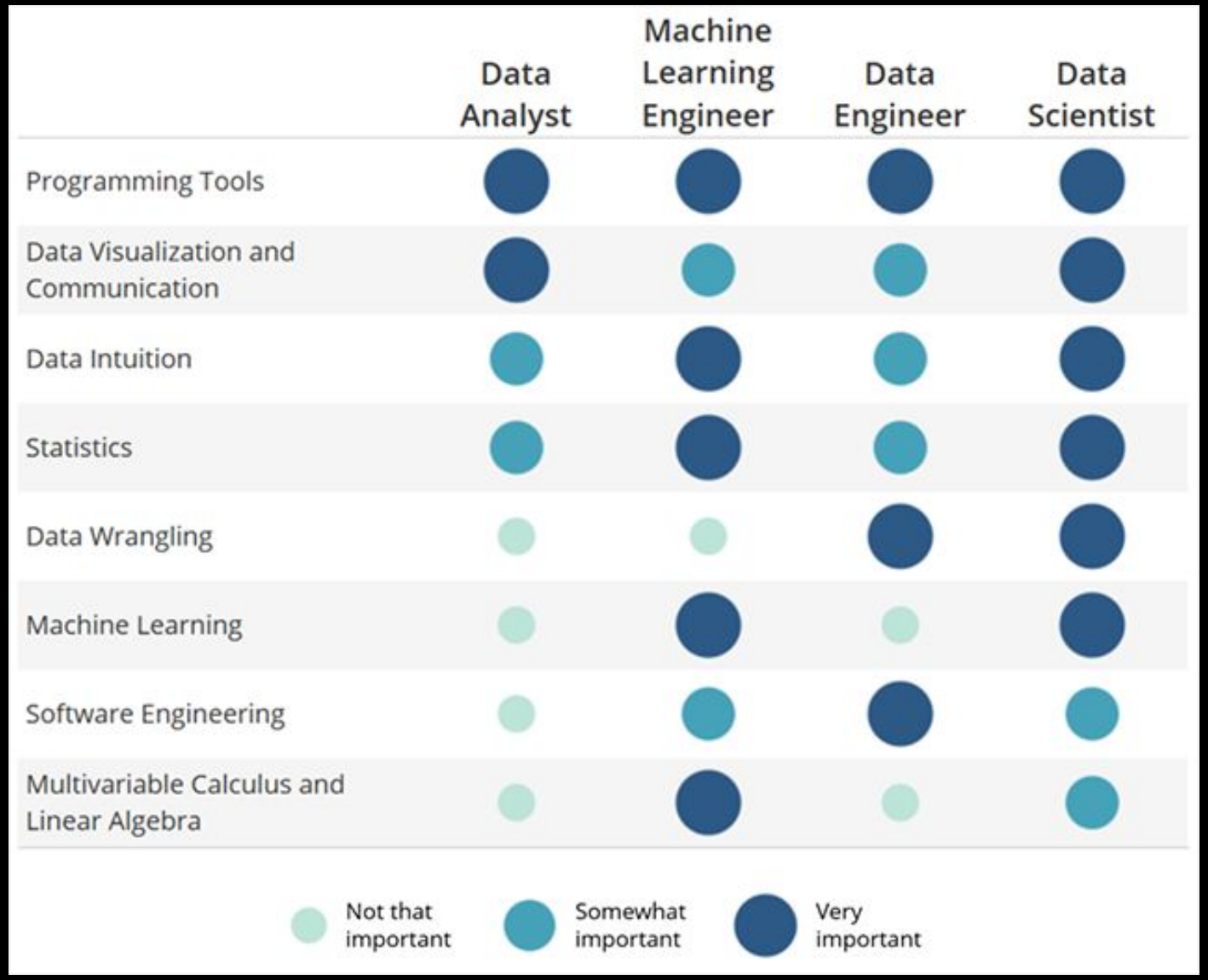

Cuánto Gana un Data Scientist según de la Experiencia
80.000 €
75.000 €
70.000 €
65.000 €
60.000 €
55.000 €
50.000 €
45.000 €
40.000 €
35.000 €
30.000 €
25.000 €
20.000 €
Sin Experiencia
0-3 años
Experiencia Media
4-9 años
Senior
10-20 años
Mucha Experiencia
>20 años

Fuente

¿Qué tipos de preguntas se hacen en una entrevista técnica?



Puestos de trabajo de Data Science


Data Analyst
Machine Learning Engineer
Data Engineer
Data Scientist
Programming Tools
Data Visualization and Communication
Data Intuition
Statistics
Data Wrangling
Machine Learning
Software Engineering
Multivariable Calculus and Linear Algebra
Not that important
Somewhat important
Very important

Data Analyst

SKILLS

1. **Analítica**: exploratorio, obtener insights, encontrar patrones
2. **Acceso a Bases de datos**: interno y externo a la empresa.
3. **Data Cleaning**
4. **Data Quality**
5. **Business understanding**
6. **Communication**: presentar resultados. Storytelling.
7. **Mates y estadística**
8. **Project manager**
9. **Dashboards/Reportes**: Automatizado
10. **Machine Learning**: no experto

TOOLS

1. **Lenguajes de programación**: SQL, Python, R
2. **Business intelligence**: Power BI, Tableau, Microstrategy, QlikSense, DataStudio.
3. **Relational DB**: MySQL, SQLServer, Oracle, MariaDB.
4. **Non relational DB:** Cassandra, MongoDB
5. **Excel**
6. **IDEs**: Pycharm, Anaconda, Jupyter, Rstudio
7. **Big Data**: pyspark Hive, ecosistema Cloudera

Machine Learning Engineer

SKILLS

1. **Modelos**: Regresión lineal, SVM, Kmeans…
2. **Matemáticas y estadística**: álgebra, probabilidad, estadística, cálculo, distribuciones…
3. **Tratamiento de imágenes**
4. **Redes neuronales**
5. **NLP**
6. **Comunicación**
7. **Conocimiento de negocio**
8. **Avanzado**: robótica, vídeo, integración con sistemas, automatización

TOOLS

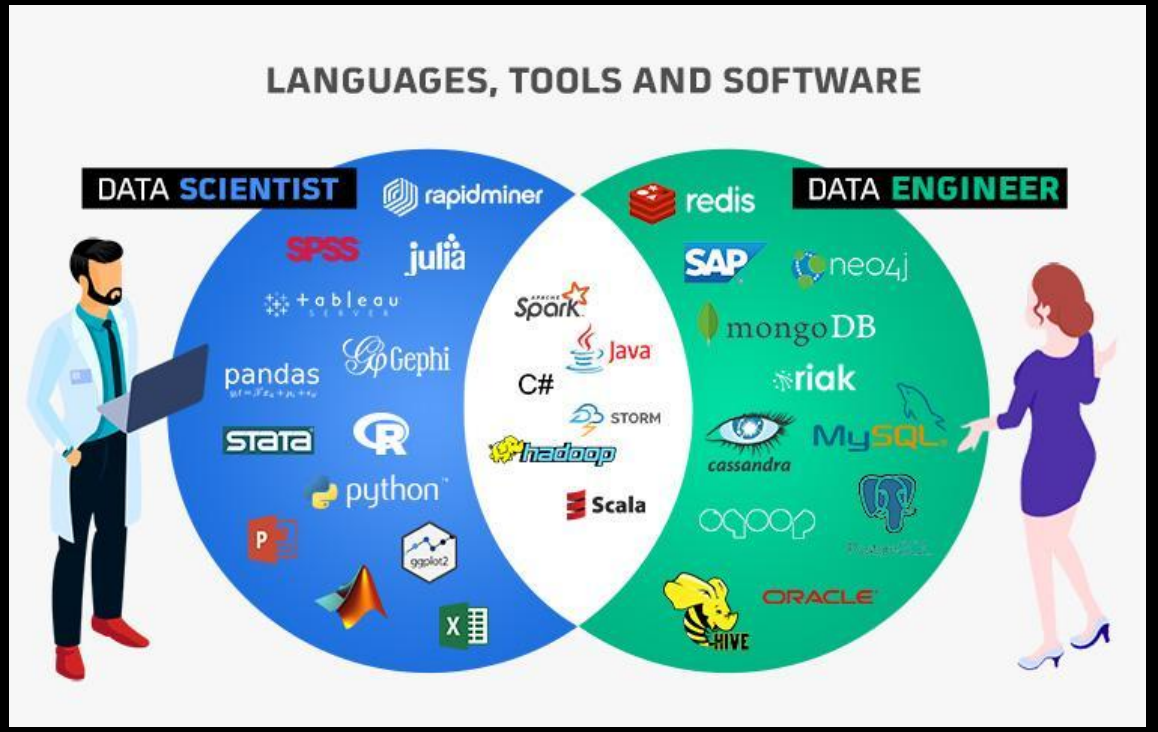
1. **Lenguajes de programación**: SQL, Python, R, C++ (comp. paralela)
2. **Deep Learning**: Tensorflow, Keras, Pytorch
3. **Relational DB**: MySQL, SQLServer, Oracle, MariaDB.
4. **Non relational DB:** Cassandra, MongoDB
5. **IDEs**: Pycharm, Anaconda, Jupyter, Rstudio
6. **Big Data**: pyspark Hive, Spark MLLib, ecosistema Cloudera
7. **Cloud**: soluciones de IA en cloud.
8. **Automation**: MLOps
9. **Otros**: MLFlow, Tensorboard

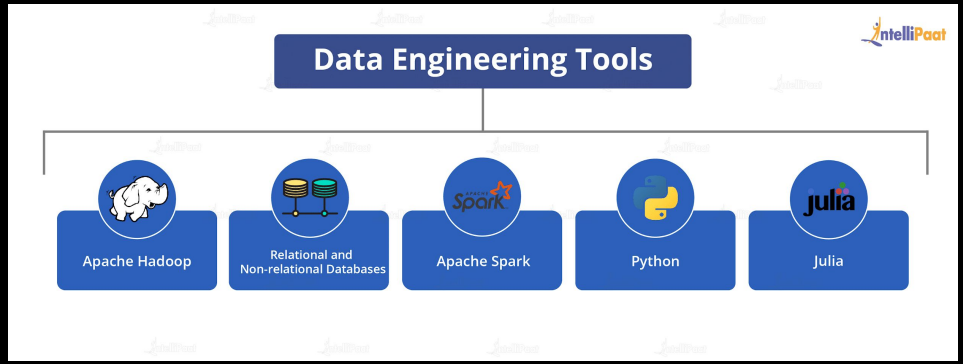
Data Engineer

Responsibilities...

ON DAILY BASIS

- **Develop, construct, test, and maintain architectures** (such as databases and large-scale processing systems)
- **Ensure architecture** will support the requirements of the business
- Discover opportunities for **data acquisition**
- **Develop data set processes** for data modeling, mining and production
- Employ a variety of languages and tools (e.g. scripting languages) to **marry systems together**
- Recommend ways to **improve data** reliability, efficiency and quality
- Conduct research to **answer industry and business questions**
- **Leverage large volumes of data** from internal and external sources to answer that business
- Employ sophisticated analytics programs, machine learning and statistical methods to **prepare data for use in predictive and prescriptive modeling**
- Explore and examine data to **find hidden patterns**
- **Automate work** through the use of predictive and prescriptive analytics
- **Tell stories to key stakeholders** based on their analysis





Next steps Data Engineer

1. **Bases de datos**: montar y gestionar bases de datos relacionales y no relacionales.
2. **Data Warehouse**: ETLs, ingestas de datos, modelado de datos para su posterior consumo.
3. **Redes**: comunicaciones, infraestructura, acceso a máquinas, bash, scripting
4. **Cloud**: IaaS, PaaS, seguridad, acceso usuarios, despliegues automáticos, migraciones a cloud.
5. **Virtualización y contenedores**: Docker y Kubernetes. Arquitecturas de microservicios.
6. **Big Data**: ecosistema Hadoop, Spark, Spark Streaming, Databricks, integración con Cloud.
7. **Automatización con pipelines**: scripting con Python, Airflow, Luigi, crontab.
8. **DevOps**: Automatización de puestas en producción.
9. **Real time**: Kafka, Spark Streaming, soluciones cloud.
10. **Otros lenguajes**: Python, SQL, Java, Scala, C, Julia.

Otros roles

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*

HIRED BY

VISA Coca-Cola

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

STATISTICIAN

'HISTORIC LEADERS OF DATA'

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*

HIRED BY

Johnson & Johnson PEPSICO

Role

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

Mindset

Logical and enthusiastic stats genius

DATA AND ANALYTICS MANAGER

DATA SCIENCE TEAM LEADER

Role

Manages a team of analysts and data scientists

Mindset

Data Wizards' Cheerleader

HIRED BY

coursera slack MOTOROLA SOLUTIONS

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*

DATABASE ADMINISTRATO

'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

tableau reddit

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*

BUSINESS ANALYST

CHANGE AGENT

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*

HIRED BY

UBER DELL ORACLE

Role

Improves business processes as intermediary between business and IT

Mindset

Resilient project juggler

Ah, entonces tengo que sacarme certificaciones y cursos

NO!

HAZ PROYECTOS. Son el FIN para aprender, los cursillos y libros el MEDIO.
Un proyecto te aporta más motivación y mayor visibilidad en tu CV