



مسیر ورود به علوم داده



Ahmad Rasti



Master of Data Science at Isfahan University of Technology



Data Scientist at Gantoum Company



Graph Neural Network Researcher



GNN Department Manager at BigDataworld Institute



Deep Learning and Neural Network Instructor



ahmad-rasti-a47483248

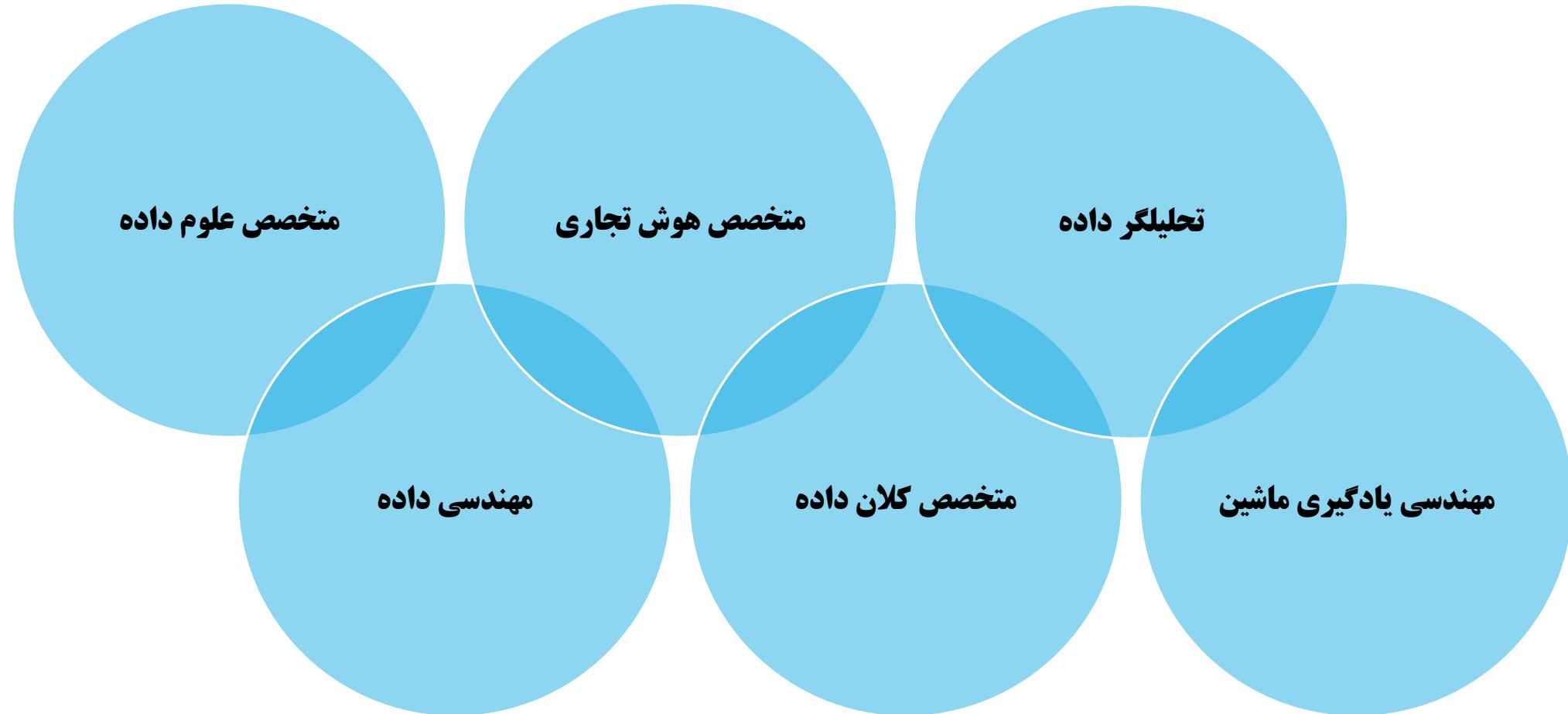


ahmadrasti11@gmail.com



github.com/ahmadrasti

انواع مسیرهای شغلی متخصص داده



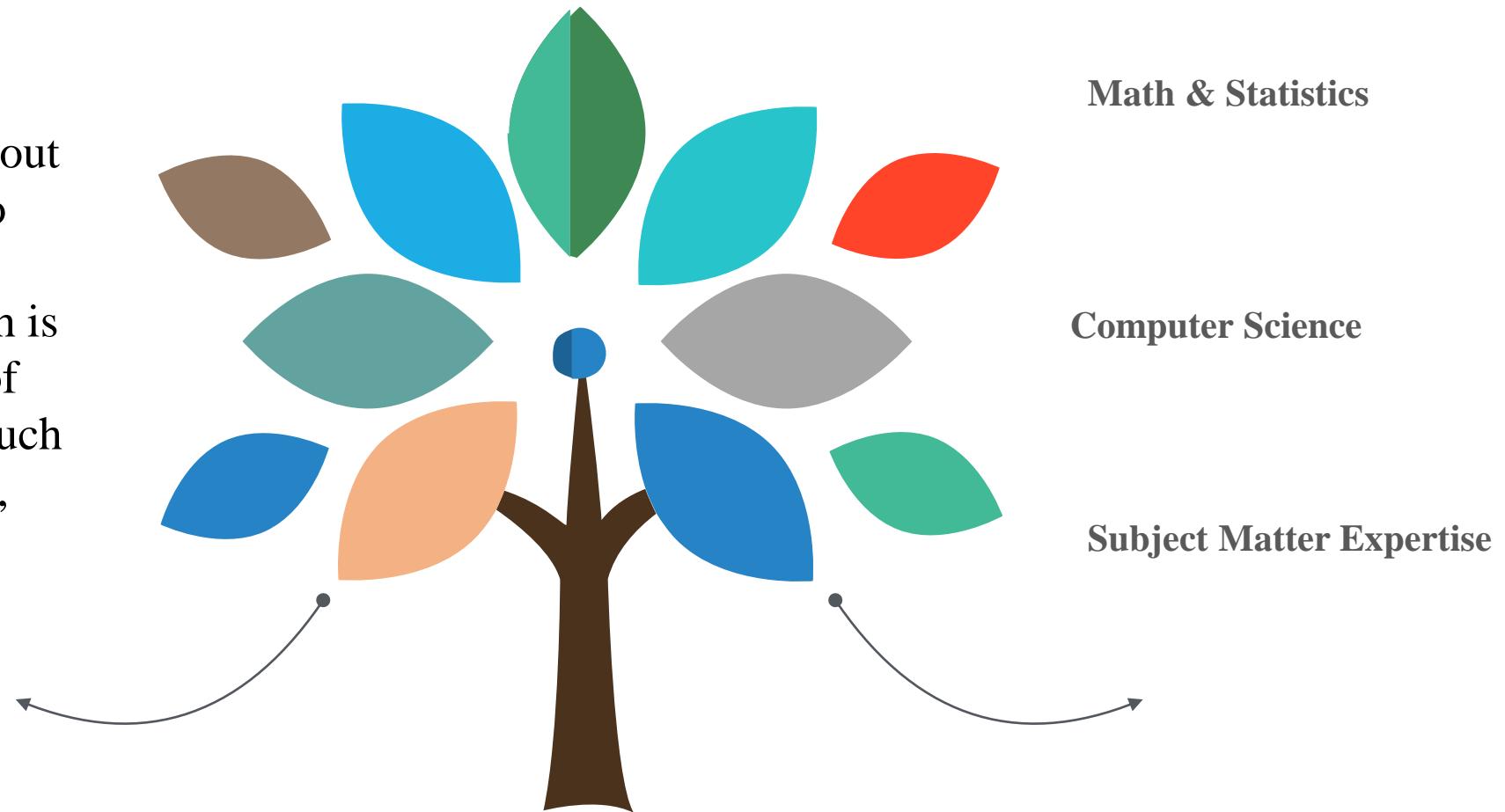


متخصص علوم داده

Data Scientist

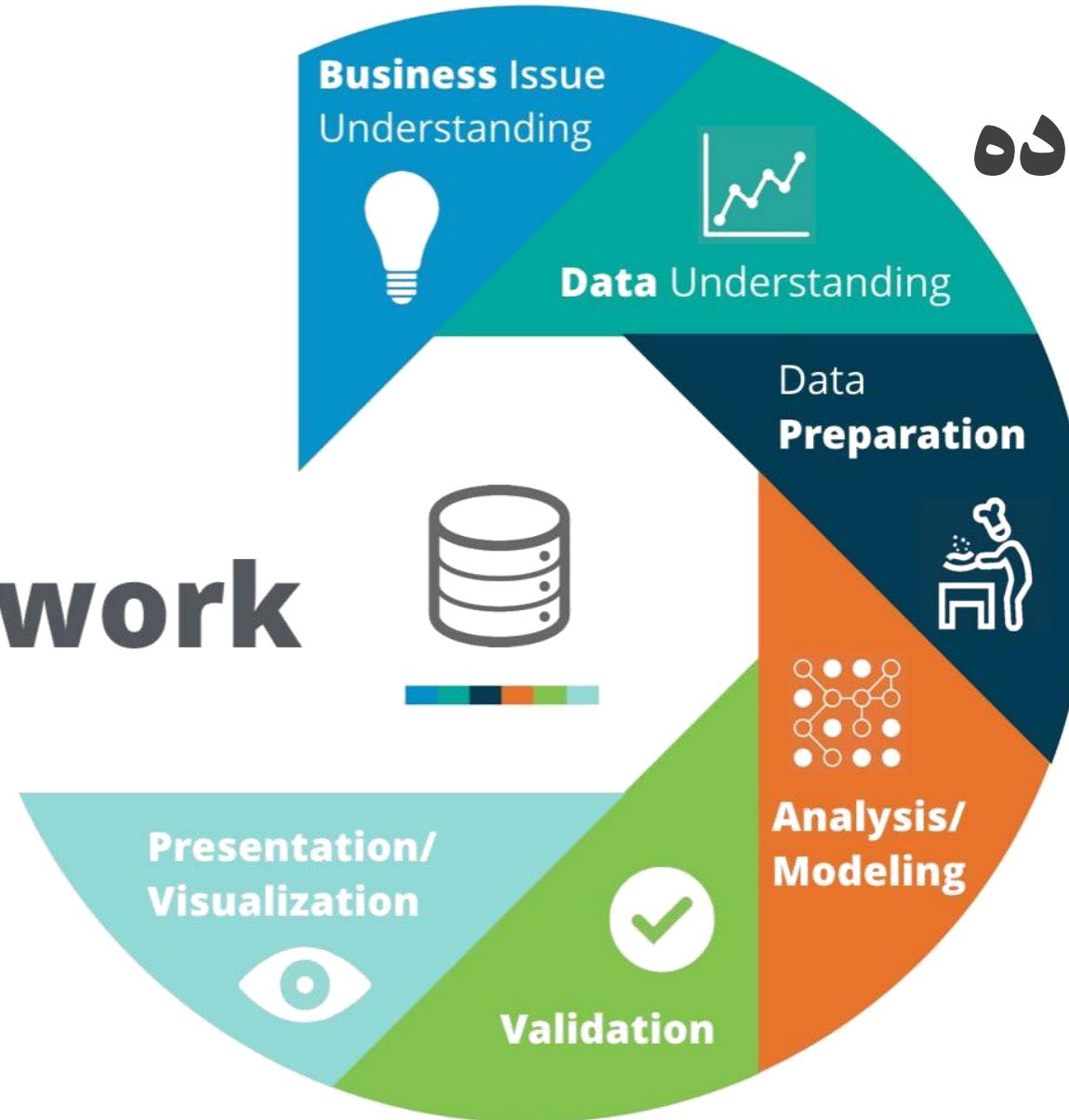
علوم داده

Data Science is an interdisciplinary field about processes and systems to extract knowledge or insights from data, which is a continuation of some of the data analysis fields such as statistics, data mining, and predictive analytics, similar to Knowledge Discovery in Databases (KDD).

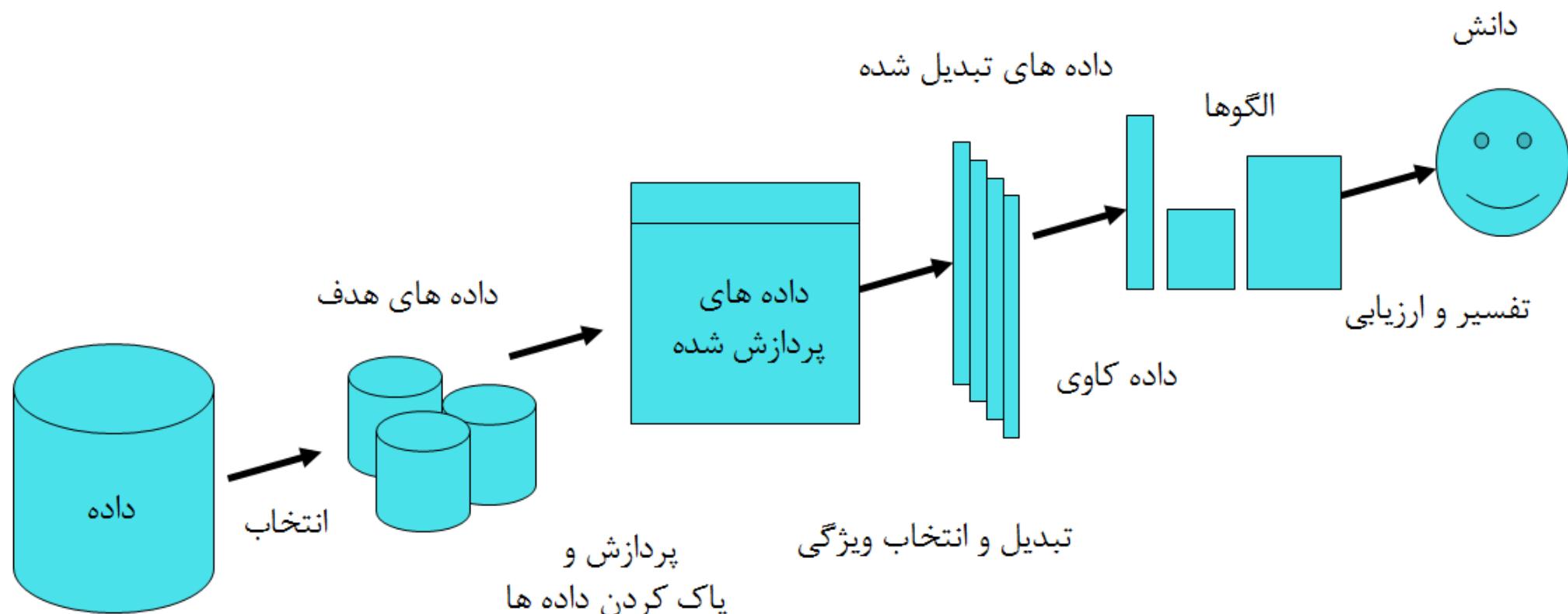


چرخه پروژه‌های علوم داده

Framework



مراحل اجرایی پروژه‌های علم داده



علم داده عبارت است از استخراج اطلاعات و دانش و **کشف الگوهای پنهان** از پایگاه داده‌های بسیار بزرگ

مهارت‌ها و ابزارهای کلیدی

ابزارهای کلیدی

- ❖ Python
- ❖ R
- ❖ Excel
- ❖ SQL
- ❖ Data Visualization Tools (Power BI, Tableau)
- ❖ Deep Learning Tools(TensorFlow, Keras,)
- ❖ Big Data Analytics (Spark, Flink ,....)

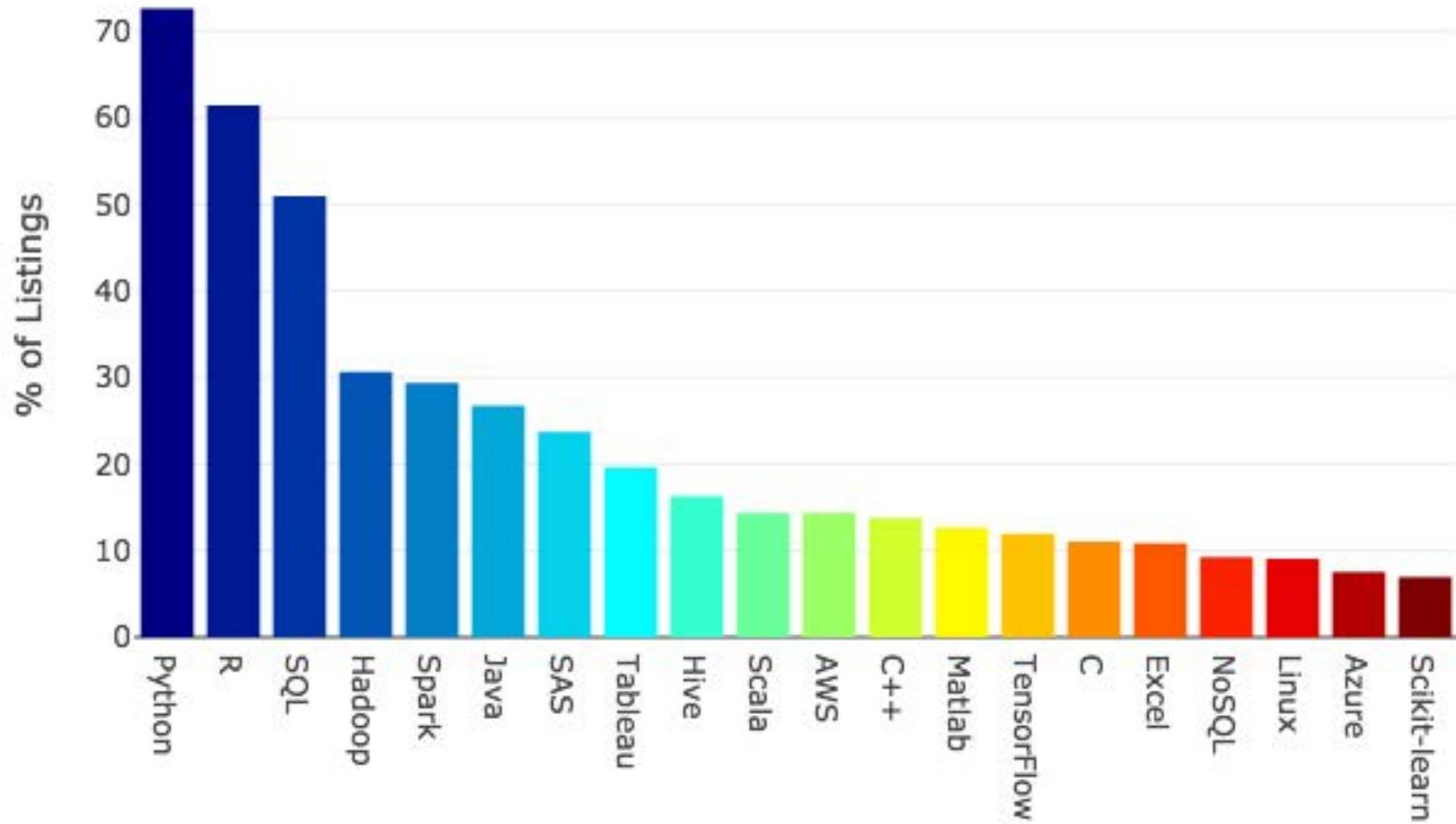
مهارت‌ها

- آمار
- برنامه‌نویسی
- یادگیری ماشین
- یادگیری عمیق
- داستان‌سرایی داده‌ها
- خلاقیت و توانایی حل مساله
- کار تیمی
- تفکر انتقادی

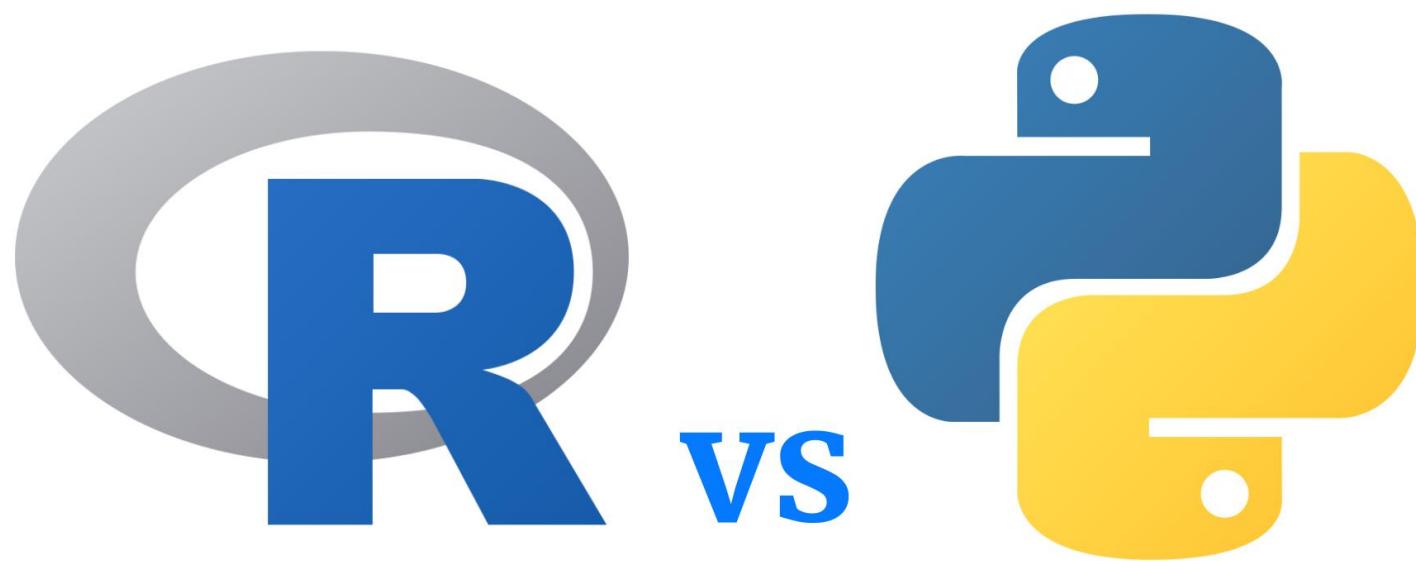


مهارت‌ها و ابزارهای کلیدی

Top 20 Technology Skills in Data Scientist Job Listings



تفاوت پایتون و R



برترین کتابخانه‌های علوم داده پایتون



TensorFlow

"TensorFlow is an open source software library for numerical computation using data flow graphs.



PyTorch

"PyTorch is a Python package that provides Deep neural networks built on a tape-based autograd system



Numpy

"NumPy is the fundamental package needed for scientific computing with Python.



Scikit-Learn

"scikit-learn is a Python module for machine learning built on NumPy, SciPy and matplotlib.



Keras

"Keras is a high-level neural networks API, written in Python and capable of running on top of TensorFlow, CNTK, or Theano.



Scipy

"SciPy is open-source software for mathematics, science, and engineering.



Pandas

"pandas is a Python package providing fast, flexible, and expressive data structures designed to make working with "relational" or "labeled" data both easy and intuitive



Matplotlib

"Matplotlib is a Python 2D plotting library which produces publication-quality figures in a variety of hardcopy formats and interactive environments across platforms.



Scrapy

"Scrapy is a fast high-level web crawling and web scraping framework, used to crawl websites and extract structured data from their pages.



مسیر یادگیری





منابع آموزشی

coursera

Explore ▾

What do you want to learn?



For Enterprise

For Students

Log In

Join for Free

Browse > Data Science > Data Analysis

Applied Data Science with Python Specialization

Gain new insights into your data . Learn to apply data science methods and techniques, and acquire analysis skills.

★★★★★ 4.5 23,305 ratings



Christopher Brooks +3 more instructors

Offered By



Enroll for Free

Starts Sep 16

Financial aid available

298,780 already enrolled



منابع آموزشی

Coursera

Explore ▾

What do you want to learn?



For Enterprise

For Students

Log In

Join for Free

Browse > Data Science > Data Analysis

IBM Data Science Professional Certificate

Kickstart your career in data science & ML. Build data science skills, learn Python & SQL, analyze & visualize data, build machine learning models. No degree or prior experience required.

★★★★★ 4.6 49,252 ratings



Rav Ahuja [+11 more instructors](#)

Offered By



Enroll for Free

Starts Sep 16

Financial aid available

52,601 already enrolled



منابع آموزشی



Categories



Search for anything

Udemy Business

Teach on Udemy



Log in

Sign up



Data Science > Python

Machine Learning A-Z™: Hands-On Python & R In Data Science

Learn to create Machine Learning Algorithms in Python and R from two Data Science experts. Code templates included.

Bestseller

4.5 ★★★★★ (150,444 ratings) 805,684 students

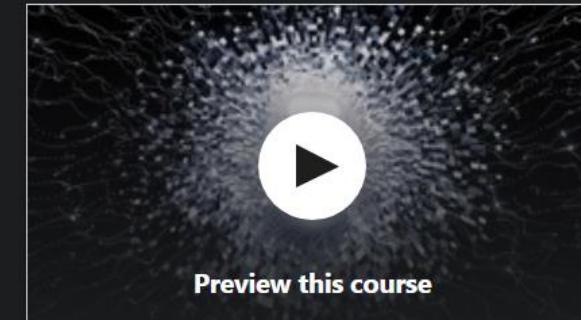
Created by Kirill Eremenko, Hadelin de Ponteves, SuperDataScience Support, Ligency Team

Last updated 9/2021 English English [Auto], French [Auto], 4 more

Wishlist

Share

Gift this course



€94.99

Add to cart

Buy now



منابع آموزشی

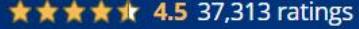
Coursera [Explore ▾](#) What do you want to learn? 

[For Enterprise](#) [For Students](#) [Log In](#) [Join for Free](#)

Browse > [Data Science](#) > [Data Analysis](#)

Data Science Specialization

Launch Your Career in Data Science. A ten-course introduction to data science, developed and taught by leading professors.

 4.5 37,313 ratings

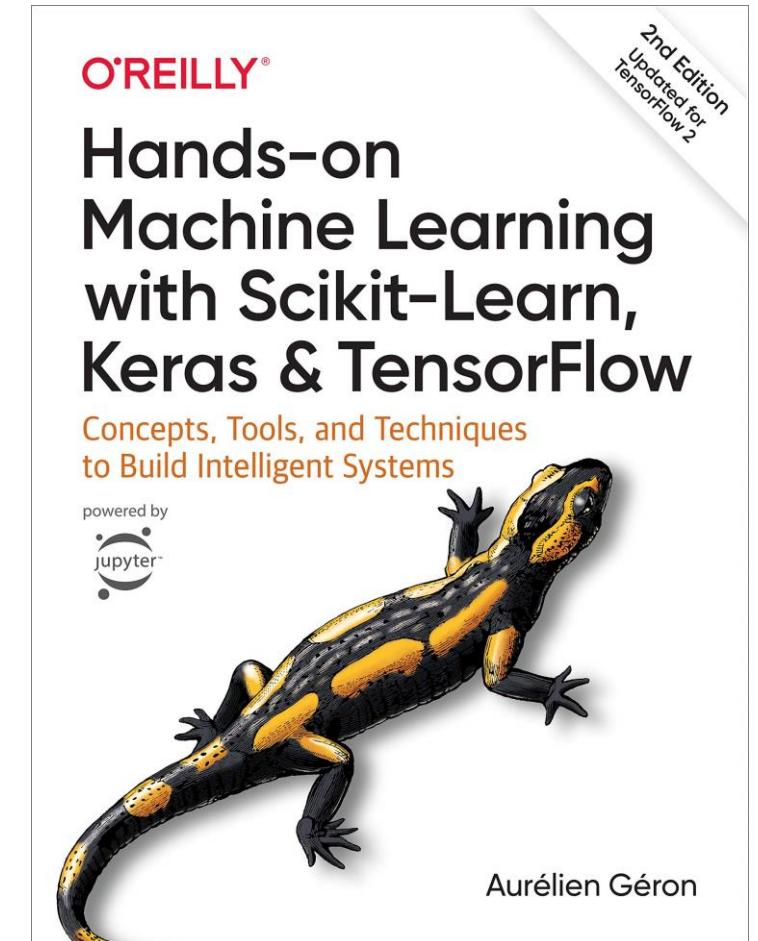
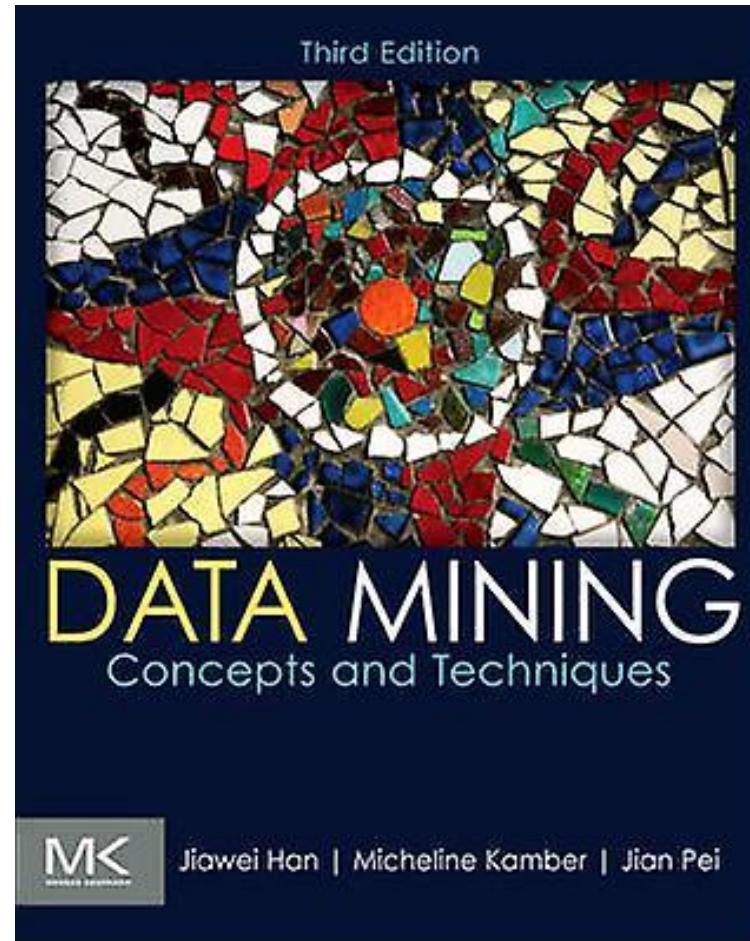
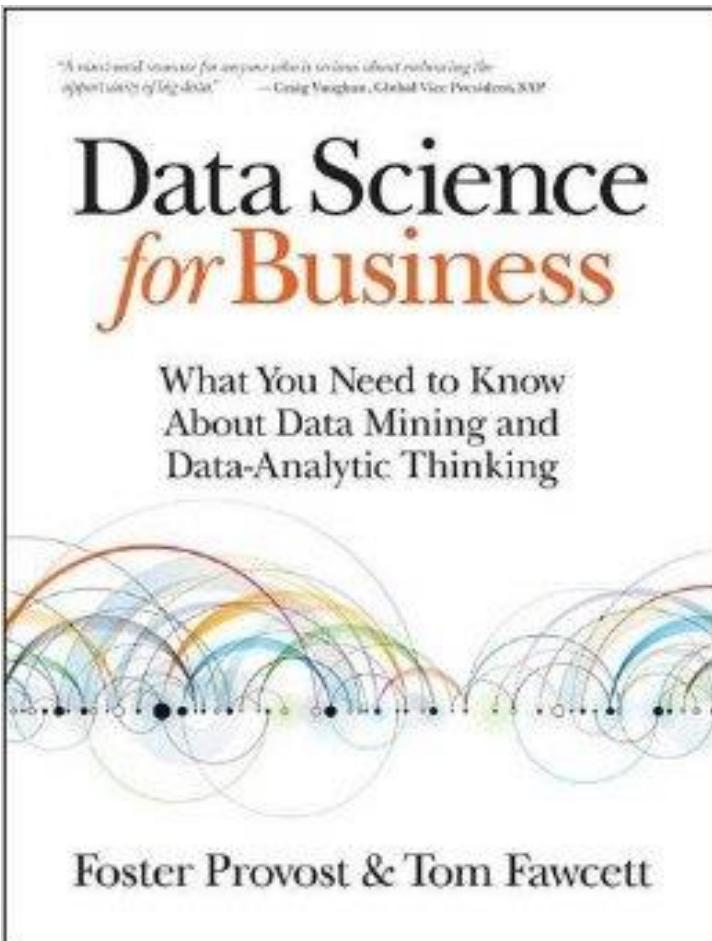
 Jeff Leek, PhD [+2 more instructors](#)

[Enroll for Free](#) Financial aid available
[Starts Sep 16](#)

440,345 already enrolled

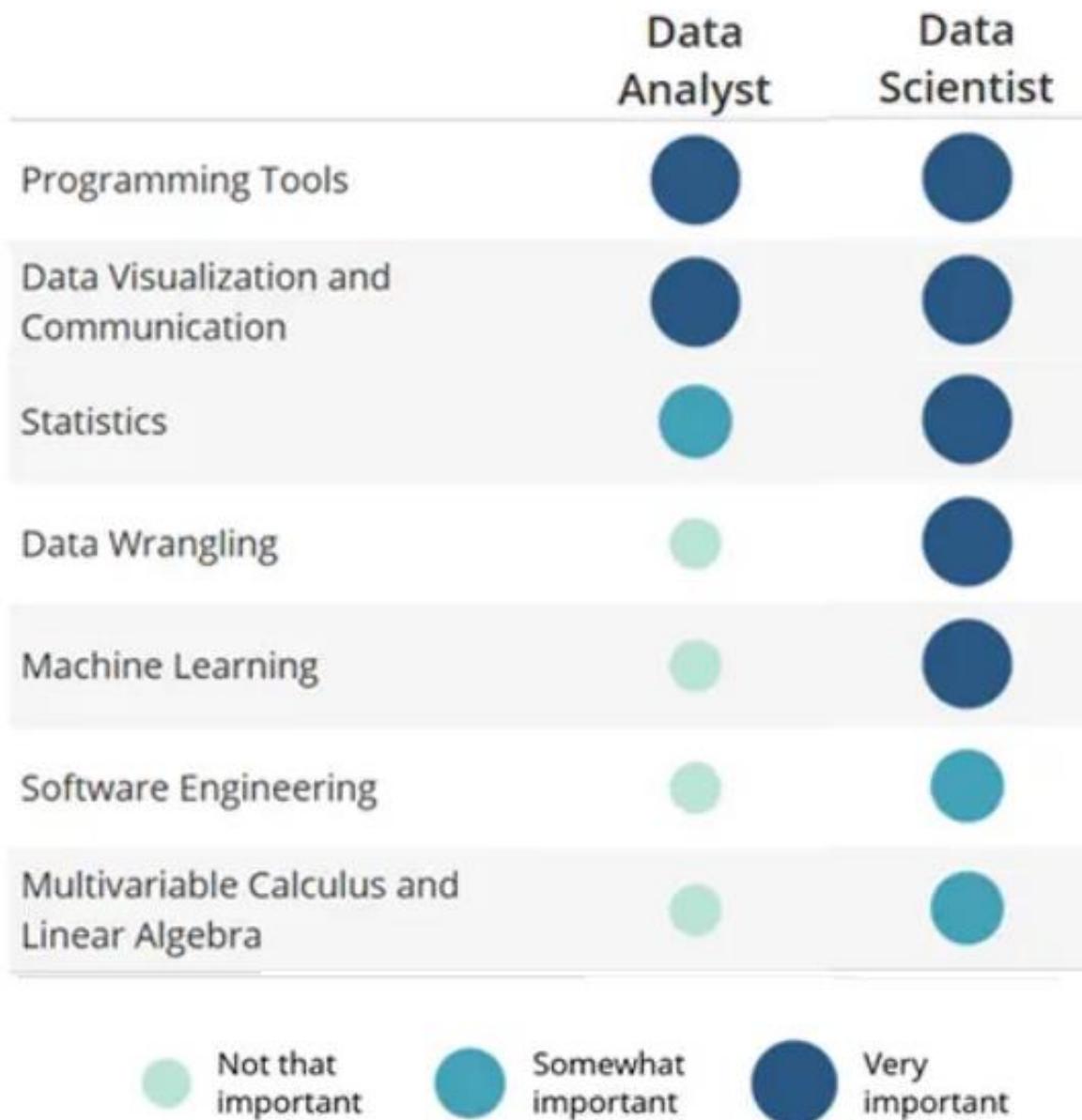


کتاب‌های مرجع

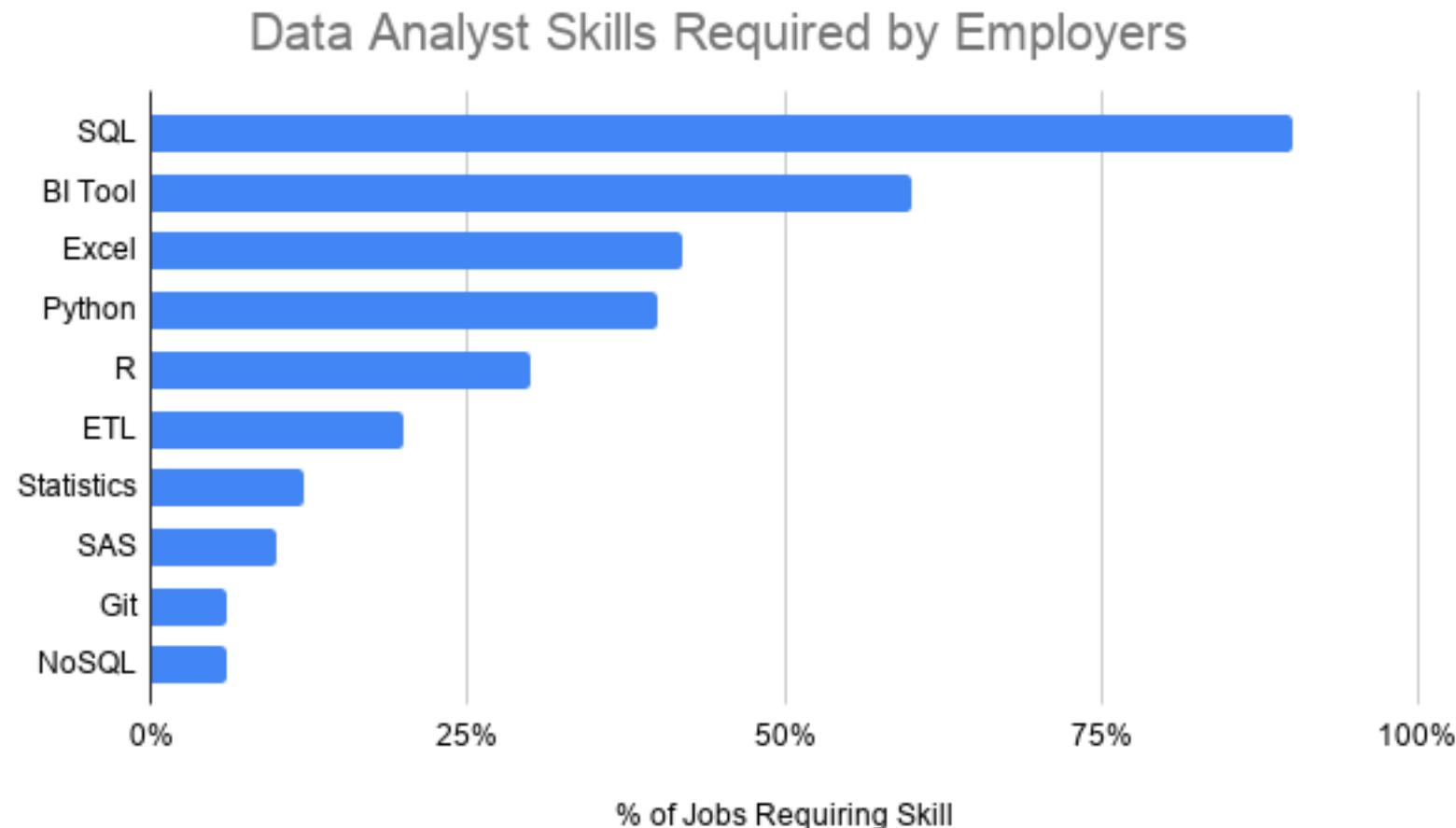




تحلیلگر داده



مهارت‌های تحلیلگر داده

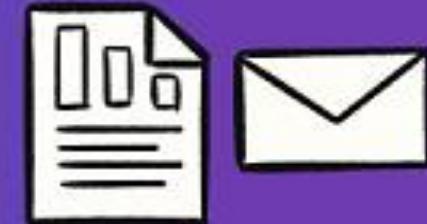


Business Intelligence

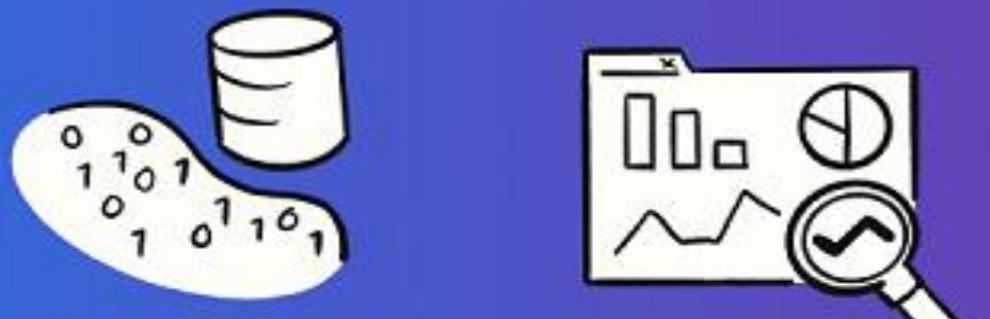


Collection

BI



Reporting



Data
Engineering

Storage



Analysis

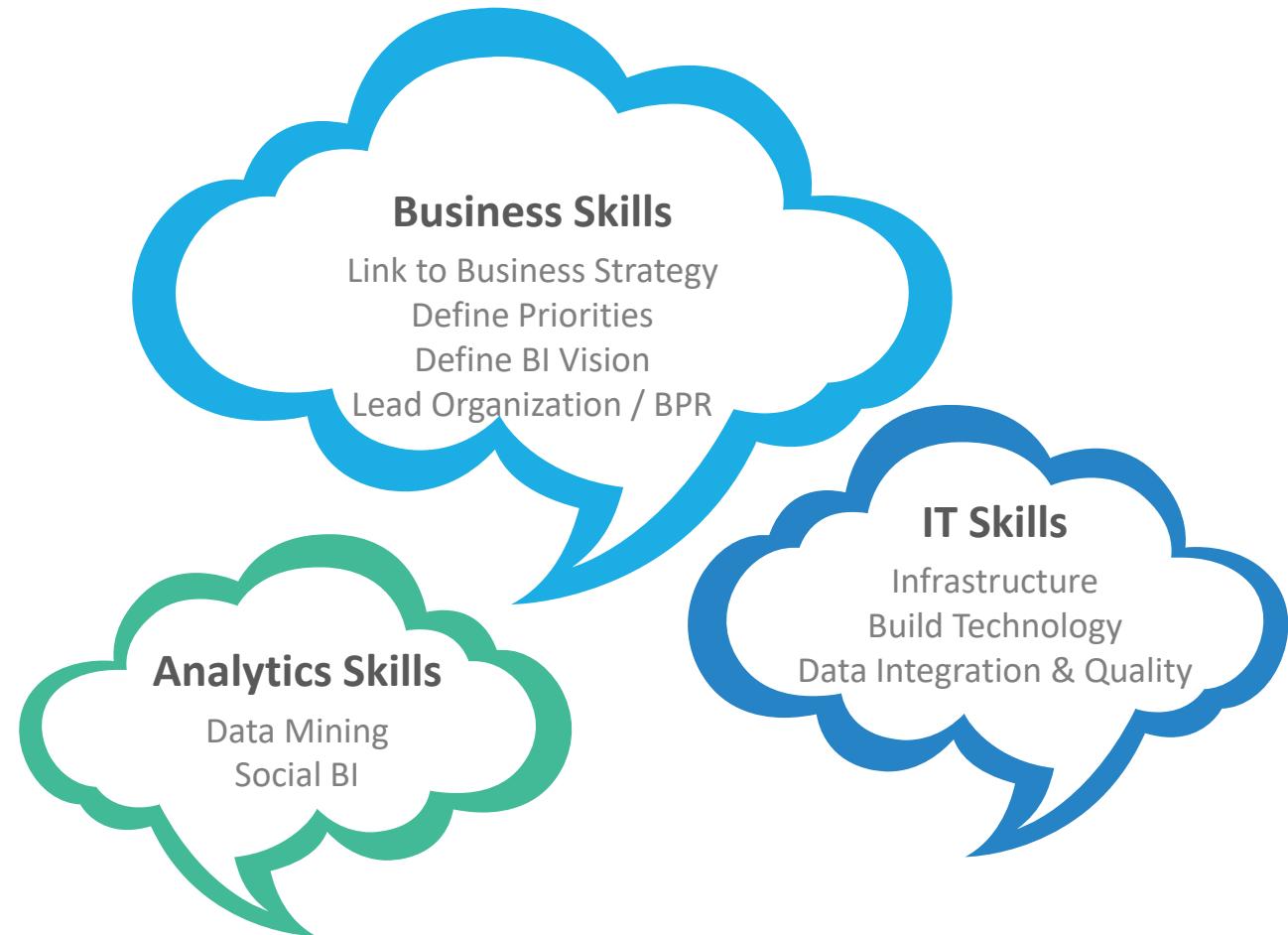
BI
Tools

متخصص هوش تجاري

Business Intelligence Expert

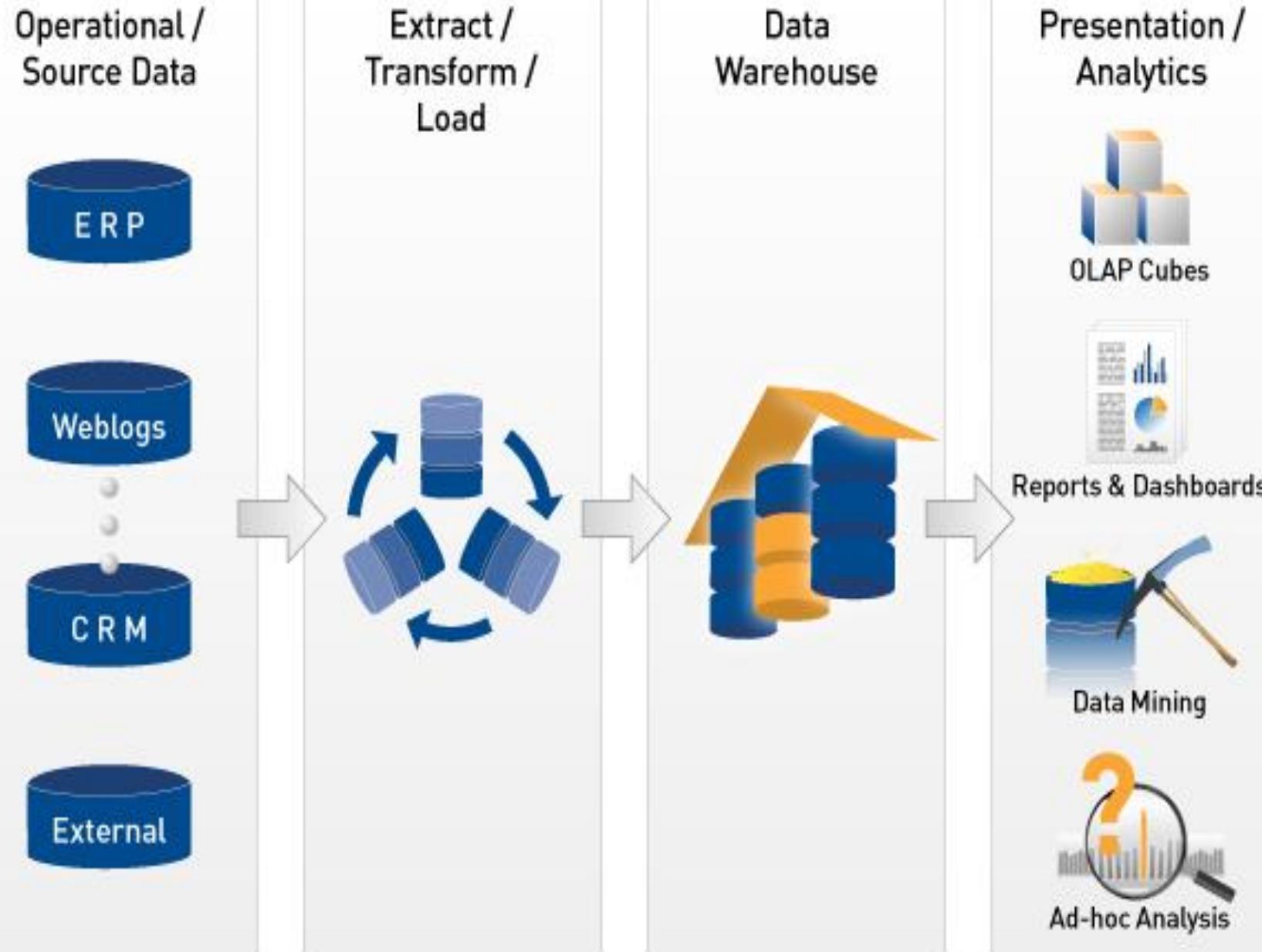
هوش تجاری

BI encompasses a wide variety of tools, applications and methodologies that enable organizations to collect data from internal systems and external sources; prepare it for analysis; develop and run queries against that data; and create reports, dashboards and data visualizations to make the analytical results available to corporate decision-makers, as well as operational workers.





فازهای پروژه‌های هوش تجاری



مهارت‌ها و ابزارهای کلیدی

ابزارهای کلیدی

- ❖ SQL
- ❖ Microsoft SQL Server BI (SSIS, SSAS, SSRS)
- ❖ Oracle BI (ODI, OBIEE)
- ❖ Data Visualization Tools (Power BI, Tableau)

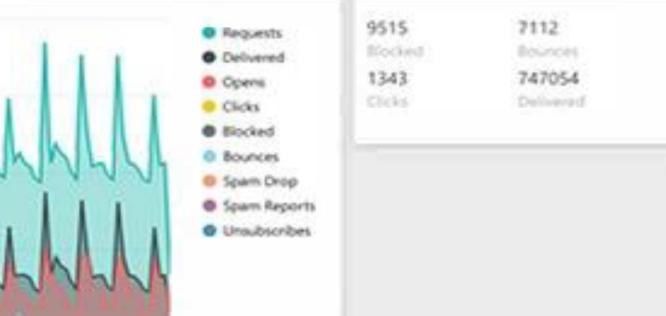
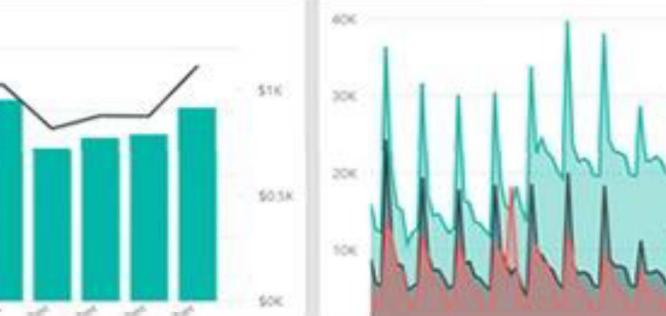
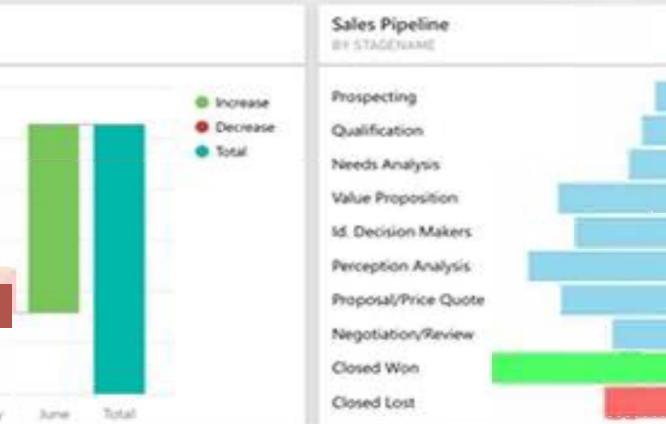
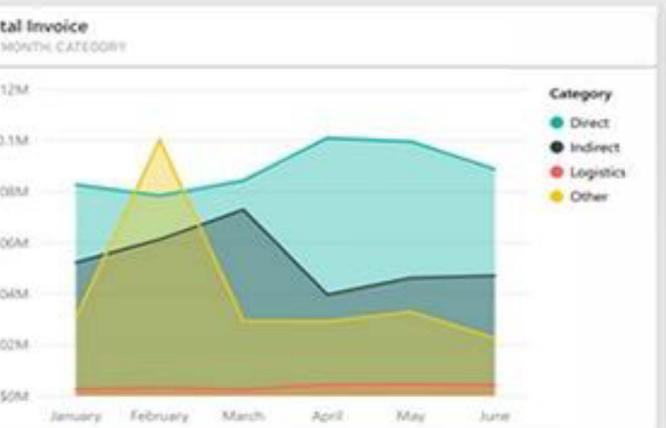
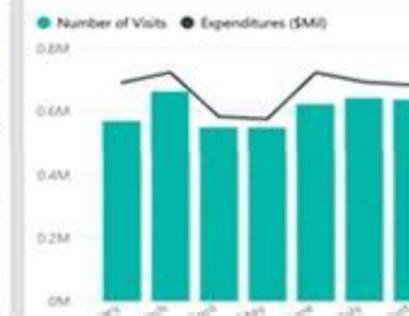
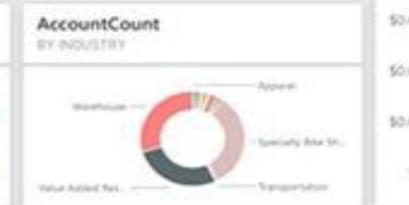
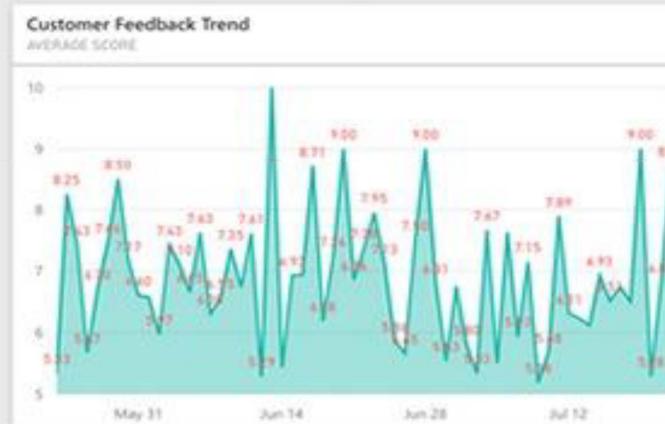
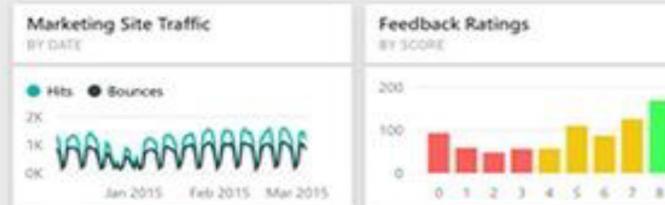
مهارت‌ها

- پایگاه داده
- معماری انبارداده
- استخراج، تبدیل و بارگذاری
- تحلیل داده و داده کاوی
- تحلیل کسب و کار
- استخراج شاخص‌های کلیدی عملکرد (KPI)
- گزارش دهنده

Executive Metrics Dashboard

Q Share Dashboard

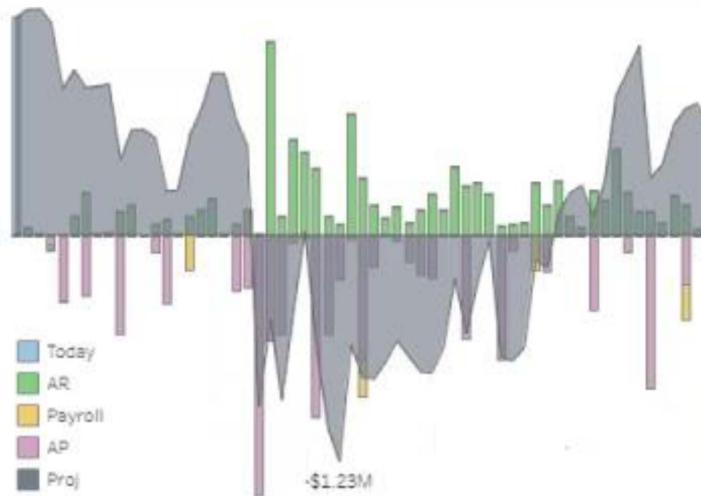
Ask a question about the data on this dashboard.



CFO Dashboard

10/5/2017 9:36:00 AM

60-day cash flow projection



Orders to Ship: \$2,802K



Hot New Opportunities: \$4,045K



AR Aging: \$4,971K



Revenue YoY

2016 \$7.90M
2017 \$8.84M
11.9% ↑



Gross Profit



EBITDA



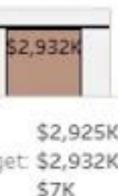
Admin



Mkt



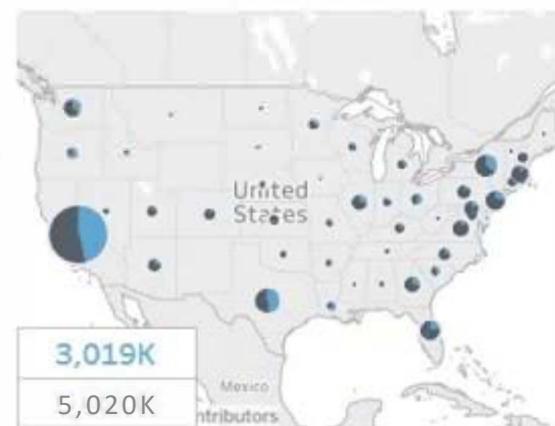
Opex



Avg Invoice Amt Trend

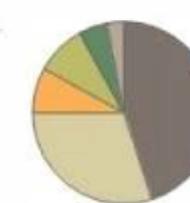


Website Visits

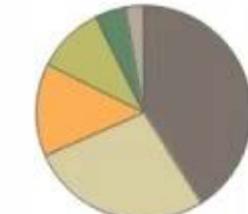


	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug
\$1.0M	\$1.1M	\$0.8M	\$0.7M	\$0.6M	\$1.7M	\$1.5M	\$1.4M	
\$0.7M	\$0.7M	\$0.5M	\$0.4M	\$0.5M	\$1.1M	\$0.8M	\$0.9M	
\$0.3M	\$0.4M	\$0.3M	\$0.2M	\$0.3M	\$0.7M	\$0.7M	\$0.5M	

Revenue Mix



2017



Debt/Equity Ratio: 4.1 %

Debt: 6P K Equity: \$1,519K



Desktop hub Executive Dash... X

Save Executive Dashboard

Region Europe SegmentDesc Bottle Shops, Co... X

Product Analysis

Product Analysis

Region

Europe	✓
Americas	
Asia	
Pacific	
Africa	

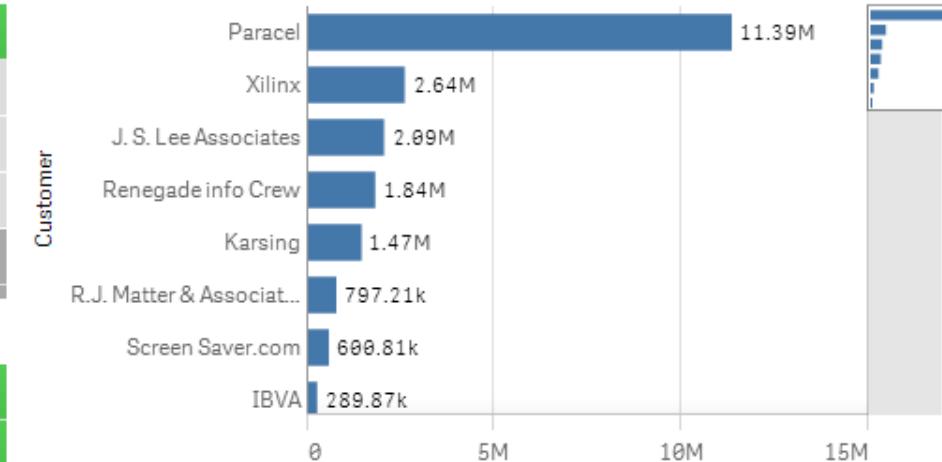
SegmentDesc

Bottle Shops	✓
Convenience Stores	✓
Restaurant & Cafes	✓

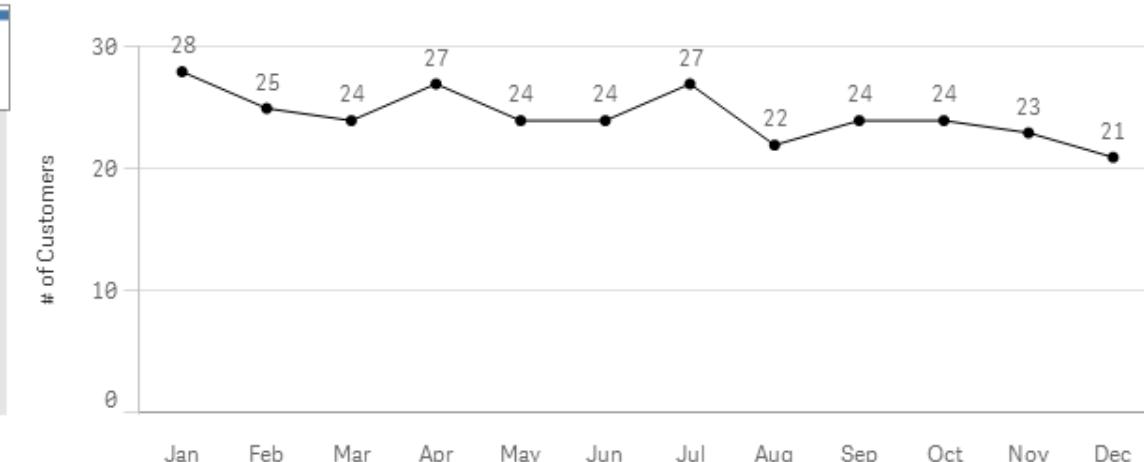
Product line

Anchovies
Bagels
Beer
Bologna
Canned Fruit

Revenue by Account



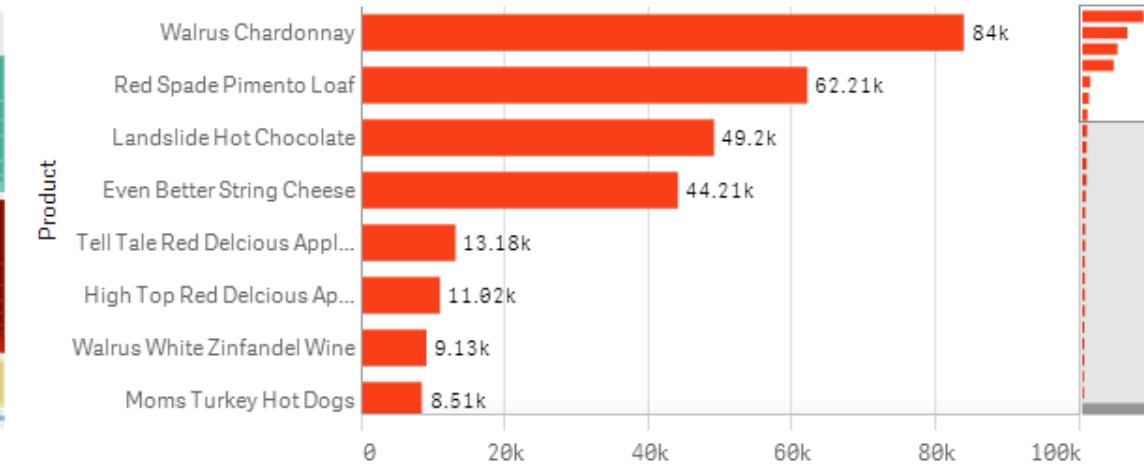
Customer count trend



Revenue by Product *



Sales volume by Product



* The data set contains negative or zero values that cannot be shown in this chart.



منابع آموزشی

udemy Categories Search for anything Udemy Business Teach on Udemy Log in Sign up

Business Analytics & Intelligence > MySQL

SQL - MySQL for Data Analytics and Business Intelligence

SQL that will get you hired – SQL for Business Analysis, Marketing, and Data Management

Bestseller 4.6 ★★★★★ (31,410 ratings) 119,955 students

Created by [365 Careers](#)

Last updated 3/2021 English English [Auto], French [Auto], [7 more](#)

[Wishlist](#)

[Share](#)

[Gift this course](#)



[Preview this course](#)

€94.99

[Add to cart](#)

[Buy now](#)



منابع آموزشی

coursera [Explore ▾](#)  [For Enterprise](#) [For Students](#) [Log In](#) [Join for Free](#)

Browse > Data Science > Data Analysis

Data Warehousing for Business Intelligence Specialization

Harness Business Data . Build a fully-optimized business data warehouse in five courses.

 4.5 2,121 ratings



Michael Mannino [+1 more instructor](#)

Offered By



University of Colorado

Boulder | Colorado Springs | Denver | Anschutz Medical Campus

Enroll for Free

Starts Sep 16

Financial aid available

46,043 already enrolled



منابع آموزشی

Udemy

Categories



Search for anything

Udemy Business

Teach on Udemy



Log in

Sign up



Business Analytics & Intelligence > Business Intelligence

Microsoft Power BI - Business Intelligence for Beginners

Beginner Course on Power BI : Business Intelligence & Data/Business Analysis with Microsoft Power BI

4.4 ★★★★★ (5,416 ratings) 24,612 students

Created by [BioticsCom King](#)

Last updated 1/2021 English English [Auto], Portuguese [Auto]

Wishlist

Share

[Gift this course](#)



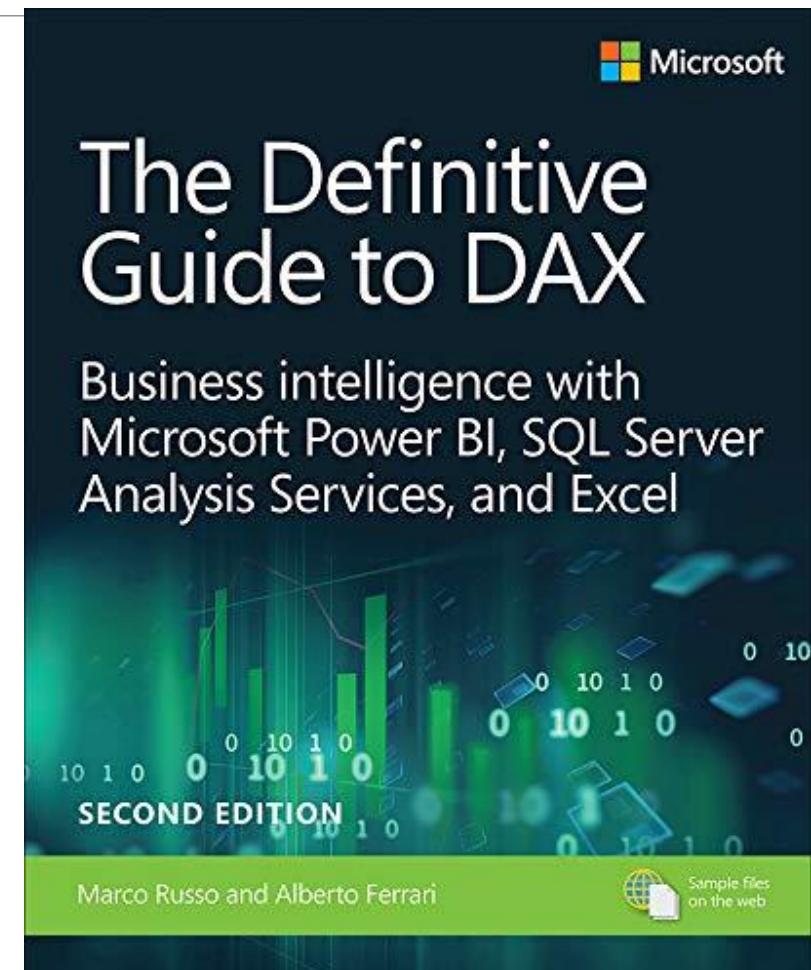
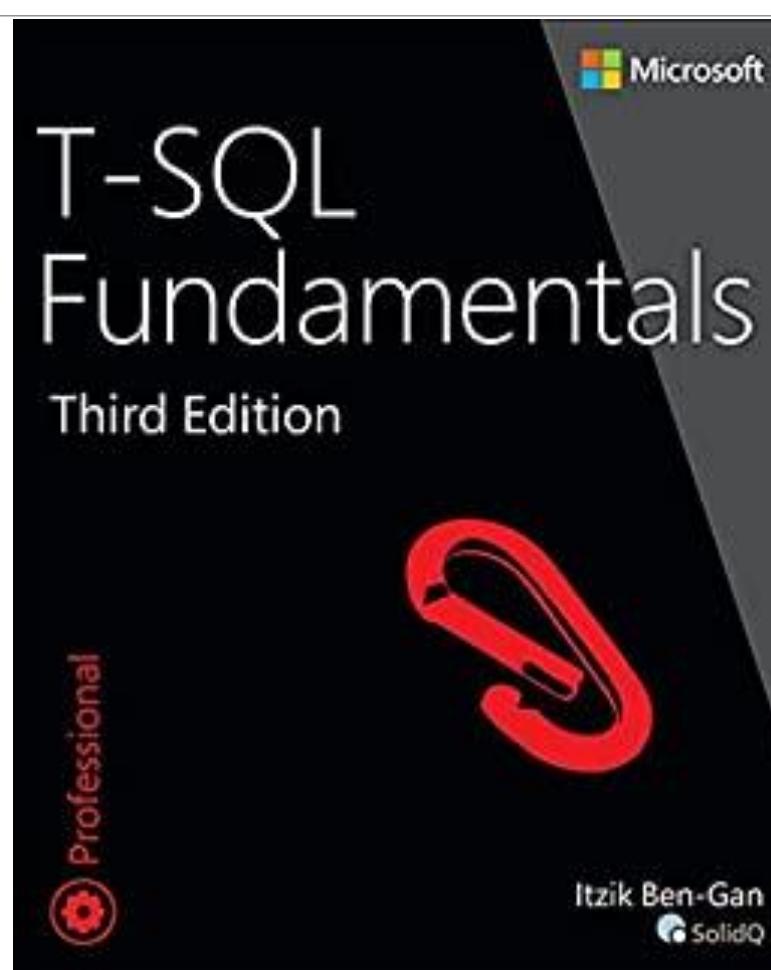
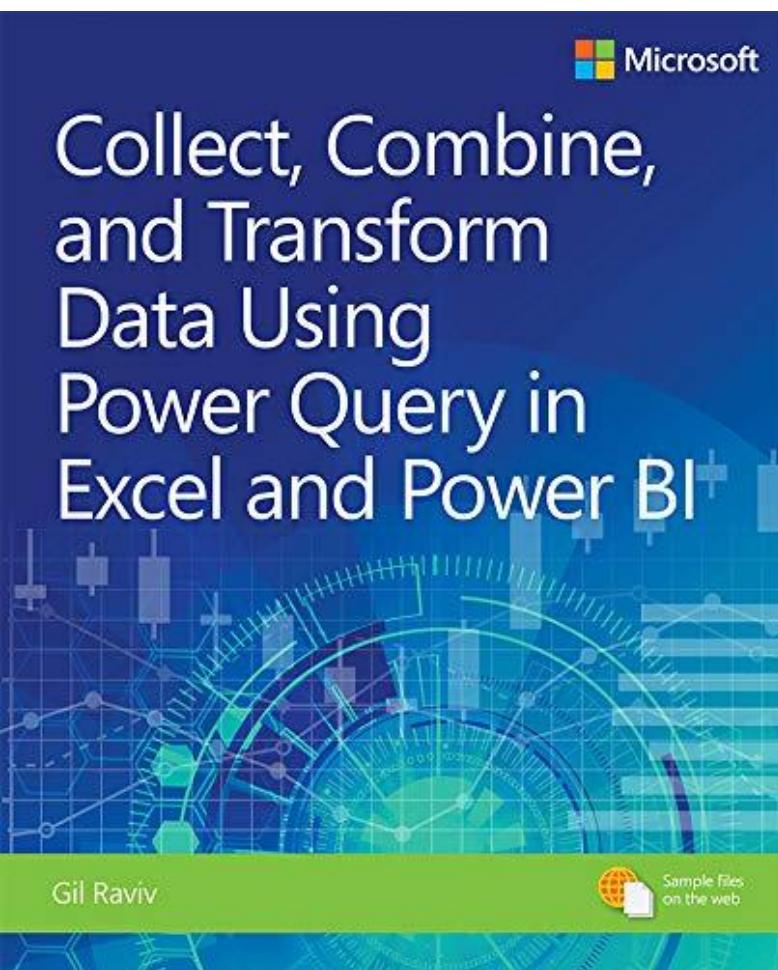
€119.99

Add to cart

Buy now



کتاب‌های مرجع





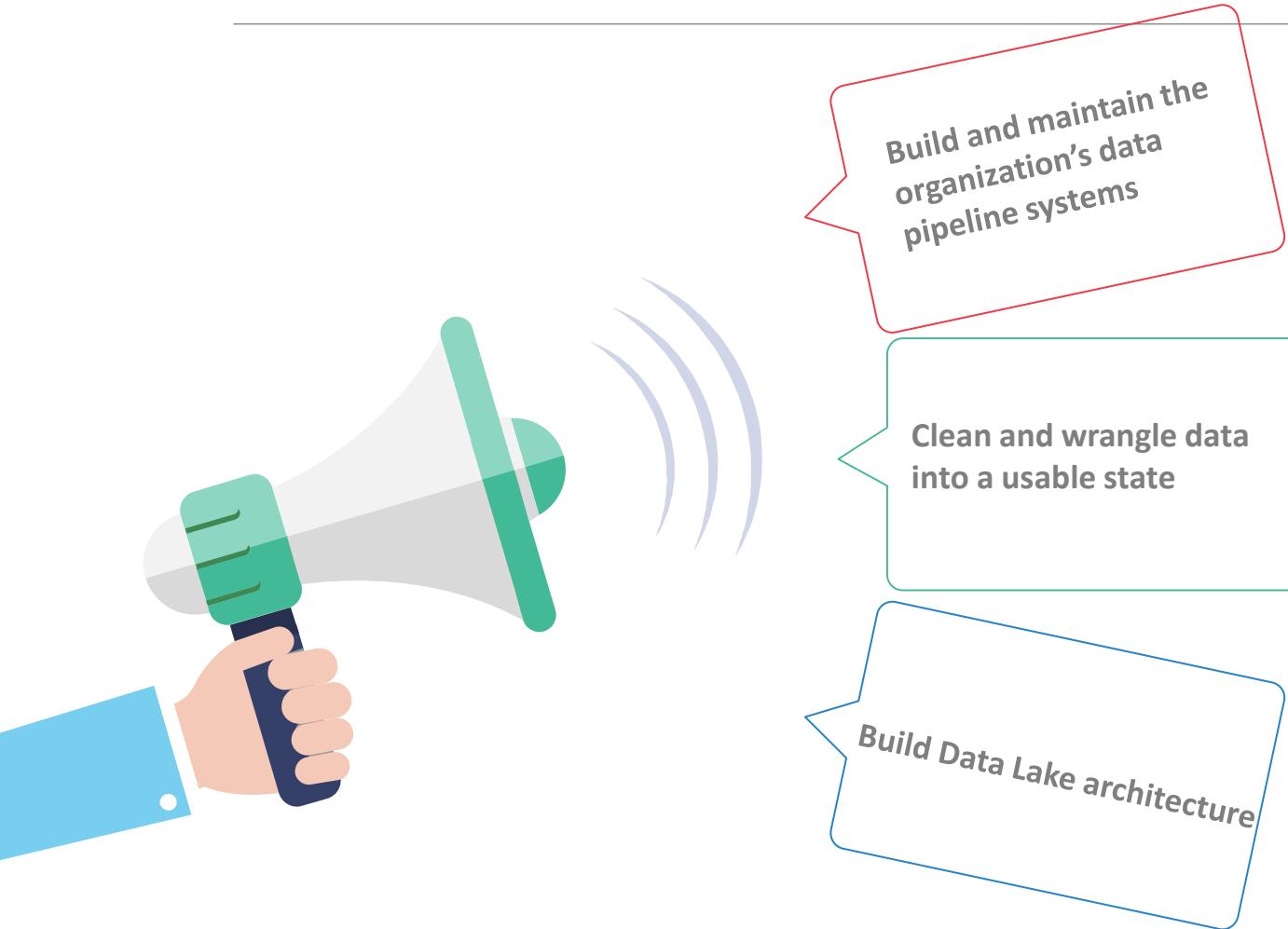
WHAT IS Data Engineering?



مهندسی داده

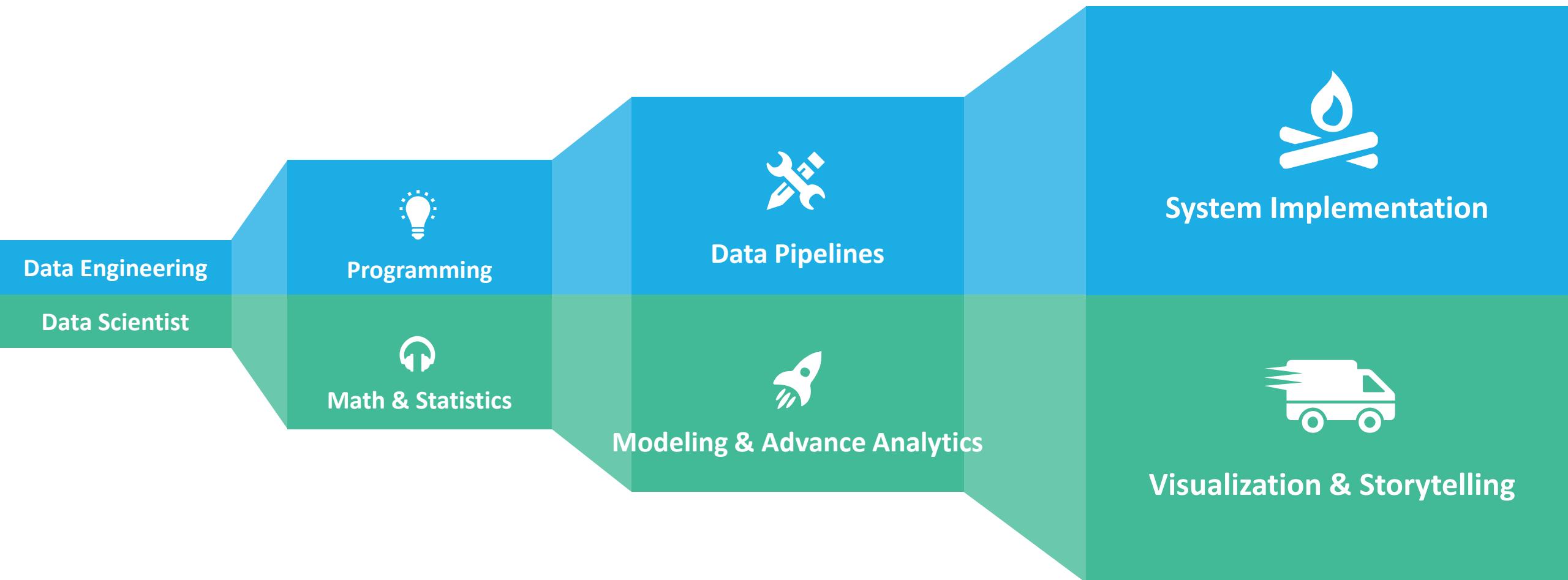
Data Engineering

مهندسی داده



Data engineers develop, maintain, test and evaluate data solutions within organizations. ... A data engineer builds large-scale data processing systems, is an expert in data warehousing solutions and should be able to work with the latest (NoSQL) database technologies.

تفاوت مهندس داده با متخصص علم داده



مهارت‌ها و ابزارهای کلیدی

ابزارهای کلیدی

❖ Linux

❖ Data Ingestion Tools (Kafka, RabbitMQ, NIFI)

❖ Databases(NoSQL & SQL)

❖ Programing (Java, Python, Scala, ...)

❖ Git

مهارت‌ها

➤ پایگاه داده

➤ پیش پردازش و پاکسازی داده

➤ راه اندازی دریاچه داده

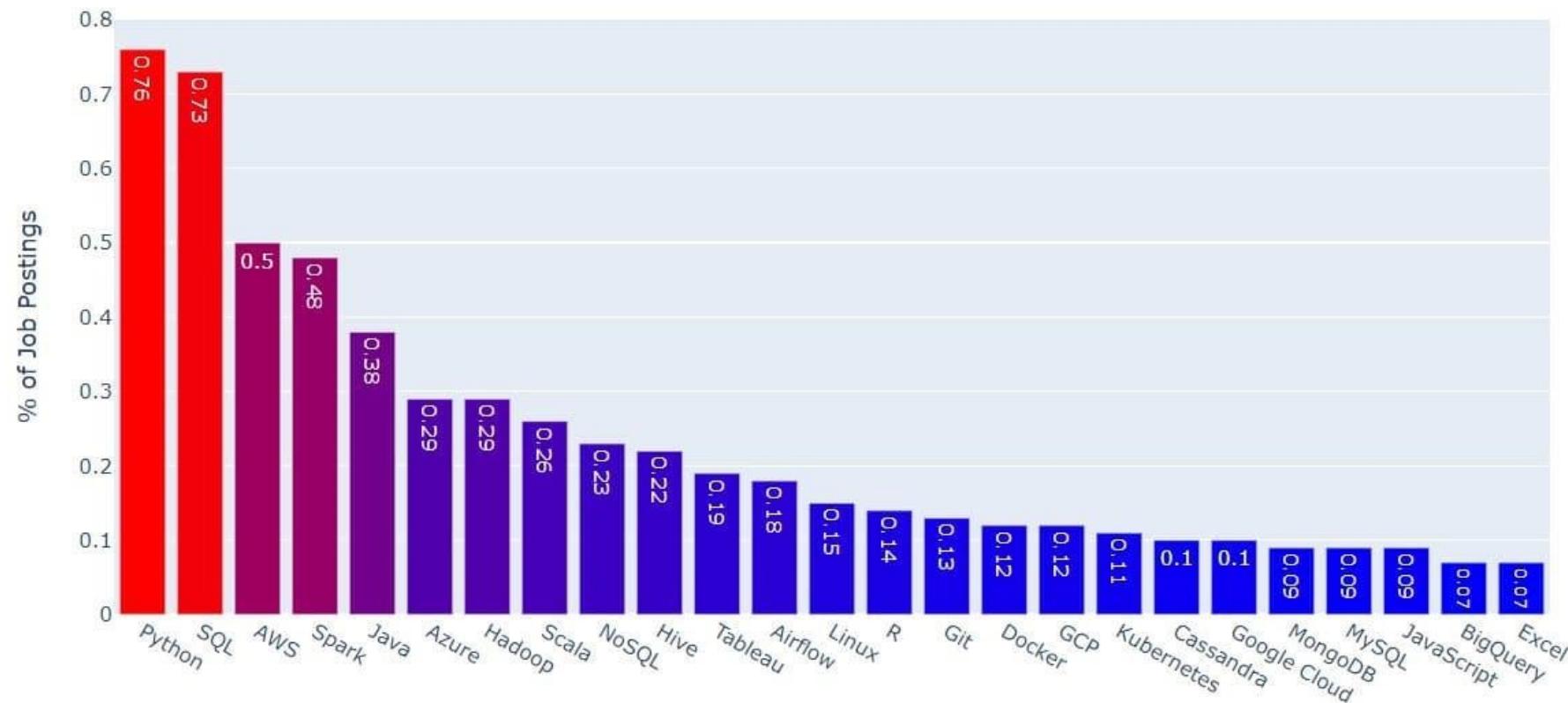
➤ نگهداری و پشتیبانی سیستم‌های نرم افزاری

➤ بهینه‌سازی سیستم

➤ توسعه چابک

مهارت‌ها و ابزارهای کلیدی

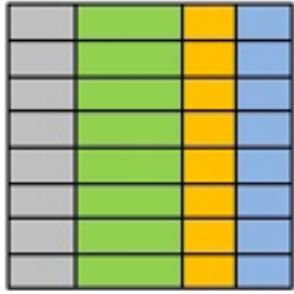
25 Most In-Demand Data Engineering Skills



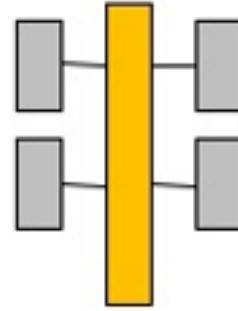


پایگاه داده غیر رابطه ای (NOSQL)

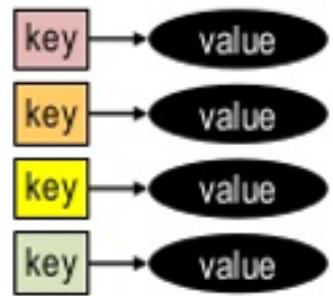
Relational



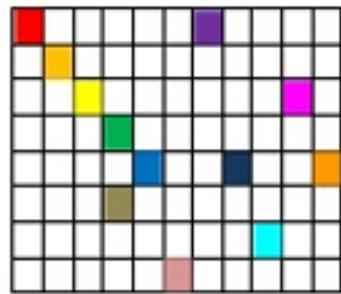
Analytical (OLAP)



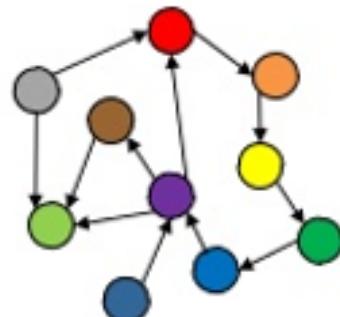
Key-Value



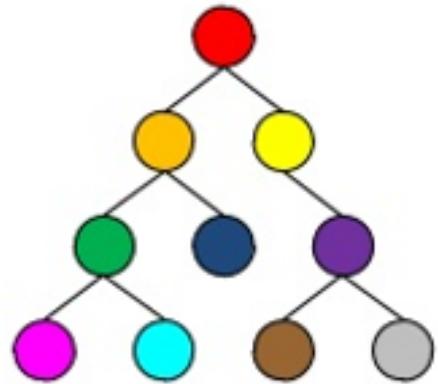
Column-Family



Graph



Document



Types of Databases



دوره‌های آموزشی

Browse > Information Technology > Data Management

IBM Data Engineering Professional Certificate

Launch your new career in Data Engineering. Master SQL, RDBMS, ETL, Data Warehousing, NoSQL, Big Data and Spark with hands-on job-ready skills.

★★★★★ 4.7 257 ratings



Rav Ahuja [+14 more instructors](#)

Enroll for Free

Starts Sep 16

Financial aid available

3,883 already enrolled

Offered By



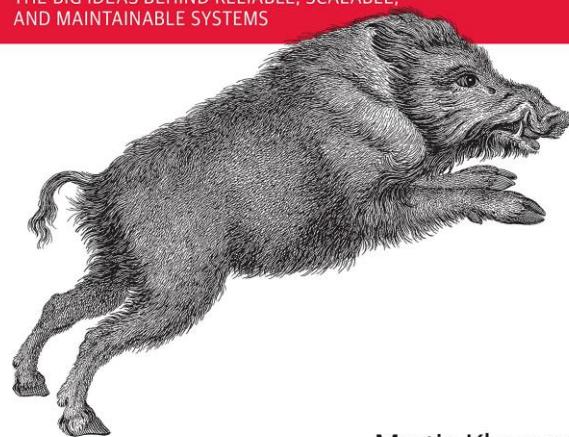


کتاب‌های مرجع

O'REILLY®

Designing Data-Intensive Applications

THE BIG IDEAS BEHIND RELIABLE, SCALABLE,
AND MAINTAINABLE SYSTEMS



Martin Kleppmann

O'REILLY®

Architecting Modern Data Platforms

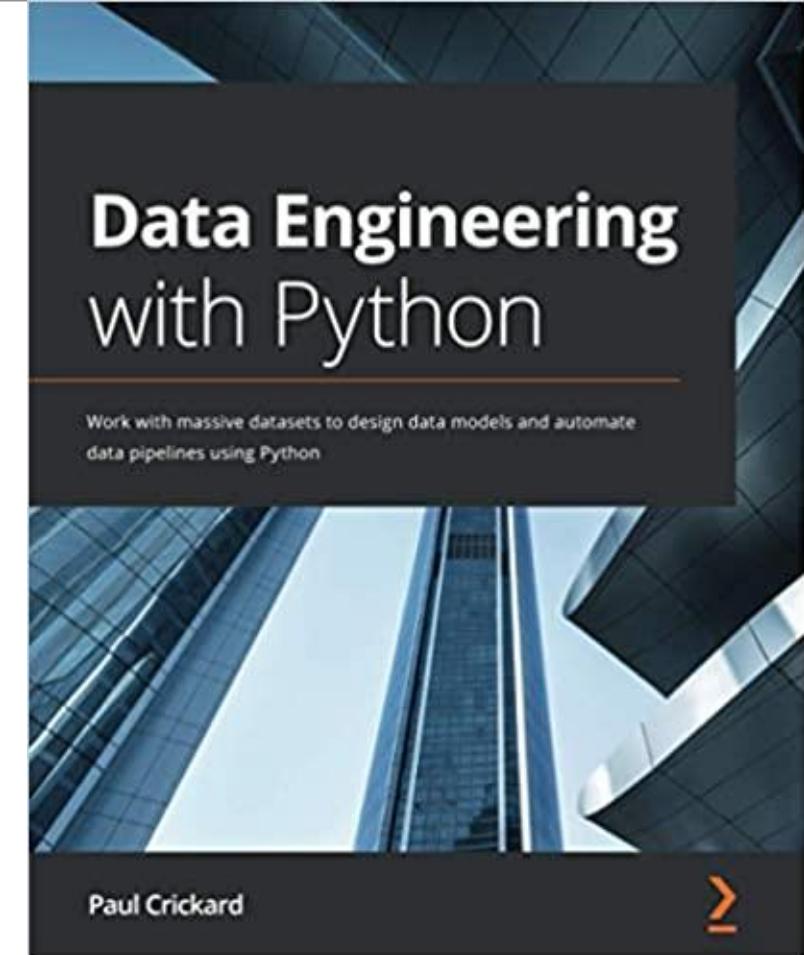
A GUIDE TO ENTERPRISE HADOOP AT SCALE

Jan Kunigk, Ian Buss,
Paul Wilkinson & Lars George



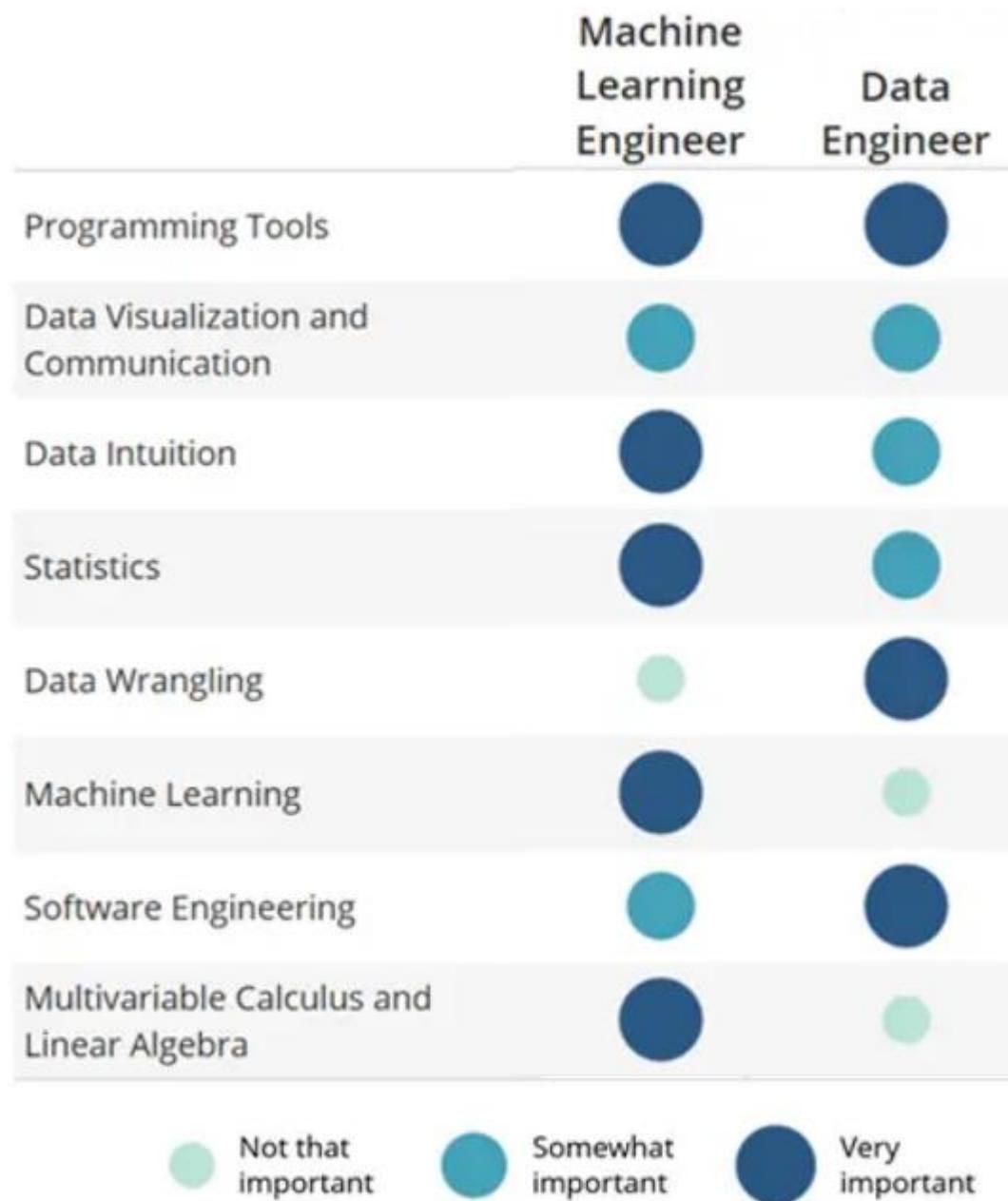
Data Engineering with Python

Work with massive datasets to design data models and automate
data pipelines using Python



Paul Crickard





مهندسی یادگیری ماشین



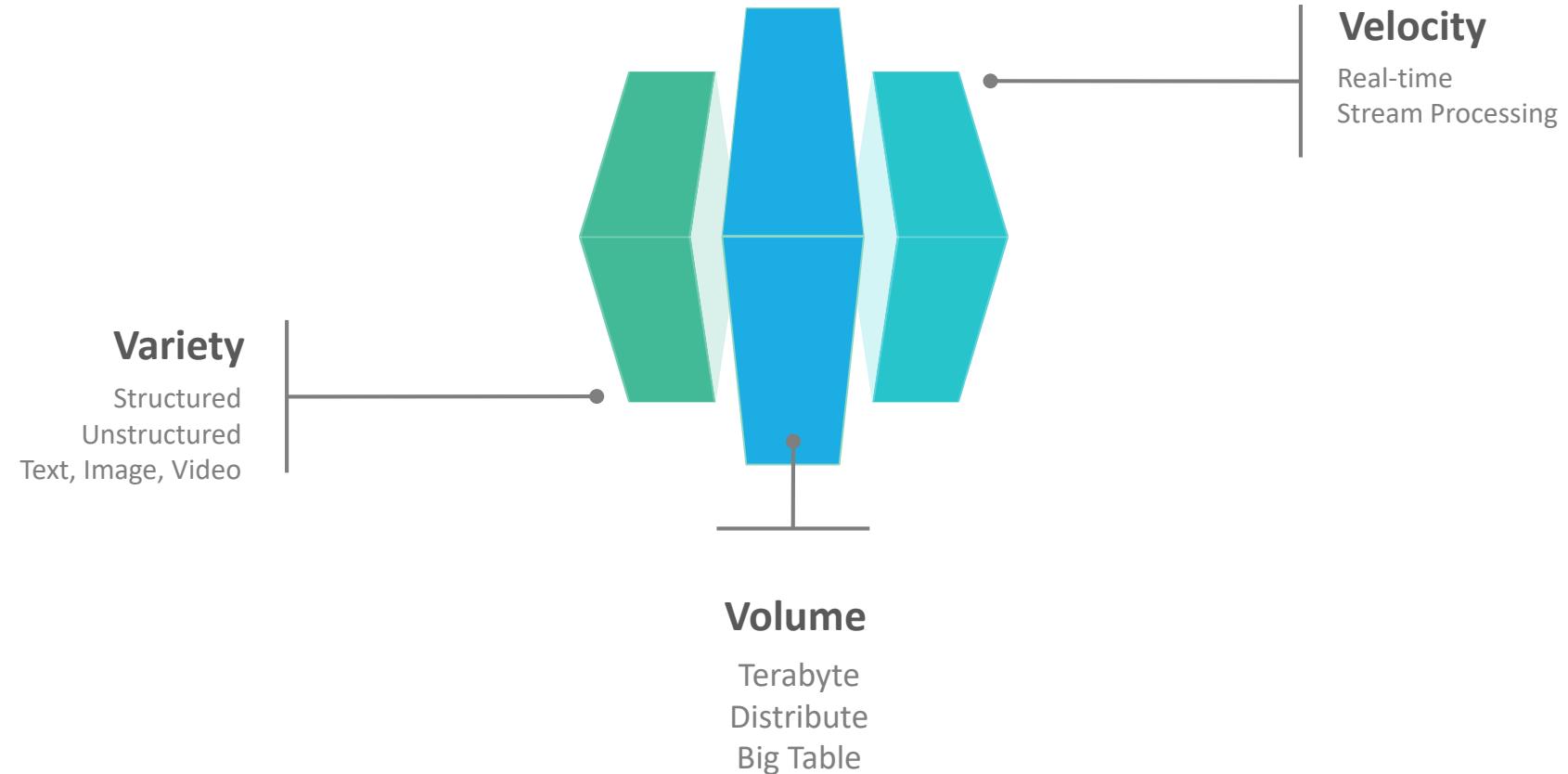
متخصص کلان داده

Big Data

کلان داده

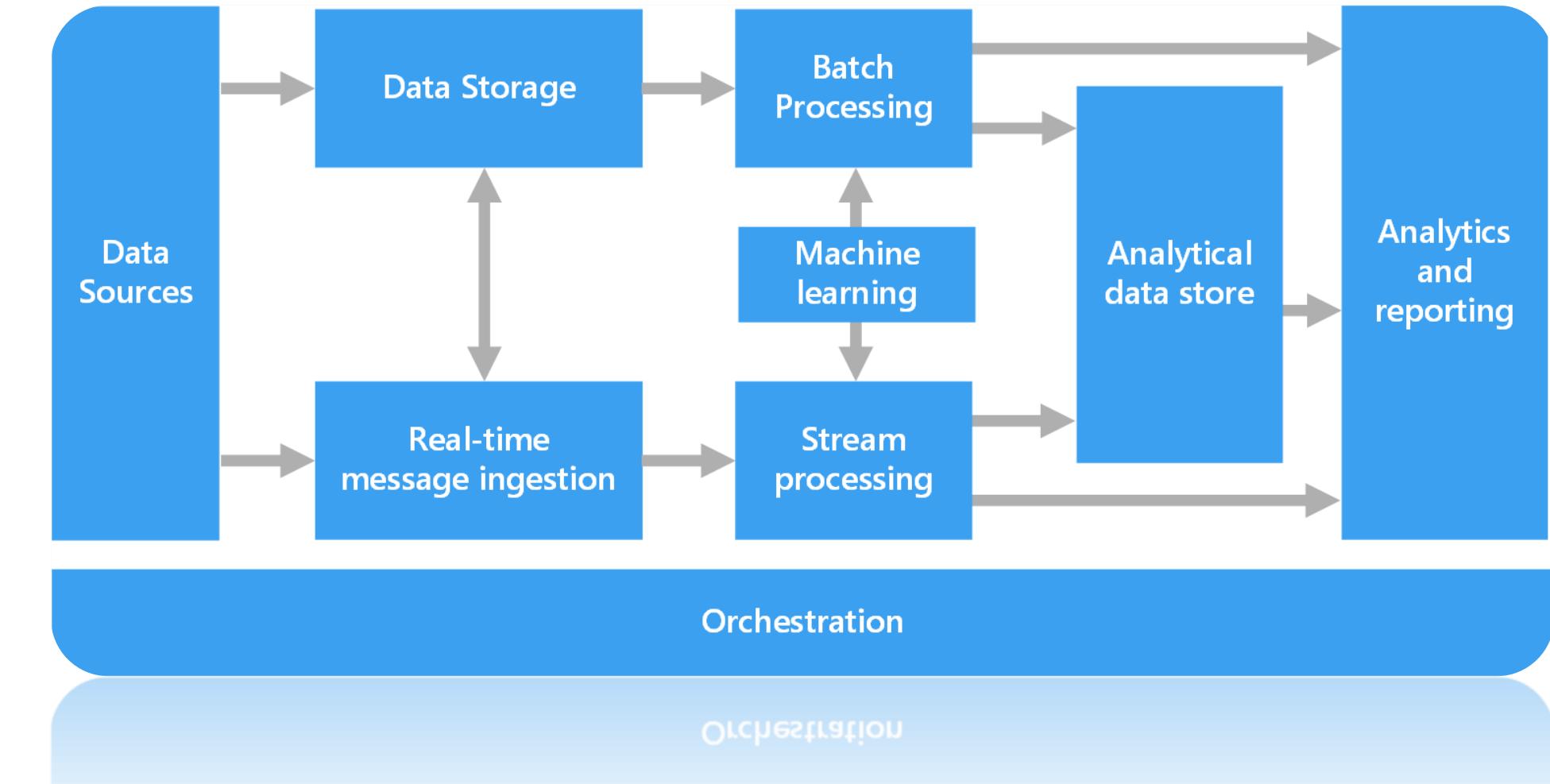
Big data is a term used to refer to data sets that are too large or complex for traditional data-processing application software to adequately deal with.

It's what organizations do with the data that matters. Big data can be analyzed for insights that lead to better decisions and strategic business moves.





مَحَارِي لَامِبِدَا



مهارت‌ها و ابزارهای کلیدی

ابزارهای کلیدی

❖ Apache Hadoop

❖ Apache Spark

❖ NoSQL

❖ Programming (Java, Python, Scala, ...)

❖ Processing Frameworks (MapReduce, Spark, Flink, ...)

❖ Linux

مهارت‌ها

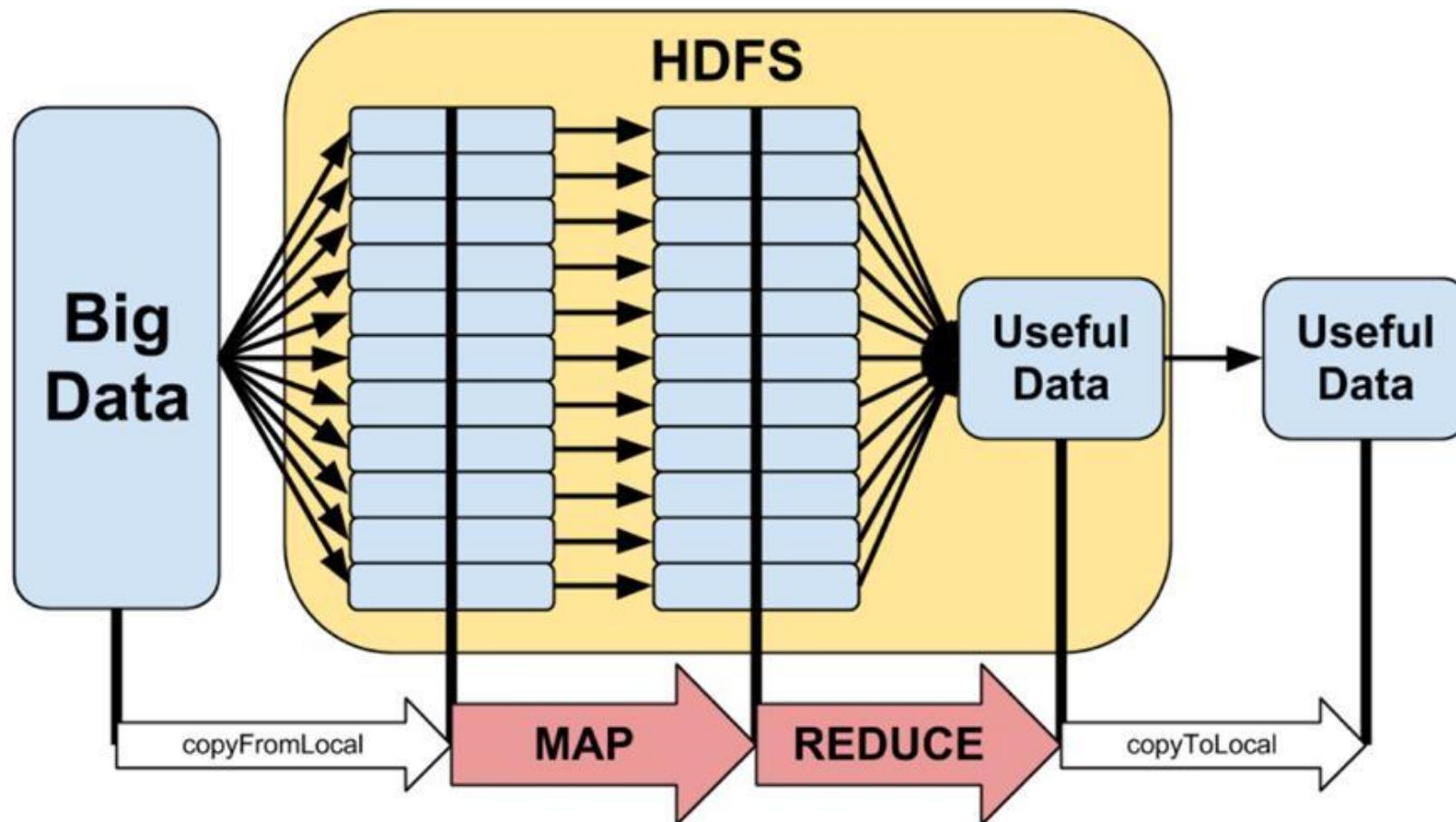
➤ توسعه سیستم مدیریت کلان داده

➤ مدیریت دادگان دسته‌ای و جریانی

➤ پردازش بلاذرنگ کلان داده

➤ پشتیبانی سیستمه

➤ تحلیل کلان داده



اکوسیستم ہڈوپ



پلتفرم پردازشی اسپارک

Streaming

MLlib

For Machine Learning

GraphX

For Graph Computing

**Spark SQL &
DataFrames**

Spark Core API

R

Python

Scala

SQL

Java



دوره‌های آموزشی

Browse > Data Science > Data Analysis

Offered By

UC San Diego

Big Data Specialization

Unlock Value in Massive Datasets. Learn fundamental big data methods in six straightforward courses.

★★★★★ 4.5 8,929 ratings



Ilkay Altintas [+2 more instructors](#)

Enroll for Free
Starts Sep 16

Financial aid available

106,694 already enrolled



دوره‌های آموزشی

udemy Categories Search for anything Udemy Business Teach on Udemy

Business Analytics & Intelligence > Big Data

The Ultimate Hands-On Hadoop: Tame your Big Data!

Hadoop tutorial with MapReduce, HDFS, Spark, Flink, Hive, HBase, MongoDB, Cassandra, Kafka + more! Over 25 technologies.

Bestseller 4.6 ★★★★★ (25,022 ratings) 138,614 students

Created by Sundog Education by Frank Kane, Frank Kane, Sundog Education Team

Last updated 7/2021 English English, French [Auto], 4 more

[Wishlist](#) [Share](#) [Gift this course](#)



Preview this course

€109.99

[Add to cart](#)

[Buy now](#)

30-Day Money-Back Guarantee

What you'll learn

- ✓ Design distributed systems that manage "big"
- ✓ Use HDFS and MapReduce for storing and

This course includes:

- ▶ 14.5 hours on-demand video



دوره‌های آموزشی

udemy Categories Search for anything Udemy Business Teach on Udemy

Software Engineering > Python

Spark and Python for Big Data with PySpark

Learn how to use Spark with Python, including Spark Streaming, Machine Learning, Spark 2.0 DataFrames and more!

Bestseller 4.6 ★★★★★ (15,981 ratings) 82,377 students

Created by [Jose Portilla](#)

Last updated 5/2020 English English [Auto], French [Auto], 8 more

[Wishlist](#) [Share](#) [Gift this course](#)



€109.99

[Add to cart](#)

[Buy now](#)

What you'll learn

- ✓ Use Python and Spark together to analyze Big Data
- ✓ Learn how to use the new Spark 2.0 DataFrame Syntax

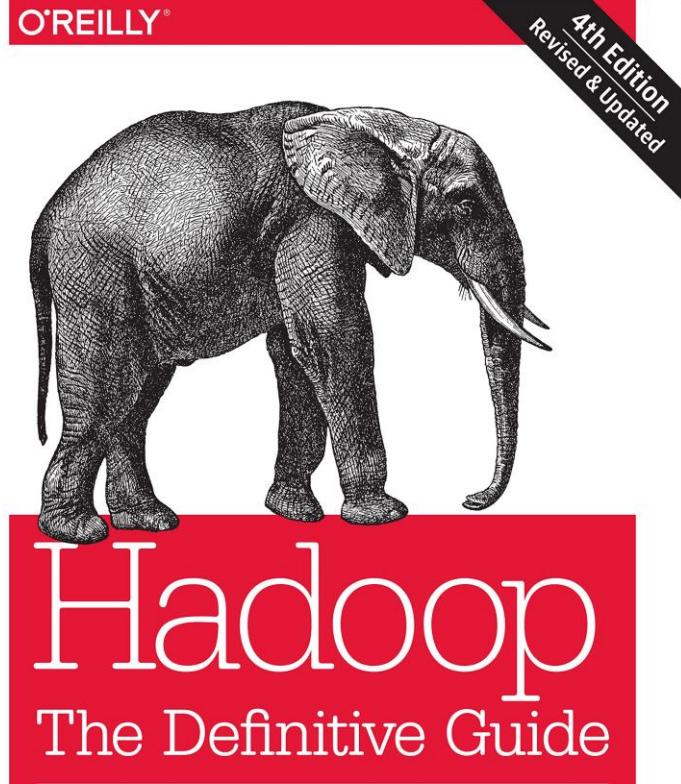
30-Day Money-Back Guarantee

This course includes:

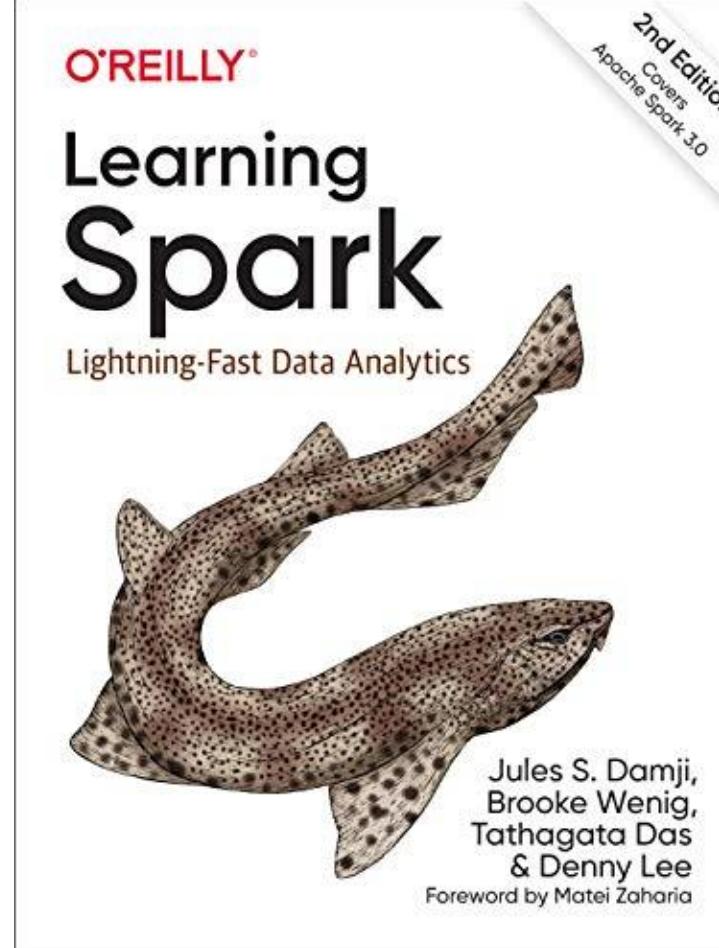
- ⌚ 10.5 hours on-demand video



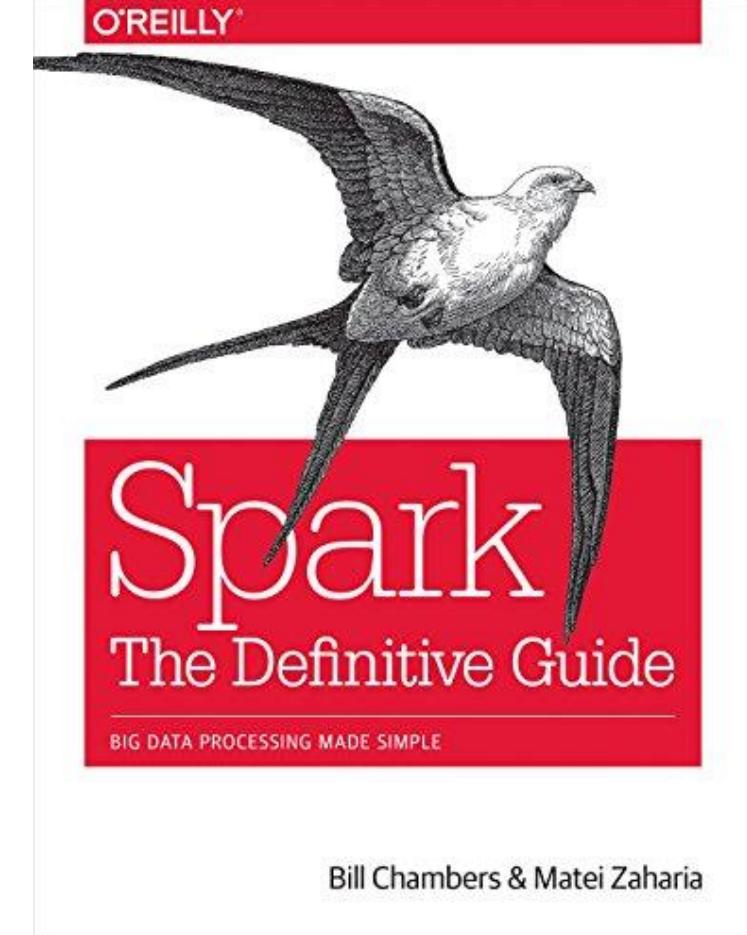
کتاب‌های مرجع



Tom White



Jules S. Damji,
Brooke Wenig,
Tathagata Das
& Denny Lee
Foreword by Matei Zaharia



Bill Chambers & Matei Zaharia

THE

END