

Hendra H Choiri

A data scientist with 7+ years of experience in the data world. Passionate in the deep learning and computer vision field, and loves to share about data science.



Hendra H Choiri
Data Scientist

Education Background



Teknik Informatika



Data Science & Technology

Introduction to Data Science

(Data Scientist: What & Why?)



Data Scientist: What & Why?

- Apa itu data science?
- Apa yang dilakukan data scientist?
- Jenis-jenis analisis
- Mengapa jadi data scientist?
- Skill yang dibutuhkan data scientist
- Pekerjaan terkait data
- Business acumen
- Profitability framework

Data Scientist: What & Why?



Apa itu data science?



Apa yang dilakukan data scientist?



Jenis-jenis analisis



Mengapa jadi data scientist?



Skill yang dibutuhkan data scientist



Pekerjaan terkait data

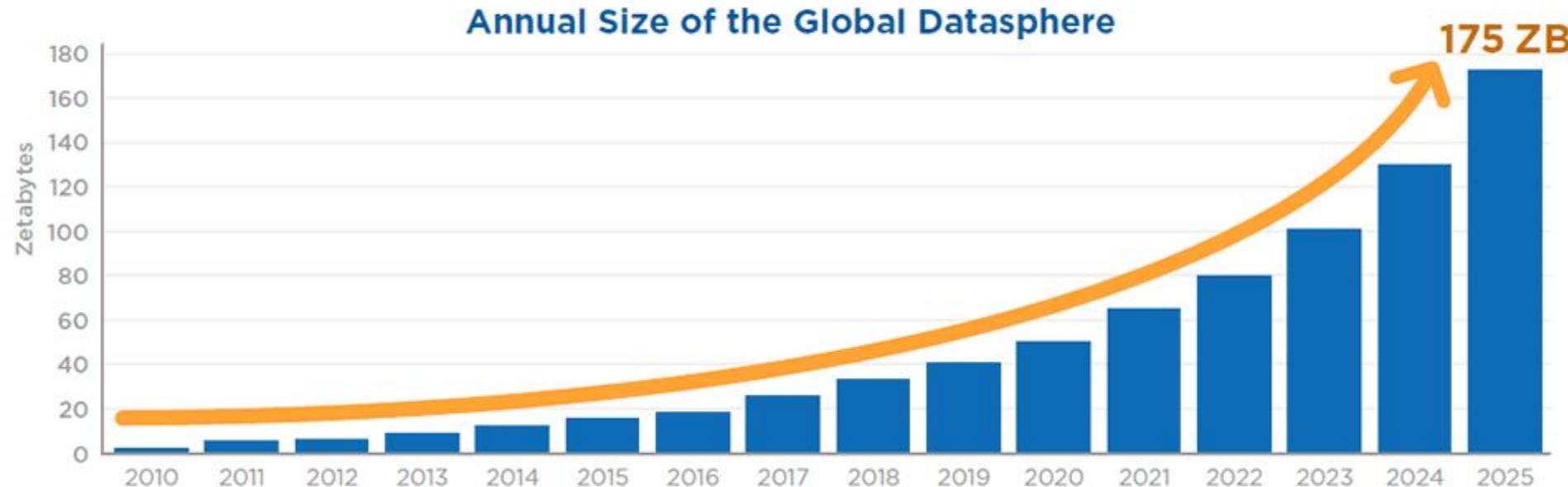
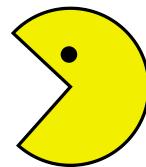


Business acumen



Profitability framework

Data Eat the World



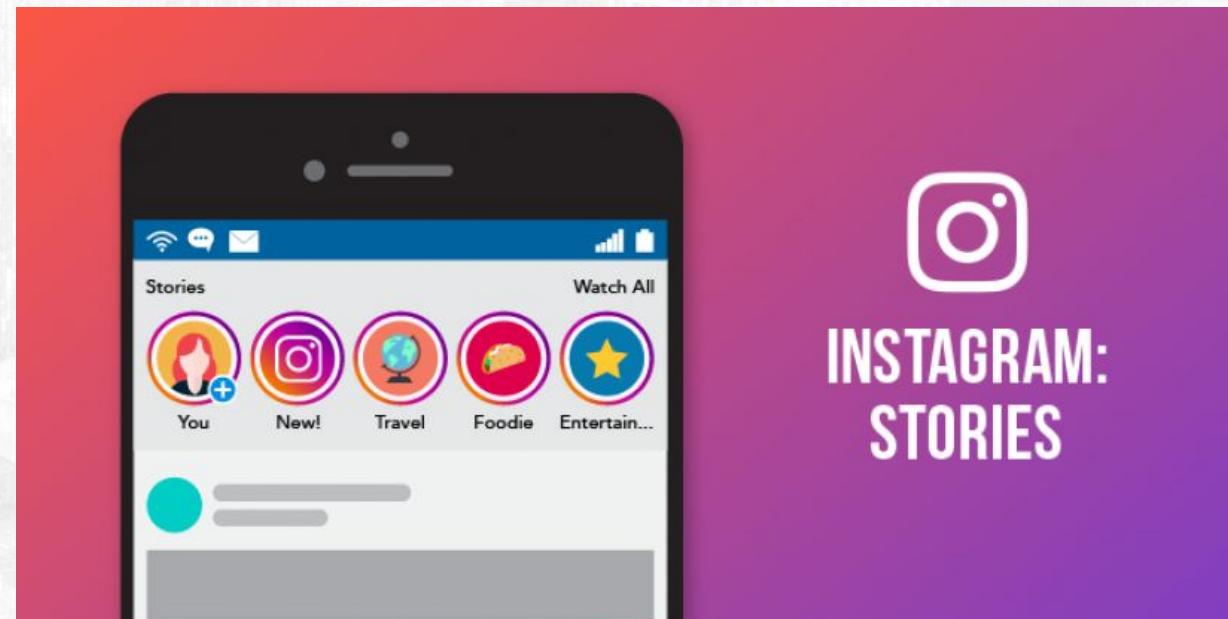
Sumber: Seagate & IDC

Diprediksi akan ada

**175 Triliun GB Data
di tahun 2025**

Tebak Tebakan...

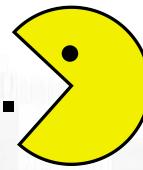
Dalam 1 menit, ada berapa IG Story yang di-share?



2021 This Is What Happens In An Internet Minute

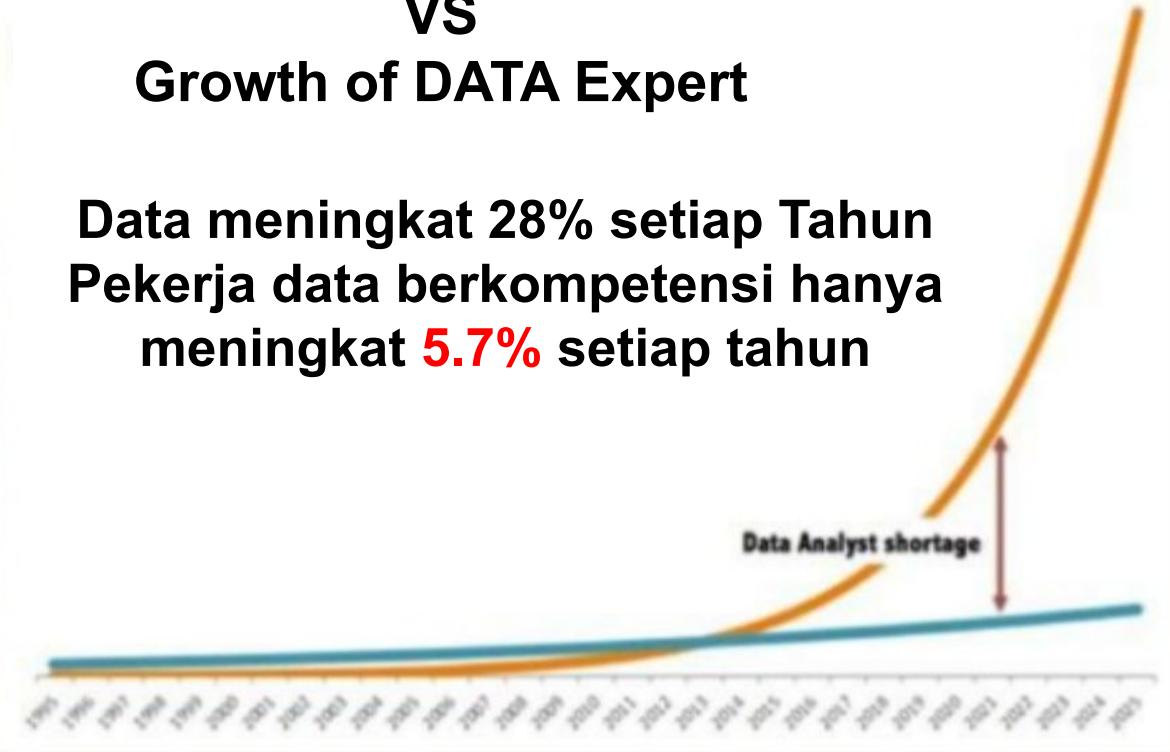


Tantangan selanjutnya ...



Growth of DATA VS Growth of DATA Expert

Data meningkat 28% setiap Tahun
Pekerja data berkompetensi hanya
meningkat **5.7%** setiap tahun



Source:
<https://www.slideshare.net/ryanorban/how-to-become-a-data-scientist>

Di tahun 2025, diprediksikan ...

Growing job demand:

1. Data Analysts and Scientists
2. AI and Machine Learning Specialists
3. Big Data Specialists
4. Digital Marketing and Strategy Specialists
5. Process Automation Specialists
6. Business Development Professionals
7. Digital Transformation Specialists
8. Information Security Analysts
9. Software and Applications Developers
10. Internet of Things Specialists

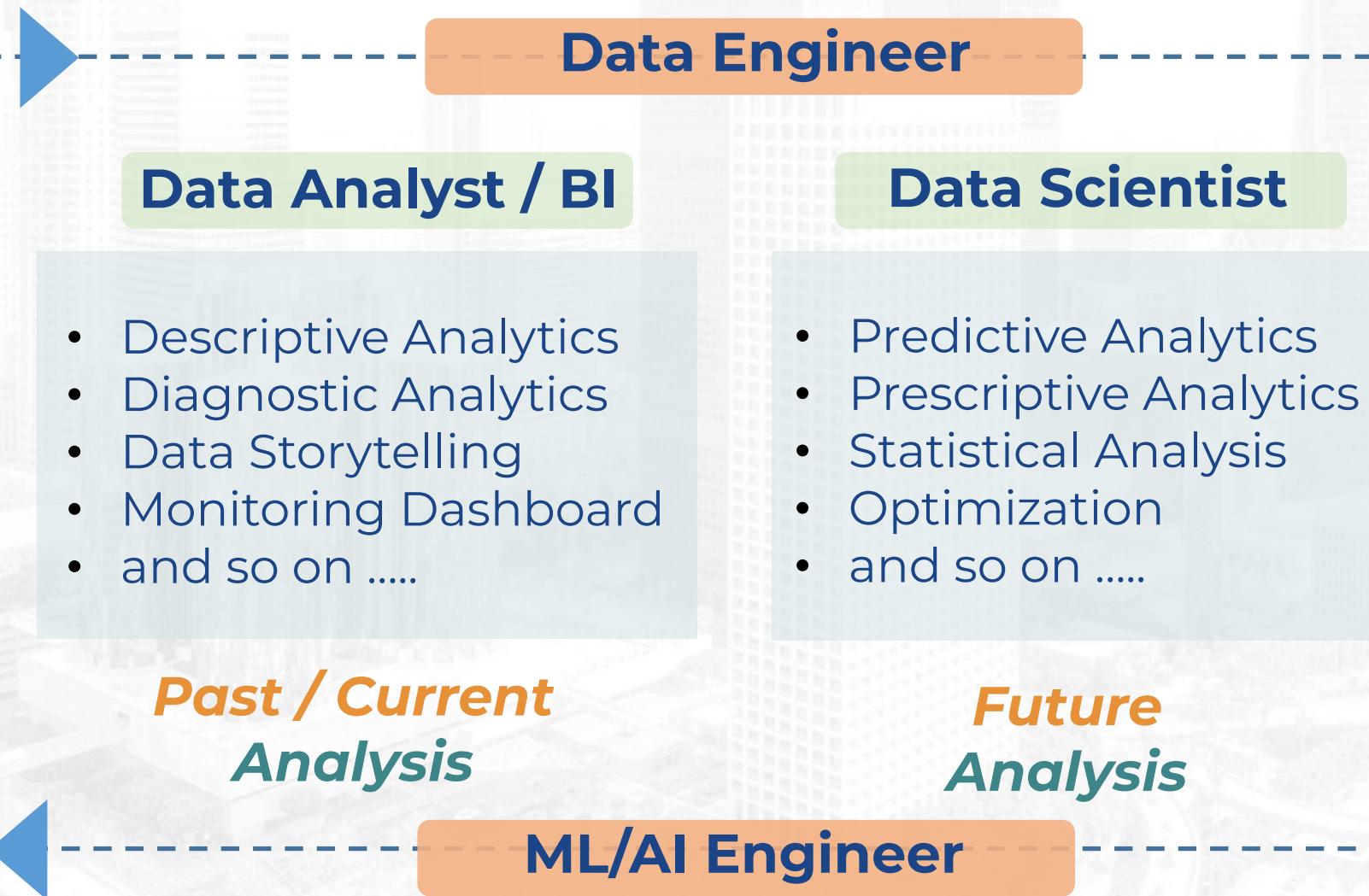
85% harus memiliki skill dan pengetahuan pada segmen digital dan data

Decreasing job demand:

1. Data Entry Clerks
2. Administrative and Executive Secretaries
3. Accounting, Bookkeeping and Payroll Clerks
4. Accountants and Auditors
5. Assembly and Factory Workers
6. Business Services and Administration Managers
7. Client Information and Customer Service Workers
8. General and Operations Managers
9. Mechanics and Machinery Repairers
10. Material-Recording and Stock-Keeping Clerks

80% merupakan pekerjaan yang bersifat *repetitive* dan *daily business as usual*

Pekerjaan Teknikal Terkait Data



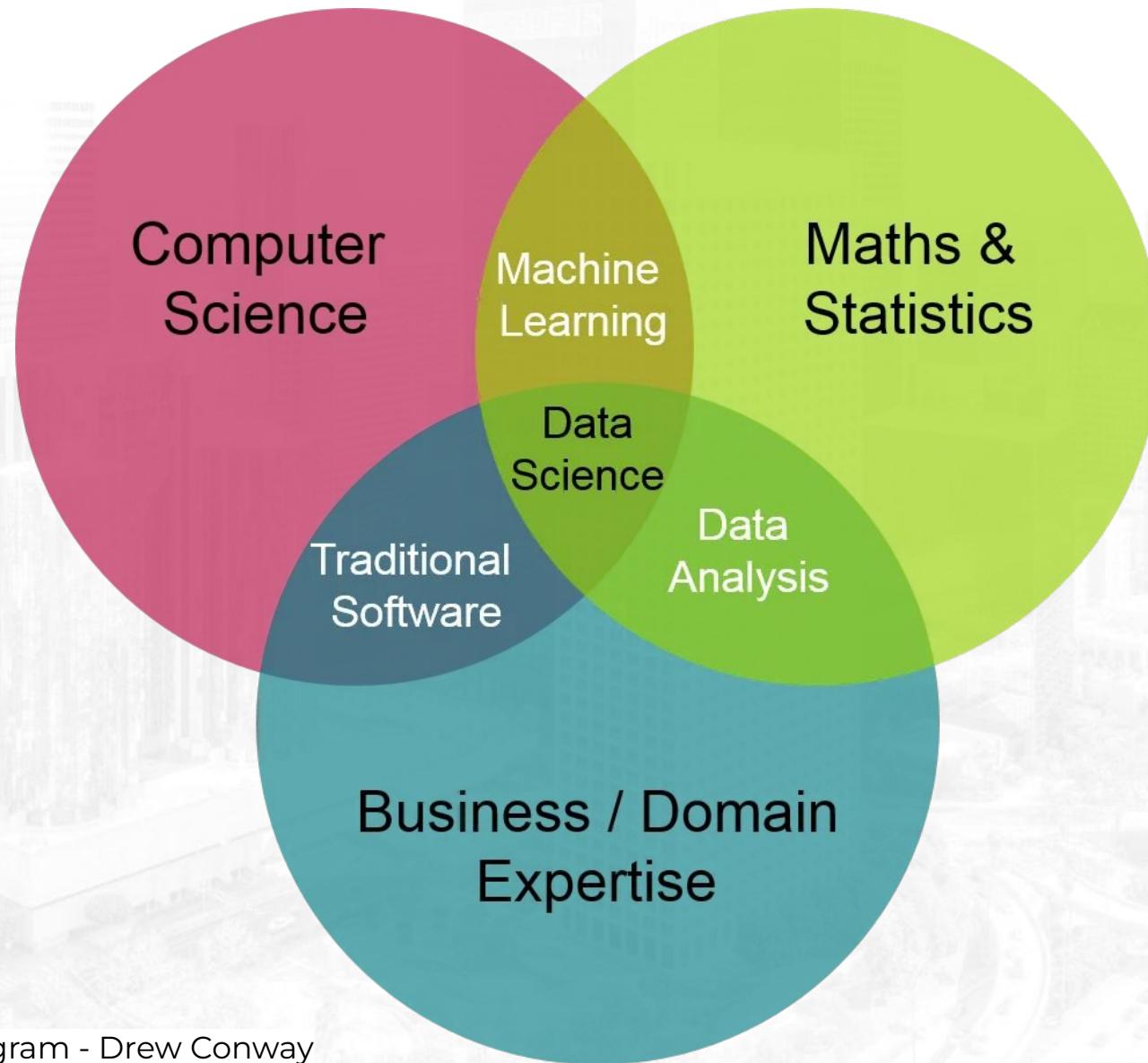
Data Science:

Ilmu yang khusus mempelajari **DATA**, khususnya data kuantitatif (numerik)

Data Scientist:

Profesi yang bertugas membuat solusi dari sebuah permasalahan menggunakan **DATA**

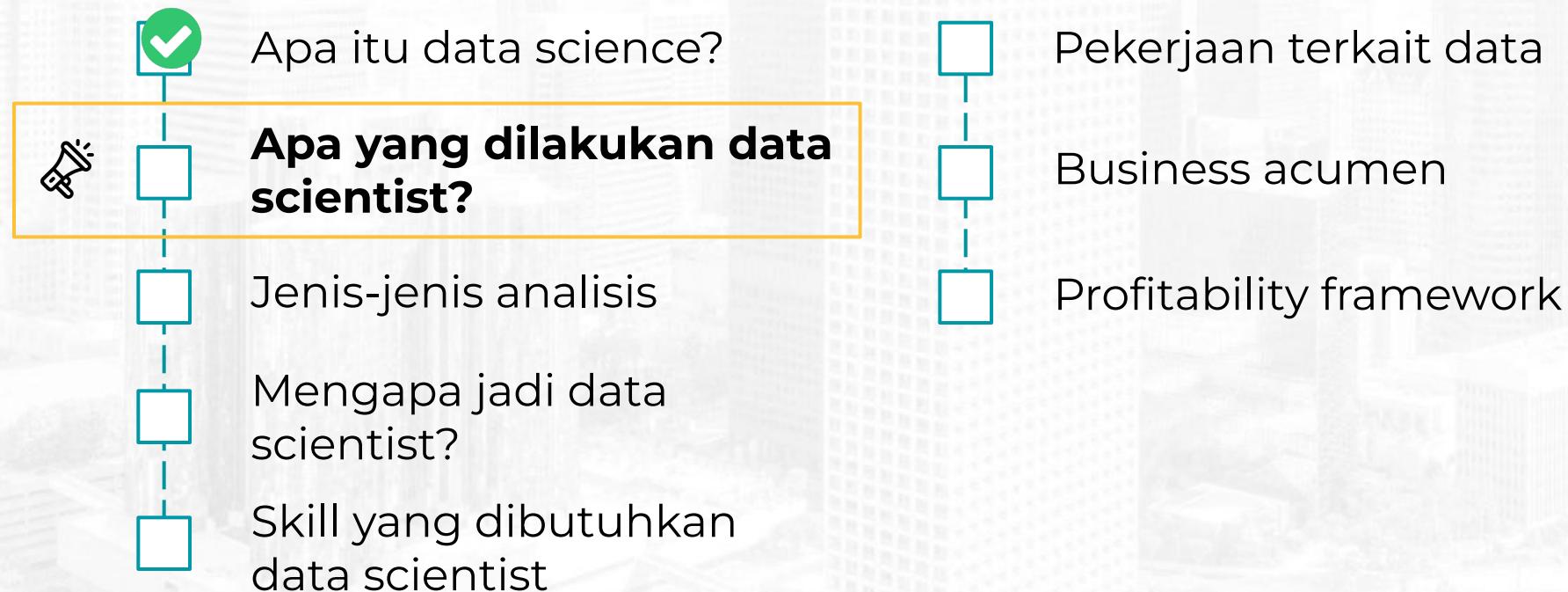
Data Science = Programming + Statistics + Business Domain



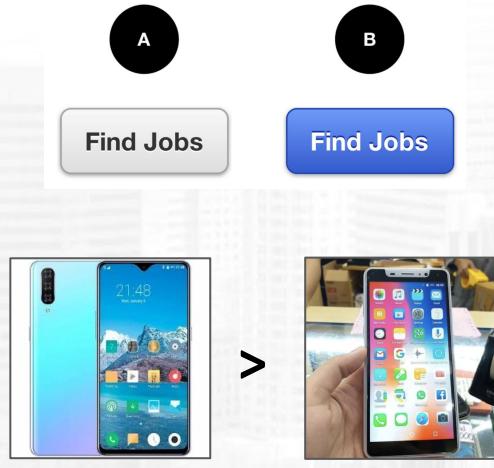
"A *data scientist* is someone who is better at *statistics* than any *programmer* and better at *programming* than any *statistician*"

Outline Pembelajaran

Data Scientist: What & Why?



Contoh Hasil Pekerjaan Data Scientist: Insights



“Di page website ini, **tombol jenis B 23% berpotensi lebih banyak diklik** dibanding tombol jenis A”

A/B Test



“Di e-commerce kita, kategori **otomotif dan elektronik** lebih banyak dibeli user **laki-laki**. Sedangkan kategori **kecantikan dan produk bayi** lebih banyak dibeli user **perempuan**”

Observasi

Contoh Hasil Pekerjaan Data Scientist: Machine Learning Models (Basic)

Menentukan suatu permohonan pinjaman modal diterima/ditolak

Classification

Mendeteksi apakah suatu transaksi adalah fraud

Memberikan rekomendasi jurusan kuliah berdasarkan nilai rapor

Memprediksi harga rumah berdasarkan spesifikasi & lokasi

Regression

Memprediksi jumlah transaksi suatu kategori produk di beberapa bulan ke depan

Forecasting

Memprediksi trend harian jumlah kasus positif COVID19

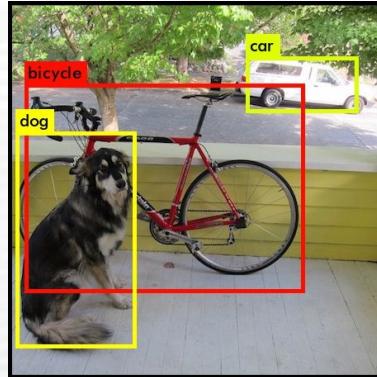
Segmentasi users menjadi beberapa grup yang masing-masing punya karakteristik tertentu

Clustering

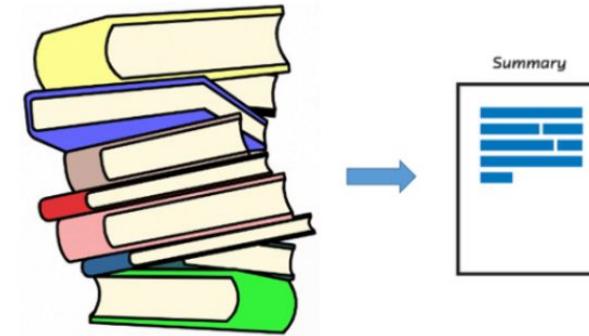
Menentukan titik-titik penting penjemputan penumpang ojol

Contoh Hasil Pekerjaan Data Scientist: Machine Learning Models (Advanced)

object detection



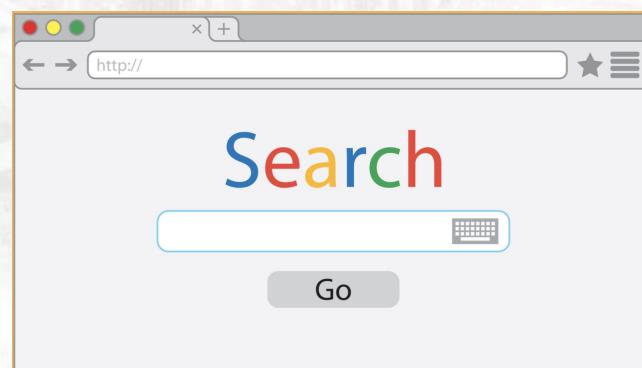
text summarization



route optimizer



search engine

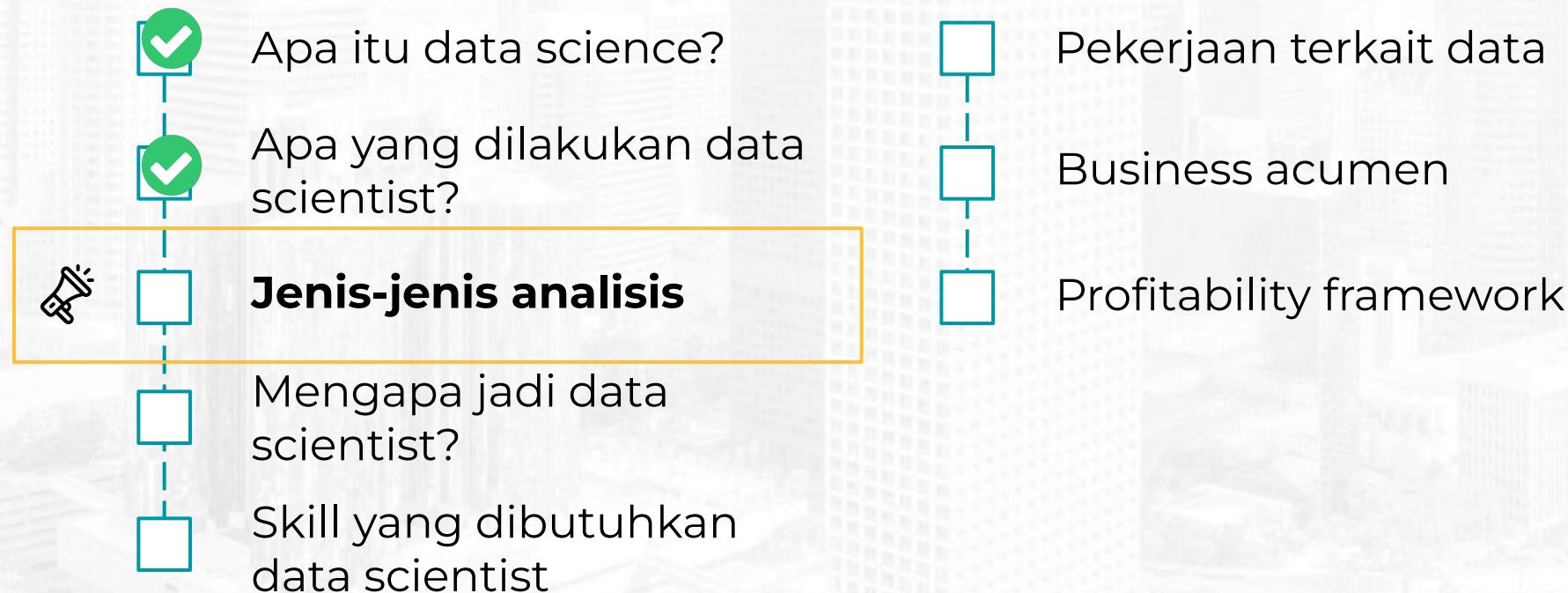


recommender system

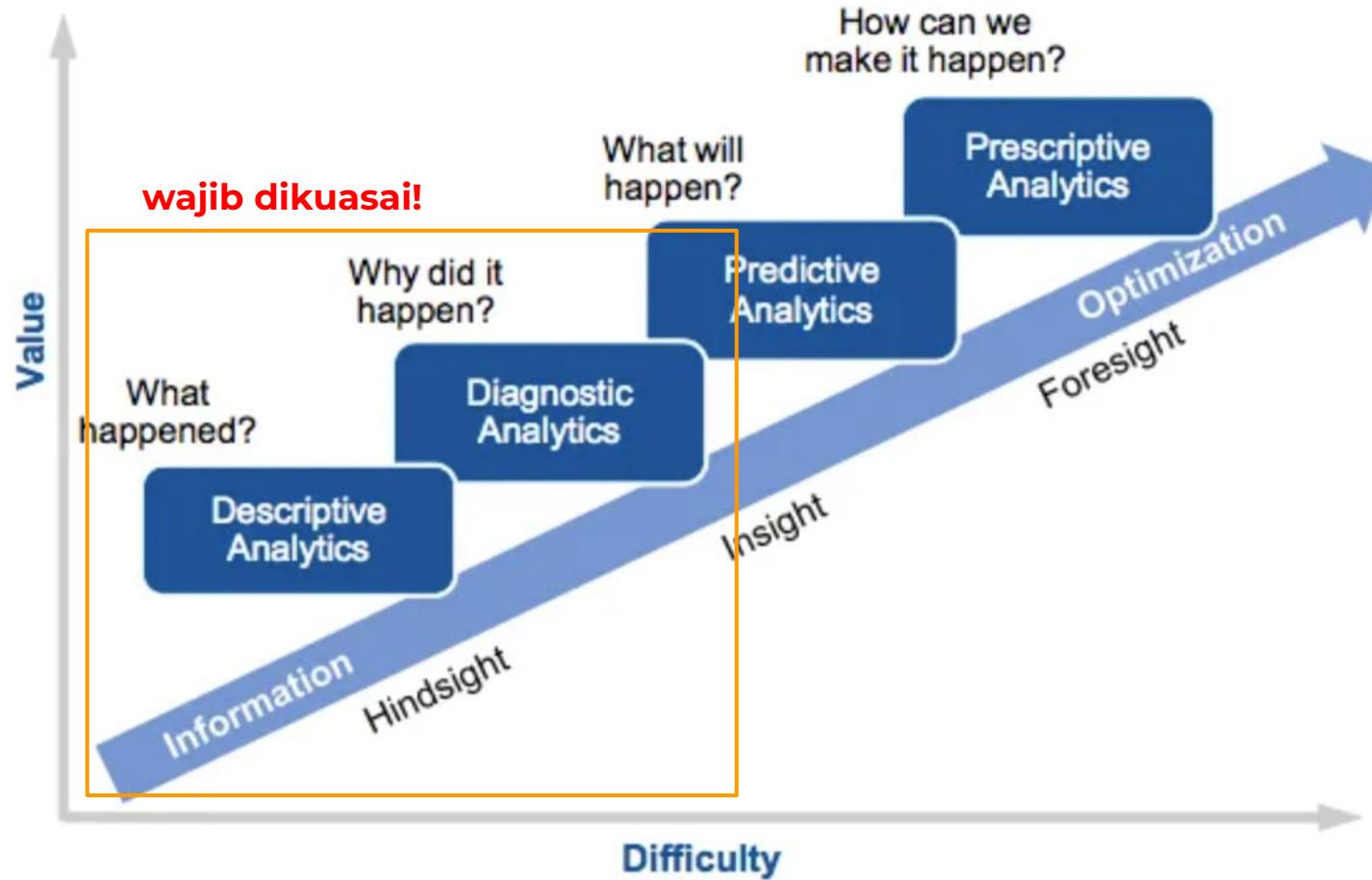


Outline Pembelajaran

Data Scientist: What & Why?



Gartner Analytic Ascendancy Model



Contoh Skenario ...

Descriptive Analytics

(apa yang terjadi?)

Data Pengunjung dan Transaksi E-commerce



"Jumlah **pengunjung** web e-commerce kita **naik drastis** bulan Juni, tetapi kenaikan jumlah **transaksi** kita masih **stagnan**"

Diagnostic Analytics

(mengapa itu terjadi?)

"Ditemukan data duplikat"

"Banyak pengunjung yang ternyata adalah bot"

"Ada salah satu produk kita yang sedang viral!"

"Kita sedang promo besar-besaran, tapi ternyata kode promo banyak yang gagal"

"Banyak pengunjung yang datang, tapi kesulitan menemukan produk yang cocok"

Contoh Skenario ...

Diagnostic Analytics

(mengapa itu terjadi?)

“Banyak pengunjung yang ternyata adalah bot”

“Banyak pengunjung yang baru datang, tapi kesulitan menemukan produk yang cocok”

Predictive Analytics

(prediksi apa yang akan terjadi)

→ “Develop sistem untuk memprediksi apakah suatu pengunjung adalah bot”

→ “Develop recommendation system (memprediksi apa produk yang mau dibeli oleh pengunjung tersebut)”

Prescriptive Analytics

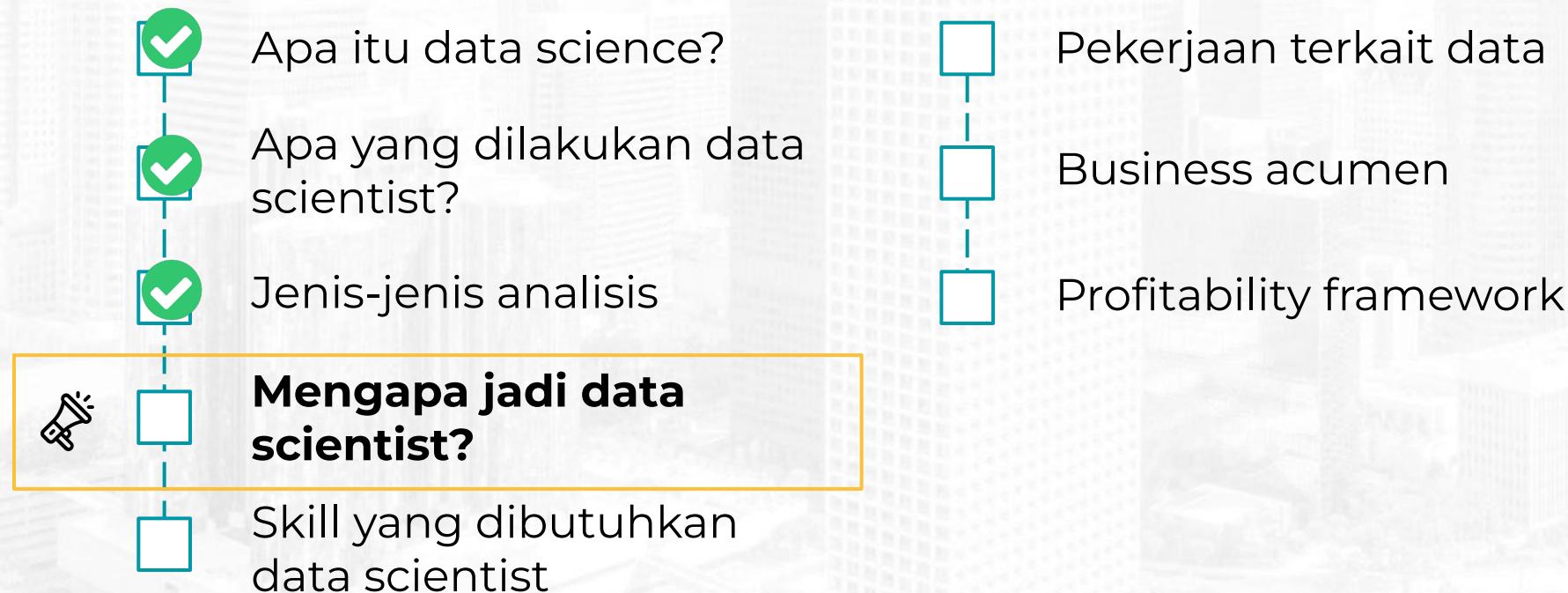
(follow up apa yang perlu dilakukan?)

→ “Beri restriksi untuk pengunjung yang diduga adalah bot”

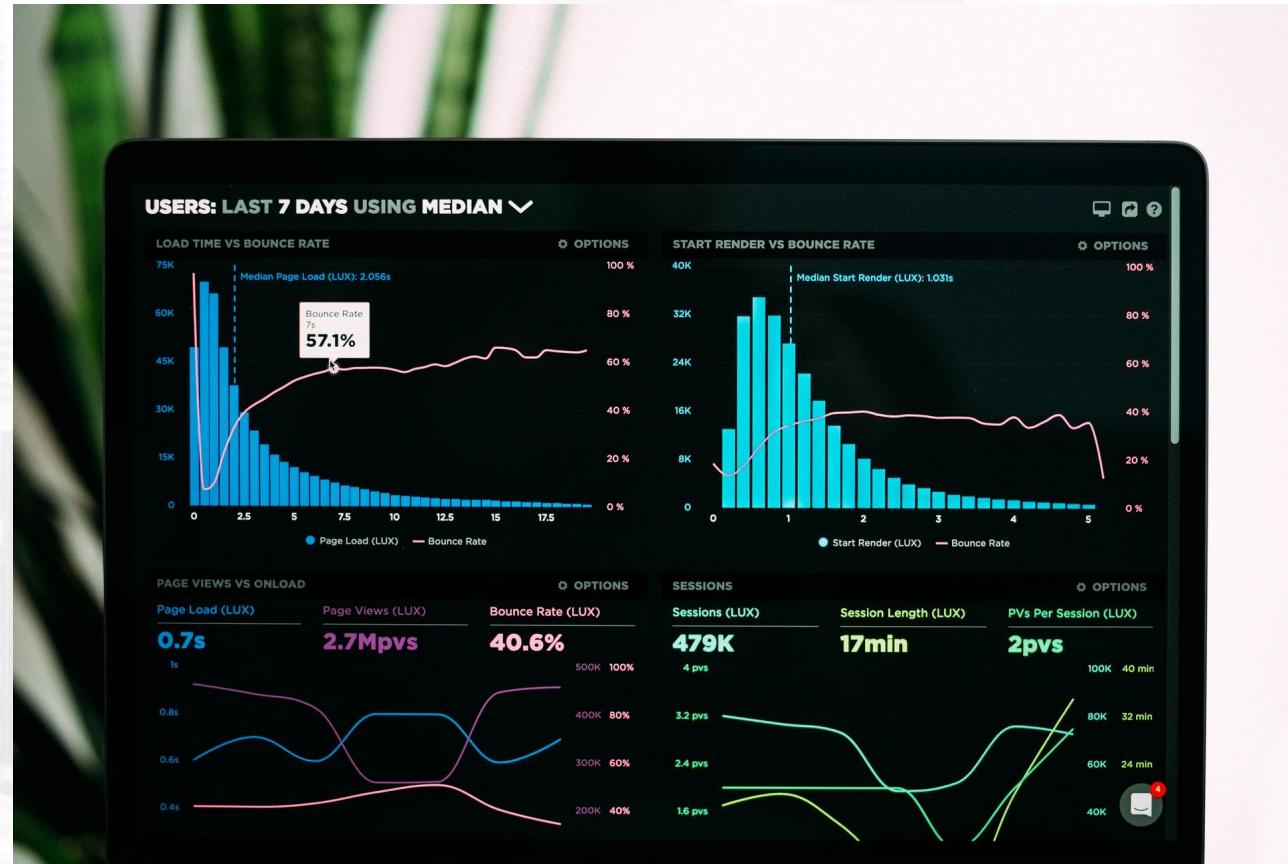
→ “Berikan rekomendasi produk yang tepat pada pengunjung agar mau beli”

Outline Pembelajaran

Data Scientist: What & Why?

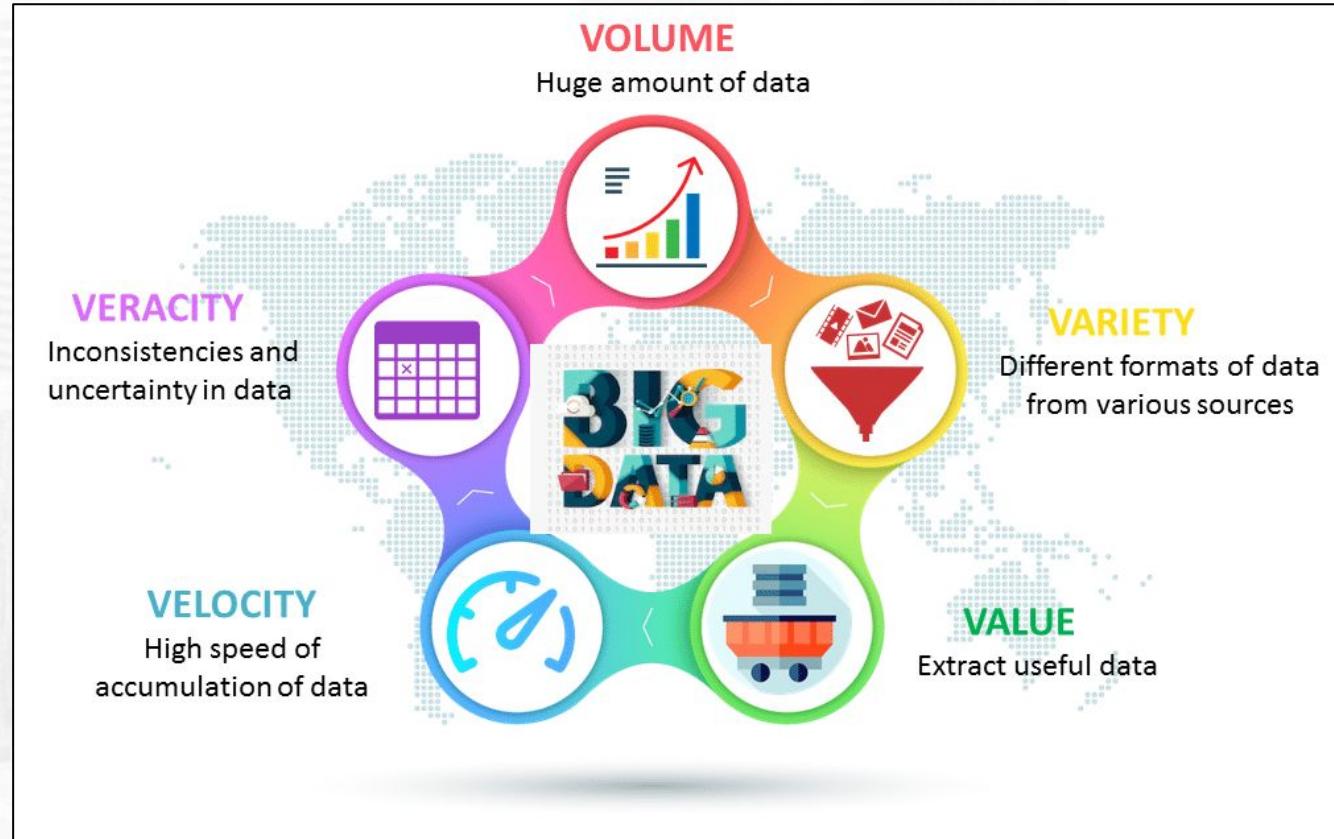


Kekuatan Data Terbukti Ampuh



Data Dapat Dimanfaatkan untuk Menyelesaikan Berbagai Problem

Perkembangan Teknologi Big Data



Data yang Super Banyak dan Kompleks Sudah Gampang Diakses

Prospek Karir yang Bagus



The Sexiest Job of the 21st Century

(menurut Harvard Business Review)

Demand Tinggi, Supply Masih Rendah

(banyak dicari perusahaan)

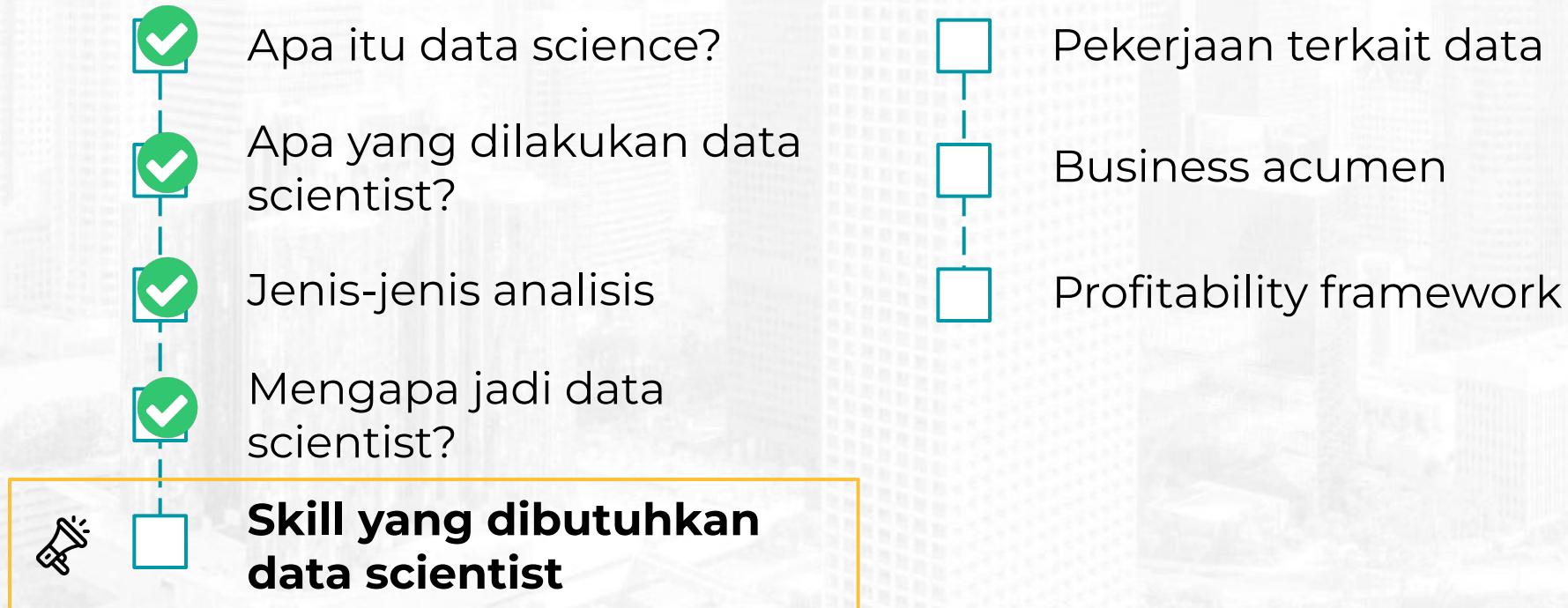
Bisa Dipelajari Siapa Saja



Banyak Sarana Belajar yang Terjangkau
(salah satunya ya **Rakamin Academy** ini ^^)

Outline Pembelajaran

Data Scientist: What & Why?



Technical Skill Utama



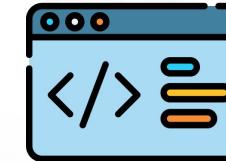
Analytics

Problem Solving, Data Exploration



Visualization

Right Plot for The Right Purpose



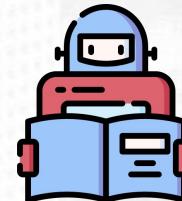
Programming

SQL, Python



Statistics

Uni/Bi/Multi-variate,
Hypothesis Testing



Machine Learning

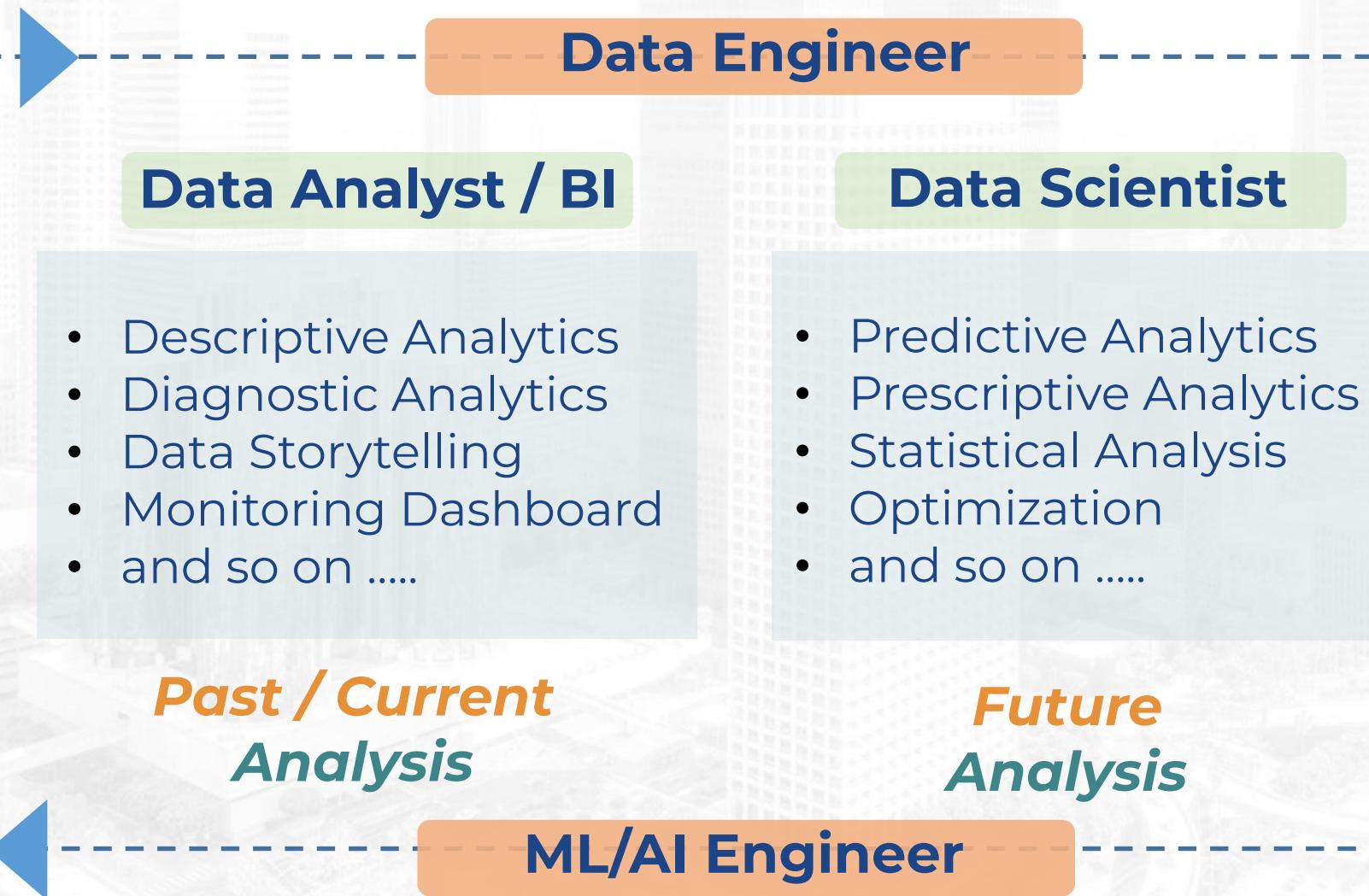
Supervised/Unsupervised, Optimization,
Deep Learning

Data Scientist: What & Why?

-  Apa itu data science?
-  Apa yang dilakukan data scientist?
-  Jenis-jenis analisis
-  Mengapa jadi data scientist?
-  Skill yang dibutuhkan data scientist



Pekerjaan Teknikal Terkait Data

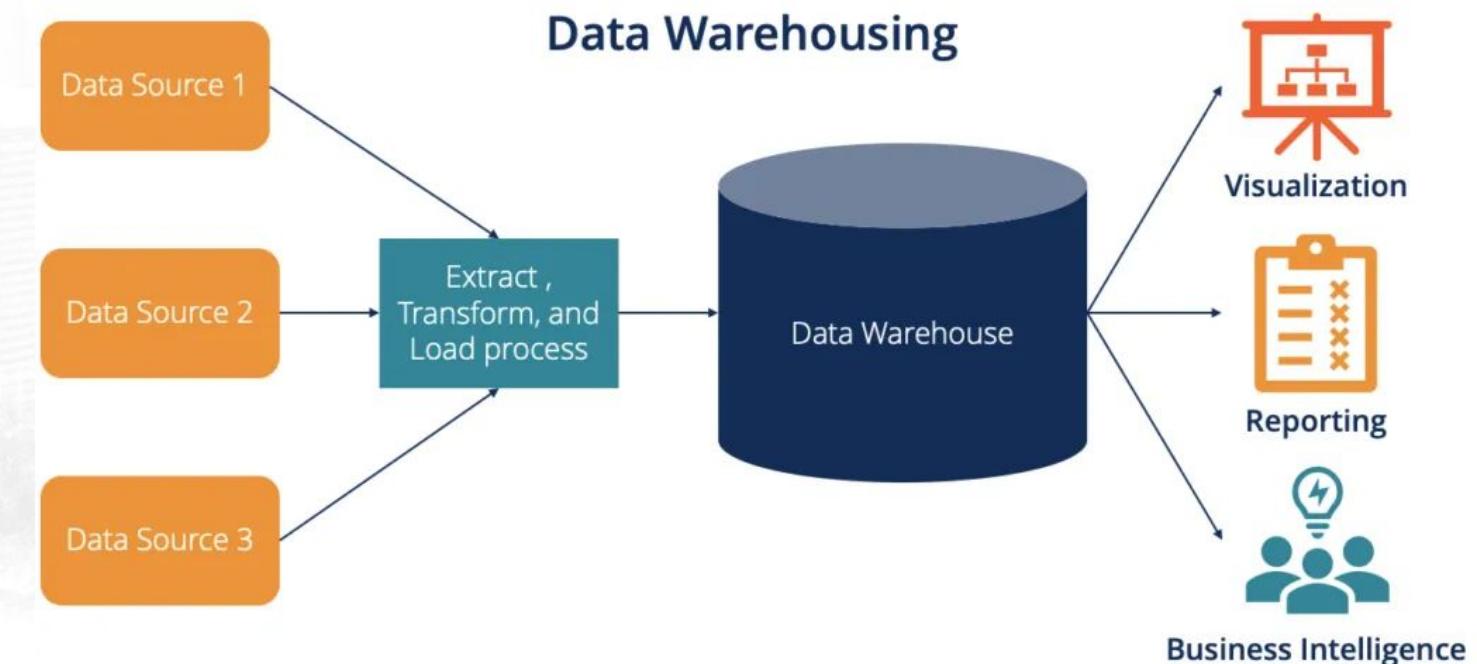


Data Engineer

Data Engineer adalah orang yang bertugas membangun **infrastruktur data** (database, pipeline, data warehouse, dsb) yang berkaitan dengan volume data dalam jumlah besar.

Pekerjaan terkait:

- Data Warehouse Engineer
- Data Wrangler



Data Analyst

Data Analyst bertugas untuk **menggali insight** dari data dan memberikan rekomendasi untuk pengambilan keputusan (fokus pada descriptive & diagnostic analytics)

Pekerjaan terkait:

- Business Intelligence
- Business Analyst
- Decision Scientist

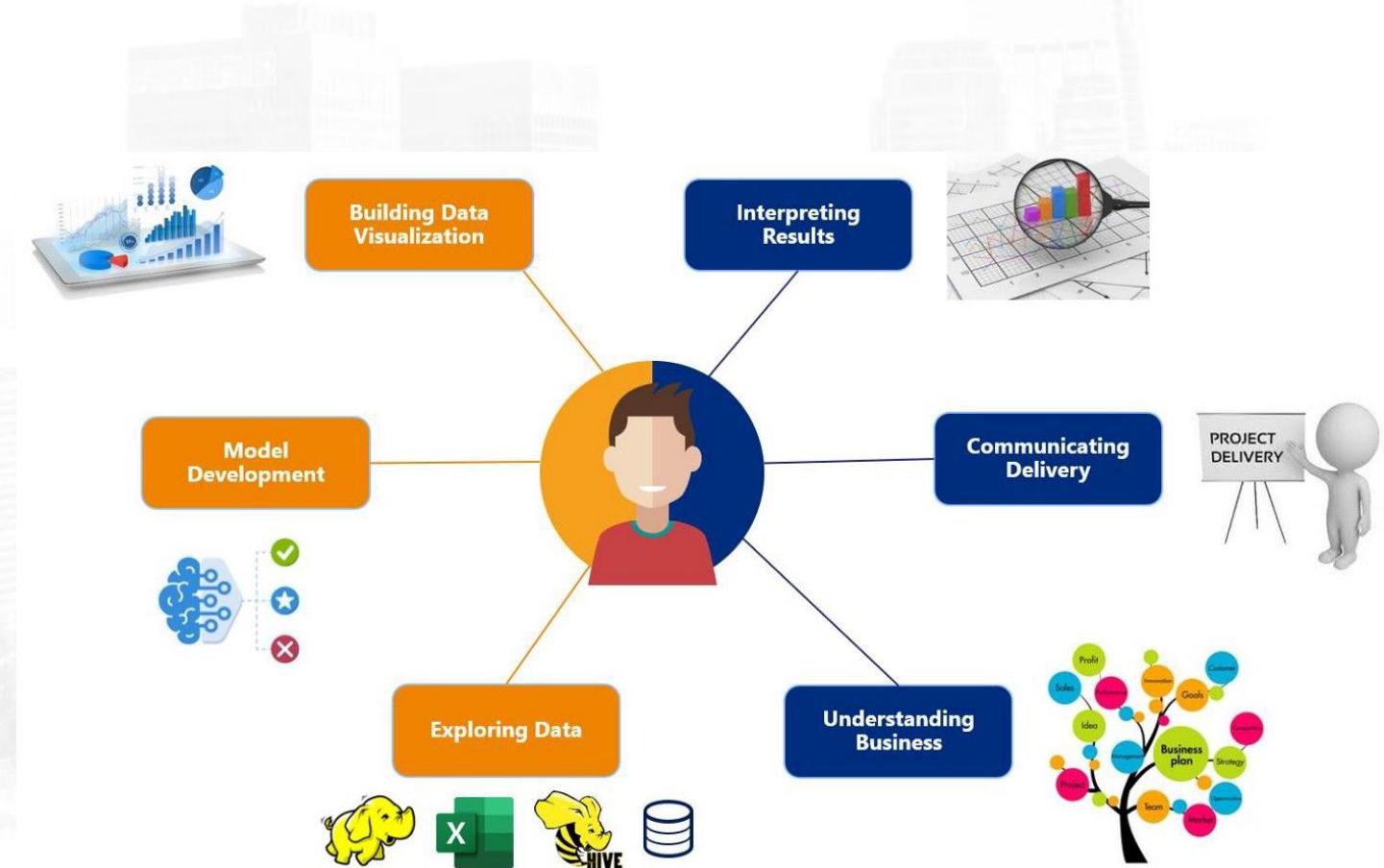


Data Scientist

Data Scientist bertugas membuat solusi dari sebuah permasalahan menggunakan DATA. Biasanya solusinya berupa machine learning model, optimisasi, atau predictive analytics lainnya.

Pekerjaan terkait:

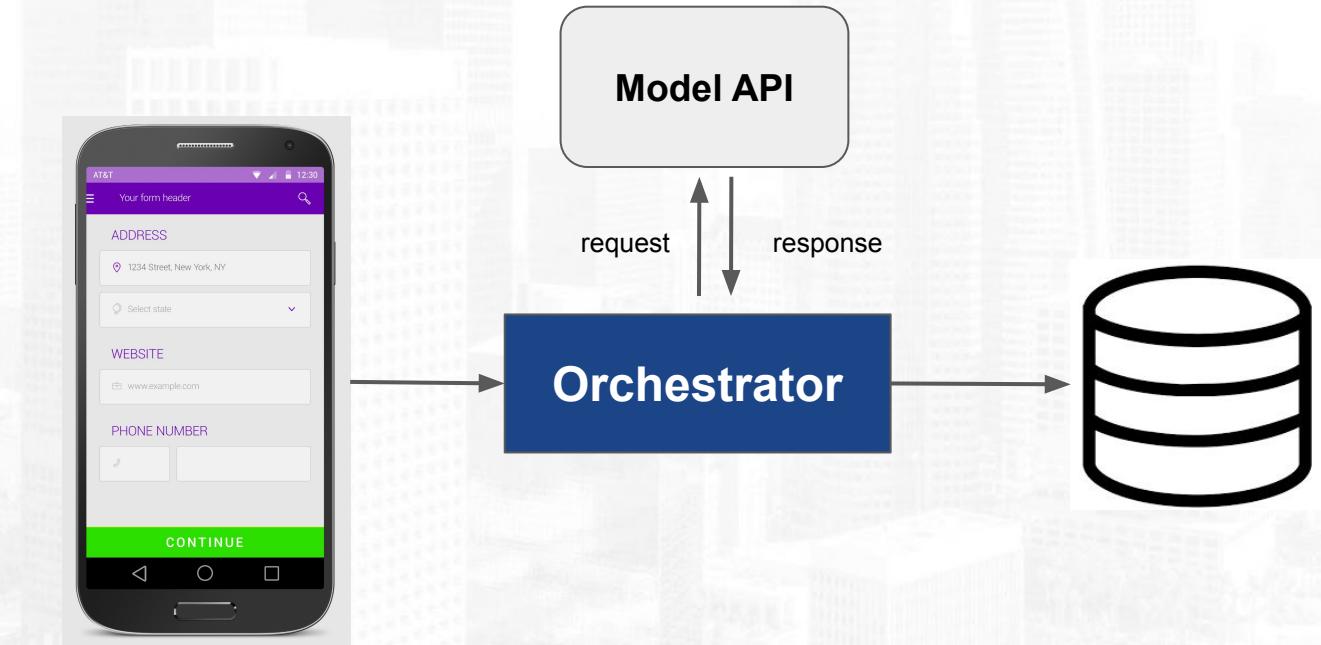
- ML Scientist
- AI Researcher
- Deep Learning Specialist



AI/ML Engineer

AI/ML Engineer bertugas mengintegrasikan model yang dibuat data scientist dengan produk/aplikasi

(tools yang digunakan sangat variatif, tergantung perusahaan)



Outline Pembelajaran

Data Scientist: What & Why?

- Apa itu data science?
 - Apa yang dilakukan data scientist?
 - Jenis-jenis analisis
 - Mengapa jadi data scientist?
 - Skill yang dibutuhkan data scientist
- Pekerjaan terkait data
 -  **Business acumen**
 - Profitability framework

Dell Data Maturity Model

Data Aware	Data Proficient	Data Savvy	Data Driven
Melakukan kompilasi berbagai data menjadi standardized reports secara manual	Pemrosesan data secara otomatis, dan pemahaman akan pentingnya data	Menggunakan data untuk pengambilan keputusan	Menggunakan data untuk semua proses bisnis. No data → no decision

Apa itu Business Acumen?

Business Acumen (ketajaman bisnis) adalah kemampuan untuk memahami dan menghadapi berbagai situasi bisnis, segala risiko dan peluangnya, dengan tajam dan akurat untuk mendapatkan hasil yang menguntungkan

Mengapa Data Scientist Butuh Business Acumen?

Identify
Business Problem



Translate into
Data Solution



Validate The
Impact

Apa yang Akan Terjadi Jika Data Scientist Tidak Punya Business Acumen yang Baik?

1. Solusi yang dibuat ternyata tidak relevan dengan masalahnya

Contoh

- **Problem:**

Banyak ditemukan produk palsu di suatu e-commerce, berpotensi membuat customer kehilangan kepercayaan

- **Solusi:**

Membuat model yang mendekripsi review produk yang menyatakan produknya palsu, lalu menghapus review tersebut agar revenue dari penjual tersebut tidak berkurang, dan bisa mengelak terkait adanya produk palsu.

Apa yang akan terjadi?

Justru makin banyak customer yang tertipu dengan produk palsu, para pemilik review pun akan sadar dan komplain karena reviewnya dihapus

Apa yang Akan Terjadi Jika Data Scientist Tidak Punya Business Acumen yang Baik?

2. Solusi yang dibuat tidak terpakai

Contoh

- **Problem:**

Banyak ditemukan produk palsu di suatu e-commerce, berpotensi membuat customer kehilangan kepercayaan

- **Solusi:**

Membuat model yang dapat mendeteksi produk palsu, tetapi bingung setelah itu modelnya dipakai untuk apa

Apa yang akan terjadi?

***Masalah tidak terselesaikan,
pekerjaan sia-sia, padahal model yang sudah dibuat sebenarnya bisa dioptimalkan***

Apa yang Akan Terjadi Jika Data Scientist Tidak Punya Business Acumen yang Baik?

3. Solusi yang dibuat justru menimbulkan masalah baru

Contoh

- **Problem:**

Banyak ditemukan produk palsu di suatu e-commerce, berpotensi membuat customer kehilangan kepercayaan

- **Solusi:**

Membuat model yang dapat mendeteksi produk palsu, lalu ketika ada customer yang membelinya, transaksi akan langsung dibatalkan secara otomatis

Apa yang akan terjadi?

Banyak customer yang komplain karena transaksi tiba-tiba dibatalkan

Bagaimana seharusnya?

Contoh

- **Problem:**
Banyak ditemukan produk palsu di suatu e-commerce, berpotensi membuat customer kehilangan kepercayaan
- **Solusi:**
Membuat model yang dapat mendeteksi produk palsu, lalu produk tersebut secara otomatis dihapus dari e-commerce dengan pesan bahwa produk yang dijual telah melanggar peraturan

Apa yang akan terjadi?

Kepercayaan customer meningkat

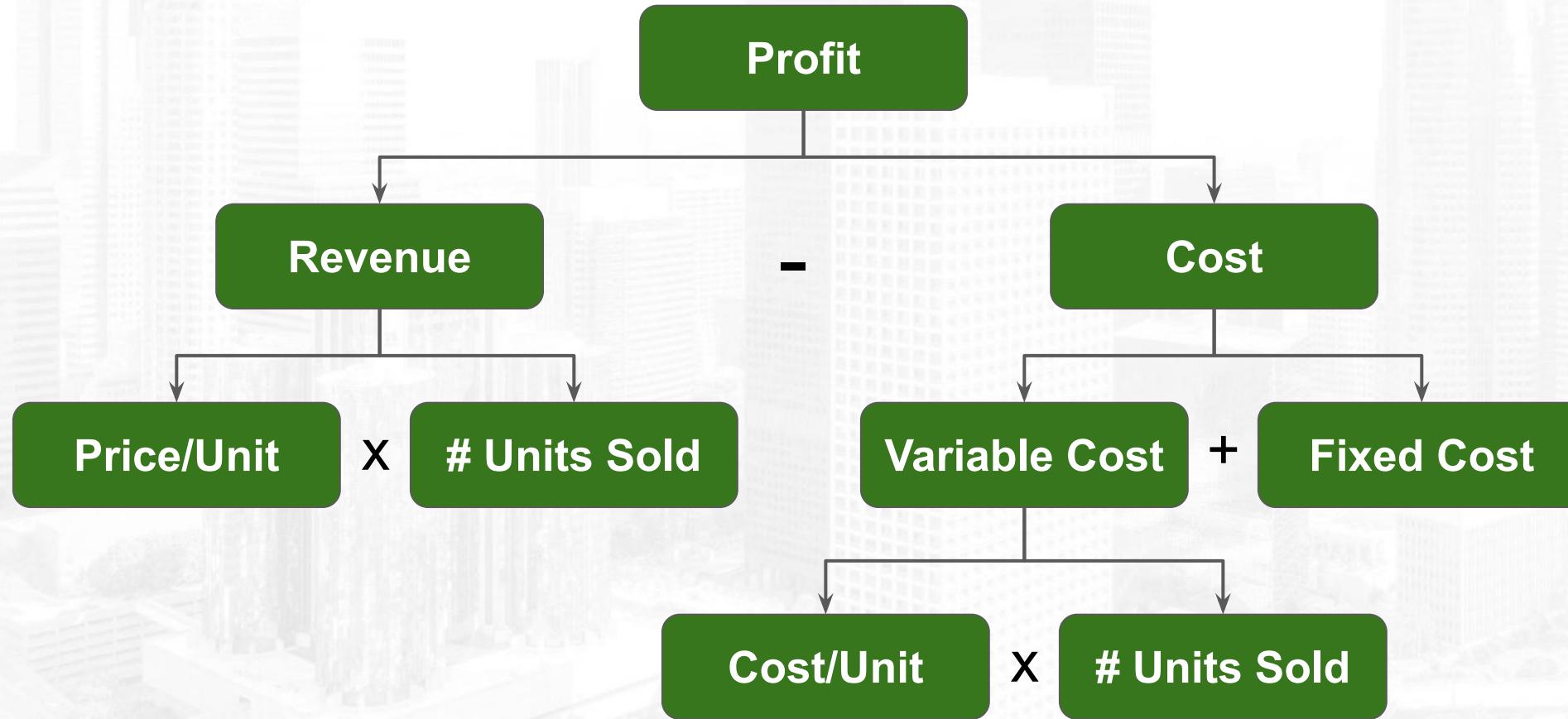
→ makin banyak yang berbelanja karena merasa aman

Outline Pembelajaran

Data Scientist: What & Why?

-  Apa itu data science?
 -  Apa yang dilakukan data scientist?
 -  Jenis-jenis analisis
 -  Mengapa jadi data scientist?
 -  Skill yang dibutuhkan data scientist
-  Pekerjaan terkait data
 - **Business acumen**
 -   **Profitability framework**

Profitability Framework



Bagaimana Cara Meningkatkan Revenue?

$$\text{Revenue} = \text{Price} \times \#\text{Units Sold}$$

Seberapa jauh kita bisa menaikkan harga tanpa menurunkan demand?

Price Optimization

- Pricing discrimination
- Dynamic pricing
- Subscription

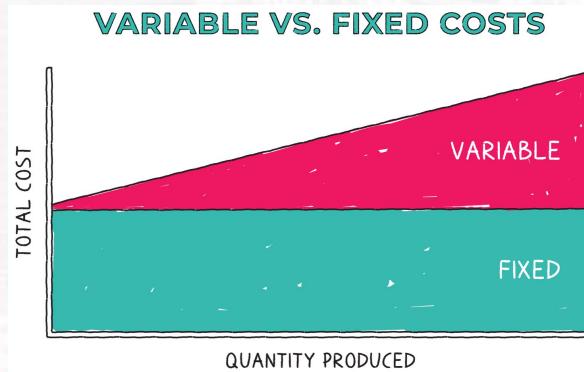
Bagaimana cara agar customer beli lebih banyak atau lebih sering?

Product Recommendation

- Analyze customer's interests
- Up-selling/Cross-selling
- Bundling option

Bagaimana Cara Mengurangi Cost?

Cost = Variable Cost + Fixed Cost

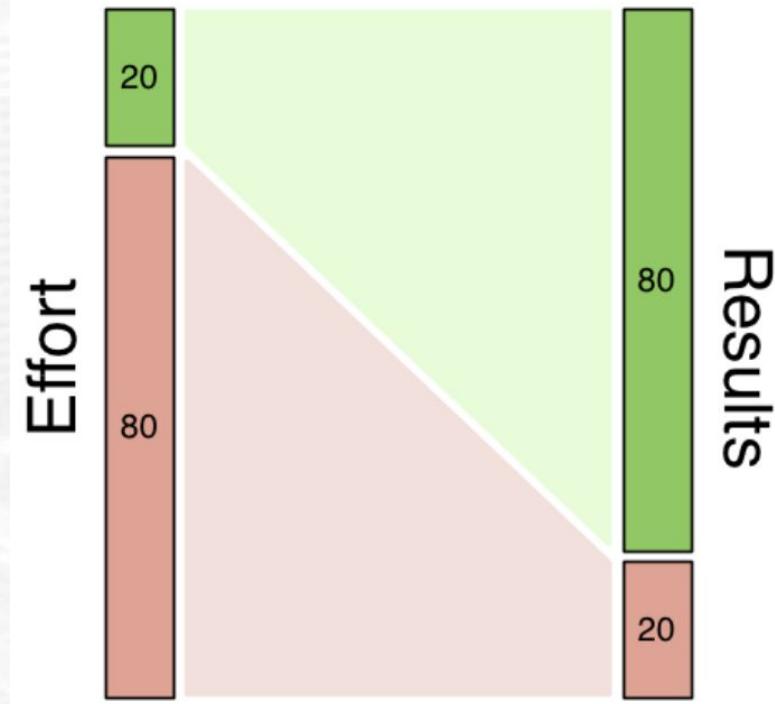


Bagaimana cara menekan cost?

- Automation
- Promo optimization
- Targeted marketing

Prinsip Pareto

“80% output itu berasal dari 20% input”



Contoh:

- **80% revenue** berasal dari **20% produk**
- **80% produk** berasal dari **20% kategori**
- **80% transaksi** berasal dari **20% pelanggan**
- **80% komplain** berasal dari **20% unit usaha**
- **80% penjualan** berasal dari **20% biaya marketing**

Cari tahu dan fokus pada 20% input tersebut!

Cara Menumuhkan Business Acumen

1. Pahami business flow di perusahaan
2. Jalin komunikasi yang baik dengan business stakeholders
3. Pertimbangkan business impact untuk setiap pengambilan keputusan



**Terima
Kasih!**