

Introduction to Big Data

oleh Randi Eka Sanjaya



Curriculum Vitae

Pendidikan

- S1 Teknik Industri, Universitas Atma Jaya Yogyakarta (2015-2020)
- S2 Teknik Industri, Universitas Atma Jaya Yogyakarta (2022-2024)

Karir

- Junior Data Scientist, PT. Solusi 247 Jakarta (2021-2023)
- Data Scientist, PT. Solusi 247 Jakarta (2023-Sekarang)

Proyek dan Aktivitas

- Proyek Big Data Analytics di salah satu BUMN (2021 - 2022)
- Proyek Data Warehouse and Business Intelligence di salah satu Kementerian (2022)
- Big Data Software Developer (2022-Sekarang)
- Keynote Speaker pada workshop internal Client
- Mentor pemegang



Apa itu data?

- Dalam penggunaannya sehari hari, data merupakan **koleksi nilai diskrit maupun kontinyu yang mengungkapkan informasi.**
- Data mendeskripsikan kuantitas, kualitas, fakta, dan statistik serta unit basis lainnya, atau serangkaian simbol yang dapat diinterpretasikan lebih jauh secara formal.
- Data yang kita jumpai biasanya terorganisir kedalam suatu struktur, seperti tabel

Attributes (Dimension; Features; Variables)

Objects (Samples, Individuals)

ID	Height	Weight	Age
Student 1	189 cm	81 kg	24
Student 2	210 cm	90 kg	26
Student 3	191 cm	92 kg	27
...
Student N	162 cm	71 kg	21

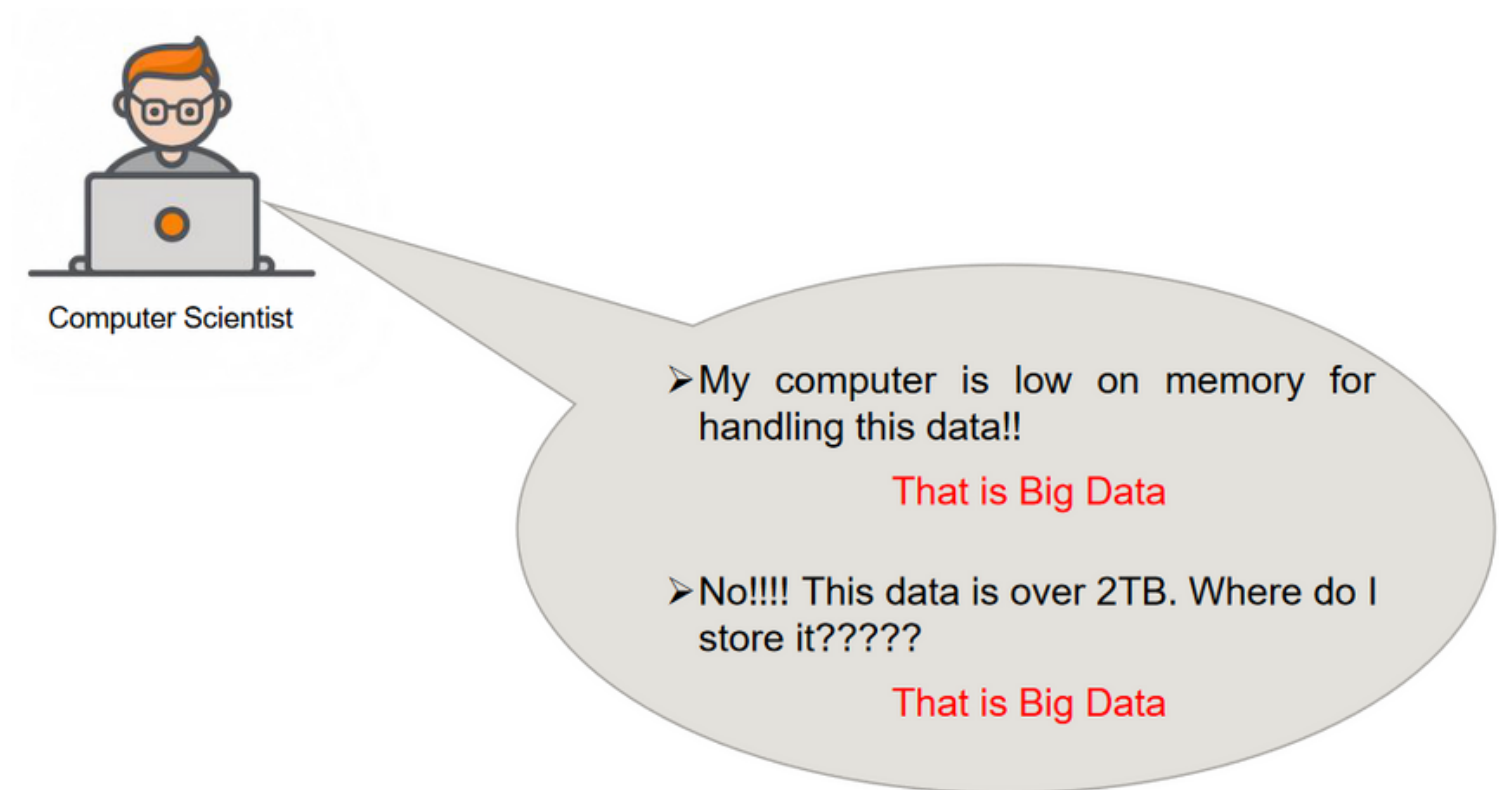


Kemudian, bagaimana dengan Big Data?

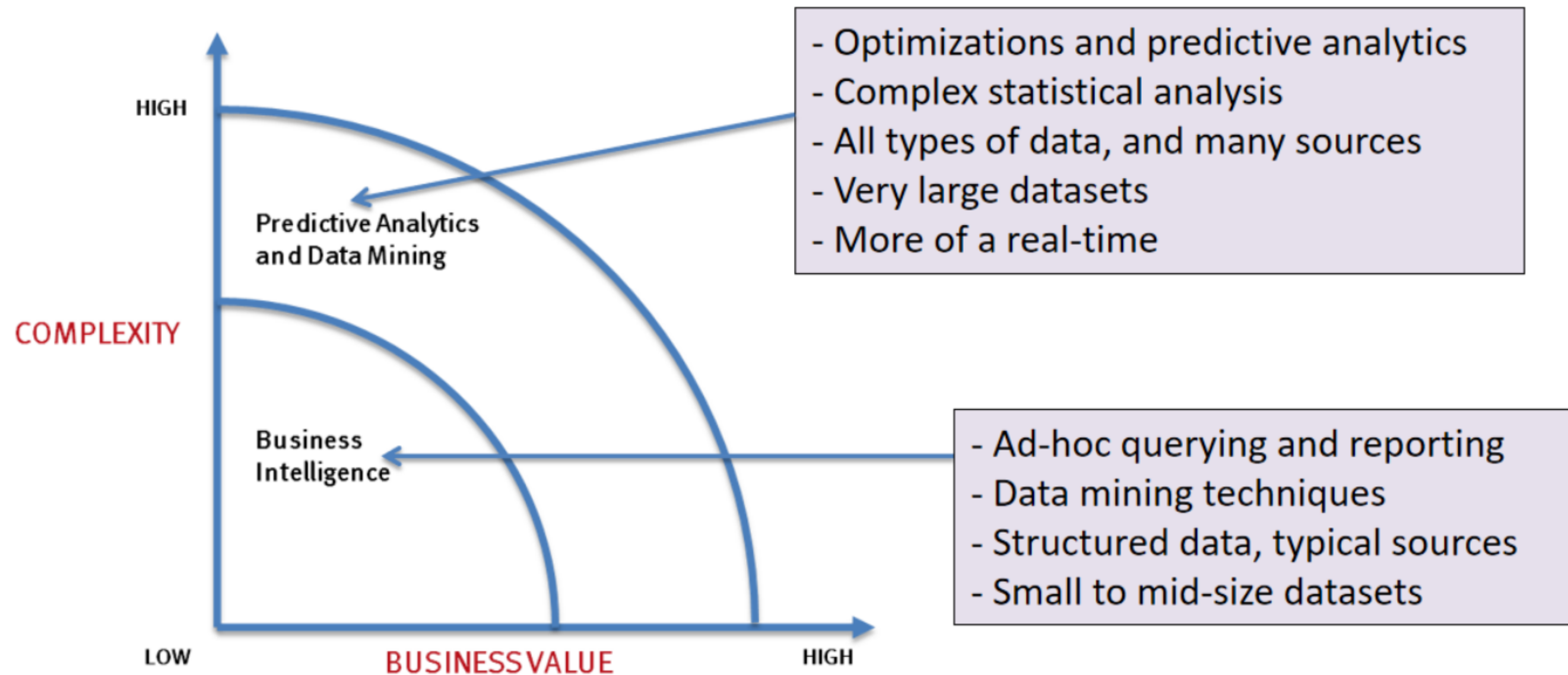
- Dalam hal ini, Big Data secara sederhana dapat diartikan sebagai **set data yang memiliki jumlah data yang besar dan beragam, baik dalam bentuk data terstruktur, semi-terstruktur, maupun tidak terstruktur.**
- Big data ditujukan pada semua hal tentang data, dari bagaimana data dikumpulkan, diproses, di analisis, untuk menghasilkan nilai tambah melalui data-driven insights dan keputusan.
- Big Data merupakan informasi yang tidak dapat dipahami ketika digunakan dalam jumlah yang kecil
- Trend terhadap data yang semakin besar dikarenakan terdapat informasi tambahan yang dapat diperoleh ketika menganalisis suatu set data yang besar, dibandingkan dengan dataset kecil yang terpisah dengan besar data yang sama.

Bagaimana kita tahu data kita merupakan Big Data?

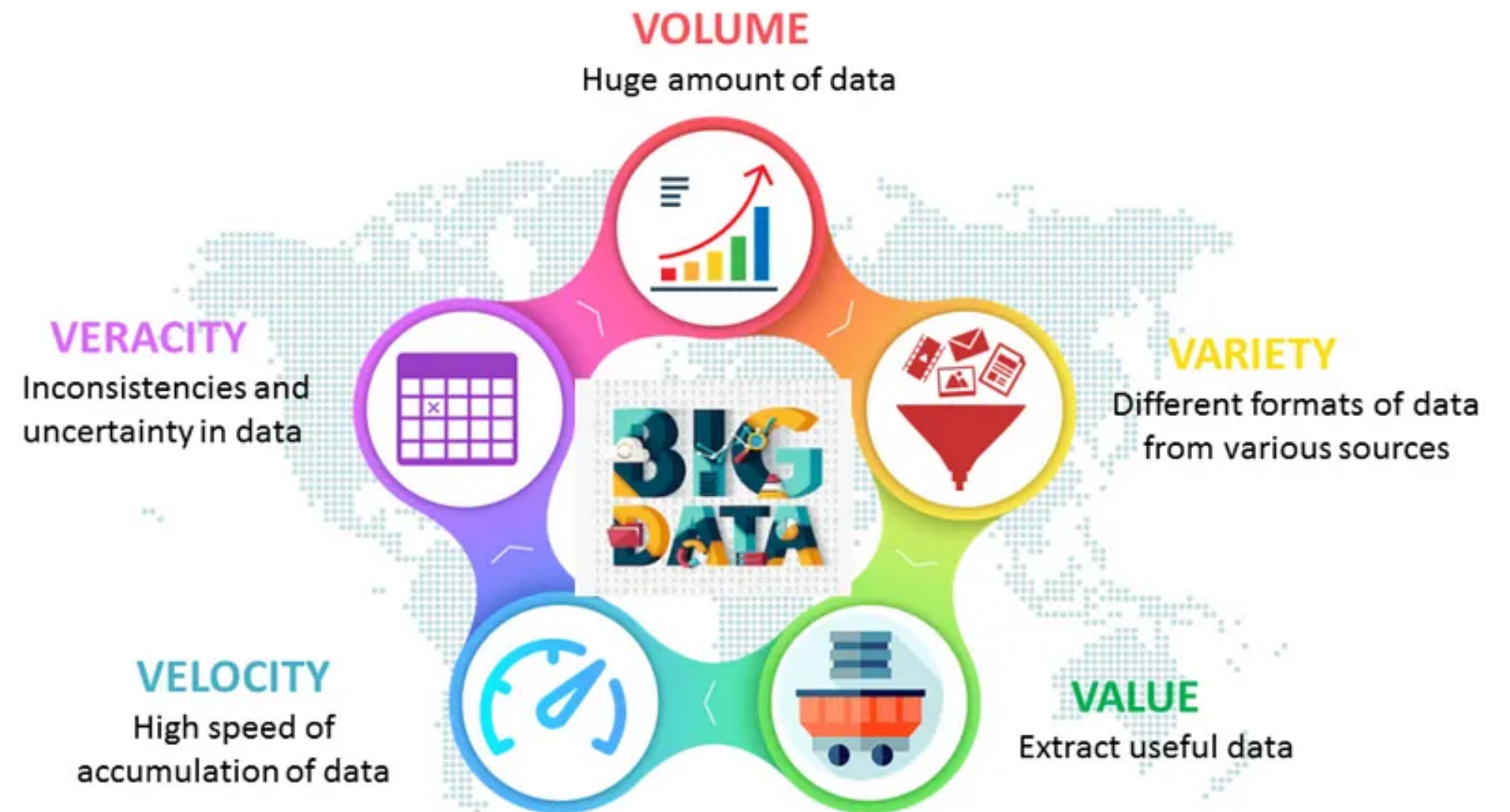
- Big Data merupakan istilah untuk koleksi set data yang sangat besar dan kompleks, hingga sulit untuk diolah menggunakan database management tools ataupun aplikasi pemrosesan data tradisional.
- Singkatnya, **jika anda mengalami kendala dalam melakukan pemrosesan dan analisis data dengan komputer anda, hal ini dapat disebabkan oleh Big Data**



Apa yang mendorong pertumbuhan Big Data?



Apa Saja Karakteristik Big Data?



<https://medium.com/@analyticsemergingindia/understanding-the-5-vs-of-big-data-unveiling-its-characteristics-27ec4a0f35a6>

Big Data Tools

Mari kita coba saja dengan Hands on, klik [disini](#)






Apa keuntungan dan kelemahan Big Data?

Keuntungan

- Meningkatkan kualitas pengambilan keputusan
- Meningkatkan efisiensi
- Meningkatkan pentargetan konsumen
- Menambah sumber pendapatan
- Memperoleh keunggulan kompetitif


Kelemahan

- Permasalahan privasi
 - Risiko pelanggaran data
 - Tantangan dari segi teknis meningkat
 - Kesulitan dalam mengintegrasikan sumber data
 - Kompleksitas Analisis
- 



Big Data Best Practice

Guna secara efektif mengelola dan menutlisasi Big Data, organisasi dapat mengikuti beberapa kaidah berikut:

- Mendefinisikan tujuan bisnis dengan jelas.
 - Hanya mengumpulkan dan menyimpan data yang relevan.
 - Memastikan kualitas data terjaga.
 - Menggunakan teknologi dan tools yang tepat.
 - Menerapkan data security dan kebijakan privasi.
 - Meningkatkan adaptasi terhadap Machine Learning dan Artificial Intelligence.
 - Berfokus pada visualisasi data.
- 

Aplikasi Big Data dan Studi Kasus dalam Teknik Industri

Area Supply Chain:

- Monitoring dan perencanaan stok
- Analisa anomali stok opname
- Segmentasi jenis transportasi
- Analisa kontrak transportasi
- Segmentasi gudang
- Evaluasi kinerja distributor / gudang
- Optimalisasi least cost distribution (Network Optimization)
- Mitigasi Risiko
- Estimasi Safety Stock
- Greenfield Analysis
- Digital Twins



Aplikasi Big Data dan Studi Kasus dalam Teknik Industri

Area Market Intelligence:

- Segmentasi konsumen
- Analisa perilaku konsumen
- Analisa profil 360 konsumen
- Analisa informasi permintaan dan dampak terhadap harga
- Prediksi harga komoditi
- Analisa market share
- Prediksi penentuan harga jual





Sumber

- [Introduction to Big Data by Korea University in 2019](#)
- [Introduction to Big Data by Joan Jarmack](#)
- [Introduction to Big Data by Kent State University](#)
- [What is Big Data? Introduction, Uses, and Applications by Analytics Vidhya](#)
- [Big Data: An Introduction](#)
- [Big data quality framework: a holistic approach to continuous quality management](#)
- [A Must-Read Guide on How to Work with PySpark on Google Colab for Data Scientists! by Analytics Vidhya](#)