



Introduction to Data Science

Bern Jonathan
Data Scientist at Female Daily
Network



Question

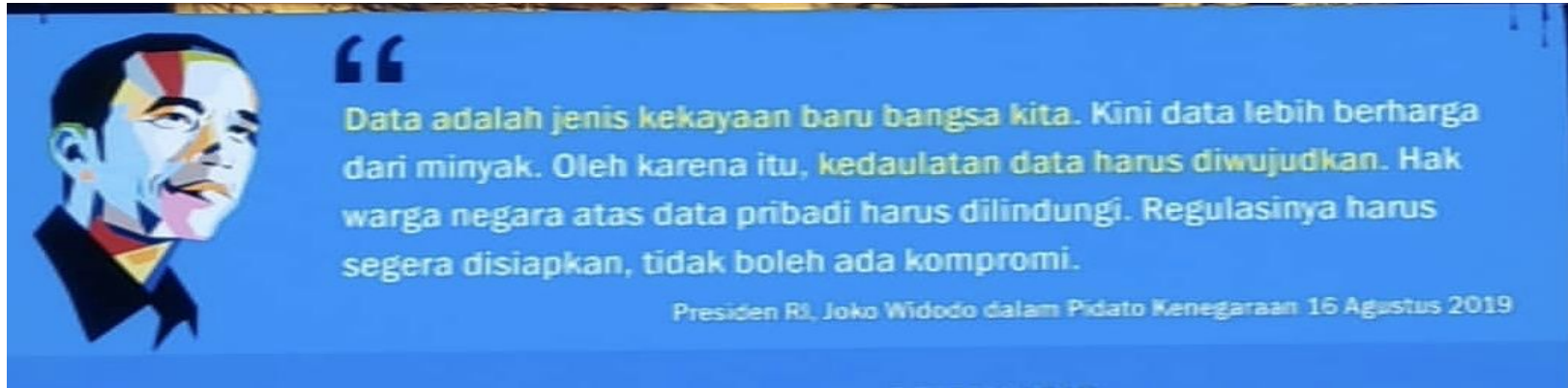
- What is Data Science ?
- Who is Data Scientist?
- What Data Scientist do?
- Data Scientist Type?
- Why we need them?

Answer it at <http://bit.ly/kepodsunai>



<http://bit.ly/demodsunai>

Industry 4.0





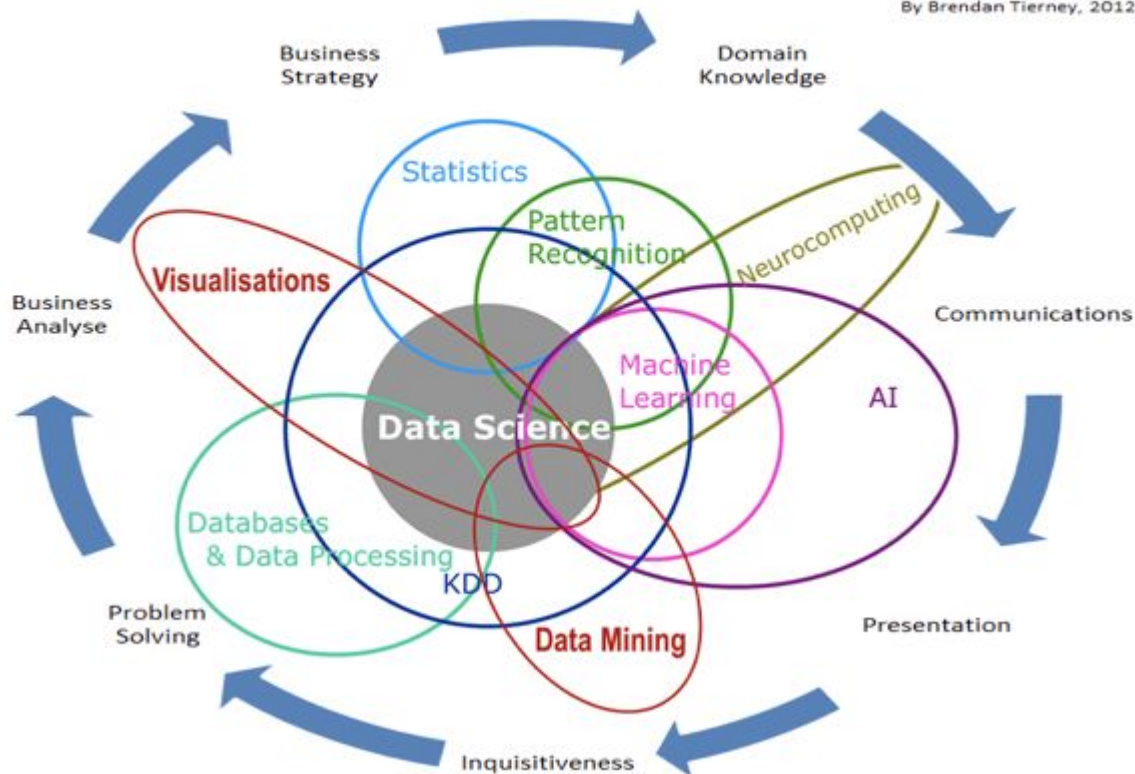
Data science?

Data science is an interdisciplinary field that uses scientific methods, processes, algorithms and systems to extract knowledge and insights from data in various forms, both structured and unstructured, similar to data mining. ~Wikipedia

Data science is the field of study that combines domain expertise, programming skills, and knowledge of math and statistics to extract meaningful insights from data. Data science practitioners apply machine learning algorithms to numbers, text, images, video, audio, and more to produce artificial intelligence (AI) systems that perform tasks which ordinarily require human intelligence. In turn, these systems generate insights that analysts and business users translate into tangible business value.
(Data Robot)

Data Science Is Multidisciplinary

By Brendan Tierney, 2012



So Data Science is?

Data Science is apply Data with Science

-Bern Jonathan





Who is Data Scientist?

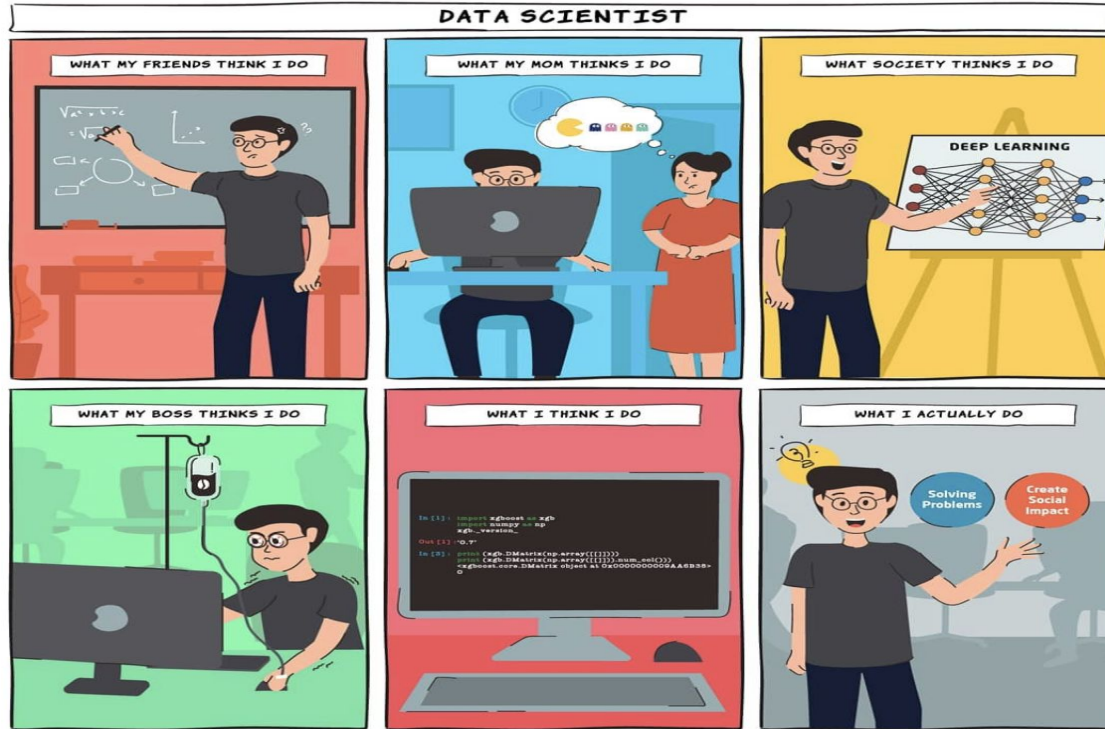
Simply:

A man who does data science things.

Expertly:

Build a Social Impact!

Data Science?





(((Josh Wills)))

@josh_wills



Follow

Data Scientist (n.): Person who is better at statistics than any software engineer and better at software engineering than any statistician.

RETWEETS

1,506

LIKES

1,087



12:55 AM - 4 May 2012



48




1.5K



1.1K



Data Scientist



[All](#) [Images](#) [Videos](#) [News](#) [Shopping](#) [More](#) [Settings](#) [Tools](#)

About 40,000 results (0.60 seconds)

VFS Digital Design/Flickr In 2012, Harvard Business Review named **data scientist** the "sexiest job of the 21st century." More recently, Glassdoor named it the "best job of the year" for 2016. "It isn't a big surprise," Dr. Andrew Chamberlain, Glassdoor's chief **economist**, told Business Insider. Feb 25, 2016

[How much money you earn in the 'sexiest job of the 21st century ...](https://www.businessinsider.com/how-much-money-you-earn-in-the-sexiest-job-of-the-21st-century)
[https://www.businessinsider.com/how-much-money-you-earn-in-the-sexiest-job-of-the-2...](https://www.businessinsider.com/how-much-money-you-earn-in-the-sexiest-job-of-the-21st-century)

[About this result](#) [Feedback](#)

People also ask

What are 21st century jobs?

What is the sexiest career?

How can I be a data scientist?

Is data science a Fad?

[Feedback](#)

[Data Scientist: The Sexiest Job of the 21st Century](https://hbr.org/2012/10/data-scientist-the-sexiest-job-of-the-21st-century)
<https://hbr.org/2012/10/data-scientist-the-sexiest-job-of-the-21st-century>



Data Team

Data Scientist

also known as Data Managers, statisticians.



A data scientist will be able to take data science projects from end to end. They can help store large amounts of data, create predictive modelling processes and present the findings.

Skills: Mathematics, Programming, Communication



Will use programmes such as:
SQL, Python, R

Data Engineers

also known as database administrators and data architects.



They are versatile generalists who use computer science to help process large datasets. They typically focus on coding, cleaning up data sets, and implementing requests that come from data scientists.

Skills: Programming, Mathematics, Big data



Will use programmes such as:
Hadoop, NoSQL, and Python

Data Analysts

also known as business Analysts.



They typically help people from across the company understand specific queries with charts.

Skills: Statistics, Communication, Business knowledge



Will use programmes such as:
Excel, Tableau, SQL

What Data Science do?



A typical data science process looks like this, which can be modified for specific use case:

- Understand the business
- Collect & explore the data
- Prepare & process the data
- Build & validate the models
- Deploy & monitor the performance

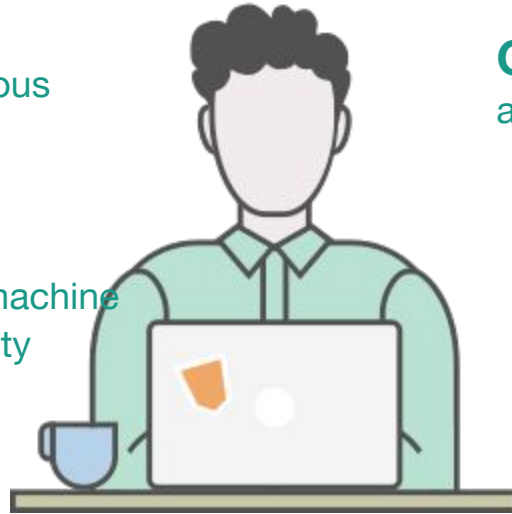
WHAT DOES A GOOD DATA SCIENTIST LOOK LIKE?

Inquistive – skeptical and curious

Coding – is good at coding, hacking, and general programming

Knowledgeable – knows machine learning, statistics, and probability

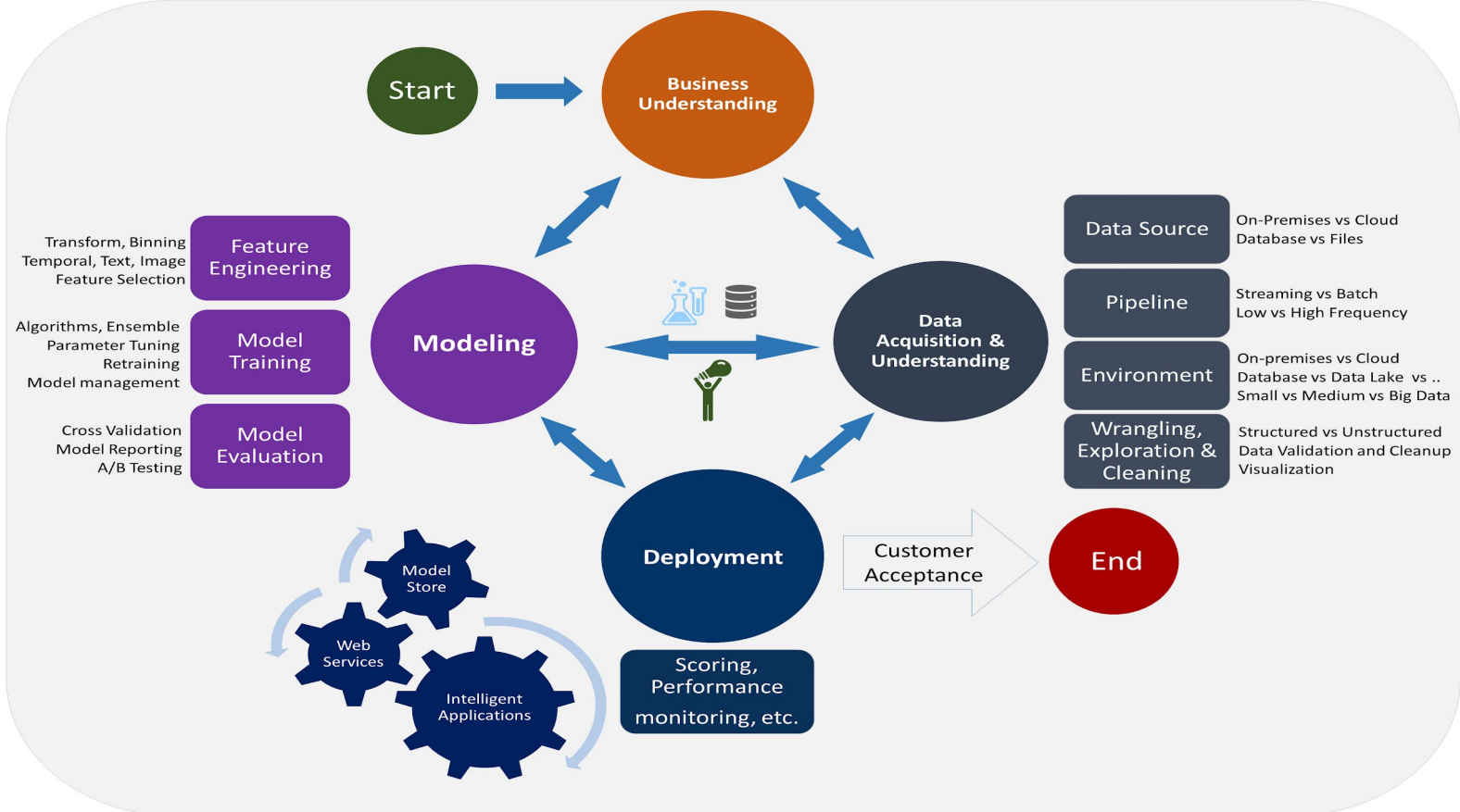
Product Oriented – knows how to build data products and visualizations to make data understandable to mere mortals



Scientific Method – Creates hypotheses, tests them, and updates understanding

Domain Knowledge – understands the business and how to tell the relevant story from business data. Able to find answers to *known unknowns*.

Data Science Lifecycle

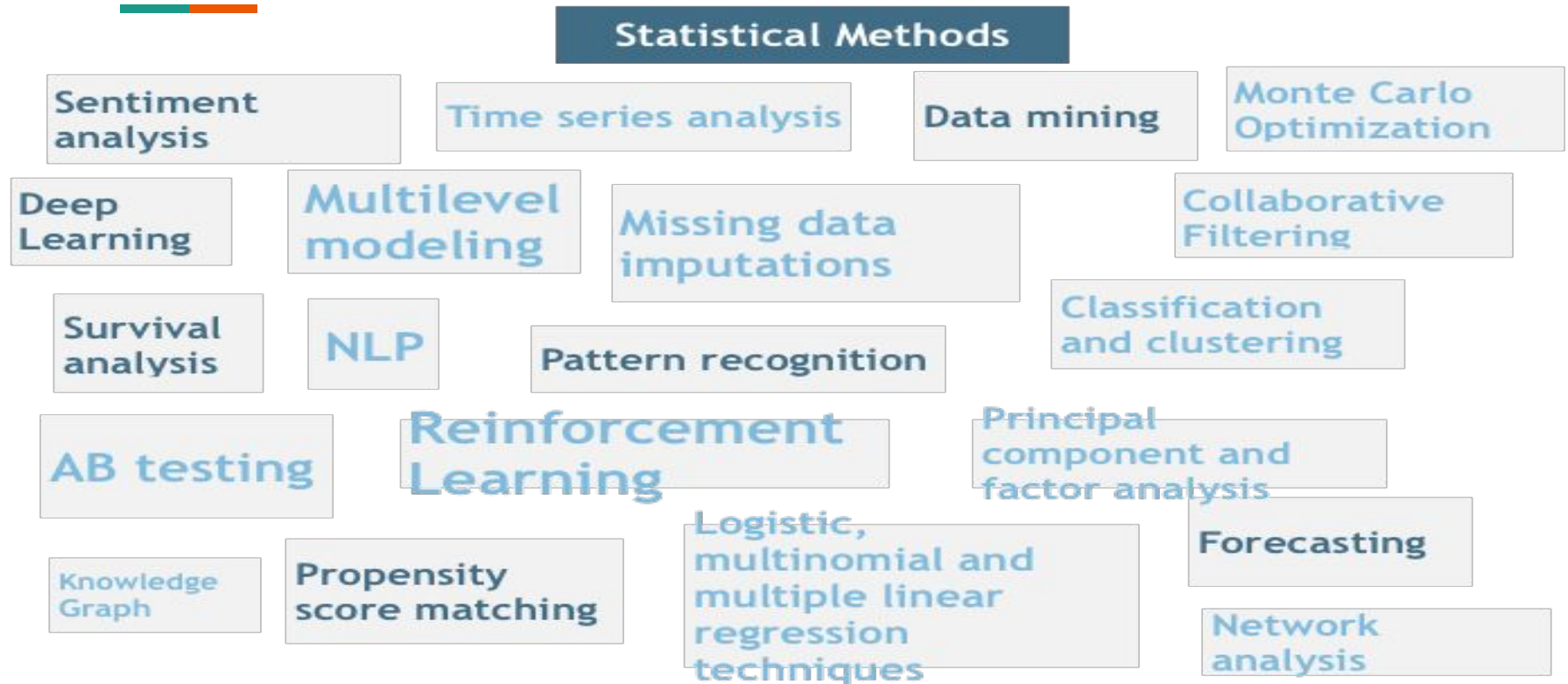


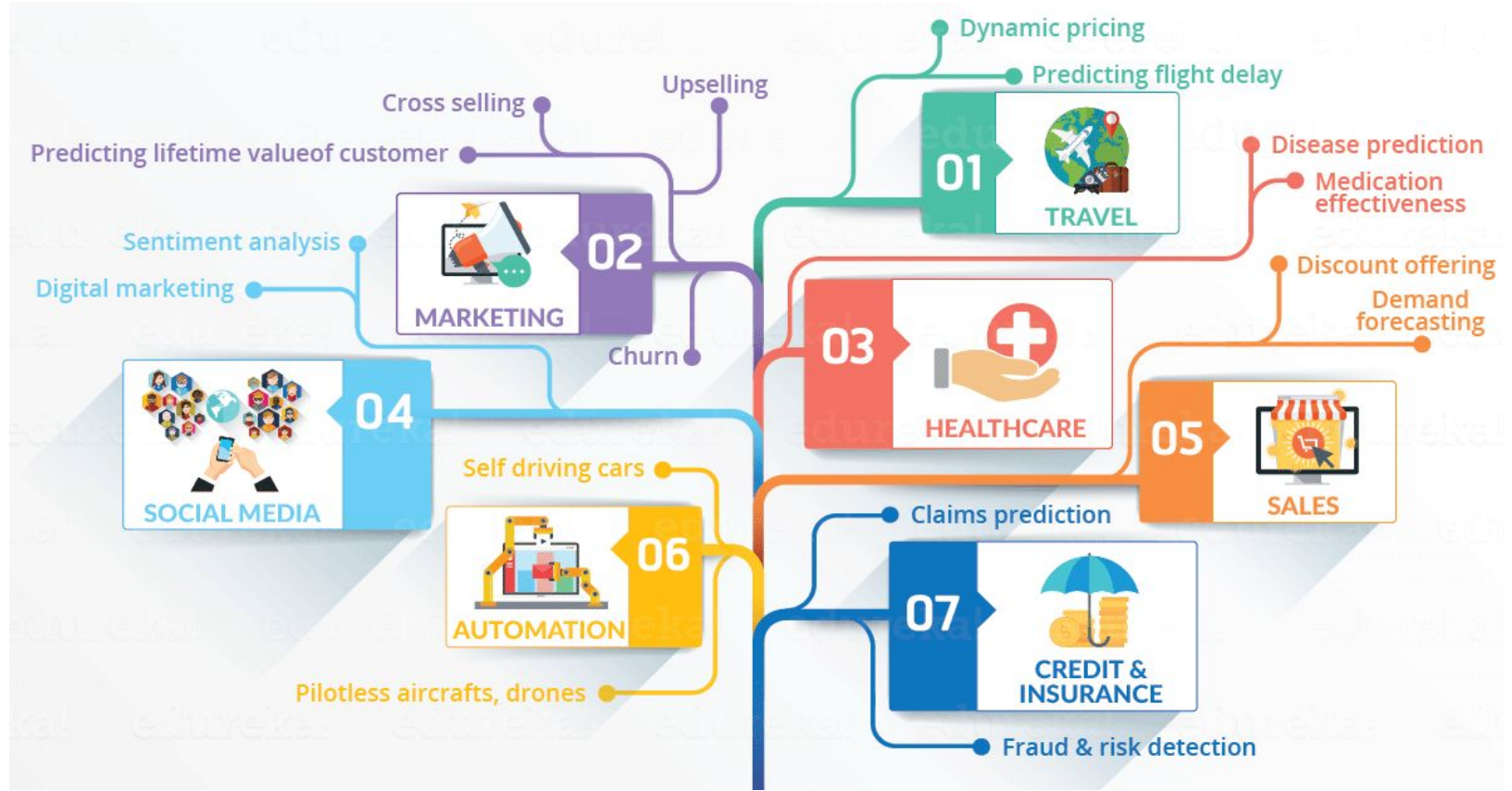


Commonly in Data Science:

1. Supervised Learning (Classification & Regression)
2. Unsupervised Learning (Clustering & Anomaly Detection)
3. Reinforcement Learning
4. Deep Learning

What's in the DataScience Methods

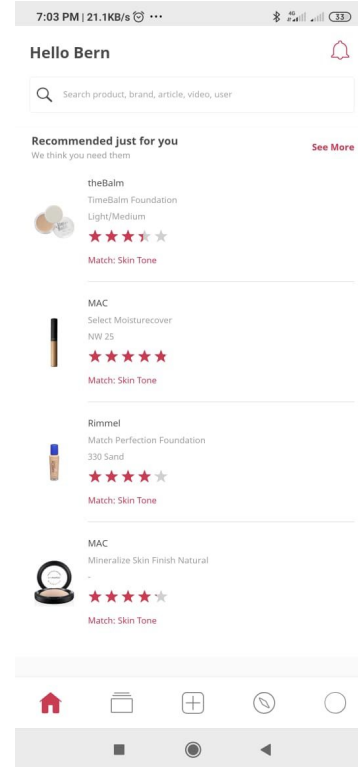
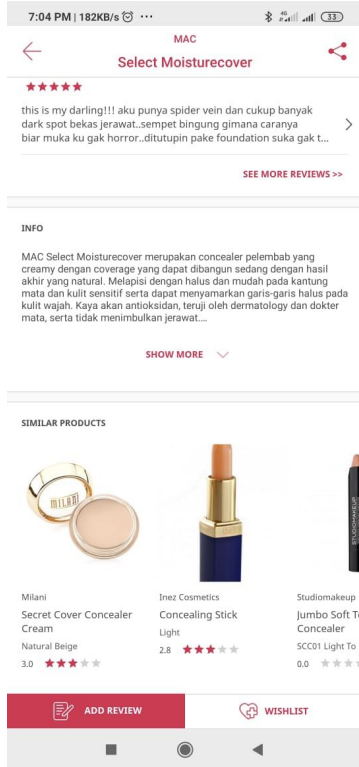






What Data Science do in Female Daily Network

Recommendation Engine





Female Daily Language Processing (On-Going)

aku pake yg biru awalnya.. kotoran keangkat semua sih dua2nya sama aja.
tapi aku kek nya lebih suka pink karena pas abis make lebih seger yg pink..
baunya entah kenapa aku suka2 aja wkwwkkwkwkwkwkwk overall
okelahh gudd sukaa

Click to add text

Aku pakai ini untuk double cleansing sebelum pakai faceial wash. Untuk ngehapus foundation/bb cream yang diset dengan compact powder dan loose powder, harus sampai 3 kapas, atau kalau mau cepat kapasnya diganti sama tissue basah. Sama kalau untuk eyemakeup kaya eyeliner dan maskara, agak susah buat bersihinnya, walaupun eyeliner maskaranya bukan yang waterproof. Kalau untuk bersihin cuma sunblock/sunscreen cukup 1 kali aja. Di aku sih nggak bikin jerawat, tapi jadi agak kering jadi hati-hati yang kulitnya dry

Ini pembersih muka enak banget menurutku. Lumayan ngebersihin bekas make up di wajah. Dulu pas jaman masih awal kuliah suka banget pakai face wash ini. Sayangnya sekarang udah sulit di temukan di minimarket minimarket.

aku pake yg biru awalnya.. kotoran keangkat semua sih dua2nya sama aja.
tapi aku kek nya lebih suka pink karena pas abis make lebih seger yg pink..
baunya entah kenapa aku suka2 aja wkwwkkwkwkwkwkwk overall
okelahh gudd sukaa

Aku lebih suka beli yang serie O.P.I mini ketimbang yang full size, mungkin karena jarang dipakai, biasanya aku pakai kalau lagi masuk monthly period hihhihihi.., jadi sebulan rata-rata cuma pakai seminggu doank. Sering kejadian kalau beli yang full size, produk masih banyak, formulanya udah berubah jadi lebih sticky *curcol. Balik lagi ke si O.P.I Mini Mariah Carey ini, warnanya cakep-cakep, looks so glam, ke empatnya cocok buat di mix and match satu sama lain wajib masuk daftar koleksi deh buat yang suka warna-warna yang elegant, kemasannya imut lagi.

Kemasannya lucu. Botol mini. Teksturnya agak kental. Cak perlu berkali-kali oles, warnanya sudah "keluar". Warnanya cantik-cantik dan harganya terjangkau. Overall saya suka dengan set Mini Four Pack ini



NER

produk pertama yg dipake dr **the ordinary BRAND** cocok banget muka jadi lbh bersih di kombinasi sama **mandelic acid INGREDIENTS** + **hada labo BRAND** moisturizer

trus always pake sunblock di siang hari cocok pake banget 😍😍😍😍😍😍

Typo Correction (Jaksel type)

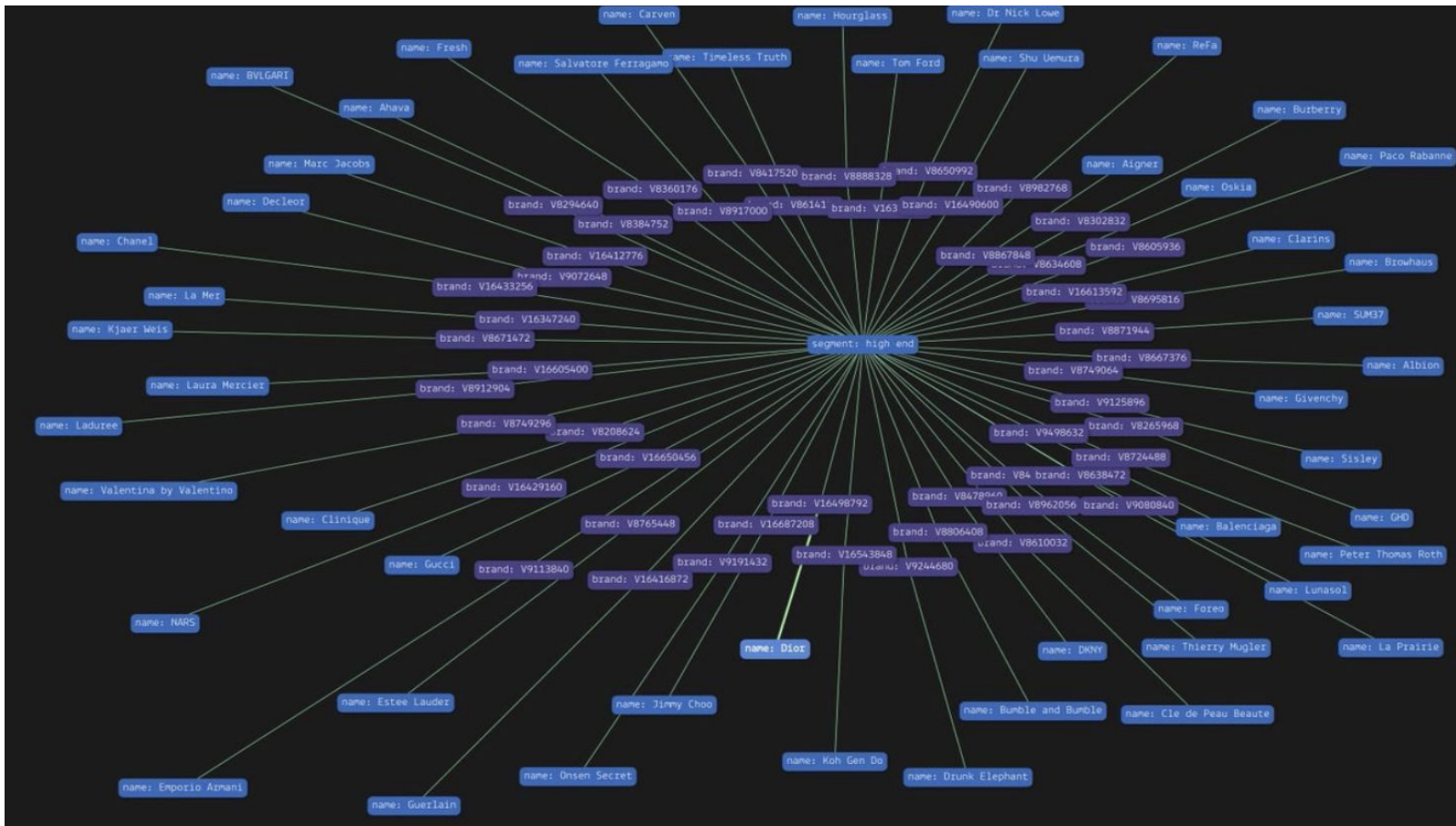
A lot of Typo = Can't Understand :(

```
correct_text('acctually, produk ini tuh ga enak banget asli. kualitasna jelwk banget sis')
```

```
'actually, produk ini tuh ga enak banget asli. kualitasnya jelek banget sis'
```



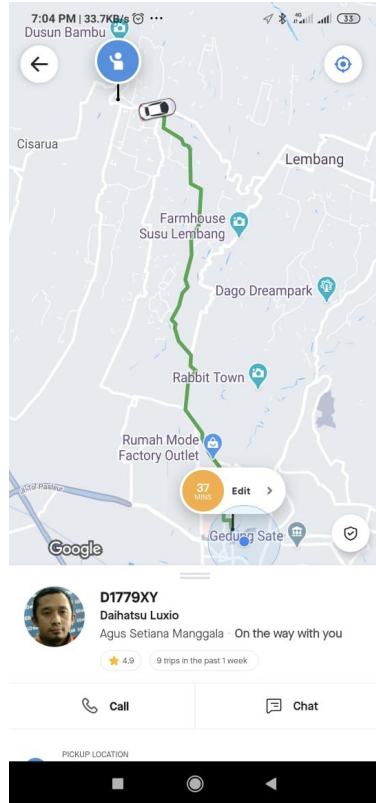
Knowledge Graph



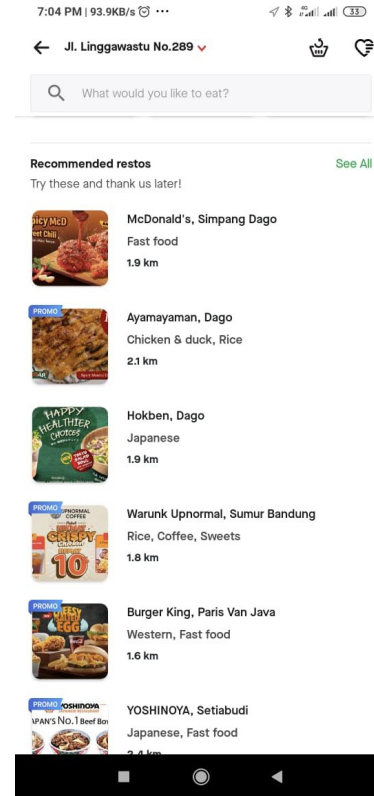
Deep Learning Image Recognition



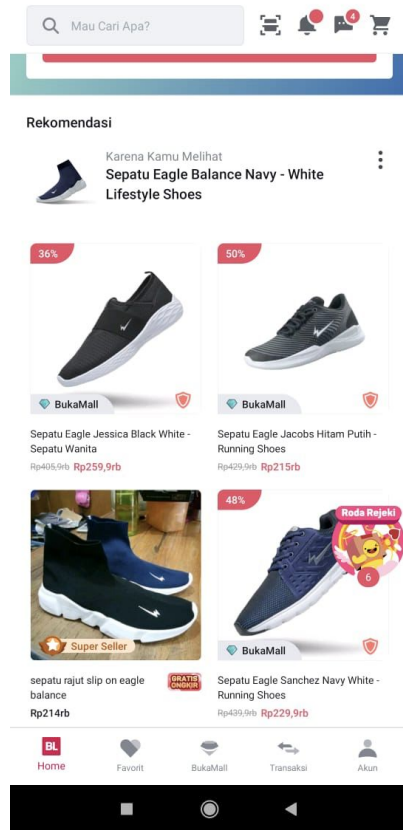
Bang,
Abang Searah
Pulang ga?



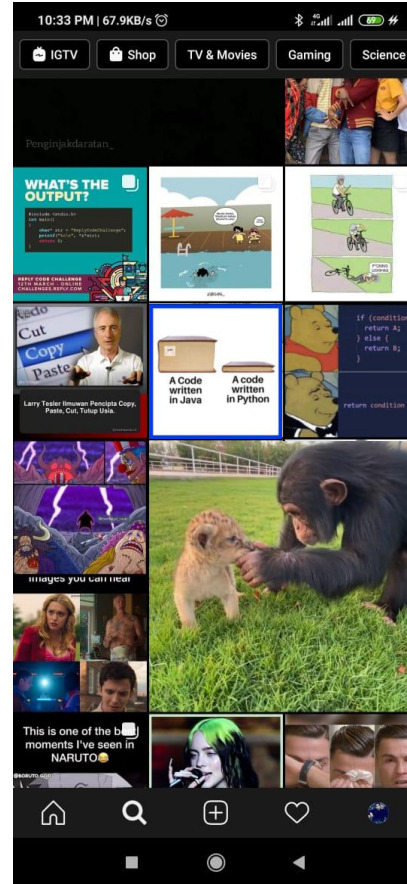
Hmmm Makan
apa ya?



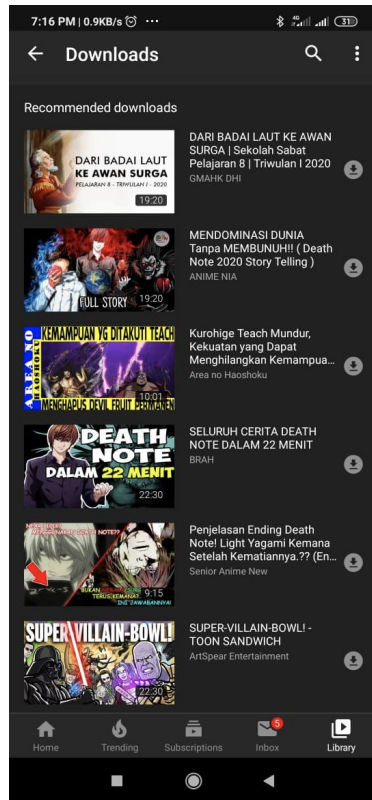
Mau Beli
Sepatu nih



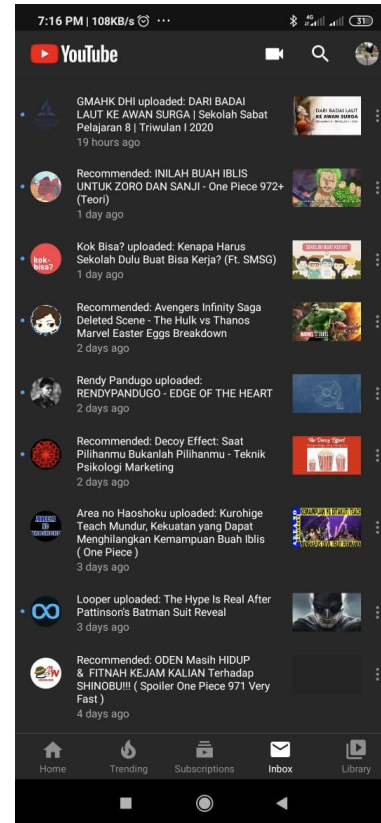
You are what your
Instagram
Discovery
Looks like



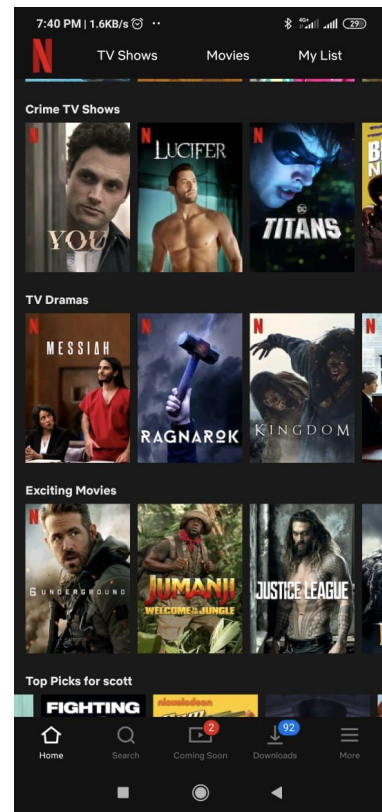
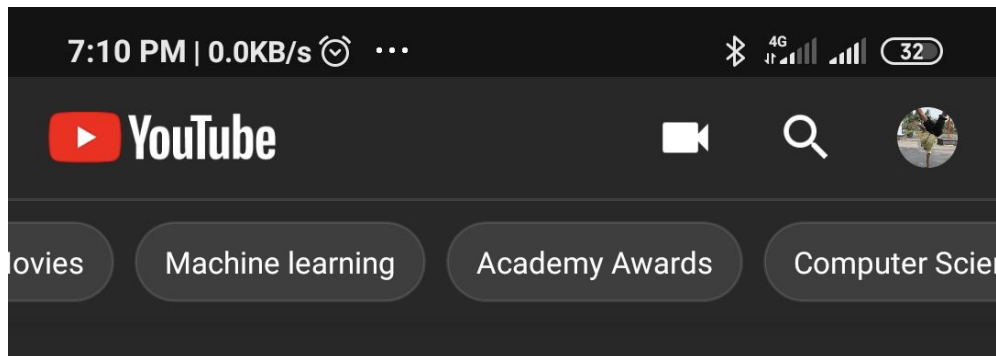
Hmm
Download apa
Ya



Nonton Apa ya?



Genre





Data Science is **NOT** only about Machine Learning!

The Kind of Data Scientist

1. Data science for humans
2. Data science for machines



Data Science for Humans



Data science for humans the consumers of the output are decision makers like executives, product managers, designers, or clinicians.

They want to draw conclusions from data in order to make decisions such as which content to license, which sales lead to follow, which medicine is less likely to cause an allergic reaction, which webpage design will lead to more engagement or more purchases, which marketing email will yield higher revenue, or which specific part of a product user experience is suboptimal and needs attention.

These data scientists design, define, and implement metrics, run and interpret experiments, create dashboards, draw causal inferences, and generate recommendations from modeling and measurement.

Nearly with

Data Analysis

Data Visualization

Data Story Telling

Business Acumen

Presentation

Predict Outcome



Data Science for Machines



Data science for machines: *here the consumers of the output are computers which consume data in the form of training data, models, and algorithms.*

Examples of the work products of these data scientists are: recommendation systems which recommend what shirt a customer might like or what medicine a physician should consider prescribing based on a designed optimization function, such as optimizing for customer clicks or for minimizing readmission rates to the hospital.

Depending on the engineering background of these data scientists, these work products are either deployed directly to the production system, or if they are prototypes they are handed off to software engineers to help implement, optimize and scale them.

Nearly With

Automation Modeling

Artificial Intelligence

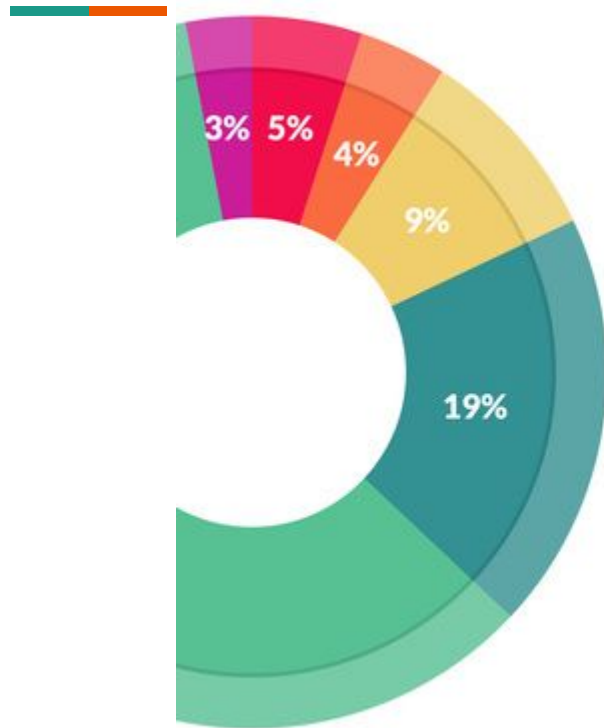
ETL

Data Engineering

Software Engineering

Architecture Optimization





What data scientists spend the most

- *Building training sets: 3%*
- *Cleaning and organizing data: 60%*
- *Collecting data sets; 19%*
- *Mining data for patterns: 9%*
- *Refining algorithms: 4%*
- *Other: 5%*



Why Data Science is Important?

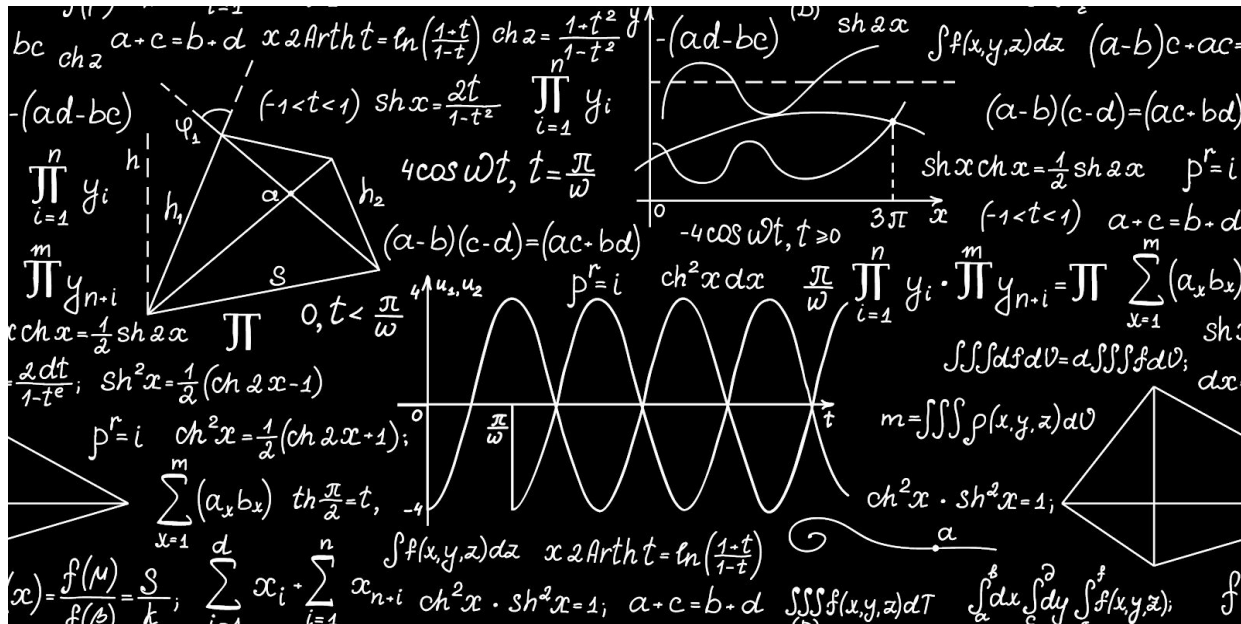
Every business has data but its business value depends on how much they know about the data they have.

Data Science has gained importance in recent times because it can help businesses to increase business value of its available data which in turn can help them to take competitive advantage against their competitors.

It can help us to know our customers better, it can help us to optimize our processes, it can help us to take better decisions. Because of data science, data has become strategic asset.

Aspects in Data Science

Step 1. Statistics, Math, Linear Algebra



Step 2. Programming (Python)



“Data Scientist is a person who is better at statistics than any programmer and better at programming than any statistician.”

Josh Wills -

*Head of Data Engineering at
Slack*





Demo

Lulus Tepat Waktu ga kalian?

<http://bit.ly/demodsunai>



<http://13.229.82.68:1949/check?nim={nimnya}>

<http://13.229.82.68:1949/check?nim=1581009>



Thanks

Follow us

 @femaledailynetwork

Get Female Daily apps also at:



Bern Jonathan
Data Scientist
at Female Daily Network

 Bern Jonathan  bernjonathan

 bernjonath  bernjonathan