

zenius



Kampus
Merdeka
INDONESIA JAYA

Introduction to Data Science

18 April 2022

Data Warehousing, Analysis, and Visualization for
Business Insights

Program Studi Independen Bersertifikat
Zenius Bersama Kampus Merdeka



Quick Intro

Rahadian Rizki Prayoga

Education:

- Sekolah Tinggi Ilmu Statistik, Major : Stats, Minor : Economics

Roles:

- **Data Analytics Lead - Enterprise Wholesale Div., Telkom Indonesia**
- Lead Data Scientist, Mamikos
- Vice Lead Big Data Analytics, Sinarmas Bank Tbk
- Senior Data Scientist, Akseleran
- Guest Lecturer - AI Subject, Universitas Gadjah Mada



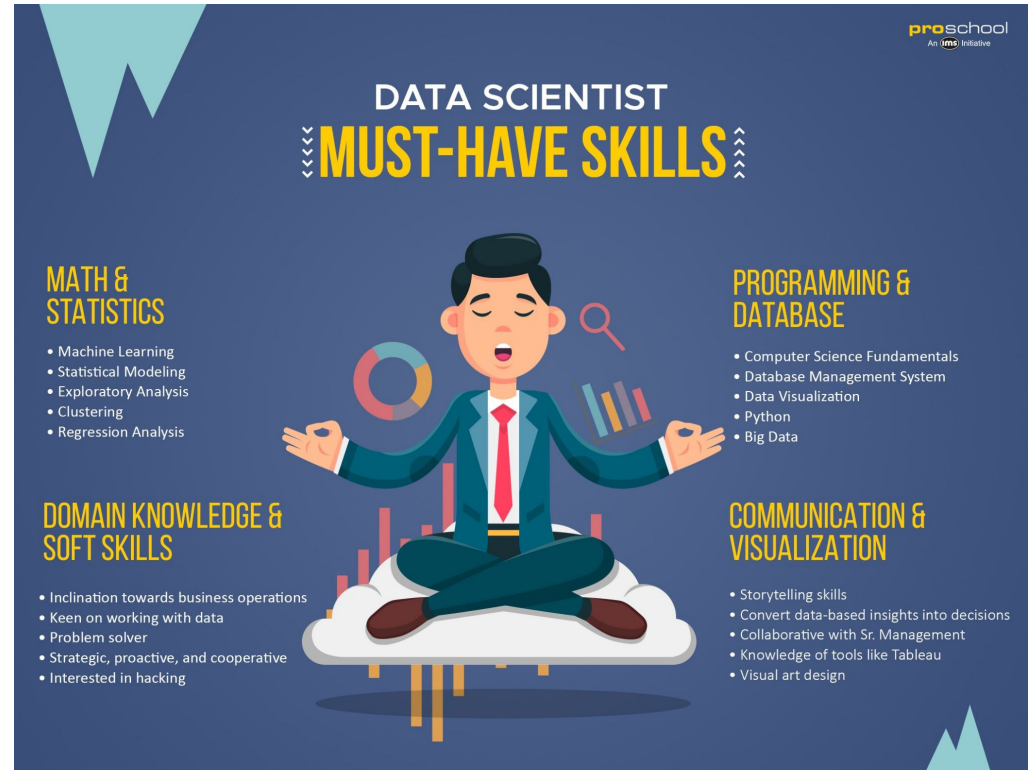
<https://www.linkedin.com/in/rahadianrizki/>

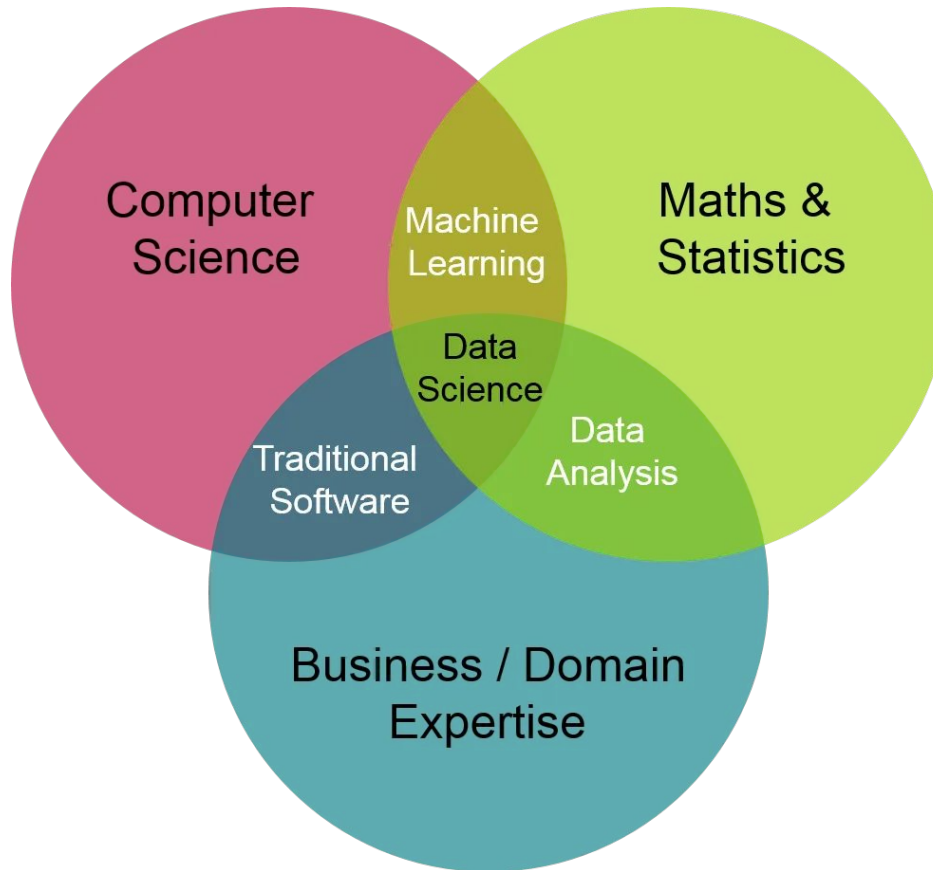
1. **What is Data Science ?**
2. **Why Data Science ?**
3. **Data Product**
4. **Data Science Methodologies**

What is Data Science ?

Data Science ?

Data Scientist ?





SPOTLIGHT ON BIG DATA

Spotlight

ARTWORK Tamar Cohen, Andrew J. Buboltz
2011, silk screen on a page from a high school
yearbook, 8.5" x 12"

Data Scientist: *The Sexiest Job of the 21st Century*

**Meet the people who
can coax treasure out of
messy, unstructured data.**
by Thomas H. Davenport
and D.J. Patil

70 Harvard Business Review October 2012



Source : <https://hbr.org/2012/10/data-scientist-the-sexiest-job-of-the-21st-century>

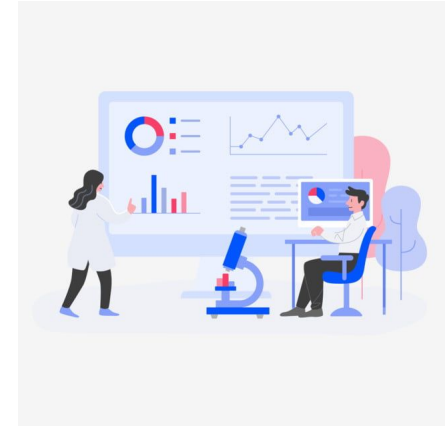
Data science **combines multiple fields**, including **statistics, scientific methods, artificial intelligence (AI), and data analysis**, with an aim **to extract value from data**.

Those **who practice data science are called data scientists**, and they combine a range of skills to analyze data collected from the web, smartphones, customers, sensors, and other sources to derive actionable insights.

So, it could be concluded that :

Data Science : The Field

Data Scientist : The Person



Source : <https://www.oracle.com/data-science/what-is-data-science/>

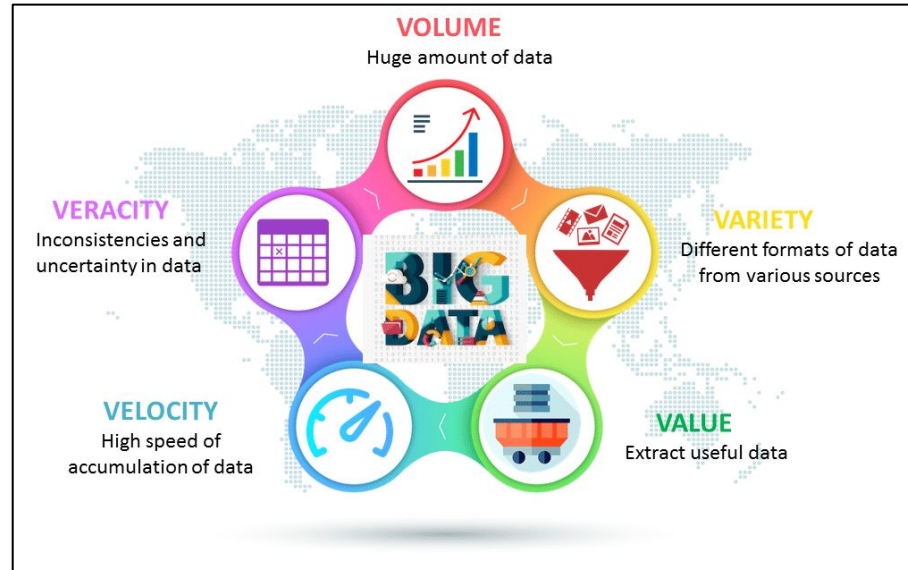
Data science encompasses **preparing data for analysis**, including **cleansing, aggregating, and manipulating** the data to **perform advanced data analysis**. **Analytic applications** and **data scientists** can then **review the results** to **uncover patterns** and **enable business leaders** to **draw informed insights**.



Source : <https://ischoolonline.berkeley.edu/data-science/what-is-data-science/>

Why Data Science ?

Why Data Science ?



Why Data Science ?

2017 This Is What Happens In An Internet Minute



2021 This Is What Happens In An Internet Minute



Why Data Science ?

Data Science enables enterprises to **measure, track, and record performance metrics** for **facilitating enterprise-wide enhanced decision making**.

Companies can analyze trends to make critical decisions to engage customers better, enhance company performance, and increase profitability.

Data Science helps organizations identify and refine target audiences by combining existing data with other data points for developing useful insights. Data Science also helps recruiters by combining data points to identify candidates that best fit their company needs.

Data Science Use Cases



Data Product

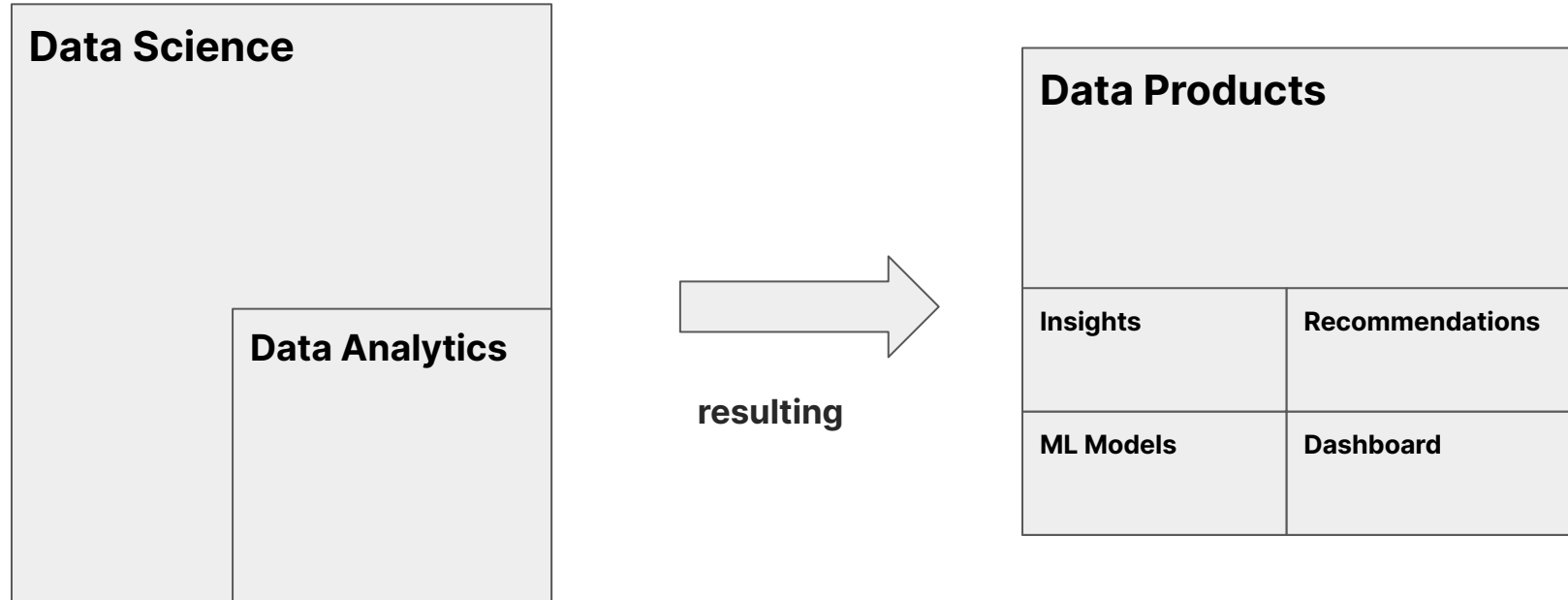
Data Product

DJ Patil, former United States Chief Data Scientist, defined a **data product** as “a product that facilitates an end goal through the use of data”

(from his book Data Jujitsu: The Art of Turning Data into Product, 2012).



We can say that :



Insights

A

Find Jobs

B

Find Jobs

“In this page **Button B 33% is clicked more often** compared to **Button A**”



>



“In our website, **a product that has no background 40% sold more than ones with background**”

A/B
Testing



“In our website, **automotive and electronics category** is visited more by **men**. On the other hand, **baby products and beauty** is visited more by **women**”

Analysis

Quoting a Data Scientist from a unicorn startup on 2019 :

“We split 2 type of data products into **machine analytics** and **human analytics**.”

Machine Analytics

Type of data products that relies dominantly on the usage of unstructured data, and the form of them frequently delivered as an application and machine is the pure decision maker

Example :

- Image Recognition
- Product Recommendation
- Fraud Detection

Human Analytics

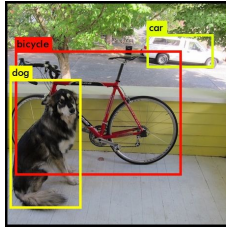
Type of data products that combines both the balance of unstructured and structured data to find insights and human will act as the decision maker

Example :

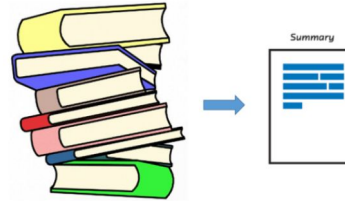
- Demand Forecasting
- Customer Lifetime Behavior Analysis

Data Product Example

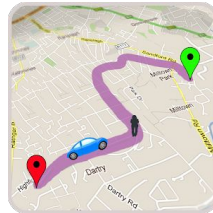
**Image
Recognition**



Text Mining



**Route
Optimizer**



Search Engine

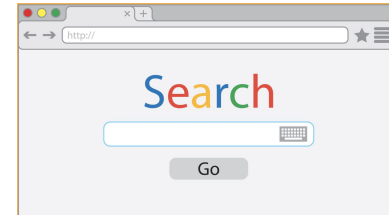


Image Recognition

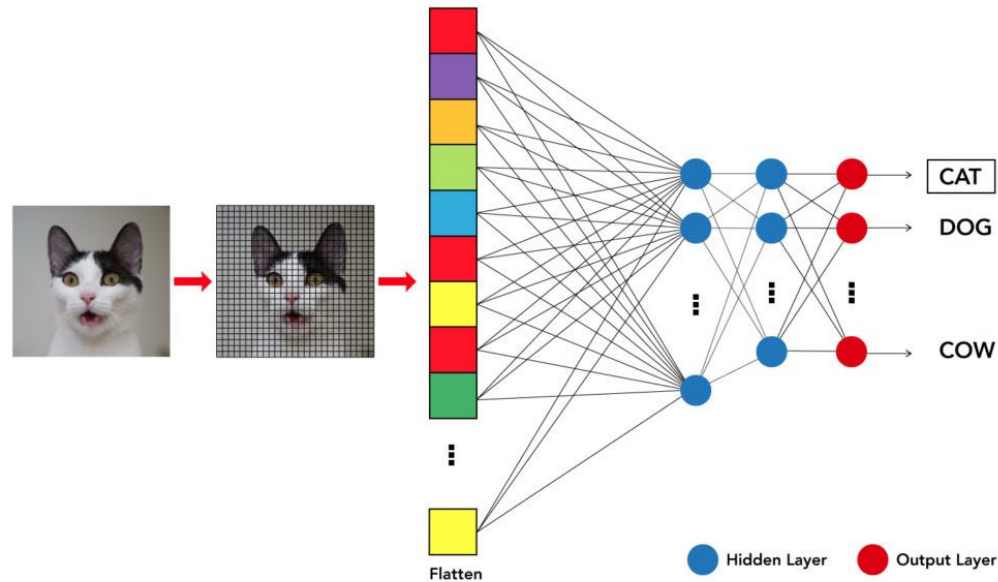
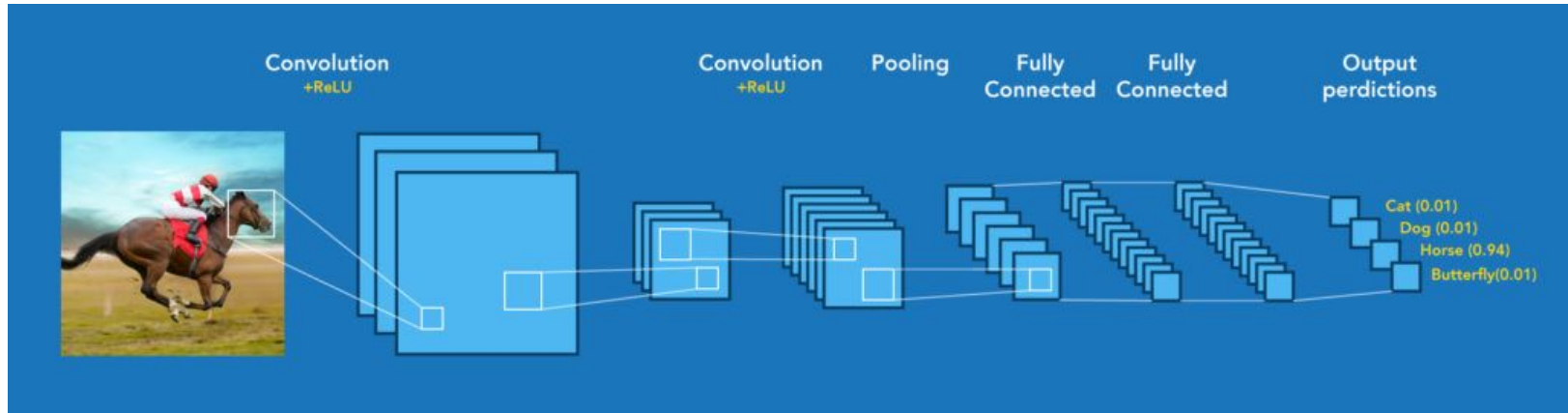
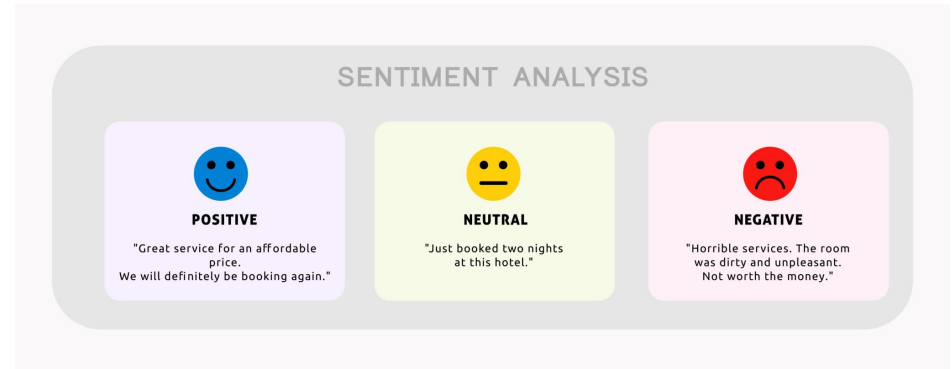
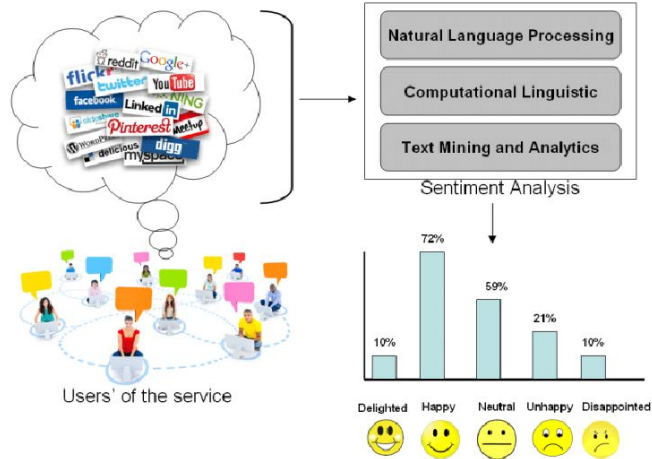


Image Recognition



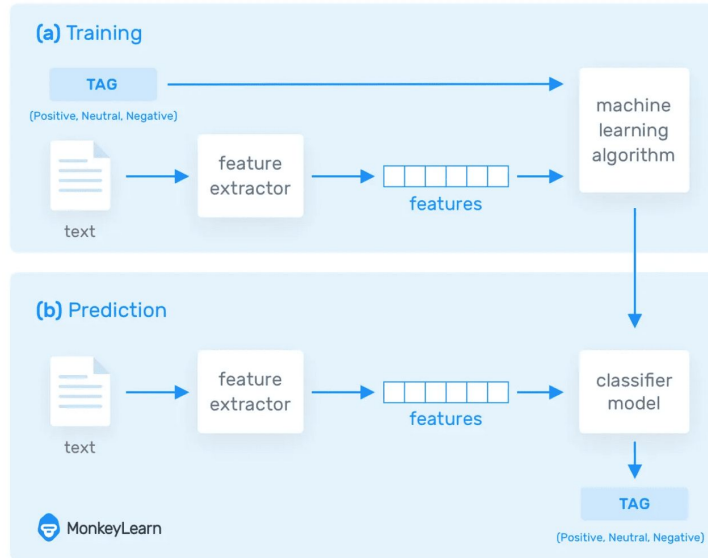
Text Mining

Sentiment Analysis



Text Mining

Sentiment Analysis

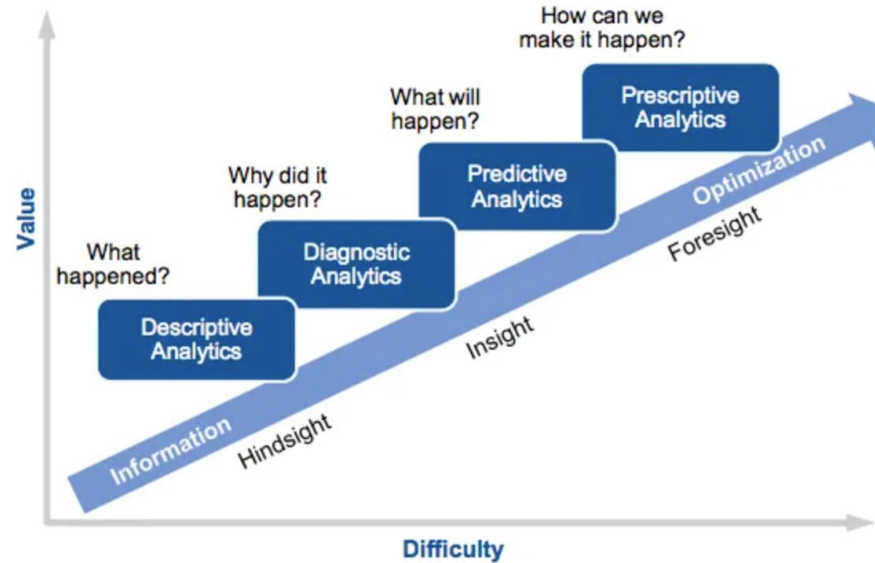


Source : <https://monkeylearn.com/blog/sentiment-analysis-machine-learning/>

Pop Quiz! Can you mention one of data science use cases around us?

Data Science Methodologies

Type Of Analytics



Source : Gartner Analytics Ascendancy Model

<https://www.clickz.com/how-can-ai-allow-marketers-to-predict-the-future/112268/gartner-analytic-ascendancy-model/>

<https://www.gartner.com/en/topics/data-and-analytics>

```
graph TD; BU[Business Understanding] <--> DU[Data Understanding]; DU --> DP[Data Preparation]; DP <--> M[Modeling]; M --> E[Evaluation]; E --> D[Deployment]; D --> BU; D((Data))
```

The diagram illustrates the Machine Learning lifecycle as a continuous, iterative process. It is represented by a large circle with a thick grey arrow indicating a clockwise flow. The cycle consists of six main stages, each in a blue rounded rectangle:

- Business Understanding**: The starting point, connected to Data Understanding by a double-headed arrow.
- Data Understanding**: Connected to Business Understanding and Data Preparation.
- Data Preparation**: Connected to Data Understanding and Modeling by a double-headed arrow.
- Modeling**: Connected to Data Preparation and Evaluation.
- Evaluation**: Connected to Modeling and Deployment.
- Deployment**: The final stage, connected to Evaluation and back to Business Understanding, completing the cycle.

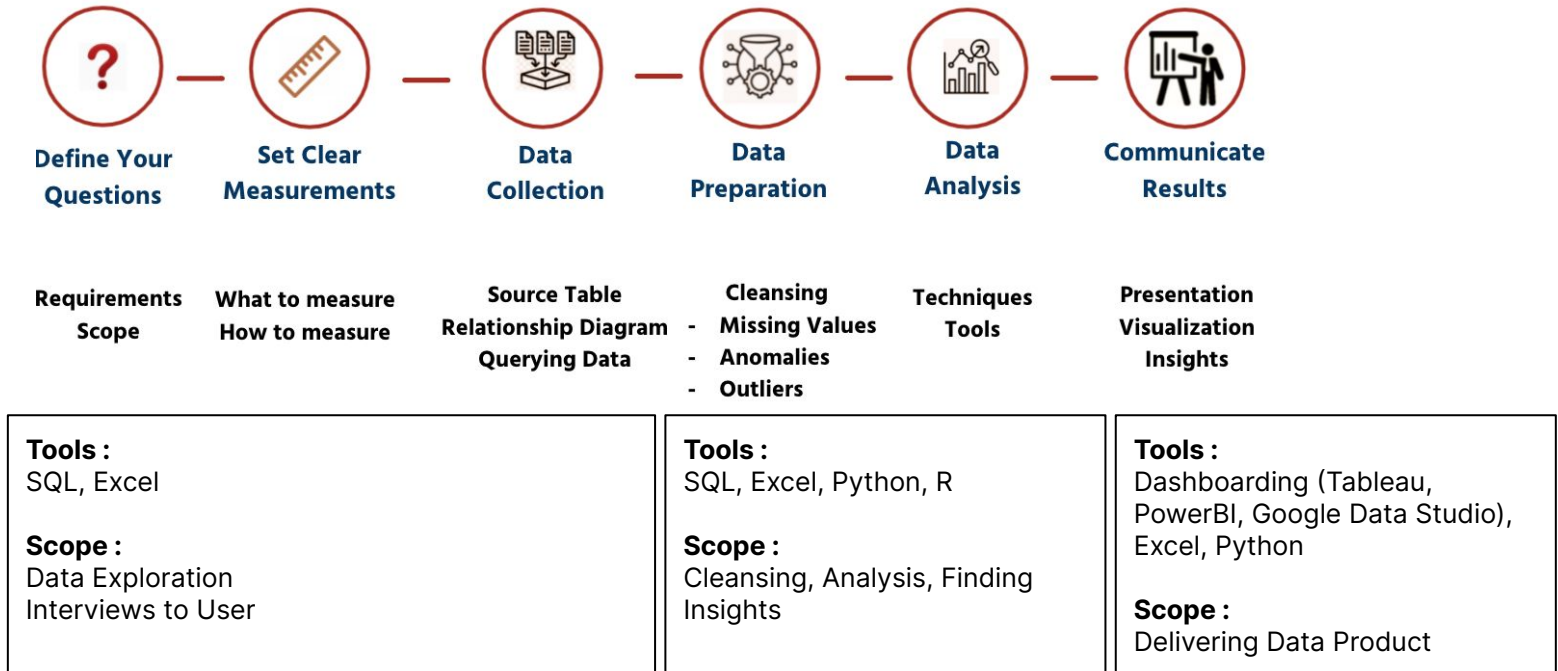
In the center of the cycle is a database icon (three stacked cylinders) labeled **Data**. A curved arrow points from this central Data icon towards the Evaluation stage, indicating that data is used for evaluation.

<https://developer.ibm.com/articles/introduction-watson-studio/>

CRISP-DM (Detail)

Business Understanding	Data Understanding	Data Preparation	Modeling	Evaluation	Deployment
Determine Business Objectives Background Business Objectives Business Success Criteria	Collect Initial Data <i>Initial Data Collection Report</i> Describe Data <i>Data Description Report</i> Explore Data <i>Data Exploration Report</i> Verify Data Quality <i>Data Quality Report</i>	Select Data <i>Rationale for Inclusion/Exclusion</i> Clean Data <i>Data Cleaning Report</i> Construct Data <i>Derived Attributes Generated Records</i> Integrate Data <i>Merged Data</i> Format Data <i>Reformatted Data</i> <i>Dataset</i> <i>Dataset Description</i>	Select Modeling Techniques <i>Modeling Technique Modeling Assumptions</i> Generate Test Design <i>Test Design</i> Build Model <i>Parameter Settings Models Model Descriptions</i> Assess Model <i>Model Assessment Revised Parameter Settings</i>	Evaluate Results <i>Assessment of Data Mining Results w.r.t. Business Success Criteria</i> <i>Approved Models</i> Review Process <i>Review of Process</i> Determine Next Steps <i>List of Possible Actions Decision</i>	Plan Deployment <i>Deployment Plan</i> Plan Monitoring and Maintenance <i>Monitoring and Maintenance Plan</i> Produce Final Report <i>Final Report Final Presentation</i> Review Project <i>Experience Documentation</i>
Assess Situation <i>Inventory of Resources Requirements, Assumptions, and Constraints Risks and Contingencies Terminology Costs and Benefits</i>					
Determine Data Mining Goals <i>Data Mining Goals Data Mining Success Criteria</i>					
Produce Project Plan <i>Project Plan Initial Assessment of Tools and Techniques</i>					

Data Analytics Workflow



Thanks!
Any Questions?

zenius



Kampus
Merdeka
INDONESIA JAYA

