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

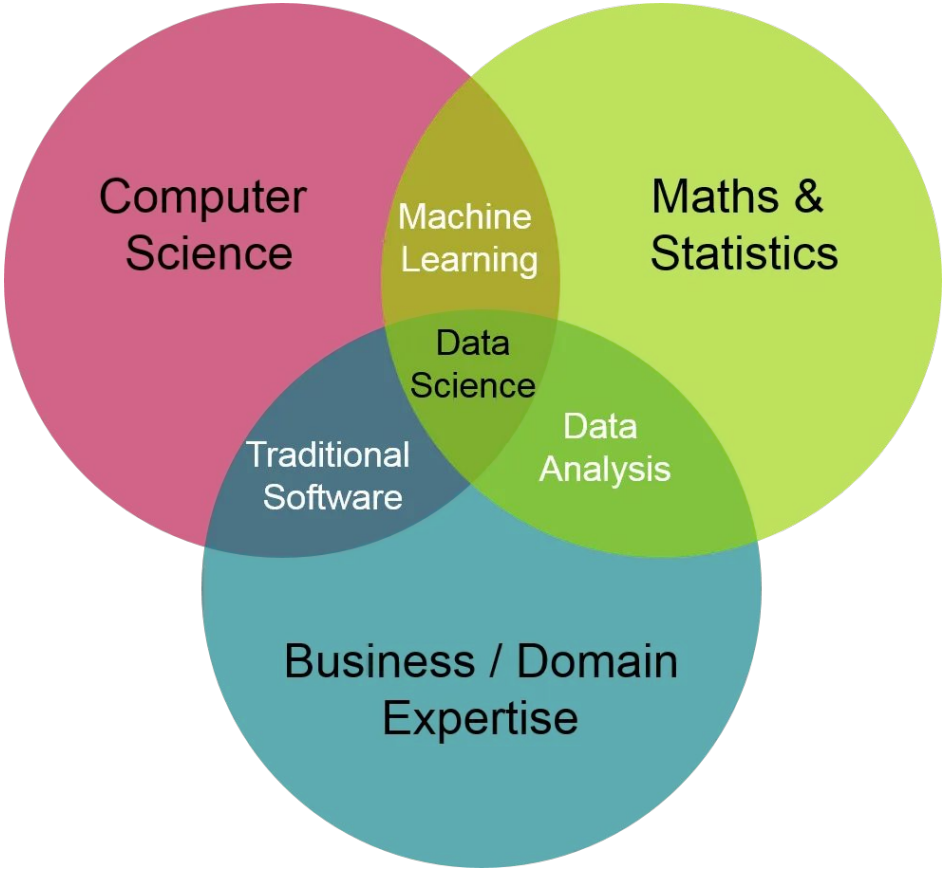
https://www.linkedin.com/in/rahadianrizki/

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.

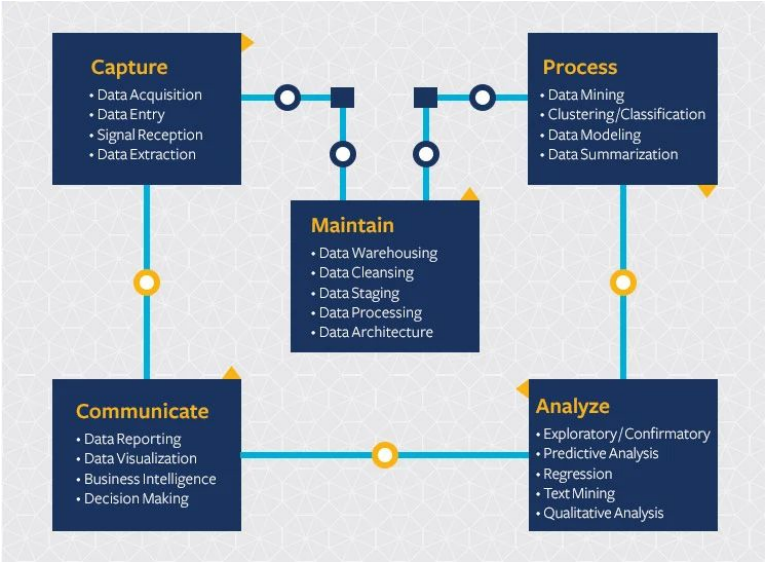
So, it could be concluded that :

Data Science : The Field

Data Scientist : The Person

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

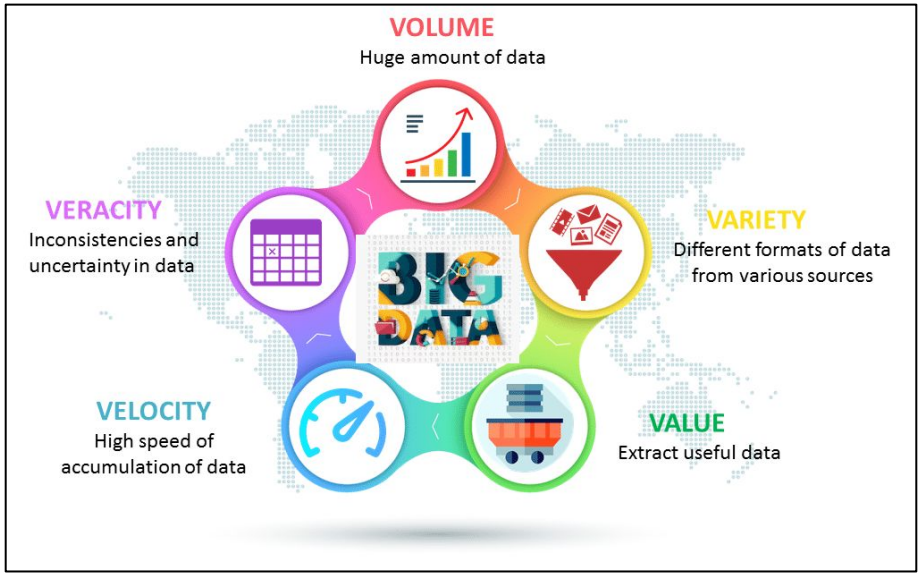
Data science encompasses **preparing data for analysis**, including **cleansing**, **aggregating**, and **manipulating** the data to p**erform advanced data analysis**. **Analytic applications** and **data scientists** can then **review the results** to **uncover patterns** and **enable business leaders**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



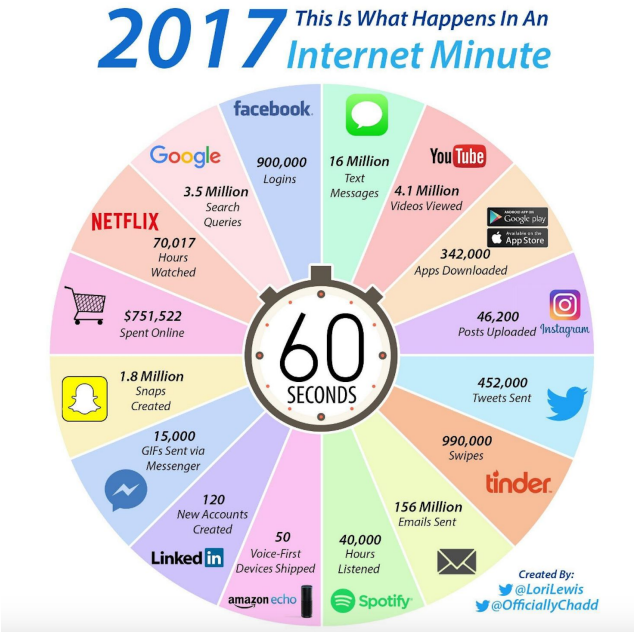
Source : https://ischoolonline.berkeley.edu/data-science/what-is-data-science/

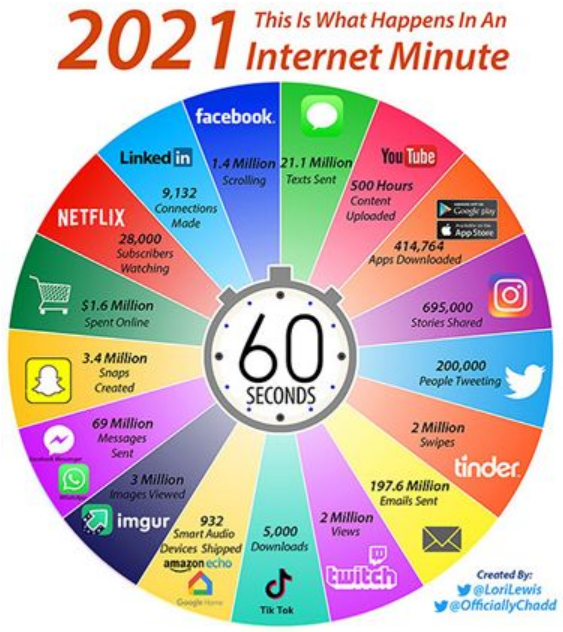
Why Data Science ?

Why Data Science ?



Why Data Science ?





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.

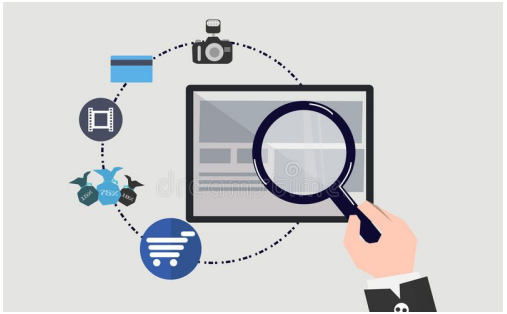
Data Science Use Cases

Data Product

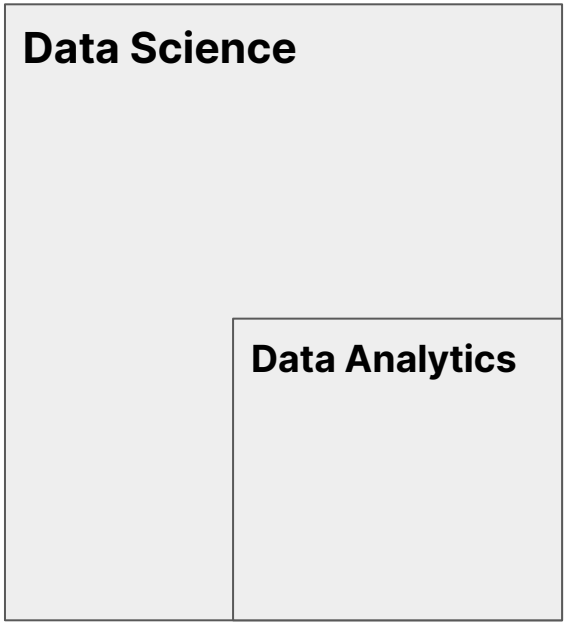
Data Product

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).



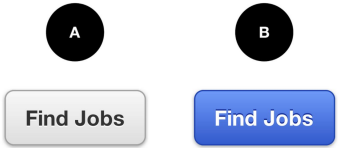
We can say that :



resulting

Data Products	
Insights	Recommendations
ML Models	Dashboard

Insights



"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

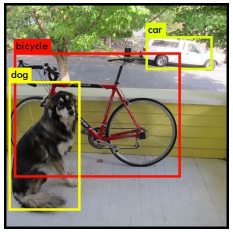
Type of data products that combines both the balance of unstructured and structured data to find insights and human will act as the decision maker

Example :

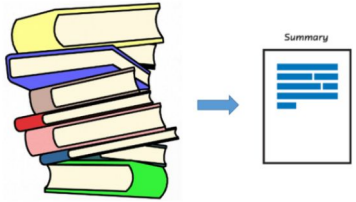
- Demand Forecasting
- Customer Lifetime Behavior Analysis

Data Product Example

Image Recognition



Text Mining



Route Optimizer

Search Engine

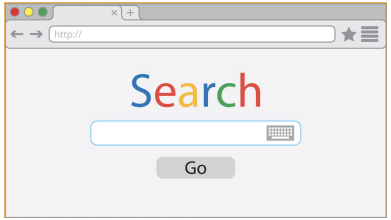


Image Recognition

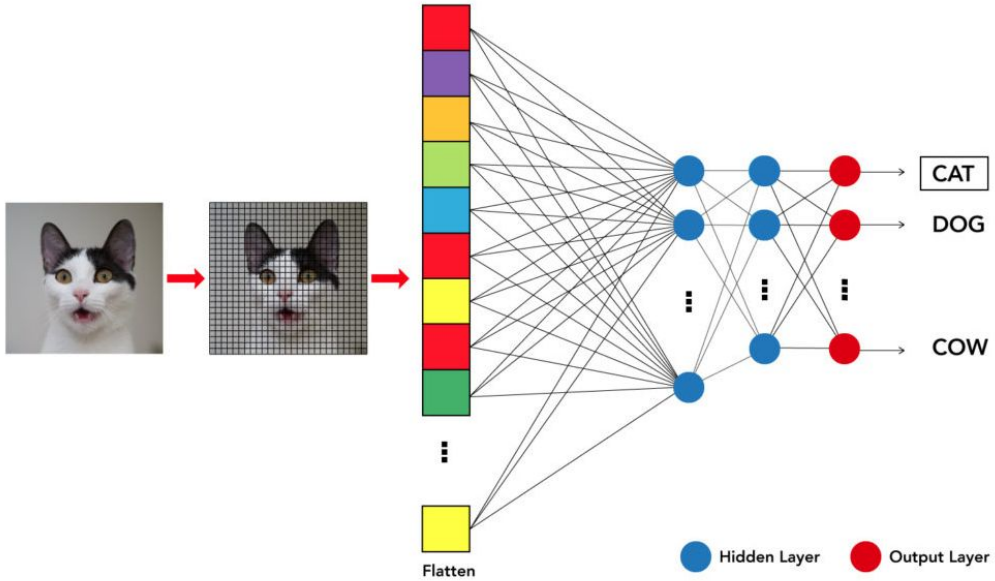
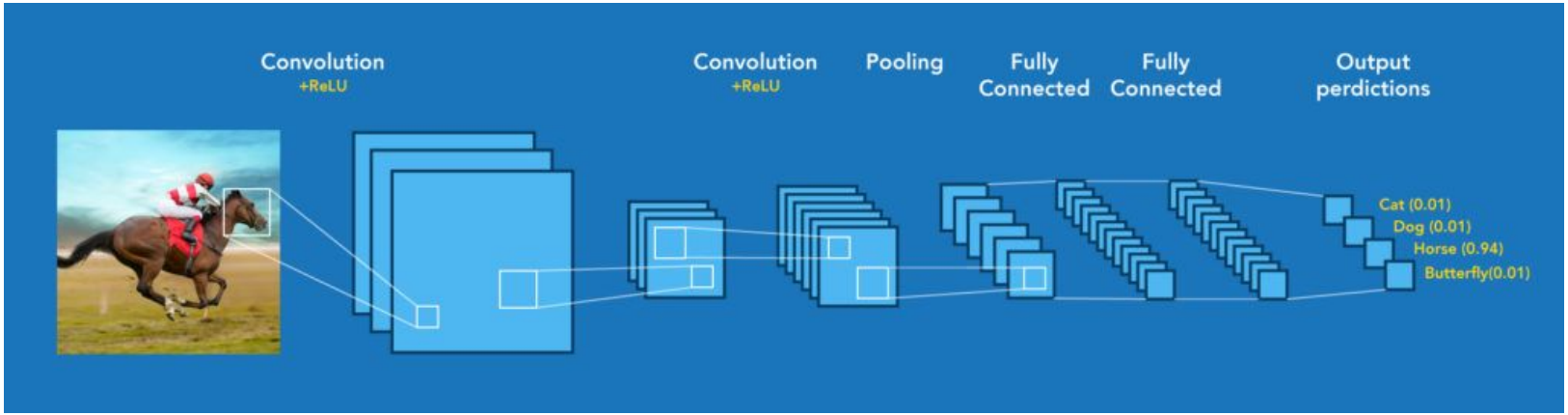
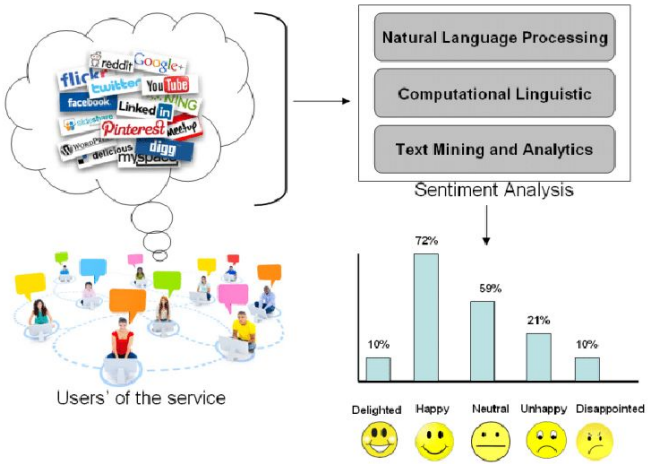


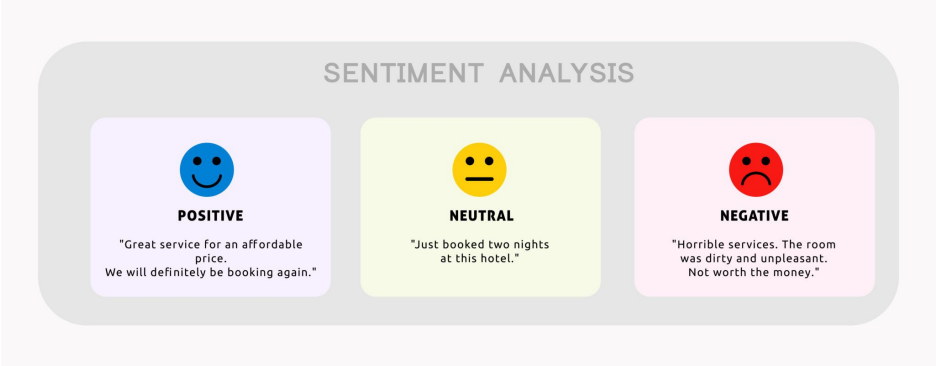
Image Recognition



Text Mining

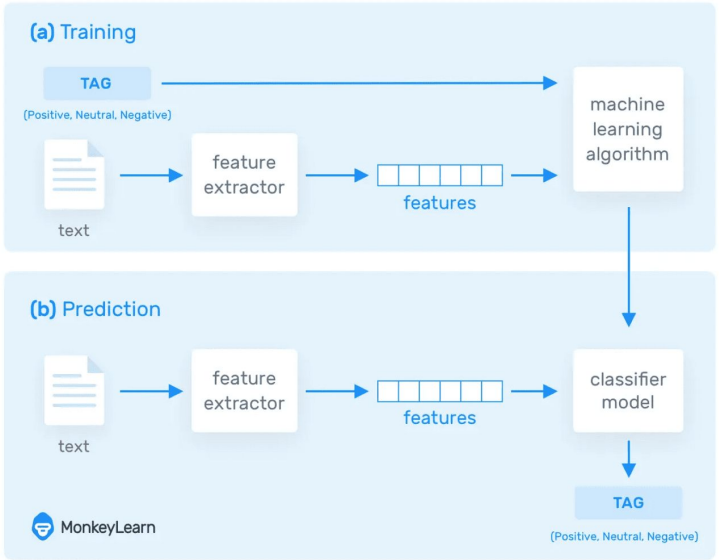
Sentiment Analysis





Text Mining

Sentiment Analysis

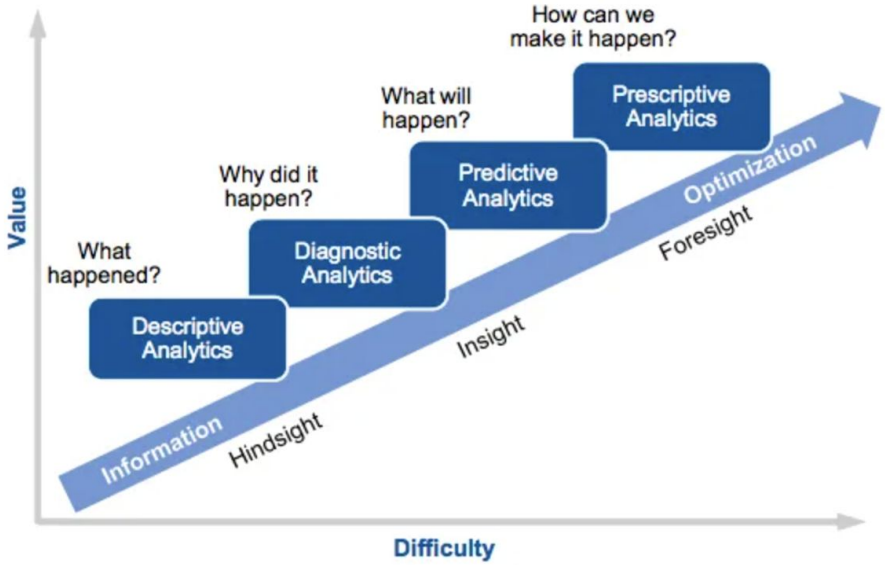


Source : https://monkeylearn.com/blog/sentiment-analysis-machine-learning/

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

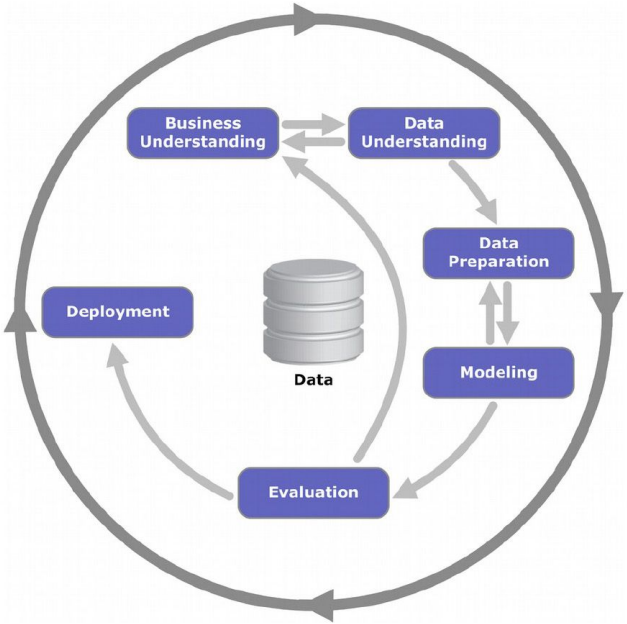
Data Science Methodologies

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

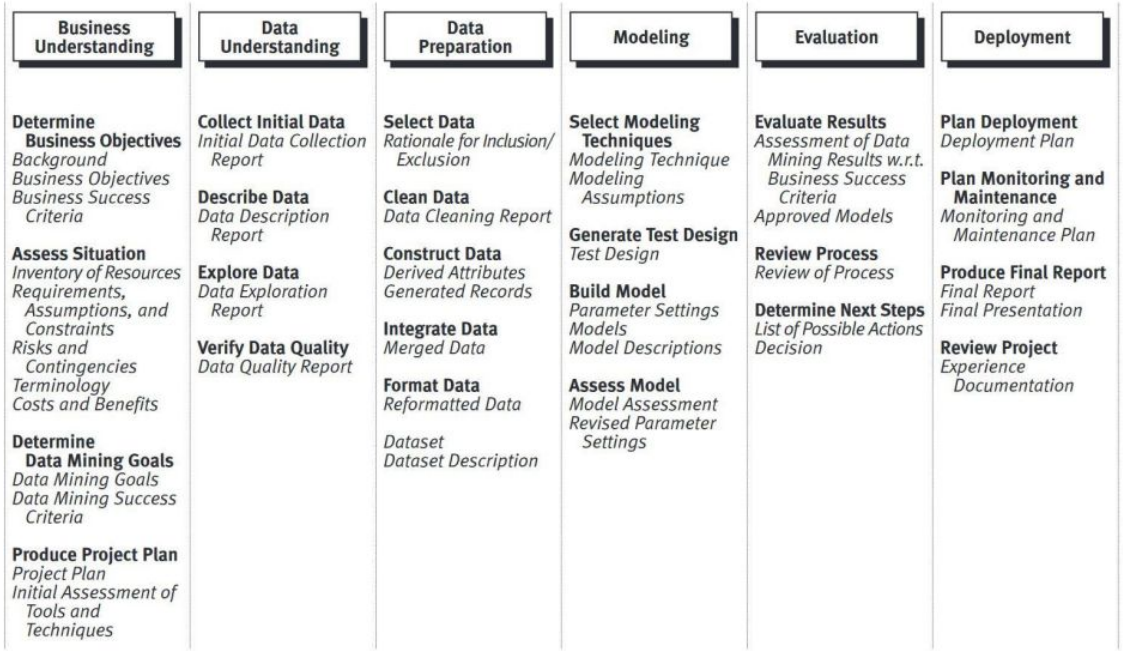
CRISP-DM



Source : CRISP-DM on IBM Watson

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

CRISP-DM (Detail)



Data Analytics Workflow

Tools : SQL, Excel **Scope :** Data Exploration Interviews to User	**Tools :** SQL, Excel, Python, R **Scope :** Cleansing, Analysis, Finding Insights	**Tools :** Dashboarding (Tableau, PowerBI, Google Data Studio), Excel, Python **Scope :** Delivering Data Product

Thanks!
Any Questions?