

A large, solid pink circle is centered on the slide, serving as a background for the main title text.

DATA SCIENCE BOOTCAMP

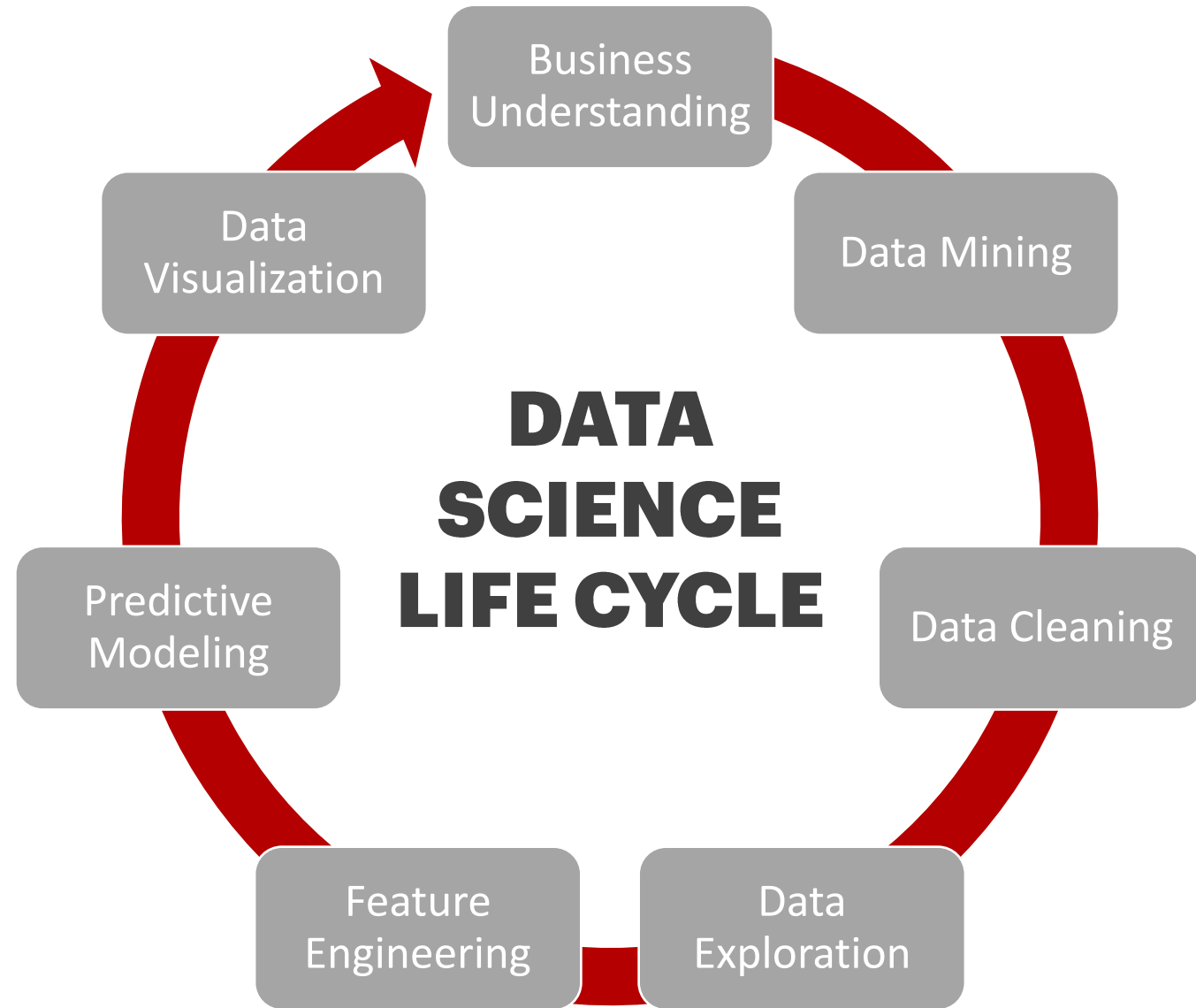
Session 2: Data Science Methodology and Python Syntax

SESSION 2 OUTLINE

- **DATA SCIENCE METHODOLOGY**
- **PYTHON SYNTAX**

SESSION 2 OUTLINE

- **DATA SCIENCE METHODOLOGY**
- **PYTHON SYNTAX**



DATA SCIENCE ROLES

Data Engineer

Key Technical Skills

- ETL/ELT Development (Informatica, Talend, Python)
- Data Management (SQL, Spark, Hadoop)
- Data Orchestration (Python, platform-specific services)
- Data Pipeline Deployment, Monitoring & Maintenance
- Business Intelligence and Data Visualization (Tableau, Power BI, Qlik)

Machine Learning Engineer

Key Technical Skills

- Statistical Programming/Data Mining (R, Python)
- Machine Learning (R, Python)
- Model Deployment, Monitoring & Maintenance

Data Scientist

Key Technical Skills

- Statistical Programming/Data Mining (R, Python)
- Machine Learning (R, Python)
- Statistics, Applied Mathematics

DATA SCIENCE **PROCESS**



DATA ENGINEERING



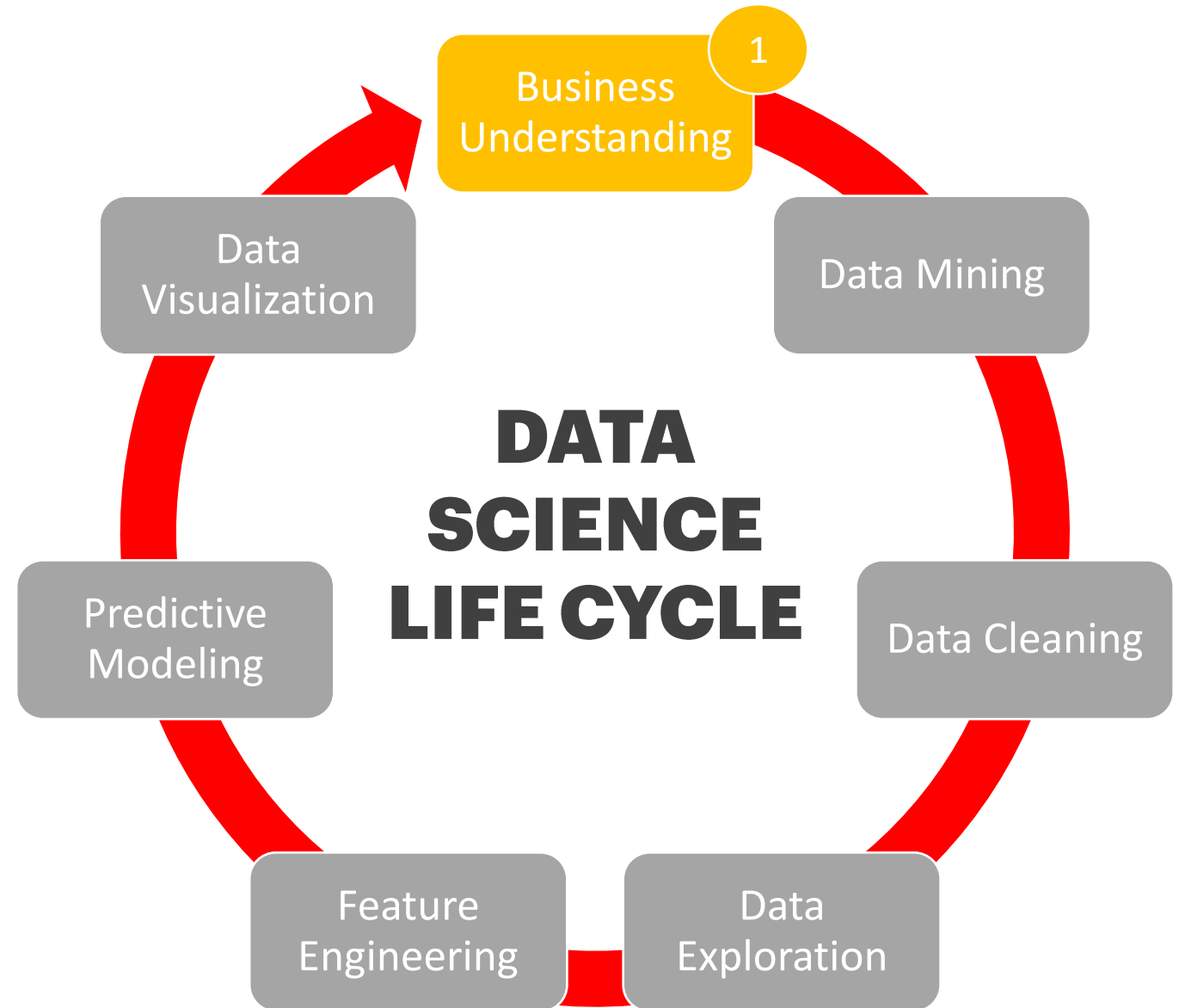
COLLABORATION



ML DEVELOPMENT

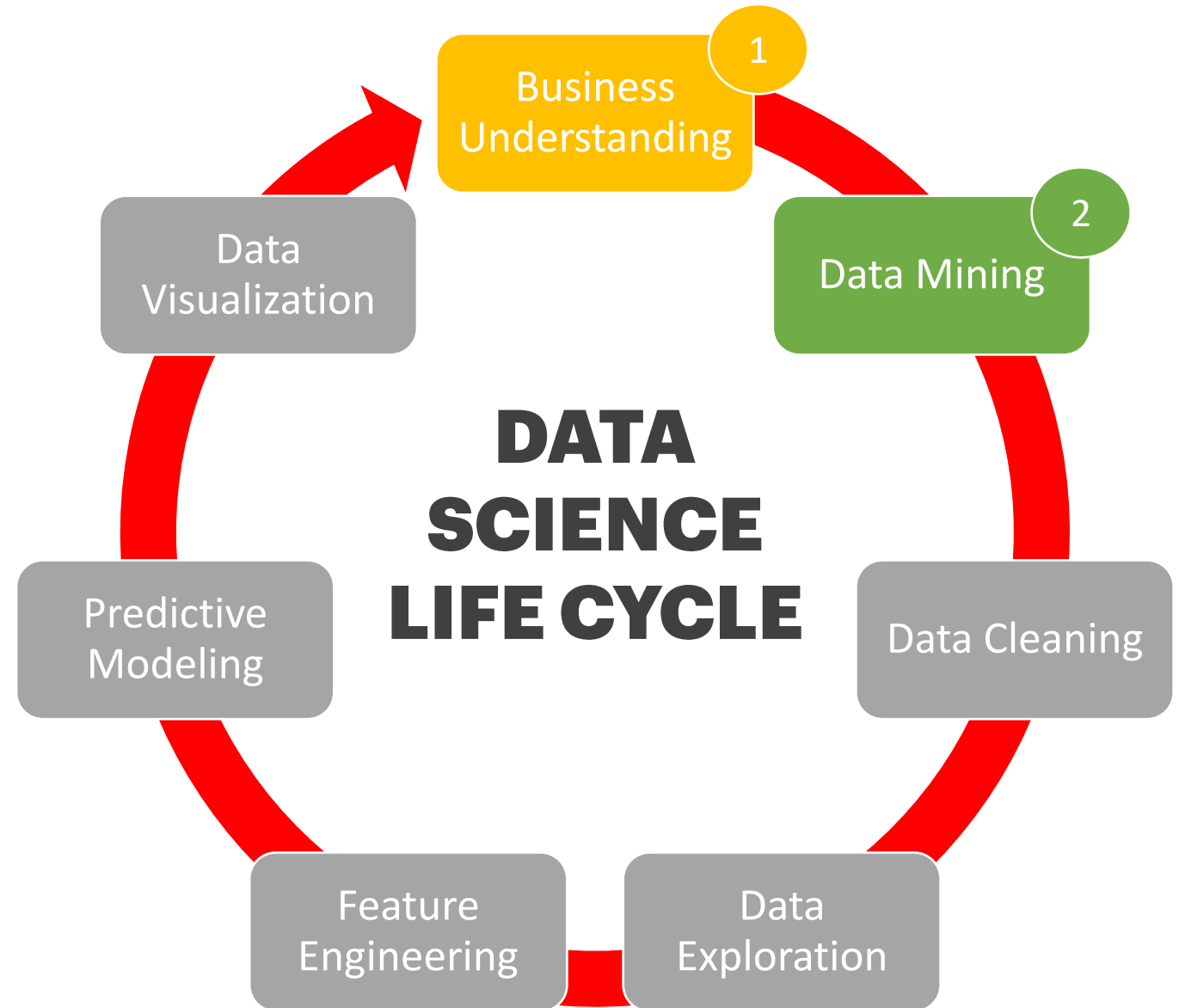
BUSINESS UNDERSTANDING

Ensures that every decision made is supported by concrete data and that it is guaranteed (with a high probability) to achieve results.



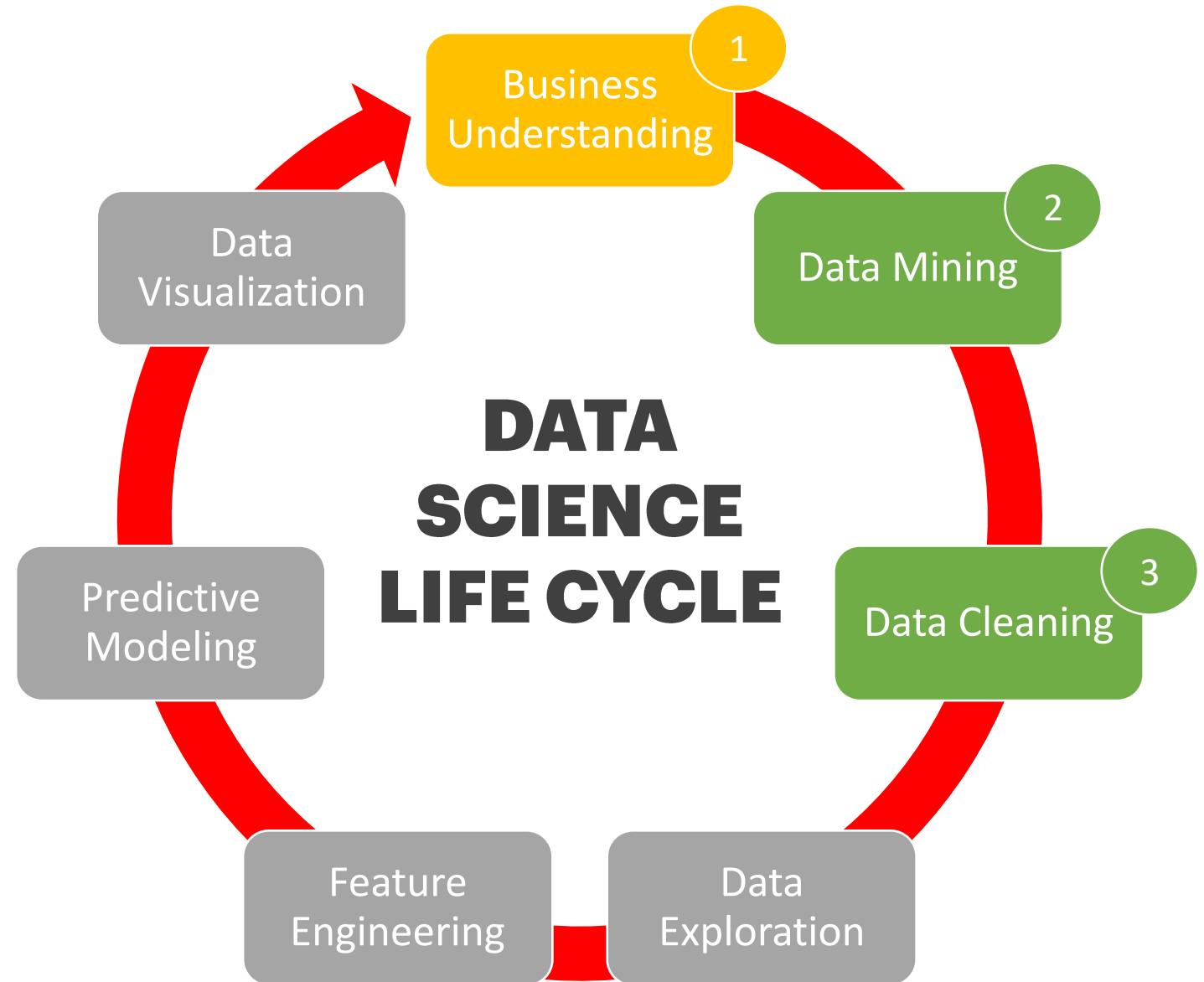
DATA MINING

Data mining is the process of gathering data from different sources.



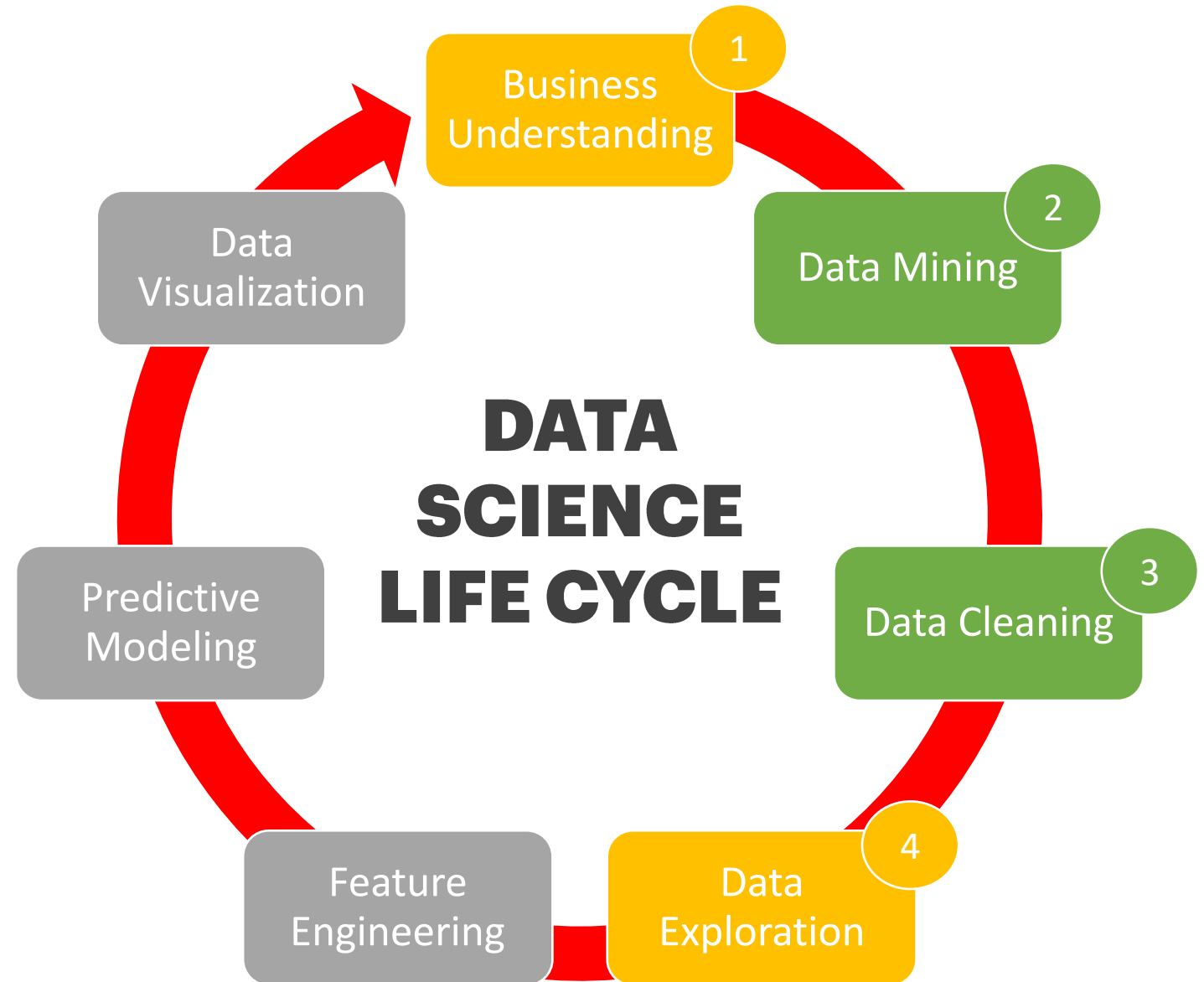
DATA CLEANING

Most time-consuming step of all.



DATA EXPLORATION

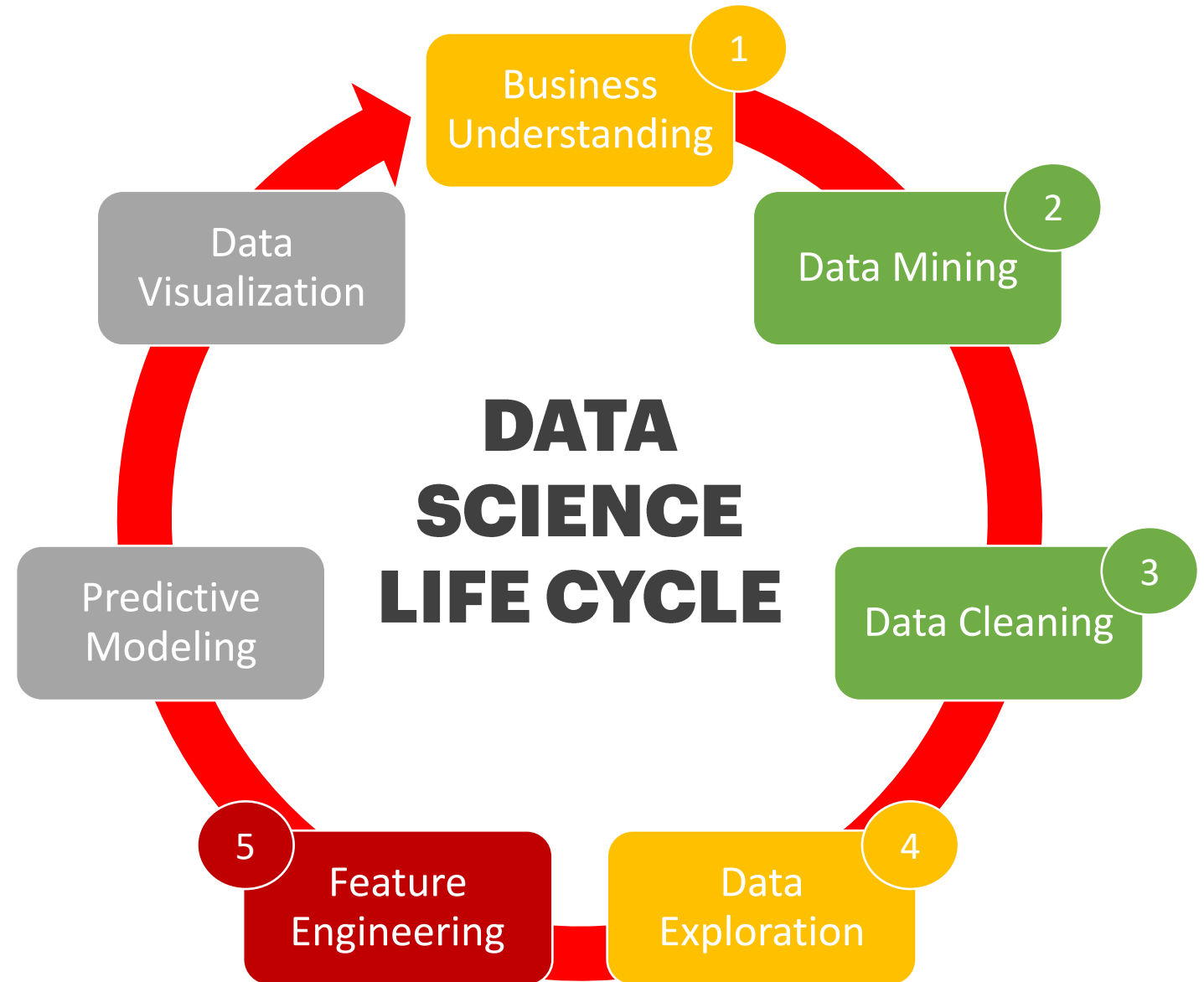
This stage is like the brainstorming of data analysis.



FEATURE ENGINEERING

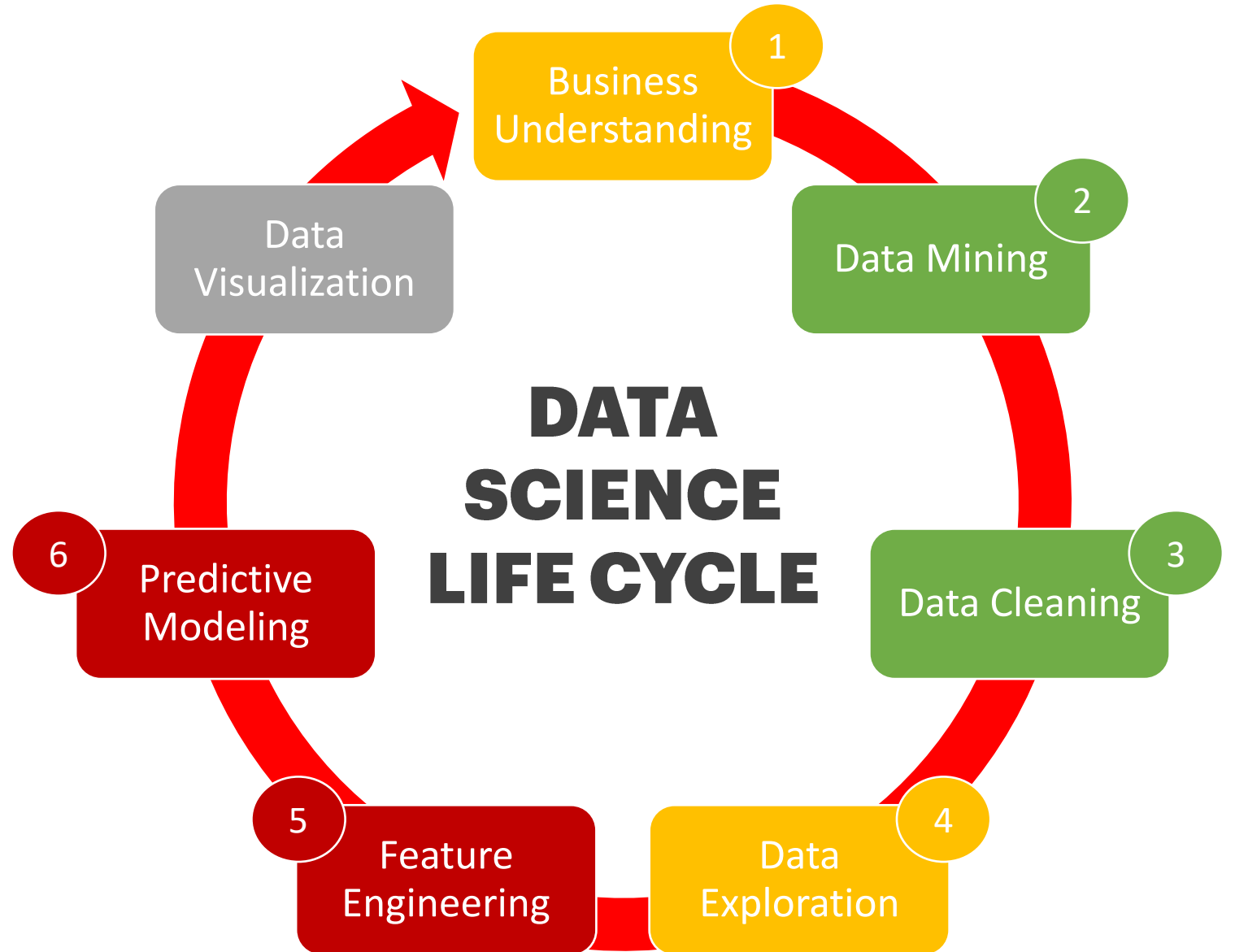
In machine learning, a feature is a measurable property or attribute of a phenomenon being observed.

Two tasks: Feature selection and Feature engineering



PREDICTIVE MODELING

This is where machine learning finally comes into the project.



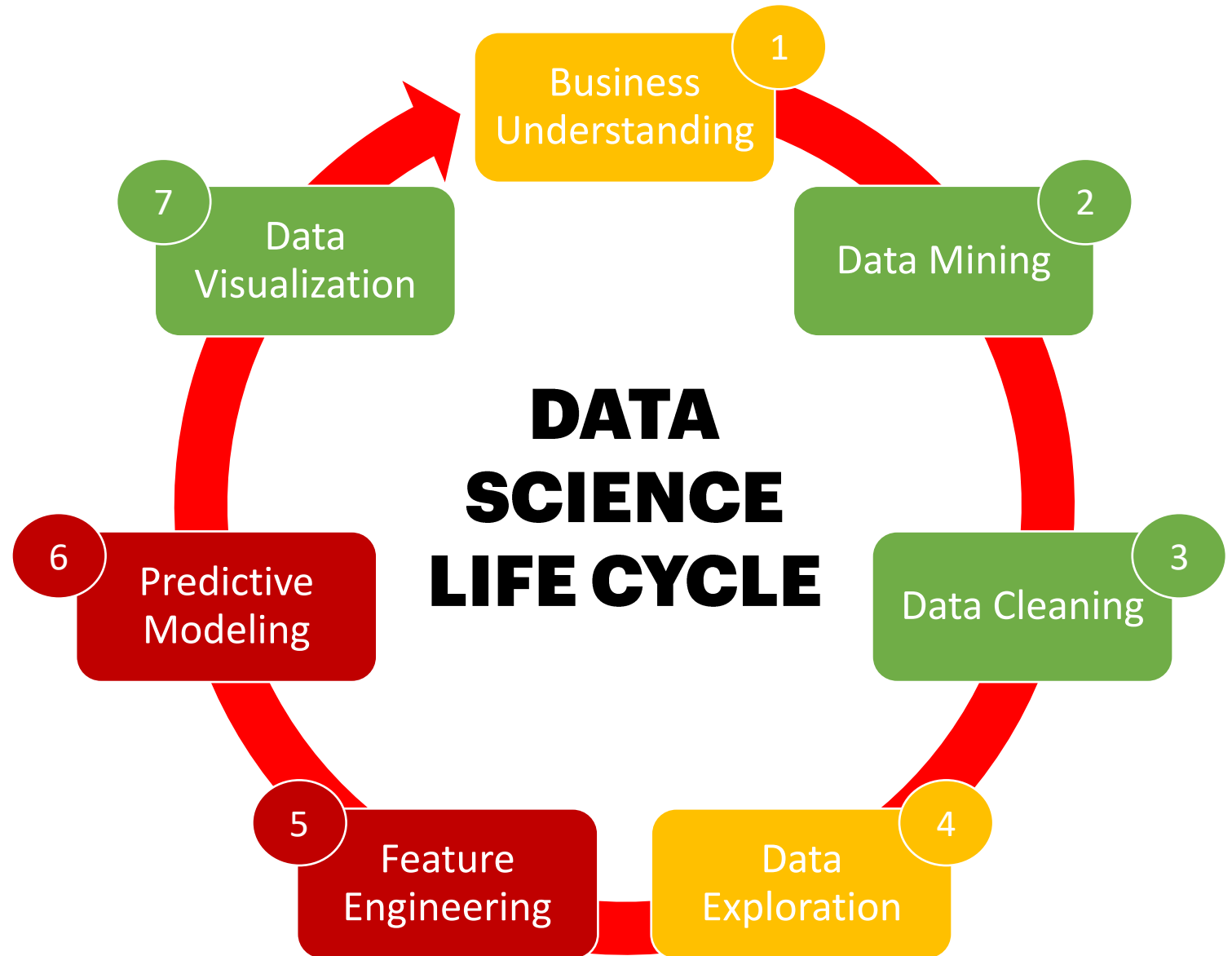
 **DATA ENGINEERING**

 **COLLABORATION**

 **ML DEVELOPMENT**

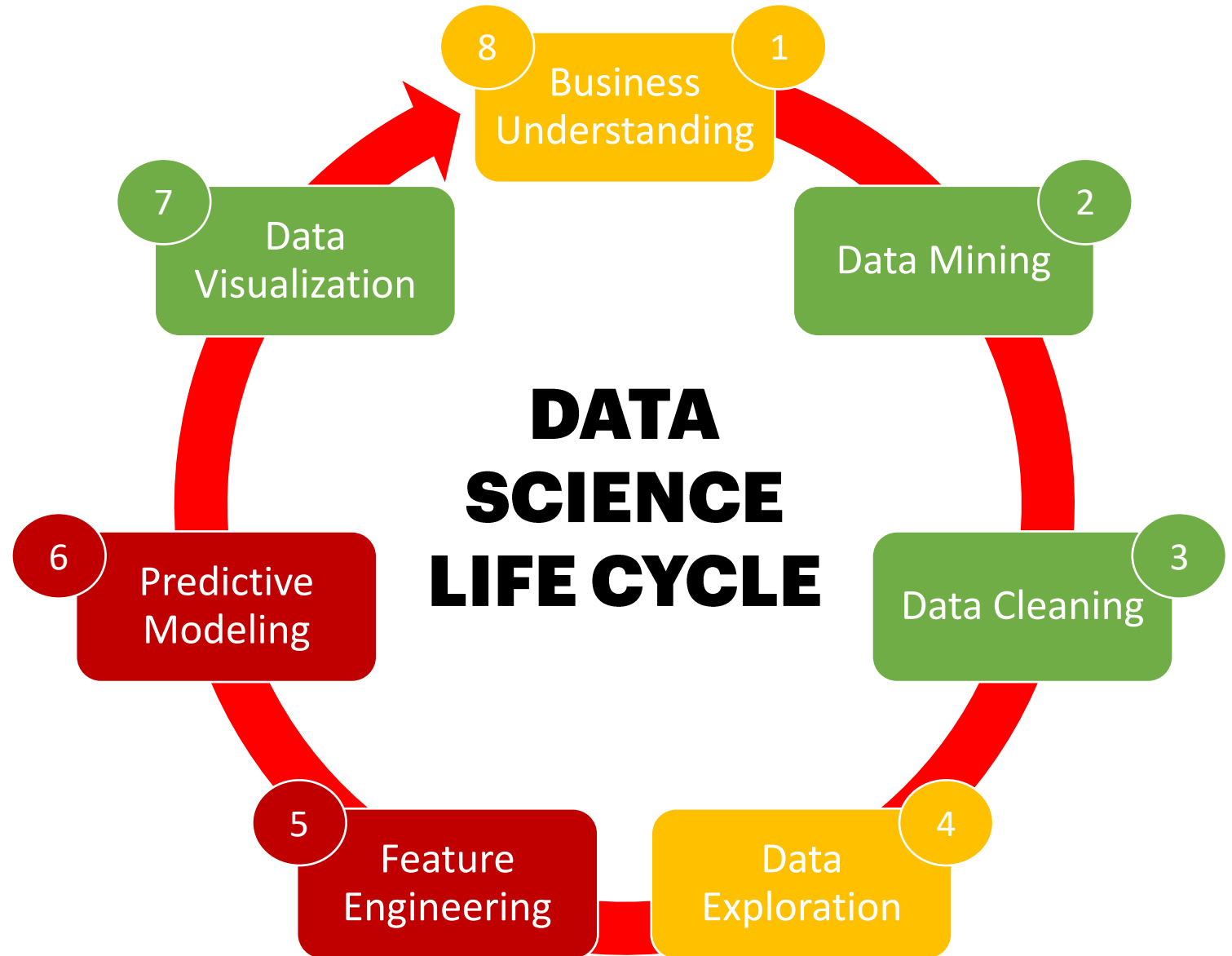
DATA VISUALIZATION

This is where we derived the intended insights from the model.

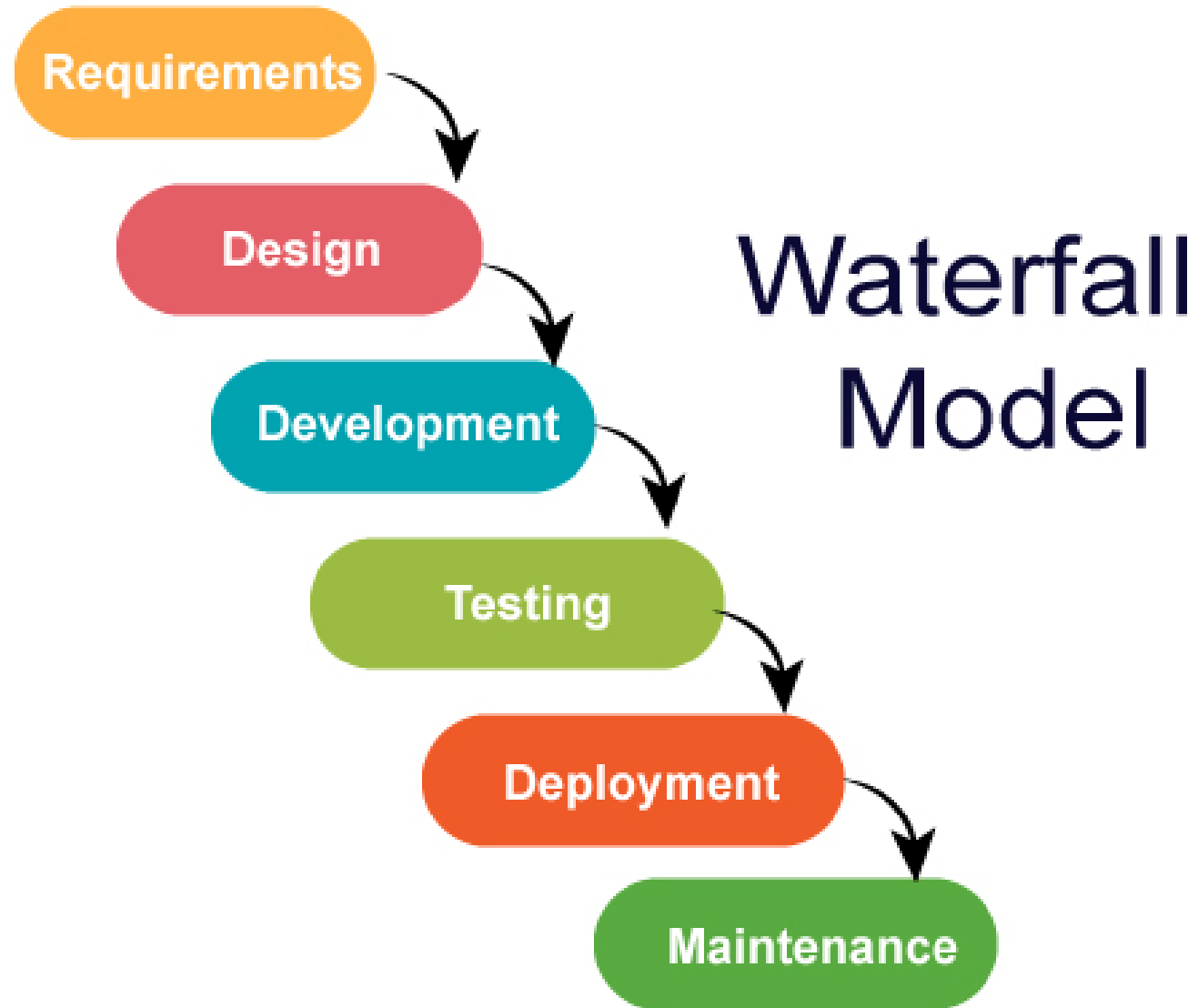


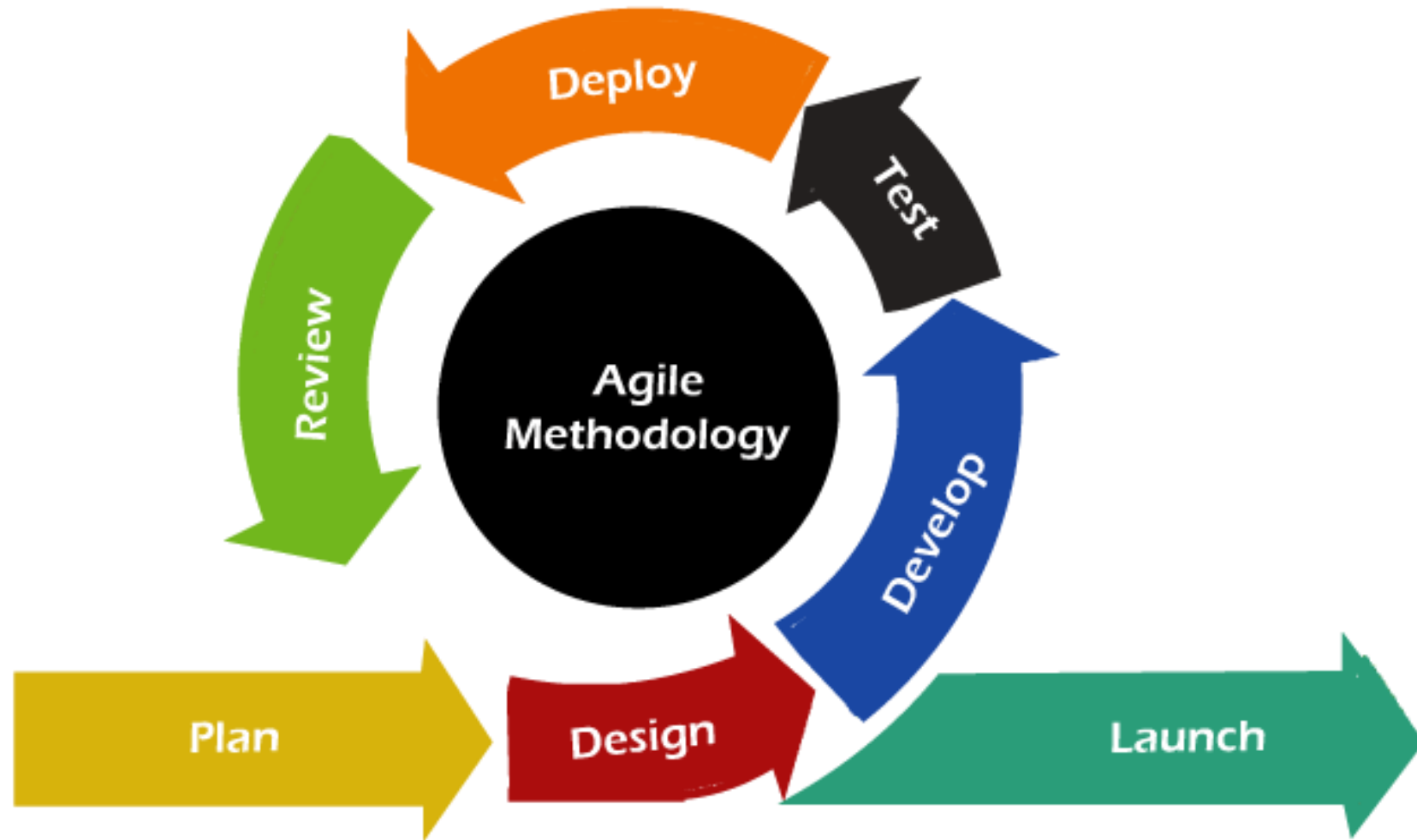
BUSINESS UNDERSTANDING

This is where we evaluate the success of the model if it relates to the original business understanding.



PROJECT MANAGEMENT







Waterfall Model

1



2



3



Iterative (Agile methodology)

1

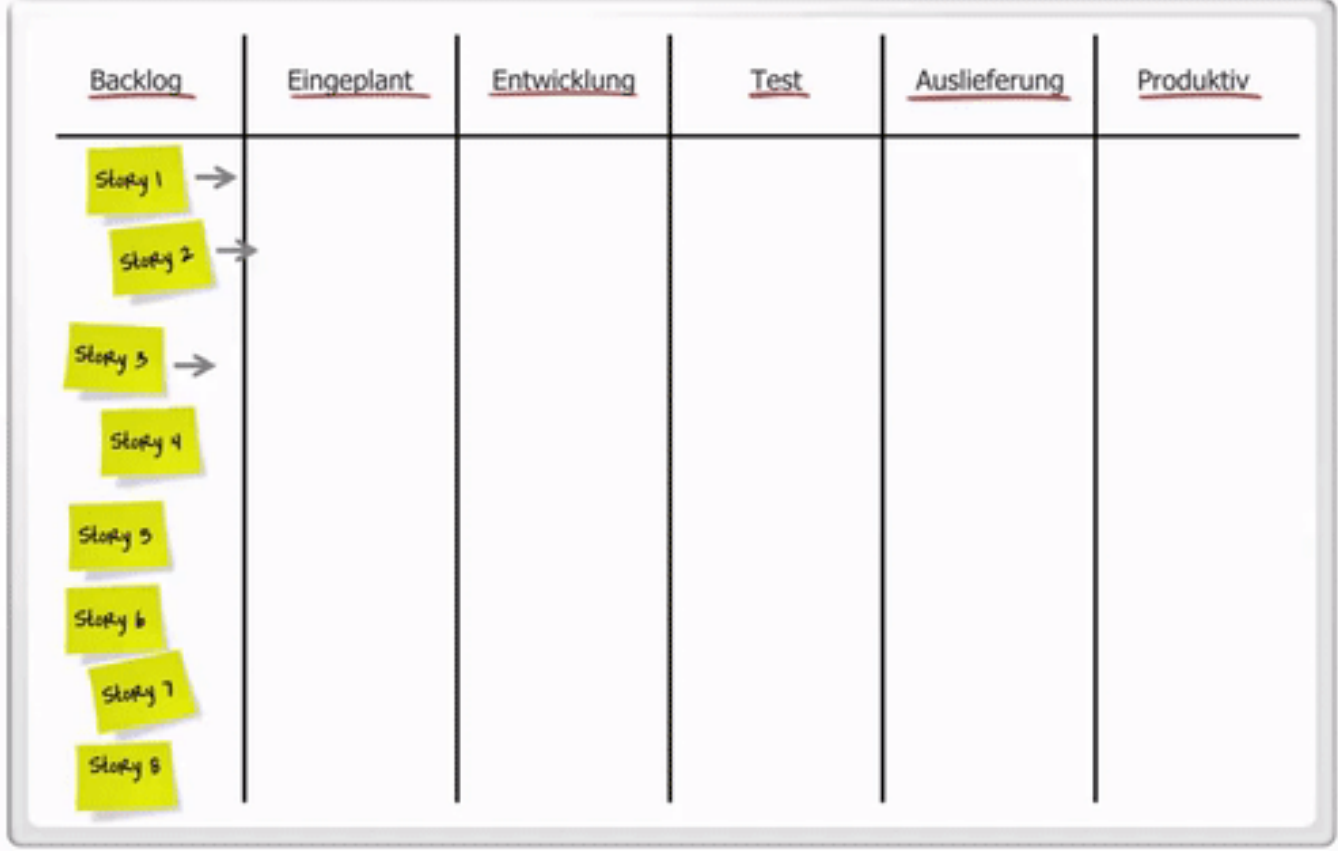


2

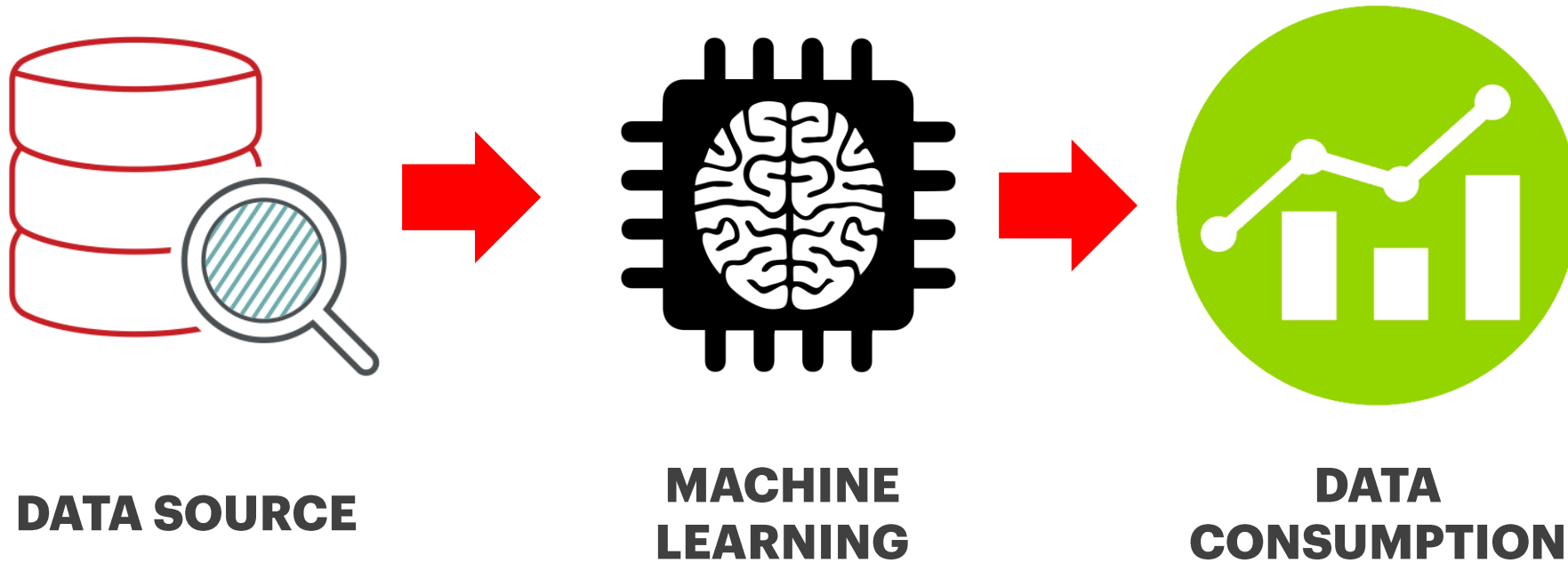


3

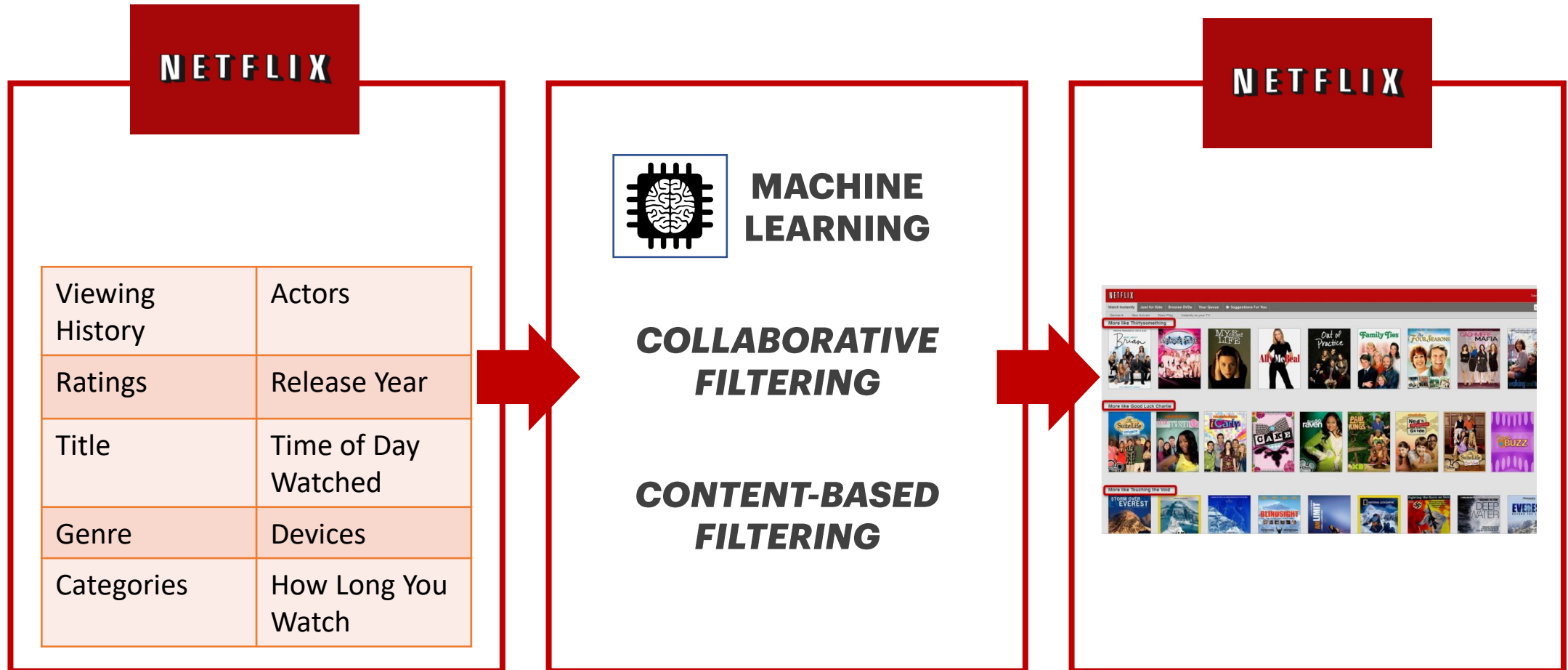


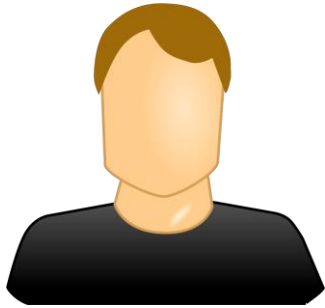


HIGH LEVEL PROCESS



HIGH LEVEL PROCESS EXAMPLE





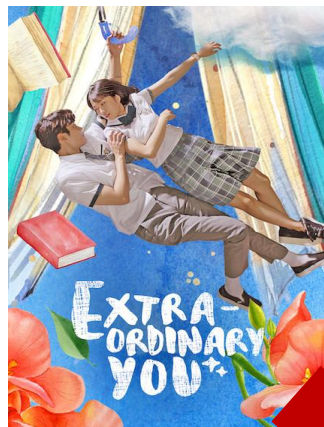
COLLABORATIVE FILTERING



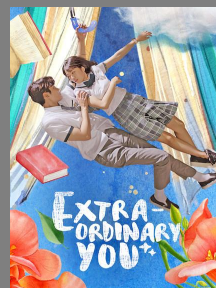
RECOMMENDATION:



WHATIS.TECHTARGET.COM



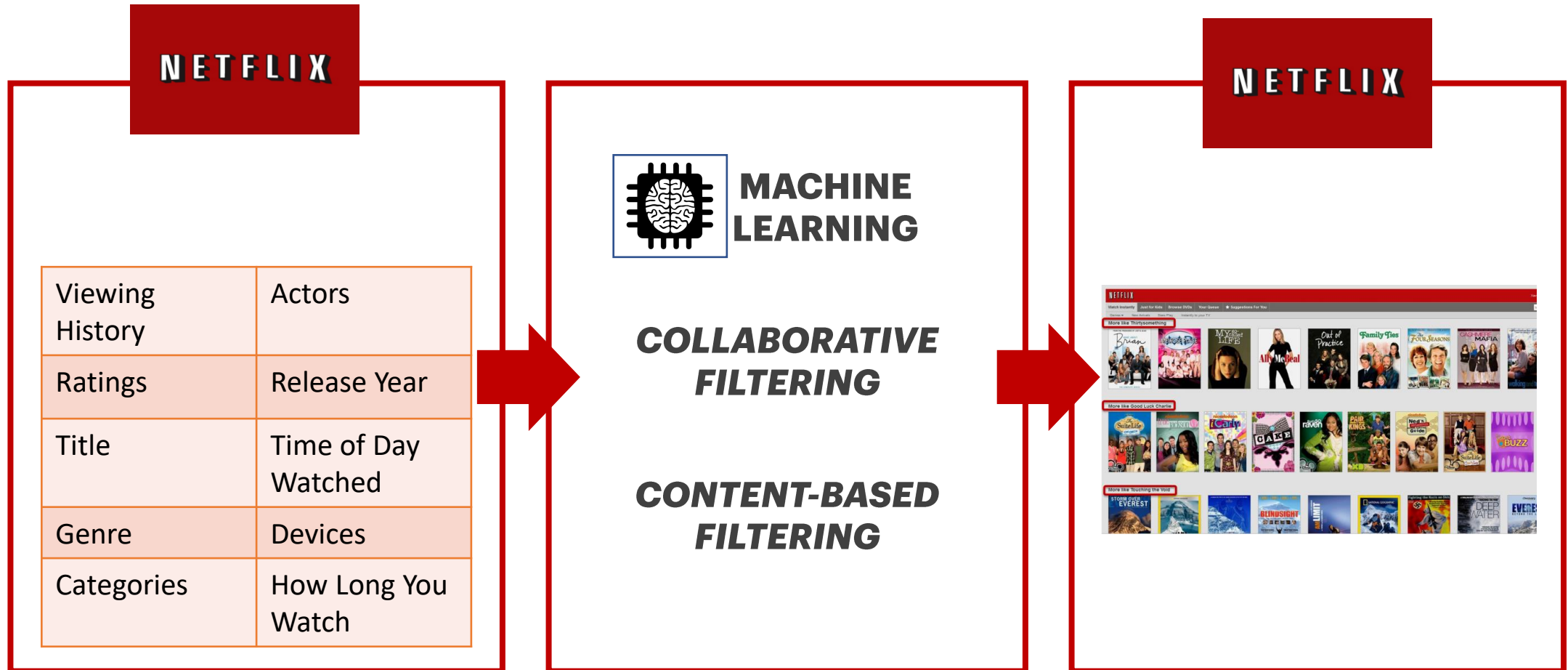
RECOMMENDATION:



CONTENT-BASED FILTERING

[DEVELOPERS.GOOGLE.COM](https://developers.google.com)

HIGH LEVEL PROCESS EXAMPLE



DATA SCIENCE APPLICATIONS

Sentiment Analysis - R Project

Sentiment

Positive



Negative



Emotion



Good



Well

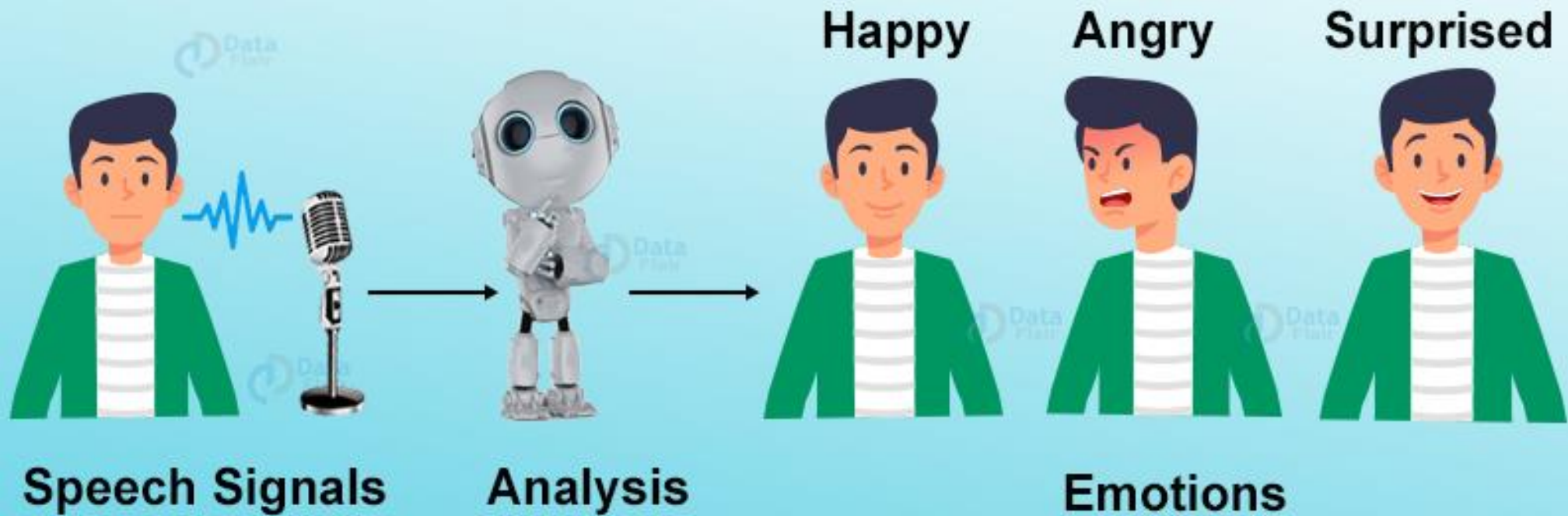


Sad



Angry

Python Project - Speech Emotion Recognition



DATA-FLAIR.TRAINING



Python Project - Gender and Age Detection



Initialize protocol
buffer and
model



Load
networks



Capture video
stream

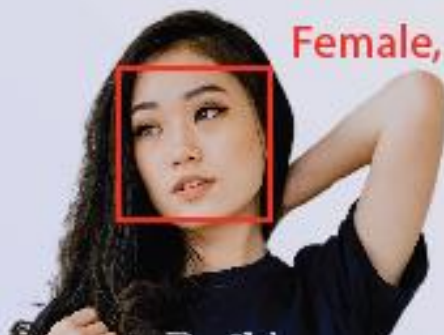


Read stream
and highlight
faces



Feed input to
network and make a
forward pass

Add result
text
to image



Female, Age 23



Male, Age 27

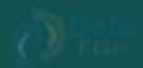


DATA-FLAIR.TRAINING

Python Project - Driver Drowsiness Detection System



*"Drowsiness is the biggest reason for road accidents and
Data Science is the best remedy for it"*



DATA-FLAIR.TRAINING

DATA SCIENCE IN DIFFERENT INDUSTRIES

Data Science in Healthcare



**Genetics &
Genomics**



**Data
Management**



**Drug
Discovery**



**Virtual
Assistance**



**Predictive
Analysis**



**Image
Analysis**

**GOOGLE IMAGES: "DATA
SCIENCE MEDICINE"**

Data Science in Finance



Risk Analytics



Managing Customer Data



Fraud Detection



Real-time Analytics



Consumer Analytics

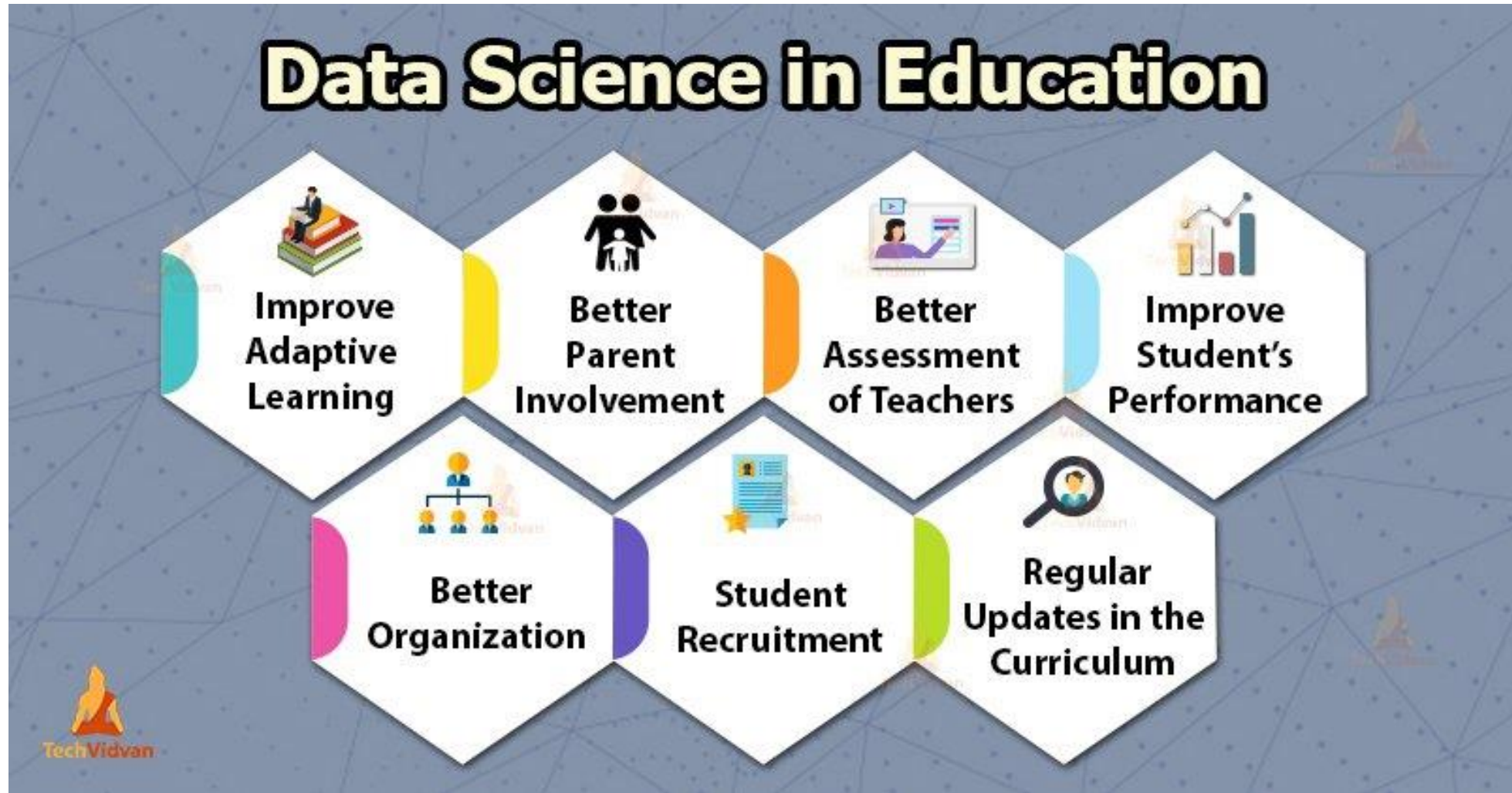


Algorithmic Trading



**GOOGLE IMAGES: "DATA
SCIENCE FINANCE"**

Data Science in Education



GOOGLE IMAGES: "DATA
SCIENCE EDUCATION"

Uses of Data Science in Marketing



Targeted Advertising



Internet Search



Channel optimization



Customer segmentation



Real-time interaction
and analytics



Sentiment analysis



Predictive analytics



Marketing budget
optimization

**GOOGLE IMAGES: "DATA
SCIENCE MARKETING"**



Data Science Uses in Telecom



Product
Optimization



Increased
Network Security



Predictive
Analytics



Fraud Detection



Price Optimization



Real-time
Analytics



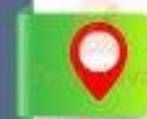
Preventing
Customer Churn



Targeted
Marketing



CLV Prediction



Location-based
Promotion

*GOOGLE IMAGES: "DATA
SCIENCE TELCO"*

SESSION 2 OUTLINE

- **DATA SCIENCE METHODOLOGY**
- **PYTHON SYNTAX**

CARE AND SHARE!

A large, solid pink circle is centered on the slide, serving as a background for the main title text.

DATA SCIENCE BOOTCAMP

Session 2: Data Science Methodology and Python Syntax