Data Mining SS25, Business Understanding & Economy

22.04.2025

Lukas Bader

Data Mining SS25

Seite 1

Fakultät für Informatik und Ingenieurwissenschaften
Institut für Data Science, Engineering, and Analytics

Technology
Arts Sciences
TH Köln

Agenda

- Business Understanding

- Economic Indicators

- Introduction to the data on GitHub

Technology
Arts Sciences

TH Köln

Business Understanding

22.04.2025

Lukas Bader

Data Mining SS25

Seite 3

Fakultät für Informatik und Ingenieurwissenschaften
Institut für Data Science, Engineering, and Analytics

Technology
Arts Sciences

TH Köln

Business Understanding

CRISP-DM

1. Precise description of the business problem

2. Determination of analytic targets and requirements for data analysis:
   - Methods
   - Features
   - …

3. Creating the project plan

22.04.2025

Lukas Bader

Data Mining SS25

Seite 4

Fakultät für Informatik und Ingenieurwissenschaften
Institut für Data Science, Engineering, and Analytics

Technology
Arts Sciences
TH Köln
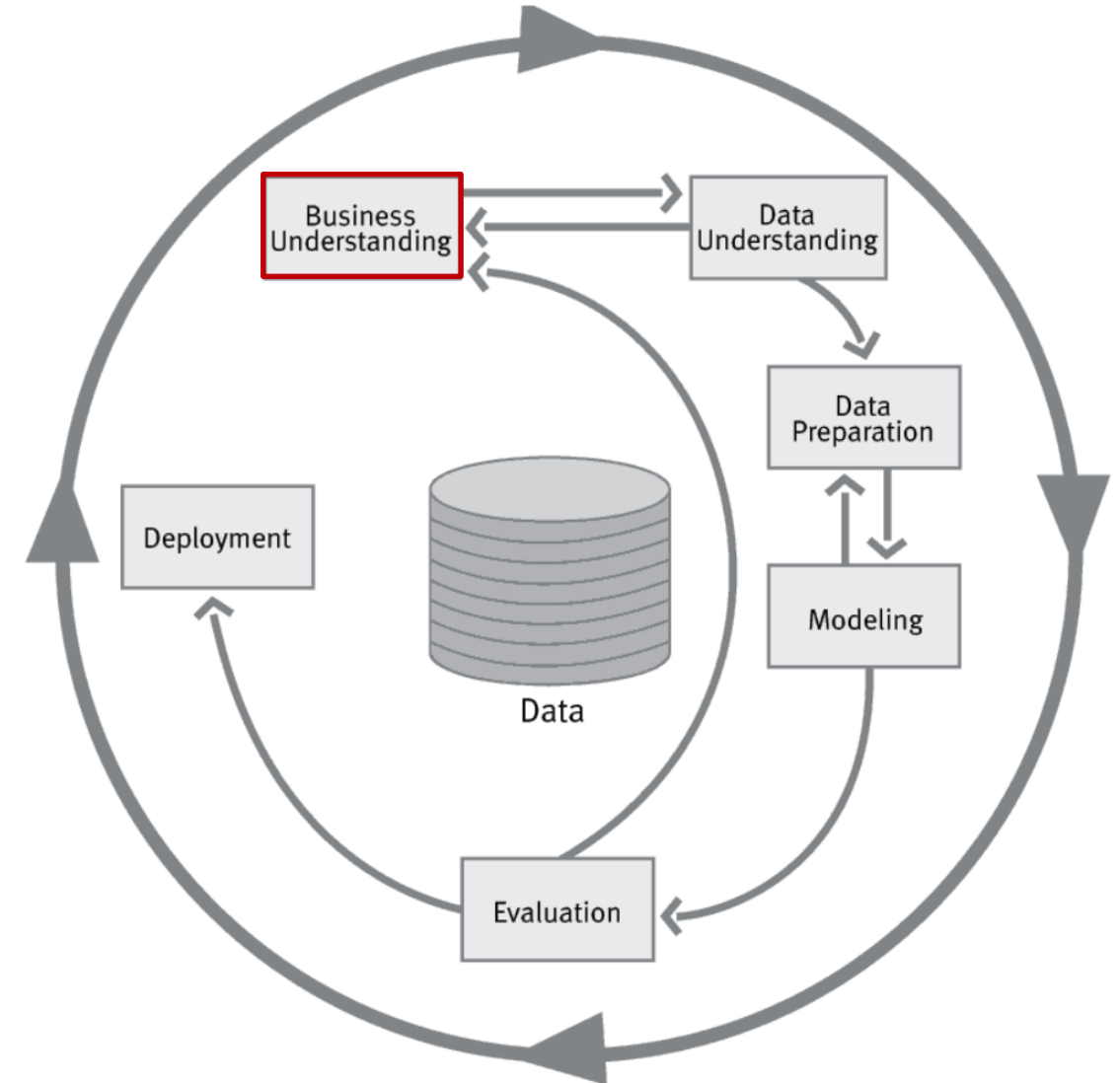
Business Understanding

Business Goal

Identify and prioritize attractive international markets for industrial tools using external data

- What is the company aiming to achieve?   → Strategically expand into new or existing markets with high sales potential

- Which features are relevant?   → Economic strength (GDP, growth), foreign trade, steel consumption, demographics, …

- What defines an "attractive" market?   → High industrial potential, stable conditions, and a large market size.

- Which data sources will we use?   → World Bank, UN, OECD, trade statistics, industry reports.

- What will the outcome look like?   → Country ranking & clustering
  → Identification of focus markets

Technology
Arts Sciences
TH Köln

Business Understanding

CRISP-DM

Project Phase	Time Allocation	Description
Definition of the Business Problem	10% – 20%	Clarify goals, requirements, and stakeholder expectations
Identification of Relevant Data Sources	20% – 30%	Select and access suitable data for the task
Data Preparation	50% – 70%	Clean, transform, and integrate the data for modeling
Modeling	10% – 20%	Develop and train statistical or machine learning models
Evaluation and Success Measurement	10% – 20%	Assess the performance using relevant metrics
Deployment and Implementation	5% – 10%	Integrate the model into production or decision-making processes

Excel list with questions for each milestone

22.04.2025

Seite 7

Lukas Bader

Data Mining SS25

Fakultät für Informatik und Ingenieurwissenschaften
Institut für Data Science, Engineering, and Analytics

Technology
Arts Sciences

TH Köln

Economic Indicators

Lukas Bader

Data Mining SS25

Fakultät für Informatik und Ingenieurwissenschaften
Institut für Data Science, Engineering, and Analytics

Technology
Arts Sciences

TH Köln

Economic Indicators

Overview

- Economic performance

- Investment climate

- Trading structure

- Industry-related indicators

- Demographics & labor market

22.04.2025

Lukas Bader

Data Mining SS25

Seite 9

Fakultät für Informatik und Ingenieurwissenschaften
Institut für Data Science, Engineering, and Analytics

**Technology
Arts Sciences**

TH Köln

Economic performance

GDP – Bruttoinlandsprodukt (BIP), [Link](#)

The absolute value of all goods and services produced in a country in one year, in US dollars, local currency or index of a base year

How is it calculated?

Perhaps the best known approach is GDP = C + I + G + (X - M)

- Consumer spending (C, private & public)

- Investments (I)

- Government spending (G)

- Exports minus imports (X-M, net exports)

Why is it important?

- Shows the economic size and global importance of a country.

- Enables comparisons of economic performance between countries.

- Useful for international market strategies.

Versions:
- Real GDP
- GDP per capita
- GDP growth rate
- GDP by Sector (Agriculture, Industry, Service)

[World Bank](#)

Task: Compare the GPD (current US$) & GDP per capita (current US$) for Germany, China, USA.

Technology
Arts Sciences
TH Köln

Economic performance

Consumer Price Index (CPI) / Inflation Rate

CPI: An index that measures the average price change of a "basket of goods" of typical consumer goods and services - e.g. food, rent, transportation.

How is it calculated?

Statistical offices select a representative basket of goods and observe price trends over time. The index is often standardized to a base year (e.g. 2010 / 2015 = 100).

Inflation rate: $\frac{CPI_{year\ X} - CPI_{year\ X-1}}{CPI_{year\ X-1}}$ x 100

Why is it important?

- CPI is used for calculating the inflation rate.

- Both indicators help to assess the economic stability of a country.

- Important for companies to adjust cost and price strategies on international markets.

- Particularly relevant for market entry decisions and long-term planning.

Task: Compare Inflation, consumer prices (annual %) for Germany, France, Spain, Italy for 2022

Technology
Arts Sciences
TH Köln

Economic performance

Average Wages / Wage Levels & Labor Cost Index

Average Wages : Average income per employee, measured as gross wages or net wages - usually per month, quarter or year.

Labor Cost Index: An index (base is often 2015) that measures the development of average labor costs per hour worked. It includes:

- Wages and salaries
- Employer's social security contributions
- Other ancillary personnel costs

Why is it important?

- Both indicators are central to the cost analysis of foreign markets.
- Particularly relevant for industries with a high proportion of manufacturing, such as August Rüggeberg.
- Can help to identify locations with low-cost but qualified labor.
- Use to identify markets with the potential for high-priced tools.

Task: Check the site of the International Labor Organization

Technology
Arts Sciences
TH Köln

Investment climate

Capital investment (by sector, if available) / Foreign Direct Investment (FDI)

Capital investment : Total investments in durable production goods such as machinery, equipment, buildings, etc. - by companies, the state and private households. (As a percentage of GDP or in absolute figures (USD billion))

- Industry: modernization, capacity expansion → particularly relevant for tool requirements (robotics, …)

- High investment volume = economic focus on the future

- Increasing investment in the industry → growing demand for B2B products

Foreign Direct Investment (FDI): Long-term investments by foreign companies in a country, e.g. through the establishment of production facilities, shareholdings or subsidiaries.

- Strong indicator of a market's international attractiveness

- Signal of economic openness and political stability

- FDI in the industrial sector can indicate strong demand for machinery and tools

Task: How is the availability of the foreign direct investments feature on the World Bank website? Check the following page and test here.

Technology
Arts Sciences
TH Köln

Investment climate

Macroeconomic stability & general conditions

Freedom from Corruption Index (0–100): Measures the perceived level of corruption in the public sector.

- High levels of corruption can lead to inefficient processes, legal uncertainty and higher market entry costs.

Trade Freedom Index (0–100): Assesses the freedom of cross-border trade in goods and services. Takes into account customs duties, non-tariff trade barriers and bureaucracy.

- Countries with a high degree of trade freedom enable easier export/import.

Economic Freedom Index (0–100): Overall indicator on economic freedom, consisting of sub-indices such as:

(Property protection, labor market regulation, State quota, Freedom in the financial and investment sector)

- High value → stable and transparent framework conditions for investment, production and trade.

Fragile State Index (0–120): Measures a country's susceptibility to instability, conflict or the collapse of state structures.

- Relevance for supply chain security, payment defaults, contract risks.

Shadow Economy (% of GDP): Estimate of the share of the informal (unrecorded) economy in GDP.

- Makes valid market analyses and official use of statistics more difficult + Potentially increased risk in business transactions and compliance.
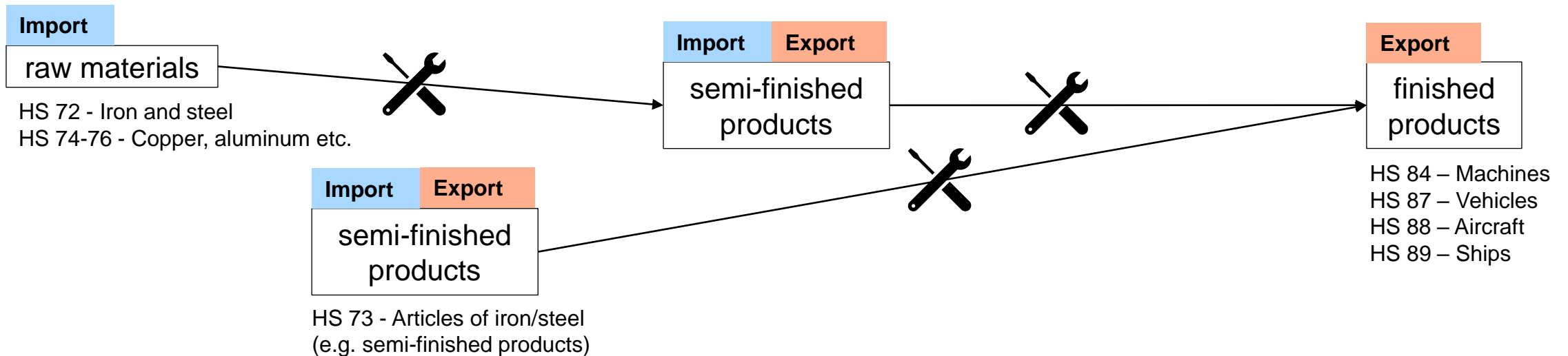
**Technology
Arts Sciences**

TH Köln

Trading structure

Export and Import values for specific goods

Trade values: High import and export values for specific products in the metal sector show a manual manufacturing added value with the use of tools for metalworking.

- Standardized system for classifying products is the HS-Codes system, [Link](#)

- Trade database is UN Comtrade, [Link](#)

Task: Go to the UN Comtrade website and tell me the value of train rails exported by Germany in 2023. Which country has imported the most? Do the same for China:

Import

raw materials

HS 72 - Iron and steel
HS 74-76 - Copper, aluminum etc.

Import **Export**

semi-finished products

Import **Export**

semi-finished products

HS 73 - Articles of iron/steel
(e.g. semi-finished products)

Export

finished products

HS 84 – Machines
HS 87 – Vehicles
HS 88 – Aircraft
HS 89 – Ships

22.04.2025

Lukas Bader

Data Mining SS25

Seite 15

Fakultät für Informatik und Ingenieurwissenschaften
Institut für Data Science, Engineering, and Analytics

Technology
Arts Sciences

TH Köln

Demographics & labor market

Population Size / Labor Force / Unemployment Rate

Population Size : Total population of a country. It makes sense to filter for people who are of working age.

- Higher population = potentially more customers, more employees, more investment opportunities.

- In combination with other indicators (e.g. GDP per capita), purchasing power and depth of industrialization can be better classified.

Labor Force: Number of people of working age who are working or actively looking for work.

- Relevant for sales potential in the B2B sector, as larger industrial labor markets generate more demand for tools.

Unemployment Rate: Proportion of the working-age population without employment but actively seeking work.

- A high ratio can indicate economic weakness or low demand.

- On the other hand, it can indicate a large available labor force potential for new investments.

Task: Which country had the largest labor force in 2024? Check the ILOSTAT

A large and active working population is a positive indicator for industries with high labor intensity.

The combination of population, labor market size and wage levels provides a complete picture of location attractiveness and market size.

Technology
Arts Sciences
TH Köln

Industry-related indicators

Crude Steel Production / Steel Use by Country and Sector / Steel Trade Flows, Link

Crude Steel Production: Production of crude steel - i.e. unprocessed steel that is further processed in subsequent processes.

- Shows the importance of the steel industry in the country
- Indicates possible local value chains
- Countries with high production can also be suppliers or export markets

Steel Use by Country and Sector: Detailed breakdown of where steel is used in the country.

- The greater the proportion of industrial end users, the higher the demand for machining tools

Steel Trade Flows: Imports and exports of steel products (already in UN Comtrade Trade data)

Task: Who were the top 3 producers of crude steel in 2024? Where does germany rank?

- Steel consumption and processing are direct proxies for industrial value creation.
- This is particularly relevant for a company that supplies tools for metalworking.
- Focus countries with a high demand for steel, strong production and industrial processing are particularly attractive.

Technology
Arts Sciences
TH Köln

Introduction to the data on GitHub

Lukas Bader

Data Mining SS25

Fakultät für Informatik und Ingenieurwissenschaften
Institut für Data Science, Engineering, and Analytics

**Technology
Arts Sciences
TH Köln**

Introduction to the data on GitHub

Crude Steel Production / Steel Use by Country and Sector / Steel Trade Flows, [Link](#)

- The Global Economy, Indicator overview and groups

- Trade data UN Comtrade

- Steel data

22.04.2025

Lukas Bader

Data Mining SS25

Seite 19

Fakultät für Informatik und Ingenieurwissenschaften
Institut für Data Science, Engineering, and Analytics

**Technology
Arts Sciences**

TH Köln