

# Business Intelligence

## BI-Project

Prof. Dr. Bastian Amberg  
(summer term 2024)

# Your Task

## General Task Description

Sie schauen sich die möglichen Problemstellungen an.

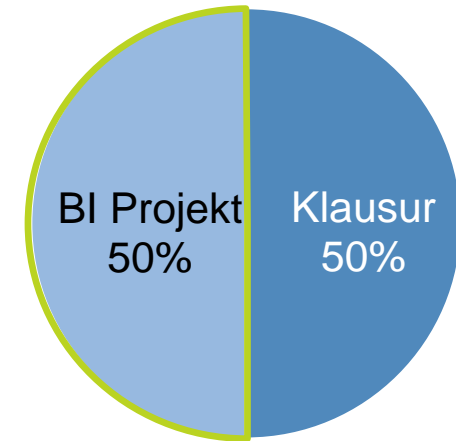
Sie bilden selbstständig eine Gruppe von 4 Personen.

Kommunikation der Gruppenzusammensetzung und der gewählten Problemstellung bis spätestens kommenden **Mittwoch, den 12. Juni 2024**.

(an [bastian.amberg@fu-berlin.de](mailto:bastian.amberg@fu-berlin.de), Betreff „BI Projektgruppe“)

### Aufgabe:

- Die ausgewählte Problemstellung mit Data-Mining-Verfahren bearbeiten (und exemplarisch lösen)
- Bearbeitung der Problemstellung analog zu CRISP-DM
- Lösung und Visualisierung
- **Präsentation** und **Dokumentation** erarbeiten



### 1/3: Präsentation\*

am 17.7. und 19.7.,  
20 Minuten,  
vorherige Abgabe (16.7., 23.59 Uhr)

### 2/3: Kurzdokumentation

Gliederung gemäß CRISP-DM,  
max. 12 Seiten (3 pro Person)

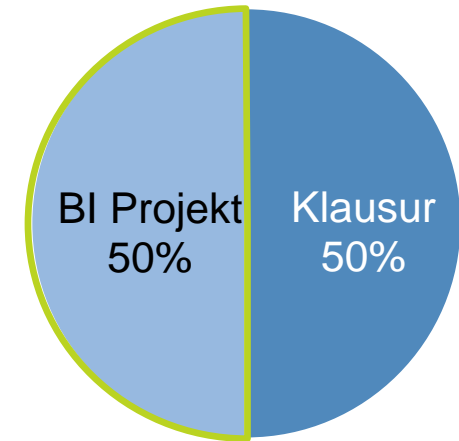
\*Die Reihenfolge der Präsentationen wird vorher gemeinsam ausgelost.

# Your Task

## Detailed Description

- Erstellen und aussagekräftige Evaluation eines **prädiktiven Modells** zum (exemplarischen) Lösen der ausgewählten Problemstellung
- Bearbeitung mit Hilfe des CRISP-DM Prozesses  
(siehe auch Zusatzquellen Chapmen et al. (2000), Wirth & Hipp (2000), in Blackboard)
- Nachvollziehbare, durchdachte Durchführung der Schritte **Project Understanding, Data Understanding, Data Preparation, Modeling, Evaluation**
- Dabei Gegenüberstellung der Ergebnisse von verschiedenartig erstellten Modellen aus dem Bereich **Supervised Learning** (~ 2),  
z.B. Decision Trees, (Multinomial) Logistic Regression, Support Vector Machines, (Deep) Neural Network, Naive Bayes
- Sowie Vergleich mit einem einfachen alternativen Ansatz, der eine Baseline-Performance bei der Bewertung vorgibt
- **Präsentation** des Projektergebnisses und **Dokumentation**, die den durchgeführten Prozess und die erzielten Ergebnisse veranschaulicht sowie reflektiert  
*Mit der Präsentation soll auch die Bearbeitungsphase des Projektes abgeschlossen sein. Die Dokumentation ist die (anschließende) saubere, schriftliche Aufbereitung des gesamten Projektes.*
- Abgaben jeweils über die angelegte Arbeitsgruppe in Blackboard (Präsentation, Dokumentation, sowie den digitalen Anhang, d.h. Code oder andere Dokumente zur Nachvollziehbarkeit)  
*Mittels ChatGPT (oder anderer KI) generierte Inhalte sind als solche zu kennzeichnen.*

Ref.



### 1/3: Präsentation

am 17.7. und 19.7.,  
20 Minuten,  
vorherige Abgabe (16.7., 23.59 Uhr)

### 2/3: Kurzdokumentation

Gliederung gemäß CRISP-DM,  
max. 12 Seiten (3 pro Person)

Kommunikation der Gruppenzusammensetzung und der gewählten Problemstellung bis **Mittwoch, den 12. Juni 2024**.  
(an [bastian.amberg@fu-berlin.de](mailto:bastian.amberg@fu-berlin.de), Betreff „BI Projektgruppe“)



# Your Projects

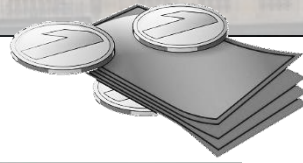
Projects from Kaggle.com



## Costa Rican Household Poverty Prediction

Set of household characteristics from a representative sample of households  
Make sure the right people are given enough aid

*Goal: Predict the level of need (income level)*



>> Crime Mapping



## Crime Classification in Los Angeles

4 years of crime reports from all across Los Angeles

*Goal: Predict the category of crime that occurred given a certain time and location*

## West Nile Virus Prediction in Chicago

7 years of weather, location, testing, and spraying data

*Goal: predict the presence of West Nile virus for a given time, location, and species*



## Crime Classification in San Francisco

12 years of crime reports from all across San Francisco

*Goal: Predict the category of crime that occurred given a certain time and location*

Ref.