

# Chicago Trip planen: Wann und wo?

Vincent Merkel, Benjamin Esch, Christof Warsinsky, Eduardo Stein Mössner.

10. November 2023

## Inhaltsverzeichnis

<b>1 Die Fragestellung.</b>	<b>1</b>
1.1 Konkretere fragen um unseren Projekt zu leiten. . . . .	1
<b>2 Die Daten.</b>	<b>1</b>
2.1 Chicago Crime Dataset aus das Data Portal der Stadt Chicago).	1
<b>3 Projektverlauf</b>	<b>2</b>
3.1 CRISP-DM Modell. . . . .	2
3.2 CRISP-DM Modell anwendung. . . . .	2
3.2.1 Verstehen was wir suchen. . . . .	2
3.2.2 Die Daten verstehen. . . . .	2
3.2.3 Die Daten vorbereiten. . . . .	2
3.2.4 Modellieren. . . . .	2
3.2.5 Überprüfen. . . . .	2
3.2.6 Präsentieren. . . . .	3
<b>4 Projektdurchführung.</b>	<b>3</b>
4.1 Daten importieren. . . . .	3
4.1.1 Daten aus das Chicago Data Portal importieren . . . .	3

## 1 Die Fragestellung.

Chicago Trip planen: Wann und Wo? Genau das haben wir uns gefragt und mit dieser Hauptfrage im Hinterkopf sind wir durch den ganzen Projektverlauf gegangen.

Um diese Hauptfrage möglichst gut zu beantworten haben wir diese in vier Unterfragen unterteilt.

## **1.1 Konkrete Leitfragen**

Wo ist es in Chicago am gefährlichsten?

Wann ist es sicherer?

Gibt es Jahreszeiten in die es unsicherer wird?

War es früher sicherer oder gefährlicher?

## **2 Die Daten.**

### **2.1 Chicago Crime Dataset aus das Data Portal der Stadt Chicago).**

Dieser Datensatz enthält die gemeldeten Straftaten, die in der Stadt Chicago von 2001 bis heute stattgefunden haben. Die Daten werden aus dem CLEAR-System (Citizen Law Enforcement Analysis and Reporting) des Chicago Police Department extrahiert.

## **3 Projektverlauf**

### **3.1 CRISP-DM Modell.**

„Cross Industry Standard Process for Data Mining“. Es handelt sich um ein standardisiertes Prozessmodell, das für das Data Mining anwendbar ist, um Datenbestände nach Mustern, Trends und Zusammenhängen zu durchsuchen.

[<https://www.bigdata-insider.de/>]

### **3.2 CRISP-DM Modell anwendung.**

Die Bearbeitung dieses Projekts wurde unter Anwendung des CRISP-DM Prozessmodell durchgeführt.

Dafür wurden die nächsten wichtige Schritte beachtet.

#### **3.2.1 Verstehen was wir suchen.**

Als aller erstes ist absolut unverzichtbar die Fragestellung genau zu verstehen und genau wissen was man sucht um zu wissen wie man vorgehen soll.

### **3.2.2 Die Daten verstehen.**

Um die Fragestellung zu beantworten ist es von aller Notwendigkeit ein Datensatz zu finden der die benötigte Daten enthält um auf dieser Basis die Antwort auf die unterschiedliche Fragen zu kommen.

### **3.2.3 Die Daten vorbereiten.**

Unter Anwendung von Python 3 müssen wir sicherstellen das die Daten Clean und Nutzbar sind.

### **3.2.4 Modellieren.**

Durch Python 3 Code müssen wir ein Programm schreiben der die Daten modelliert und entsprechende Antworten auf unsere Fragestellungen liefert.

### **3.2.5 Überprüfen.**

Es ist von großer Wichtigkeit die Ergebnisse zu überprüfen, denn diese müssen der Realität gerechtfertigt sein.

### **3.2.6 Präsentieren.**

Zum Schluss, nachdem sichergestellt wurde das die Ergebnisse die vorausgesetzte Qualität entsprechen, werden die Ergebnisse präsentiert und die entsprechende Antworten auf die Fragestellung wird in klarer Form vorgestellt.

## **4 Projektdurchführung.**

### **4.1 Daten importieren.**

#### **4.1.1 Daten aus das Chicago Data Portal importieren**

Für die Durchführung dieser Projektarbeit wurde das Chicago-Crime Dataset aus die Quelle als CSV heruntergeladen.

Mit den folgenden Link kommt man an das entsprechende Dataset:

<https://data.cityofchicago.org/Public-Safety/Crimes-2001-to-Present/ijzp-q8t2/data>

```
import pandas as pd

df = pd.read_csv('C:\\Users\\User\\OneDrive\\DHBW\\Erste Semester\\Introduction\\data\\data.csv')

df
```