



# Data Science Education On The Rise

Erstellung einer weltweiten Studiengangs-Übersicht für  
die Studienrichtung “Data Science”

---

# Team



Carole Mattmann



Luca Casuscelli



Jonas Zürcher



Silvan Leibacher

---

## Agenda

1. Fragestellung, Vorgehen und Tools
2. Prozessschritte CIP
3. Beantwortung Fragestellung
4. Schlusswort, Diskussion, Ausblick

---

## Fragestellungen

1. Welches Masterstudienangebot für Data Science existiert in Europa?
2. Welche Universität in Asien, hat das beste internationale Ranking und bietet einen Studiengang in Data Science an?
3. Wie verändert sich das Studienangebot im Bereich Data Science in der Region Amerika im Verlauf der Zeit?

---

## Vorgehen

1. Crawlen der Daten
2. Bereinigung und Transformation mit Pandas
3. Vereinigung der Daten in Tableau Prep zu einem Datenset
4. Import des Datensets in eine SQL-Datenbank

The background features a blue silhouette of a human head in profile, facing left. Overlaid on this are various financial and data-related graphics: a red line graph with peaks and valleys, a bar chart with orange bars, a grid of blue dots, and a pattern of binary code (0s and 1s) in blue. A magnifying glass icon is visible near the bottom left, and a small white icon resembling a hamburger menu is in the top right corner.

# Prozessschritte zur Erstellung des Übersichtsdatensets



---

# Collection und Extraction

mit Python

Es bestehen zwei vom Aufbau her fast identische Crawler. Die beiden Webseiten der Studiengänge stammen vom selben Anbieter und sind genau gleich aufgebaut.

1

Data-Crawling über alle Data Science Studiengänge aus zwei Quellen:

<https://www.bachelorstudies.com/Data-Science/>, <https://www.masterstudies.com/Data-Science/>

2

Data-Crawling über die Rankings aller Hochschulen:

<http://www.webometrics.info/en/world>

---

# Data Cleaning

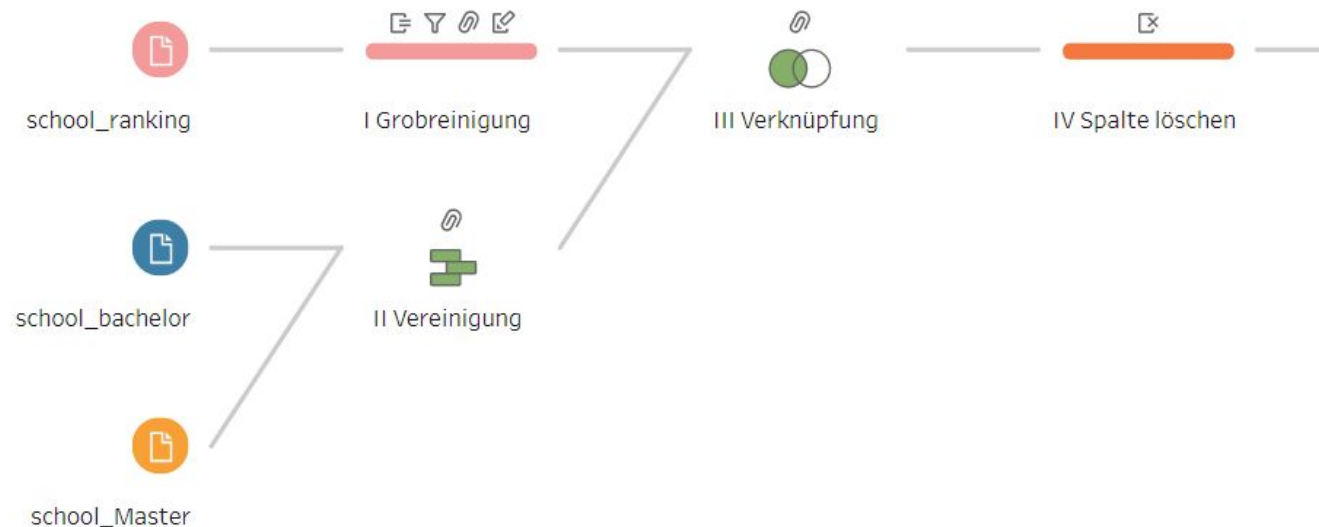
mit Pandas (Python)

- Die Anwendung von Pandas und damit der erste Teil der Daten Bereinigung erfolgt bereits in den Crawler Skripten.
- Der Bereinigungs-Prozess wurde in der Funktion “cleanse\_data(dataframe)” definiert. Dieser wurde nur bei den Studiengangs-Daten angewendet, damit mehr Bereinigungs-schritte mit Tableau Prep übrigbleiben.



# Cleaning, Integration & Transformation

In Tableau Prep



---

# Cleaning, Integration & Transformation

Von drei csv-Dateien in Tableau Prep

1

Grobbereinigung der  
"school\_ranking"-Daten

2

Vereinigung & Verknüpfung der Datensets,  
redundante Spalten löschen

3

Transformation Lokalisierung & Sprache

6

Transformation Abschlusstyp, Dauer,  
Schulstart

7

Kopfzeile & Spaltenreihenfolge  
anpassen

8

Export der Datei als csv und tde

---

# Transfer & Import auf die Datenbank

Microsoft SQL Server 2017 + Microsoft SQL Server Management Studio

1

Datenbank erstellen

2

Datenset importieren

3

Kontrolle des Imports

The background features a blue silhouette of a human head in profile, facing left. Overlaid on this are various financial and data-related graphics: a red line graph with peaks and valleys, several orange and red bar charts, and a grid of blue dots. Faint icons of a magnifying glass, a laptop, and a person are visible at the bottom. The overall color palette is dominated by blue, orange, and red.

# Beantwortung der Fragestellungen

---

**Fragestellung 1: Welches Masterstudienangebot für Data Science existiert in Europa? (Stand: 20.03.2020)**

- In Tableau wurden die Daten gefiltert, dass nur noch Hochschulen von Europa mit einem Master Programm angezeigt werden.
- **102 Hochschulen bieten Data Science Studiengänge auf Stufe Master an.**
- **Einige Hochschulen bieten mehrere Studiengänge an.**
- **Die Hochschule Luzern fehlt leider in der Liste.**

---

**Fragestellung 2: Welche Universität in Asien, hat das beste internationale Ranking und bietet einen Studiengang in Data Science an? (Stand: 20.03.2020)**

- In Tableau wurden die Daten gefiltert, dass nur noch Hochschulen von Asien mit einem Data Science Programm und einem vorhandenen Hochschul-Ranking angezeigt werden
- **Die City University of Hong Kong das beste Ranking mit dem Gesamt-Rang von 186.**
- **Die City University of Hong Kong bietet drei Studiengänge an.**

---

**Fragestellung 3: Wie verändert sich das Studienangebot im Bereich Data Science in der Region Amerika im Verlauf der Zeit? (Zeitintervall: 20.03.2020-09.04.2020)**

- Die zwei Übersichtsdatensets von den beiden Zeitpunkten werden verwendet.
- Ein neues Datensets wird im Tableau Prep erstellt.
- Das neue Datenset beinhaltet nur Zugänge und Abgänge bei den Studiengängen.



---

**Fragestellung 3: Wie verändert sich das Studienangebot im Bereich Data Science in der Region Amerika im Verlauf der Zeit? (Zeitintervall: 20.03.2020-09.04.2020)**

- **Das Studienangebot hat sich für den Kontinent Nordamerika im Zeitabschnitt nur leicht verändert.**
- **Hinzugekommen ist der Studiengang „MS in Health Data Analytics (On-Campus)“ der University of Louisville.**
- **Weggefallen ist der Studiengang „Master of Science in Computer Science Data Analytics» der Nova Southeastern University.**

The background is a complex digital collage. On the left, a blue silhouette of a human head in profile is visible. Overlaid on this and the rest of the image are various financial and data-related graphics: orange and red bar charts, a fluctuating orange line graph, and a red line graph with white circular markers. A specific data point on the red line is labeled '+11,00.00'. The background also features a grid of blue dots, binary code (0s and 1s), and various financial icons like dollar signs, coins, and a magnifying glass. The overall color palette is dominated by blue, orange, and red.

# Schlusswort

---

## Diskussion

- Nur die gecrawlten Daten der Studiengänge wurden mit Pandas gereinigt. Es wäre sinnvoll diese bei einer Weiterverfolgung des Projekts **bei allen gecrawlten Daten durchzuführen**.
- Problem könnte aber in Zukunft entstehen, wenn neue Hochschulen Data Science Studiengänge einführen (v.a. Verknüpfungsschritt im Tableau Prep). Die Datensets der Studiengänge und der Rankings werden über den Hochschulnamen verknüpft.
- Die Namen derselben Hochschulen unterscheiden sich leider in nicht reproduzierbarer Weise und führt zu Problemen im Matching. → **Für eine 100 % Übereinstimmung müsste man alle Datensets mittels einer klar zuweisbaren ID oder örtlichen Adresse zur Verfügung haben.**

---

## Ausblick

- Der nächste Schritt in diesem Projekt wäre die Verknüpfung der verschiedenen Datensets zu optimieren.
- Das Übersichtsdatenset könnte zur Entwicklung einer Applikation verwendet werden.
- Die Applikation könnte bei Studienplanung, Vergleichen von Partneruniversitäten für ein Auslandssemester helfen oder auch als Führungsinstrument für die Studiengangsleitungen dienen.



**Vielen Dank!**