Projektname

Datenbericht

Letzte Änderung: 24.08.2021

# Rohdaten

## Übersichtstabelle der Rohdatensätze

|  |  |  |
| --- | --- | --- |
| **Datensatz Name** | **Quelle** | **Speicherort** |
| Dataset 1 | Name/kurze Beschreibung der Datenquelle | Link und/oder kurze Beschreibung des Ortes, wo die Daten abgelegt wurden |
| Dataset 2 | … | … |

Falls zutreffend: Entity-Relationship Diagram

## Details Dataset 1

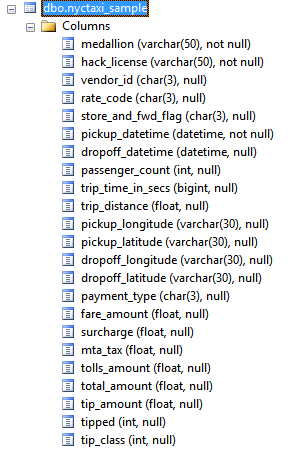
* Beschreibung, welche Informationen die Daten enthalten
* Details zur Datenquelle
* Informationen zur Datenbeschaffung: Beschreibung und eventuell Verweise auf Ressourcen (Download-Scripts, Tools, Online-Services, …)
* Rechtliche Aspekte zur Nutzung der Daten, Lizenzen etc.
* Data Governance-Aspekte: Kategorisierung der Daten aufgrund Geschäftsinterner Vorgaben, z.B. Öffentlich, Geschäftsrelevant, Persönlich
* Wie kann auf die Daten zugegriffen werden: Beschreibung, Scripts, Tools, …
* Falls zutreffend: Unterteilung in abhängige (Zielgrösse, Regressand) und unabhängige (Regressor) Variablen
* …

### Datenkatalog

Beispiel eines Data Dictionary

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Spaltenindex | Spaltenname | Datentyp | Werte (Wertebereich, ev. Validierungsregeln) | Kurze Beschreibung | Primary/ Foreign Key |
| 1 |  |  |  |  |  |
| 2 |  |  |  |  |  |
| … |  |  |  |  |  |

Beispiel eines Schemas von Daten in einer relationalen Datenbank:



## Datenqualität Dataset 1

Explorative Datenanalyse

* Beschreibung der Vorgehensweise und Verweise auf verwendete Ressourcen (Scripts, Tools, Software, …) und Output-Dateien -> Nachvollziehbarkeit und Sicherstellung der Reproduzierbarkeit
* Zusammenfassende Ergebnisse: Visualisationen und statistische Kenngrössen
* Erkenntnisse und relevante Zusammenhänge im Bezug auf die Modellierungs- und Analyseziele
* Einschätzung der Datenqualität

## Details Dataset 2

* …

### Datenkatalog Dataset 2

### Datenqualität Dataset 2

# Prozessierte Daten

## Übersichtstabelle der Prozessierten Daten

|  |  |  |
| --- | --- | --- |
| Name | Input-Datensätze | Speicherort |
| Processed Dataset 1 | Dataset 1, Dataset 2 | Link und/oder kurze Beschreibung des Ortes, wo die Daten abgelegt wurden |
| Processed Dataset 2 | Dataset 2 |  |

## Details Processed Dataset 1

* Beschreibung, welche Informationen die Daten enthalten
* Details und Gründe zu den Prozessierungsschritten -> Nachvollziehbarkeit und Sicherstellung der Reproduzierbarkeit
* Wie kann auf die Daten zugegriffen werden: Beschreibung, Scripts, Tools, …
* …

## Details Processed Dataset 2

* Beschreibung, welche Informationen die Daten enthalten
* Details und Gründe zu den Prozessierungsschritten
* Wie kann auf die Daten zugegriffen werden: Beschreibung, Scripts, Tools, …
* …