DB Travel Data

1. **Um was für Daten handelt es sich?**

* Die Daten bestehen aus einem Tagessatz mit allen Zügen der Deutschen Bahn, die an dem 20.06.19 in Frankfurt eintreffen oder dort wieder abfahren.
* Dort werden Gründe für Verspätungen angegeben und die Zeit, wie lange sich die Fahrt verspätet.

**Genutzte Spalten des Datensatzes**

|  |  |  |
| --- | --- | --- |
| **Kürzel** | **Beschreibung** | **Für weitere Bearbeitung benötigt?!** |
| TAA | Warnungen, warum kommt es zur Verspätung |  |
| TA | Verspätete Abfahrtszeit |  |
| TIN | Zug Modell |  |
| TIR | Route |  |
| TSI | ID der Station (Hier immer 8000105 (Frankfurt)) |  |
| TIM | Kommt der Zug an oder fährt er ab. |  |
| TIL | Abfrageparameter um die Daten zu bekommen |  |
| TIRE | Fahrtziel des Zuges |  |
| TIP | Gleis auf dem der Zug ankommt |  |
| TIT | Abfahrtszeit |  |
| TID | Datum |  |
| TSC | Bahnhof Name einschließlich seiner ID hier immer (Frankfurt%238000105) |  |
| TAc | Minuten, die der Zug verspätet eintrifft ( TA - TIT ) |  |

1. **Wie sehen meine Daten überhaupt aus?**
2. **Muss ich meine Daten vor der Analyse vorverarbeiten?**
3. **Welche Analysen sind für meinen Datensatz geeignet und was will ich damit herausfinden?**
4. **Wie stelle ich die Ergebnisse meiner Analysen da, damit auch meine Kommilitonen etwas daraus lernen können?** (in max. 15 min)