

T Square
Project Step 3: Cleaning & Showcase

Timo Scheerer, Timo Büchert

Cleaning Pipeline



- 1. Attribute Cleaning
 - 1. City names
 - Align different pronounciations of city names, use longer option
 - □ "Frankfurt(M)" & "Frankfurt(Main)" → "Frankfurt(Main)"
 - 2. Coordinates of identical stations
 - Use the middle
- 2. Entity cleaning
 - 1. Removal of all irrelevant stations that are not used in any transfer option
 - "Marburg Taubenweg"
 - "Hamburg-Altona"























