**Data description**

The attributes of subway lines and stations in Beijing/Shanghai/Shenzhen (by 2020) are shared in this dataset.

* l**ines\_[bj/sh/sz].csv**: list of subway lines.

|  |  |  |
| --- | --- | --- |
| **Column** | **Definition** | **Data type** |
| nid | identication of the subway line | int |
| name | name of the subway line | text |
| full\_names | full names of the line that distinguishes the direction (marked by the starting and terminal station) | text |

**Data example:**

$ head -5 lines\_bj.csv

nid,name,full\_names

1,LINE 1,"{""LINE 1 (PINGGUOYUAN-SIHUIDONG)"",""LINE 1 (SIHUIDONG-PINGGUOYUAN)""}"

2,LINE 2,"{""LINE 2 (XIZHIMEN-XIZHIMEN)"",""LINE 2 (JISHUITAN-JISHUITAN)""}"

3,LINE 4,"{""LINE 4 (TIANGONGYUAN-ANHEQIAO NORTH)"",""LINE 4 (ANHEQIAO NORTH-TIANGONGYUAN)""}"

4,LINE 5,"{""LINE 5 (TIANTONGYUAN NORTH-SONGJIAZHUANG)"",""LINE 5 (SONGJIAZHUANG-TIANTONGYUAN NORTH)""}"

* **stations\_[bj/sh/sz].csv**: list of subway stations.

|  |  |  |
| --- | --- | --- |
| **Column** | **Definition** | **Data type** |
| sid | identication of the subway station | int |
| name | name of the subway station | text |
| linesid | list of ids of subway lines where the station is located | int[] |
| lines | list of names of subway lines where the station is located | text[] |
| lat | latitude of the station | double |
| lng | longitude of the station | double |

**Data example:**

$ head -5 stations\_bj.csv

sid,name,linesid,lines,lat,lng

1,MUXIDI,{1},"{""LINE 1""}",39.90614419,116.3313326

2,GUCHENG,{1},"{""LINE 1""}",39.90615316,116.1844542

3,TIAN'ANMENDONG,{1},"{""LINE 1""}",39.90633242,116.3952964

4,TIAN'ANMENXI,{1},"{""LINE 1""}",39.90606169,116.3854103

* **Eudistance\_[bj/sh/sz].csv**: Euclidean distance between each station pair. The distance is the geodesic distance calculated based on the latitude and longitude of the two stations

| **Column** | **Definition** | **Data type** |
| --- | --- | --- |
| id | row index | int |
| sid1 | starting station id | int |
| sid2 | terminal station id | int |
| Eudistance | Euclidean distance between stations sid1 and sid2 (in meters) | double |

**Data example:**

$ head -5 Eudistance\_2019.csv

id,sid1,sid2,Eudistance

0,1,1,0.0

1,1,2,12559.66674557

2,1,3,5469.61603355

3,1,4,4624.23073008