

DBKDA-2016 Tutorial II:

Hands-On Exercise 2

1. Download the file `city.tsv` from <http://www.smiffy.de/dbkda-2016/city.tsv> and store it on your disk.
2. Import the file into a data frame with name `'city.all'`.
3. Substitute the „NULL“-values for the concerned columns by NA values.
4. Change the datatype of the columns `Longitude`, `Latitude` and `Population` to numeric.
5. Formulate the following queries:
 - How many cities are in the data frame?
 - Return the cities which reside on the equator.
 - Return all cities from Brasilia with coordinates, ordered by their population.
 - How many cities can be found in the USA
 - In which states (Province) can cities with the name `'Springfield'` be found?
6. Add the city of Roskilde (population: 46701; coordinates: 55.65, 12.083333) to the data frame.
7. Formulate a query, which returns Roskilde from the data frame.
8. Datastructure conversion:
Generate the following data frame:

```
to.convert<-city.frame[city.frame$Country=="DK",  
                        c('Name', 'Population')]
```


and convert it into a named vector (population as vector values, cityname as vector element names)
9. Graphics:
 - Create a plot, which shows the coordinates (longitude, latitude) of all cities.