

Sensordownload

Auf <http://archive.luftdaten.info/> werden täglich die Messungen der verschiedenen Sensoren hochgeladen. Dabei existiert für jeden Sensor eine eigene csv-Datei, welche durch Anklicken des jeweiligen Links heruntergeladen werden kann. Die angegebenen Links setzen sich aus Datum, Sensortyp und Sensor-ID zusammen - daher ist anhand des Links zunächst nicht klar, an welchem Ort ein Sensor installiert ist. Somit müssen zunächst die IDs der in München installierten Sensoren bestimmt werden.

Identifizierung der Sensoren in München

Durch Anklicken eines Links werden die Daten des jeweiligen Sensors heruntergeladen - diese Daten enthalten dann u.a. Latitude und Longitude des Sensors. Unser Vorgehen ist daher zunächst für den ersten Tag der Analyse, den 25.07.2019, sämtliche für diesen Tag verfügbaren Sensordaten herunterzuladen, um anschließend herauszufinden welche der Sensoren sich in München befinden. Die IDs dieser Sensoren werden dann in einer Textdatei abgespeichert und für den gesamten Beobachtungszeitraum (25.07.2019 - 14.08.2019) verwendet.

Die Sensoren für den 25.07.2019 stehen unter <http://archive.luftdaten.info/2019-07-25/> zum Download bereit. Der Download aller dort angegebenen Sensordaten kann mit dem R-Skript **download_all_sensors.R** durchgeführt werden, in welchem alle notwendigen Schritte erklärt werden. Da jeder Sensortyp andere Merkmale misst, werden die Sensoren automatisch in Ordnern je nach Sensortyp gespeichert, um die Daten später möglichst effizient einlesen zu können.

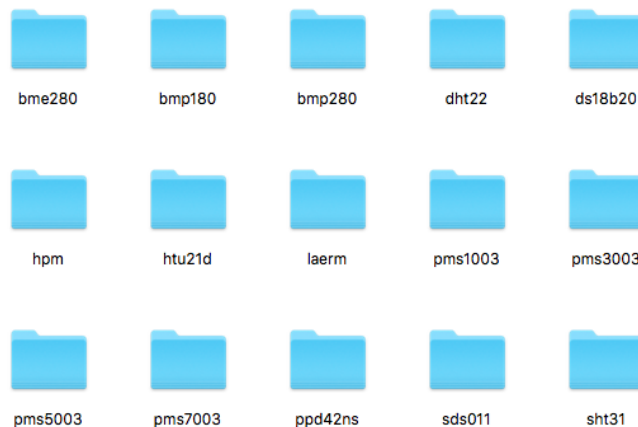


Figure 1: Für jeden Sensortyp wird ein eigener Ordner angelegt

Nachdem die Daten heruntergeladen wurden wird geprüft, welche der Sensoren sich in München befinden. Dies geschieht mithilfe des R-Skripts **extract_munich_sensors.R**. Dazu werden die Koordinaten der Stadtgrenze von München verwendet, welche zuvor mithilfe des R-Skripts **get_munich_boundary.R** bestimmt wurden. Alle Sensoren welche nicht innerhalb der Stadtgrenze liegen werden gelöscht. Die verbliebenen Sensoren gehören nur noch vier verschiedenen Sensortypen an, d.h. die anderen Typen kommen in München nicht vor. Diese Sensortypen sind bme280 (Luftdruck), bmp180 (Luftdruck), dht22 (Temperatur & Luftfeuchtigkeit), und sds011 (Feinstaub). Die identifizierten Sensoren werden im Verzeichnis *workfiles* als eindeutige Typ-ID Kombinationen (z.B. bme280_sensor_10093) in einer Textdatei gespeichert, da dies der Zusammensetzung der Links auf der Website entspricht.

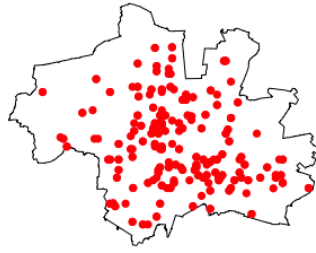


Figure 2: Messorte der Sensoren in München

Download der Sensoren in München

Da nun die Sensoren in München anhand ihrer Typ-ID Kennung identifiziert wurden, kann ein Download dieser Sensoren für den Beobachtungszeitraum (25.07.2019 - 14.08.2019) durchgeführt werden. Im R-Skript **download_specific_sensors.R** wird der Beobachtungszeitraum spezifiziert. Außerdem muss ein Verzeichnis angegeben werden, in welchem sich die zuvor erstellte Textdatei befindet (*workfiles*). Im gleichen Verzeichnis wird das Ergebnis des Downloads gespeichert, eine csv-Datei für jeden Sensortyp mit allen Daten der Sensoren dieses Typs (d.h. insgesamt vier Dateien). Im anderen spezifizierten Verzeichnis (*data*) wird für jeden Sensor eine eigene csv-Datei gespeichert.

Nachdem die zuvor erstellte Textdatei eingelesen wurde, kann der Download für die gesamte Periode durchgeführt werden, wobei der aktuelle Downloadfortschritt angezeigt wird. Die Daten werden zunächst wie beschrieben für jeden Sensor einzeln im Verzeichnis *data* gespeichert, wobei wieder analog für jeden Sensortyp ein eigener Ordner existiert.

Im letzten Teil des Skripts werden dann je Sensortyp alle Sensoren zu einer großen Tabelle zusammengefügt und anschließend jeweils als csv-Datei im *workfiles*-Verzeichnis gespeichert.

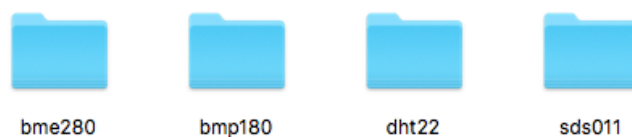


Figure 3: In München vorhandene Sensortypen

Die heruntergeladenen Daten können nun bearbeitet werden. Für jeden Sensortyp existiert eine csv-Datei im *workfiles*-Verzeichnis und für jeden Sensor eine eigene Datei im *data*-Verzeichnis im jeweiligen Sensortyp-Ordner. Die individuellen Sensordateien im *data*-Verzeichnis sind für die weitere Analyse nicht mehr von Nöten und können daher bei Bedarf wieder gelöscht werden. Das nächste Zwischenziel ist es die Daten zu bereinigen, sowie die Messungen der vier Sensortypen in einer Tabelle zu verknüpfen.