

# Исходные данные

## 1. Данные стран ЕС о концентрации загрязнителей

### 1.1. Источник

<https://discomap.eea.europa.eu/map/fme/AirQualityExport.htm>

### 1.2. Способ выгрузки

Для выгрузки файлов csv был разработан модуль [aqreport\\_loader](#).

### 1.3. Место хранения

<https://disk.yandex.ru/d/SVvFER5hMgcKtA>

### 1.4. Примечание

Общий объем выгруженных данных превышает 300Гб, поэтому они были заархивированы. На каждый загрязнитель приходится 1-2 архива с именем соответствующем коду загрязнителя. Поскольку расчет Индекса качества воздуха выполняется для каждого загрязнителя независимо, можно в процессе работы распаковывать по одному архиву.

Для предварительной оценки данных рядом с архивами размещена папка [samples](#).

### 1.5. Описание данных

#### 1.5.1. *Общее описание файлов и папок*

Данные разбиты по заархивированным папкам. Название папки соответствует коду загрязнителя.

Данные хранятся в csv файлах. Название файла содержит информацию о параметрах данных в следующем формате:

**[Код страны]\_[код загрязнителя]\_[внутренний код станции]\_[год]\_timeseries.csv**

Например: **AT\_1\_49911\_2013\_timeseries.csv**

Данные о станциях (расположении, измерительном оборудовании и проч) содержатся в файле **[PanEuropean\\_metadata.csv](#)**.

### 1.5.2. Коды и описание загрязнителей

| Код  | Наим-е            | Особенности загрязнителя   |
|------|-------------------|--|
| 1    | SO <sub>2</sub>   | Сжигание ископаемого топлива, особенно угля и нефти. Крупные заводы и энергетические предприятия выбрасывают около 90% SO <sub>2</sub> в атмосферу.<br><br>Естественный источник – извержения вулканов, лесные пожары.<br><br>Маркер PM <sub>2.5</sub>   |
| 5    | PM <sub>10</sub>  | Частицы размером от 2,5 до 10 микрон.<br><br>Сжигание топлива, отходов, добыча полезных ископаемых, строительство<br><br>Естественные источники – морская соль, пыль в пустынях.   |
| 7    | O <sub>3</sub>    | Ground level ozon. Формируется в результате реакции между NO <sub>x</sub> , CO, CH <sub>3</sub> и VOC (volatile organic compounds) под воздействием солнца.<br><br>Концентрация возрастает с увеличением высоты над уровнем моря.<br><br>Метан 80 % - с.х., а также отходы, добыча угля, LDGT.<br><br>Аммиак (NH <sub>3</sub> ) – 90% - с.х. |
| 8    | NO <sub>2</sub>   | Сжигание топлива, биомассы, производственные циклы<br><br>Естественный источник – небольшой вклад, молнии, пожары, биологические процессы  |
| 10   | CO                | Транспорт (особенно без катализаторов).<br>Естественный источник – незначителен.<br><br>Может реагировать с другими загрязнителями, образуя ground level ozon  |
| 6001 | PM <sub>2.5</sub> | Частицы размером менее 2,5 микрон.<br><br>Сжигание топлива, отходов, добыча полезных ископаемых, строительство.<br><br>Также формируется в результате реакции между NO <sub>x</sub> , CO, NH <sub>3</sub> и VOC (volatile organic compounds).<br><br>Около 40% от транспорта.<br><br>Естественные источники – морская соль, пыль в пустынях. |

### 1.5.3. Описание структуры данных в csv файлах timeseries

| Столбец                  | Тип      | Описание  |
|--------------------------|----------|---|
| Countrycode              | String   | Country iso code  |
| Namespace                | String   | Unique namespace as provided by the country   |
| AirQualityNetwork        | String   | Network identifier  |
| AirQualityStation        | String   | Localid of the station  |
| AirQualityStationEoICode | String   | Unique station identifier as used in the past AirBase system  |
| Samplingpoint            | String   | Localid of the samplingpoint  |
| SamplingProcess          | String   | Localid of the samplingprocess  |
| Sample                   | String   | Localid of the sample (also known as the feature of interest)   |
| AirPollutant             | String   | Short name of pollutant. Full list:<br><a href="http://dd.eionet.europa.eu/vocabulary/aq/pollutant/view">http://dd.eionet.europa.eu/vocabulary/aq/pollutant/view</a>  |
| AirPollutantCode         | String   | Reference (URL) to the definition of the pollutant in data dictionary   |
| AveragingTime            | String   | Defines the time for which the measure have been taken (hour, day, etc)<br><br><i>Примечание: при первом осмотре обнаружен только min hour</i>  |
| Concentration            | Value    | The measured value/concentration  |
| UnitOfMeasurement        | String   | Defines the unit of the concentration<br><br><i>Примечание: при первом осмотре единицы измерения единственны для загрязнителя и совпадают с требованиями к таковым</i>  |
| DateTimeBegin            | Datetime | Defines the start time (yyyy-mm-dd hh:mm:ss Z) of the measurement (includes timezone)<br><br><i>Примечание: проверить, что дата начала и конца измерения всегда отличаются на час</i>   |
| DateTimeEnd              | Datetime | Defines the end time (yyyy-mm-dd hh:mm:ss Z) of the measurement (includes timezone)   |
| Validity                 | Integer  | The validity flag for the measurement. See <a href="http://dd.eionet.europa.eu/vocabulary/aq/observationvalidity/view">http://dd.eionet.europa.eu/vocabulary/aq/observationvalidity/view</a><br><br><i>Примечание: брать с Validity &gt; 0</i>              |
| Verification             | Integer  | The verification flag for the measurement. See <a href="http://dd.eionet.europa.eu/vocabulary/aq/observationverification/view">http://dd.eionet.europa.eu/vocabulary/aq/observationverification/view</a><br><br><i>Примечание: брать с Verification = 1</i> |

**Зеленый** – нужные столбцы;

**серый** – совершенно ненужные столбцы;

**оранжевый** – столбцы, требующие проверки предположения;

**синий** – столбцы для фильтрации пригодных данных.

#### 1.5.4. Описание структуры данных в csv файле metadata

| Столбец                  | Тип      | Описание  |
|--------------------------|----------|---|
| Countrycode              | String   | Country iso code  |
| Timezone                 | String   | Timezone used by the country to report dates  |
| Namespace                | String   | Namespace of network  |
| AirQualityNetwork        | String   | Network identifier  |
| AirQualityStation        | String   | Localid of station  |
| AirQualityStationEoICode | String   | Station EoI code as used in the past AirBase system   |
| AirQualityStationNatCode | String   | National code of station  |
| SamplingPoint            | String   | Localid of samplingpoint  |
| SamplingProcess          | String   | Localid of process  |
| Sample                   | String   | Localid of sample (also known as feature of interest)   |
| AirPollutantCode         | String   | Reference (URL) to pollutant definition in data dictionary  |
| ObservationDateBegin     | Datetime | Observation's start time (yyyy-mm-dd hh:mm:ss)  |
| ObservationDateEnd       | Datetime | Observation's end time (yyyy-mm-dd hh:mm:ss)  |
| Projection               | String   | Reference to projection   |
| Longitude                | Decimal  | Longitude of samplingpoint  |
| Latitude                 | Decimal  | Latitude of samplingpoint   |
| Altitude                 | Integer  | Altitude of samplingpoint (m)   |
| MeasurementType          | String   | Reference to measurement type (full list at <a href="http://dd.eionet.europa.eu/vocabulary/aq/measurementtype/view">http://dd.eionet.europa.eu/vocabulary/aq/measurementtype/view</a> )                       |
| AirQualityStationType    | String   | Short name of station type (full list at <a href="http://dd.eionet.europa.eu/vocabulary/aq/stationclassification/view">http://dd.eionet.europa.eu/vocabulary/aq/stationclassification/view</a> )              |
| AirQualityStationArea    | String   | Reference to area code. Full list at <a href="http://dd.eionet.europa.eu/vocabulary/aq/areaclassification/view">http://dd.eionet.europa.eu/vocabulary/aq/areaclassification/view</a>                          |
| EquivalenceDemonstrated  | String   | Reference to demonstrated equivalence. Full list at <a href="http://dd.eionet.europa.eu/vocabulary/aq/equivalencedemonstrated/view">http://dd.eionet.europa.eu/vocabulary/aq/equivalencedemonstrated/view</a> |
| MeasurementEquipment     | String   | Refernce (URL) to measurement equipment   |
| InLetHeight              | Integer  | Height of inlet (m). Value of -999 indicate unknown   |
| BuildingDistance         | Integer  | Distance to building (m). Value of -999 indicate unknown  |
| KerbDistance             | Integer  | Distance to kerp (m). Value of -999 indicate unknown  |

**Зеленый** – нужные столбцы;

**серый** – совершенно ненужные столбцы;