**MIS Monthly Report**

**Useful info**

* historical state files location - \10.2.100.63\g$\ETL\03\_SCADA\OperationalData
* state files status list web page in reporting software website - <https://reporting.wrldc.in/POSOCOUI/FileUpload/OperationalDataFileList>
* demand met = restricted demand = catered demand (incl aux consumption)
* unrestricted demand = restricted demand + load shedding (scheduled + unscheduled)

**Workflow**

* Get the files info from 'file\_mapping' sheet
* Iterate through each file
* Open file and store in openpyxl object
* Get the entity metrics info in 'meas\_info' sheet
* Get address for each entity metric and store in database table in the format (timestamp, entity\_tag, metric\_name, value)

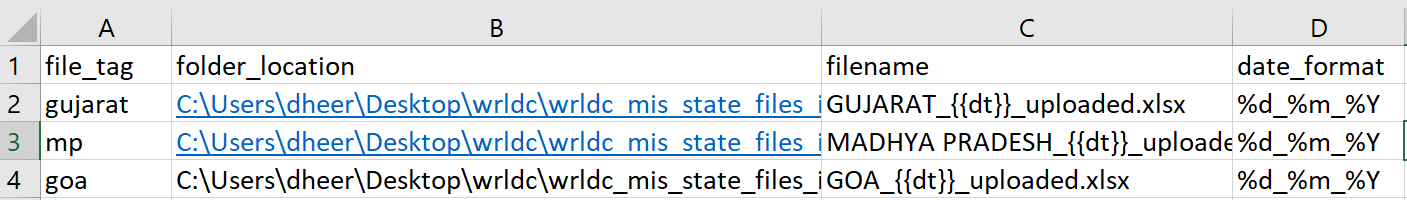
**Aggregations and equations while reading data from file**

* The address column may have multiple cells
* The aggregation strategy can be any one of 'average', 'sum' or 'equation'
* If the aggregation strategy is 'equation', then the metric value will be the evaluated expression by substituting the values in equation placeholders ({{0}} will be substituted by first address value, {{1}} will be substituted by second address value etc)

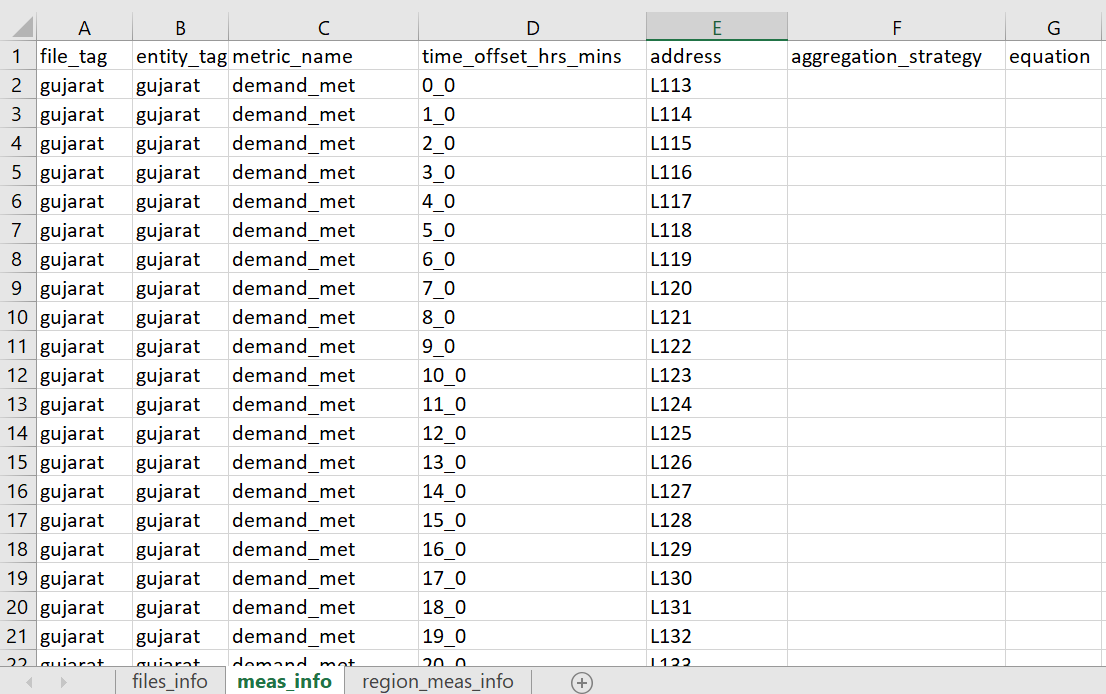
**Links**

* read excel sheet cell value by address from xlsx file using openpyxl python - <https://stackoverflow.com/questions/22613272/how-to-access-the-real-value-of-a-cell-using-the-openpyxl-module-for-python>
* marge multiple columns to make timestamp- <https://stackoverflow.com/questions/37955332/python-adding-hours-to-pandas-timestamp>
* convert multiple columns to single column using melt (long data frame to short)- https://pandas.pydata.org/docs/reference/api/pandas.melt.html

**Files info location:**

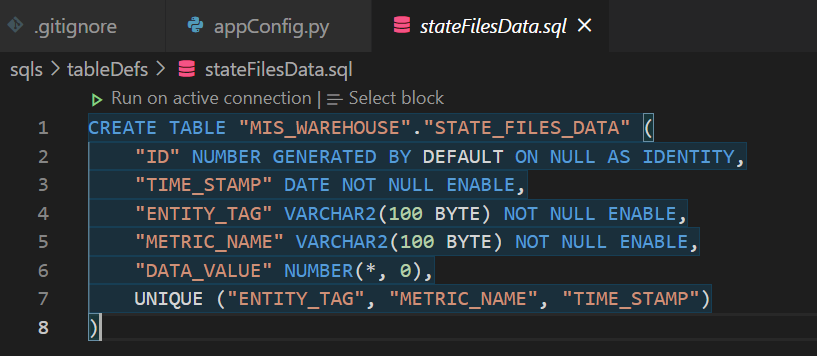


**Measurements info location:**

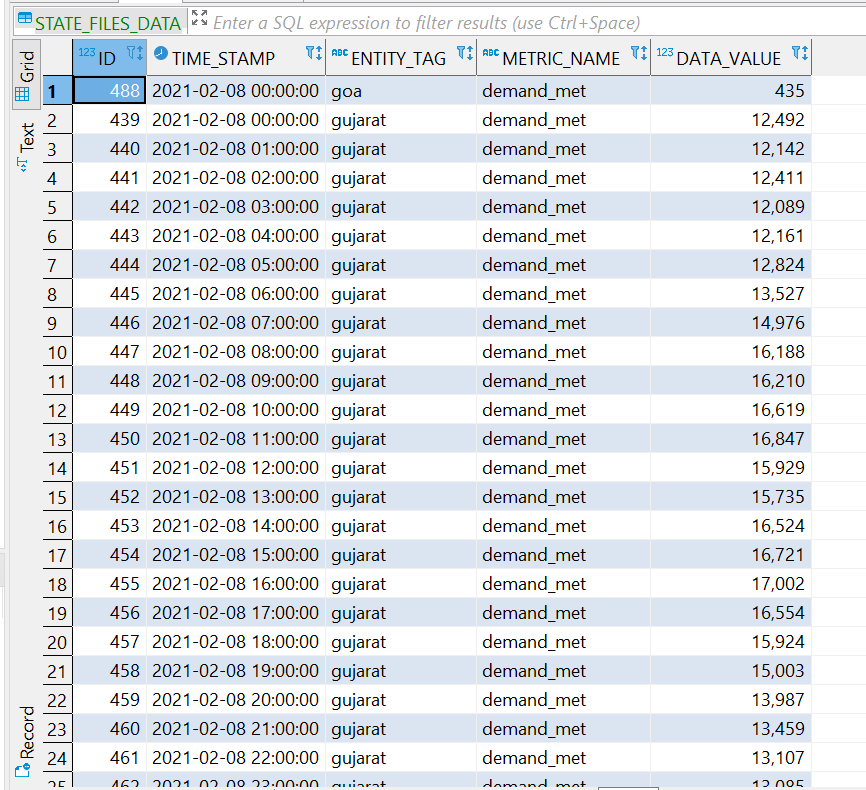


**Database Design:**

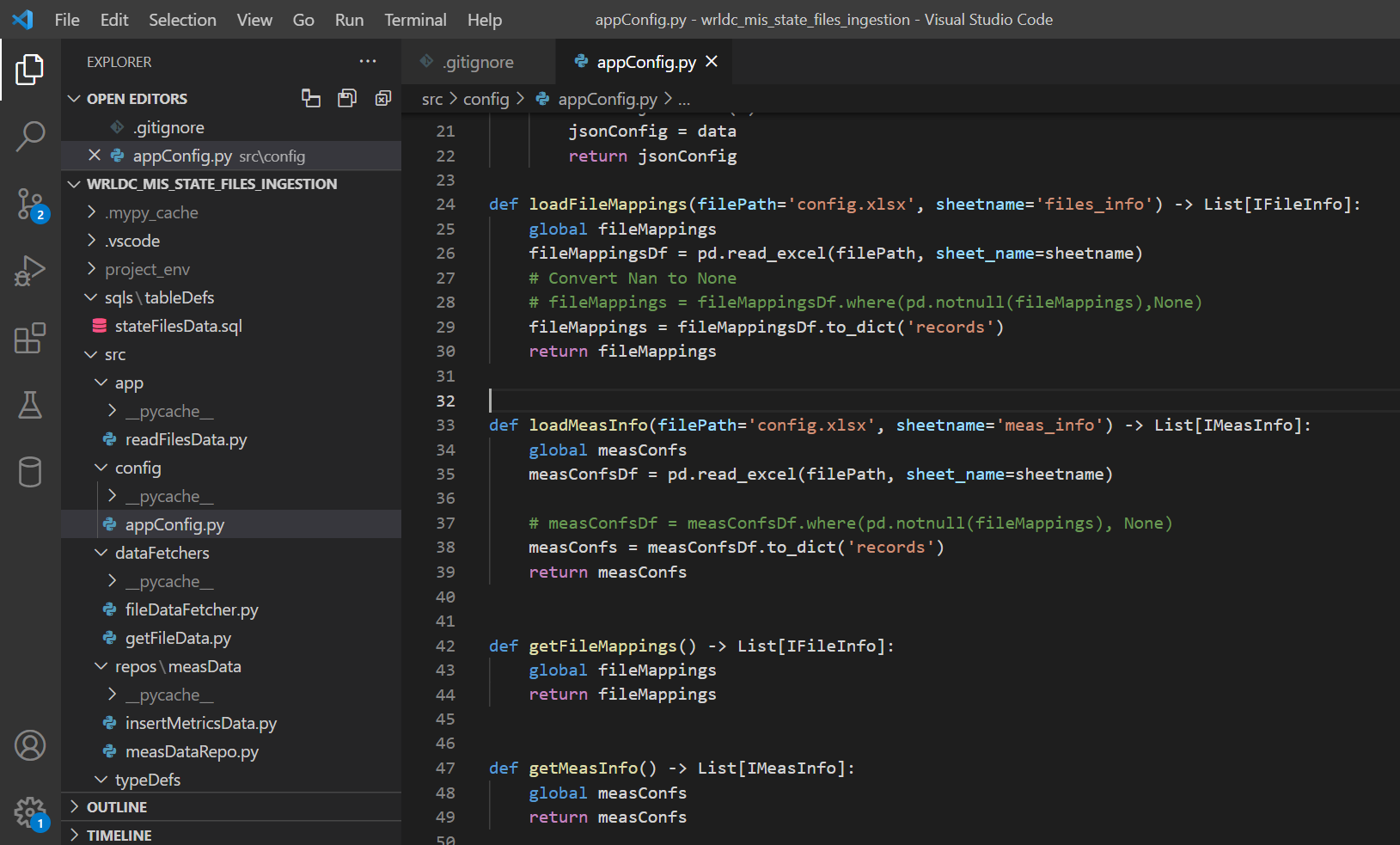
**Table Structure:**

****

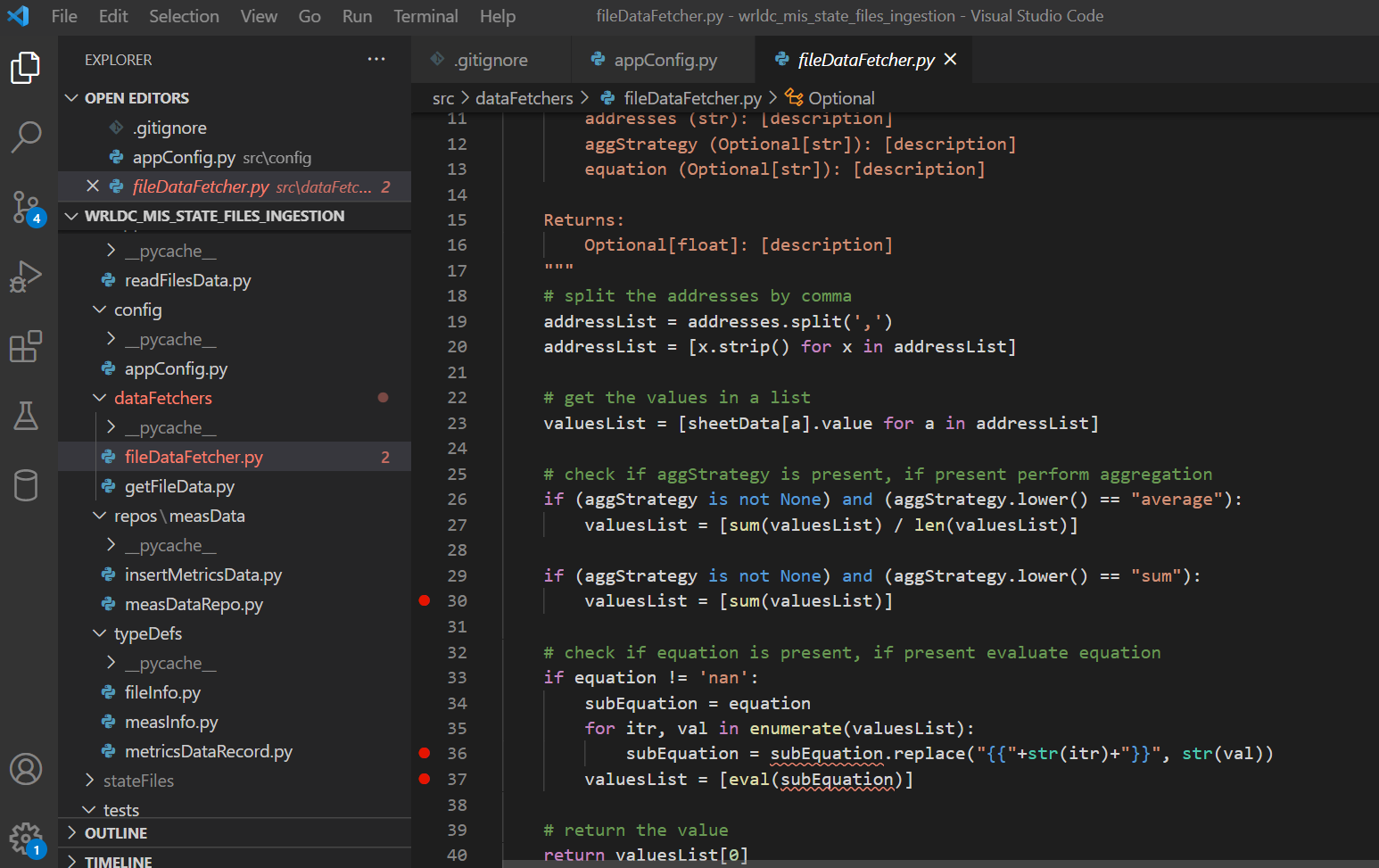
**Data Sample:**



**Code Workflow:**



**To accommodate the special cases like Goa and Madhya Pradesh:**



**To handle the tests cases:**

