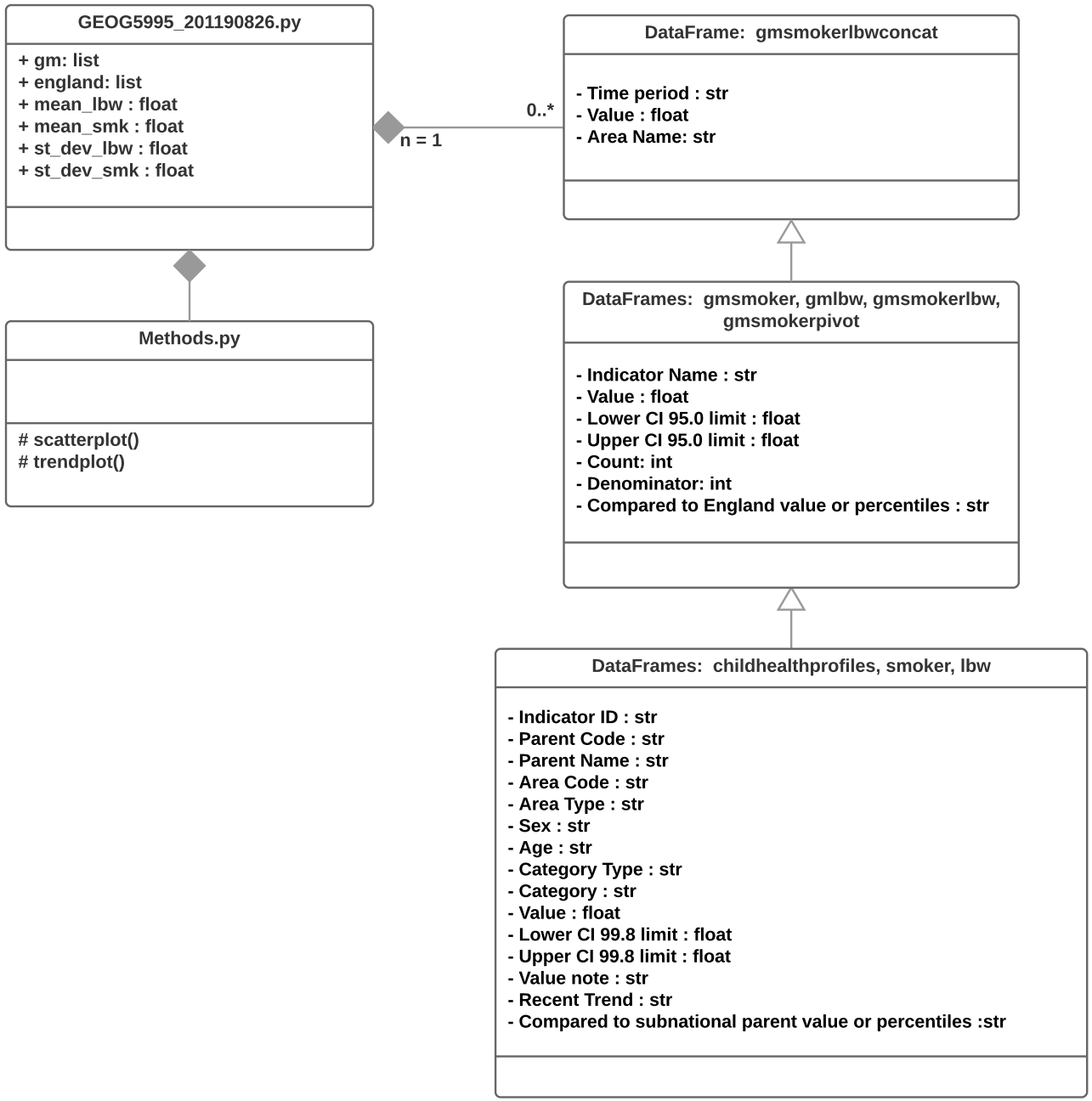
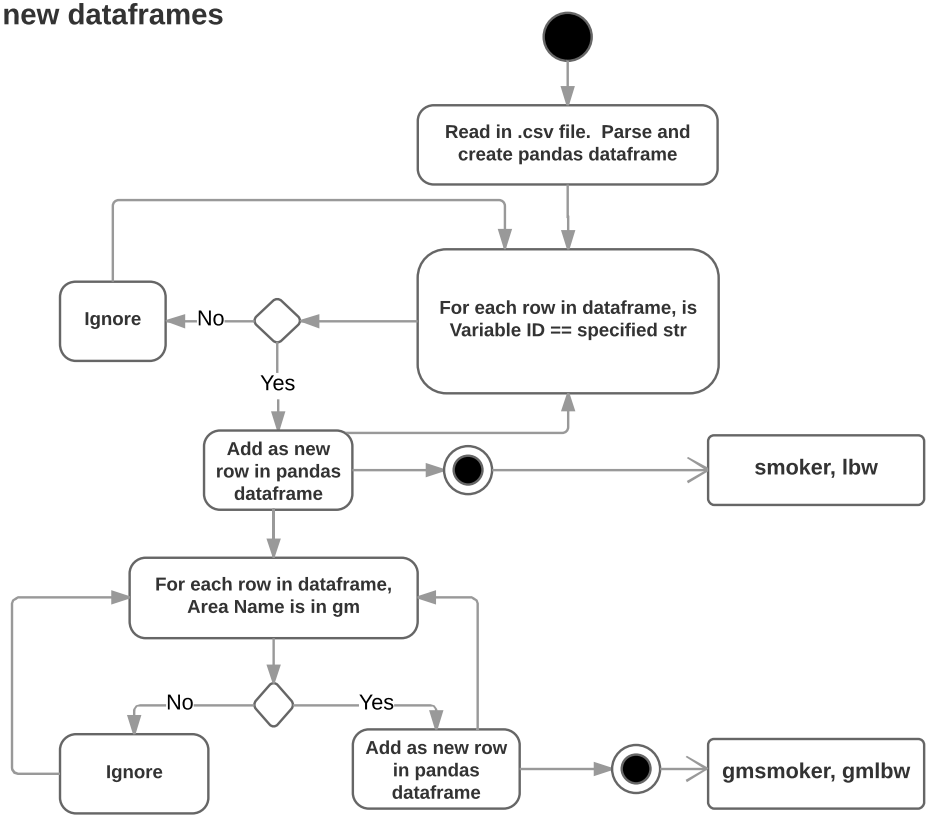
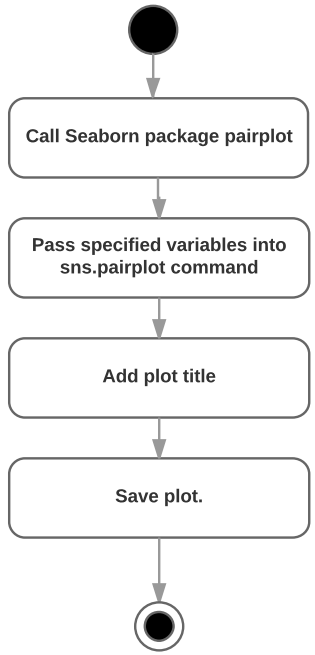

GEOG5995_201190826.py
+ gm: list
+ england: list
+ mean_lbw : float
+ mean_smk : float
+ st_dev_lbw : float
+ st_dev_smk : float
0..*
n = 1
DataFrame: gmsmokerlbwconcat
- Time period : str
- Value : float
- Area Name: str
Methods.py
scatterplot()
trendplot()
DataFrames: gmsmoker, gmlbw, gmsmokerlbw, gmsmokerpivot
- Indicator Name : str
- Value : float
- Lower CI 95.0 limit : float
- Upper CI 95.0 limit : float
- Count: int
- Denominator: int
- Compared to England value or percentiles : str
DataFrames: childhealthprofiles, smoker, lbw
- Indicator ID : str
- Parent Code : str
- Parent Name : str
- Area Code : str
- Area Type : str
- Sex : str
- Age : str
- Category Type : str
- Category : str
- Value : float
- Lower CI 99.8 limit : float
- Upper CI 99.8 limit : float
- Value note : str
- Recent Trend : str
- Compared to subnational parent value or percentiles :str

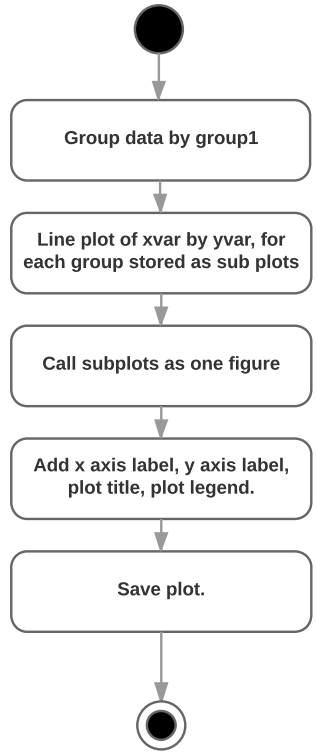
Select variables and cases into new dataframes



scatterplot(x, y, data, hue, size, scattertitle, fname2)



trendplot (data, group1, xvar, yvar, xlab, ylab, plottitle, fname)



Notes to UML

Two functions scatterplot() and trendplot() are defined (as shown above) in a separate script and called from within the main script.

The script manipulates several data frames, placing variables into different dataframes for easier visualisation and to be writtne to an excel workbook containing only those data of interest from a larger dataset. The activity diagram for sleecting variables and cases is shown above but other manipulations are omited. Standard functions from within the Pandas library are used to concatenate, pivot and append variables selected out of the larger dataframe using a list object as a reference.

In total eight dataframes are produced, these are written out to an excel workbook. All are exported as tidy objects (Wickham, Hadley . "Tidy Data."*Journal of Statistical Software*[Online], 59.10 (2014): 1 - 23. Web. 30 Nov. 2017).