

1. Data: Pensioners.csv is a simulated data for pensioner of country XXX between 2015 and 2040. The variable of concern is the LastPension and the gender, all data are sorted by year.

	A	B	C	D	E	F
1	year	RecNo	LASTPENSION	age	TYPEOFTERMINATIION	GENDER
2	2015	1	14072.38	57	4	2
3	2015	2	14290.69	60	4	2
4	2015	3	19261.51	78	8	2
5	2015	4	4330.9	46	7	1
6	2015	5	11875.97	56	4	1
7	2015	6	1822.16	93	8	2
8	2015	7	2468.18	80	3	2
9	2015	8	18625.69	74	8	2
10	2015	9	1500	91	8	2
11	2015	10	12788	67	8	2
12	2015	11	3192.7	76	3	2
13	2015	12	14097	55	4	2
14	2015	13	13288.35	75	4	2

2. Write a Python program to read in this file and process it row by row. Using the aggregator you created in last homework to compute the annual indicators of following:
  - a. Total pension
  - b. Total pension of all males
  - c. Total pension of all females
  - d. Average age of all pensioners
  - e. Number of all pensioners
  - f. Average pension amount (defined as total\_pension/number\_of\_pensioners)
  - g. Same as (f.) but for male only
  - h. Same as (f.) but for female only
3. Hint:
  - a. Since the data is already sorted by year, you don't need to sort it.
  - b. Do not read in the whole dataset, try to process everything row by row. (imaging what will happen if this file is 20G or 1T in size)
  - c. Create as many aggregators as you may need
  - d. Output all indicators on the last row of year block
  - e. What function can you add to the aggregator class so that you only need to create them once and reuse them?