Data Mining Assignment 1

Group members:
Y. Sai Dheeraj (S20150010056)
M. Shariq Suhail (S20150010034)

OS Details
OS name: Ubuntu
OS version: 18.04.1 LTS
Compiler used: gcc
gcc version: 7.3.0

Algorithm implemented: FP-growth frequent pattern mining algorithm.

Input 1:

1,A,B,D,E
2,B,C,E
3,A,B,D,E
4,A,B,C,E
5,A,B,C,D,E
6,B,C,D

Output for minimum support count = 3 (50%)

```
Output after Running FP growth algorithm

B  6
E  5
E, B  5
A  4
A, B  4
A, E  4
A, B, E  4
D  4
D, E  3
D, A  3
D, B  4
D, E, A  3
D, E, B  3
D, A, B  3
D, E, A, B  3
C  4
C, E  3
C, B  4
C, E, B  3
```

Output for minimum support count = 2 (33.33%)

```
Output after Running FP growth algorithm

B  6
E  5
E, B  5
A  4
A, B  4
A, E  4
A, B, E  4
D  4
D, E  3
D, A  3
D, B  4
D, E, A  3
D, E, B  3
D, A, B  3
D, E, A, B  3
C,
C, B  4
C, E  3
C, E, B  3
C, A  2
C, A, B  2
C, A, E  2
C, A, B, E  2
C, D  2
C, D, B  2
```

Input 2:

Data file: groceries_subset.csv

Output for minimum support count = 150 (15%)

```
whole milk  269
rolls/buns  222
other vegetables  186
soda  159
```

Output for minimum support count = 50 (5%)

```
whole milk  269
rolls/buns  222
rolls/buns, whole milk  68
other vegetables  186
other vegetables, whole milk  72
soda  159
bottled water  132
yogurt  127
yogurt, whole milk  56
root vegetables  110
tropical fruit  97
citrus fruit  95
newspapers  89
pastry  78
sausage  78
shopping bags  78
coffee  76
bottled beer  76
canned beer  76
curd  75
whipped/sour cream  74
frankfurter  73
beef  69
fruit/vegetable juice  67
margarine  56
pork  53
domestic eggs  53
brown bread  50
```

Note:
1. The numbers printed in the output are the frequencies of the items.
2. Please update the values of the macros used in the code appropriately before execution inorder to avoid segmentation faults.