In [8]:
```
# importing pandas module
import pandas as pd
### making data frame
df = pd.read_csv("C:/Users/Admin/Desktop/Sem 1/_births_and_deaths.csv")
#print(df)
df2 = pd.DataFrame(df)
#df2
# Prints no. of rows and columns of a DataFrame
print("The size of the dataframe is")
print(df2.shape)
# Prints the first 5 rows of a DataFrame as default
print("The first five rows of dataframe is")
print(df2.head(5))

# prints first 5 rows and every column which replicates df.head()
print("Another way of displaying the first five rows of dataframe is")
print(df.iloc[0:5,:])

#### STAT ANALYSIS
# computes various summary statistics, excluding NaN values
print(df2.describe())
# for computing correlations
print(df2.corr())
print(df2.sort_index())
# computes numerical data ranks
df2.rank()
```

```
The size of the dataframe is
(52, 5)
The first five rows of dataframe is
  Quarter  Male Live Births  Female Live Births  Male Deaths  Female Deaths
0  2000Q1              7639                7139         3346           3070
1  2000Q2              7365                6866         3372           3178
2  2000Q3              7174                6843         3675           3511
3  2000Q4              6979                6600         3357           3151
4  2001Q1              7496                7232         3231           3070
Another way of displaying the first five rows of dataframe is
  Quarter  Male Live Births  Female Live Births  Male Deaths  Female Deaths
0  2000Q1              7639                7139         3346           3070
1  2000Q2              7365                6866         3372           3178
2  2000Q3              7174                6843         3675           3511
3  2000Q4              6979                6600         3357           3151
4  2001Q1              7496                7232         3231           3070
       Male Live Births  Female Live Births  Male Deaths  Female Deaths
count         52.000000           52.000000    52.000000      52.000000
mean        7639.750000         7263.673077  3549.961538    3556.788462
std          506.576548          445.757682   272.253844     329.419890
min         6713.000000         6438.000000  3103.000000    3070.000000
25%         7275.000000         6864.250000  3354.250000    3319.250000
50%         7635.000000         7307.000000  3498.500000    3487.500000
75%         8037.250000         7544.250000  3680.000000    3699.750000
max         8756.000000         8212.000000  4149.000000    4287.000000
                    Male Live Births  Female Live Births  Male Deaths  \
Male Live Births            1.000000            0.963614     0.137545
Female Live Births          0.963614            1.000000     0.197679
Male Deaths                 0.137545            0.197679     1.000000
Female Deaths               0.125531            0.178861     0.950433

                    Female Deaths
Male Live Births         0.125531
Female Live Births       0.178861
Male Deaths              0.950433
Female Deaths            1.000000
   Quarter  Male Live Births  Female Live Births  Male Deaths  Female Deaths
0   2000Q1              7639                7139         3346           3070
1   2000Q2              7365                6866         3372           3178
2   2000Q3              7174                6843         3675           3511
3   2000Q4              6979                6600         3357           3151
4   2001Q1              7496                7232         3231           3070
5   2001Q2              7101                6796         3481           3392
6   2001Q3              6873                6783         3914           4000
7   2001Q4              6863                6655         3357           3380
8   2002Q1              6891                6757         3265           3258
9   2002Q2              6713                6438         3523           3365
10  2002Q3              7061                6634         3959           4003
11  2002Q4              6912                6615         3276           3416
12  2003Q1              7425                6859         3103           3160
13  2003Q2              6881                6555         3403           3444
14  2003Q3              7125                6842         4073           4113
15  2003Q4              7389                7058         3441           3273
16  2004Q1              7650                7370         3245           3204
17  2004Q2              7383                7053         3411           3374
18  2004Q3              7505                6972         3874           4104
19  2004Q4              7206                6934         3545           3662
```

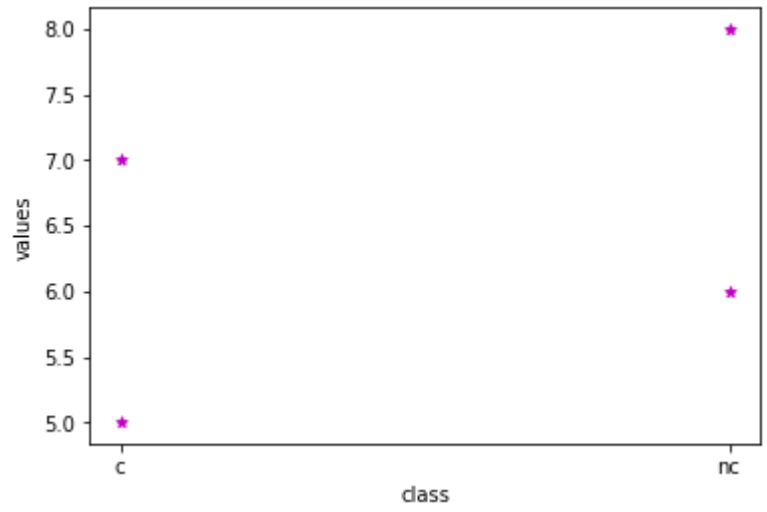
20	2005Q1	7499	7063	3129	3145
21	2005Q2	7509	7298	3485	3493
22	2005Q3	7281	6825	3577	3690
23	2005Q4	7257	7013	3240	3275
24	2006Q1	7800	7458	3319	3352
25	2006Q2	7492	7124	3363	3482
26	2006Q3	7631	7340	3966	4185
27	2006Q4	7317	7031	3276	3302
28	2007Q1	8602	7934	3450	3436
29	2007Q2	8093	7664	3403	3444
30	2007Q3	8068	7653	3870	3965
31	2007Q4	8250	7780	3552	3402
32	2008Q1	7948	7796	3342	3317
33	2008Q2	8597	8053	3617	3625
34	2008Q3	8262	7854	4075	4198
35	2008Q4	8295	7538	3501	3513
36	2009Q1	8092	7468	3300	3320
37	2009Q2	8019	7436	3534	3520
38	2009Q3	8264	8047	3999	3994
39	2009Q4	7737	7480	3647	3650
40	2010Q1	8756	8212	3289	3211
41	2010Q2	8107	7518	3539	3513
42	2010Q3	8084	7838	3993	3951
43	2010Q4	7957	7425	3402	3540
44	2011Q1	8370	7884	3559	3609
45	2011Q2	7676	7425	3573	3698
46	2011Q3	7961	7563	4079	4247
47	2011Q4	7469	7055	3612	3705
48	2012Q1	8027	7684	3496	3401
49	2012Q2	7804	7468	3695	3611
50	2012Q3	7639	7316	4149	4287
51	2012Q4	7773	7467	3716	3744

Out[8]:

	Quarter	Male Live Births	Female Live Births	Male Deaths	Female Deaths
0	1.0	27.5	24.0	13.0	1.5
1	2.0	16.0	14.0	17.0	6.0
2	3.0	11.0	12.0	39.0	28.0
3	4.0	7.0	3.0	14.5	4.0
4	5.0	22.0	25.0	3.0	1.5
5	6.0	9.0	9.0	24.0	19.0
6	7.0	3.0	8.0	44.0	45.0
7	8.0	2.0	6.0	14.5	18.0
8	9.0	5.0	7.0	6.0	9.0
9	10.0	1.0	1.0	28.0	16.0
10	11.0	8.0	5.0	45.0	46.0
11	12.0	6.0	4.0	7.5	22.0
12	13.0	19.0	13.0	1.0	5.0
13	14.0	4.0	2.0	19.5	24.5
14	15.0	10.0	11.0	49.0	48.0
15	16.0	18.0	21.0	22.0	10.0
16	17.0	29.0	29.0	5.0	7.0
17	18.0	17.0	19.0	21.0	17.0
18	19.0	24.0	16.0	43.0	47.0
19	20.0	12.0	15.0	31.0	37.0
20	21.0	23.0	22.0	2.0	3.0
21	22.0	25.0	26.0	25.0	27.0
22	23.0	14.0	10.0	35.0	38.0
23	24.0	13.0	17.0	4.0	11.0
24	25.0	33.0	33.0	11.0	15.0
25	26.0	21.0	23.0	16.0	26.0
26	27.0	26.0	28.0	46.0	49.0
27	28.0	15.0	18.0	7.5	12.0
28	29.0	51.0	49.0	23.0	23.0
29	30.0	43.0	42.0	19.5	24.5
30	31.0	40.0	41.0	42.0	43.0
31	32.0	45.0	44.0	32.0	21.0
32	33.0	35.0	45.0	12.0	13.0
33	34.0	50.0	51.0	37.0	35.0
34	35.0	46.0	47.0	50.0	50.0

	Quarter	Male Live Births	Female Live Births	Male Deaths	Female Deaths
35	36.0	48.0	39.0	27.0	29.5
36	37.0	42.0	35.5	10.0	14.0
37	38.0	38.0	32.0	29.0	31.0
38	39.0	47.0	50.0	48.0	44.0
39	40.0	31.0	37.0	38.0	36.0
40	41.0	52.0	52.0	9.0	8.0
41	42.0	44.0	38.0	30.0	29.5
42	43.0	41.0	46.0	47.0	42.0
43	44.0	36.0	30.5	18.0	32.0
44	45.0	49.0	48.0	33.0	33.0
45	46.0	30.0	30.5	34.0	39.0
46	47.0	37.0	40.0	51.0	51.0
47	48.0	20.0	20.0	36.0	40.0
48	49.0	39.0	43.0	26.0	20.0
49	50.0	34.0	35.5	40.0	34.0
50	51.0	27.5	27.0	52.0	52.0
51	52.0	32.0	34.0	41.0	41.0

In [10]:
```
# importing pandas module
import pandas as pd
import matplotlib.pyplot as plt
### making data frame
df1 = pd.read_csv("C:/Users/Admin/Desktop/Sem 1/_births_and_deaths.csv")
#print(df)
df2 = pd.DataFrame(df)
df = pd.DataFrame({"a":["c", "nc", "c", "nc"],
"b":[5, 6, 7, 8]})
# plot a histogram
#df['1965'].hist(bins=10)
# shows presence of a lot of outliers/extreme values
#df.boxplot(column='1975', by = '1965')
# plotting points as a scatter plot
x = df["a"]
y = df["b"]
plt.scatter(x, y, label= "stars", color= "m",
marker= "*", s=30)
# x-axis label
plt.xlabel('class')
# frequency label
plt.ylabel('values')
# function to show the plot
plt.show()
```



In [ ]:

In [ ]: