# Question 1

Step 1

Step 2 and 3

```
   user_id  age gender  occupation zip_code
0        1   24      M  technician    85711
1        2   53      F       other    94043
2        3   23      M      writer    32067
3        4   24      M  technician    43537
4        5   33      F       other    15213
```

Step 4

```
Mean age per occupation:
       occupation   mean_age
0   administrator  38.746835
1          artist  31.392857
2          doctor  43.571429
3        educator  42.010526
4        engineer  36.388060
5   entertainment  29.222222
6       executive  38.718750
7      healthcare  41.562500
8       homemaker  32.571429
9          lawyer  36.750000
10      librarian  40.000000
11      marketing  37.615385
12           none  26.555556
13          other  34.523810
14     programmer  33.121212
15        retired  63.071429
16       salesman  35.666667
17      scientist  35.548387
18        student  22.081633
19     technician  33.148148
20         writer  36.311111
```

Step 5

```
Male ratio per occupation (sorted descending):
occupation
doctor           1.000000
engineer         0.970149
technician       0.962963
retired          0.928571
programmer       0.909091
executive        0.906250
scientist        0.903226
entertainment    0.888889
lawyer           0.833333
salesman         0.750000
educator         0.726316
student          0.693878
other            0.657143
marketing        0.615385
writer           0.577778
none             0.555556
```

```
administrator    0.544304
artist           0.535714
librarian        0.431373
healthcare       0.312500
homemaker        0.142857
dtype: float64
```

Step 6

```
For each occupation, the minimum and maximum ages:
               min_age  max_age
occupation
administrator       21       70
artist              19       48
doctor              28       64
educator            23       63
engineer            22       70
entertainment       15       50
executive           22       69
healthcare          22       62
homemaker           20       50
lawyer              21       53
librarian           23       69
marketing           24       55
none                11       55
other               13       64
programmer          20       63
retired             51       73
salesman            18       66
scientist           23       55
student              7       42
technician          21       55
writer              18       60
```

Step 7

```
For each combination of occupation and sex, the mean age:
                         mean_age
occupation    gender
administrator F       40.638889
              M       37.162791
artist        F       30.307692
              M       32.333333
doctor        M       43.571429
educator      F       39.115385
              M       43.101449
engineer      F       29.500000
              M       36.600000
entertainment F       31.000000
              M       29.000000
executive     F       44.000000
              M       38.172414
healthcare    F       39.818182
              M       45.400000
homemaker     F       34.166667
              M       23.000000
lawyer        F       39.500000
              M       36.200000
librarian     F       40.000000
              M       40.000000
marketing     F       37.200000
              M       37.875000
none          F       36.500000
              M       18.600000
other         F       35.472222
```

```
               M       34.028986
programmer     F       32.166667
               M       33.216667
retired        F       70.000000
               M       62.538462
salesman       F       27.000000
               M       38.555556
scientist      F       28.333333
               M       36.321429
student        F       20.750000
               M       22.669118
technician     F       38.000000
               M       32.961538
writer         F       37.631579
               M       35.346154
```

Step 8

```
For each occupation the percentage of women and men:
occupation     gender
administrator  F          45.569620
               M          54.430380
artist         F          46.428571
               M          53.571429
doctor         M         100.000000
educator       F          27.368421
               M          72.631579
engineer       F           2.985075
               M          97.014925
entertainment  F          11.111111
               M          88.888889
executive      F           9.375000
               M          90.625000
healthcare     F          68.750000
               M          31.250000
homemaker      F          85.714286
               M          14.285714
lawyer         F          16.666667
               M          83.333333
librarian      F          56.862745
               M          43.137255
marketing      F          38.461538
               M          61.538462
none           F          44.444444
               M          55.555556
other          F          34.285714
               M          65.714286
programmer     F           9.090909
               M          90.909091
retired        F           7.142857
               M          92.857143
salesman       F          25.000000
               M          75.000000
scientist      F           9.677419
               M          90.322581
student        F          30.612245
               M          69.387755
technician     F           3.703704
               M          96.296296
writer         F          42.222222
               M          57.777778
dtype: float64
```

# Question 2

Step 1

Step 2 and 3

```
Euro12 data:
             Team  Goals  Shots on target  Shots off target Shooting Accuracy  \
0         Croatia      4               13                12             51.9%
1  Czech Republic      4               13                18             41.9%
2         Denmark      4               10                10             50.0%
3         England      5               11                18             50.0%
4          France      3               22                24             37.9%

  % Goals-to-shots  Total shots (inc. Blocked)  Hit Woodwork  Penalty goals  \
0            16.0%                          32             0              0
1            12.9%                          39             0              0
2            20.0%                          27             1              0
3            17.2%                          40             0              0
4             6.5%                          65             1              0

   Penalties not scored  ...  Saves made  Saves-to-shots ratio  Fouls Won  \
0                     0  ...          13                 81.3%         41
1                     0  ...           9                 60.1%         53
2                     0  ...          10                 66.7%         25
3                     0  ...          22                 88.1%         43
4                     0  ...           6                 54.6%         36

  Fouls Conceded  Offsides  Yellow Cards  Red Cards  Subs on  Subs off  \
0             62         2             9          0        9         9
1             73         8             7          0       11        11
2             38         8             4          0        7         7
3             45         6             5          0       11        11
4             51         5             6          0       11        11

   Players Used
0            16
1            19
2            15
3            16
4            19

[5 rows x 35 columns]

(16, 35)
```

Step 4

```
Onlt the goals column:
    Goals
0       4
1       4
2       4
3       5
4       3
5      10
6       5
7       6
8       2
9       2
```

```
10      6
11      1
12      5
13     12
14      5
15      2
```

Step 5

```
Teams participated in the Euro2012:
['Croatia' 'Czech Republic' 'Denmark' 'England' 'France' 'Germany'
 'Greece' 'Italy' 'Netherlands' 'Poland' 'Portugal' 'Republic of Ireland'
 'Russia' 'Spain' 'Sweden' 'Ukraine']
Total teams:  16
```

Step 6

```
Number of columns in the dataset:  35
```

Step 7

```
Discipline dataframe:
                   Team  Yellow Cards  Red Cards
0               Croatia             9          0
1        Czech Republic             7          0
2               Denmark             4          0
3               England             5          0
4                France             6          0
5               Germany             4          0
6                Greece             9          1
7                 Italy            16          0
8           Netherlands             5          0
9                Poland             7          1
10             Portugal            12          0
11  Republic of Ireland             6          1
12               Russia             6          0
13                Spain            11          0
14               Sweden             7          0
15              Ukraine             5          0
```

Step 8

```
Sorted discipline:
                   Team  Yellow Cards  Red Cards
6                Greece             9          1
9                Poland             7          1
11  Republic of Ireland             6          1
7                 Italy            16          0
10             Portugal            12          0
13                Spain            11          0
0               Croatia             9          0
1        Czech Republic             7          0
14               Sweden             7          0
4                France             6          0
12               Russia             6          0
3               England             5          0
8           Netherlands             5          0
15              Ukraine             5          0
2               Denmark             4          0
5               Germany             4          0
```

```
<ipython-input-16-2b310c00d462>:1: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_
```

```
guide/indexing.html#returning-a-view-versus-a-copy
  discipline.sort_values(by=['Red Cards', 'Yellow Cards'], ascending = False, inplace =
True)
```

Step 9

```
Mean yellow cards per team:
                    mean_yellow_cards
Team
Croatia                           9.0
Czech Republic                    7.0
Denmark                           4.0
England                           5.0
France                            6.0
Germany                           4.0
Greece                            9.0
Italy                            16.0
Netherlands                       5.0
Poland                            7.0
Portugal                         12.0
Republic of Ireland               6.0
Russia                            6.0
Spain                            11.0
Sweden                            7.0
Ukraine                           5.0
```

Step 10

```
Teams with more than 6 goals:
       Team  Goals  Shots on target  Shots off target Shooting Accuracy  \
5   Germany     10               32                32             47.8%
13    Spain     12               42                33             55.9%

   % Goals-to-shots  Total shots (inc. Blocked)  Hit Woodwork  Penalty goals  \
5             15.6%                          80             2              1
13            16.0%                         100             0              1

    Penalties not scored  ...  Saves made  Saves-to-shots ratio  Fouls Won  \
5                      0  ...          10                62.6%         63
13                     0  ...          15                93.8%        102

   Fouls Conceded  Offsides  Yellow Cards  Red Cards  Subs on  Subs off  \
5              49        12             4          0       15        15
13             83        19            11          0       17        17

    Players Used
5             17
13            18

[2 rows x 35 columns]
```

Step 11

```
Teams starting with a G:
      Team  Goals  Shots on target  Shots off target Shooting Accuracy  \
5  Germany     10               32                32             47.8%
6   Greece      5                8                18             30.7%

  % Goals-to-shots  Total shots (inc. Blocked)  Hit Woodwork  Penalty goals  \
5            15.6%                          80             2              1
6            19.2%                          32             1              1

   Penalties not scored  ...  Saves made  Saves-to-shots ratio  Fouls Won  \
5                     0  ...          10                62.6%         63
```

```
6                    1  ...          13                65.1%         67

  Fouls Conceded  Offsides  Yellow Cards  Red Cards  Subs on  Subs off  \
5             49        12             4          0       15        15
6             48        12             9          1       12        12

   Players Used
5            17
6            20

[2 rows x 35 columns]
```

Step 12

```
First 7 columns from euro12:
                   Team  Goals  Shots on target  Shots off target  \
0               Croatia      4               13                12
1        Czech Republic      4               13                18
2               Denmark      4               10                10
3               England      5               11                18
4                France      3               22                24
5               Germany     10               32                32
6                Greece      5                8                18
7                 Italy      6               34                45
8           Netherlands      2               12                36
9                Poland      2               15                23
10             Portugal      6               22                42
11  Republic of Ireland      1                7                12
12               Russia      5                9                31
13                Spain     12               42                33
14               Sweden      5               17                19
15              Ukraine      2                7                26

   Shooting Accuracy % Goals-to-shots  Total shots (inc. Blocked)
0              51.9%            16.0%                          32
1              41.9%            12.9%                          39
2              50.0%            20.0%                          27
3              50.0%            17.2%                          40
4              37.9%             6.5%                          65
5              47.8%            15.6%                          80
6              30.7%            19.2%                          32
7              43.0%             7.5%                         110
8              25.0%             4.1%                          60
9              39.4%             5.2%                          48
10             34.3%             9.3%                          82
11             36.8%             5.2%                          28
12             22.5%            12.5%                          59
13             55.9%            16.0%                         100
14             47.2%            13.8%                          39
15             21.2%             6.0%                          38
```

Step 13

```
All columns except last 3 columns from euro12:
                   Team  Goals  Shots on target  Shots off target  \
0               Croatia      4               13                12
1        Czech Republic      4               13                18
2               Denmark      4               10                10
3               England      5               11                18
4                France      3               22                24
5               Germany     10               32                32
6                Greece      5                8                18
7                 Italy      6               34                45
8           Netherlands      2               12                36
9                Poland      2               15                23
```

| | | | | |
|---|---|---|---|---|
| 10 | Portugal | 6 | 22 | 42 |
| 11 | Republic of Ireland | 1 | 7 | 12 |
| 12 | Russia | 5 | 9 | 31 |
| 13 | Spain | 12 | 42 | 33 |
| 14 | Sweden | 5 | 17 | 19 |
| 15 | Ukraine | 2 | 7 | 26 |

| | Shooting Accuracy % | Goals-to-shots | Total shots (inc. Blocked) | \ |
|---|---|---|---|---|
| 0 | 51.9% | 16.0% | 32 | |
| 1 | 41.9% | 12.9% | 39 | |
| 2 | 50.0% | 20.0% | 27 | |
| 3 | 50.0% | 17.2% | 40 | |
| 4 | 37.9% | 6.5% | 65 | |
| 5 | 47.8% | 15.6% | 80 | |
| 6 | 30.7% | 19.2% | 32 | |
| 7 | 43.0% | 7.5% | 110 | |
| 8 | 25.0% | 4.1% | 60 | |
| 9 | 39.4% | 5.2% | 48 | |
| 10 | 34.3% | 9.3% | 82 | |
| 11 | 36.8% | 5.2% | 28 | |
| 12 | 22.5% | 12.5% | 59 | |
| 13 | 55.9% | 16.0% | 100 | |
| 14 | 47.2% | 13.8% | 39 | |
| 15 | 21.2% | 6.0% | 38 | |

| | Hit Woodwork | Penalty goals | Penalties not scored | ... | Clean Sheets | \ |
|---|---|---|---|---|---|---|
| 0 | 0 | 0 | 0 | ... | 0 | |
| 1 | 0 | 0 | 0 | ... | 1 | |
| 2 | 1 | 0 | 0 | ... | 1 | |
| 3 | 0 | 0 | 0 | ... | 2 | |
| 4 | 1 | 0 | 0 | ... | 1 | |
| 5 | 2 | 1 | 0 | ... | 1 | |
| 6 | 1 | 1 | 1 | ... | 1 | |
| 7 | 2 | 0 | 0 | ... | 2 | |
| 8 | 2 | 0 | 0 | ... | 0 | |
| 9 | 0 | 0 | 0 | ... | 0 | |
| 10 | 6 | 0 | 0 | ... | 2 | |
| 11 | 0 | 0 | 0 | ... | 0 | |
| 12 | 2 | 0 | 0 | ... | 0 | |
| 13 | 0 | 1 | 0 | ... | 5 | |
| 14 | 3 | 0 | 0 | ... | 1 | |
| 15 | 0 | 0 | 0 | ... | 0 | |

| | Blocks | Goals conceded | Saves made | Saves-to-shots ratio | Fouls Won | \ |
|---|---|---|---|---|---|---|
| 0 | 10 | 3 | 13 | 81.3% | 41 | |
| 1 | 10 | 6 | 9 | 60.1% | 53 | |
| 2 | 10 | 5 | 10 | 66.7% | 25 | |
| 3 | 29 | 3 | 22 | 88.1% | 43 | |
| 4 | 7 | 5 | 6 | 54.6% | 36 | |
| 5 | 11 | 6 | 10 | 62.6% | 63 | |
| 6 | 23 | 7 | 13 | 65.1% | 67 | |
| 7 | 18 | 7 | 20 | 74.1% | 101 | |
| 8 | 9 | 5 | 12 | 70.6% | 35 | |
| 9 | 8 | 3 | 6 | 66.7% | 48 | |
| 10 | 11 | 4 | 10 | 71.5% | 73 | |
| 11 | 23 | 9 | 17 | 65.4% | 43 | |
| 12 | 8 | 3 | 10 | 77.0% | 34 | |
| 13 | 8 | 1 | 15 | 93.8% | 102 | |
| 14 | 12 | 5 | 8 | 61.6% | 35 | |
| 15 | 4 | 4 | 13 | 76.5% | 48 | |

| | Fouls Conceded | Offsides | Yellow Cards | Red Cards |
|---|---|---|---|---|
| 0 | 62 | 2 | 9 | 0 |
| 1 | 73 | 8 | 7 | 0 |
| 2 | 38 | 8 | 4 | 0 |
| 3 | 45 | 6 | 5 | 0 |

```
4               51         5              6          0
5               49        12              4          0
6               48        12              9          1
7               89        16             16          0
8               30         3              5          0
9               56         3              7          1
10              90        10             12          0
11              51        11              6          1
12              43         4              6          0
13              83        19             11          0
14              51         7              7          0
15              31         4              5          0

[16 rows x 32 columns]
```

Step 14

```
Shooting accuracy from England, Italy and Russia:
       Team Shooting Accuracy
3   England             50.0%
7     Italy             43.0%
12   Russia             22.5%
```

# Question 3

Step 1

```
Series 1:
 0     4
1     3
2     3
3     2
4     3
     ..
95    2
96    3
97    3
98    3
99    1
Length: 100, dtype: int64
Series 2:
 0     2
1     2
2     1
3     2
4     2
     ..
95    1
96    2
97    1
98    2
99    2
Length: 100, dtype: int64
Series 3:
 0     18589
1     10167
2     10414
3     25928
4     14552
      ...
95    13936
96    18983
97    28013
98    20548
99    24990
Length: 100, dtype: int64
```

Step 3

```
   0  1      2
0  4  2  18589
1  3  2  10167
2  3  1  10414
3  2  2  25928
4  3  2  14552
```

Step 4

```
After column name changed:
    bedrs  bathrs  price_sqr_meter
0      4       2            18589
1      3       2            10167
2      3       1            10414
```

```
3      2       2            25928
4      3       2            14552
```

Step 5

```
Bigcolumn dataframe:
 0         4
1         3
2         3
3         2
4         3
      ...
95    13936
96    18983
97    28013
98    20548
99    24990
Length: 300, dtype: int64
```

Step 6

```
Bigcolumn index:
Int64Index([ 0,  1,  2,  3,  4,  5,  6,  7,  8,  9,
            ...
            90, 91, 92, 93, 94, 95, 96, 97, 98, 99],
           dtype='int64', length=300)
```

Step 7

```
After index reset:
 0          4
1          3
2          3
3          2
4          3
       ...
295    13936
296    18983
297    28013
298    20548
299    24990
Length: 300, dtype: int64
```

# Question 4

Step 1

Step 2

```
Data:
    Yr  Mo  Dy    RPT    VAL    ROS    KIL    SHA   BIR    DUB    CLA    MUL  \
0  61   1   1  15.04  14.96  13.17   9.29    NaN  9.87  13.67  10.25  10.83
1  61   1   2  14.71    NaN  10.83   6.50  12.62  7.67  11.50  10.04   9.79
2  61   1   3  18.50  16.88  12.33  10.13  11.17  6.17  11.25    NaN   8.50
3  61   1   4  10.58   6.63  11.75   4.58   4.54  2.88   8.63   1.79   5.83
4  61   1   5  13.33  13.25  11.42   6.17  10.71  8.21  11.92   6.54  10.92

     CLO    BEL    MAL
0  12.58  18.50  15.04
1   9.67  17.54  13.83
2   7.67  12.75  12.71
3   5.88   5.46  10.88
4  10.34  12.92  11.83
```

Step 3

```
     RPT    VAL    ROS    KIL    SHA   BIR    DUB    CLA    MUL    CLO    BEL  \
0  15.04  14.96  13.17   9.29    NaN  9.87  13.67  10.25  10.83  12.58  18.50
1  14.71    NaN  10.83   6.50  12.62  7.67  11.50  10.04   9.79   9.67  17.54
2  18.50  16.88  12.33  10.13  11.17  6.17  11.25    NaN   8.50   7.67  12.75
3  10.58   6.63  11.75   4.58   4.54  2.88   8.63   1.79   5.83   5.88   5.46
4  13.33  13.25  11.42   6.17  10.71  8.21  11.92   6.54  10.92  10.34  12.92

     MAL       Date
0  15.04 2061-01-01
1  13.83 2061-01-02
2  12.71 2061-01-03
3  10.88 2061-01-04
4  11.83 2061-01-05
```

Step 4

```
Year 2061 fix:
      RPT    VAL    ROS    KIL    SHA   BIR    DUB    CLA    MUL    CLO    BEL  \
0  15.04  14.96  13.17   9.29    NaN  9.87  13.67  10.25  10.83  12.58  18.50
1  14.71    NaN  10.83   6.50  12.62  7.67  11.50  10.04   9.79   9.67  17.54
2  18.50  16.88  12.33  10.13  11.17  6.17  11.25    NaN   8.50   7.67  12.75
3  10.58   6.63  11.75   4.58   4.54  2.88   8.63   1.79   5.83   5.88   5.46
4  13.33  13.25  11.42   6.17  10.71  8.21  11.92   6.54  10.92  10.34  12.92

     MAL       Date
0  15.04 1961-01-01
1  13.83 1961-01-02
2  12.71 1961-01-03
3  10.88 1961-01-04
4  11.83 1961-01-05
```

Step 5

```
After setting right dates as index:
              RPT    VAL    ROS    KIL    SHA   BIR    DUB    CLA    MUL  \
Date
```

```
1961-01-01  15.04  14.96  13.17   9.29    NaN  9.87  13.67  10.25  10.83
1961-01-02  14.71    NaN  10.83   6.50  12.62  7.67  11.50  10.04   9.79
1961-01-03  18.50  16.88  12.33  10.13  11.17  6.17  11.25    NaN   8.50
1961-01-04  10.58   6.63  11.75   4.58   4.54  2.88   8.63   1.79   5.83
1961-01-05  13.33  13.25  11.42   6.17  10.71  8.21  11.92   6.54  10.92

              CLO    BEL    MAL
Date
1961-01-01  12.58  18.50  15.04
1961-01-02   9.67  17.54  13.83
1961-01-03   7.67  12.75  12.71
1961-01-04   5.88   5.46  10.88
1961-01-05  10.34  12.92  11.83
```

Step 6

```
Missing values per location:
 RPT    6
VAL    3
ROS    2
KIL    5
SHA    2
BIR    0
DUB    3
CLA    2
MUL    3
CLO    1
BEL    0
MAL    4
dtype: int64
```

Step 7

```
Non-missing values in total:
 78857
```

Step 8

```
Mean windspeeds of the windspeeds over all the locations and all the times:
 10.22788376428218
```

Step 9

```
Min, max and mean windspeeds and standard deviations of the windspeeds at each location:
       min       std       mean    max
RPT  0.67  5.618413  12.362987  35.80
VAL  0.21  5.267356  10.644314  33.37
ROS  1.50  5.008450  11.660526  33.84
KIL  0.00  3.605811   6.306468  28.46
SHA  0.13  4.936125  10.455834  37.54
BIR  0.00  3.968683   7.092254  26.16
DUB  0.00  4.977555   9.797343  30.37
CLA  0.00  4.499449   8.495053  31.08
MUL  0.00  4.166872   8.493590  25.88
CLO  0.04  4.503954   8.707332  28.21
BEL  0.13  5.835037  13.121007  42.38
MAL  0.67  6.699794  15.599079  42.54
```

Step 10

```
day_stats: min, max and mean windspeed and standard deviations of the windspeeds across
all the locations at each day:
           min        max       mean       std
Day
```

```
1    2.402500  29.994167  10.532620  4.418214
2    3.222500  26.236667  10.390365  4.514749
3    3.165000  21.517500   9.920201  3.939575
4    2.598333  21.715000   9.806661  4.015167
5    2.282500  23.771667  10.114915  3.917657
6    2.062500  23.283333  10.015224  4.453977
7    2.840833  25.624167  10.101673  4.395235
8    2.802500  20.981667   9.902049  4.065742
9    2.890000  24.764167  10.399803  4.197974
10   3.086667  23.653333  10.529325  4.491358
11   3.212500  24.160000  10.560683  4.481103
12   2.325833  25.039167  10.495718  4.871524
13   3.037500  23.613333   9.889074  4.289324
14   3.226667  25.355833  10.372138  4.528647
15   2.498333  26.139167  10.073495  4.417735
16   3.035000  24.403333  10.288977  4.590038
17   3.295833  28.841667  10.374973  4.746188
18   2.926667  21.880000  10.152459  4.396713
19   2.567500  24.673333  10.220950  4.376312
20   3.017500  23.721667  10.068786  4.209427
21   2.287500  23.703333   9.774354  4.289666
22   2.750833  24.384167  10.006636  4.350735
23   2.514167  23.600000  10.505355  3.999849
24   2.234167  23.663333  10.396698  4.136423
25   2.590000  25.467500  10.435988  4.261661
26   2.653333  23.975000   9.927810  4.319961
27   2.496667  26.843333  10.345127  4.636004
28   2.485000  23.921667  10.209516  4.166212
29   3.373333  22.349167  10.341877  4.254877
30   3.650000  22.736667  10.629337  4.044755
31   2.773333  22.740000  10.410893  4.227855
```

Step 11

```
Average windspeed for January month for each location:
 RPT    14.847325
VAL    12.914560
ROS    13.299624
KIL     7.199498
SHA    11.667734
BIR     8.054839
DUB    11.819355
CLA     9.512047
MUL     9.543208
CLO    10.053566
BEL    14.550520
MAL    18.028763
dtype: float64
```

Step 12

```
Downsample to a yearly frequency for each location:
             RPT        VAL        ROS       KIL        SHA       BIR  \
Year
1961  12.299583  10.351796  11.362369  6.958227  10.881763  7.729726
1962  12.246923  10.110438  11.732712  6.960440  10.657918  7.393068
1963  12.813452  10.836986  12.541151  7.330055  11.724110  8.434712
1964  12.363661  10.920164  12.104372  6.787787  11.454481  7.570874
1965  12.451370  11.075534  11.848767  6.858466  11.024795  7.478110
1966  13.461973  11.557205  12.020630  7.345726  11.805041  7.793671
1967  12.737151  10.990986  11.739397  7.143425  11.630740  7.368164
1968  11.835628  10.468197  11.409754  6.477678  10.760765  6.067322
1969  11.166356   9.723699  10.902000  5.767973   9.873918  6.189973
1970  12.600329  10.726932  11.730247  6.217178  10.567370  7.609452
1971  11.273123   9.095178  11.088329  5.241507   9.440329  6.097151
```

```
1972  12.463962  10.561311  12.058333  5.929699   9.430410  6.358825
1973  11.828466  10.680493  10.680493  5.547863   9.640877  6.548740
1974  13.643096  11.811781  12.336356  6.427041  11.110986  6.809781
1975  12.008575  10.293836  11.564712  5.269096   9.190082  5.668521
1976  11.737842  10.203115  10.761230  5.109426   8.846339  6.311038
1977  13.099616  11.144493  12.627836  6.073945  10.003836  8.586438
1978  12.504356  11.044274  11.380000  6.082356  10.167233  7.650658

            DUB        CLA       MUL        CLO        BEL        MAL
Year
1961   9.733923   8.858788  8.647652   9.835577  13.502795  13.680773
1962  11.020712   8.793753  8.316822   9.676247  12.930685  14.323956
1963  11.075699  10.336548  8.903589  10.224438  13.638877  14.999014
1964  10.259153   9.467350  7.789016  10.207951  13.740546  14.910301
1965  10.618712   8.879918  7.907425   9.918082  12.964247  15.591644
1966  10.579808   8.835096  8.514438   9.768959  14.265836  16.307260
1967  10.652027   9.325616  8.645014   9.547425  14.774548  17.135945
1968   8.859180   8.255519  7.224945   7.832978  12.808634  15.017486
1969   8.564493   7.711397  7.924521   7.754384  12.621233  15.762904
1970   9.609890   8.334630  9.297616   8.289808  13.183644  16.456027
1971   8.385890   6.757315  7.915370   7.229753  12.208932  15.025233
1972   9.704508   7.680792  8.357295   7.515273  12.727377  15.028716
1973   8.482110   7.614274  8.245534   7.812411  12.169699  15.441096
1974  10.084603   9.896986  9.331753   8.736356  13.252959  16.947671
1975   8.562603   7.843836  8.797945   7.382822  12.631671  15.307863
1976   9.149126   7.146202  8.883716   7.883087  12.332377  15.471448
1977  11.523205   8.378384  9.098192   8.821616  13.459068  16.590849
1978   9.489342   8.800466  9.089753   8.301699  12.967397  16.771370
```

Step 13

```
Downsample to a monthly frequency for each location:
             RPT        VAL        ROS       KIL        SHA        BIR  \
Month
1      14.841333  11.988333  13.431613  7.736774  11.072759   8.588065
2      16.269286  14.975357  14.441481  9.230741  13.852143  10.937500
3      10.890000  11.296452  10.752903  7.284000  10.509355   8.866774
4      10.722667   9.427667   9.998000  5.830667   8.435000   6.495000
5       9.860968   8.850000  10.818065  5.905333   9.490323   6.574839
...          ...        ...        ...       ...        ...        ...
8       9.645161   8.259355   9.032258  4.502903   7.368065   5.935161
9      10.913667  10.895000  10.635000  5.725000  10.372000   9.278333
10      9.897742   8.670968   9.295806  4.721290   8.525161   6.774194
11     16.151667  14.802667  13.508000  7.317333  11.475000   8.743000
12     16.175484  13.748065  15.635161  7.094839  11.398710   9.241613

             DUB        CLA        MUL        CLO        BEL        MAL
Month
1      11.184839   9.245333   9.085806  10.107419  13.880968  14.703226
2      11.890714  11.846071  11.821429  12.714286  18.583214  15.411786
3       9.644194   9.829677  10.294138  11.251935  16.410968  15.720000
4       6.925333   7.094667   7.342333   7.237000  11.147333  10.278333
5       7.604000   8.177097   8.039355   8.499355  11.900323  12.011613
...          ...        ...        ...        ...        ...        ...
8       5.650323   5.417742   7.241290   5.536774  10.466774  12.054194
9      10.790333   9.583000  10.069333   8.939000  15.680333  19.391333
10      8.115484   7.337742   8.297742   8.243871  13.776774  17.150000
11     11.492333   9.657333  10.701333  10.676000  17.404667  20.723000
12     12.077419  10.194839  10.616774  11.028710  13.859677  21.371613

[216 rows x 12 columns]
```

Step 14

```
Downsample to a weekly frequency for each location:
```

| Week | RPT | VAL | ROS | KIL | SHA | BIR |
|---|---|---|---|---|---|---|
| 1 | 15.040000 | 14.960000 | 13.170000 | 9.290000 | NaN | 9.870000 |
| 1 | 13.541429 | 11.486667 | 10.487143 | 6.417143 | 9.474286 | 6.435714 |
| 1 | 12.468571 | 8.967143 | 11.958571 | 4.630000 | 7.351429 | 5.072857 |
| 1 | 13.204286 | 9.862857 | 12.982857 | 6.328571 | 8.966667 | 7.417143 |
| 1 | 19.880000 | 16.141429 | 18.225714 | 12.720000 | 17.432857 | 14.828571 |
| ... | ... | ... | ... | ... | ... | ... |
| 12 | 14.934286 | 11.232857 | 13.941429 | 5.565714 | 10.215714 | 8.618571 |
| 12 | 20.740000 | 19.190000 | 17.034286 | 9.777143 | 15.287143 | 12.774286 |
| 12 | 16.758571 | 14.692857 | 14.987143 | 6.917143 | 11.397143 | 7.272857 |
| 12 | 11.155714 | 8.008571 | 13.172857 | 4.004286 | 7.825714 | 6.290000 |
| 12 | 14.951429 | 11.801429 | 16.035714 | 6.507143 | 9.660000 | 8.620000 |

| Week | DUB | CLA | MUL | CLO | BEL | MAL |
|---|---|---|---|---|---|---|
| 1 | 13.670000 | 10.250000 | 10.830000 | 12.580000 | 18.500000 | 15.040000 |
| 1 | 11.061429 | 6.616667 | 8.434286 | 8.497143 | 12.481429 | 13.238571 |
| 1 | 7.535714 | 6.820000 | 5.712857 | 7.571429 | 11.125714 | 11.024286 |
| 1 | 9.257143 | 7.875714 | 7.145714 | 8.124286 | 9.821429 | 11.434286 |
| 1 | 15.528571 | 15.160000 | 14.480000 | 15.640000 | 20.930000 | 22.530000 |
| ... | ... | ... | ... | ... | ... | ... |
| 12 | 9.642857 | 7.685714 | 9.011429 | 9.547143 | 11.835714 | 18.728571 |
| 12 | 14.437143 | 12.488571 | 13.870000 | 14.082857 | 18.517143 | 23.061429 |
| 12 | 10.208571 | 7.967143 | 9.168571 | 8.565714 | 11.102857 | 15.562857 |
| 12 | 7.798571 | 8.667143 | 7.151429 | 8.072857 | 11.845714 | 18.977143 |
| 12 | 13.708571 | 10.477143 | 10.868571 | 11.471429 | 12.947143 | 26.844286 |

[940 rows x 12 columns]

Step 15

| Week | min | max | mean | std |
|---|---|---|---|---|
| 01 | 7.895000 | 18.020000 | 12.311667 | 2.909304 |
| 02 | 6.167143 | 13.458571 | 9.640496 | 2.575846 |
| 03 | 4.624286 | 13.017143 | 8.391310 | 2.842770 |
| 04 | 7.150000 | 13.392857 | 9.925556 | 2.183014 |
| 05 | 12.357143 | 22.340000 | 16.729028 | 2.988786 |
| 06 | 9.311429 | 18.582857 | 12.918214 | 3.030879 |
| 07 | 10.770000 | 22.152857 | 15.569405 | 3.305013 |
| 08 | 8.281429 | 19.234286 | 12.590655 | 2.902097 |
| 09 | 8.315714 | 17.005714 | 12.920714 | 2.492463 |
| 10 | 7.300000 | 16.444286 | 10.751290 | 2.313917 |
| 11 | 7.494286 | 18.057143 | 11.617877 | 2.945308 |
| 12 | 7.815714 | 18.577143 | 11.831429 | 3.109227 |
| 13 | 5.505714 | 17.312857 | 9.403571 | 3.148063 |
| 14 | 7.528333 | 14.024286 | 10.773651 | 1.728760 |
| 15 | 5.798571 | 12.547143 | 8.938095 | 2.207989 |
| 16 | 4.458571 | 9.828571 | 6.573690 | 1.837791 |
| 17 | 8.227143 | 13.595714 | 10.643214 | 1.838235 |
| 18 | 5.515714 | 11.021429 | 8.193690 | 1.998838 |
| 19 | 10.090000 | 17.552857 | 13.620238 | 2.145580 |
| 20 | 4.314286 | 8.702857 | 6.670972 | 1.374801 |
| 21 | 4.061429 | 11.411429 | 7.300000 | 2.351056 |
| 22 | 4.780000 | 12.702857 | 8.064008 | 2.344159 |
| 23 | 5.708571 | 12.960000 | 8.978690 | 2.271478 |
| 24 | 5.115714 | 12.114286 | 8.172500 | 2.081590 |
| 25 | 6.698571 | 16.155714 | 10.533690 | 2.820740 |
| 26 | 5.572857 | 16.660000 | 10.036071 | 3.213309 |
| 27 | 7.201429 | 15.904286 | 10.340813 | 2.563225 |
| 28 | 6.832857 | 13.797143 | 9.196647 | 2.078194 |
| 29 | 8.198571 | 16.054286 | 11.594643 | 2.312691 |
| 30 | 2.852857 | 7.987143 | 5.314048 | 1.695954 |
| 31 | 5.918571 | 14.267143 | 9.520357 | 2.282987 |

| | | | | |
|---|---|---|---|---|
| 32 | 6.077143 | 13.218571 | 9.377857 | 2.037304 |
| 33 | 5.900000 | 12.315000 | 9.351726 | 1.956986 |
| 34 | 9.011429 | 17.435714 | 13.237262 | 2.512534 |
| 35 | 7.215714 | 16.328571 | 11.718929 | 2.644762 |
| 36 | 4.321429 | 11.472857 | 6.804048 | 2.053783 |
| 37 | 6.387143 | 12.648571 | 9.065476 | 2.038477 |
| 38 | 10.588571 | 18.920000 | 14.296667 | 2.942733 |
| 39 | 4.985714 | 12.870000 | 7.977381 | 2.279029 |
| 40 | 8.762857 | 16.580000 | 11.704921 | 2.617872 |
| 41 | 5.204286 | 11.682857 | 8.090833 | 2.294498 |
| 42 | 6.554286 | 16.944286 | 10.174821 | 3.229020 |
| 43 | 11.114286 | 23.695714 | 15.876905 | 3.559088 |
| 44 | 7.778571 | 17.778571 | 11.994266 | 3.020417 |
| 45 | 7.064286 | 18.178571 | 10.990595 | 3.129724 |
| 46 | 4.744286 | 14.255714 | 8.477381 | 2.962276 |
| 47 | 3.948571 | 14.381429 | 6.750238 | 3.059382 |
| 48 | 5.000000 | 12.328571 | 8.290020 | 2.348823 |
| 49 | 7.310000 | 16.630000 | 11.318333 | 2.960049 |
| 50 | 9.088571 | 17.150000 | 13.507619 | 2.783435 |
| 51 | 7.007143 | 13.675714 | 10.988452 | 2.135537 |
| 52 | 6.457143 | 16.815714 | 10.121190 | 3.215027 |

# Question 5

Step 1

Step 2 and 3

```
Chipo:
    order_id  quantity                              item_name  \
0         1         1           Chips and Fresh Tomato Salsa
1         1         1                                   Izze
2         1         1                         Nantucket Nectar
3         1         1  Chips and Tomatillo-Green Chili Salsa
4         2         2                           Chicken Bowl

                                choice_description item_price
0                                              NaN     $2.39
1                                    [Clementine]     $3.39
2                                         [Apple]     $3.39
3                                              NaN     $2.39
4  [Tomatillo-Red Chili Salsa (Hot), [Black Beans...    $16.98
```

Out[119]:

| | order_id | quantity | item_name | choice_description | item_price |
|---|---|---|---|---|---|
| **0** | 1 | 1 | Chips and Fresh Tomato Salsa | NaN | $2.39 |
| **1** | 1 | 1 | Izze | [Clementine] | $3.39 |
| **2** | 1 | 1 | Nantucket Nectar | [Apple] | $3.39 |
| **3** | 1 | 1 | Chips and Tomatillo-Green Chili Salsa | NaN | $2.39 |
| **4** | 2 | 2 | Chicken Bowl | [Tomatillo-Red Chili Salsa (Hot), [Black Beans... | $16.98 |

Step 4

```
First 10 chipo entries:
    order_id  quantity                              item_name  \
0         1         1           Chips and Fresh Tomato Salsa
1         1         1                                   Izze
2         1         1                         Nantucket Nectar
3         1         1  Chips and Tomatillo-Green Chili Salsa
4         2         2                           Chicken Bowl
5         3         1                           Chicken Bowl
6         3         1                         Side of Chips
7         4         1                         Steak Burrito
8         4         1                       Steak Soft Tacos
9         5         1                         Steak Burrito

                                choice_description item_price
0                                              NaN     $2.39
1                                    [Clementine]     $3.39
2                                         [Apple]     $3.39
3                                              NaN     $2.39
4  [Tomatillo-Red Chili Salsa (Hot), [Black Beans...    $16.98
5  [Fresh Tomato Salsa (Mild), [Rice, Cheese, Sou...    $10.98
6                                              NaN     $1.69
7  [Tomatillo Red Chili Salsa, [Fajita Vegetables...    $11.75
8  [Tomatillo Green Chili Salsa, [Pinto Beans, Ch...     $9.25
9  [Fresh Tomato Salsa, [Rice, Black Beans, Pinto...     $9.25
```

Step 5

```
Number of observations in Chipo:  4622
```

Step 6

```
Number of columns in Chipo:  5
```

Step 7

```
All the column names:
 ['order_id' 'quantity' 'item_name' 'choice_description' 'item_price']
```

Step 8

```
Dataset indexing type: RangeIndex(start=0, stop=4622, step=1)
```

Step 9

```
Most ordered item name: Chicken Bowl
```

Step 10

```
Most ordered item count:  761
```

Step 11

```
Most ordered item from choice description column:  Rice
```

Step 12

```
Total items that were ordered:  2497
```

Step 13

```
Item price type:  object
```

| | order_id | quantity | item_name | choice_description | item_price |
|---|---|---|---|---|---|
| 0 | 1 | 1 | Chips and Fresh Tomato Salsa | [] | 2.39 |
| 1 | 1 | 1 | Izze | [Clementine] | 3.39 |
| 2 | 1 | 1 | Nantucket Nectar | [Apple] | 3.39 |
| 3 | 1 | 1 | Chips and Tomatillo-Green Chili Salsa | [] | 2.39 |
| 4 | 2 | 2 | Chicken Bowl | [Tomatillo-Red Chili Salsa (Hot), Black Beans,... | 16.98 |

```
Item price type:  float64
```

Step 14

```
Revenue for the period in the dataset:  39237.02 $
```

Step 15

```
Total number of orders:  1834
```

Step 16

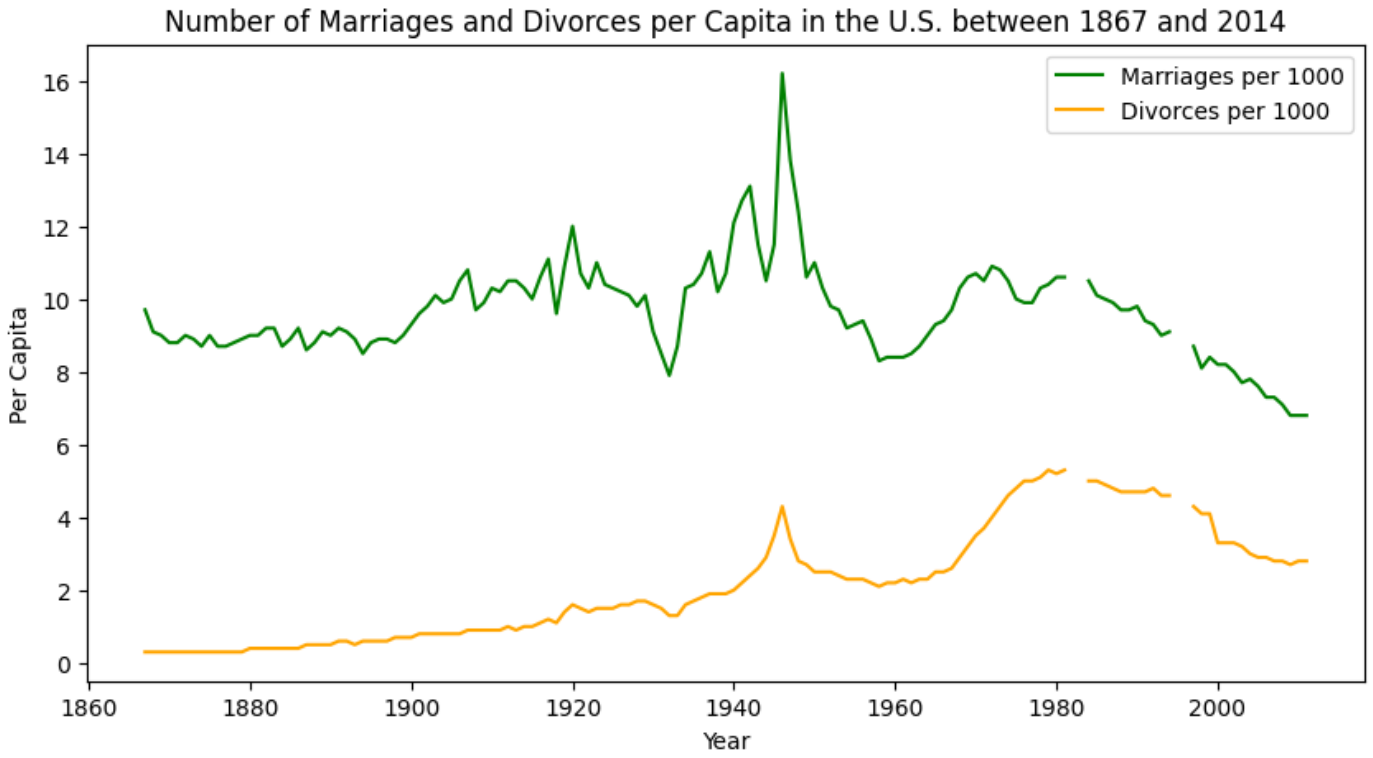
```
Average revenue amount per order:  21.39423118865867 $
```

Step 17

```
Total number of different items sold:  50
```

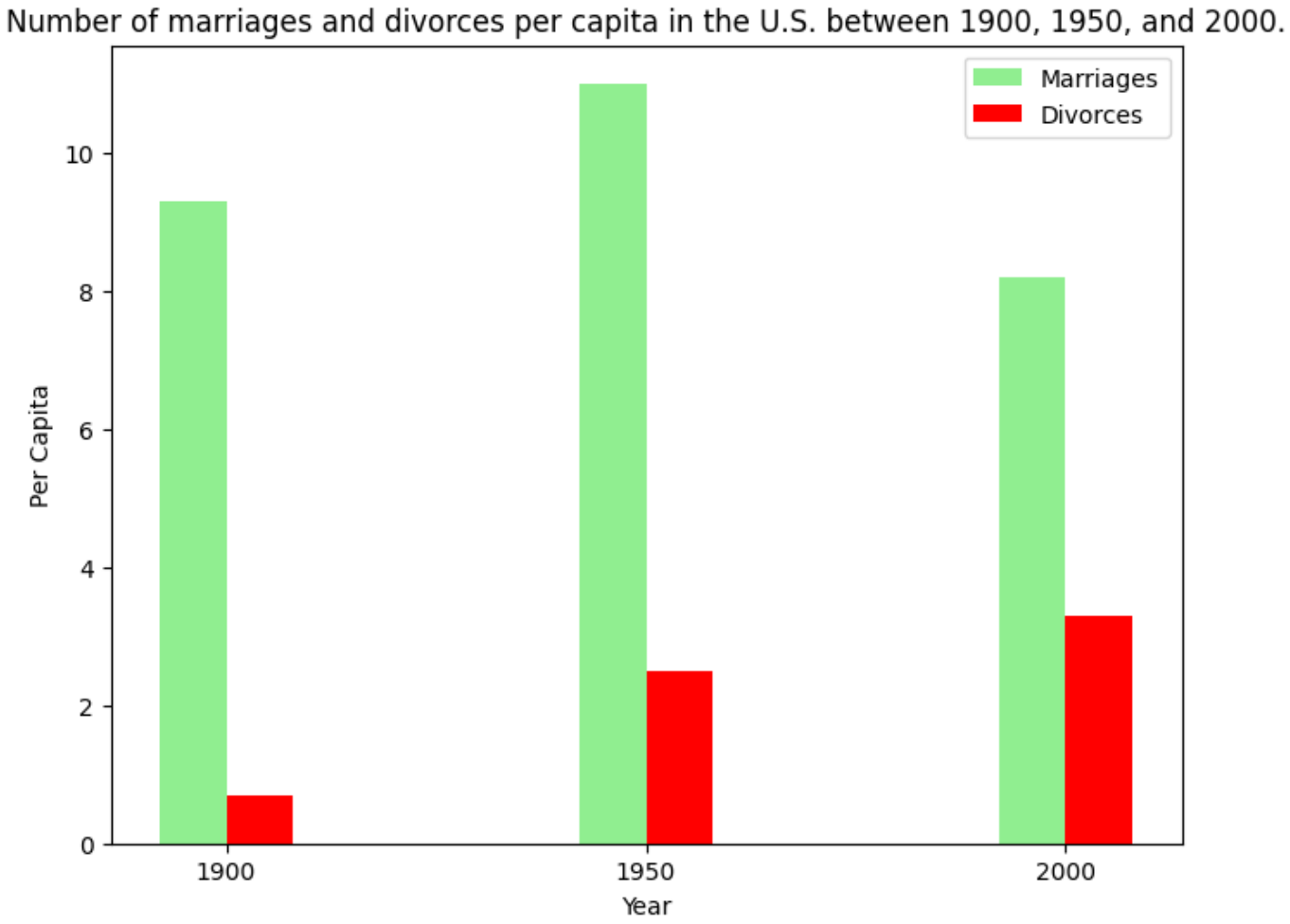
# Question 6

```
   Year  Marriages  Divorces  Population  Marriages_per_1000  \
0  1867   357000.0   10000.0    36970000                 9.7
1  1868   345000.0   10000.0    37885000                 9.1
2  1869   348000.0   11000.0    38870000                 9.0
3  1870   352000.0   11000.0    39905000                 8.8
4  1871   359000.0   12000.0    41010000                 8.8

   Divorces_per_1000
0                0.3
1                0.3
2                0.3
3                0.3
4                0.3
```



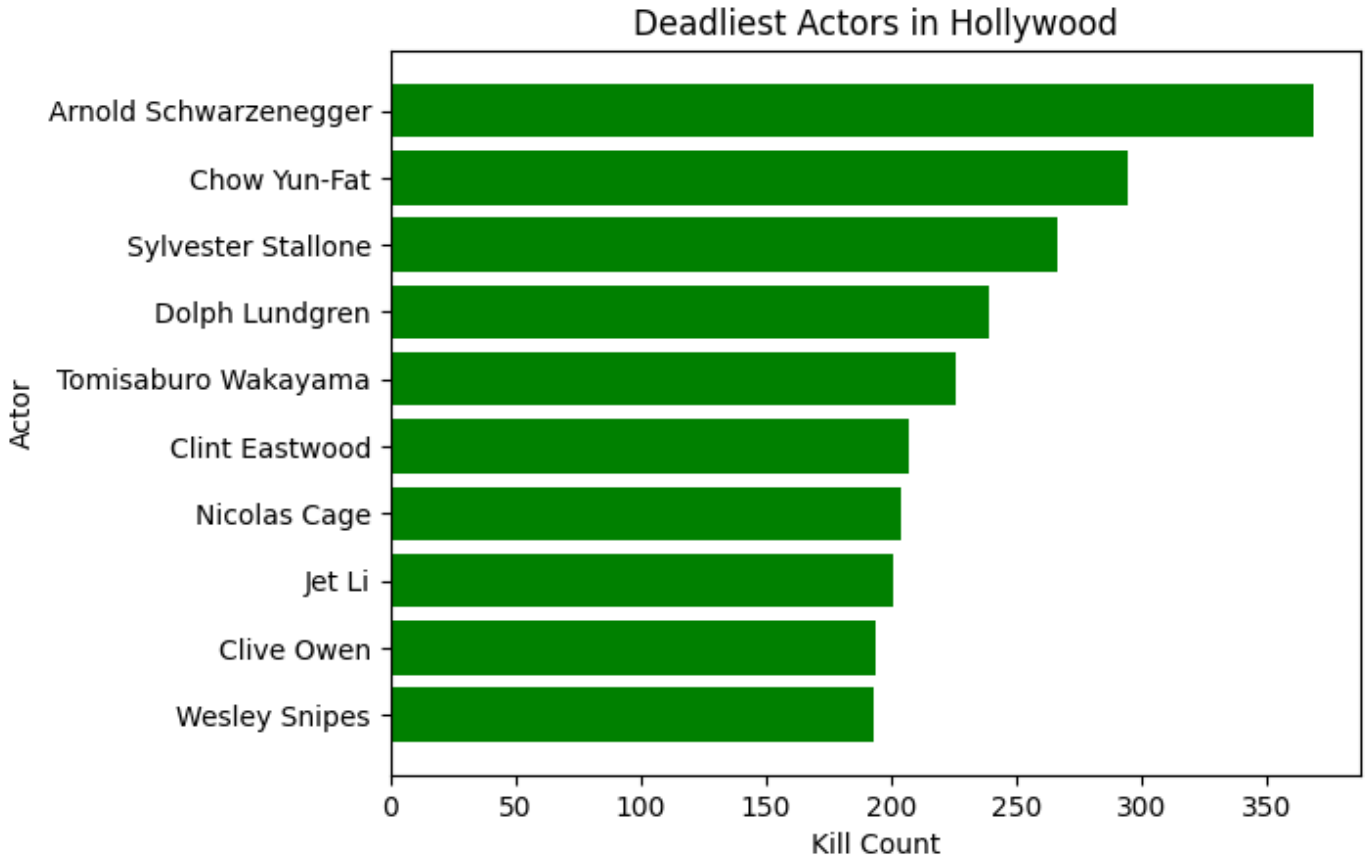
# Question 7

```
   Year  Marriages  Divorces  Population  Marriages_per_1000  \
0  1867   357000.0   10000.0    36970000                 9.7
1  1868   345000.0   10000.0    37885000                 9.1
2  1869   348000.0   11000.0    38870000                 9.0
3  1870   352000.0   11000.0    39905000                 8.8
4  1871   359000.0   12000.0    41010000                 8.8

   Divorces_per_1000
0                0.3
1                0.3
2                0.3
3                0.3
4                0.3
```



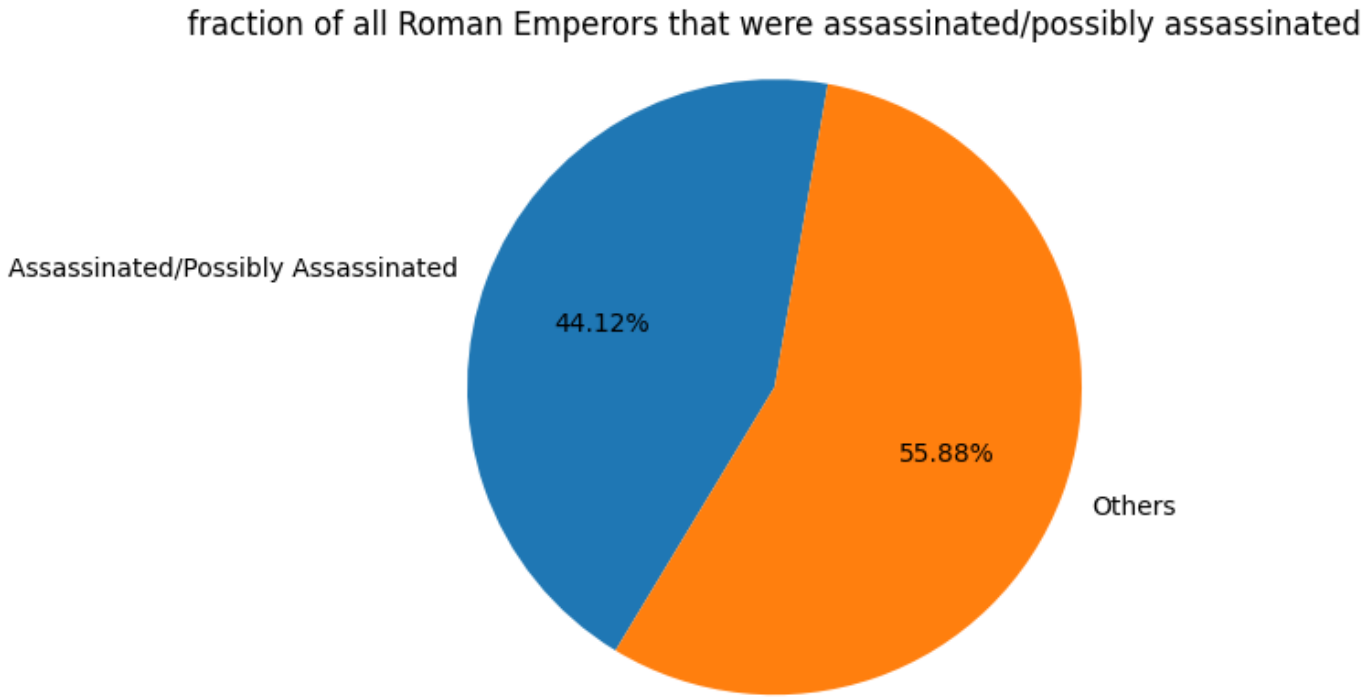
# Question 8

| | Actor | Count |
|---|---|---|
| 0 | Arnold Schwarzenegger | 369 |
| 1 | Chow Yun-Fat | 295 |
| 2 | Clint Eastwood | 207 |
| 3 | Clive Owen | 194 |
| 4 | Dolph Lundgren | 239 |

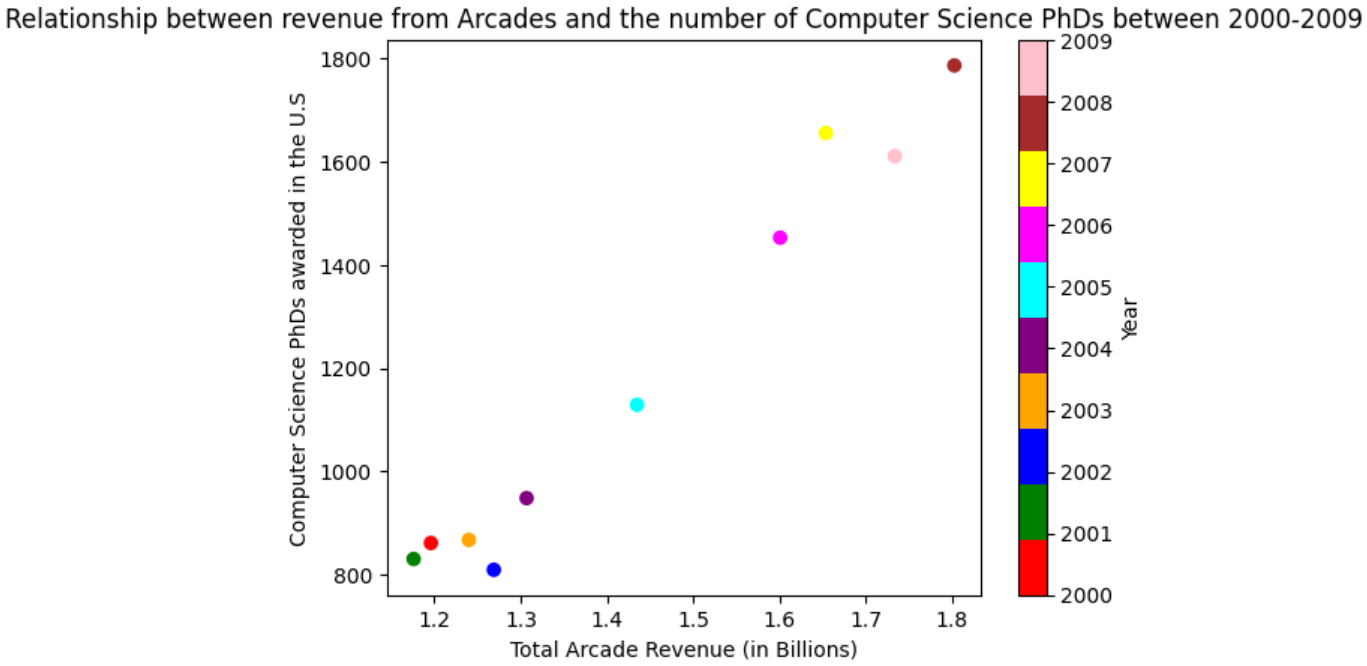


# Question 9

| | Emperor | Length_of_Reign | Cause_of_Death |
|---|---|---|---|
| 0 | Augustus | 40.58 | Possibly assassinated |
| 1 | Tiberius | 22.50 | Possibly assassinated |
| 2 | Caligula | 4.83 | Assassinated |
| 3 | Claudius | 13.75 | Possibly assassinated |
| 4 | Nero | 13.67 | Suicide |



# Question 10


Relationship between revenue from Arcades and the number of Computer Science PhDs between 2000-2009
Computer Science PhDs awarded in the U.S
1800
1600
1400
1200
1000
800
1.2
1.3
1.4
1.5
1.6
1.7
1.8
Total Arcade Revenue (in Billions)
Year
2009
2008
2007
2006
2005
2004
2003
2002
2001
2000