In [ ]:
```
#TASK 1
#CONVERT DATA TYPE OF DATE COLUMN TO DATA TIME FORMAT

#TASK 2
#ADD A NEW COLUMN "YEAR" TO THE DATA FRAME WHICH CONTAINS YEAR ONLY

#TASK 2B
#ADD NEW COLUMN "MONTH" TO THE SECOND COLUMN OF THE DATA FRAME WHICH CONTAIN MONTH ONLY

#TASK 3
#REMOVE THE COLUMN CODE AND HOUSES_SOLD FROM THE DATA FRAME

#TASK 4
#SHOW ALL THE RECORD WHERE NO OF CRIME IS ZERO, HOW MANY OF SUCH RECORD ARE THERE

#TASK 5
#WHAT IS THE MINIMUM AND MAXIMUM "AVERAGE_PRICE" IN ENGLAND PER YEAR

#TASK 6
#WHAT IS THE MAXIMUM AND MINIMUM NUMBER OF CRIMES RECORDED PER AREA

#TASK 7
#SHOW THE TOTAL COUNTS OF RECORD FOR EACH AREA WHERE THE AVERAGE PRICE IS LESS THAN 100000
```

In [2]:
```
import pandas as pd

data= pd.read_csv("C:\\Users\\Joseph\\Desktop\\housing_in_london_monthly_variables.csv")
data.head(5)
```

Out[2]:

	date	area	average_price	code	houses_sold	no_of_crimes	borough_flag
0	1995-01-01	city of london	91449	E09000001	17.0	NaN	1
1	1995-02-01	city of london	82203	E09000001	7.0	NaN	1
2	1995-03-01	city of london	79121	E09000001	14.0	NaN	1
3	1995-04-01	city of london	77101	E09000001	7.0	NaN	1
4	1995-05-01	city of london	84409	E09000001	10.0	NaN	1

In [3]:
```
#TASK 1
#CONVERT DATA TYPE OF DATE COLUMN TO DATA TIME FORMAT
data["date"]= pd.to_datetime(data["date"])
data.head(2)
```

Out[3]:

	date	area	average_price	code	houses_sold	no_of_crimes	borough_flag
0	1995-01-01	city of london	91449	E09000001	17.0	NaN	1
1	1995-02-01	city of london	82203	E09000001	7.0	NaN	1

In [4]:
```
data.dtypes
```

Out[4]:
```
date             datetime64[ns]
area                     object
average_price             int64
code                     object
houses_sold             float64
no_of_crimes            float64
borough_flag              int64
dtype: object
```

In [5]:
```
#TASK 2
#ADD A NEW COLUMN "YEAR" TO THE DATA FRAME WHICH CONTAINS YEAR ONLY

data["year"]= data["date"].dt.year

data.head(2)
```

Out[5]:

	date	area	average_price	code	houses_sold	no_of_crimes	borough_flag	year
0	1995-01-01	city of london	91449	E09000001	17.0	NaN	1	1995
1	1995-02-01	city of london	82203	E09000001	7.0	NaN	1	1995

In [6]:
```
#TASK 2B
#ADD NEW COLUMN "MONTH" TO THE SECOND COLUMN OF THE DATA FRAME WHICH CONTAIN MONTH ONLY

data.insert(1, "month", data["date"].dt.month)
data.head(2)
```

Out[6]:

	date	month	area	average_price	code	houses_sold	no_of_crimes	borough_flag	year
0	1995-01-01	1	city of london	91449	E09000001	17.0	NaN	1	1995
1	1995-02-01	2	city of london	82203	E09000001	7.0	NaN	1	1995

In [7]:
```
#TASK 3
#REMOVE THE COLUMN CODE AND HOUSES_SOLD FROM THE DATA FRAME

data.drop(["code", "houses_sold"], axis= 1, inplace= True)
data.head(2)
```

Out[7]:

	date	month	area	average_price	no_of_crimes	borough_flag	year
0	1995-01-01	1	city of london	91449	NaN	1	1995
1	1995-02-01	2	city of london	82203	NaN	1	1995

In [9]:
```
#TASK 4
#SHOW ALL THE RECORD WHERE NO OF CRIME IS ZERO, HOW MANY OF SUCH RECORD ARE THERE

data["no_of_crimes"].value_counts()
#the record shows 104 times where no of crime is zero
```

Out[9]:
```
0.0       104
2039.0     11
1956.0     11
2276.0     10
2114.0     10
         ...
1947.0      1
3493.0      1
3614.0      1
3241.0      1
6405.0      1
Name: no_of_crimes, Length: 2669, dtype: int64
```

In [10]:
```
data[(data["no_of_crimes"]== 0)]
```

Out[10]:

	date	month	area	average_price	no_of_crimes	borough_flag	year
72	2001-01-01	1	city of london	284262	0.0	1	2001
73	2001-02-01	2	city of london	198137	0.0	1	2001
74	2001-03-01	3	city of london	189033	0.0	1	2001
75	2001-04-01	4	city of london	205494	0.0	1	2001
76	2001-05-01	5	city of london	223459	0.0	1	2001
...	...	...	...	...	...	...	...
178	2009-11-01	11	city of london	397909	0.0	1	2009
179	2009-12-01	12	city of london	411955	0.0	1	2009
180	2010-01-01	1	city of london	464436	0.0	1	2010
181	2010-02-01	2	city of london	490525	0.0	1	2010
182	2010-03-01	3	city of london	498241	0.0	1	2010

104 rows × 7 columns

In [ ]:
```
#the record shows 104 times where no of crime is zero
```

In [11]:
```
#TASK 5
#WHAT IS THE MINIMUM AND MAXIMUM "AVERAGE_PRICE" IN ENGLAND PER YEAR

data.head(2)
```

Out[11]:

	date	month	area	average_price	no_of_crimes	borough_flag	year
0	1995-01-01	1	city of london	91449	NaN	1	1995
1	1995-02-01	2	city of london	82203	NaN	1	1995

In [18]:
```
eng_data= data[(data["area"]== "england")]
```

In [20]:
```
#maximum average price
eng_data.groupby("year").average_price.max()
```

Out[20]:
```
year
1995     53901
1996     55755
1997     61564
1998     65743
1999     75071
2000     84191
2001     95992
2002    119982
2003    138985
2004    160330
2005    167244
2006    182031
2007    194764
2008    191750
2009    174136
2010    180807
2011    177335
2012    180129
2013    188544
2014    203639
2015    219582
2016    231922
2017    242628
2018    248620
2019    250410
2020    247355
Name: average_price, dtype: int64
```

In [21]:
```
#minimum average price
eng_data.groupby("year").average_price.min()
```

Out[21]:
```
year
1995     52788
1996     52333
1997     55789
1998     61659
1999     65522
2000     75219
2001     84245
2002     96215
2003    121610
2004    139719
2005    158572
2006    166544
2007    181824
2008    165795
2009    159340
2010    174458
2011    173046
2012    174161
2013    176816
2014    188265
2015    202856
2016    220361
2017    231593
2018    240428
2019    243281
2020    247355
Name: average_price, dtype: int64
```

In [26]:
```
#TASK 6
#WHAT IS THE MAXIMUM AND MINIMUM NUMBER OF CRIMES RECORDED PER AREA

#maximum number of crimes
data.groupby("area").no_of_crimes.max()
```

Out[26]:
```
area
barking and dagenham      2049.0
barnet                    2893.0
bexley                    1914.0
brent                     2937.0
bromley                   2637.0
camden                    4558.0
city of london              10.0
croydon                   3263.0
ealing                    3401.0
east midlands                NaN
east of england              NaN
enfield                   2798.0
england                      NaN
greenwich                 2853.0
hackney                   3466.0
hammersmith and fulham    2645.0
haringey                  3199.0
harrow                    1763.0
havering                  1956.0
hillingdon                2819.0
hounslow                  2817.0
inner london                 NaN
islington                 3384.0
kensington and chelsea    2778.0
kingston upon thames      1379.0
lambeth                   4701.0
lewisham                  2813.0
london                       NaN
merton                    1623.0
newham                    3668.0
north east                   NaN
north west                   NaN
outer london                 NaN
redbridge                 2560.0
richmond upon thames      1551.0
south east                   NaN
south west                   NaN
southwark                 3821.0
sutton                    1425.0
tower hamlets             3316.0
waltham forest            2941.0
wandsworth                3051.0
west midlands                NaN
westminster               7461.0
yorks and the humber         NaN
Name: no_of_crimes, dtype: float64
```

In [27]:
```
#minimum number of crimes
data.groupby("area").no_of_crimes.min()
```

Out[27]:
```
area
barking and dagenham      1217.0
barnet                    1703.0
bexley                     860.0
brent                     1850.0
bromley                   1441.0
camden                    2079.0
city of london               0.0
croydon                   2031.0
ealing                    1871.0
east midlands                NaN
east of england              NaN
enfield                   1635.0
england                      NaN
greenwich                 1513.0
hackney                   1870.0
hammersmith and fulham    1323.0
haringey                  1536.0
harrow                     937.0
havering                  1130.0
hillingdon                1445.0
hounslow                  1529.0
inner london                 NaN
islington                 1871.0
kensington and chelsea    1347.0
kingston upon thames       692.0
lambeth                   2381.0
lewisham                  1675.0
london                       NaN
merton                     819.0
newham                    2130.0
north east                   NaN
north west                   NaN
outer london                 NaN
redbridge                 1487.0
richmond upon thames       700.0
south east                   NaN
south west                   NaN
southwark                 2267.0
sutton                     787.0
tower hamlets             1646.0
waltham forest            1575.0
wandsworth                1582.0
west midlands                NaN
westminster               3504.0
yorks and the humber         NaN
Name: no_of_crimes, dtype: float64
```

In [28]:
```
#TASK 7
#SHOW THE TOTAL COUNTS OF RECORD FOR EACH AREA WHERE THE AVERAGE PRICE IS LESS THAN 100000

data.head(2)
```

Out[28]:

	date	month	area	average_price	no_of_crimes	borough_flag	year
0	1995-01-01	1	city of london	91449	NaN	1	1995
1	1995-02-01	2	city of london	82203	NaN	1	1995

In [30]:
```
price_data= data[(data["average_price"] < 100000)]
```

In [31]:
```
price_data.head(2)
```

Out[31]:

	date	month	area	average_price	no_of_crimes	borough_flag	year
0	1995-01-01	1	city of london	91449	NaN	1	1995
1	1995-02-01	2	city of london	82203	NaN	1	1995

In [32]:
```
price_data["area"].value_counts()
```

Out[32]:
```
north east              112
north west              111
yorks and the humber    110
east midlands            96
west midlands            94
england                  87
barking and dagenham     85
south west               78
east of england          76
newham                   72
bexley                   64
waltham forest           64
lewisham                 62
havering                 60
south east               59
greenwich                59
croydon                  57
enfield                  54
sutton                   54
hackney                  53
redbridge                52
southwark                48
tower hamlets            47
outer london             46
hillingdon               44
lambeth                  41
hounslow                 41
brent                    40
london                   39
merton                   35
haringey                 33
bromley                  33
inner london             31
ealing                   31
kingston upon thames     30
harrow                   30
wandsworth               26
barnet                   25
islington                19
city of london           11
Name: area, dtype: int64
```

In [ ]: