In [1]:
```
import pandas as pd                          # Importing Libraries
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
import warnings
warnings.filterwarnings('ignore')
```

In [2]:
```
price= pd.read_csv('CSUSHPISA.csv')     # House Prices
price
```

Out[2]:

	DATE	CSUSHPISA
0	1987-01-01	63.965
1	1987-02-01	64.424
2	1987-03-01	64.735
3	1987-04-01	65.131
4	1987-05-01	65.563
...	...	...
433	2023-02-01	296.958
434	2023-03-01	298.210
435	2023-04-01	300.214
436	2023-05-01	302.657
437	2023-06-01	304.635

438 rows × 2 columns

In [3]:
```
unemp_rate=pd.read_csv('U2RATE.csv')     # Unemployement
unemp_rate
```

Out[3]:

	DATE	U2RATE
0	1967-01-01	1.6
1	1967-02-01	1.6
2	1967-03-01	1.5
3	1967-04-01	1.6
4	1967-05-01	1.6
...	...	...
675	2023-04-01	1.6
676	2023-05-01	1.8
677	2023-06-01	1.7
678	2023-07-01	1.6
679	2023-08-01	1.7

680 rows × 2 columns

In [4]:
```
#New Houses for Sale by Stage of Construction, Not Started
```

```
constr_not_startd= pd.read_csv('NHFSEPNTS.csv')  # thousands of units
constr_not_startd
```

Out[4]:

	DATE	NHFSEPNTS
0	1999-01-01	38.0
1	1999-02-01	38.0
2	1999-03-01	36.0
3	1999-04-01	41.0
4	1999-05-01	39.0
...	...	...
290	2023-03-01	90.0
291	2023-04-01	93.0
292	2023-05-01	94.0
293	2023-06-01	97.0
294	2023-07-01	108.0

295 rows × 2 columns

In [5]:
```
#New Houses for Sale by Stage of Construction, Under Construction
undr_constrtn = pd.read_csv('NHFSEPUCS.csv')
undr_constrtn
```

Out[5]:

	DATE	NHFSEPUCS
0	1999-01-01	178.0
1	1999-02-01	180.0
2	1999-03-01	185.0
3	1999-04-01	180.0
4	1999-05-01	184.0
...	...	...
290	2023-03-01	273.0
291	2023-04-01	267.0
292	2023-05-01	266.0
293	2023-06-01	262.0
294	2023-07-01	254.0

295 rows × 2 columns

In [6]:
```
# New Houses for Sale by Stage of Construction, Completed
cnstr_cmplt= pd.read_csv('NHFSEPCS.csv')
cnstr_cmplt
```

Out[6]:

	DATE	NHFSEPCS
0	1999-01-01	68.0
1	1999-02-01	67.0

2	1999-03-01	68.0
3	1999-04-01	69.0
4	1999-05-01	72.0
...	...	...
290	2023-03-01	70.0
291	2023-04-01	70.0
292	2023-05-01	66.0
293	2023-06-01	69.0
294	2023-07-01	75.0

295 rows × 2 columns

In [7]:
```
df_under_compl= pd.merge(cnstr_cmplt,undr_constrtn,on='DATE',how='inner')     #joining
df_under_compl
```

Out[7]:

	DATE	NHFSEPCS	NHFSEPUCS
0	1999-01-01	68.0	178.0
1	1999-02-01	67.0	180.0
2	1999-03-01	68.0	185.0
3	1999-04-01	69.0	180.0
4	1999-05-01	72.0	184.0
...	...	...	...
290	2023-03-01	70.0	273.0
291	2023-04-01	70.0	267.0
292	2023-05-01	66.0	266.0
293	2023-06-01	69.0	262.0
294	2023-07-01	75.0	254.0

295 rows × 3 columns

In [8]:
```
df_not_unemp =pd.merge(constr_not_startd,unemp_rate,on='DATE',how='inner')
df_not_unemp
```

Out[8]:

	DATE	NHFSEPNTS	U2RATE
0	1999-01-01	38.0	2.0
1	1999-02-01	38.0	2.0
2	1999-03-01	36.0	1.9
3	1999-04-01	41.0	1.9
4	1999-05-01	39.0	1.9
...	...	...	...
290	2023-03-01	90.0	1.8
291	2023-04-01	93.0	1.6

292	2023-05-01	94.0	1.8
293	2023-06-01	97.0	1.7
294	2023-07-01	108.0	1.6

295 rows × 3 columns

In [9]:
```
join_df = pd.merge(df_under_compl,df_not_unemp,on='DATE',how='inner')
join_df
```

Out[9]:

	DATE	NHFSEPCS	NHFSEPUCS	NHFSEPNTS	U2RATE
0	1999-01-01	68.0	178.0	38.0	2.0
1	1999-02-01	67.0	180.0	38.0	2.0
2	1999-03-01	68.0	185.0	36.0	1.9
3	1999-04-01	69.0	180.0	41.0	1.9
4	1999-05-01	72.0	184.0	39.0	1.9
...	...	...	...	...	...
290	2023-03-01	70.0	273.0	90.0	1.8
291	2023-04-01	70.0	267.0	93.0	1.6
292	2023-05-01	66.0	266.0	94.0	1.8
293	2023-06-01	69.0	262.0	97.0	1.7
294	2023-07-01	75.0	254.0	108.0	1.6

295 rows × 5 columns

In [10]:
```
join_data_df = pd.merge(join_df,price,on='DATE',how='inner')
```

In [11]:
```
#renaming column name for better understing the data & aligning Data to center(for bette
join_data_df.rename(columns = {'NHFSEPCS':'Const_complt', 'NHFSEPUCS':'un_constr','NHFSE
join_data_df
```

Out[11]:

	DATE	Const_complt	un_constr	Cnstr_not_Strtd	Unemploy_Rate	Price_fact
0	1999-01-01	68.0	178.0	38.0	2.0	93.207
1	1999-02-01	67.0	180.0	38.0	2.0	93.670
2	1999-03-01	68.0	185.0	36.0	1.9	94.216
3	1999-04-01	69.0	180.0	41.0	1.9	94.784
4	1999-05-01	72.0	184.0	39.0	1.9	95.343
...	...	...	...	...	...	...
289	2023-02-01	69.0	278.0	90.0	1.7	296.958
290	2023-03-01	70.0	273.0	90.0	1.8	298.210
291	2023-04-01	70.0	267.0	93.0	1.6	300.214
292	2023-05-01	66.0	266.0	94.0	1.8	302.657
293	2023-06-01	69.0	262.0	97.0	1.7	304.635

294 rows × 6 columns

In [12]:
```
join_data_df.describe()   #calculating some statistical data like percentile, mean and s
```

Out[12]:

	Const_complt	un_constr	Cnstr_not_Strtd	Unemploy_Rate	Price_fact
count	294.000000	294.000000	294.000000	294.000000	294.000000
mean	81.894558	184.258503	52.615646	3.123129	171.417405
std	40.068019	67.904529	22.246952	1.585905	49.479289
min	31.000000	70.000000	22.000000	1.500000	93.207000
25%	56.000000	123.250000	36.250000	2.025000	140.161000
50%	75.000000	185.000000	47.000000	2.600000	164.796500
75%	87.000000	230.000000	66.500000	3.700000	189.398250
max	194.000000	338.000000	102.000000	13.200000	304.817000

In [13]:
```
join_data_df.info()  # prints information about the data
```

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 294 entries, 0 to 293
Data columns (total 6 columns):
 #   Column           Non-Null Count  Dtype  
---  ------           --------------  -----  
 0   DATE             294 non-null    object 
 1   Const_complt     294 non-null    float64
 2   un_constr        294 non-null    float64
 3   Cnstr_not_Strtd  294 non-null    float64
 4   Unemploy_Rate    294 non-null    float64
 5   Price_fact       294 non-null    float64
dtypes: float64(5), object(1)
memory usage: 16.1+ KB
```

In [14]:
```
join_data_df.duplicated().sum()   # in the DataFrame are duplicated and not
```

Out[14]:
```
0
```

In [15]:
```
join_data_df.isnull().sum()    #Cheecking there null value in DataSet
```

Out[15]:
```
DATE               0
Const_complt       0
un_constr          0
Cnstr_not_Strtd    0
Unemploy_Rate      0
Price_fact         0
dtype: int64
```

In [16]:
```
join_data_df.nunique()         #checking the number of unique values for each column.
```

Out[16]:
```
DATE               294
Const_complt       109
un_constr          162
Cnstr_not_Strtd     73
Unemploy_Rate       55
Price_fact         294
dtype: int64
```

Exploratory data analysis (EDA)

In [17]:
```
join_data_df['year'] = pd.DatetimeIndex(join_data_df['DATE']).year      # Creating New co
```

In [18]:
```
join_data_df
```

Out[18]:

	DATE	Const_complt	un_constr	Cnstr_not_Strtd	Unemploy_Rate	Price_fact	year
0	1999-01-01	68.0	178.0	38.0	2.0	93.207	1999
1	1999-02-01	67.0	180.0	38.0	2.0	93.670	1999
2	1999-03-01	68.0	185.0	36.0	1.9	94.216	1999
3	1999-04-01	69.0	180.0	41.0	1.9	94.784	1999
4	1999-05-01	72.0	184.0	39.0	1.9	95.343	1999
...	...	...	...	...	...	...	...
289	2023-02-01	69.0	278.0	90.0	1.7	296.958	2023
290	2023-03-01	70.0	273.0	90.0	1.8	298.210	2023
291	2023-04-01	70.0	267.0	93.0	1.6	300.214	2023
292	2023-05-01	66.0	266.0	94.0	1.8	302.657	2023
293	2023-06-01	69.0	262.0	97.0	1.7	304.635	2023

294 rows × 7 columns

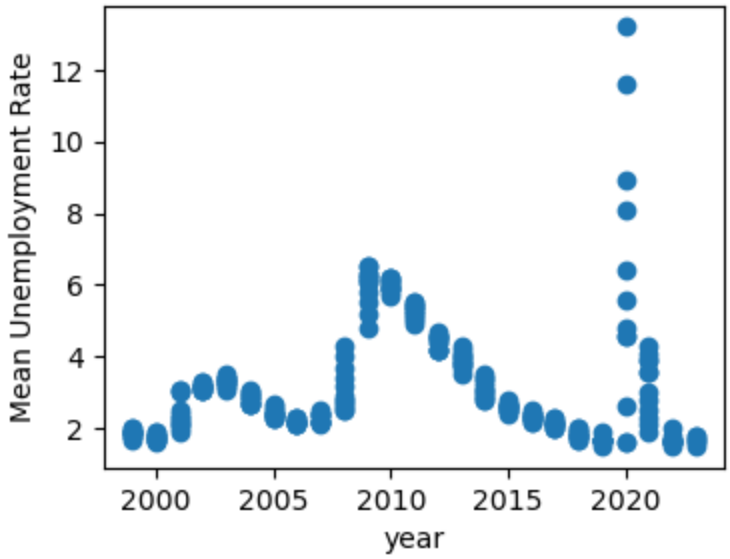
In [19]:
```
join_df_mean = join_data_df.groupby(by='year', as_index=False).mean()   # Mean of the Dat
join_df_mean
```

Out[19]:

	year	Const_complt	un_constr	Cnstr_not_Strtd	Unemploy_Rate	Price_fact
0	1999	71.500000	183.000000	41.333333	1.875000	96.365333
1	2000	84.250000	180.250000	39.500000	1.766667	104.768417
2	2001	77.583333	182.333333	41.833333	2.416667	113.179500
3	2002	82.500000	199.000000	46.000000	3.183333	122.278500
4	2003	80.833333	213.916667	53.666667	3.316667	133.731000
5	2004	89.916667	243.250000	62.166667	2.850000	150.440083
6	2005	106.750000	279.500000	81.416667	2.433333	171.736750
7	2006	144.083333	318.166667	91.083333	2.191667	183.447417
8	2007	185.416667	267.166667	78.500000	2.291667	179.918833
9	2008	179.083333	190.416667	56.000000	3.116667	164.057167
10	2009	128.000000	115.750000	35.083333	5.958333	148.544583
11	2010	86.500000	97.083333	27.750000	5.991667	144.674167
12	2011	65.500000	77.166667	24.666667	5.266667	139.260000
13	2012	44.500000	77.750000	23.583333	4.425000	140.994667
14	2013	40.000000	99.083333	29.500000	3.916667	154.520417
15	2014	50.250000	117.750000	32.250000	3.116667	164.699333
16	2015	51.583333	127.333333	37.166667	2.583333	172.182417
17	2016	58.333333	146.583333	38.083333	2.333333	180.927250

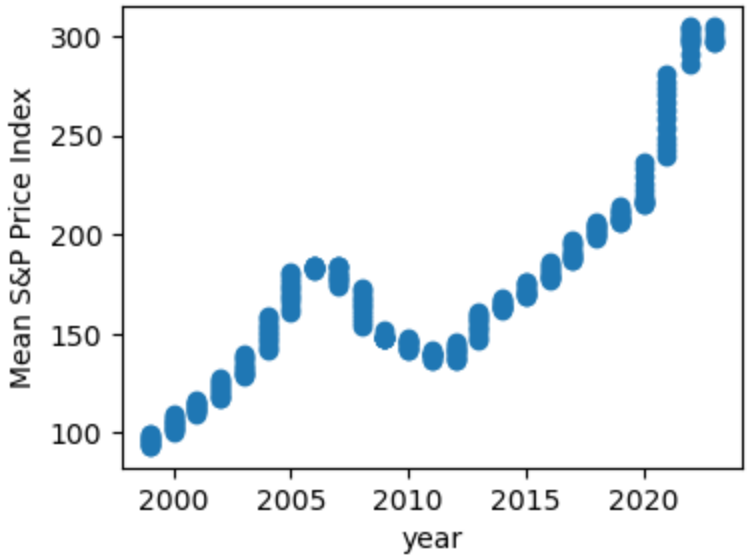
18	2017	62.250000	165.500000	47.416667	2.133333	191.402000
19	2018	66.500000	191.083333	56.166667	1.841667	202.484750
20	2019	77.500000	198.166667	54.666667	1.691667	209.473417
21	2020	60.583333	184.166667	59.083333	6.133333	222.155250
22	2021	34.416667	225.250000	89.500000	3.166667	260.066667
23	2022	44.333333	298.666667	96.833333	1.675000	298.478917
24	2023	68.500000	272.000000	91.666667	1.683333	299.879167

In [20]:
```
plt.figure(figsize=(4,3))
plt.scatter(join_data_df.year,join_data_df.Unemploy_Rate)
plt.xlabel('year')
plt.ylabel('Mean Unemployment Rate')
plt.show()
```



the first peak in unemployment rate came in the year 2021 and started to decrease slowly from 2022 to 2023 became lowest in the year 2023

In [21]:
```
plt.figure(figsize=(4,3))
plt.scatter(join_data_df.year,join_data_df.Price_fact)
plt.xlabel('year')
plt.ylabel('Mean S&P Price Index')
plt.show()
```



The price of house kept on increasing from year 2003 to 2005 and started gradually decreasing from 2007 to 2012 (in the same year the unemployment rate started increasing), after 2020 there was large difference in the price of house.

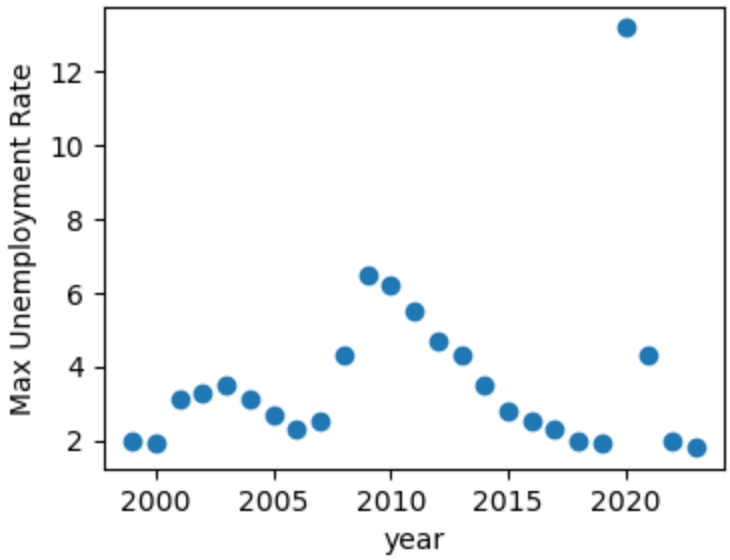
In [22]:
```
df_max=join_data_df.groupby(by='year', as_index=False).max()    # max values of each fact
df_max
```

Out[22]:

	year	DATE	Const_complt	un_constr	Cnstr_not_Strtd	Unemploy_Rate	Price_fact
0	1999	1999-12-01	77.0	191.0	47.0	2.0	99.844
1	2000	2000-12-01	90.0	188.0	42.0	1.9	109.140
2	2001	2001-12-01	82.0	191.0	45.0	3.1	116.456
3	2002	2002-12-01	86.0	204.0	52.0	3.3	127.623
4	2003	2003-12-01	86.0	232.0	61.0	3.5	140.179
5	2004	2004-12-01	97.0	257.0	70.0	3.1	159.330
6	2005	2005-12-01	110.0	306.0	95.0	2.7	180.910
7	2006	2006-12-01	166.0	338.0	100.0	2.3	184.364
8	2007	2007-12-01	194.0	285.0	85.0	2.5	184.598
9	2008	2008-12-01	191.0	227.0	69.0	4.3	173.132
10	2009	2009-12-01	161.0	139.0	41.0	6.5	151.506
11	2010	2010-12-01	96.0	109.0	31.0	6.2	147.395
12	2011	2011-12-01	75.0	84.0	28.0	5.5	141.522
13	2012	2012-12-01	53.0	86.0	25.0	4.7	145.503
14	2013	2013-12-01	41.0	114.0	37.0	4.3	160.993
15	2014	2014-12-01	56.0	123.0	35.0	3.5	168.052
16	2015	2015-12-01	55.0	139.0	39.0	2.8	176.545
17	2016	2016-12-01	61.0	152.0	43.0	2.5	185.726
18	2017	2017-12-01	65.0	178.0	52.0	2.3	197.170
19	2018	2018-12-01	75.0	203.0	69.0	2.0	206.149

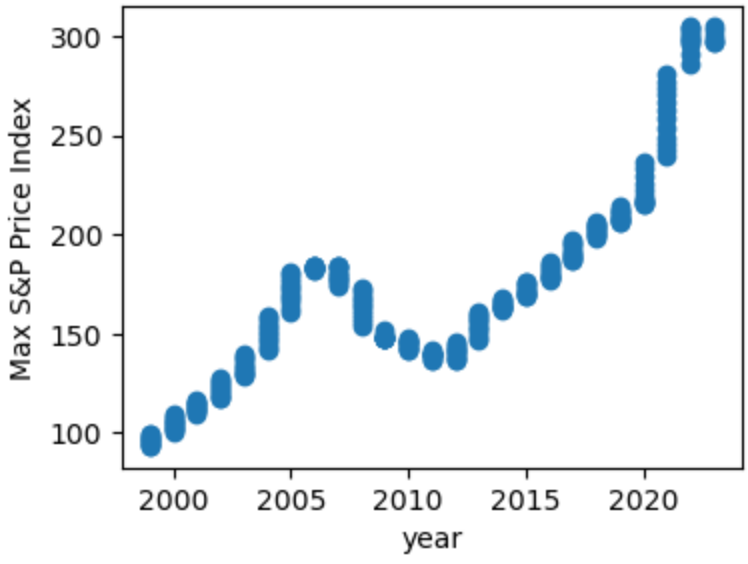
20	2019	2019-12-01	81.0	211.0	59.0	1.9	213.906
21	2020	2020-12-01	78.0	199.0	68.0	13.2	236.433
22	2021	2021-12-01	40.0	264.0	98.0	4.3	281.266
23	2022	2022-12-01	66.0	318.0	102.0	2.0	304.817
24	2023	2023-06-01	70.0	286.0	97.0	1.8	304.635

In [23]:
```
plt.figure(figsize=(4,3))
plt.scatter(df_max.year,df_max.Unemploy_Rate)
plt.xlabel('year')
plt.ylabel('Max Unemployment Rate')
plt.show()
```



Maximum Unemployelent Rate was seen in the year 2021

In [24]:
```
plt.figure(figsize=(4,3))
plt.scatter(join_data_df.year,join_data_df.Price_fact)
plt.xlabel('year')
plt.ylabel('Max S&P Price Index')
plt.show()
```

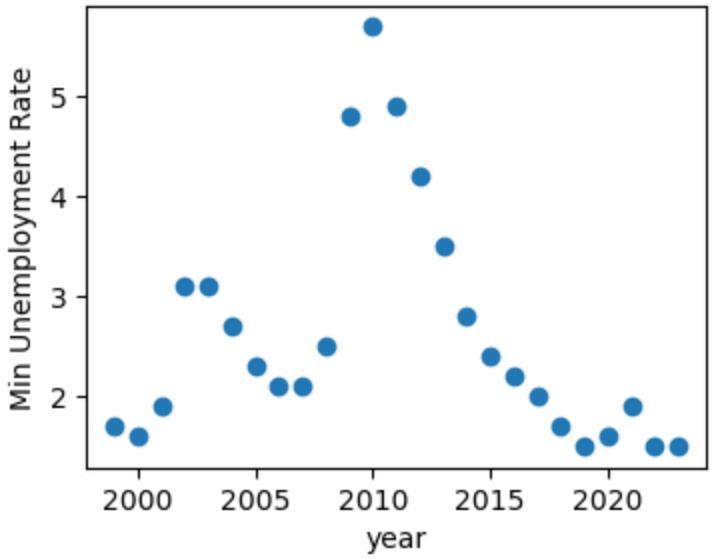


In [25]:
```
df_min=join_data_df.groupby(by='year', as_index=False).min()   # min values of each facto
df_min
```

Out[25]:

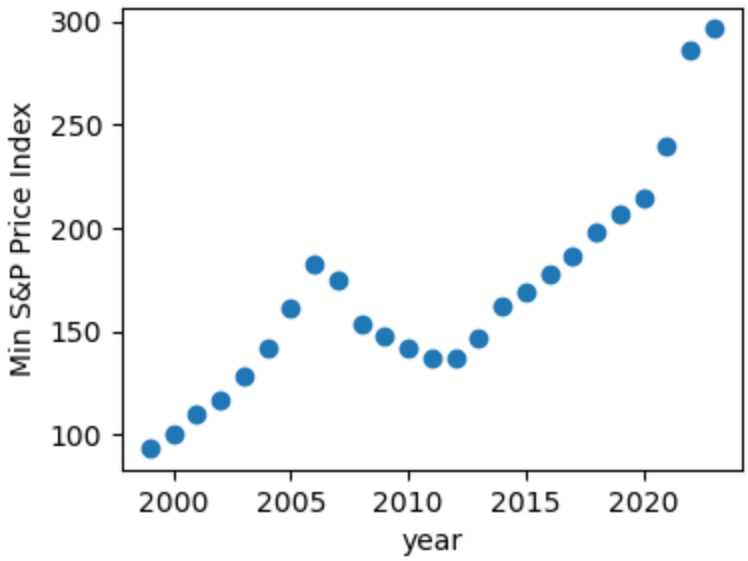
	year	DATE	Const_complt	un_constr	Cnstr_not_Strtd	Unemploy_Rate	Price_fact
0	1999	1999-01-01	67.0	175.0	36.0	1.7	93.207
1	2000	2000-01-01	79.0	170.0	37.0	1.6	100.551
2	2001	2001-01-01	75.0	170.0	39.0	1.9	109.846
3	2002	2002-01-01	77.0	190.0	40.0	3.1	117.143
4	2003	2003-01-01	77.0	202.0	51.0	3.1	128.460
5	2004	2004-01-01	82.0	231.0	54.0	2.7	141.646
6	2005	2005-01-01	102.0	264.0	70.0	2.3	161.288
7	2006	2006-01-01	113.0	290.0	80.0	2.1	182.320
8	2007	2007-01-01	169.0	234.0	72.0	2.1	174.342
9	2008	2008-01-01	166.0	145.0	42.0	2.5	153.618
10	2009	2009-01-01	96.0	106.0	29.0	4.8	147.695
11	2010	2010-01-01	77.0	86.0	26.0	5.7	142.061
12	2011	2011-01-01	56.0	70.0	23.0	4.9	136.675
13	2012	2012-01-01	40.0	71.0	22.0	4.2	136.533
14	2013	2013-01-01	37.0	87.0	22.0	3.5	146.827
15	2014	2014-01-01	43.0	112.0	29.0	2.8	161.927
16	2015	2015-01-01	48.0	112.0	36.0	2.4	168.634
17	2016	2016-01-01	55.0	139.0	36.0	2.2	177.272
18	2017	2017-01-01	59.0	153.0	45.0	2.0	186.800
19	2018	2018-01-01	59.0	178.0	52.0	1.7	198.294
20	2019	2019-01-01	74.0	190.0	51.0	1.5	206.495
21	2020	2020-01-01	40.0	172.0	52.0	1.6	214.904
22	2021	2021-01-01	32.0	188.0	75.0	1.9	239.413
23	2022	2022-01-01	31.0	268.0	92.0	1.5	285.708
24	2023	2023-01-01	66.0	262.0	86.0	1.5	296.601

In [26]:
```
plt.figure(figsize=(4,3))
plt.scatter(df_min.year,df_min.Unemploy_Rate)
plt.xlabel('year')
plt.ylabel('Min Unemployment Rate')
plt.show()
```

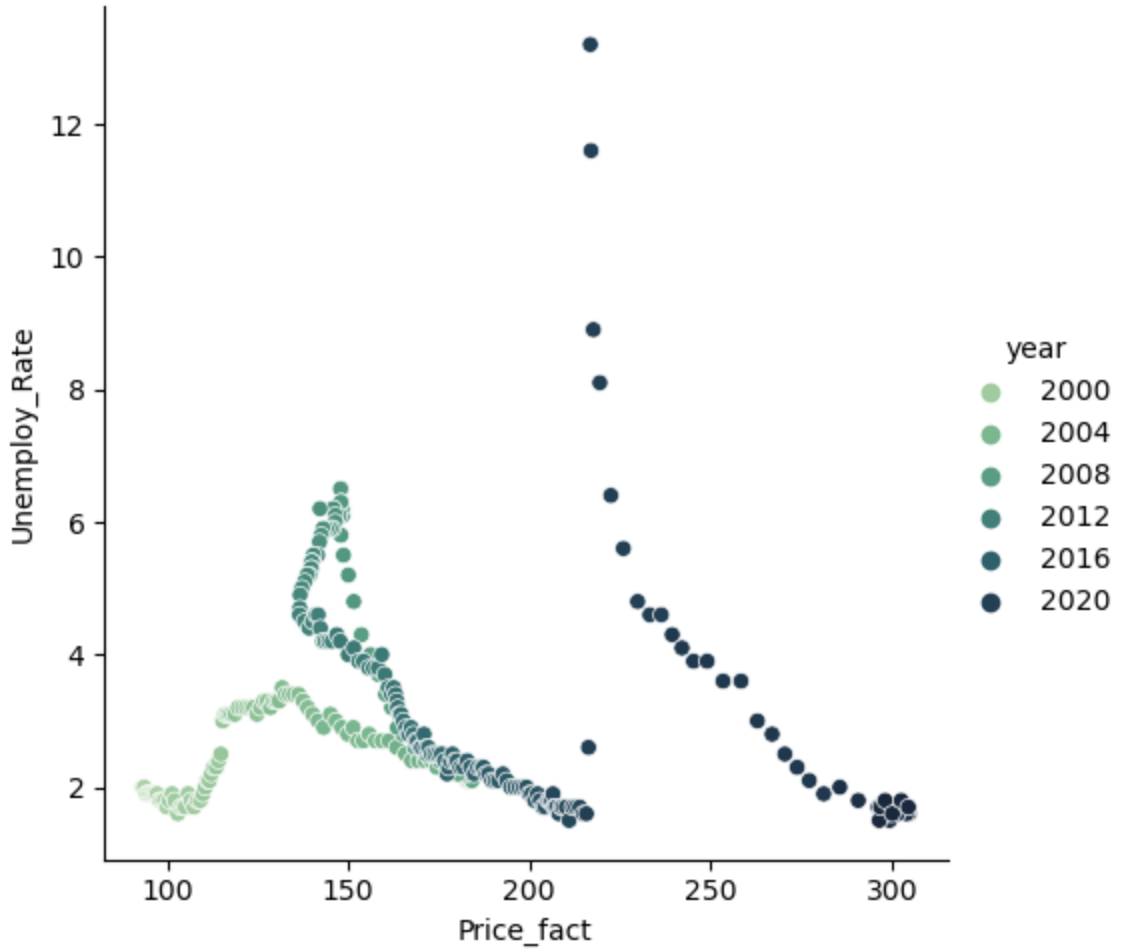


Minimum Unemployment Rate was seen in the year 2019

In [27]:
```
plt.figure(figsize=(4,3))
plt.scatter(df_min.year,df_min.Price_fact)
plt.xlabel('year')
plt.ylabel('Min S&P Price Index')
plt.show()
```



In [28]:
```
sns.relplot(x="Price_fact", y="Unemploy_Rate", hue="year", palette="ch:r=-.5,l=.75", dat
```



In [29]:
```
data_new=pd.read_csv('HNFSEPUSSA.csv')  # Total number of houses for sale
data_new.head(12)
```

Out[29]:

	DATE	HNFSEPUSSA
0	1963-01-01	235.0
1	1963-02-01	238.0
2	1963-03-01	242.0
3	1963-04-01	246.0
4	1963-05-01	248.0
5	1963-06-01	253.0
6	1963-07-01	254.0
7	1963-08-01	264.0
8	1963-09-01	257.0
9	1963-10-01	274.0
10	1963-11-01	258.0
11	1963-12-01	264.0

In [30]:
```
data_new.describe()
```

Out[30]:

	HNFSEPUSSA
count	727.000000
mean	313.906465

std	85.437649
min	142.000000
25%	253.500000
50%	312.000000
75%	362.000000
max	572.000000

In [31]:
```
data_new['year']=  pd.DatetimeIndex(data_new['DATE']).year
```

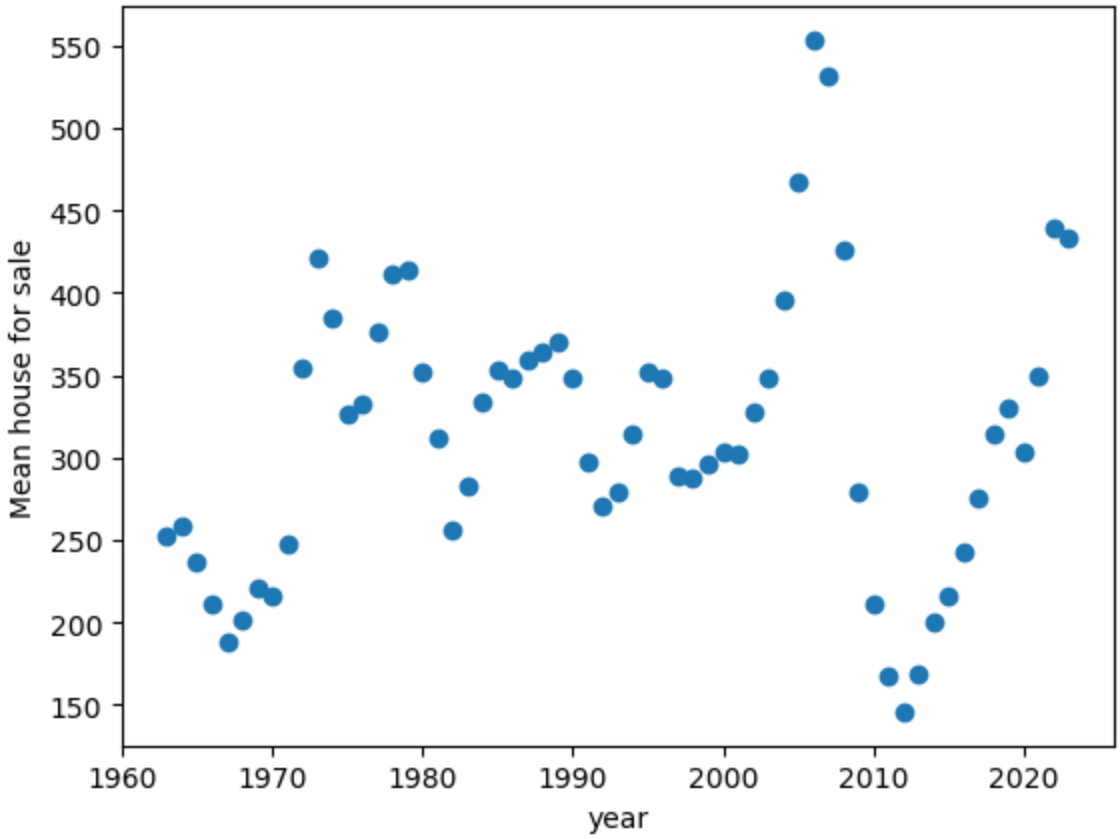
In [32]:
```
data_new1=data_new.groupby(by='year',as_index=False).mean()
data_new1
```

Out[32]:

	year	HNFSEPUSSA
0	1963	252.750000
1	1964	258.750000
2	1965	236.583333
3	1966	211.666667
4	1967	187.583333
...	...	...
56	2019	330.333333
57	2020	303.833333
58	2021	349.166667
59	2022	439.833333
60	2023	432.857143

61 rows × 2 columns

In [33]:
```
plt.scatter(data_new1.year,data_new1.HNFSEPUSSA)
plt.xlabel('year')
plt.ylabel('Mean house for sale')
plt.show()
```



Maximum number of houses were available in 2007 and least in 2012

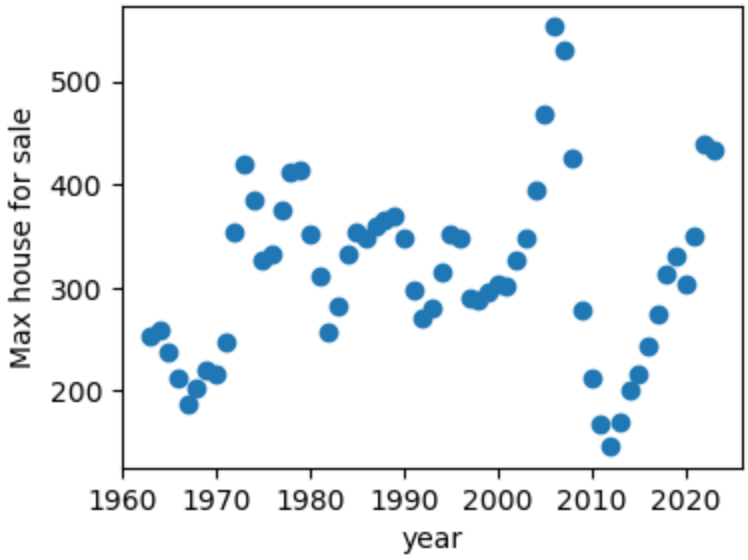
In [34]:
```
data_max=data_new1.groupby(by='year',as_index=False).max()
data_max
```

Out[34]:

	year	HNFSEPUSSA
0	1963	252.750000
1	1964	258.750000
2	1965	236.583333
3	1966	211.666667
4	1967	187.583333
...	...	...
56	2019	330.333333
57	2020	303.833333
58	2021	349.166667
59	2022	439.833333
60	2023	432.857143

61 rows × 2 columns

In [35]:
```
plt.figure(figsize=(4,3))
plt.scatter(data_max.year,data_max.HNFSEPUSSA)
plt.xlabel('year')
plt.ylabel('Max house for sale')
plt.show()
```



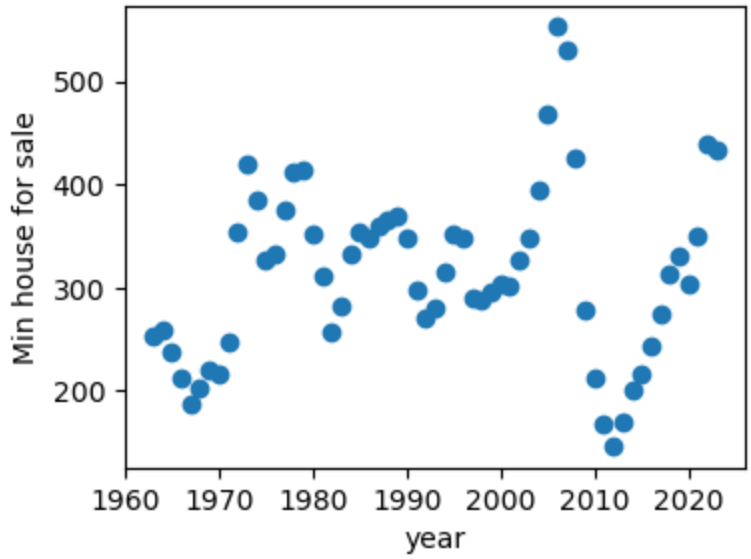
In [36]:
```
data_min=data_new1.groupby(by='year',as_index=False).min()
data_min
```

Out[36]:

	year	HNFSEPUSSA
0	1963	252.750000
1	1964	258.750000
2	1965	236.583333
3	1966	211.666667
4	1967	187.583333
...	...	...
56	2019	330.333333
57	2020	303.833333
58	2021	349.166667
59	2022	439.833333
60	2023	432.857143

61 rows × 2 columns

In [37]:
```
plt.figure(figsize=(4,3))
plt.scatter(data_min.year,data_min.HNFSEPUSSA)
plt.xlabel('year')
plt.ylabel('Min house for sale')
plt.show()
```



In [38]:
```
join_data_df.corr()
```

Out[38]:

	Const_complt	un_constr	Cnstr_not_Strtd	Unemploy_Rate	Price_fact	year
Const_complt	1.000000	0.369077	0.259954	-0.033380	-0.189004	-0.439836
un_constr	0.369077	1.000000	0.923796	-0.525463	0.449616	-0.061550
Cnstr_not_Strtd	0.259954	0.923796	1.000000	-0.396526	0.686356	0.218150
Unemploy_Rate	-0.033380	-0.525463	-0.396526	1.000000	-0.169878	0.050090
Price_fact	-0.189004	0.449616	0.686356	-0.169878	1.000000	0.843770
year	-0.439836	-0.061550	0.218150	0.050090	0.843770	1.000000

In [39]:
```
data_df_new=pd.merge(data_new,join_data_df,on='DATE',how ='inner')
```

In [40]:
```
data_df_new
```

Out[40]:

	DATE	HNFSEPUSSA	year_x	Const_complt	un_constr	Cnstr_not_Strtd	Unemploy_Rate	Price_fact	year_y
0	1999-01-01	284.0	1999	68.0	178.0	38.0	2.0	93.207	1999
1	1999-02-01	285.0	1999	67.0	180.0	38.0	2.0	93.670	1999
2	1999-03-01	289.0	1999	68.0	185.0	36.0	1.9	94.216	1999
3	1999-04-01	290.0	1999	69.0	180.0	41.0	1.9	94.784	1999
4	1999-05-01	295.0	1999	72.0	184.0	39.0	1.9	95.343	1999
...	...	...	...	...	...	...	...	...	...
289	2023-02-01	437.0	2023	69.0	278.0	90.0	1.7	296.958	2023
290	2023-03-01	433.0	2023	70.0	273.0	90.0	1.8	298.210	2023
291	2023-04-01	430.0	2023	70.0	267.0	93.0	1.6	300.214	2023

292	2023-05-01	426.0	2023	66.0	266.0	94.0	1.8	302.657	2023
293	2023-06-01	428.0	2023	69.0	262.0	97.0	1.7	304.635	2023

294 rows × 9 columns

In [41]:
```
data_df_new=data_df_new.drop(columns=['year_x','year_y','Const_complt','un_constr','Cnst
data_df_new
```

Out[41]:

	DATE	HNFSEPUSSA	Unemploy_Rate	Price_fact
0	1999-01-01	284.0	2.0	93.207
1	1999-02-01	285.0	2.0	93.670
2	1999-03-01	289.0	1.9	94.216
3	1999-04-01	290.0	1.9	94.784
4	1999-05-01	295.0	1.9	95.343
...	...	...	...	...
289	2023-02-01	437.0	1.7	296.958
290	2023-03-01	433.0	1.8	298.210
291	2023-04-01	430.0	1.6	300.214
292	2023-05-01	426.0	1.8	302.657
293	2023-06-01	428.0	1.7	304.635

294 rows × 4 columns

In [42]:
```
data_df_new.rename(columns={'HNFSEPUSSA':'ttl_homes_avlbl_for_sale'},inplace=True)
data_df_new
```

Out[42]:

	DATE	ttl_homes_avlbl_for_sale	Unemploy_Rate	Price_fact
0	1999-01-01	284.0	2.0	93.207
1	1999-02-01	285.0	2.0	93.670
2	1999-03-01	289.0	1.9	94.216
3	1999-04-01	290.0	1.9	94.784
4	1999-05-01	295.0	1.9	95.343
...	...	...	...	...
289	2023-02-01	437.0	1.7	296.958
290	2023-03-01	433.0	1.8	298.210
291	2023-04-01	430.0	1.6	300.214
292	2023-05-01	426.0	1.8	302.657
293	2023-06-01	428.0	1.7	304.635

294 rows × 4 columns

In [43]:
```
data_df_new.drop(columns=['DATE'],axis=1,inplace=True)
```

In [44]:
```
data_df_new
```

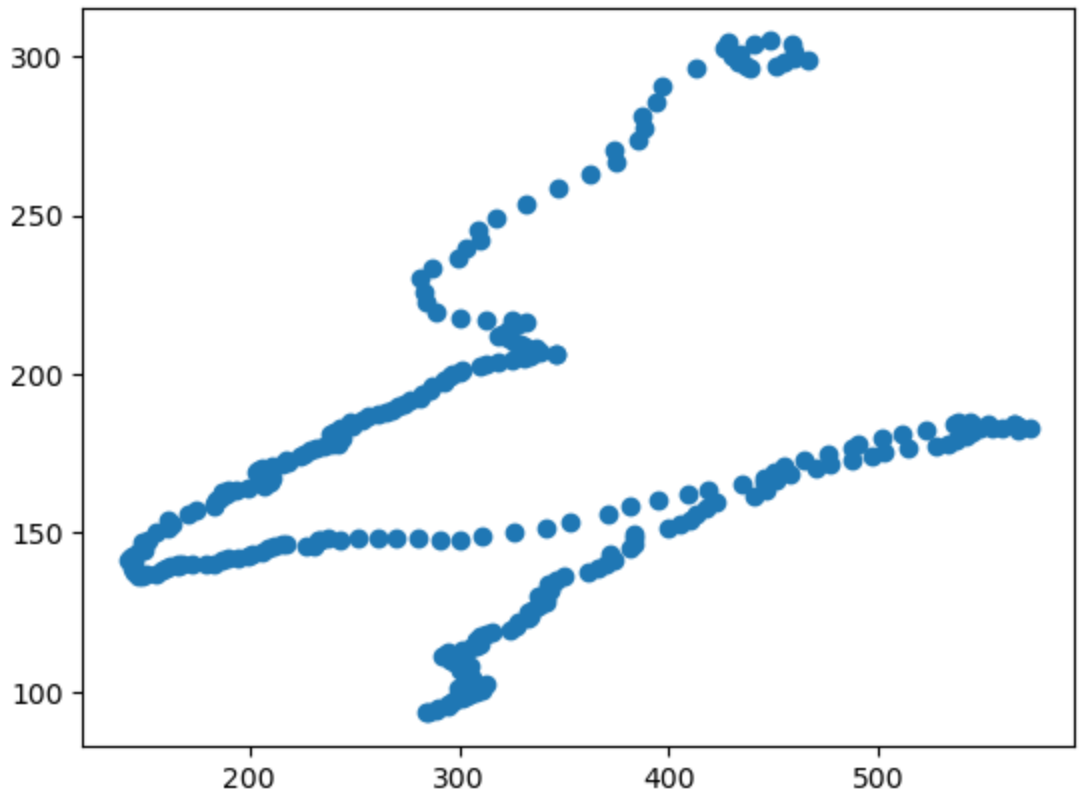
Out[44]:

	ttl_homes_avlbl_for_sale	Unemploy_Rate	Price_fact
0	284.0	2.0	93.207
1	285.0	2.0	93.670
2	289.0	1.9	94.216
3	290.0	1.9	94.784
4	295.0	1.9	95.343
...	...	...	...
289	437.0	1.7	296.958
290	433.0	1.8	298.210
291	430.0	1.6	300.214
292	426.0	1.8	302.657
293	428.0	1.7	304.635

294 rows × 3 columns

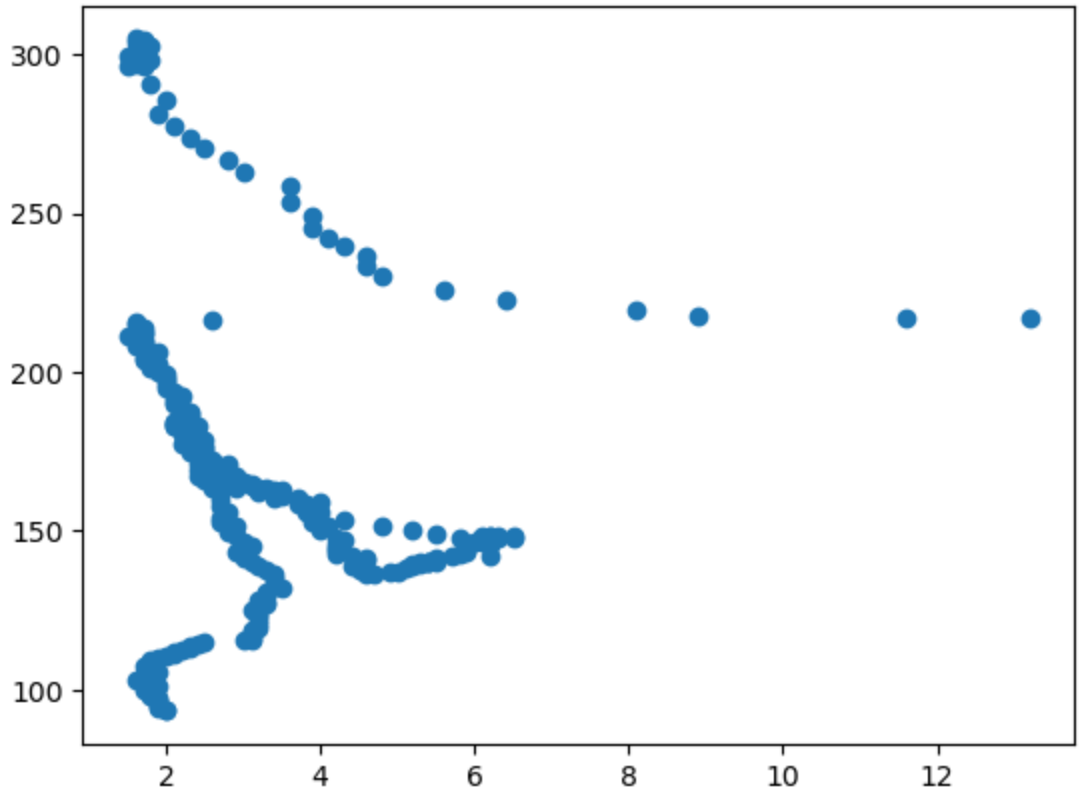
In [45]:
```
plt.scatter(data_df_new.ttl_homes_avlbl_for_sale,data_df_new.Price_fact)
```

Out[45]:
```
<matplotlib.collections.PathCollection at 0x1f1bd5afb20>
```



In [46]:
```
plt.scatter(data_df_new.Unemploy_Rate,data_df_new.Price_fact)
```

Out[46]:
```
<matplotlib.collections.PathCollection at 0x1f1bd558a00>
```



Building Data Science Model

In [47]:
```
X=data_df_new.drop(columns=['Price_fact'],axis=1)
Y=data_df_new['Price_fact']
```

In [48]:
```
from sklearn.model_selection import train_test_split
x_train,x_test,y_train,y_test = train_test_split(X,Y,test_size=0.2)
```

In [49]:
```
from sklearn.linear_model import LinearRegression
```

In [50]:
```
lin_reg =LinearRegression()
```

In [51]:
```
lin_reg.fit(x_train,y_train)
```

Out[51]:
```
▾ LinearRegression
LinearRegression()
```

In [52]:
```
Y_predict =lin_reg.predict(x_test)
Y_predict
```

Out[52]:
```
array([153.41606285, 172.37906649, 167.7940575 , 163.1885028 ,
       176.69664417, 172.82896734, 178.72432773, 185.62325703,
       171.28809403, 156.31359967, 192.23420933, 170.10797781,
       143.98557452, 173.14863302, 172.14680434, 177.76533069,
       171.47404406, 213.15803869, 171.14497565, 151.90687822,
       151.15785739, 182.40605453, 168.26798453, 156.61097941,
       210.5244527 , 143.84245614, 170.86814162, 171.92568516,
       171.46290109, 171.82713866, 156.25196256, 171.92568516,
       177.01630985, 155.70769723, 156.12590941, 166.26258692,
       175.48657952, 159.31246193, 166.04668845, 191.20835447,
       201.15908195, 175.01265248, 171.62830542, 146.25492296,
       198.20583028, 159.04573217, 188.55422277, 212.23247057,
       176.37697849, 172.68584896, 144.32752614, 175.46255334,
```

```
       192.11163666, 194.99977074, 159.81703895, 171.45349836,
       193.66539317, 147.0373727 , 212.23247057])
```

In [53]:
```
from sklearn.metrics import r2_score
```

In [54]:
```
score=[]
for i in range(1000):
    x_train,x_test,y_train,y_test = train_test_split(X,Y,test_size=0.2,random_state=i)
    lr=LinearRegression()
    lr.fit(x_train,y_train)
    yprd=lr.predict(x_test)
    score.append(r2_score(y_test,yprd))
```

In [55]:
```
import numpy as np
np.argmax(score)
```

Out[55]:
```
346
```

In [56]:
```
score[np.argmax(score)]
```

Out[56]:
```
0.1922218021173141
```

Since the realtion between dependent and independent variable is not linear LR model is not giving good result Therefor we are using Random forest regressor

In [57]:
```
from sklearn.ensemble import RandomForestRegressor
regressor1 = RandomForestRegressor(n_estimators = 100,max_depth=9,random_state = 0)
regressor1.fit(x_train, y_train)
```

Out[57]:
```
RandomForestRegressor(max_depth=9, random_state=0)
```

In [58]:
```
Y_pred = regressor1.predict(x_test)

r2_score(y_test,Y_pred)
```

Out[58]:
```
0.7087528973730306
```

In [59]:
```
scoree=[]
for i in range(1000):
    x_train,x_test,y_train,y_test = train_test_split(X,Y,test_size=0.2,random_state=i)
    regressor = RandomForestRegressor(n_estimators = 100, random_state = 0)
    regressor.fit(x_train, y_train)
    yprd=regressor.predict(x_test)
    scoree.append(r2_score(y_test,yprd))
```

In [60]:
```
np.argmax(scoree)
```

Out[60]:
```
193
```

In [61]:
```
scoree[np.argmax(scoree)]
```

Out[61]:
```
0.935544507957901
```

In [63]:
```
!pip install -U notebook-as-pdf
!pyppeteer-install
```

```
Collecting notebook-as-pdf
  Downloading notebook_as_pdf-0.5.0-py3-none-any.whl (6.5 kB)
Collecting pyppeteer
  Downloading pyppeteer-1.0.2-py3-none-any.whl (83 kB)
     ---------------------------------------- 83.4/83.4 kB 671.0 kB/s eta 0:00:00
Requirement already satisfied: nbconvert in c:\users\lenovo\anaconda3\lib\site-packages
(from notebook-as-pdf) (6.5.4)
Collecting PyPDF2
  Downloading pypdf2-3.0.1-py3-none-any.whl (232 kB)
     ---------------------------------------- 232.6/232.6 kB 1.3 MB/s eta 0:00:00
Requirement already satisfied: nbclient>=0.5.0 in c:\users\lenovo\anaconda3\lib\site-pac
kages (from nbconvert->notebook-as-pdf) (0.5.13)
Requirement already satisfied: pygments>=2.4.1 in c:\users\lenovo\anaconda3\lib\site-pac
kages (from nbconvert->notebook-as-pdf) (2.15.1)
Requirement already satisfied: pandocfilters>=1.4.1 in c:\users\lenovo\anaconda3\lib\sit
e-packages (from nbconvert->notebook-as-pdf) (1.5.0)
Requirement already satisfied: beautifulsoup4 in c:\users\lenovo\anaconda3\lib\site-pack
ages (from nbconvert->notebook-as-pdf) (4.12.2)
Requirement already satisfied: packaging in c:\users\lenovo\anaconda3\lib\site-packages
(from nbconvert->notebook-as-pdf) (23.0)
Requirement already satisfied: nbformat>=5.1 in c:\users\lenovo\anaconda3\lib\site-packa
ges (from nbconvert->notebook-as-pdf) (5.7.0)
Requirement already satisfied: entrypoints>=0.2.2 in c:\users\lenovo\anaconda3\lib\site-
packages (from nbconvert->notebook-as-pdf) (0.4)
Requirement already satisfied: jupyter-core>=4.7 in c:\users\lenovo\anaconda3\lib\site-p
ackages (from nbconvert->notebook-as-pdf) (5.3.0)
Requirement already satisfied: defusedxml in c:\users\lenovo\anaconda3\lib\site-packages
(from nbconvert->notebook-as-pdf) (0.7.1)
Requirement already satisfied: mistune<2,>=0.8.1 in c:\users\lenovo\anaconda3\lib\site-p
ackages (from nbconvert->notebook-as-pdf) (0.8.4)
Requirement already satisfied: traitlets>=5.0 in c:\users\lenovo\anaconda3\lib\site-pack
ages (from nbconvert->notebook-as-pdf) (5.7.1)
Requirement already satisfied: jinja2>=3.0 in c:\users\lenovo\anaconda3\lib\site-package
s (from nbconvert->notebook-as-pdf) (3.1.2)
Requirement already satisfied: MarkupSafe>=2.0 in c:\users\lenovo\anaconda3\lib\site-pac
kages (from nbconvert->notebook-as-pdf) (2.1.1)
Requirement already satisfied: bleach in c:\users\lenovo\anaconda3\lib\site-packages (fr
om nbconvert->notebook-as-pdf) (4.1.0)
Requirement already satisfied: tinycss2 in c:\users\lenovo\anaconda3\lib\site-packages
(from nbconvert->notebook-as-pdf) (1.2.1)
Requirement already satisfied: jupyterlab-pygments in c:\users\lenovo\anaconda3\lib\site
-packages (from nbconvert->notebook-as-pdf) (0.1.2)
Requirement already satisfied: lxml in c:\users\lenovo\anaconda3\lib\site-packages (from
nbconvert->notebook-as-pdf) (4.9.2)
Requirement already satisfied: certifi>=2021 in c:\users\lenovo\anaconda3\lib\site-packa
ges (from pyppeteer->notebook-as-pdf) (2023.5.7)
Requirement already satisfied: tqdm<5.0.0,>=4.42.1 in c:\users\lenovo\anaconda3\lib\site
-packages (from pyppeteer->notebook-as-pdf) (4.65.0)
Collecting websockets<11.0,>=10.0
  Downloading websockets-10.4-cp310-cp310-win_amd64.whl (101 kB)
     ---------------------------------------- 101.4/101.4 kB 2.9 MB/s eta 0:00:00
Requirement already satisfied: appdirs<2.0.0,>=1.4.3 in c:\users\lenovo\anaconda3\lib\si
te-packages (from pyppeteer->notebook-as-pdf) (1.4.4)
Requirement already satisfied: importlib-metadata>=1.4 in c:\users\lenovo\anaconda3\lib
\site-packages (from pyppeteer->notebook-as-pdf) (6.6.0)
Requirement already satisfied: urllib3<2.0.0,>=1.25.8 in c:\users\lenovo\anaconda3\lib\s
ite-packages (from pyppeteer->notebook-as-pdf) (1.26.15)
Collecting pyee<9.0.0,>=8.1.0
  Downloading pyee-8.2.2-py2.py3-none-any.whl (12 kB)
Requirement already satisfied: zipp>=0.5 in c:\users\lenovo\anaconda3\lib\site-packages
(from importlib-metadata>=1.4->pyppeteer->notebook-as-pdf) (3.15.0)
Requirement already satisfied: platformdirs>=2.5 in c:\users\lenovo\anaconda3\lib\site-p
ackages (from jupyter-core>=4.7->nbconvert->notebook-as-pdf) (2.5.2)
Requirement already satisfied: pywin32>=300 in c:\users\lenovo\anaconda3\lib\site-packag
es (from jupyter-core>=4.7->nbconvert->notebook-as-pdf) (305.1)
Requirement already satisfied: nest-asyncio in c:\users\lenovo\anaconda3\lib\site-packag
```

```
es (from nbclient>=0.5.0->nbconvert->notebook-as-pdf) (1.5.6)
Requirement already satisfied: jupyter-client>=6.1.5 in c:\users\lenovo\anaconda3\lib\site-packages (from nbclient>=0.5.0->nbconvert->notebook-as-pdf) (8.1.0)
Requirement already satisfied: jsonschema>=2.6 in c:\users\lenovo\anaconda3\lib\site-packages (from nbformat>=5.1->nbconvert->notebook-as-pdf) (4.17.3)
Requirement already satisfied: fastjsonschema in c:\users\lenovo\anaconda3\lib\site-packages (from nbformat>=5.1->nbconvert->notebook-as-pdf) (2.16.2)
Requirement already satisfied: colorama in c:\users\lenovo\appdata\roaming\python\python310\site-packages (from tqdm<5.0.0,>=4.42.1->pyppeteer->notebook-as-pdf) (0.4.3)
Requirement already satisfied: soupsieve>1.2 in c:\users\lenovo\anaconda3\lib\site-packages (from beautifulsoup4->nbconvert->notebook-as-pdf) (2.4)
Requirement already satisfied: six>=1.9.0 in c:\users\lenovo\appdata\roaming\python\python310\site-packages (from bleach->nbconvert->notebook-as-pdf) (1.14.0)
Requirement already satisfied: webencodings in c:\users\lenovo\anaconda3\lib\site-packages (from bleach->nbconvert->notebook-as-pdf) (0.5.1)
Requirement already satisfied: attrs>=17.4.0 in c:\users\lenovo\anaconda3\lib\site-packages (from jsonschema>=2.6->nbformat>=5.1->nbconvert->notebook-as-pdf) (22.1.0)
Requirement already satisfied: pyrsistent!=0.17.0,!=0.17.1,!=0.17.2,>=0.14.0 in c:\users\lenovo\anaconda3\lib\site-packages (from jsonschema>=2.6->nbformat>=5.1->nbconvert->notebook-as-pdf) (0.18.0)
Requirement already satisfied: python-dateutil>=2.8.2 in c:\users\lenovo\anaconda3\lib\site-packages (from jupyter-client>=6.1.5->nbclient>=0.5.0->nbconvert->notebook-as-pdf) (2.8.2)
Requirement already satisfied: tornado>=6.2 in c:\users\lenovo\anaconda3\lib\site-packages (from jupyter-client>=6.1.5->nbclient>=0.5.0->nbconvert->notebook-as-pdf) (6.2)
Requirement already satisfied: pyzmq>=23.0 in c:\users\lenovo\anaconda3\lib\site-packages (from jupyter-client>=6.1.5->nbclient>=0.5.0->nbconvert->notebook-as-pdf) (25.0.2)
Installing collected packages: pyee, websockets, PyPDF2, pyppeteer, notebook-as-pdf
Successfully installed PyPDF2-3.0.1 notebook-as-pdf-0.5.0 pyee-8.2.2 pyppeteer-1.0.2 websockets-10.4
[INFO] Starting Chromium download.

  0%|          | 0.00/137M [00:00<?, ?b/s]
  0%|          | 20.5k/137M [00:00<19:28, 117kb/s]
  0%|          | 51.2k/137M [00:00<19:27, 117kb/s]
  0%|          | 81.9k/137M [00:00<19:20, 118kb/s]
  0%|          | 133k/137M [00:00<11:42, 195kb/s]
  0%|          | 164k/137M [00:00<11:38, 196kb/s]
  0%|          | 256k/137M [00:01<06:24, 356kb/s]
  0%|          | 328k/137M [00:01<05:11, 438kb/s]
  0%|          | 481k/137M [00:01<03:13, 704kb/s]
  0%|          | 614k/137M [00:01<02:37, 866kb/s]
  1%|          | 860k/137M [00:01<01:45, 1.29Mb/s]
  1%|          | 1.19M/137M [00:01<01:14, 1.83Mb/s]
  1%|1         | 1.60M/137M [00:01<00:56, 2.40Mb/s]
  2%|1         | 2.12M/137M [00:01<00:44, 3.06Mb/s]
  2%|1         | 2.63M/137M [00:01<00:37, 3.61Mb/s]
  2%|2         | 3.07M/137M [00:02<00:35, 3.79Mb/s]
  3%|2         | 3.46M/137M [00:02<00:36, 3.61Mb/s]
  3%|2         | 3.83M/137M [00:02<00:45, 2.96Mb/s]
  3%|3         | 4.56M/137M [00:02<00:33, 4.00Mb/s]
  4%|3         | 5.15M/137M [00:02<00:29, 4.44Mb/s]
  4%|4         | 5.70M/137M [00:02<00:28, 4.66Mb/s]
  5%|4         | 6.23M/137M [00:02<00:27, 4.80Mb/s]
  5%|4         | 6.73M/137M [00:02<00:26, 4.84Mb/s]
  5%|5         | 7.23M/137M [00:02<00:27, 4.71Mb/s]
  6%|5         | 7.72M/137M [00:03<00:27, 4.75Mb/s]
  6%|6         | 8.22M/137M [00:03<00:28, 4.60Mb/s]
  6%|6         | 8.69M/137M [00:03<00:31, 4.02Mb/s]
  7%|6         | 9.11M/137M [00:03<00:43, 2.94Mb/s]
  7%|7         | 9.98M/137M [00:03<00:30, 4.15Mb/s]
  8%|7         | 10.5M/137M [00:03<00:29, 4.28Mb/s]
  8%|8         | 11.0M/137M [00:03<00:29, 4.22Mb/s]
  8%|8         | 11.5M/137M [00:04<00:28, 4.40Mb/s]
  9%|8         | 12.0M/137M [00:04<00:27, 4.47Mb/s]
```

```
 9%|9         | 12.6M/137M [00:04<00:25, 4.97Mb/s]
10%|9         | 13.3M/137M [00:04<00:22, 5.40Mb/s]
10%|#         | 13.9M/137M [00:04<00:21, 5.63Mb/s]
11%|#         | 14.5M/137M [00:04<00:22, 5.50Mb/s]
11%|#1        | 15.1M/137M [00:04<00:22, 5.35Mb/s]
11%|#1        | 15.6M/137M [00:04<00:30, 4.04Mb/s]
12%|#2        | 16.5M/137M [00:04<00:23, 5.12Mb/s]
12%|#2        | 17.1M/137M [00:05<00:28, 4.19Mb/s]
13%|#2        | 17.6M/137M [00:05<00:27, 4.27Mb/s]
13%|#3        | 18.3M/137M [00:05<00:24, 4.88Mb/s]
14%|#3        | 18.8M/137M [00:05<00:23, 5.01Mb/s]
14%|#4        | 19.4M/137M [00:05<00:23, 5.06Mb/s]
15%|#4        | 20.0M/137M [00:05<00:22, 5.23Mb/s]
15%|#4        | 20.5M/137M [00:05<00:23, 4.86Mb/s]
15%|#5        | 21.0M/137M [00:05<00:23, 4.85Mb/s]
16%|#5        | 21.5M/137M [00:06<00:28, 4.06Mb/s]
16%|#6        | 22.1M/137M [00:06<00:28, 4.07Mb/s]
16%|#6        | 22.6M/137M [00:06<00:33, 3.41Mb/s]
17%|#6        | 23.1M/137M [00:06<00:30, 3.69Mb/s]
17%|#7        | 23.6M/137M [00:06<00:28, 4.01Mb/s]
18%|#7        | 24.0M/137M [00:06<00:28, 4.02Mb/s]
18%|#7        | 24.6M/137M [00:06<00:25, 4.46Mb/s]
18%|#8        | 25.2M/137M [00:06<00:23, 4.71Mb/s]
19%|#8        | 25.7M/137M [00:07<00:23, 4.83Mb/s]
19%|#9        | 26.3M/137M [00:07<00:21, 5.04Mb/s]
20%|#9        | 26.8M/137M [00:07<00:30, 3.63Mb/s]
20%|##        | 27.5M/137M [00:07<00:25, 4.37Mb/s]
20%|##        | 28.0M/137M [00:07<00:24, 4.47Mb/s]
21%|##        | 28.5M/137M [00:07<00:25, 4.29Mb/s]
21%|##1       | 29.0M/137M [00:07<00:25, 4.21Mb/s]
22%|##1       | 29.5M/137M [00:07<00:25, 4.16Mb/s]
22%|##1       | 29.9M/137M [00:08<00:26, 4.07Mb/s]
22%|##2       | 30.3M/137M [00:08<00:27, 3.89Mb/s]
22%|##2       | 30.8M/137M [00:08<00:26, 4.01Mb/s]
23%|##2       | 31.3M/137M [00:08<00:24, 4.24Mb/s]
23%|##3       | 31.7M/137M [00:08<00:34, 3.05Mb/s]
24%|##3       | 32.6M/137M [00:08<00:24, 4.30Mb/s]
24%|##4       | 33.1M/137M [00:08<00:27, 3.79Mb/s]
25%|##4       | 34.0M/137M [00:09<00:21, 4.82Mb/s]
25%|##5       | 34.5M/137M [00:09<00:23, 4.39Mb/s]
26%|##5       | 35.1M/137M [00:09<00:25, 4.02Mb/s]
26%|##5       | 35.6M/137M [00:09<00:24, 4.17Mb/s]
26%|##6       | 36.0M/137M [00:09<00:24, 4.18Mb/s]
27%|##6       | 36.4M/137M [00:09<00:24, 4.06Mb/s]
27%|##6       | 36.9M/137M [00:09<00:23, 4.21Mb/s]
27%|##7       | 37.4M/137M [00:09<00:26, 3.73Mb/s]
28%|##7       | 37.8M/137M [00:10<00:25, 3.87Mb/s]
28%|##7       | 38.2M/137M [00:10<00:26, 3.79Mb/s]
28%|##8       | 38.6M/137M [00:10<00:25, 3.86Mb/s]
28%|##8       | 39.0M/137M [00:10<00:27, 3.58Mb/s]
29%|##8       | 39.5M/137M [00:10<00:24, 4.03Mb/s]
29%|##9       | 40.0M/137M [00:10<00:23, 4.06Mb/s]
29%|##9       | 40.4M/137M [00:10<00:24, 4.00Mb/s]
30%|##9       | 40.8M/137M [00:10<00:24, 3.90Mb/s]
30%|###       | 41.2M/137M [00:10<00:25, 3.77Mb/s]
30%|###       | 41.6M/137M [00:11<00:25, 3.79Mb/s]
31%|###       | 42.0M/137M [00:11<00:28, 3.35Mb/s]
31%|###       | 42.3M/137M [00:11<00:31, 2.96Mb/s]
31%|###1      | 42.6M/137M [00:11<00:31, 2.95Mb/s]
31%|###1      | 42.9M/137M [00:11<00:34, 2.70Mb/s]
32%|###1      | 43.2M/137M [00:11<00:39, 2.35Mb/s]
32%|###1      | 43.5M/137M [00:11<00:43, 2.16Mb/s]
32%|###1      | 43.7M/137M [00:12<00:47, 1.97Mb/s]
32%|###2      | 44.0M/137M [00:12<00:41, 2.26Mb/s]
32%|###2      | 44.3M/137M [00:12<00:39, 2.37Mb/s]
33%|###2      | 44.6M/137M [00:12<00:57, 1.59Mb/s]
```

```
33%|###3      | 45.3M/137M [00:12<00:35, 2.58Mb/s]
33%|###3      | 45.6M/137M [00:12<00:34, 2.64Mb/s]
34%|###3      | 46.0M/137M [00:12<00:37, 2.42Mb/s]
34%|###3      | 46.2M/137M [00:13<00:37, 2.45Mb/s]
34%|###3      | 46.5M/137M [00:13<00:37, 2.38Mb/s]
34%|###4      | 46.8M/137M [00:13<00:40, 2.23Mb/s]
34%|###4      | 47.1M/137M [00:13<00:37, 2.41Mb/s]
35%|###4      | 47.4M/137M [00:13<00:38, 2.31Mb/s]
35%|###4      | 47.6M/137M [00:13<00:46, 1.91Mb/s]
35%|###5      | 48.2M/137M [00:13<00:31, 2.78Mb/s]
35%|###5      | 48.6M/137M [00:14<00:29, 2.97Mb/s]
36%|###5      | 48.9M/137M [00:14<00:28, 3.09Mb/s]
36%|###5      | 49.3M/137M [00:14<00:28, 3.04Mb/s]
36%|###6      | 49.7M/137M [00:14<00:26, 3.32Mb/s]
37%|###6      | 50.0M/137M [00:14<00:25, 3.39Mb/s]
37%|###6      | 50.4M/137M [00:14<00:24, 3.56Mb/s]
37%|###7      | 50.8M/137M [00:14<00:23, 3.68Mb/s]
37%|###7      | 51.3M/137M [00:14<00:21, 3.95Mb/s]
38%|###7      | 51.7M/137M [00:14<00:22, 3.77Mb/s]
38%|###8      | 52.1M/137M [00:15<00:30, 2.78Mb/s]
38%|###8      | 52.5M/137M [00:15<00:27, 3.07Mb/s]
39%|###8      | 52.9M/137M [00:15<00:25, 3.32Mb/s]
39%|###8      | 53.3M/137M [00:15<00:24, 3.41Mb/s]
39%|###9      | 53.7M/137M [00:15<00:23, 3.56Mb/s]
40%|###9      | 54.1M/137M [00:15<00:23, 3.52Mb/s]
40%|###9      | 54.5M/137M [00:15<00:23, 3.50Mb/s]
40%|####      | 54.8M/137M [00:15<00:23, 3.47Mb/s]
40%|####      | 55.3M/137M [00:15<00:21, 3.76Mb/s]
41%|####      | 55.7M/137M [00:16<00:22, 3.59Mb/s]
41%|####      | 56.0M/137M [00:16<00:23, 3.43Mb/s]
41%|####1     | 56.4M/137M [00:16<00:28, 2.81Mb/s]
42%|####1     | 57.0M/137M [00:16<00:22, 3.59Mb/s]
42%|####1     | 57.5M/137M [00:16<00:20, 3.87Mb/s]
42%|####2     | 57.9M/137M [00:16<00:22, 3.48Mb/s]
43%|####2     | 58.3M/137M [00:16<00:21, 3.58Mb/s]
43%|####2     | 58.7M/137M [00:16<00:22, 3.54Mb/s]
43%|####3     | 59.0M/137M [00:17<00:21, 3.58Mb/s]
43%|####3     | 59.4M/137M [00:17<00:21, 3.60Mb/s]
44%|####3     | 59.9M/137M [00:17<00:20, 3.80Mb/s]
44%|####4     | 60.4M/137M [00:17<00:17, 4.26Mb/s]
44%|####4     | 60.8M/137M [00:17<00:18, 4.05Mb/s]
45%|####4     | 61.3M/137M [00:17<00:24, 3.07Mb/s]
45%|####5     | 61.9M/137M [00:17<00:20, 3.74Mb/s]
46%|####5     | 62.4M/137M [00:17<00:18, 4.06Mb/s]
46%|####6     | 63.1M/137M [00:17<00:16, 4.53Mb/s]
46%|####6     | 63.6M/137M [00:18<00:16, 4.55Mb/s]
47%|####6     | 64.0M/137M [00:18<00:16, 4.50Mb/s]
47%|####7     | 64.5M/137M [00:18<00:16, 4.28Mb/s]
48%|####7     | 65.1M/137M [00:18<00:15, 4.56Mb/s]
48%|####7     | 65.5M/137M [00:18<00:15, 4.53Mb/s]
48%|####8     | 66.3M/137M [00:18<00:13, 5.09Mb/s]
49%|####8     | 66.9M/137M [00:18<00:13, 5.35Mb/s]
49%|####9     | 67.4M/137M [00:18<00:15, 4.41Mb/s]
50%|####9     | 68.2M/137M [00:19<00:13, 5.18Mb/s]
50%|#####     | 68.8M/137M [00:19<00:12, 5.46Mb/s]
51%|#####     | 69.4M/137M [00:19<00:13, 4.88Mb/s]
51%|#####1    | 69.9M/137M [00:19<00:14, 4.55Mb/s]
51%|#####1    | 70.4M/137M [00:19<00:14, 4.61Mb/s]
52%|#####1    | 70.9M/137M [00:19<00:14, 4.68Mb/s]
52%|#####2    | 71.4M/137M [00:19<00:15, 4.32Mb/s]
52%|#####2    | 71.8M/137M [00:19<00:15, 4.23Mb/s]
53%|#####2    | 72.4M/137M [00:19<00:14, 4.42Mb/s]
53%|#####3    | 72.9M/137M [00:20<00:17, 3.66Mb/s]
54%|#####3    | 73.5M/137M [00:20<00:14, 4.27Mb/s]
54%|#####4    | 74.1M/137M [00:20<00:13, 4.53Mb/s]
55%|#####4    | 74.9M/137M [00:20<00:11, 5.35Mb/s]
```

```
55%|#####5    | 75.5M/137M [00:20<00:11, 5.43Mb/s]
56%|#####5    | 76.2M/137M [00:20<00:10, 5.60Mb/s]
56%|#####6    | 76.8M/137M [00:20<00:10, 5.71Mb/s]
57%|#####6    | 77.4M/137M [00:20<00:10, 5.71Mb/s]
57%|#####7    | 78.1M/137M [00:21<00:09, 5.95Mb/s]
58%|#####7    | 78.7M/137M [00:21<00:09, 5.88Mb/s]
58%|#####7    | 79.3M/137M [00:21<00:10, 5.61Mb/s]
58%|#####8    | 79.9M/137M [00:21<00:13, 4.28Mb/s]
59%|#####8    | 80.7M/137M [00:21<00:11, 5.07Mb/s]
59%|#####9    | 81.3M/137M [00:21<00:13, 4.16Mb/s]
60%|#####9    | 81.7M/137M [00:21<00:15, 3.56Mb/s]
60%|#####9    | 82.1M/137M [00:22<00:15, 3.43Mb/s]
60%|######    | 82.5M/137M [00:22<00:15, 3.40Mb/s]
61%|######    | 82.9M/137M [00:22<00:15, 3.47Mb/s]
61%|######    | 83.3M/137M [00:22<00:15, 3.38Mb/s]
61%|######1   | 83.9M/137M [00:22<00:13, 4.01Mb/s]
62%|######1   | 84.4M/137M [00:22<00:16, 3.20Mb/s]
62%|######2   | 85.0M/137M [00:22<00:13, 3.97Mb/s]
62%|######2   | 85.5M/137M [00:22<00:12, 4.04Mb/s]
63%|######2   | 86.0M/137M [00:23<00:13, 3.82Mb/s]
63%|######3   | 86.4M/137M [00:23<00:13, 3.74Mb/s]
63%|######3   | 86.8M/137M [00:23<00:13, 3.60Mb/s]
64%|######3   | 87.3M/137M [00:23<00:12, 3.94Mb/s]
64%|######4   | 87.7M/137M [00:23<00:13, 3.75Mb/s]
64%|######4   | 88.1M/137M [00:23<00:13, 3.75Mb/s]
65%|######4   | 88.5M/137M [00:23<00:12, 3.74Mb/s]
65%|######4   | 88.9M/137M [00:23<00:13, 3.60Mb/s]
65%|######5   | 89.3M/137M [00:24<00:13, 3.46Mb/s]
66%|######5   | 89.8M/137M [00:24<00:12, 3.83Mb/s]
66%|######5   | 90.3M/137M [00:24<00:11, 4.00Mb/s]
66%|######6   | 90.7M/137M [00:24<00:12, 3.72Mb/s]
67%|######6   | 91.1M/137M [00:24<00:16, 2.77Mb/s]
67%|######6   | 91.4M/137M [00:24<00:15, 2.85Mb/s]
67%|######6   | 91.7M/137M [00:24<00:17, 2.64Mb/s]
67%|######7   | 92.0M/137M [00:25<00:17, 2.50Mb/s]
67%|######7   | 92.3M/137M [00:25<00:17, 2.51Mb/s]
68%|######7   | 92.7M/137M [00:25<00:19, 2.30Mb/s]
68%|######7   | 92.9M/137M [00:25<00:21, 2.05Mb/s]
68%|######8   | 93.3M/137M [00:25<00:18, 2.33Mb/s]
69%|######8   | 93.8M/137M [00:25<00:14, 2.89Mb/s]
69%|######8   | 94.1M/137M [00:25<00:17, 2.44Mb/s]
69%|######8   | 94.4M/137M [00:25<00:17, 2.47Mb/s]
69%|######9   | 94.7M/137M [00:26<00:17, 2.42Mb/s]
69%|######9   | 94.9M/137M [00:26<00:23, 1.82Mb/s]
70%|######9   | 95.2M/137M [00:26<00:27, 1.54Mb/s]
70%|######9   | 95.5M/137M [00:26<00:22, 1.87Mb/s]
70%|######9   | 95.7M/137M [00:26<00:24, 1.65Mb/s]
70%|#######   | 95.9M/137M [00:27<00:28, 1.42Mb/s]
70%|#######   | 96.2M/137M [00:27<00:23, 1.70Mb/s]
70%|#######   | 96.4M/137M [00:27<00:23, 1.74Mb/s]
71%|#######   | 96.6M/137M [00:27<00:23, 1.74Mb/s]
71%|#######   | 97.1M/137M [00:27<00:15, 2.63Mb/s]
71%|#######1  | 97.6M/137M [00:27<00:12, 3.23Mb/s]
72%|#######1  | 98.0M/137M [00:27<00:11, 3.25Mb/s]
72%|#######1  | 98.4M/137M [00:27<00:14, 2.61Mb/s]
72%|#######2  | 98.8M/137M [00:28<00:12, 2.98Mb/s]
72%|#######2  | 99.1M/137M [00:28<00:14, 2.55Mb/s]
73%|#######2  | 99.4M/137M [00:28<00:17, 2.16Mb/s]
73%|#######2  | 99.7M/137M [00:28<00:16, 2.21Mb/s]
73%|#######2  | 99.9M/137M [00:28<00:16, 2.30Mb/s]
73%|#######3  | 100M/137M [00:28<00:16, 2.17Mb/s]
73%|#######3  | 100M/137M [00:28<00:16, 2.16Mb/s]
74%|#######3  | 101M/137M [00:28<00:16, 2.15Mb/s]
74%|#######3  | 101M/137M [00:29<00:13, 2.66Mb/s]
74%|#######4  | 101M/137M [00:29<00:14, 2.54Mb/s]
74%|#######4  | 102M/137M [00:29<00:10, 3.36Mb/s]
```

```
75%|#######4  | 102M/137M [00:29<00:11, 3.09Mb/s]
75%|#######5  | 103M/137M [00:29<00:09, 3.57Mb/s]
76%|#######5  | 103M/137M [00:29<00:07, 4.42Mb/s]
76%|#######5  | 104M/137M [00:29<00:07, 4.42Mb/s]
76%|#######6  | 104M/137M [00:29<00:07, 4.39Mb/s]
77%|#######6  | 105M/137M [00:29<00:07, 4.26Mb/s]
77%|#######7  | 106M/137M [00:30<00:06, 4.71Mb/s]
78%|#######7  | 107M/137M [00:30<00:05, 5.86Mb/s]
78%|#######8  | 107M/137M [00:30<00:06, 4.30Mb/s]
79%|#######8  | 108M/137M [00:30<00:08, 3.44Mb/s]
79%|#######9  | 108M/137M [00:30<00:08, 3.34Mb/s]
79%|#######9  | 109M/137M [00:30<00:08, 3.21Mb/s]
80%|#######9  | 109M/137M [00:31<00:09, 3.00Mb/s]
80%|#######9  | 109M/137M [00:31<00:09, 2.93Mb/s]
80%|########  | 110M/137M [00:31<00:09, 2.81Mb/s]
80%|########  | 110M/137M [00:31<00:09, 2.76Mb/s]
80%|########  | 110M/137M [00:31<00:09, 2.69Mb/s]
81%|########  | 110M/137M [00:31<00:12, 2.08Mb/s]
81%|########1 | 111M/137M [00:31<00:10, 2.57Mb/s]
81%|########1 | 111M/137M [00:32<00:09, 2.62Mb/s]
82%|########1 | 112M/137M [00:32<00:08, 2.97Mb/s]
82%|########1 | 112M/137M [00:32<00:08, 2.94Mb/s]
82%|########2 | 112M/137M [00:32<00:08, 2.76Mb/s]
82%|########2 | 113M/137M [00:32<00:09, 2.47Mb/s]
83%|########2 | 113M/137M [00:32<00:08, 2.78Mb/s]
83%|########2 | 113M/137M [00:32<00:09, 2.51Mb/s]
83%|########2 | 114M/137M [00:33<00:12, 1.90Mb/s]
83%|########3 | 114M/137M [00:33<00:10, 2.17Mb/s]
83%|########3 | 114M/137M [00:33<00:11, 2.04Mb/s]
84%|########3 | 114M/137M [00:33<00:10, 2.10Mb/s]
84%|########3 | 115M/137M [00:33<00:10, 2.07Mb/s]
84%|########3 | 115M/137M [00:33<00:09, 2.34Mb/s]
84%|########4 | 115M/137M [00:33<00:09, 2.34Mb/s]
84%|########4 | 115M/137M [00:33<00:08, 2.51Mb/s]
85%|########4 | 116M/137M [00:33<00:07, 2.67Mb/s]
85%|########4 | 116M/137M [00:34<00:08, 2.35Mb/s]
85%|########4 | 116M/137M [00:34<00:08, 2.33Mb/s]
85%|########5 | 117M/137M [00:34<00:08, 2.40Mb/s]
85%|########5 | 117M/137M [00:34<00:08, 2.41Mb/s]
86%|########5 | 117M/137M [00:34<00:06, 2.84Mb/s]
86%|########5 | 118M/137M [00:34<00:06, 2.90Mb/s]
86%|########6 | 118M/137M [00:34<00:06, 3.07Mb/s]
86%|########6 | 118M/137M [00:34<00:06, 2.95Mb/s]
87%|########6 | 119M/137M [00:34<00:05, 3.58Mb/s]
87%|########7 | 119M/137M [00:35<00:04, 3.86Mb/s]
88%|########7 | 120M/137M [00:35<00:04, 4.17Mb/s]
88%|########7 | 120M/137M [00:35<00:04, 3.85Mb/s]
88%|########8 | 121M/137M [00:35<00:06, 2.64Mb/s]
89%|########8 | 121M/137M [00:35<00:04, 3.73Mb/s]
89%|########9 | 122M/137M [00:35<00:04, 3.49Mb/s]
89%|########9 | 122M/137M [00:35<00:04, 3.52Mb/s]
90%|########9 | 123M/137M [00:36<00:04, 3.40Mb/s]
90%|########9 | 123M/137M [00:36<00:04, 3.11Mb/s]
90%|######### | 123M/137M [00:36<00:04, 3.04Mb/s]
90%|######### | 124M/137M [00:36<00:04, 2.89Mb/s]
91%|######### | 124M/137M [00:36<00:04, 2.92Mb/s]
91%|######### | 124M/137M [00:36<00:06, 1.91Mb/s]
91%|#########1| 125M/137M [00:37<00:05, 2.35Mb/s]
91%|#########1| 125M/137M [00:37<00:04, 2.44Mb/s]
92%|#########1| 125M/137M [00:37<00:04, 2.39Mb/s]
92%|#########1| 126M/137M [00:37<00:04, 2.60Mb/s]
92%|#########2| 126M/137M [00:37<00:04, 2.52Mb/s]
92%|#########2| 127M/137M [00:37<00:03, 3.25Mb/s]
93%|#########2| 127M/137M [00:37<00:02, 3.44Mb/s]
93%|#########3| 127M/137M [00:37<00:02, 3.42Mb/s]
93%|#########3| 128M/137M [00:37<00:02, 3.51Mb/s]
```

```
 94%|#########3| 128M/137M [00:38<00:02, 3.38Mb/s]
 94%|#########3| 129M/137M [00:38<00:02, 2.99Mb/s]
 94%|#########4| 129M/137M [00:38<00:02, 3.49Mb/s]
 95%|#########4| 130M/137M [00:38<00:01, 3.92Mb/s]
 95%|#########5| 130M/137M [00:38<00:01, 4.00Mb/s]
 95%|#########5| 131M/137M [00:38<00:01, 3.13Mb/s]
 96%|#########5| 131M/137M [00:38<00:01, 2.93Mb/s]
 96%|#########5| 131M/137M [00:39<00:01, 2.76Mb/s]
 96%|#########6| 132M/137M [00:39<00:01, 2.89Mb/s]
 97%|#########6| 132M/137M [00:39<00:01, 3.40Mb/s]
 97%|#########6| 133M/137M [00:39<00:01, 2.94Mb/s]
 97%|#########7| 133M/137M [00:39<00:01, 3.05Mb/s]
 97%|#########7| 133M/137M [00:39<00:01, 2.80Mb/s]
 98%|#########7| 134M/137M [00:39<00:01, 2.68Mb/s]
 98%|#########7| 134M/137M [00:39<00:01, 2.72Mb/s]
 98%|#########8| 134M/137M [00:40<00:00, 2.73Mb/s]
 98%|#########8| 135M/137M [00:40<00:00, 2.76Mb/s]
 99%|#########8| 135M/137M [00:40<00:00, 2.82Mb/s]
 99%|#########8| 135M/137M [00:40<00:00, 2.97Mb/s]
 99%|#########9| 136M/137M [00:40<00:00, 3.18Mb/s]
 99%|#########9| 136M/137M [00:40<00:00, 3.28Mb/s]
100%|#########9| 136M/137M [00:40<00:00, 3.01Mb/s]
100%|#########9| 137M/137M [00:40<00:00, 3.31Mb/s]
100%|##########| 137M/137M [00:40<00:00, 3.35Mb/s]
[INFO] Beginning extraction
[INFO] Chromium extracted to: C:\Users\Lenovo\AppData\Local\pyppeteer\pyppeteer\local-ch
romium\588429
```

In [ ]: