

In [3]:

```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

```
# 데이터 출처 : https://www.kaggle.com/sudalairajkumar/cryptocurrencypricehistory
```

```
import pandas as pd
import glob
import os
```

```
input_file = r'D:\workspace\cakd3\programming\07\dataset\coin' # csv파일들이 있는 디렉토리 위치
output_file = r'D:\workspace\cakd3\programming\07\dataset\coin\coin_total.csv' # 병합하고 저장하려는 파일명
```

```
allFile_list = glob.glob(os.path.join(input_file, 'coin_*')) # glob함수로 sales_로 시작하는 파일들을 모은다
```

```
allData = [] # 읽어 들인 csv파일 내용을 저장할 빈 리스트를 하나 만든다
for file in allFile_list:
    df = pd.read_csv(file) # for구문으로 csv파일들을 읽어 들인다
    allData.append(df) # 빈 리스트에 읽어 들인 내용을 추가한다
```

```
dataCombine = pd.concat(allData, axis=0, ignore_index=True) # concat함수를 이용해서 리스트의 내용을 병합
```

```
# axis=0은 수직으로 병합함. axis=1은 수평.
```

```
#ignore_index=True는 인덱스 값이 기존 순서를 무시하고 순서대로 정렬되도록 한다.
```

```
dataCombine.to_csv(output_file, index=False) # to_csv함수로 저장한다. 인덱스를 빼려면 False로 설정
```

일련번호, 코인명, 상징, 날짜, 높은가치, 낮은가치, 시가, 종가, USD거래량, 코인 시가총액(미래 활용 가능성)

In [206]:

```
str = ['Aave', 'BinanceCoin', 'Bitcoin', 'Cardano', 'ChainLink', 'Cosmos', 'CryptocomCoin', 'Dogecoin',
      'EOS', 'Ethereum', 'Iota', 'Litecoin', 'Monero', 'NEM', 'Polkadot', 'Solana', 'Stellar']
df = pd.read_csv(f'dataset/coin/coin_{str[0]}.csv')
df
```

Out[206]:

	SNo	Name	Symbol	Date	High	Low	Open	Close	Volume	Market
0	1	Aave	AAVE	2020-10-05 23:59:59	55.11	49.79	52.68	53.22	0.00	89,128,12
1	2	Aave	AAVE	2020-10-06 23:59:59	53.40	40.73	53.29	42.40	583,091.46	71,011,44
2	3	Aave	AAVE	2020-10-07 23:59:59	42.41	35.97	42.40	40.08	682,834.19	67,130,03
3	4	Aave	AAVE	2020-10-08 23:59:59	44.90	36.70	39.89	43.76	1,658,816.92	220,265,14
4	5	Aave	AAVE	2020-10-09 23:59:59	47.57	43.29	43.76	46.82	815,537.66	235,632,20
...
270	271	Aave	AAVE	2021-07-02 23:59:59	234.48	215.95	230.81	233.95	256,639,522.76	3,001,542,33
271	272	Aave	AAVE	2021-07-03 23:59:59	265.11	228.49	234.21	259.66	385,377,799.45	3,331,546,05
272	273	Aave	AAVE	2021-07-04 23:59:59	289.00	248.29	259.40	277.04	427,571,943.31	3,555,054,09
273	274	Aave	AAVE	2021-07-05 23:59:59	317.39	263.43	277.11	307.83	793,140,860.06	3,950,268,56
274	275	Aave	AAVE	2021-07-06 23:59:59	346.71	308.00	308.00	316.90	988,705,452.84	4,066,775,91

275 rows × 10 columns

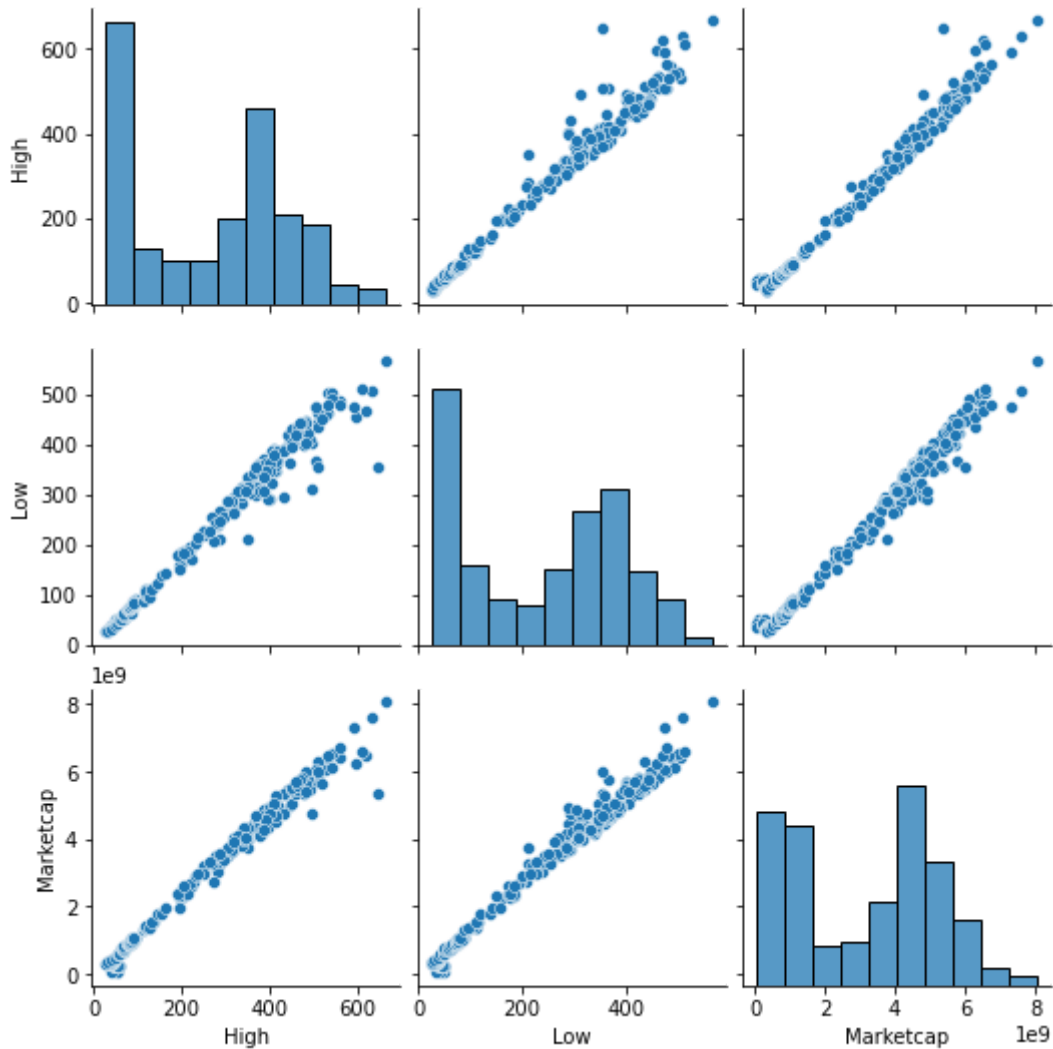


In [208]:

```
# 높고 낮은 가치와 시가총액의 상관관계  
sns.pairplot(df[['High', 'Low', 'Marketcap']])
```

Out[208]:

<seaborn.axisgrid.PairGrid at 0x2b942ae0e50>



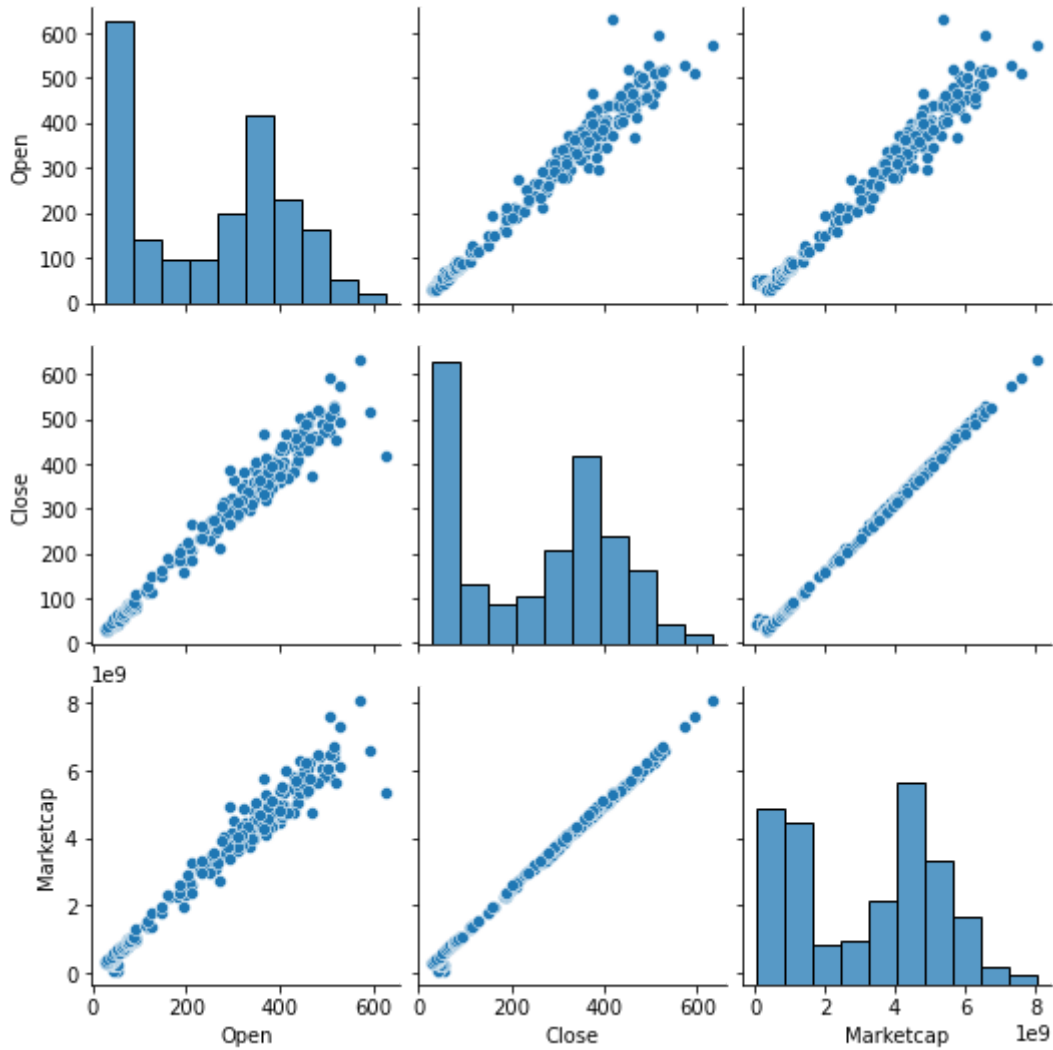
In [210]:

```
# 시가& 종가와 시가총액의 상관관계
sns.pairplot(df[['Open', 'Close', 'Marketcap']])

# 종가와 시가총액의 상관관계가 매우 높음.
# 상장종목 별로 당일 종가에 상장주식 수를 곱하여 산출한 값이 시가총액이기 때문.
```

Out[210]:

<seaborn.axisgrid.PairGrid at 0x2b944b0ff70>



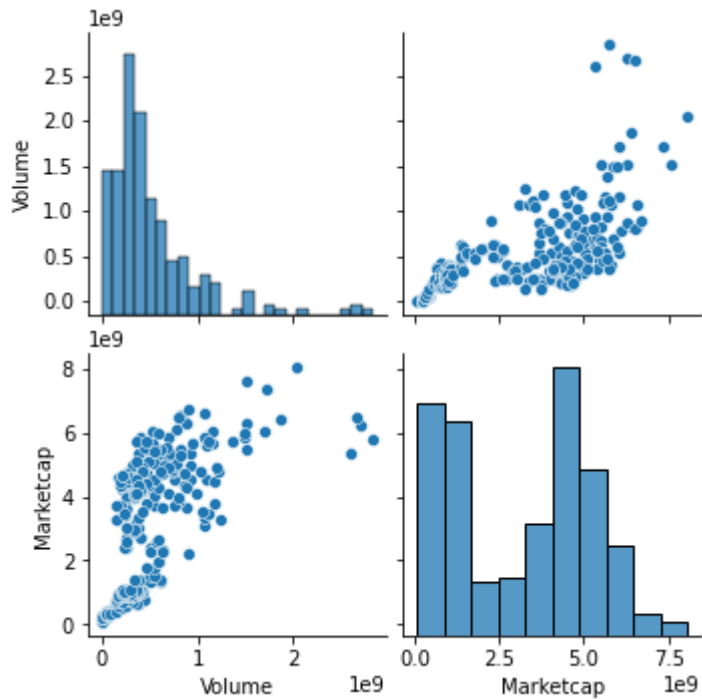
In [211]:

```
# usd 거래량과 시가총액의 상관관계
sns.pairplot(df[['Volume', 'Marketcap']])

# usd 거래량과 시가총액의 상관관계는 위의 상관관계들보다 상대적으로 낮음.
# 시가총액의 금액이 커질수록 거래량은 줄어듦.
```

Out[211]:

<seaborn.axisgrid.PairGrid at 0x2b944ab23d0>



In [202]:

```

str = ['Aave', 'BinanceCoin', 'Bitcoin', 'Cardano', 'ChainLink', 'Cosmos', 'CryptocomCoin', 'Dogecoin',
      'EOS', 'Ethereum', 'Iota', 'Litecoin', 'Monero', 'NEM', 'Polkadot', 'Solana', 'Stellar']
df = pd.read_csv(f'dataset/coin/coin_{str[0]}.csv')

df_list = []
df_list1 = []
df_list2 = []
for i in range(17):
    df = pd.read_csv(f'dataset/coin/coin_{str[i]}.csv')

    # 결측값 확인
    # df.isnull().sum() # null 값 없음
    # df.info()

    # 필요한 컬럼만 추출
    df = df.reindex(['Name', 'Marketcap'], axis=1)

    # 코인의 시가총액 값을 내림차순으로 정렬.
    df = df.sort_values(['Marketcap'], ascending=False)

    # 코인의 시가총액 값이 가장 큰 행만 추출.
    df = df.iloc[0]

    # 행과 열 바꾸기
    df = pd.DataFrame(df).transpose()

    # 각각의 리스트에 값 추가.
    df_list1.append(df.iloc[0,0])
    df_list2.append(df.iloc[0,1])

df_list.append(df_list1)
df_list.append(df_list2)

# 리스트로 출력된 값들을 데이터프레임 형식으로 다시 변환.
df_list = pd.DataFrame(df_list).transpose()
# 칼럼이름 설정
df_list = df_list.rename(columns = {0: 'Name', 1: 'Marketcap'})
# 코인 시가총액 값을 기준으로 내림차순
df_list = df_list.sort_values(['Marketcap'], ascending=False)
df_list

# 인덱스 번호 초기화
df_list = df_list.reset_index(drop = True)
df_list

```

Out[202]:

	Name	Marketcap
0	Bitcoin	1,186,364,044,140.27
1	Ethereum	482,881,900,490.93
2	Binance Coin	103,672,165,463.46
3	Dogecoin	88,680,824,303.07
4	Cardano	73,772,243,227.06
5	Polkadot	44,973,584,932.56

	Name	Marketcap
6	Litecoin	25,796,522,602.14
7	Chainlink	21,871,752,692.13
8	EOS	17,769,451,250.20
9	Stellar	16,854,691,072.30
10	NEM	16,584,479,998.20
11	Solana	15,243,420,387.39
12	IOTA	14,915,876,743.60
13	Monero	8,658,968,281.96
14	Aave	8,063,209,781.54
15	Crypto.com Coin	6,236,131,115.20
16	Cosmos	6,203,965,905.60

In [200]:

```
# <test>

str = ['Aave', 'BinanceCoin', 'Bitcoin', 'Cardano', 'ChainLink', 'Cosmos', 'CryptocomCoin', 'Dogecoin',
      'EOS', 'Ethereum', 'Iota', 'Litecoin', 'Monero', 'NEM', 'Polkadot', 'Solana', 'Stellar']
df = pd.read_csv(f'dataset/coin/coin_{str[5]}.csv')

df = df.reindex(['Name', 'Marketcap'], axis=1)
df = df.sort_values(['Marketcap'], ascending=False)
df = df.iloc[0]

# 행과 열 바꾸기
df = pd.DataFrame(df).transpose()

df
```

Out[200]:

	Name	Marketcap
785	Cosmos	6,203,965,905.60