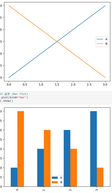
```
[In]: import numpy as np
      import matplotlib as mpl
```

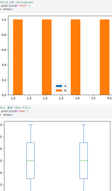
Pandas plot()

Pandas plot() 함수는 데이터프레임 또는 시리즈 객체를 이용하여는 matplotlib 기반의 Pandas Wrapper(래퍼) 시각화를 수행합니다. plot() 함수는 다양한 유형의 플롯을 생성할 수 있으며, 사용하기 쉬우며, 시각적 매력 요구를 모두 충족 하고 훌륭 하고 빠른 속도를 제공합니다.

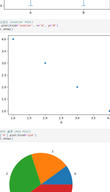
```
[In]: # 데이터프레임 생성
      df = pd.DataFrame(
      {'A': [1, 2, 3, 4, 5],
       'B': [2, 3, 4, 5, 6]}
      )
      df = df.set(drop=True)

      # 라인 플롯
      df.plot()
      plt.show()
```




```
[In]: # 막 플롯 (bar 차트)
      df.plot(kind='bar')
      plt.show()
```

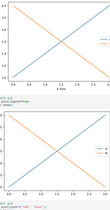



```
[In]: # 히스토그램 (histogram)
      df.plot(kind='hist')
      plt.show()
```




```
[In]: # 막대 플롯 (bar 차트)
      df.plot(kind='bar')
      plt.show()
```

```
[In]: # 산점도 (scatter 차트)
      df.plot(kind='scatter', 'A', 'B')
      plt.show()
```




```
[In]: # 파이 플롯 (Pie 차트)
      df.plot()
      plt.show()
```

추가 설정

```
[In]: # 제목 설정
      df.plot(title='My Plot')
      plt.show()
```

```
[In]: # 라벨링 설정
      df.plot(xlabel='X Axis', ylabel='Y Axis')
      plt.show()
```

```
[In]: # 그리드 설정
      df.plot(grid=True)
      plt.show()
```

Gapminder 파일 불러오기 (출처 : Kaggle)

```
[In]: !wget -q https://raw.githubusercontent.com/dsrscientist/dataset1/master/country_data.csv
      df.head(5)
```

```
[Out]:   country year population continent life_exp  gdp_cap
       0  Afghan  1952   8425333    Asia   28.801  779.4453
       1  Afghan  1957   9240934    Asia   30.332  820.8530
       2  Afghan  1962  10267083    Asia   31.997  853.1007
       3  Afghan  1967  11537966    Asia   34.020  836.2138
       4  Afghan  1972  13079460    Asia   36.088  739.9811
```

```
[In]: df.dtypes
      df.info()
```

```
[Out]: <class 'pandas.core.frame.DataFrame'>
       RangeIndex: 1704 entries, 0 to 1703
       Data columns (total 6 columns):
        #   Column     Non-Null Count  Dtype
       ---  ------     --------------  -----
        0   country    1704 non-null   object
        1   year       1704 non-null   int64
        2   population 1704 non-null   int64
        3   continent  1704 non-null   object
        4   life_exp   1704 non-null   float64
        5   gdp_cap    1704 non-null   float64
       dtypes: float64(2), int64(2), object(2)
       memory usage: 79.9687 KB
```

```
[In]: # country 컬럼 유니크 리스트 보기
      df['country'].unique()
```

```
[Out]: array(['Afghan', 'Albania', 'Algeria', 'Angola', 'Argentina',
              'Australia', 'Austria', 'Bahrain', 'Bangladesh', 'Belgium',
              'Benin', 'Bolivia', 'Botswana', 'Brazil', 'Bulgaria',
              'Burkina F', 'Burma', 'Burundi', 'CAMBODIA', 'Cameroon',
              'Canada', 'Central', 'Chad', 'Chile', 'China', 'Colombia',
              'Congo(Dem.', 'Congo(Rep.', 'Costa Rica', 'Croatia', 'Cuba',
              'Czech Rep.', 'Denmark', 'Dominican', 'Ecuador', 'Egypt',
              'Eq.Guinea', 'Ethiopia', 'Finland', 'France', 'Gabon',
              'Gambia', 'Germany', 'Ghana', 'Greece', 'Guatemala', 'Guinea',
              'Haiti', 'Honduras', 'Hungary', 'India', 'Indonesia',
              'Iran', 'Iraq', 'Ireland', 'Israel', 'Italy', 'Ivory Coa.',
              'Jamaica', 'Japan', 'Jordan', 'Kenya', 'Korea, Dem.',
              'Korea, Rep.', 'Kuwait', 'Lebanon', 'Lesotho', 'Liberia',
              'Libya', 'Madagascar', 'Malawi', 'Malaysia', 'Mali', 'Mauritania',
              'Mauritius', 'Mexico', 'Mongolia', 'Morocco', 'Mozambique',
              'Nepal', 'Netherlands', 'New Zealand', 'Nicaragua', 'Niger',
              'Nigeria', 'Norway', 'Oman', 'Pakistan', 'Panama', 'Paraguay',
              'Peru', 'Philippines', 'Poland', 'Portugal', 'Puerto Ric',
              'Reunion', 'Romania', 'Rwanda', 'Sao Tome and', 'Saudi Arabia',
              'Senegal', 'Sierra Leone', 'Singapore', 'Slovakia', 'Slovenia',
              'Somalia', 'South Africa', 'Spain', 'Sri Lanka', 'Sudan',
              'Swaziland', 'Sweden', 'Switzerland', 'Syria', 'Taiwan',
              'Tanzania', 'Thailand', 'Togo', 'Trinidad and Tobago', 'Tunisia',
              'Turkey', 'Uganda', 'United Kingd', 'United States', 'Uruguay',
              'Venezuela', 'Vietnam', 'West Bank and Gaza', 'Yemen, Rep.',
              'Zambia', 'ZIMBABWE', 'Bruguadirio'],
       dtype=object)
```

특정 나라의 시간에 따른 기대 수명 변화 시각화

```
[In]: # 한국 데이터만 따로 분리
      df_korea = df[df['country'] == 'Korea, Rep.']

      # 시각화
      plt.scatter(df_korea['year'], df_korea['life_exp'], color='lightblue')
      plt.title('Life Expectancy in South Korea Over Years', color='orange')
      plt.xlabel('Year')
      plt.ylabel('Life Expectancy')
      plt.grid(True)
      plt.show()
```

2007년 대륙별 평균 GDP 시각화

```
[In]: # 데이터 준비
      df_2007 = df[df['year'] == 2007]

      # 대륙별 평균 GDP 계산
      continent_gdp = df_2007.groupby('continent')['gdp_cap'].mean()

      # 막대 차트 그리기
      plt.bar(continent_gdp.index, continent_gdp.values, color='green')
      plt.title('Average GDP per Capita by Continent (2007)', color='green')
      plt.xlabel('Continent')
      plt.ylabel('GDP per Capita')
      plt.xticks(rotation=45)
      plt.tight_layout()
      plt.show()
```

나라별 기대 수명과 GDP의 관계 시각화

```
[In]: # 데이터 준비
      df_2007 = df[df['year'] == 2007]

      # 산점도 그리기
      plt.scatter(df_2007['gdp_cap'], df_2007['life_exp'], color='lightblue')
      plt.title('GDP per Capita vs Life Expectancy (2007)', color='orange')
      plt.xlabel('GDP per Capita')
      plt.ylabel('Life Expectancy')
      plt.grid(True)
      plt.show()
```


Pandas 리뷰

IMDB Movie rating 정확 평점 및 관련된 데이터이터를 포함하고 있습니다.

```
[In]: IMDB Movie rating 데이터를 불러오고 데이터를 보기 합니다.
      !wget -q https://raw.githubusercontent.com/dsrscientist/dataset3/master/imdb_movies.csv
      df.head(5)
```

```
[Out]:   Rank                Movie name  Year Certificate  Runtime in min  ...  Genre  Metascore  Gross in $ K M  Rating from IMDb
       0    1  The Shawshank Redemption  1994          R             142  ...  Drama       80.0           28.34              9.3
       1    2                 The Godfather  1972          R             175  ...  Crime, Drama  100.0          134.97              9.2
       2    3                The Godfather  1974         UA             202  ...  Crime, Drama   90.0          57.30              9.0
       3    4              The Dark Knight  2008      PG-13             152  ...  Action, Adventure, Drama  84.0  534.86              9.0
       4    5  The Lord of the Rings: The Return of the King  2003      PG-13             201  ...  Action, Adventure, Drama  94.0  377.85              8.9

       [5 rows x 10 columns]
```

```
[In]: df.info()
      df.describe
      df.describe(include='object')
      df.describe(include='bool')
      df.describe(include='all')
      df.shape
      df.columns
      df.T
      df.sort_values(by='Rating from IMDb', ascending=False)
      df.sort_values(by=['Rating from IMDb', 'Runtime in min'], ascending=[False, True])
      df.sort_values(by=['Rating from IMDb', 'Runtime in min'], ascending=[False, True], inplace=True)
      df.sort_values(by=['Rating from IMDb', 'Runtime in min'], ascending=[False, True], inplace=True, ignore_index=True)
      df.head(5)
      df.tail(5)
      df.sample(5)
      df.sample(5, frac=0.5)
      df.sample(5, frac=0.5, replace=True)
      df.sample(5, frac=0.5, replace=True, random_state=1)
      df.nunique()
      df.nunique(axis=1)
      df.isnull()
      df.isnull().sum()
      df.isnull().sum().sum()
      df.notnull()
      df.notnull().sum()
      df.notnull().sum().sum()
      df.dropna()
      df.dropna(axis=1)
      df.dropna(axis=1, how='all')
      df.dropna(axis=1, how='any')
      df.dropna(axis=1, how='any', inplace=True)
      df.dropna(axis=1, how='any', inplace=True, ignore_index=True)
      df.fillna(0)
      df.fillna(0, inplace=True)
      df.fillna(method='ffill')
      df.fillna(method='bfill')
      df.fillna(method='ffill', limit=2)
      df.fillna(method='bfill', limit=2)
      df.fillna(method='ffill', limit=2, inplace=True)
      df.fillna(method='bfill', limit=2, inplace=True)
      df.fillna(method='ffill', limit=2, inplace=True, ignore_index=True)
      df.fillna(method='bfill', limit=2, inplace=True, ignore_index=True)
      df.fillna(method='ffill', limit=2, inplace=True, ignore_index=True, axis=1)
      df.fillna(method='bfill', limit=2, inplace=True, ignore_index=True, axis=1)
      df.fillna(method='ffill', limit=2, inplace=True, ignore_index=True, axis=1, inplace=True)
      df.fillna(method='bfill', limit=2, inplace=True, ignore_index=True, axis=1, inplace=True)
      df.fillna(method='ffill', limit=2, inplace=True, ignore_index=True, axis=1, inplace=True, ignore_index=True)
      df.fillna(method='bfill', limit=2, inplace=True, ignore_index=True, axis=1, inplace=True, ignore_index=True)
      df.fillna(method='ffill', limit=2, inplace=True, ignore_index=True, axis=1, inplace=True, ignore_index=True, inplace=True)
      df.fillna(method='bfill', limit=2, inplace=True, ignore_index=True, axis=1, inplace=True, ignore_index=True, inplace=True)
      df.fillna(method='ffill', limit=2, inplace=True, ignore_index=True, axis=1, inplace=True, ignore_index=True, inplace=True, ignore_index=True)
      df.fillna(method='bfill', limit=2, inplace=True, ignore_index=True, axis=1, inplace=True, ignore_index=True, inplace=True, ignore_index=True)
      df.fillna(method='ffill', limit=2, inplace=True, ignore_index=True, axis=1, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True)
      df.fillna(method='bfill', limit=2, inplace=True, ignore_index=True, axis=1, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True)
      df.fillna(method='ffill', limit=2, inplace=True, ignore_index=True, axis=1, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True)
      df.fillna(method='bfill', limit=2, inplace=True, ignore_index=True, axis=1, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True)
      df.fillna(method='ffill', limit=2, inplace=True, ignore_index=True, axis=1, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True)
      df.fillna(method='bfill', limit=2, inplace=True, ignore_index=True, axis=1, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True)
      df.fillna(method='ffill', limit=2, inplace=True, ignore_index=True, axis=1, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True)
      df.fillna(method='bfill', limit=2, inplace=True, ignore_index=True, axis=1, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True)
      df.fillna(method='ffill', limit=2, inplace=True, ignore_index=True, axis=1, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True)
      df.fillna(method='bfill', limit=2, inplace=True, ignore_index=True, axis=1, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True)
      df.fillna(method='ffill', limit=2, inplace=True, ignore_index=True, axis=1, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True)
      df.fillna(method='bfill', limit=2, inplace=True, ignore_index=True, axis=1, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True)
      df.fillna(method='ffill', limit=2, inplace=True, ignore_index=True, axis=1, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True)
      df.fillna(method='bfill', limit=2, inplace=True, ignore_index=True, axis=1, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True)
      df.fillna(method='ffill', limit=2, inplace=True, ignore_index=True, axis=1, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True)
      df.fillna(method='bfill', limit=2, inplace=True, ignore_index=True, axis=1, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True)
      df.fillna(method='ffill', limit=2, inplace=True, ignore_index=True, axis=1, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True)
      df.fillna(method='bfill', limit=2, inplace=True, ignore_index=True, axis=1, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True)
      df.fillna(method='ffill', limit=2, inplace=True, ignore_index=True, axis=1, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True)
      df.fillna(method='bfill', limit=2, inplace=True, ignore_index=True, axis=1, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True)
      df.fillna(method='ffill', limit=2, inplace=True, ignore_index=True, axis=1, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True)
      df.fillna(method='bfill', limit=2, inplace=True, ignore_index=True, axis=1, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True)
      df.fillna(method='ffill', limit=2, inplace=True, ignore_index=True, axis=1, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True)
      df.fillna(method='bfill', limit=2, inplace=True, ignore_index=True, axis=1, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True)
      df.fillna(method='ffill', limit=2, inplace=True, ignore_index=True, axis=1, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True)
      df.fillna(method='bfill', limit=2, inplace=True, ignore_index=True, axis=1, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True)
      df.fillna(method='ffill', limit=2, inplace=True, ignore_index=True, axis=1, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True)
      df.fillna(method='bfill', limit=2, inplace=True, ignore_index=True, axis=1, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True)
      df.fillna(method='ffill', limit=2, inplace=True, ignore_index=True, axis=1, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True)
      df.fillna(method='bfill', limit=2, inplace=True, ignore_index=True, axis=1, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True)
      df.fillna(method='ffill', limit=2, inplace=True, ignore_index=True, axis=1, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True)
      df.fillna(method='bfill', limit=2, inplace=True, ignore_index=True, axis=1, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True)
      df.fillna(method='ffill', limit=2, inplace=True, ignore_index=True, axis=1, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True)
      df.fillna(method='bfill', limit=2, inplace=True, ignore_index=True, axis=1, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True)
      df.fillna(method='ffill', limit=2, inplace=True, ignore_index=True, axis=1, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True)
      df.fillna(method='bfill', limit=2, inplace=True, ignore_index=True, axis=1, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True, inplace=True, ignore_index=True
```