Unemployment Data Analysis in USA

October 31, 2022

```python
[73]: import numpy as np
      import pandas as pd
      import matplotlib.pyplot as plt
      import seaborn as sns
      %matplotlib inline
```

```python
[75]: unemp=pd.read_csv("C:\\Users\\demar\\Desktop\\Rahel's
      ↪docs\\unemployment_rate_data.csv")
```

```python
[80]: unemp.head()
```

```
[80]:         date  unrate  unrate_men  unrate_women  unrate_16_to_17  \
      0  1/1/1948     4.0         4.2           3.5             10.8
      1  2/1/1948     4.7         4.7           4.8             15.0
      2  3/1/1948     4.5         4.5           4.4             13.2
      3  4/1/1948     4.0         4.0           4.1              9.9
      4  5/1/1948     3.4         3.3           3.4              6.4

         unrate_18_to_19  unrate_20_to_24  unrate_25_to_34  unrate_35_to_44  \
      0              9.6              6.6              3.6              2.6
      1              9.5              8.0              4.0              3.2
      2              9.3              8.6              3.5              3.2
      3              8.1              6.8              3.5              3.1
      4              7.2              6.3              2.8              2.5

         unrate_45_to_54  unrate_55_over
      0              2.7             3.6
      1              3.4             4.0
      2              2.9             3.5
      3              2.9             3.2
      4              2.3             2.9
```

```python
[81]: unemp.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 887 entries, 0 to 886
Data columns (total 11 columns):
 #   Column              Non-Null Count  Dtype
```

```
 ---   ------              --------------   -----
  0    date               887 non-null     object
  1    unrate             887 non-null     float64
  2    unrate_men         887 non-null     float64
  3    unrate_women       887 non-null     float64
  4    unrate_16_to_17    887 non-null     float64
  5    unrate_18_to_19    887 non-null     float64
  6    unrate_20_to_24    887 non-null     float64
  7    unrate_25_to_34    887 non-null     float64
  8    unrate_35_to_44    887 non-null     float64
  9    unrate_45_to_54    887 non-null     float64
  10   unrate_55_over     887 non-null     float64
dtypes: float64(10), object(1)
memory usage: 76.4+ KB
```

[82]: `unemp.describe()`

[82]:

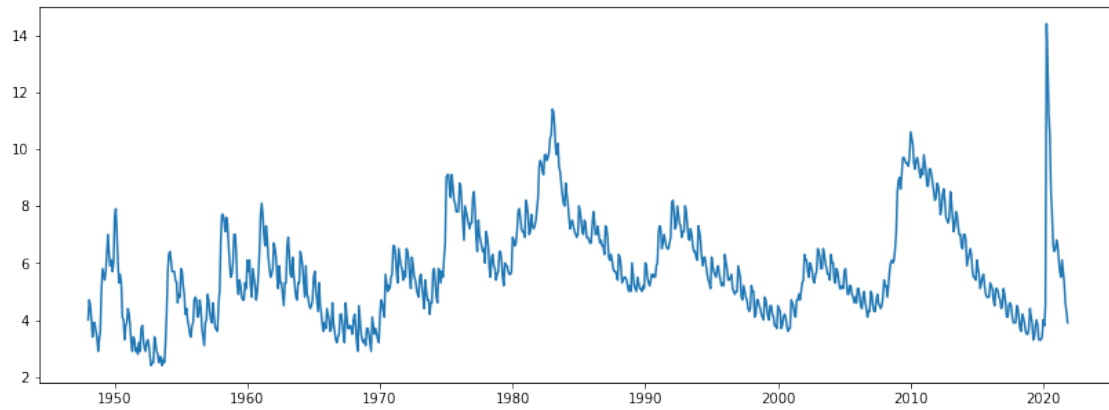|	unrate	unrate_men	unrate_women	unrate_16_to_17	unrate_18_to_19 \
count	887.000000	887.000000	887.000000	887.000000	887.000000
mean	5.763134	5.633709	6.028749	17.943517	14.824803
std	1.740101	1.954639	1.608252	5.018894	4.047867
min	2.400000	1.900000	2.600000	5.700000	5.200000
25%	4.500000	4.300000	4.900000	14.700000	12.300000
50%	5.500000	5.300000	5.800000	17.800000	14.600000
75%	6.800000	6.700000	7.000000	20.900000	17.000000
max	14.400000	13.300000	15.700000	35.800000	33.300000

	unrate_20_to_24	unrate_25_to_34	unrate_35_to_44	unrate_45_to_54 \
count	887.000000	887.000000	887.000000	887.000000
mean	9.345660	5.532582	4.242954	3.867193
std	2.800988	1.923600	1.443626	1.352247
min	3.300000	2.000000	1.600000	1.500000
25%	7.700000	4.200000	3.200000	2.950000
50%	9.100000	5.200000	4.000000	3.600000
75%	10.800000	6.700000	5.000000	4.500000
max	25.000000	14.300000	11.300000	12.100000

	unrate_55_over
count	887.000000
mean	3.838782
std	1.241579
min	1.500000
25%	3.000000
50%	3.600000
75%	4.400000
max	13.400000

```
[83]: unemp['date']=pd.to_datetime(unemp['date'])
      unemp.describe()
```

[83]:
	unrate	unrate_men	unrate_women	unrate_16_to_17	unrate_18_to_19 \
count	887.000000	887.000000	887.000000	887.000000	887.000000
mean	5.763134	5.633709	6.028749	17.943517	14.824803
std	1.740101	1.954639	1.608252	5.018894	4.047867
min	2.400000	1.900000	2.600000	5.700000	5.200000
25%	4.500000	4.300000	4.900000	14.700000	12.300000
50%	5.500000	5.300000	5.800000	17.800000	14.600000
75%	6.800000	6.700000	7.000000	20.900000	17.000000
max	14.400000	13.300000	15.700000	35.800000	33.300000

	unrate_20_to_24	unrate_25_to_34	unrate_35_to_44	unrate_45_to_54 \
count	887.000000	887.000000	887.000000	887.000000
mean	9.345660	5.532582	4.242954	3.867193
std	2.800988	1.923600	1.443626	1.352247
min	3.300000	2.000000	1.600000	1.500000
25%	7.700000	4.200000	3.200000	2.950000
50%	9.100000	5.200000	4.000000	3.600000
75%	10.800000	6.700000	5.000000	4.500000
max	25.000000	14.300000	11.300000	12.100000

	unrate_55_over
count	887.000000
mean	3.838782
std	1.241579
min	1.500000
25%	3.000000
50%	3.600000
75%	4.400000
max	13.400000

```
[84]: # unemployment rate

      plt.figure(figsize=(14,5))
      plt.plot(unemp['date'],unemp['unrate'])
```
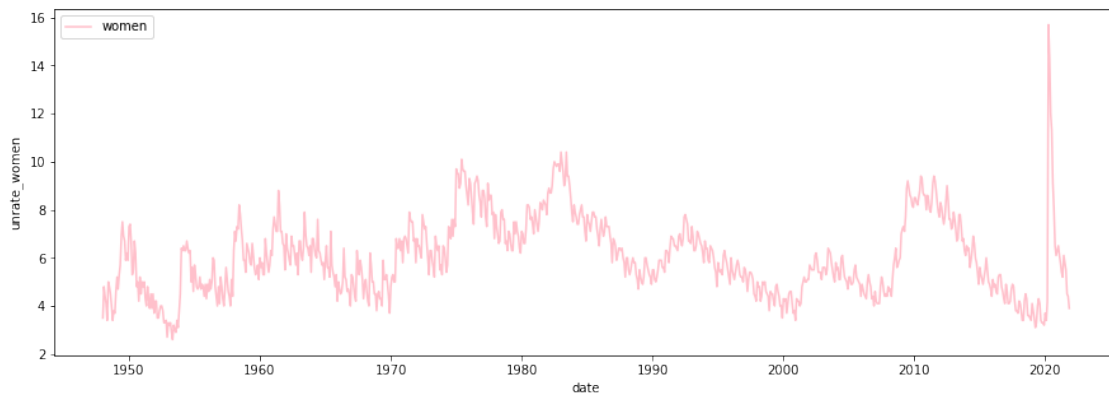
[84]: [<matplotlib.lines.Line2D at 0x19aaebaaf70>]

```
[43]: #Women unemployement rate
      plt.figure(figsize=(15,5))
      sns.lineplot(data=unemp,x='date',y='unrate_women',color='pink',label='women')
```

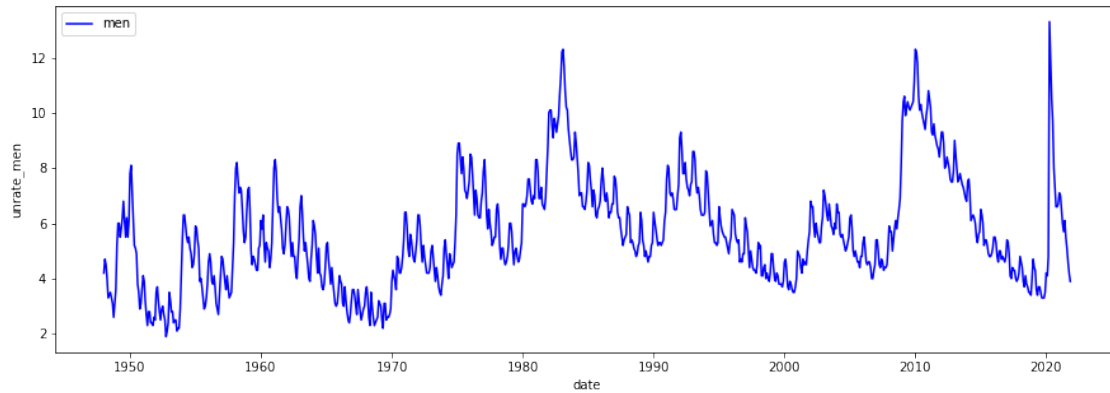[43]: <AxesSubplot:xlabel='date', ylabel='unrate_women'>


```
[85]: #Men unemployment rate

      plt.figure(figsize=(15,5))
      sns.lineplot(data=unemp,x='date',y='unrate_men',color='blue',label='men')
```

[85]: <AxesSubplot:xlabel='date', ylabel='unrate_men'>

```
[31]: #unemployment between different age group.

      plt.figure(figsize=(15,5))
      sns.
        ↪lineplot(data=unemp,x='date',y='unrate_16_to_17',color='brown',label='16-17')
      sns.
        ↪lineplot(data=unemp,x='date',y='unrate_18_to_19',color='Green',label='18-19')
      sns.
        ↪lineplot(data=unemp,x='date',y='unrate_20_to_24',color='yellow',label='20-24')
      sns.lineplot(data=unemp,x='date',y='unrate_35_to_44',color='red',label='35-44')
      sns.
        ↪lineplot(data=unemp,x='date',y='unrate_45_to_54',color='purple',label='45-54')
      sns.lineplot(data=unemp,x='date',y='unrate_55_over',color='blue',label='over␣
        ↪55')
```

[31]: <AxesSubplot:xlabel='date', ylabel='unrate_16_to_17'>
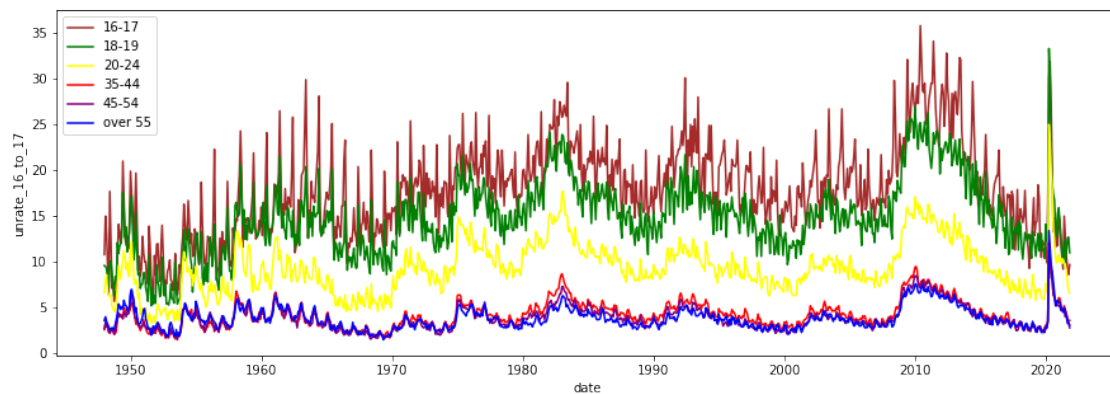
[35]:
```
#USA recession in 1973 -1975 that caused by middle east crisis

plt.figure(figsize=(15,5))
sns.lineplot(data=unemp[(unemp.date >='01-01-1973')&(unemp.date␣
 ↪<='01-01-1975')],x='date',y='unrate')
```
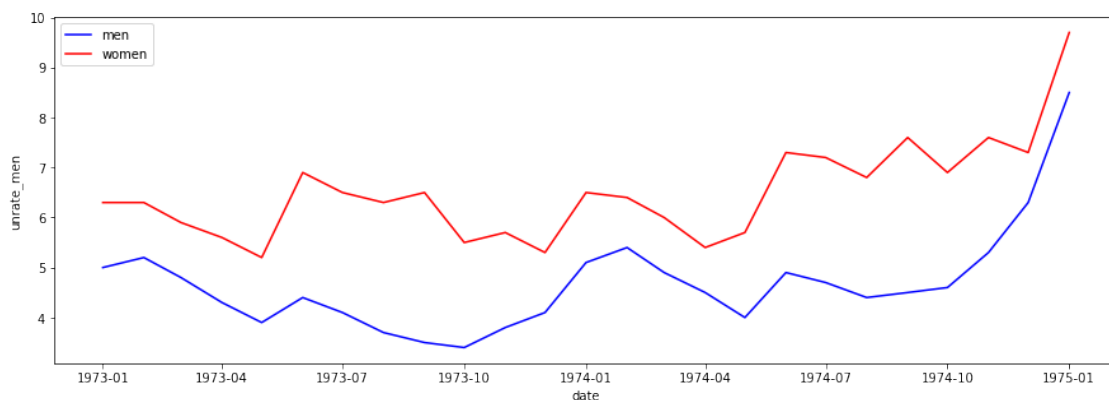
[35]: <AxesSubplot:xlabel='date', ylabel='unrate'>
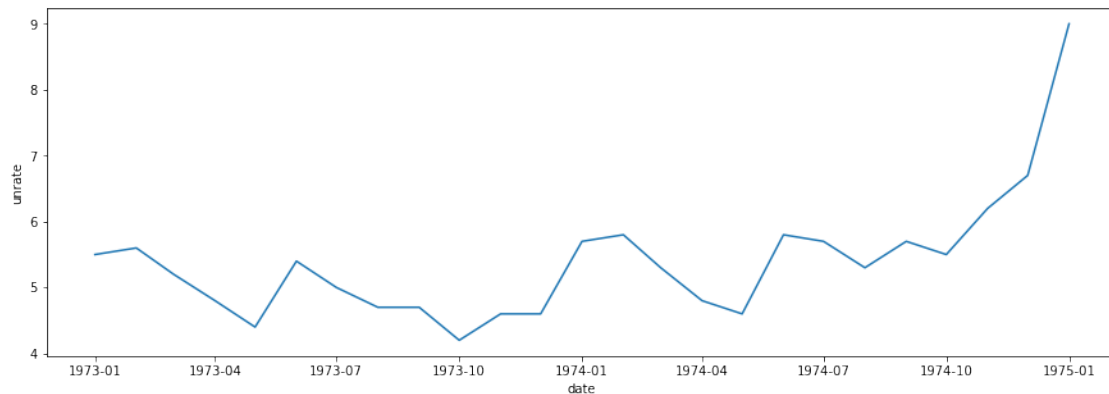
[36]:
```
#USA unemployement rate between Women and men from 1973-1975
plt.figure(figsize=(15,5))
sns.lineplot(data=unemp[(unemp.date >='01-01-1973')&(unemp.date␣
 ↪<='01-01-1975')],x='date',y='unrate_men',color='blue',label='men')
sns.lineplot(data=unemp[(unemp.date >='01-01-1973')&(unemp.date␣
 ↪<='01-01-1975')],x='date',y='unrate_women',color='red',label='women')
```

[36]: <AxesSubplot:xlabel='date', ylabel='unrate_men'>

```
[46]: # USA unemployement rate between different age groups in 1973-1975.

      plt.figure(figsize=(15,5))
      plt.figure(figsize=(15,5))
      sns.lineplot(data=unemp[(unemp.date >='01-01-1973')&(unemp.date␣
       ↪<='01-01-1975')],x='date',y='unrate_16_to_17',color='blue',label='16-17')
      sns.lineplot(data=unemp[(unemp.date >='01-01-1973')&(unemp.date␣
       ↪<='01-01-1975')],x='date',y='unrate_18_to_19',color='red',label='18-19')
      sns.lineplot(data=unemp[(unemp.date >='01-01-1973')&(unemp.date␣
       ↪<='01-01-1975')],x='date',y='unrate_20_to_24',color='black',label='20-24')
      sns.lineplot(data=unemp[(unemp.date >='01-01-1973')&(unemp.date␣
       ↪<='01-01-1975')],x='date',y='unrate_35_to_44',color='purple',label='35-44')
      sns.lineplot(data=unemp[(unemp.date >='01-01-1973')&(unemp.date␣
       ↪<='01-01-1975')],x='date',y='unrate_45_to_54',color='green',label='45-54')
      sns.lineplot(data=unemp[(unemp.date >='01-01-1973')&(unemp.date␣
       ↪<='01-01-1975')],x='date',y='unrate_55_over',color='yellow',label='over 55')
```

[46]: <AxesSubplot:xlabel='date', ylabel='unrate_16_to_17'>

<Figure size 1080x360 with 0 Axes>
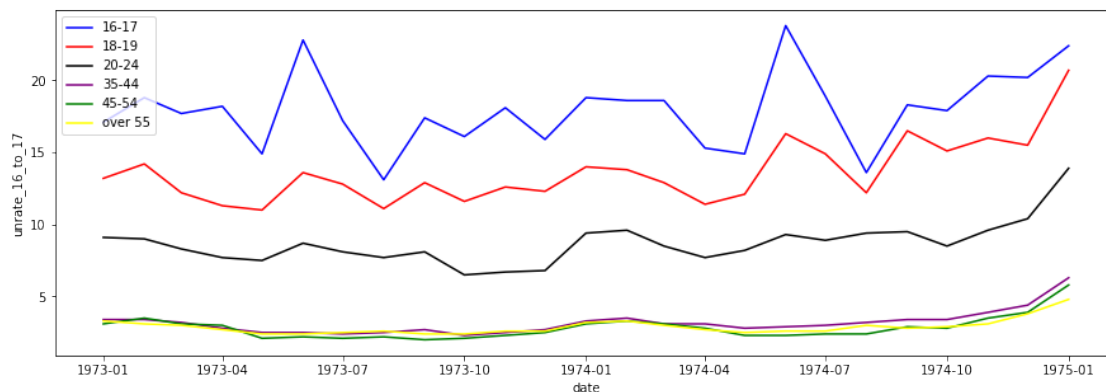

```
[86]: #unemployement rate during early 90's recession.
      recession=unemp[(unemp.date >='01-01-1986')&(unemp.date <='01-01-1990')]
      recession.head()
```

[86]:   date unrate unrate_men unrate_women unrate_16_to_17 \
      456 1986-01-01   7.3   7.5   7.2   21.2
      457 1986-02-01   7.8   8.0   7.6   22.5
      458 1986-03-01   7.5   7.6   7.2   20.9
      459 1986-04-01   7.0   7.0   6.9   21.3
      460 1986-05-01   7.0   6.8   7.1   20.6

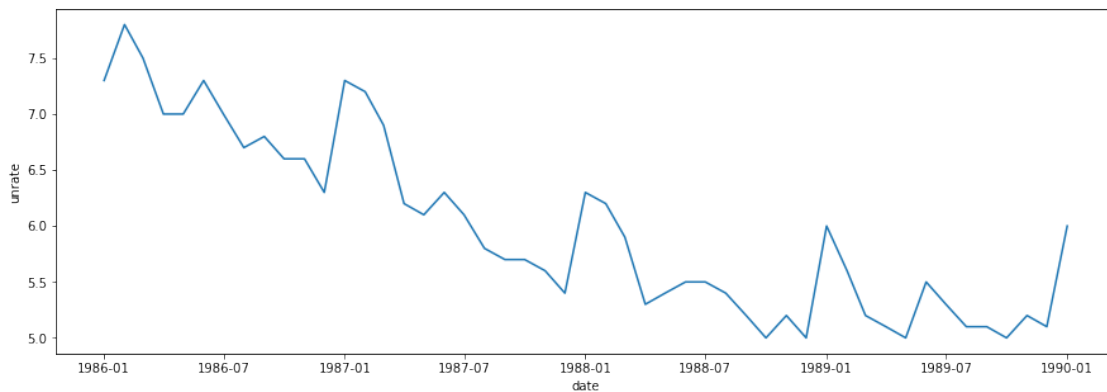   unrate_18_to_19 unrate_20_to_24 unrate_25_to_34 unrate_35_to_44 \

	unrate_45_to_54	unrate_55_over
456	17.8	11.6	7.1	5.2
457	18.5	11.9	7.8	5.8
458	17.4	11.1	7.4	5.6
459	16.9	10.6	6.9	5.1
460	17.3	11.6	6.7	4.9

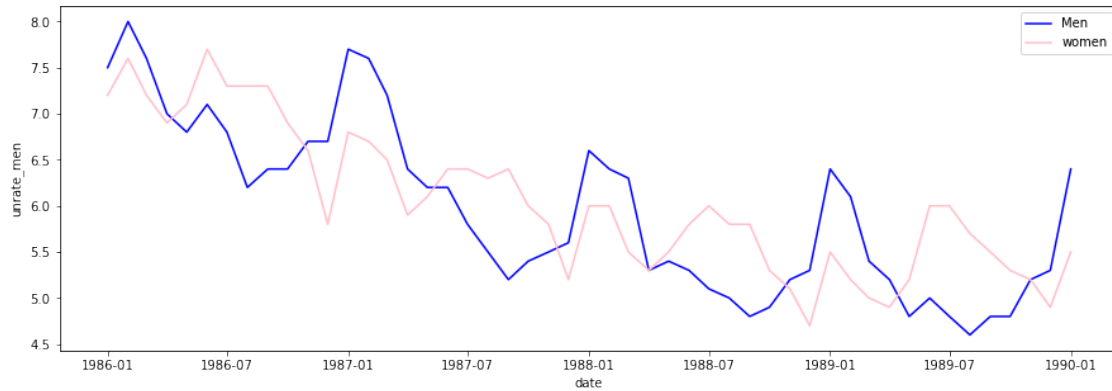|	unrate_45_to_54	unrate_55_over
456	5.2	4.3
457	4.8	4.8
458	4.9	4.7
459	4.4	4.0
460	4.4	3.4

```
[65]: plt.figure(figsize=(15,5))
      sns.lineplot(data=recession,x='date',y='unrate')
```

```
[65]: <AxesSubplot:xlabel='date', ylabel='unrate'>
```



```
[66]: plt.figure(figsize=(15,5))
      sns.lineplot(data=recession,x='date',y='unrate_men',label='Men',color='blue')
      sns.
        ↪lineplot(data=recession,x='date',y='unrate_women',label='women',color='pink')
```

```
[66]: <AxesSubplot:xlabel='date', ylabel='unrate_men'>
```

```
[67]: #USA unemployment rate between age 16-55 during early 90's recession.

      plt.figure(figsize=(15,5))
      sns.
       ↪lineplot(data=recession,x='date',y='unrate_16_to_17',color='brown',label='16-17')
      sns.
       ↪lineplot(data=recession,x='date',y='unrate_18_to_19',color='Green',label='18-19')
      sns.
       ↪lineplot(data=recession,x='date',y='unrate_20_to_24',color='yellow',label='20-24')
      sns.
       ↪lineplot(data=recession,x='date',y='unrate_35_to_44',color='red',label='35-44')
      sns.
       ↪lineplot(data=recession,x='date',y='unrate_45_to_54',color='purple',label='45-54')
      sns.
       ↪lineplot(data=recession,x='date',y='unrate_55_over',color='blue',label='over␣
       ↪55')
```

[67]: <AxesSubplot:xlabel='date', ylabel='unrate_16_to_17'>
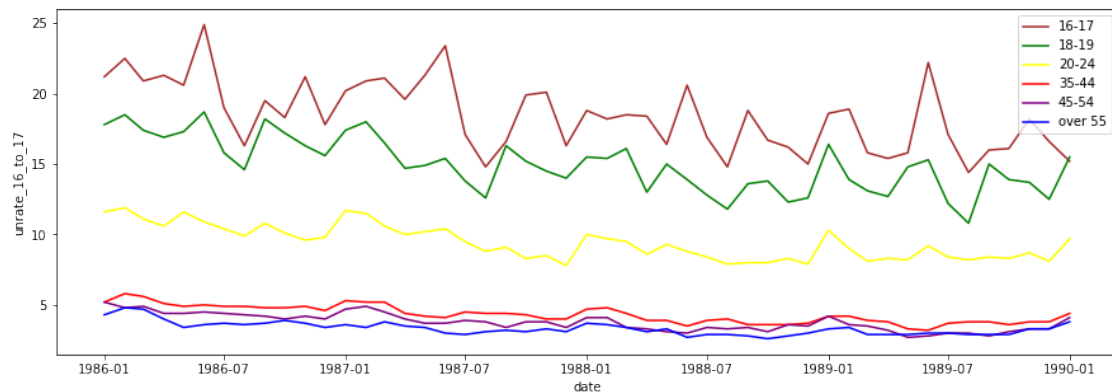
```
[87]: #USA unemployemet rate during Grater recession from 2007 to 2010

      recession=unemp[(unemp.date >='01-01-2007')&(unemp.date <='01-01-2010')]
      recession.head()
```

```
[87]:          date  unrate  unrate_men  unrate_women  unrate_16_to_17  \
      708 2007-01-01     5.0         5.4           4.6             15.9
      709 2007-02-01     4.9         5.4           4.2             17.0
      710 2007-03-01     4.5         5.0           4.1             16.6
      711 2007-04-01     4.3         4.5           4.1             16.2
      712 2007-05-01     4.3         4.4           4.1             17.8

           unrate_18_to_19  unrate_20_to_24  unrate_25_to_34  unrate_35_to_44  \
      708             14.6              9.0              5.4              3.8
      709             14.7              8.0              5.3              3.6
      710             13.3              7.6              4.7              3.5
      711             14.5              7.3              4.3              3.3
      712             15.4              7.4              4.2              3.0

           unrate_45_to_54  unrate_55_over
      708              3.4             3.7
      709              3.6             3.3
      710              3.3             3.3
      711              3.1             2.9
      712              2.9             2.9
```

```
[70]: recession.describe()
```

```
[70]:             unrate  unrate_men  unrate_women  unrate_16_to_17  unrate_18_to_19  \
      count  37.000000   37.000000     37.000000        37.000000        37.000000
      mean    6.662162    7.178378      6.056757        21.759459        18.467568
      std     2.134030    2.577050      1.675732         4.494222         4.343197
      min     4.300000    4.300000      4.100000        15.900000        12.500000
      25%     4.800000    5.000000      4.600000        17.800000        14.600000
      50%     6.000000    5.900000      5.700000        21.300000        17.500000
      75%     9.000000   10.200000      7.300000        24.200000        22.700000
      max    10.600000   12.300000      9.200000        32.100000        26.900000

             unrate_20_to_24  unrate_25_to_34  unrate_35_to_44  unrate_45_to_54  \
      count        37.000000        37.000000        37.000000        37.000000
      mean         11.186486         6.902703         5.437838         4.913514
      std           3.022155         2.387408         2.076101         1.853250
      min           7.300000         4.200000         3.000000         2.900000
      25%           8.400000         4.800000         3.700000         3.400000
      50%          10.500000         5.700000         4.700000         3.900000
      75%          13.900000         9.800000         7.500000         7.000000
```
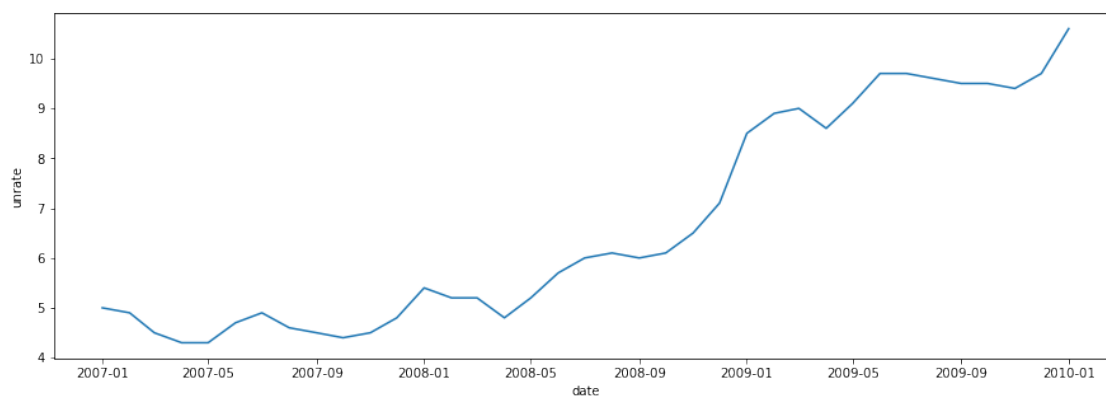
```
max       17.100000        11.000000        9.400000        8.500000

       unrate_55_over
count       37.000000
mean         4.589189
std          1.624087
min          2.900000
25%          3.200000
50%          3.900000
75%          6.300000
max          7.500000
```

[71]:
```python
#USA unemployment rate during grater recession(2007-2010.
plt.figure(figsize=(15,5))
sns.lineplot(data=recession[(recession.date>='01-01-2007')&(recession.date␣
  ↪<='01-01-2010')],x='date',y='unrate')
```
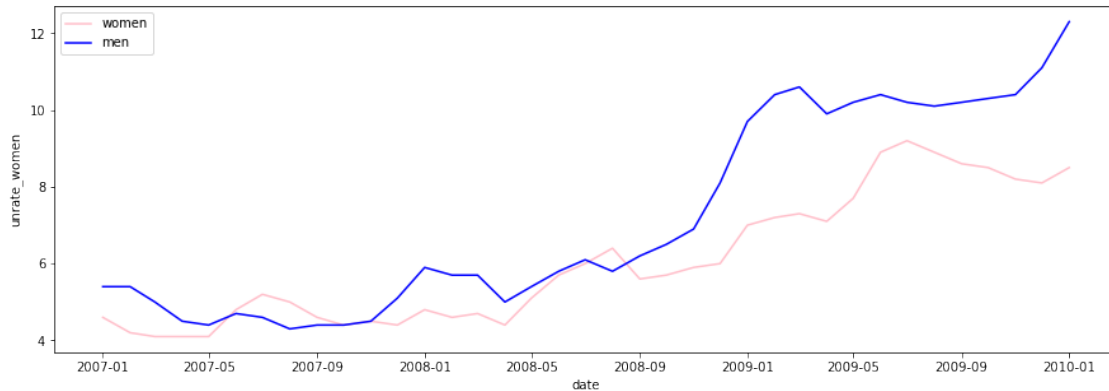
[71]: <AxesSubplot:xlabel='date', ylabel='unrate'>

[78]:
```python
# Women and Men uneployment rate during grate recession(2007-2009)

plt.figure(figsize=(15,5))
sns.
  ↪lineplot(data=recession,x='date',y='unrate_women',label='women',color='pink')
sns.lineplot(data=recession,x='date',y='unrate_men',label='men',color='blue')
```

[78]: <AxesSubplot:xlabel='date', ylabel='unrate_women'>

```
[72]:  #Unemployment rate from age 16 to over 55 during grate recession(2007-2010)

       plt.figure(figsize=(15,5))
       sns.
        ↪lineplot(data=recession,x='date',y='unrate_16_to_17',color='brown',label='16-17')
       sns.
        ↪lineplot(data=recession,x='date',y='unrate_18_to_19',color='Green',label='18-19')
       sns.
        ↪lineplot(data=recession,x='date',y='unrate_20_to_24',color='yellow',label='20-24')
       sns.
        ↪lineplot(data=recession,x='date',y='unrate_35_to_44',color='red',label='35-44')
       sns.
        ↪lineplot(data=recession,x='date',y='unrate_45_to_54',color='purple',label='45-54')
       sns.
        ↪lineplot(data=recession,x='date',y='unrate_55_over',color='blue',label='over␣
        ↪55')
```

[72]: <AxesSubplot:xlabel='date', ylabel='unrate_16_to_17'>
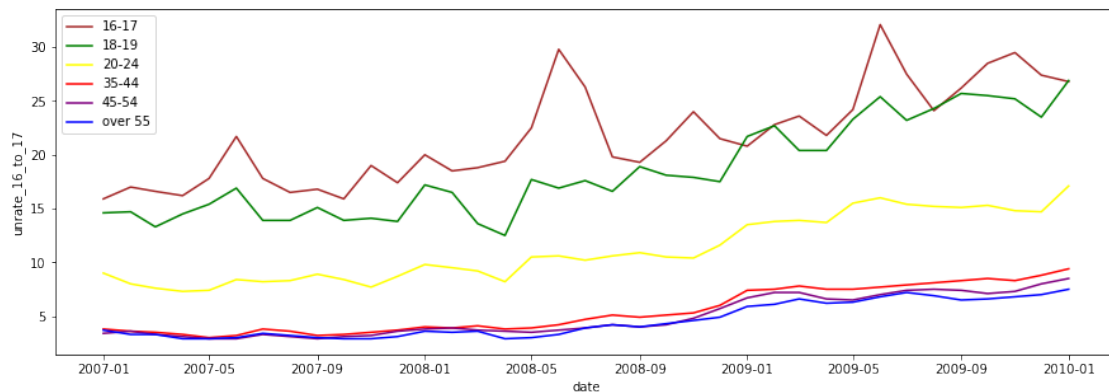
```
[105]:  # Unemployment rate in USA during covid_19 outbreak.

        covid_19=unemp[(unemp.date >= '01-01-2019')&(unemp.date <='01-01-2021')]
        covid_19.head()
```

```
[105]:          date  unrate  unrate_men  unrate_women  unrate_16_to_17  \
        852 2019-01-01     4.4         4.7           4.1             14.6
        853 2019-02-01     4.1         4.4           3.8             17.0
        854 2019-03-01     3.9         4.3           3.5             12.9
        855 2019-04-01     3.3         3.6           3.1             14.3
        856 2019-05-01     3.4         3.4           3.3             14.4

             unrate_18_to_19  unrate_20_to_24  unrate_25_to_34  unrate_35_to_44  \
        852             13.3              8.6              4.4              3.3
        853             12.5              7.5              4.4              2.9
        854             11.8              7.1              4.1              3.0
        855             10.6              5.8              3.6              2.5
        856             12.0              7.0              3.2              2.4

             unrate_45_to_54  unrate_55_over
        852              3.3             3.5
        853              3.1             2.9
        854              3.1             2.8
        855              2.4             2.4
        856              2.3             2.5
```

```
[106]:  covid_19.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 25 entries, 852 to 876
Data columns (total 11 columns):
 #   Column           Non-Null Count  Dtype
---  ------           --------------  -----
 0   date             25 non-null     datetime64[ns]
 1   unrate           25 non-null     float64
 2   unrate_men       25 non-null     float64
 3   unrate_women     25 non-null     float64
 4   unrate_16_to_17  25 non-null     float64
 5   unrate_18_to_19  25 non-null     float64
 6   unrate_20_to_24  25 non-null     float64
 7   unrate_25_to_34  25 non-null     float64
 8   unrate_35_to_44  25 non-null     float64
 9   unrate_45_to_54  25 non-null     float64
 10  unrate_55_over   25 non-null     float64
dtypes: datetime64[ns](1), float64(10)
memory usage: 2.3 KB
```

```
[109]: covid_19.describe()
```

```
[109]:           unrate   unrate_men  unrate_women  unrate_16_to_17  unrate_18_to_19  \
       count   25.000000   25.000000     25.000000        25.000000        25.000000
       mean     5.916000    5.820000      6.020000        15.764000        15.156000
       std      3.261222    2.902872      3.689173         5.845716         5.756596
       min      3.300000    3.300000      3.100000         9.900000        10.400000
       25%      3.800000    3.600000      3.400000        11.900000        12.000000
       50%      4.100000    4.400000      4.100000        14.300000        13.100000
       75%      6.800000    7.100000      6.500000        17.000000        15.900000
       max     14.400000   13.300000     15.700000        31.900000        33.300000

              unrate_20_to_24  unrate_25_to_34  unrate_35_to_44  unrate_45_to_54  \
       count        25.000000         25.00000        25.000000         25.00000
       mean         10.324000          6.14000         4.640000          4.56800
       std           5.554974          3.40869         2.597274          2.70797
       min           5.800000          3.20000         2.400000          2.30000
       25%           6.800000          3.60000         2.800000          2.70000
       50%           7.600000          4.40000         3.100000          3.10000
       75%          10.800000          7.20000         6.100000          5.70000
       max          25.000000         14.30000        11.300000         12.10000

              unrate_55_over
       count       25.000000
       mean         4.864000
       std          3.162789
       min          2.300000
       25%          2.700000
       50%          2.900000
       75%          5.800000
       max         13.400000
```
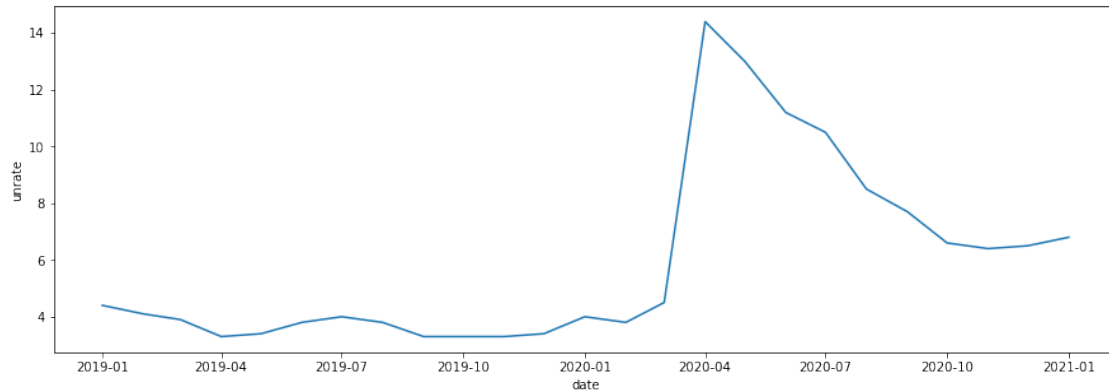
```
[113]: #Over all covid_19 unemployment rate from (01-01-2019 to 01-01-2021)
       plt.figure(figsize=(15,5))
       sns.lineplot(data=covid_19[(covid_19.date >='01-01-2019')&(covid_19.
         ↪date<='01-01-2021')],x='date',y='unrate')
```
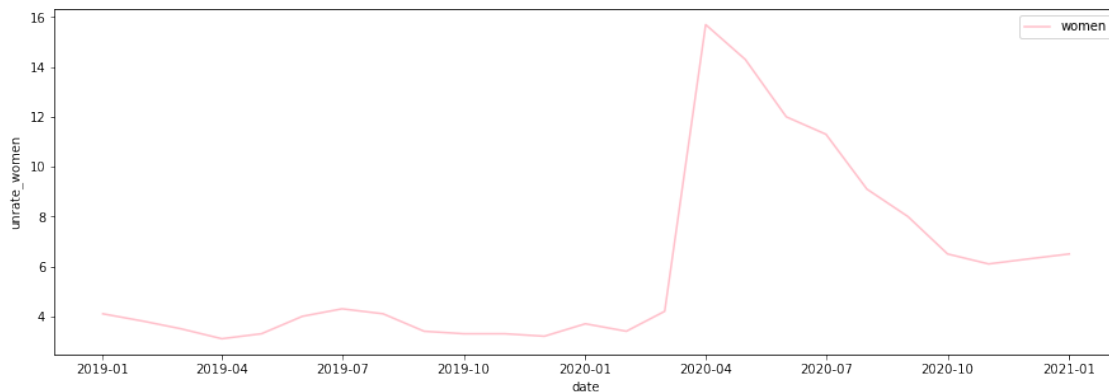
```
[113]: <AxesSubplot:xlabel='date', ylabel='unrate'>
```

[114]:
```
# Covid_19 impact on women's unemployment rate.

plt.figure(figsize=(15,5))
sns.lineplot(data=covid_19,x='date',y='unrate_women',color='pink',label='women')
```

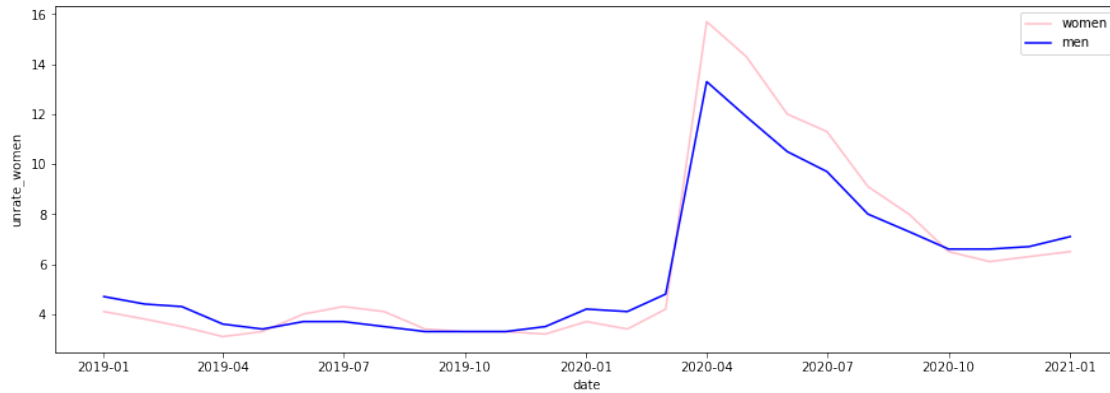[114]: `<AxesSubplot:xlabel='date', ylabel='unrate_women'>`

[115]:
```
#USA unemployment rate during covid_19 outbreak between different age group.
plt.figure(figsize=(15,5))
sns.lineplot(data=covid_19,x='date',y='unrate_women',label='women',color='pink')
sns.lineplot(data=covid_19,x='date',y='unrate_men',label='men',color='blue')
```
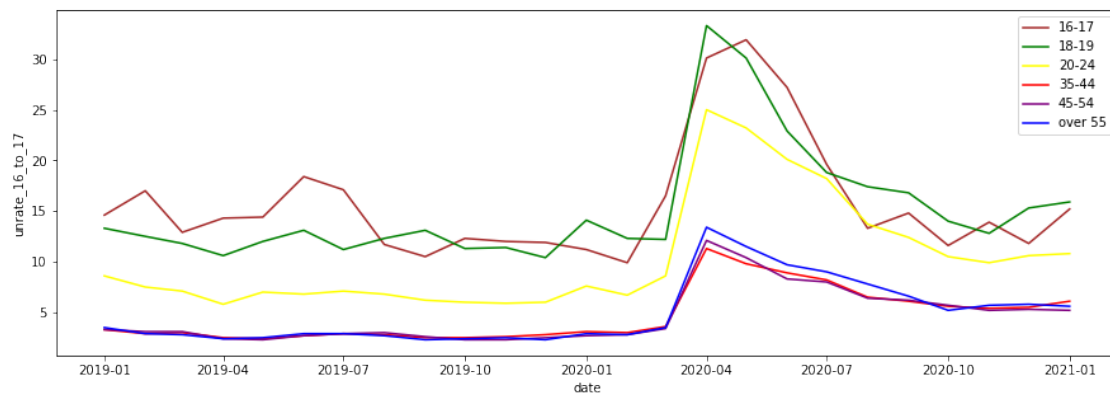
[115]: `<AxesSubplot:xlabel='date', ylabel='unrate_women'>`

15

```
[116]: plt.figure(figsize=(15,5))
       sns.
         ↪lineplot(data=covid_19,x='date',y='unrate_16_to_17',color='brown',label='16-17')
       sns.
         ↪lineplot(data=covid_19,x='date',y='unrate_18_to_19',color='Green',label='18-19')
       sns.
         ↪lineplot(data=covid_19,x='date',y='unrate_20_to_24',color='yellow',label='20-24')
       sns.
         ↪lineplot(data=covid_19,x='date',y='unrate_35_to_44',color='red',label='35-44')
       sns.
         ↪lineplot(data=covid_19,x='date',y='unrate_45_to_54',color='purple',label='45-54')
       sns.lineplot(data=covid_19,x='date',y='unrate_55_over',color='blue',label='over␣
         ↪55')
```

[116]: <AxesSubplot:xlabel='date', ylabel='unrate_16_to_17'>

[ ]: