In [2]:

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
%matplot inline
from matplotlib.pyploat import rcParams
rcParams['figure.figsize'] = 10, 6
```

```
UsageError: Line magic function `%matplot` not found.
```

In [7]:

```
dataset = pd.read_csv("C:/Users/jay/Desktop/AirPassengers.csv")
```

In [9]:

```
dataset.head()
```

Out[9]:

	Month	#Passengers
0	1949-01	112
1	1949-02	118
2	1949-03	132
3	1949-04	129
4	1949-05	121

In [10]:

```
# Parse strings to datatime type
```

In [12]:

```
dataset['Month'] = pd.to_datetime(dataset['Month'], infer_datetime_format=True)
```

In [16]:

```
indexedDataset = dataset.set_index(['Month'])
```

In [17]:

```
from datetime import datetime
indexedDataset.head(4)
```

Out[17]:

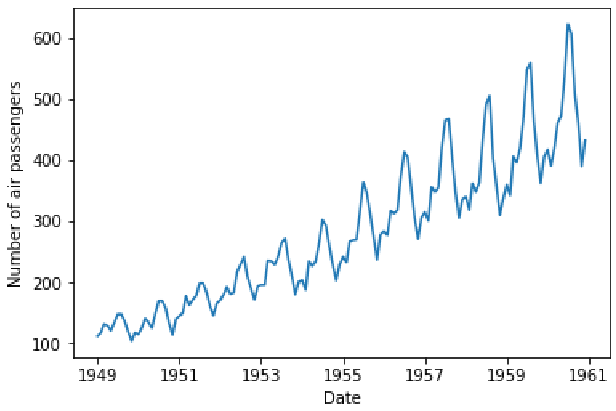
	#Passengers
Month	
1949-01-01	112
1949-02-01	118
1949-03-01	132
1949-04-01	129

In [19]:

```
## plot graph
plt.xlabel("Date")
plt.ylabel("Number of air passengers")
plt.plot(indexedDataset)
```

Out[19]:

```
[<matplotlib.lines.Line2D at 0x8d8b748>]
```



In [37]:

```
#Determing rolling statistics
rolmean = indexedDataset.rolling(window=12).mean()
rolstd = indexedDataset.rolling(window=12).std()
print(rolmean, rolstd)
```

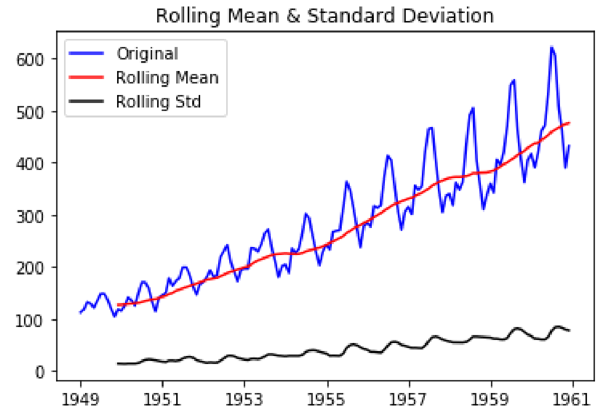
```
            #Passengers
Month
1949-01-01          NaN
1949-02-01          NaN
1949-03-01          NaN
1949-04-01          NaN
1949-05-01          NaN
1949-06-01          NaN
1949-07-01          NaN
1949-08-01          NaN
1949-09-01          NaN
1949-10-01          NaN
1949-11-01          NaN
1949-12-01   126.666667
1950-01-01   126.916667
1950-02-01   127.583333
1950-03-01   128.333333
1950-04-01   128.833333
1950-05-01   129.166667
1950-06-01   130.333333
1950-07-01   132.166667
1950-08-01   134.000000
1950-09-01   135.833333
1950-10-01   137.000000
1950-11-01   137.833333
1950-12-01   139.666667
1951-01-01   142.166667
1951-02-01   144.166667
1951-03-01   147.250000
1951-04-01   149.583333
1951-05-01   153.500000
1951-06-01   155.916667
...                 ...
1958-07-01   376.333333
1958-08-01   379.500000
1958-09-01   379.500000
1958-10-01   380.500000
1958-11-01   380.916667
1958-12-01   381.000000
1959-01-01   382.666667
1959-02-01   384.666667
1959-03-01   388.333333
1959-04-01   392.333333
1959-05-01   397.083333
```

```
1959-06-01   400.166667
1959-07-01   404.916667
1959-08-01   409.416667
1959-09-01   414.333333
1959-10-01   418.333333
1959-11-01   422.666667
1959-12-01   428.333333
1960-01-01   433.083333
1960-02-01   437.166667
1960-03-01   438.250000
1960-04-01   443.666667
1960-05-01   448.000000
1960-06-01   453.250000
1960-07-01   459.416667
1960-08-01   463.333333
1960-09-01   467.083333
1960-10-01   471.583333
1960-11-01   473.916667
1960-12-01   476.166667

[144 rows x 1 columns]             #Passengers
Month
1949-01-01          NaN
1949-02-01          NaN
1949-03-01          NaN
1949-04-01          NaN
1949-05-01          NaN
1949-06-01          NaN
1949-07-01          NaN
1949-08-01          NaN
1949-09-01          NaN
1949-10-01          NaN
1949-11-01          NaN
1949-12-01    13.720147
1950-01-01    13.453342
1950-02-01    13.166475
1950-03-01    13.686977
1950-04-01    13.822467
1950-05-01    13.663710
1950-06-01    14.760718
1950-07-01    18.135016
1950-08-01    20.797727
1950-09-01    21.928949
1950-10-01    21.315807
1950-11-01    20.067311
1950-12-01    19.070841
1951-01-01    17.439940
1951-02-01    16.781122
1951-03-01    19.349066
1951-04-01    19.425655
1951-05-01    18.744696
1951-06-01    19.942911
...                 ...
1958-07-01    59.590013
1958-08-01    65.557054
1958-09-01    65.557054
1958-10-01    65.106207
1958-11-01    64.593074
1958-12-01    64.530472
1959-01-01    63.627229
1959-02-01    61.759553
1959-03-01    61.597422
1959-04-01    60.284678
1959-05-01    60.008270
1959-06-01    63.009138
1959-07-01    71.987951
1959-08-01    80.049369
1959-09-01    81.485451
1959-10-01    79.680422
1959-11-01    74.498729
1959-12-01    69.830097
1960-01-01    66.624399
1960-02-01    61.866180
1960-03-01    61.382741
1960-04-01    60.171472
1960-05-01    60.184565
1960-06-01    65.021849
```

```
1960-07-01    77.194510
1960-08-01    83.630500
1960-09-01    84.617276
1960-10-01    82.541954
1960-11-01    79.502382
1960-12-01    77.737125

[144 rows x 1 columns]
```

In [42]:

```
#Plot rolling statistics:
orig = plt.plot(indexedDataset, color='blue',label='Original')
mean = plt.plot(rolmean, color= 'red', label= 'Rolling Mean')
std = plt.plot(rolstd, color='black', label = 'Rolling Std')
plt.legend(loc='best')
plt.title('Rolling Mean & Standard Deviation')
plt.show(block=False)
```



In [53]:

```
#Perform Dickey-Fuller test:
from statsmodels.tsa.stattools import adfuller
print('Results of Dickey-Fuller Test:')
dftest = adfuller(indexedDataset['#Passengers'], autolag='AIC')
```

```
Results of Dickey-Fuller Test:
```

In [65]:

```
dfoutput = pd.Series(dftest[0:4], index=['Test Statistic','p-value','#Lags Used','Number of Observations Used'])
for key,value in dftest[4].items():
            dfoutput['Critical Value (%s)'%key] = value
print(dfoutput)
```

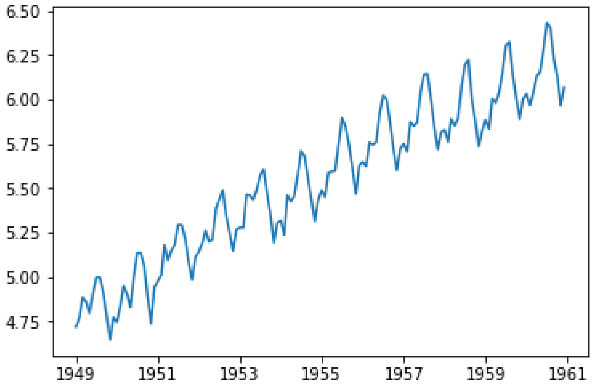
```
Test Statistic                   0.815369
p-value                          0.991880
#Lags Used                      13.000000
Number of Observations Used    130.000000
Critical Value (1%)             -3.481682
Critical Value (5%)             -2.884042
Critical Value (10%)            -2.578770
dtype: float64
```

In [67]:

```
#Estimating trend
indexedDataset_logScale = np.log(indexedDataset)
plt.plot(indexedDataset_logScale)
```

Out[67]:

```
[<matplotlib.lines.Line2D at 0xae3d198>]
```

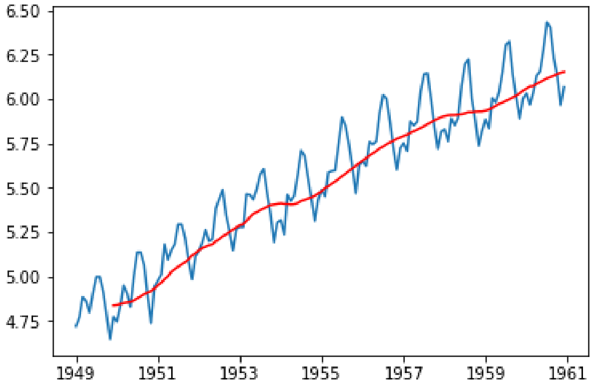


In [68]:

```
movingAverage = indexedDataset_logScale.rolling(window=12).mean()
movingSTD = indexedDataset_logScale.rolling(window=12).std()
plt.plot(indexedDataset_logScale)
plt.plot(movingAverage, color='red')
```

Out[68]:

```
[<matplotlib.lines.Line2D at 0xaeb9cc0>]
```

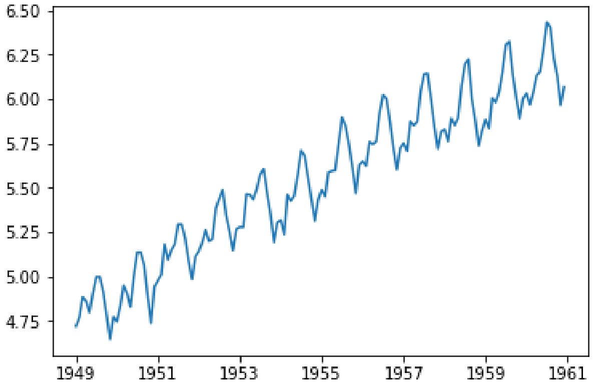


In [70]:

```
# Estimating treand
indexedDatset_logScale = np.log(indexedDataset)
plt.plot(indexedDataset_logScale)
```

Out[70]:

```
[<matplotlib.lines.Line2D at 0xd3f7278>]
```



In [9]:

```
import numpy as np
import itertools
import warnings
import matplotlib.pyplot as plt
warnings.filterwarnings("ignore")
plt.style.use('fivethirtyeight')
import statsmodels.api as sm
import pandas as pd
```

In [2]:

```
  File "<ipython-input-2-8e3189d1a776>", line 1
    pip install nbconvent
              ^
SyntaxError: invalid syntax
```

In [ ]: