In [1]:
```
import pandas as pd
import matplotlib.pyplot as plt
data = pd.read_csv('E:\Files\AirPassengers.csv')
```

In [5]:
```
data['date'][0]
```

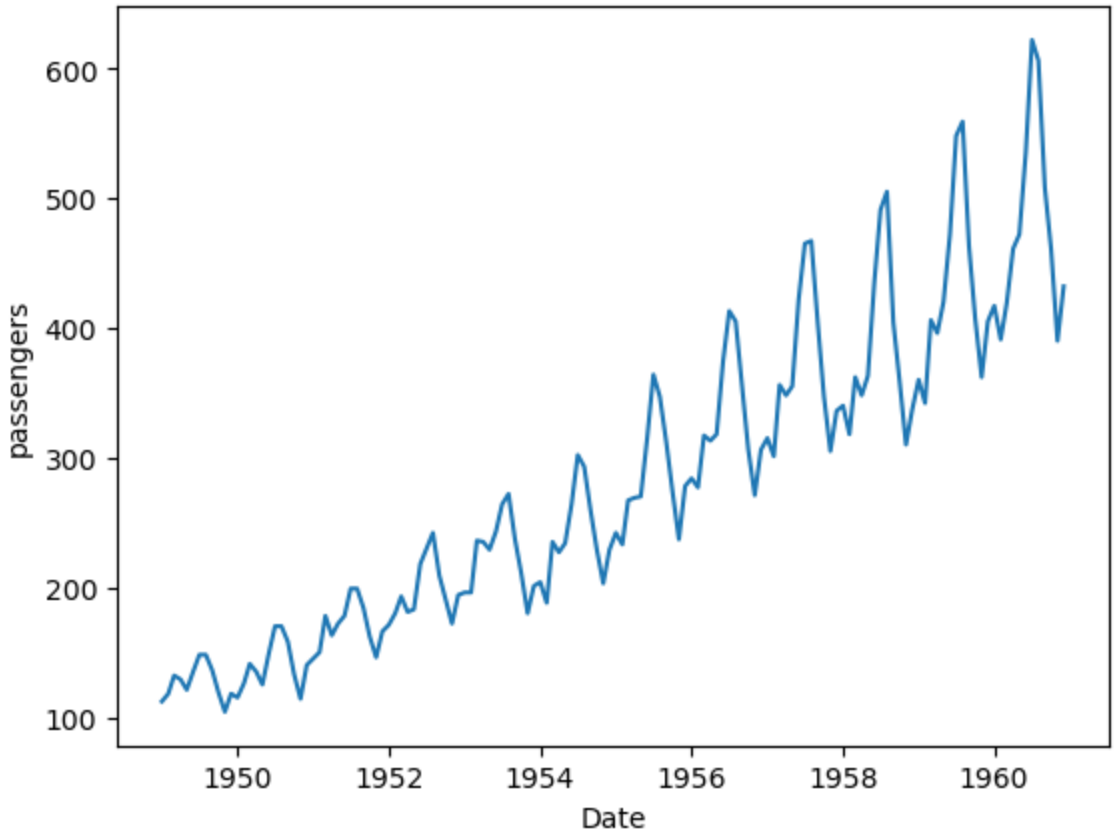
Out[5]:
```
Timestamp('1949-01-01 00:00:00')
```

In [4]:
```
data['date'] = pd.to_datetime(data['date'])
```

In [7]:
```
ts_data = data.set_index(['date'])
```

In [9]:
```
plt.xlabel("Date")
plt.ylabel("passengers")
plt.plot(ts_data)
```

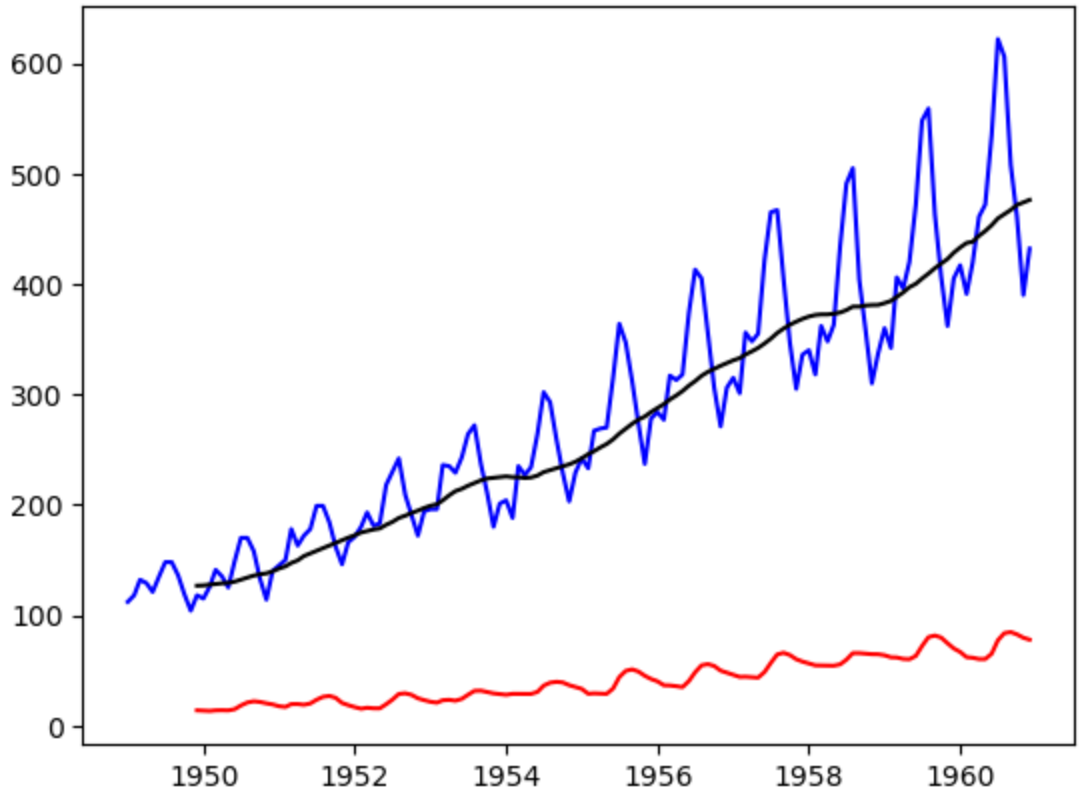
Out[9]:
```
[<matplotlib.lines.Line2D at 0x1d56aab4580>]
```



In [10]:
```
rol_mean = ts_data.rolling(window=12).mean()
rol_std = ts_data.rolling(window=12).std()

plt.plot(ts_data,c='blue')
plt.plot(rol_mean,c='black')
plt.plot(rol_std,c='red')
```

Out[10]:
```
[<matplotlib.lines.Line2D at 0x1d56c42b0a0>]
```



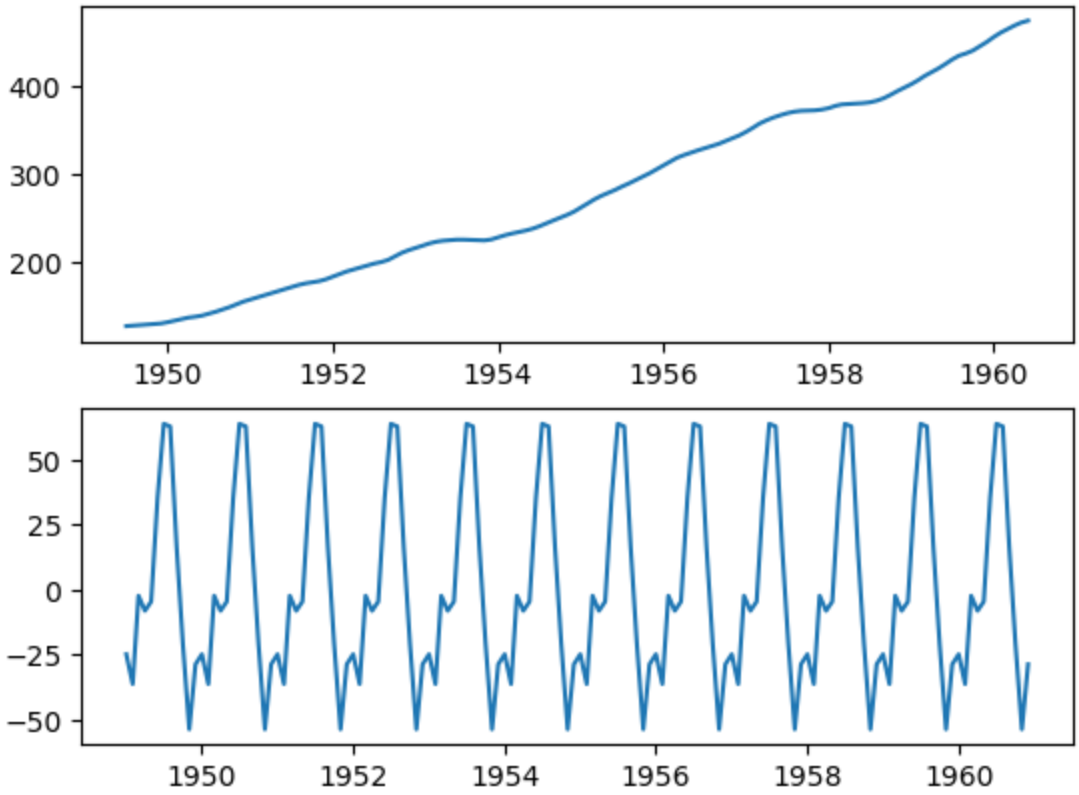
In [11]:
```
from statsmodels.tsa.stattools import adfuller
dftest = adfuller(ts_data['value'])
print(dftest)
```

```
(0.8153688792060543, 0.9918802434376411, 13, 130, {'1%': -3.4816817173418295, '5%': -2.8
840418343195267, '10%': -2.578770059171598}, 996.692930839019)
```

In [12]:
```
from statsmodels.tsa.seasonal import seasonal_decompose
decompose = seasonal_decompose(ts_data)
trend = decompose.trend
season = decompose.seasonal
```

In [13]:
```
plt.subplot(211)
plt.plot(trend)
plt.subplot(212)
plt.plot(season)
```

Out[13]:
```
[<matplotlib.lines.Line2D at 0x1d570053670>]
```



differencing

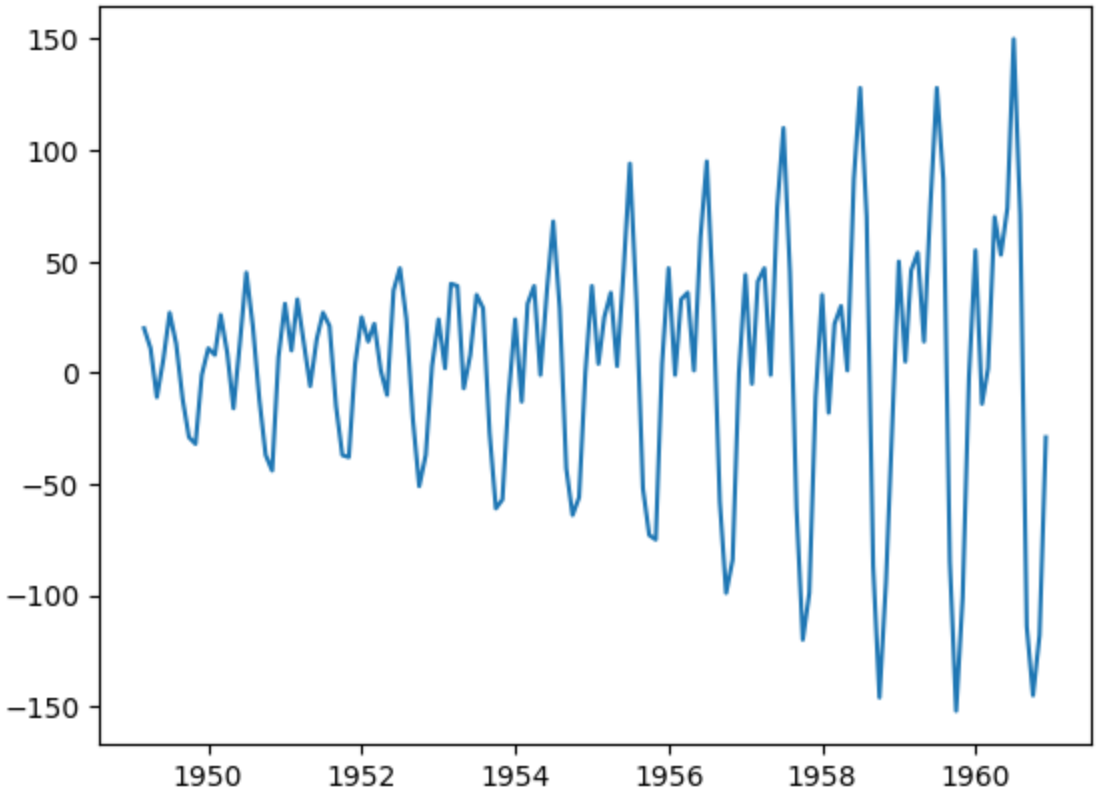
In [17]:
```
ts_data_new = ts_data['value'] - ts_data['value'].shift(2)
```

In [18]:
```
adfuller(ts_data_new.dropna())
```

Out[18]:
```
(-2.9616951355554337,
 0.03862975767698741,
 11,
 130,
 {'1%': -3.4816817173418295,
  '5%': -2.8840418343195267,
  '10%': -2.578770059171598},
 985.7309222414162)
```

In [19]:
```
plt.plot(ts_data_new.dropna())
```

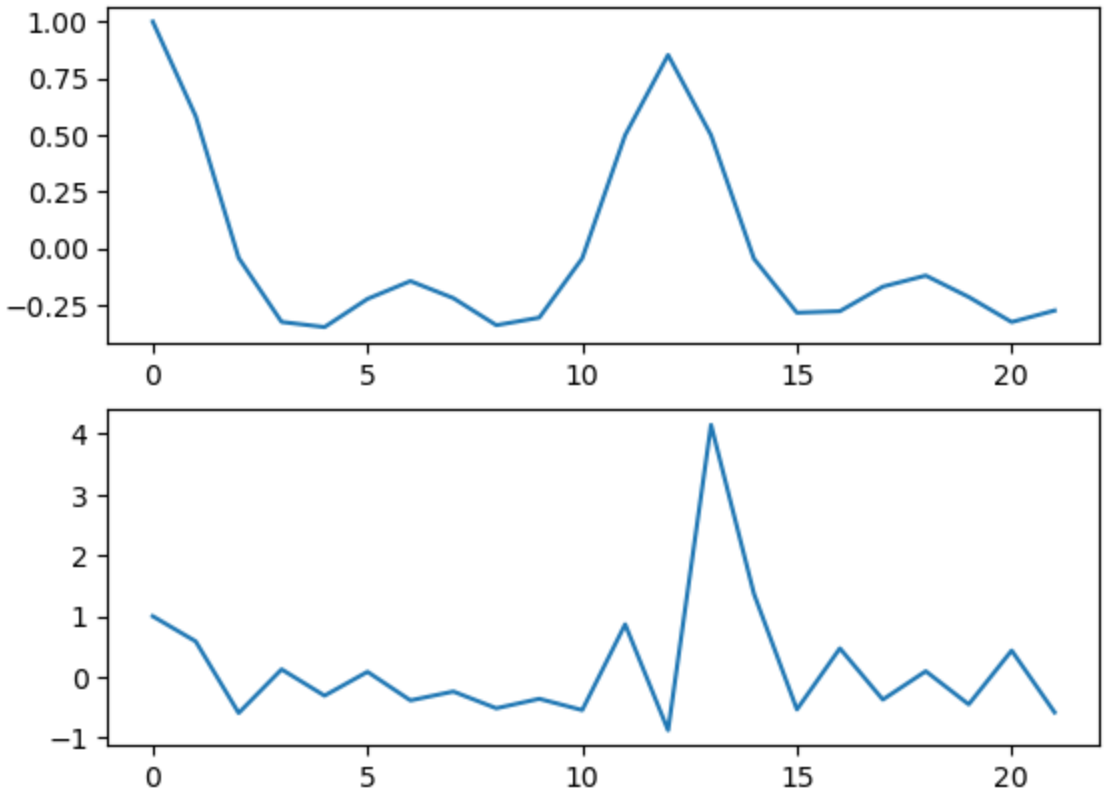
Out[19]:
```
[<matplotlib.lines.Line2D at 0x1d57023ad40>]
```



In [20]:
```
from statsmodels.tsa.stattools import acf,pacf
acf_plot = acf(ts_data_new.dropna())
pacf_plot = pacf(ts_data_new.dropna())
```

In [21]:
```
plt.subplot(211)
plt.plot(acf_plot)
plt.subplot(212)
plt.plot(pacf_plot)
```

Out[21]:
```
[<matplotlib.lines.Line2D at 0x1d5705b01c0>]
```



p = 2 q = 1 d = 2 season = 12

In [28]:
```
from statsmodels.tsa.arima.model import ARIMA
model = ARIMA(ts_data, seasonal_order=(2,2,1,12))
model_fit = model.fit()
```

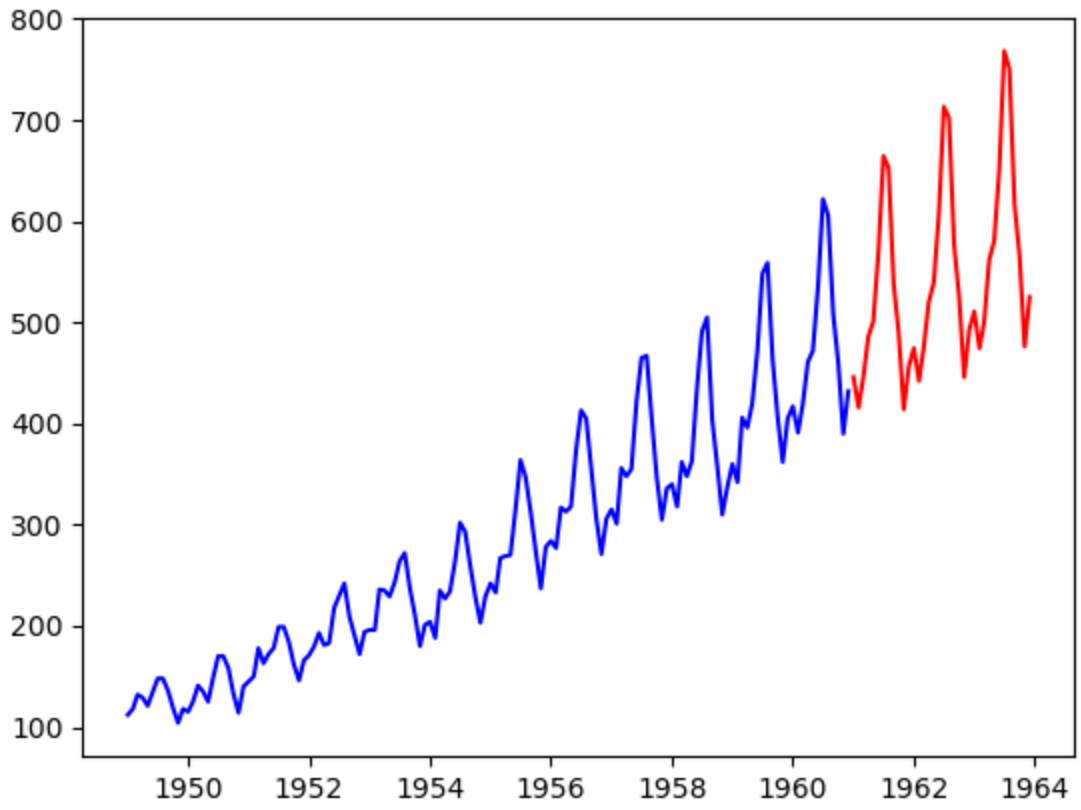
```
E:\anaconda3\lib\site-packages\statsmodels\tsa\base\tsa_model.py:471: ValueWarning: No f
requency information was provided, so inferred frequency MS will be used.
  self._init_dates(dates, freq)
```

In [32]:
```
res = model_fit.forecast(36)
```

In [33]:
```
res = pd.DataFrame(res)
res.columns = ['value']
```

In [34]:
```
plt.plot(ts_data,c='blue')
plt.plot(res,c='red')
```

Out[34]:
```
[<matplotlib.lines.Line2D at 0x1d5738000a0>]
```



In [ ]:

In [ ]: