In [1]:
```python
import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
import plotly.express as pk
```

In [2]:
```python
df =pd.read_csv("C:/Users/vijay kumar jeekuru/Downloads/weatherHistory.csv")
df.head(10)
```

Out[2]:

	Formatted Date	Summary	Precip Type	Temperature (C)	Apparent Temperature (C)	Humidity	Wind Speed (km/h)	Wind Bearing (degrees)	Visibility (km)	Pressure (millibars)	Daily Summary
0	2006-04-01 00:00:00.000 +0200	Partly Cloudy	rain	9.472222	7.388889	0.89	14.1197	251	15.8263	1015.13	Partly cloudy throughout the day.
1	2006-04-01 01:00:00.000 +0200	Partly Cloudy	rain	9.355556	7.227778	0.86	14.2646	259	15.8263	1015.63	Partly cloudy throughout the day.
2	2006-04-01 02:00:00.000 +0200	Mostly Cloudy	rain	9.377778	9.377778	0.89	3.9284	204	14.9569	1015.94	Partly cloudy throughout the day.
3	2006-04-01 03:00:00.000 +0200	Partly Cloudy	rain	8.288889	5.944444	0.83	14.1036	269	15.8263	1016.41	Partly cloudy throughout the day.
4	2006-04-01 04:00:00.000 +0200	Mostly Cloudy	rain	8.755556	6.977778	0.83	11.0446	259	15.8263	1016.51	Partly cloudy throughout the day.
5	2006-04-01 05:00:00.000 +0200	Partly Cloudy	rain	9.222222	7.111111	0.85	13.9587	258	14.9569	1016.66	Partly cloudy throughout the day.
6	2006-04-01 06:00:00.000 +0200	Partly Cloudy	rain	7.733333	5.522222	0.95	12.3648	259	9.9820	1016.72	Partly cloudy throughout the day.
7	2006-04-01 07:00:00.000 +0200	Partly Cloudy	rain	8.772222	6.527778	0.89	14.1519	260	9.9820	1016.84	Partly cloudy throughout the day.
8	2006-04-01 08:00:00.000 +0200	Partly Cloudy	rain	10.822222	10.822222	0.82	11.3183	259	9.9820	1017.37	Partly cloudy throughout the day.
9	2006-04-01 09:00:00.000 +0200	Partly Cloudy	rain	13.772222	13.772222	0.72	12.5258	279	9.9820	1017.22	Partly cloudy throughout the day.

In [3]:
```python
df.tail(10)
```

Out[3]:

	Formatted Date	Summary	Precip Type	Temperature (C)	Apparent Temperature (C)	Humidity	Wind Speed (km/h)	Wind Bearing (degrees)	Visibility (km)	Pressure (millibars)	Daily Summary
96443	2016-09-09 14:00:00.000 +0200	Partly Cloudy	rain	30.894444	29.450000	0.28	14.7798	43	15.5526	1014.66	Partly cloudy starting in the morning.
96444	2016-09-09 15:00:00.000 +0200	Partly Cloudy	rain	31.083333	29.616667	0.28	15.5043	40	16.1000	1014.17	Partly cloudy starting in the morning.
96445	2016-09-09 16:00:00.000 +0200	Partly Cloudy	rain	31.083333	29.611111	0.28	13.8943	40	16.1000	1013.97	Partly cloudy starting in the morning.
96446	2016-09-09 17:00:00.000 +0200	Partly Cloudy	rain	30.766667	29.311111	0.28	14.2163	24	15.5526	1013.83	Partly cloudy starting in the morning.
96447	2016-09-09 18:00:00.000 +0200	Partly Cloudy	rain	28.838889	27.850000	0.32	12.2038	21	16.1000	1014.07	Partly cloudy starting in the morning.
96448	2016-09-09 19:00:00.000 +0200	Partly Cloudy	rain	26.016667	26.016667	0.43	10.9963	31	16.1000	1014.36	Partly cloudy starting in the morning.
96449	2016-09-09 20:00:00.000 +0200	Partly Cloudy	rain	24.583333	24.583333	0.48	10.0947	20	15.5526	1015.16	Partly cloudy starting in the morning.
96450	2016-09-09 21:00:00.000 +0200	Partly Cloudy	rain	22.038889	22.038889	0.56	8.9838	30	16.1000	1015.66	Partly cloudy starting in the morning.
96451	2016-09-09 22:00:00.000 +0200	Partly Cloudy	rain	21.522222	21.522222	0.60	10.5294	20	16.1000	1015.95	Partly cloudy starting in the morning.
96452	2016-09-09 23:00:00.000 +0200	Partly Cloudy	rain	20.438889	20.438889	0.61	5.8765	39	15.5204	1016.16	Partly cloudy starting in the morning.

In [4]:
```python
df.isnull().sum()
```

Out[4]:
```
Formatted Date                0
Summary                       0
Precip Type                 517
Temperature (C)               0
Apparent Temperature (C)      0
Humidity                      0
Wind Speed (km/h)             0
Wind Bearing (degrees)        0
Visibility (km)               0
Pressure (millibars)          0
Daily Summary                 0
dtype: int64
```

In [5]:
```python
df.dtypes
```

Out[5]:
```
Formatted Date               object
Summary                      object
Precip Type                  object
Temperature (C)             float64
Apparent Temperature (C)    float64
Humidity                    float64
Wind Speed (km/h)           float64
Wind Bearing (degrees)        int64
Visibility (km)             float64
Pressure (millibars)        float64
Daily Summary                object
dtype: object
```

In [6]:
```python
df.describe()
```

Out[6]:

	Temperature (C)	Apparent Temperature (C)	Humidity	Wind Speed (km/h)	Wind Bearing (degrees)	Visibility (km)	Pressure (millibars)
count	96453.000000	96453.000000	96453.000000	96453.000000	96453.000000	96453.000000	96453.000000
mean	11.932678	10.855029	0.734899	10.810640	187.509232	10.347325	1003.235956
std	9.551546	10.696847	0.195473	6.913571	107.383428	4.192123	116.969906
min	-21.822222	-27.716667	0.000000	0.000000	0.000000	0.000000	0.000000
25%	4.688889	2.311111	0.600000	5.828200	116.000000	8.339800	1011.900000
50%	12.000000	12.000000	0.780000	9.965900	180.000000	10.046400	1016.450000
75%	18.838889	18.838889	0.890000	14.135800	290.000000	14.812000	1021.090000
max	39.905556	39.344444	1.000000	63.852600	359.000000	16.100000	1046.380000

In [7]:
```python
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 96453 entries, 0 to 96452
Data columns (total 11 columns):
Formatted Date              96453 non-null object
Summary                     96453 non-null object
Precip Type                 95936 non-null object
Temperature (C)             96453 non-null float64
Apparent Temperature (C)    96453 non-null float64
Humidity                    96453 non-null float64
Wind Speed (km/h)           96453 non-null float64
Wind Bearing (degrees)      96453 non-null int64
Visibility (km)             96453 non-null float64
Pressure (millibars)        96453 non-null float64
Daily Summary               96453 non-null object
dtypes: float64(6), int64(1), object(4)
memory usage: 8.1+ MB
```

In [8]:
```python
df['Formatted Date'] =pd.to_datetime(df['Formatted Date'], utc=True)
df['Formatted Date']
```

Out[8]:
```
0       2006-03-31 22:00:00+00:00
1       2006-03-31 23:00:00+00:00
2       2006-04-01 00:00:00+00:00
3       2006-04-01 01:00:00+00:00
4       2006-04-01 02:00:00+00:00
                   ...
96448   2016-09-09 17:00:00+00:00
96449   2016-09-09 18:00:00+00:00
96450   2016-09-09 19:00:00+00:00
96451   2016-09-09 20:00:00+00:00
96452   2016-09-09 21:00:00+00:00
Name: Formatted Date, Length: 96453, dtype: datetime64[ns, UTC]
```

In [9]:
```python
df=df.set_index('Formatted Date')
df.head()
```

Out[9]:

Formatted Date	Summary	Precip Type	Temperature (C)	Apparent Temperature (C)	Humidity	Wind Speed (km/h)	Wind Bearing (degrees)	Visibility (km)	Pressure (millibars)	Daily Summary
2006-03-31 22:00:00+00:00	Partly Cloudy	rain	9.472222	7.388889	0.89	14.1197	251	15.8263	1015.13	Partly cloudy throughout the day.
2006-03-31 23:00:00+00:00	Partly Cloudy	rain	9.355556	7.227778	0.86	14.2646	259	15.8263	1015.63	Partly cloudy throughout the day.
2006-04-01 00:00:00+00:00	Mostly Cloudy	rain	9.377778	9.377778	0.89	3.9284	204	14.9569	1015.94	Partly cloudy throughout the day.
2006-04-01 01:00:00+00:00	Partly Cloudy	rain	8.288889	5.944444	0.83	14.1036	269	15.8263	1016.41	Partly cloudy throughout the day.
2006-04-01 02:00:00+00:00	Mostly Cloudy	rain	8.755556	6.977778	0.83	11.0446	259	15.8263	1016.51	Partly cloudy throughout the day.

In [10]:
```python
data_coloumn=['Apparent Temperature (C)','Humidity']
df_monthly_mean=df.resample('MS').mean()
df_monthly_mean.head()
```

Out[10]:

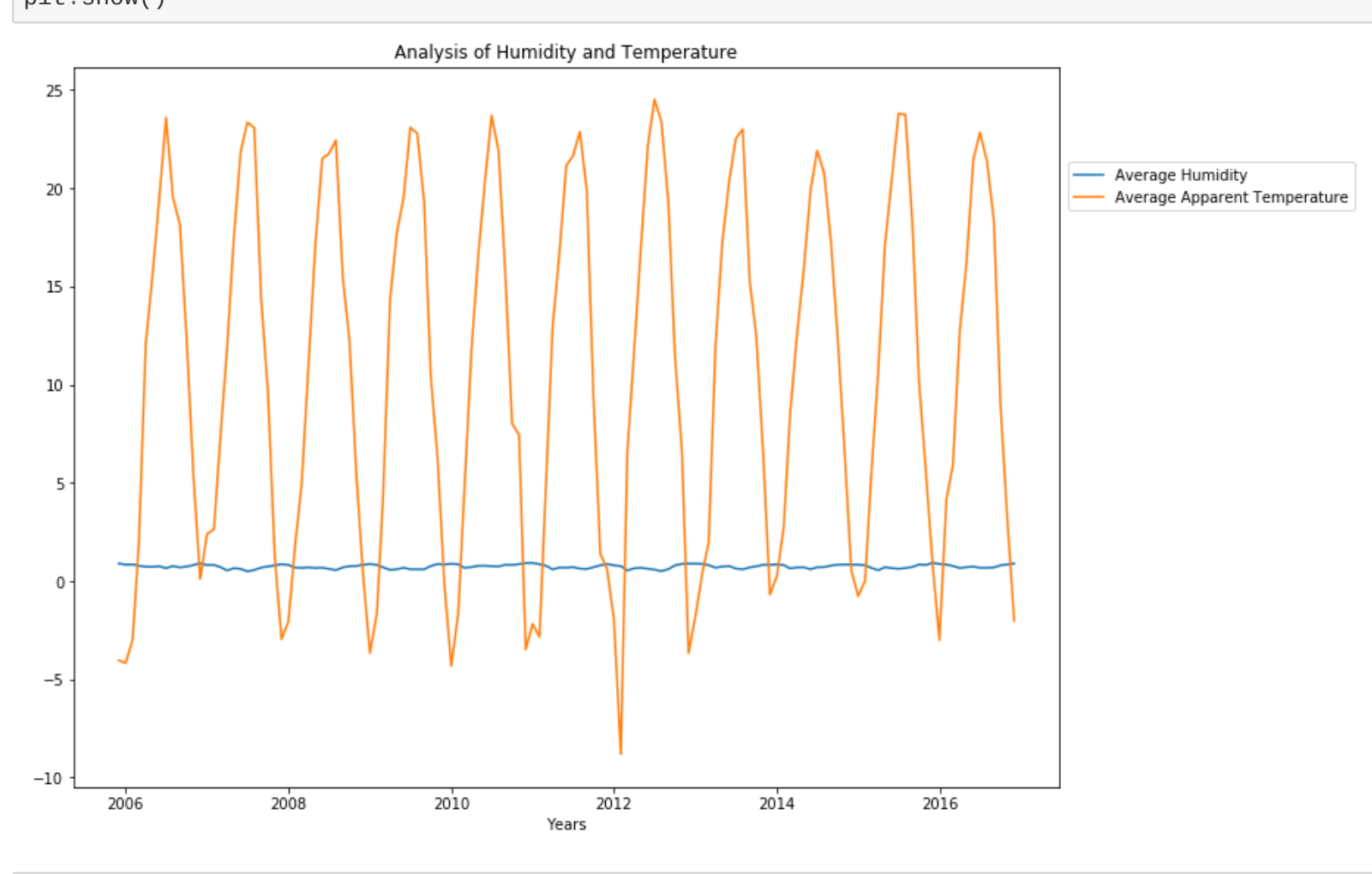
Formatted Date	Temperature (C)	Apparent Temperature (C)	Humidity	Wind Speed (km/h)	Wind Bearing (degrees)	Visibility (km)	Pressure (millibars)
2005-12-01 00:00:00+00:00	0.577778	-4.050000	0.890000	17.114300	140.000000	9.982000	1016.660000
2006-01-01 00:00:00+00:00	-1.677942	-4.173708	0.834610	8.894211	161.018817	7.894064	1021.204960
2006-02-01 00:00:00+00:00	-0.065394	-2.990716	0.843467	10.957008	197.886905	7.418794	995.183914
2006-03-01 00:00:00+00:00	4.559274	1.969780	0.778737	14.421488	195.059140	9.602590	976.436263
2006-04-01 00:00:00+00:00	12.635031	12.098827	0.728625	10.930670	191.877778	10.626760	1013.493694

In [11]:
```python
df=df.resample('MS').mean()
```

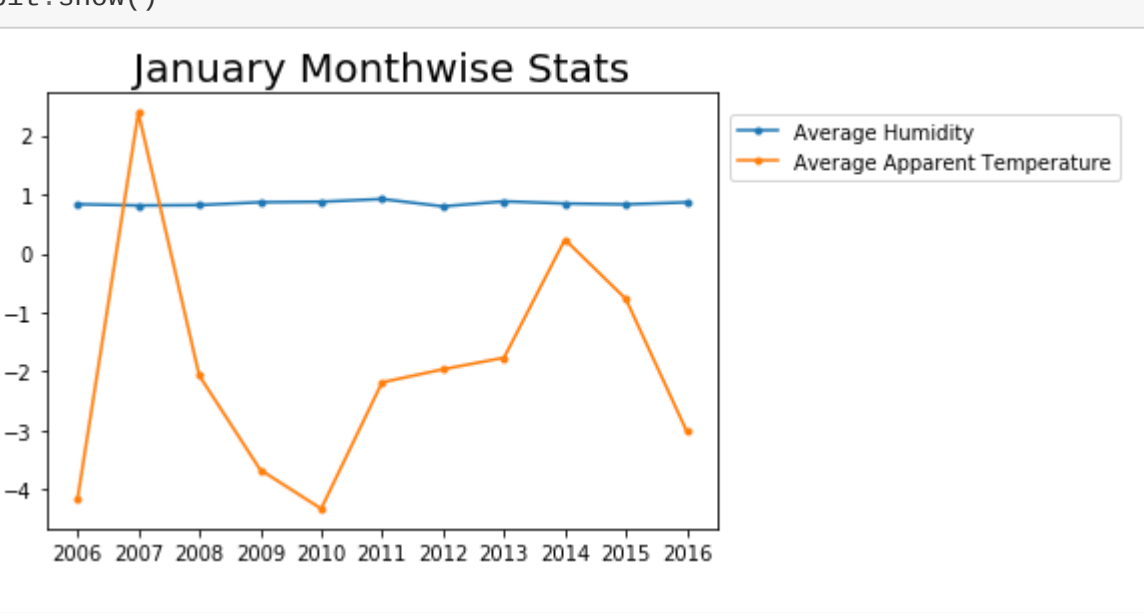
In [12]:
```python
apt=df['Apparent Temperature (C)']
```

In [13]:
```python
hmd=df['Humidity']
```

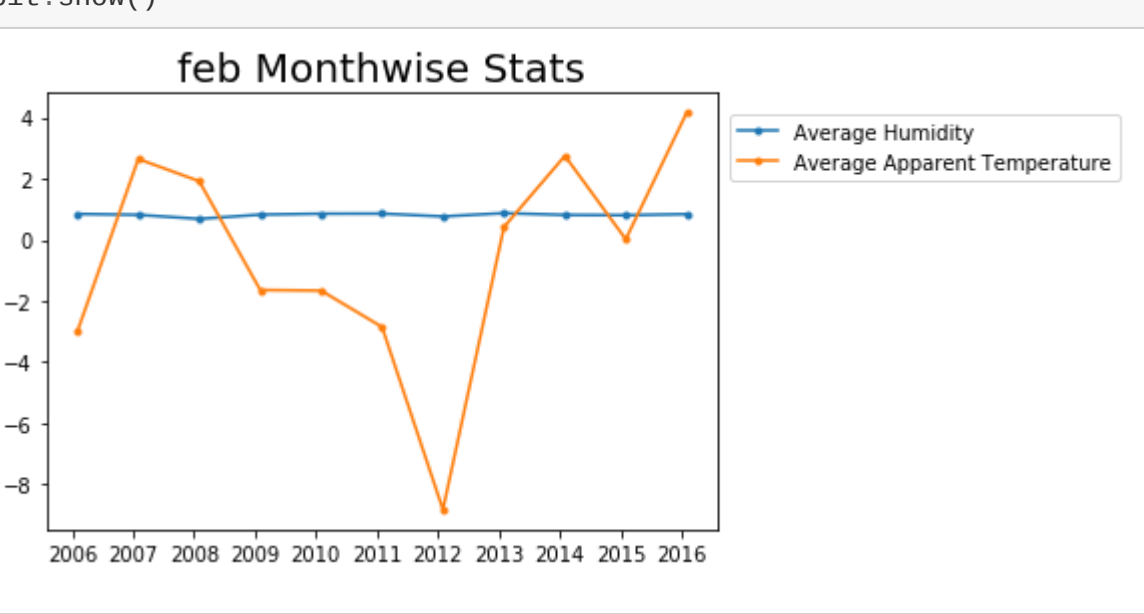
In [14]:
```python
import seaborn as sns
import warnings
warnings.filterwarnings("ignore")

plt.figure(figsize=(12,9))
plt.title("Analysis of Humidity and Temperature")
plt.plot(hmd,label="Average Humidity")
plt.plot(apt,label="Average Apparent Temperature")
plt.xlabel("Years")
plt.legend(loc=(1.01,0.8))
plt.show()
```



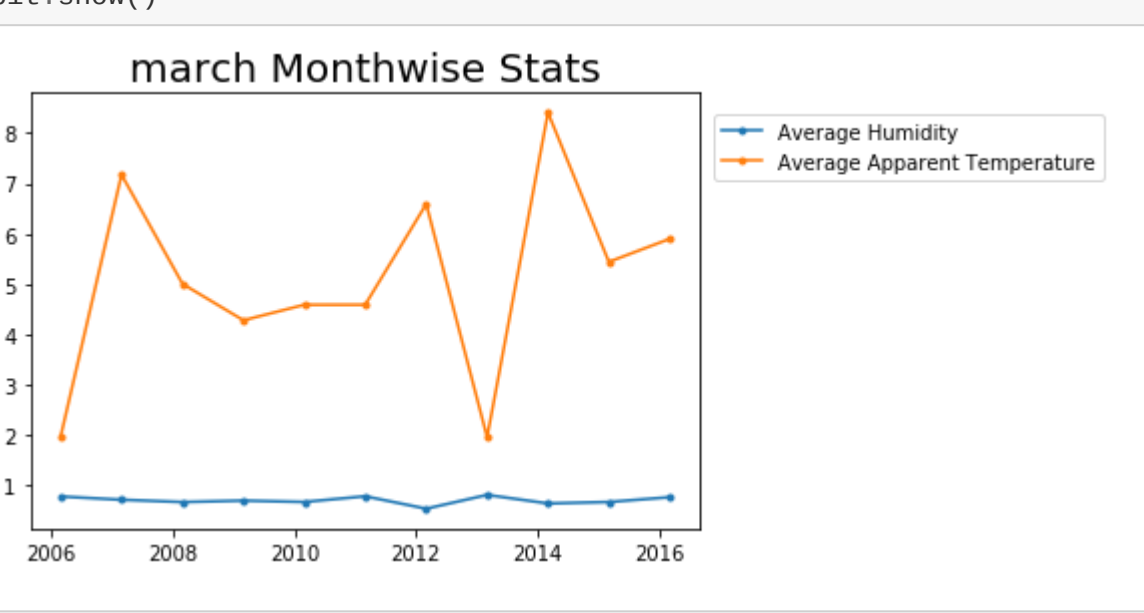
In [15]:
```python
df1 = df[df.index.month==1]
hdt=df1['Humidity']
atm=df1['Apparent Temperature (C)']
plt.title("January Monthwise Stats",fontsize=20)
plt.plot(hdt,label="Average Humidity",marker=".")
plt.plot(atm,label="Average Apparent Temperature",marker='.')
plt.legend(loc=(1.02,0.8))
plt.show()
```



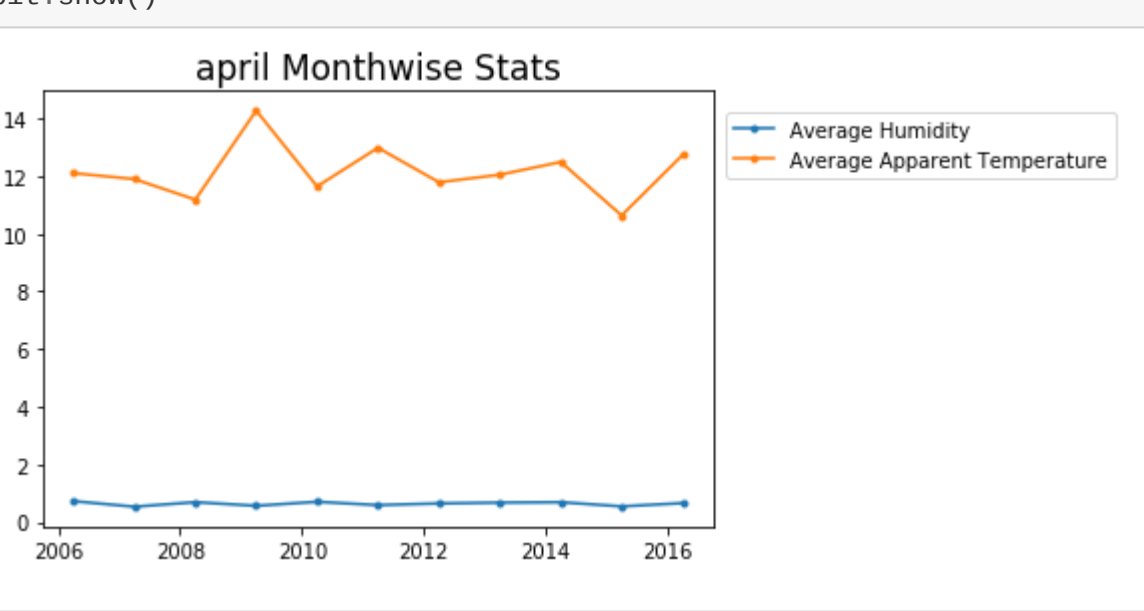
In [16]:
```python
df1 = df[df.index.month==2]
hdt=df1['Humidity']
atm=df1['Apparent Temperature (C)']
plt.title("feb Monthwise Stats",fontsize=20)
plt.plot(hdt,label="Average Humidity",marker=".")
plt.plot(atm,label="Average Apparent Temperature",marker='.')
plt.legend(loc=(1.02,0.8))
plt.show()
```



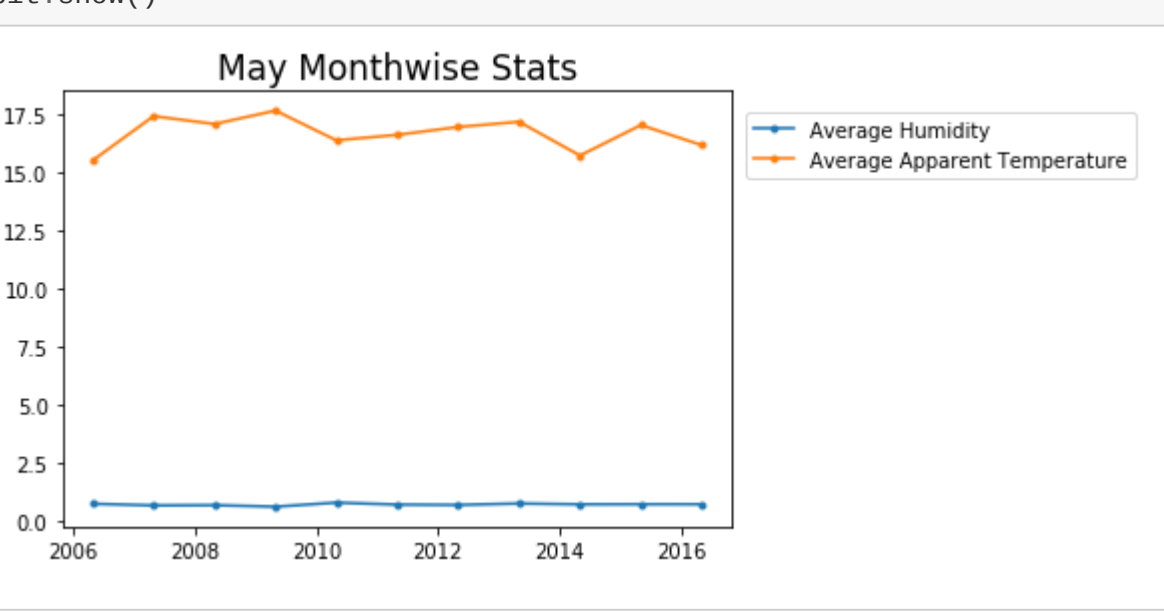
In [17]:
```python
df1 = df[df.index.month==3]
hdt=df1['Humidity']
atm=df1['Apparent Temperature (C)']
plt.title("march Monthwise Stats",fontsize=20)
plt.plot(hdt,label="Average Humidity",marker=".")
plt.plot(atm,label="Average Apparent Temperature",marker='.')
plt.legend(loc=(1.02,0.8))
plt.show()
```



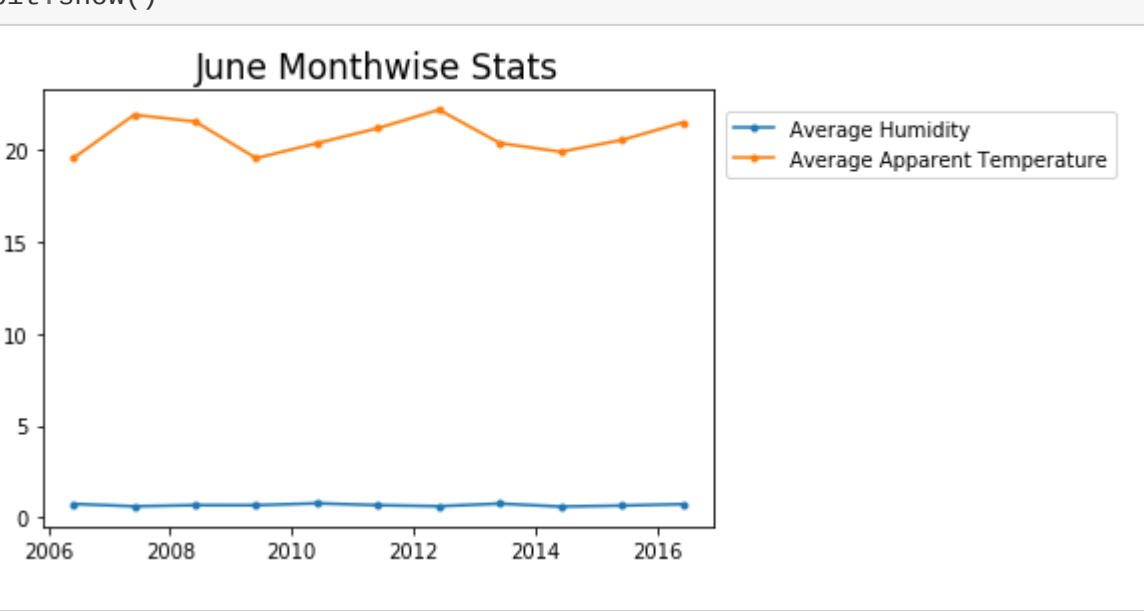
In [18]:
```python
df1 = df[df.index.month==4]
hdt=df1['Humidity']
atm=df1['Apparent Temperature (C)']
plt.title("april Monthwise Stats",fontsize=17)
plt.plot(hdt,label="Average Humidity",marker=".")
plt.plot(atm,label="Average Apparent Temperature",marker='.')
plt.legend(loc=(1.02,0.8))
plt.show()
```



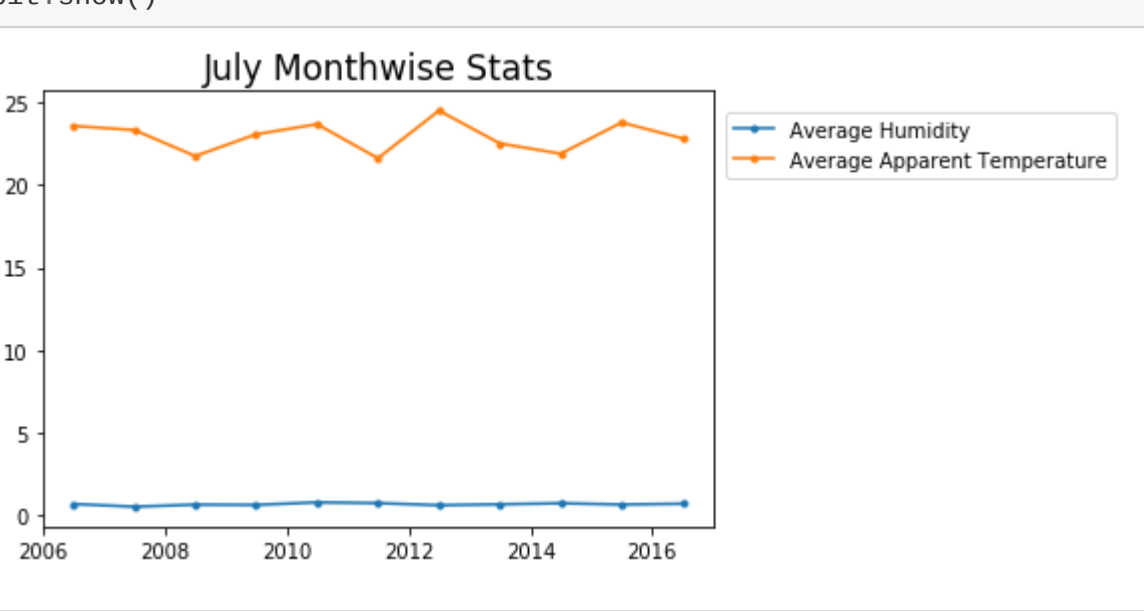
In [19]:
```python
df1 = df[df.index.month==5]
hdt=df1['Humidity']
atm=df1['Apparent Temperature (C)']
plt.title("May Monthwise Stats",fontsize=17)
plt.plot(hdt,label="Average Humidity",marker=".")
plt.plot(atm,label="Average Apparent Temperature",marker='.')
plt.legend(loc=(1.02,0.8))
plt.show()
```



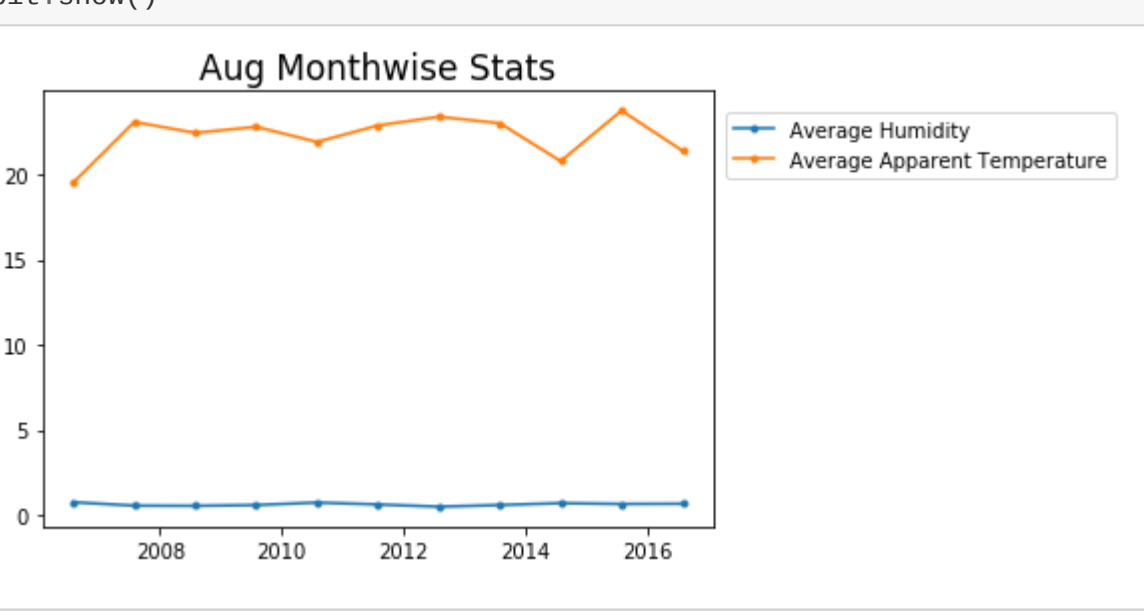
In [20]:
```python
df1 = df[df.index.month==6]
hdt=df1['Humidity']
atm=df1['Apparent Temperature (C)']
plt.title("June Monthwise Stats",fontsize=17)
plt.plot(hdt,label="Average Humidity",marker=".")
plt.plot(atm,label="Average Apparent Temperature",marker='.')
plt.legend(loc=(1.02,0.8))
plt.show()
```



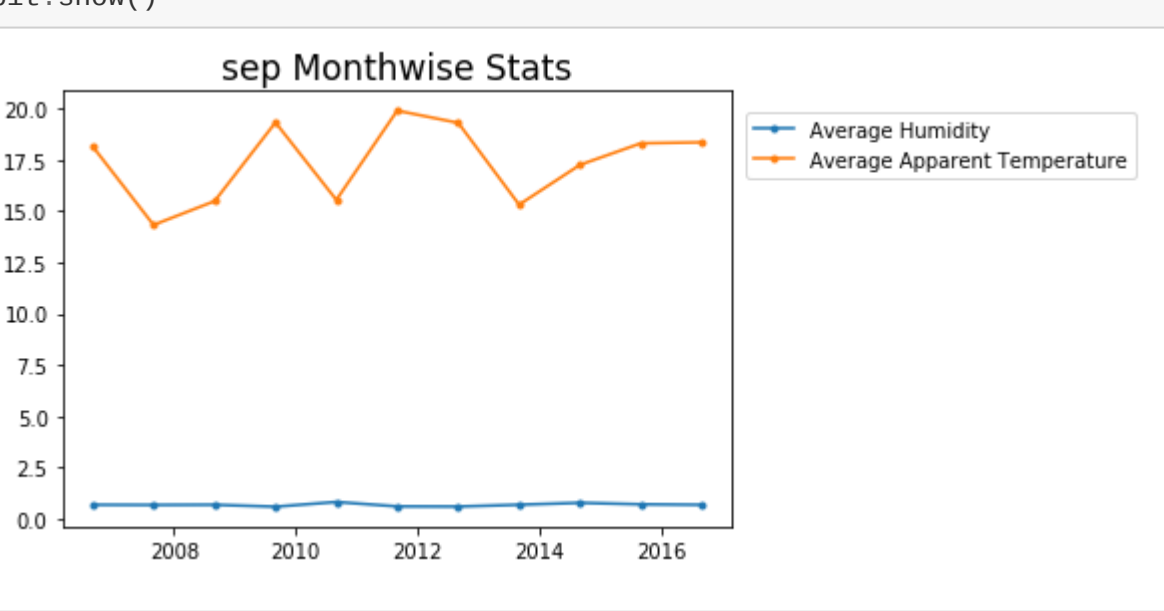
In [21]:
```python
df1 = df[df.index.month==7]
hdt=df1['Humidity']
atm=df1['Apparent Temperature (C)']
plt.title("July Monthwise Stats",fontsize=17)
plt.plot(hdt,label="Average Humidity",marker=".")
plt.plot(atm,label="Average Apparent Temperature",marker='.')
plt.legend(loc=(1.02,0.8))
plt.show()
```



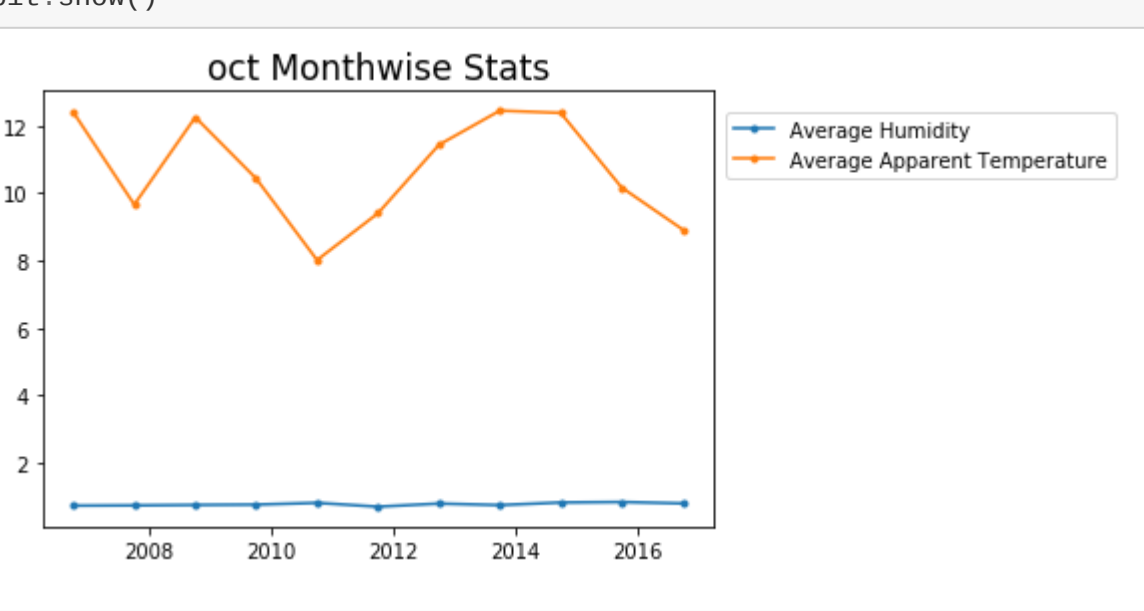
In [22]:
```python
df1 = df[df.index.month==8]
hdt=df1['Humidity']
atm=df1['Apparent Temperature (C)']
plt.title("Aug Monthwise Stats",fontsize=17)
plt.plot(hdt,label="Average Humidity",marker=".")
plt.plot(atm,label="Average Apparent Temperature",marker='.')
plt.legend(loc=(1.02,0.8))
plt.show()
```



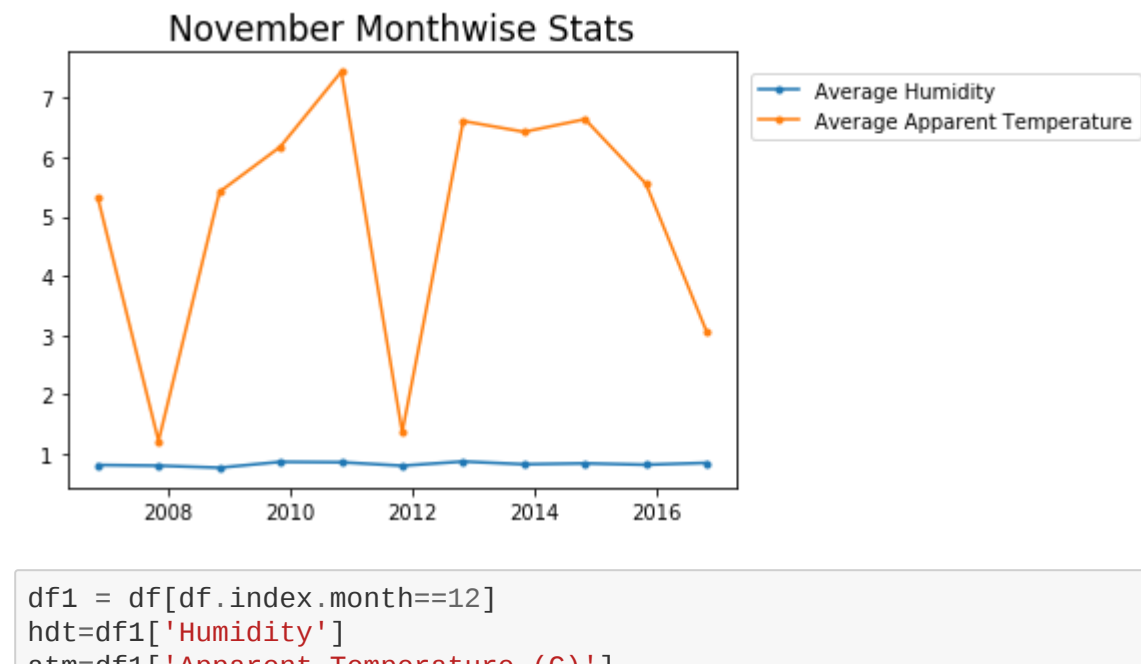
In [23]:
```python
df1 = df[df.index.month==9]
hdt=df1['Humidity']
atm=df1['Apparent Temperature (C)']
plt.title("sep Monthwise Stats",fontsize=17)
plt.plot(hdt,label="Average Humidity",marker=".")
plt.plot(atm,label="Average Apparent Temperature",marker='.')
plt.legend(loc=(1.02,0.8))
plt.show()
```



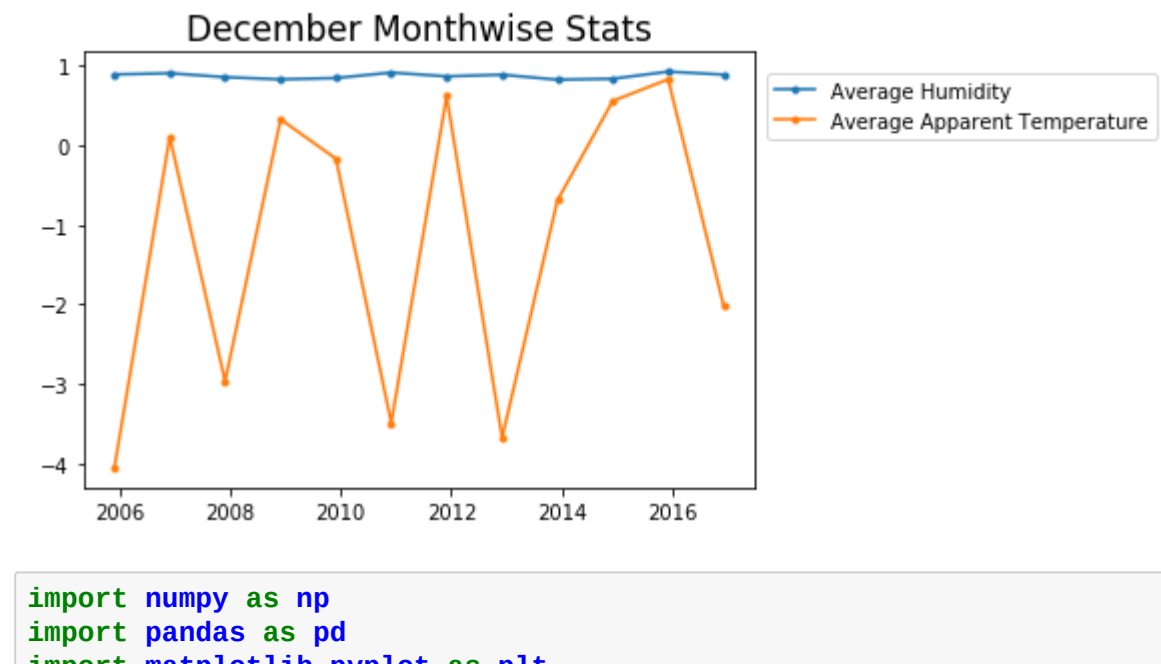
In [24]:
```python
df1 = df[df.index.month==10]
hdt=df1['Humidity']
atm=df1['Apparent Temperature (C)']
plt.title("oct Monthwise Stats",fontsize=17)
plt.plot(hdt,label="Average Humidity",marker=".")
plt.plot(atm,label="Average Apparent Temperature",marker='.')
plt.legend(loc=(1.02,0.8))
plt.show()
```



In [25]:
```python
df1 = df[df.index.month==11]
hdt=df1['Humidity']
atm=df1['Apparent Temperature (C)']
plt.title("November Monthwise Stats",fontsize=17)
plt.plot(hdt,label="Average Humidity",marker=".")
plt.plot(atm,label="Average Apparent Temperature",marker='.')
plt.legend(loc=(1.02,0.8))
plt.show()
```



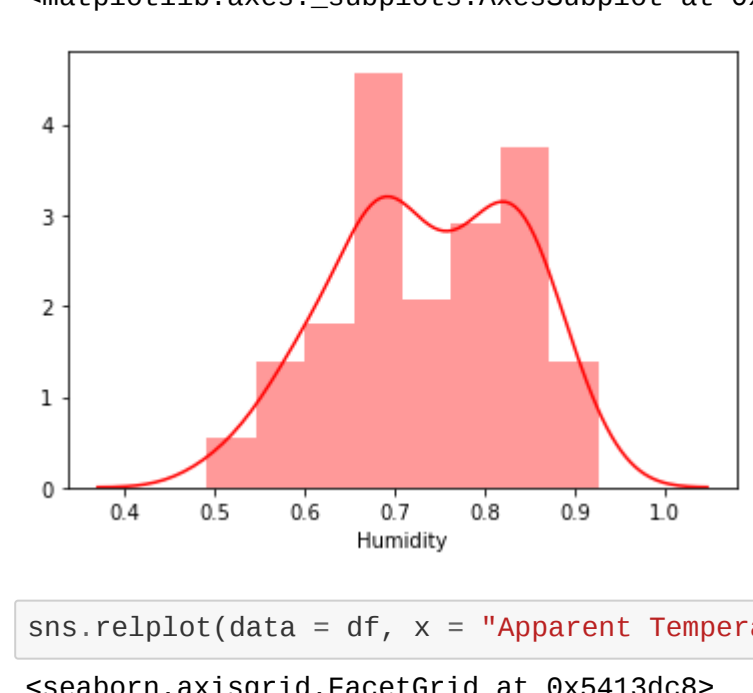
In [26]:
```python
df1 = df[df.index.month==12]
hdt=df1['Humidity']
atm=df1['Apparent Temperature (C)']
plt.title("December Monthwise Stats",fontsize=17)
plt.plot(hdt,label="Average Humidity",marker=".")
plt.plot(atm,label="Average Apparent Temperature",marker='.')
plt.legend(loc=(1.02,0.8))
plt.show()
```



In [27]:
```python
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
import sklearn as sk
```

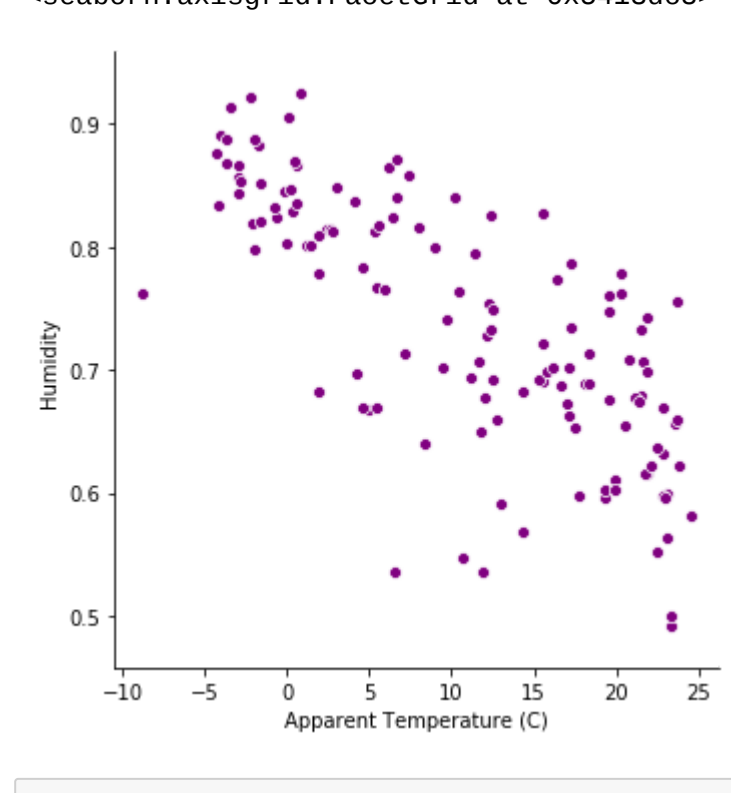
In [28]:
```python
sns.distplot(df.Humidity, color = 'red')
```

Out[28]:
```
<matplotlib.axes._subplots.AxesSubplot at 0x193f4e48>
```



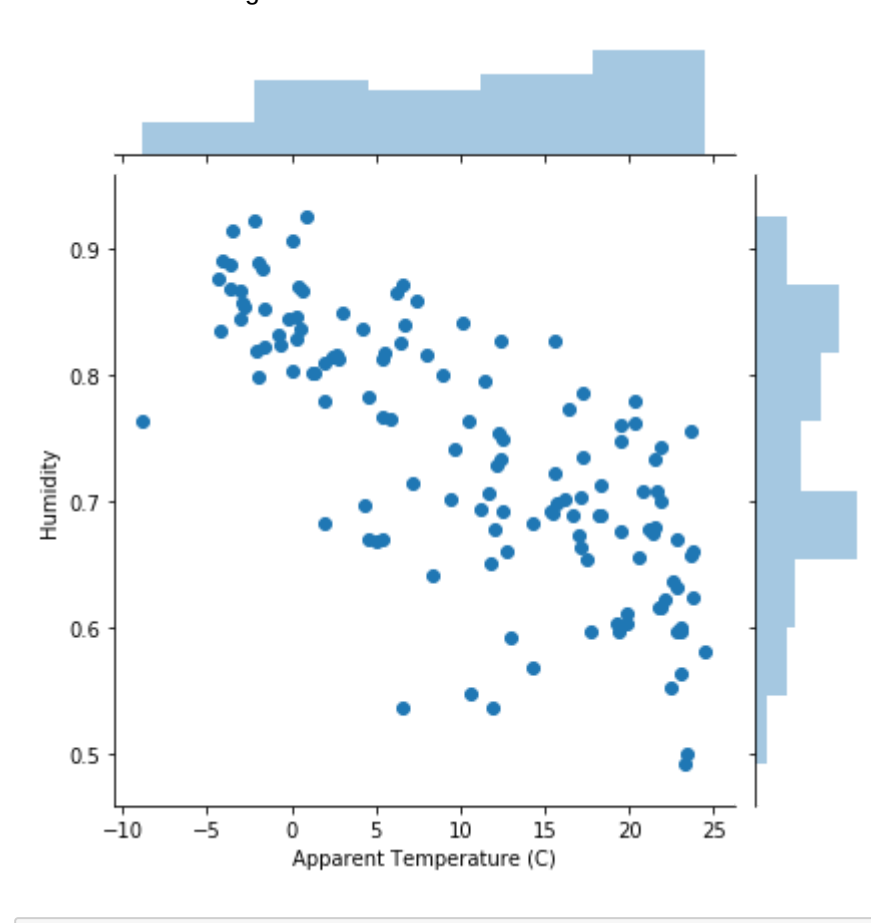
In [34]:
```python
sns.relplot(data = df, x = "Apparent Temperature (C)", y = "Humidity", color = 'purple')
```

Out[34]:
```
<seaborn.axisgrid.FacetGrid at 0x5413dc8>
```



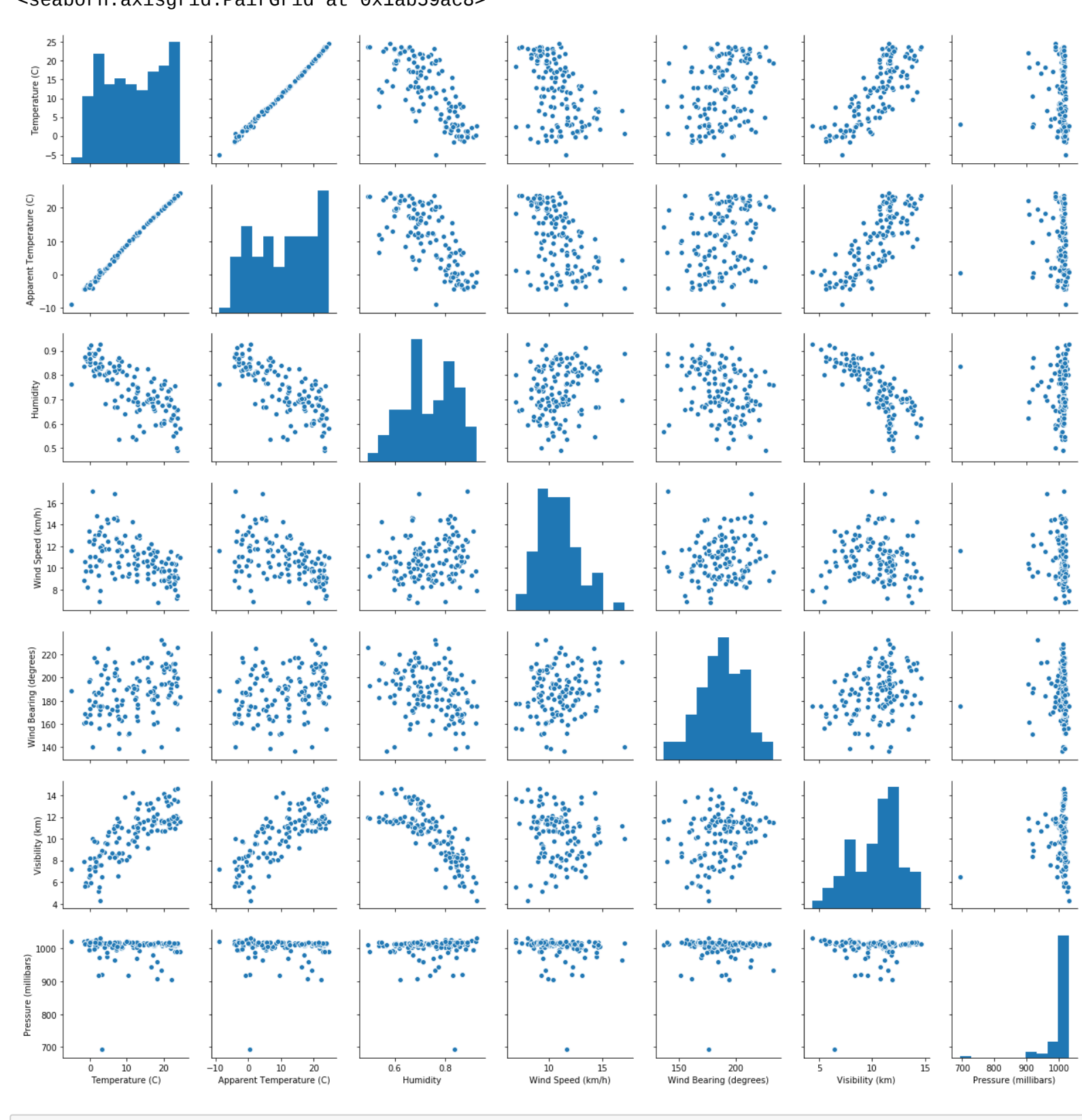
In [30]:
```python
sns.jointplot(data = df, x = "Apparent Temperature (C)", y = "Humidity")
```

Out[30]:
```
<seaborn.axisgrid.JointGrid at 0x19c4de48>
```



In [31]:
```python
sns.pairplot(df)
```

Out[31]:
```
<seaborn.axisgrid.PairGrid at 0x1ab59ac8>
```



In [32]:
```python
#Conclusion:¶
#There is No change in average humidity. The year 2009 can see an increase in average apparent temperature, then a fall in 2010, then a slight increase in 2011, then a significant drop in 2015, and then an increase in 2016.
```

In [ ]: