

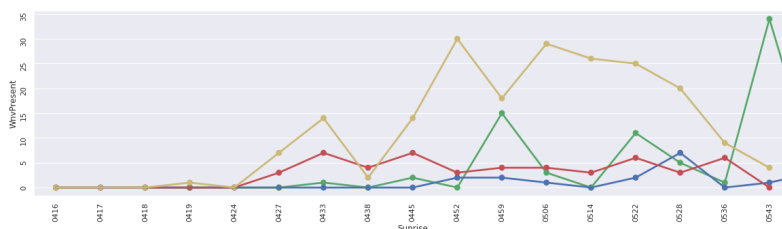
Search

From the above plot we can see that for Virus to be present the departure from the average temperature is not much (except for 2013 where it shows equal for both) which indicates the air was already containing some moisture in it.

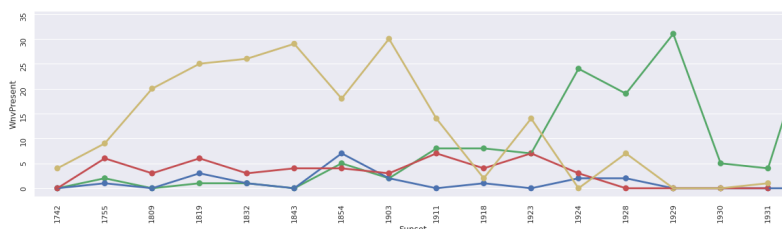
This also raises one question for us to see if the DewPoint and WetBulb are related to each other. We will use some statistical test to check this out.

The plot above confirms our assumption even further. For almost all the years the temp distribution is on the higher side where Wnv is present.

```
In [22]: # Let us see if the timings of Sunrise seem to be of any use to us.
plt.figure(figsize=(20,5))
tmp = df.groupby(['Year', 'Sunrise'])['WnvPresent'].sum().reset_index()
sns.pointplot(tmp.loc[tmp['Year']==2007, 'Sunrise'], tmp['WnvPresent'], color='g')
sns.pointplot(tmp.loc[tmp['Year']==2009, 'Sunrise'], tmp['WnvPresent'], color='b')
sns.pointplot(tmp.loc[tmp['Year']==2011, 'Sunrise'], tmp['WnvPresent'], color='r')
sns.pointplot(tmp.loc[tmp['Year']==2013, 'Sunrise'], tmp['WnvPresent'], color='y')
plt.tick_params(rotation=90)
plt.show()
```



```
In [23]: plt.figure(figsize=(20,5))
tmp = df.groupby(['Year', 'Sunset'])['WnvPresent'].sum().reset_index()
sns.pointplot(tmp.loc[tmp['Year']==2007, 'Sunset'], tmp['WnvPresent'], color='g',
              edge=True)
sns.pointplot(tmp.loc[tmp['Year']==2009, 'Sunset'], tmp['WnvPresent'], color='b')
sns.pointplot(tmp.loc[tmp['Year']==2011, 'Sunset'], tmp['WnvPresent'], color='r')
sns.pointplot(tmp.loc[tmp['Year']==2013, 'Sunset'], tmp['WnvPresent'], color='y')
plt.tick_params(rotation=90)
plt.show()
```



We are not able to deduce any meaningful pattern in the Sunrise and Sunset Feature that relates to the presence of Wnv carrying mosquitos. Hence we might need to re-consider this feature while modelling or come up with some useful feature creation steps.

```
In [24]: # Effect of Climatic Conditions
print(f'Count of distinct weather condition: {df["CodeSum"].nunique()}')
fig, ax = plt.subplots(1,1, figsize=(20,5))
tmp = pd.crosstab(df['CodeSum'], df['WnvPresent'], normalize='index')
```

Notebook

Input

Comments