dsbda-assignment-6

May 9, 2023

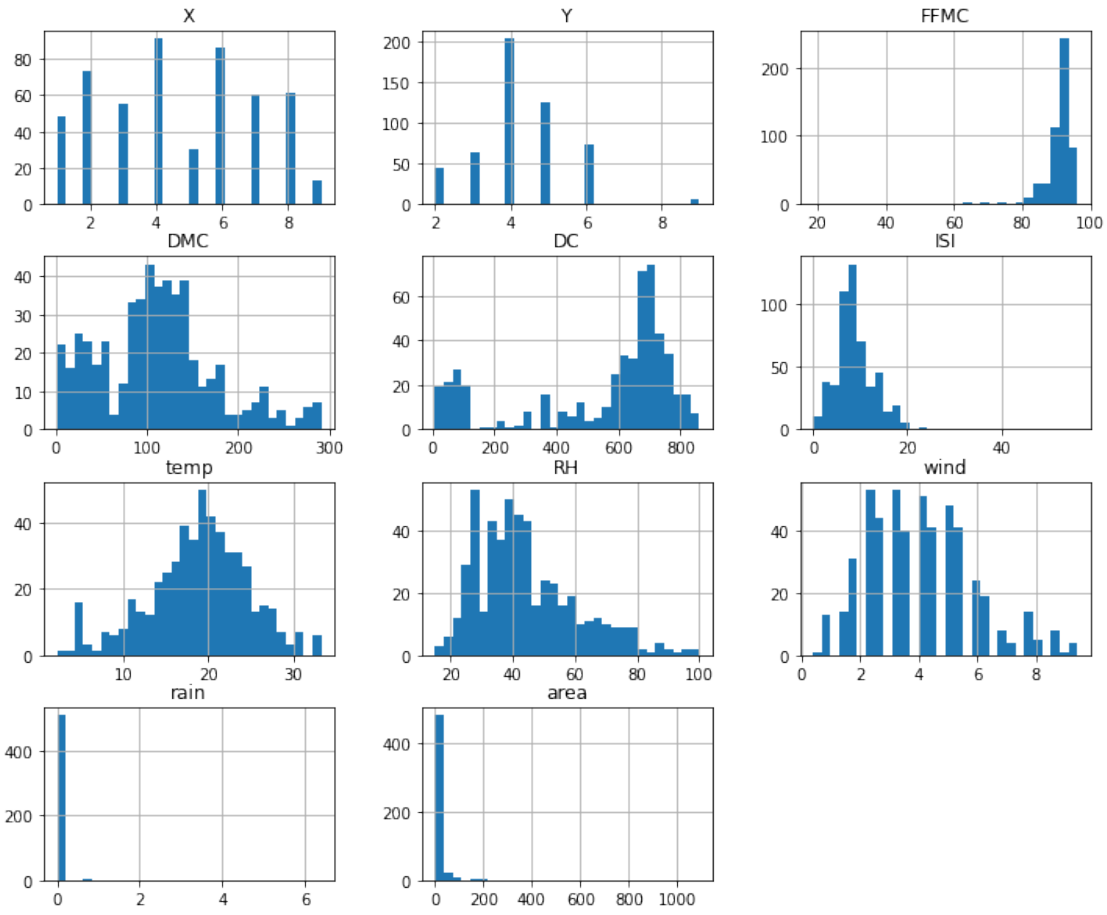
[2]:
```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

[3]:
```
dataset = pd.read_csv("./forestfires.csv")
dataset.head()
```

[3]:

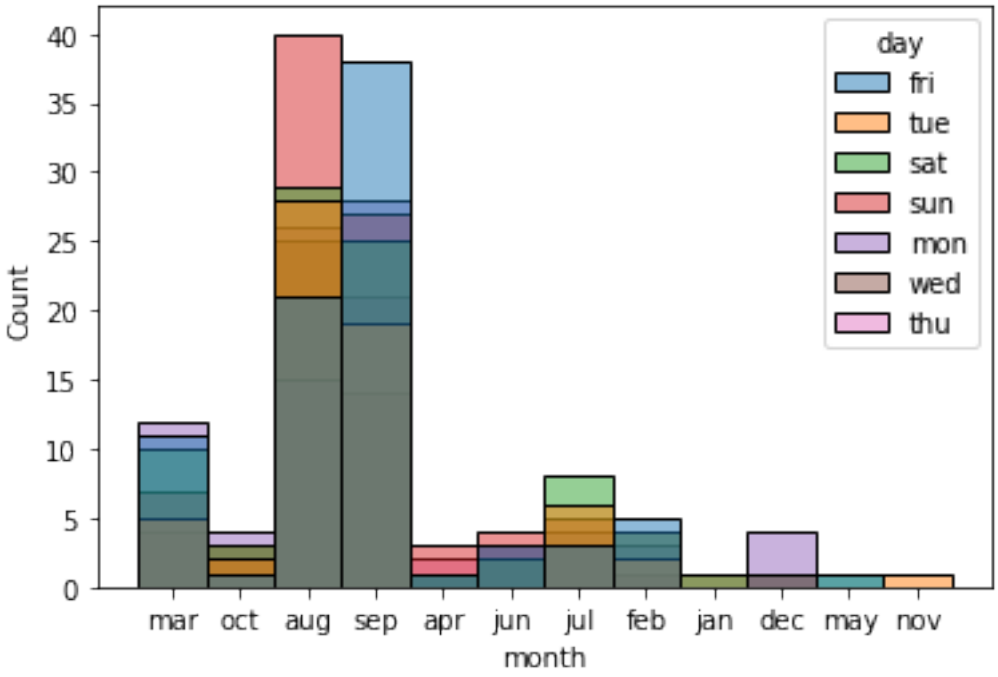
	X	Y	month	day	FFMC	DMC	DC	ISI	temp	RH	wind	rain	area
0	7	5	mar	fri	86.2	26.2	94.3	5.1	8.2	51	6.7	0.0	0.0
1	7	4	oct	tue	90.6	35.4	669.1	6.7	18.0	33	0.9	0.0	0.0
2	7	4	oct	sat	90.6	43.7	686.9	6.7	14.6	33	1.3	0.0	0.0
3	8	6	mar	fri	91.7	33.3	77.5	9.0	8.3	97	4.0	0.2	0.0
4	8	6	mar	sun	89.3	51.3	102.2	9.6	11.4	99	1.8	0.0	0.0

[4]:
```
dataset.hist(bins=30, figsize=(12,10))
plt.show()
```



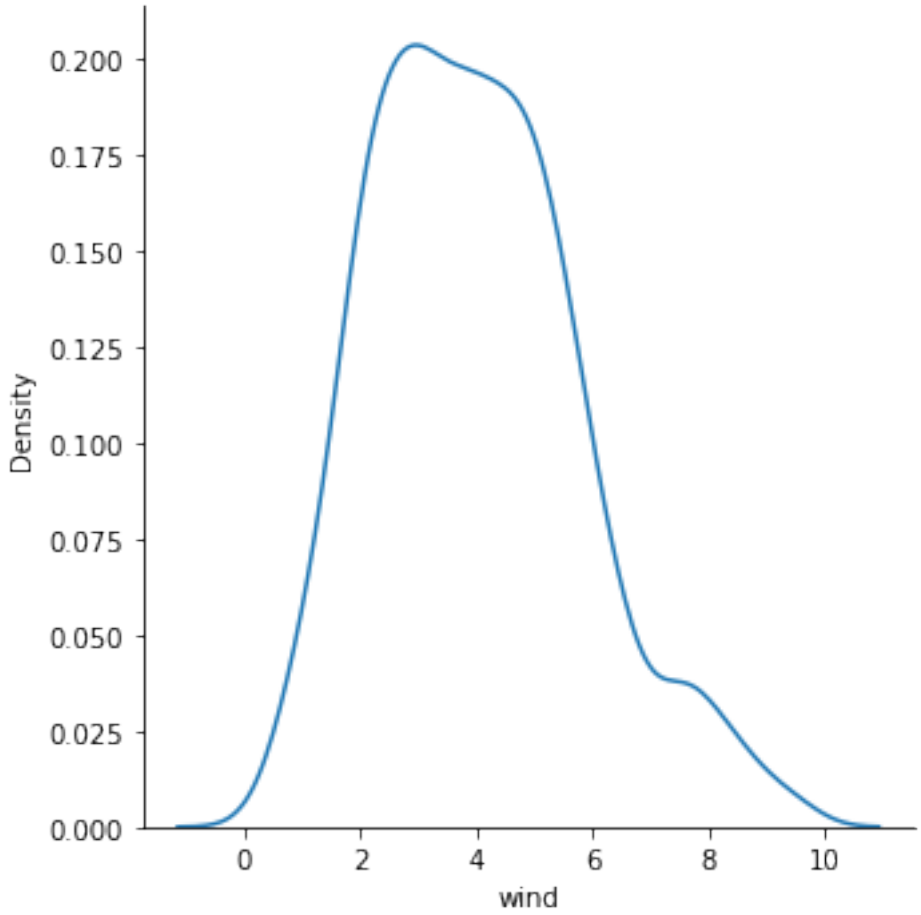
[5]:
```
# The bins parameter enables you to control the bins of the histogram (i.e.,␣
↪the number of bars).
sns.histplot(data=dataset, x="month", hue="day")
```

[5]:
```
<AxesSubplot:xlabel='month', ylabel='Count'>
```



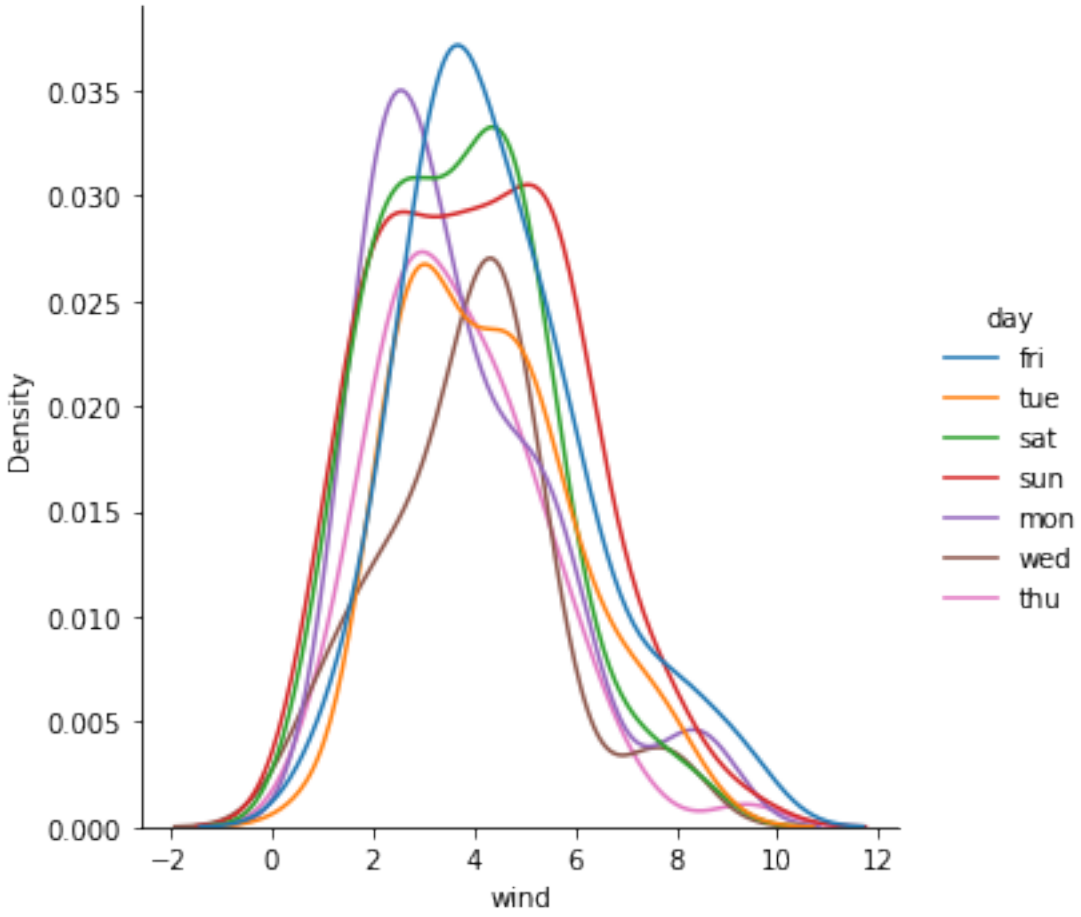
[6]:
```
sns.displot(data=dataset, x="wind", kind="kde")
```

[6]: <seaborn.axisgrid.FacetGrid at 0x7f87b033b780>



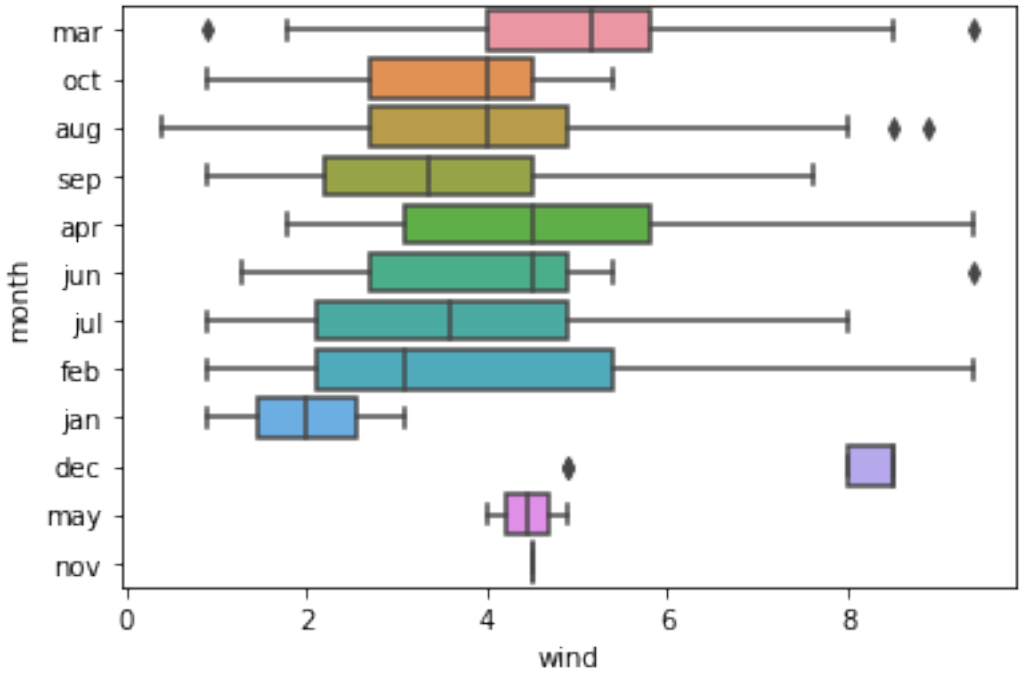
[7]:
```
sns.displot(data=dataset, x="wind", hue="day", kind="kde")
```

[7]:
```
<seaborn.axisgrid.FacetGrid at 0x7f87b010c898>
```



[8]:
```
sns.boxplot(data=dataset, x="wind", y="month")
```

[8]:
```
<AxesSubplot:xlabel='wind', ylabel='month'>
```



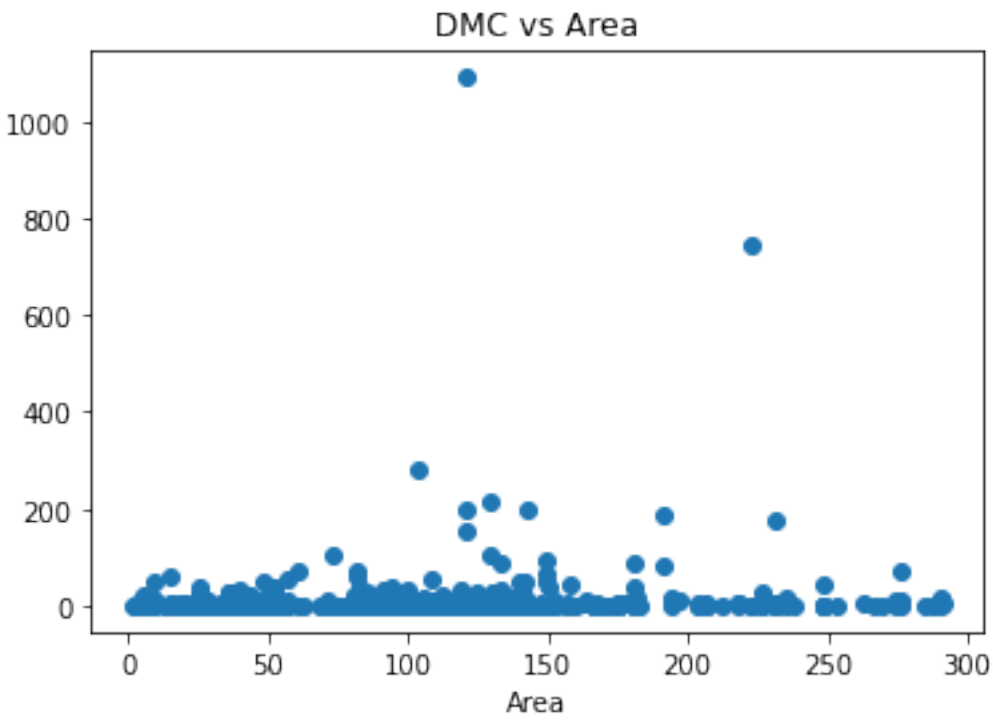
[9]:
```
# Create new column 'burned'. 0 if area = 0.0, else 1
dataset['burned'] = dataset['area'].apply(lambda x: 0.0 if x == 0.0 else 1.0)
```

[10]:
```
dataset.head()
```

[10]:

	X	Y	month	day	FFMC	DMC	DC	ISI	temp	RH	wind	rain	area	burned
0	7	5	mar	fri	86.2	26.2	94.3	5.1	8.2	51	6.7	0.0	0.0	0.0
1	7	4	oct	tue	90.6	35.4	669.1	6.7	18.0	33	0.9	0.0	0.0	0.0
2	7	4	oct	sat	90.6	43.7	686.9	6.7	14.6	33	1.3	0.0	0.0	0.0
3	8	6	mar	fri	91.7	33.3	77.5	9.0	8.3	97	4.0	0.2	0.0	0.0
4	8	6	mar	sun	89.3	51.3	102.2	9.6	11.4	99	1.8	0.0	0.0	0.0

[13]:
```
plt.scatter(dataset["DMC"],dataset["area"])
plt.title("DMC vs Area")
plt.xlabel("DMC")
plt.xlabel("Area")
plt.show()
```


DMC vs Area
1000
800
600
400
200
0
0
50
100
150
200
250
300
Area

[ ]: