In [19]:
```
import plotly
import plotly.express as px
import cufflinks as cf
import pandas as pd
import numpy as np
from plotly.offline import download_plotlyjs,init_notebook_mode,plot,iplot
init_notebook_mode(connected=True)
cf.go_offline()
import matplotlib.pyplot as plt
```

In [6]:
```
bike = pd.read_csv("train.csv")
```

In [7]:
```
bike.head()
```

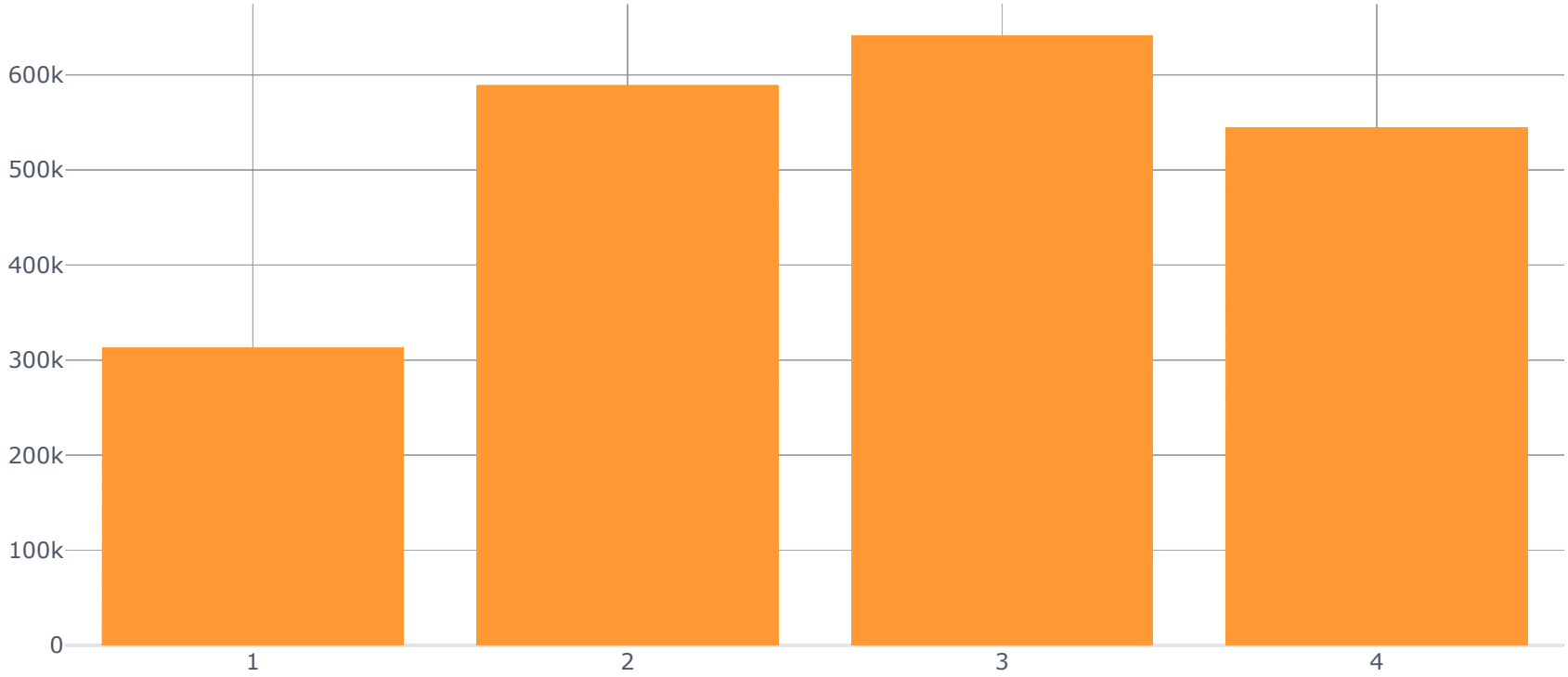
Out[7]:

	datetime	season	holiday	workingday	weather	temp	atemp	humidity	windspeed	casual	registered	count
0	2011-01-01 00:00:00	1	0	0	1	9.84	14.395	81	0.0	3	13	16
1	2011-01-01 01:00:00	1	0	0	1	9.02	13.635	80	0.0	8	32	40
2	2011-01-01 02:00:00	1	0	0	1	9.02	13.635	80	0.0	5	27	32
3	2011-01-01 03:00:00	1	0	0	1	9.84	14.395	75	0.0	3	10	13
4	2011-01-01 04:00:00	1	0	0	1	9.84	14.395	75	0.0	0	1	1

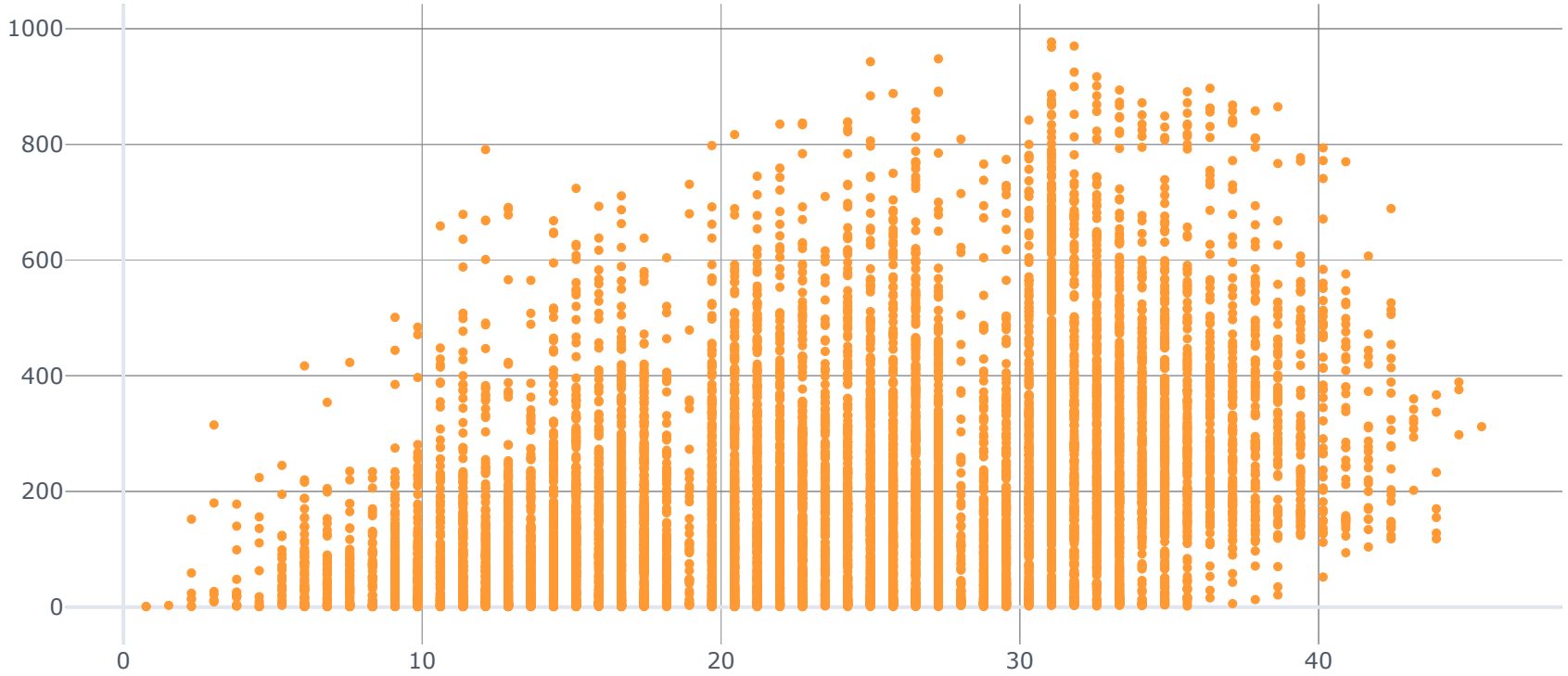
In [14]:
```
bike.weather.unique()
```

Out[14]:
```
array([1, 2, 3, 4], dtype=int64)
```

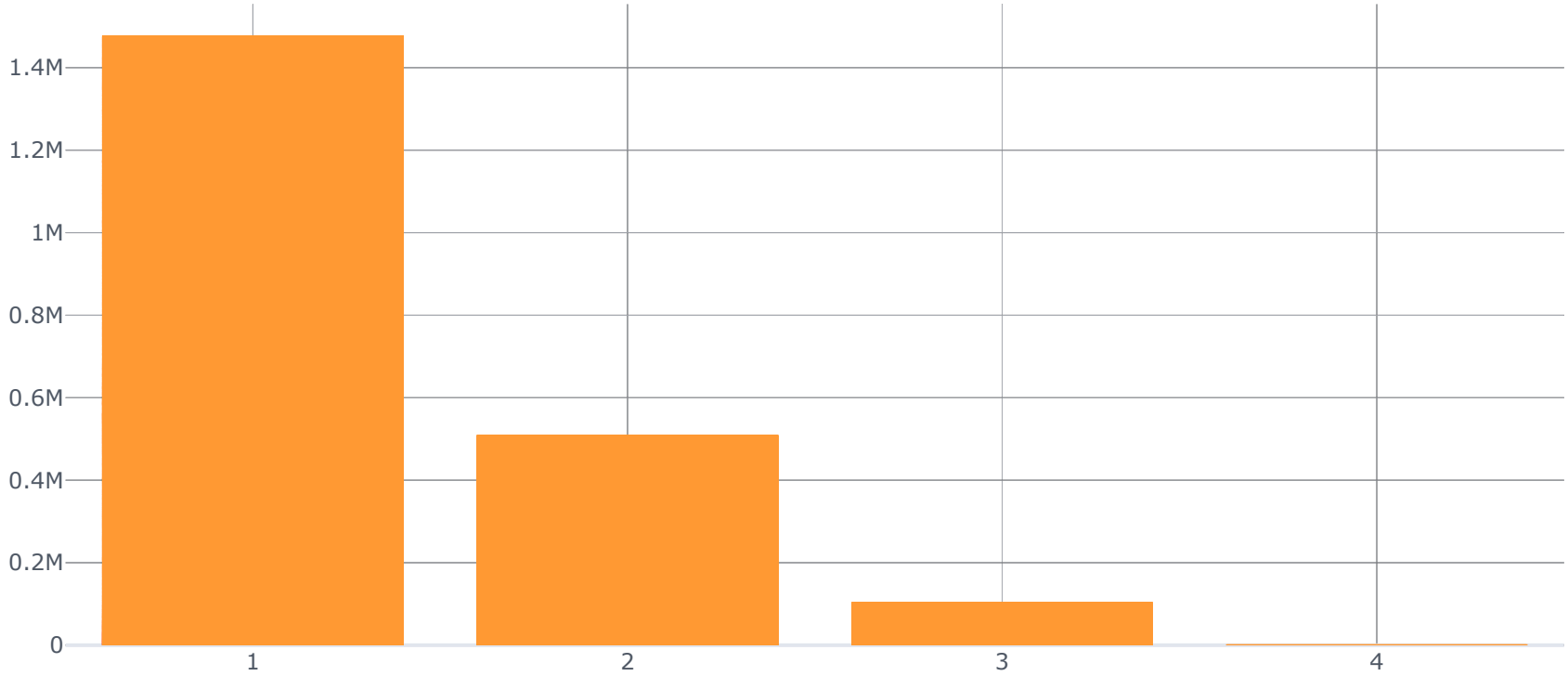
In [12]:
```
bike.iplot(kind="bar", x= "season", y= "count")
# insight : 가을(season=3) 에 자전거 대여수가 가장 높고 봄(season=1)에 가장 낮다
```



In [11]:
```
bike.iplot(kind = "scatter", x="atemp", y="count", mode="markers",size=5, color="season")
# insight : 체감온도 10도 이하와 40도 이상에서는 대여수가 낮다
```



In [21]:
```
bike.iplot(kind="bar", x="weather", y="count")
# insight : 맑은날(1) 대여대수가 가장 많고 비가 오는날(4) 대여대수가 가장 적다
```



In [22]:
```
bike.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 10886 entries, 0 to 10885
Data columns (total 12 columns):
 #   Column      Non-Null Count  Dtype  
---  ------      --------------  -----  
 0   datetime    10886 non-null  object 
 1   season      10886 non-null  int64  
 2   holiday     10886 non-null  int64  
 3   workingday  10886 non-null  int64  
 4   weather     10886 non-null  int64  
 5   temp        10886 non-null  float64
 6   atemp       10886 non-null  float64
 7   humidity    10886 non-null  int64  
 8   windspeed   10886 non-null  float64
 9   casual      10886 non-null  int64  
 10  registered  10886 non-null  int64  
 11  count       10886 non-null  int64  
dtypes: float64(3), int64(8), object(1)
memory usage: 1020.7+ KB
```

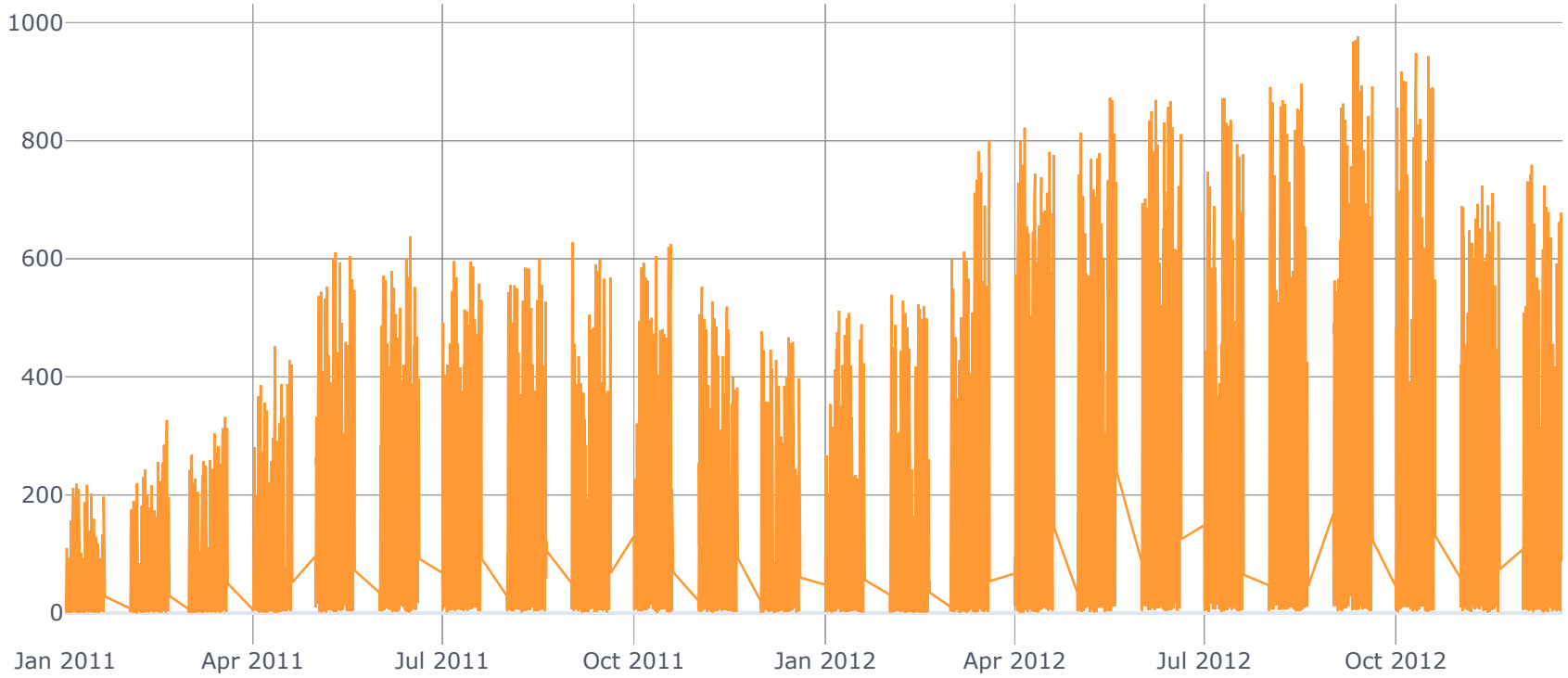
In [23]:
```
import pandas as pd
```

In [24]:
```
bike["datetime"] = pd.to_datetime(bike["datetime"])
```

In [27]:
```
bike.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 10886 entries, 0 to 10885
Data columns (total 12 columns):
 #   Column      Non-Null Count  Dtype         
---  ------      --------------  -----         
 0   datetime    10886 non-null  datetime64[ns]
 1   season      10886 non-null  int64         
 2   holiday     10886 non-null  int64         
 3   workingday  10886 non-null  int64         
 4   weather     10886 non-null  int64         
 5   temp        10886 non-null  float64       
 6   atemp       10886 non-null  float64       
 7   humidity    10886 non-null  int64         
 8   windspeed   10886 non-null  float64       
 9   casual      10886 non-null  int64         
 10  registered  10886 non-null  int64         
 11  count       10886 non-null  int64         
dtypes: datetime64[ns](1), float64(3), int64(8)
memory usage: 1020.7 KB
```

In [28]:
```
bike.iplot(x="datetime", y="count")
# insight : 2011년 대비 2012년 자전거 대여대수는 증가하는 추세를 보인다.
```



In [30]:
```
# 기온, 계절에 따른 대여량
fig = px.scatter(bike, x="temp", y="count", color="season")
fig.show()
```