BayWheel Bike Data Visualziation

In [218]:
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sb
import math
import datetime
%matplotlib notebook
```

In [219]:
```
df = pd.read_csv('201801-fordgobike-tripdata.csv')
df.head()
```

Out[219]:

	duration_sec	start_time	end_time	start_station_id	start_station_name	start_station_latitud
0	75284	2018-01-31 22:52:35.2390	2018-02-01 19:47:19.8240	120	Mission Dolores Park	37.76142
1	85422	2018-01-31 16:13:34.3510	2018-02-01 15:57:17.3100	15	San Francisco Ferry Building (Harry Bridges Pl...	37.79539
2	71576	2018-01-31 14:23:55.8890	2018-02-01 10:16:52.1160	304	Jackson St at 5th St	37.34875
3	61076	2018-01-31 14:53:23.5620	2018-02-01 07:51:20.5000	75	Market St at Franklin St	37.77379
4	39966	2018-01-31 19:52:24.6670	2018-02-01 06:58:31.0530	74	Laguna St at Hayes St	37.77643

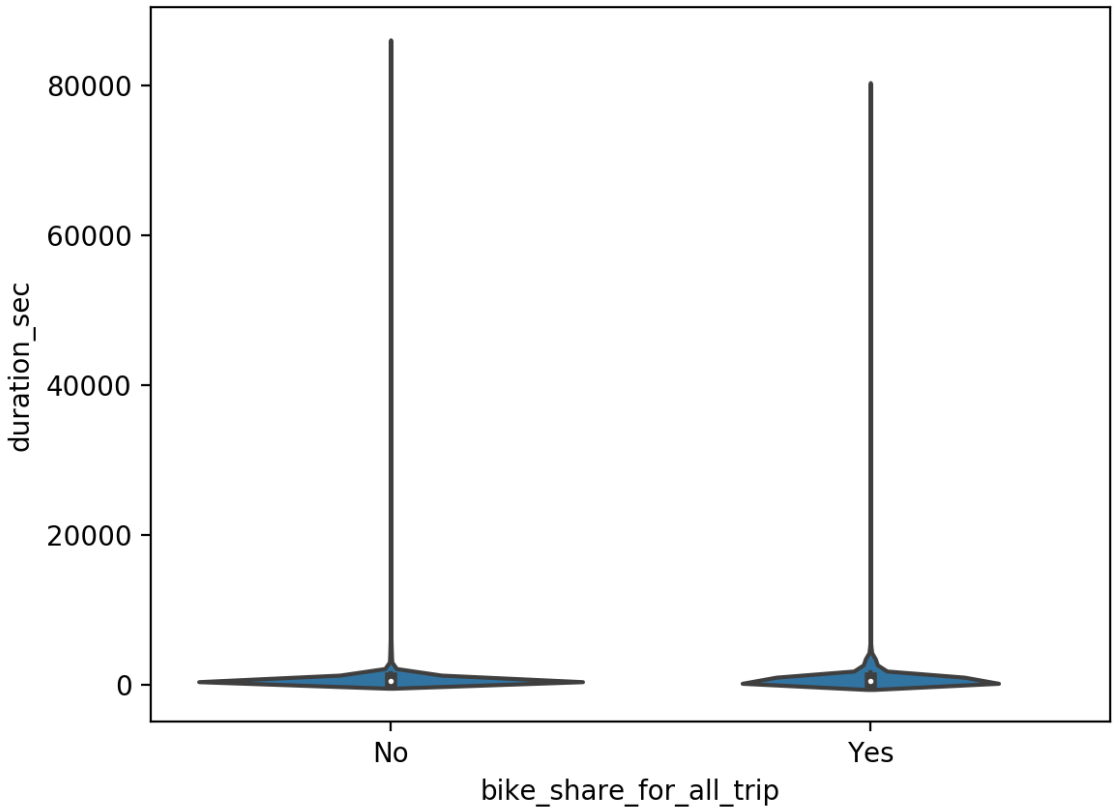
In [220]:
```
df.columns
```

Out[220]:
```
Index(['duration_sec', 'start_time', 'end_time', 'start_station_id',
       'start_station_name', 'start_station_latitude',
       'start_station_longitude', 'end_station_id', 'end_station_name',
       'end_station_latitude', 'end_station_longitude', 'bike_id', 'user_
type',
       'member_birth_year', 'member_gender', 'bike_share_for_all_trip'],
      dtype='object')
```

In [222]:
```
df.bike_share_for_all_trip.value_counts()
```

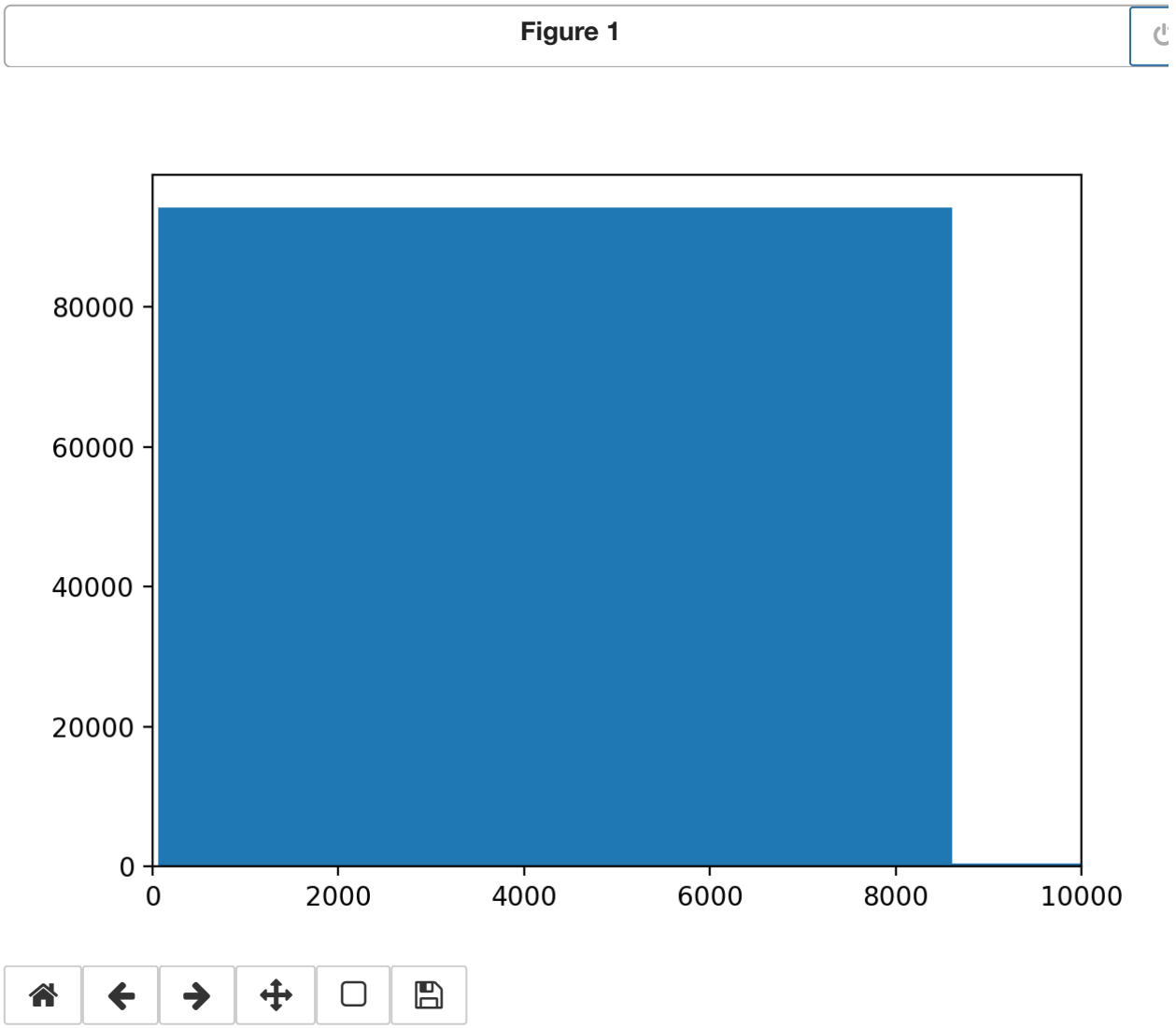
Out[222]:
```
No     88680
Yes     6122
Name: bike_share_for_all_trip, dtype: int64
```

In [227]:
```
base_color = sb.color_palette()[0]
sb.violinplot(data=df, x="bike_share_for_all_trip", y="duration_sec", color
```



Out[227]: <matplotlib.axes._subplots.AxesSubplot at 0x1a2cfc7128>

In [228]:
```
plt.hist(df.duration_sec)
plt.xlim(0, 10000)
```



Out[228]: (0, 10000)

In [134]:
```
df['duration_min'] = df['duration_sec'].apply(lambda d: math.floor(d/60)).a
```