In [2]:
```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
import plotly.offline
from plotly.offline import init_notebook_mode
init_notebook_mode(connected=True)
import cufflinks as cf
cf.go_offline()
```

In [3]:
```
record = pd.read_csv(r"C:\Users\GRACE\Desktop\TECH1M\train.csv")
record
```

```
C:\Users\GRACE\AppData\Local\Temp\ipykernel_8840\3449231188.py:1: DtypeWarning:

Columns (7) have mixed types. Specify dtype option on import or set low_memory=False.
```

Out[3]:

	Store	DayOfWeek	Date	Sales	Customers	Open	Promo	StateHoliday	SchoolHoliday
0	1	5	2015-07-31	5263	555	1	1	0	1
1	2	5	2015-07-31	6064	625	1	1	0	1
2	3	5	2015-07-31	8314	821	1	1	0	1
3	4	5	2015-07-31	13995	1498	1	1	0	1
4	5	5	2015-07-31	4822	559	1	1	0	1
...	...	...	...	...	...	...	...	...	...
1017204	1111	2	2013-01-01	0	0	0	0	a	1
1017205	1112	2	2013-01-01	0	0	0	0	a	1
1017206	1113	2	2013-01-01	0	0	0	0	a	1
1017207	1114	2	2013-01-01	0	0	0	0	a	1
1017208	1115	2	2013-01-01	0	0	0	0	a	1

1017209 rows × 9 columns

In [4]:
```
promo_in_2015 = record.loc[(record["Promo"]==1) & (record["Date"]=="2015-07-31")]
promo_in_2015
```

Out[4]:

	Store	DayOfWeek	Date	Sales	Customers	Open	Promo	StateHoliday	SchoolHoliday
0	1	5	2015-07-31	5263	555	1	1	0	1
1	2	5	2015-07-31	6064	625	1	1	0	1
2	3	5	2015-07-31	8314	821	1	1	0	1
3	4	5	2015-07-31	13995	1498	1	1	0	1
4	5	5	2015-07-31	4822	559	1	1	0	1
...	...	...	...	...	...	...	...	...	...
1110	1111	5	2015-07-31	5723	422	1	1	0	1
1111	1112	5	2015-07-31	9626	767	1	1	0	1
1112	1113	5	2015-07-31	7289	720	1	1	0	1
1113	1114	5	2015-07-31	27508	3745	1	1	0	1
1114	1115	5	2015-07-31	8680	538	1	1	0	1

1115 rows × 9 columns

In [5]:
```
without_promo_in_2015 = record.loc[(record["Promo"]==0) & (record["Date"]=="2015-01-01")]
without_promo_in_2015
```

Out[5]:

	Store	DayOfWeek	Date	Sales	Customers	Open	Promo	StateHoliday	SchoolHoliday
235265	1	4	2015-01-01	0	0	0	0	a	1
235266	2	4	2015-01-01	0	0	0	0	0	1
235267	3	4	2015-01-01	0	0	0	0	a	1
235268	4	4	2015-01-01	0	0	0	0	a	1
235269	5	4	2015-01-01	0	0	0	0	a	1
...	...	...	...	...	...	...	...	...	...
236375	1111	4	2015-01-01	0	0	0	0	a	1
236376	1112	4	2015-01-01	0	0	0	0	a	1
236377	1113	4	2015-01-01	0	0	0	0	a	1
236378	1114	4	2015-01-01	0	0	0	0	a	1
236379	1115	4	2015-01-01	0	0	0	0	a	1

1115 rows × 9 columns

In [6]:
```
promo_in_2013 = record.loc[(record["Promo"]==1) & (record["Date"]=="2013-07-31")]
promo_in_2013
```

Out[6]:

	Store	DayOfWeek	Date	Sales	Customers	Open	Promo	StateHoliday	SchoolHoliday
780830	1	3	2013-07-31	5572	624	1	1	0	1
780831	2	3	2013-07-31	8295	942	1	1	0	1
780832	3	3	2013-07-31	10895	1036	1	1	0	1
780833	4	3	2013-07-31	13403	1657	1	1	0	1
780834	5	3	2013-07-31	8061	788	1	1	0	1
...	...	...	...	...	...	...	...	...	...
781940	1111	3	2013-07-31	6428	529	1	1	0	1
781941	1112	3	2013-07-31	14127	964	1	1	0	1
781942	1113	3	2013-07-31	8423	801	1	1	0	1
781943	1114	3	2013-07-31	26468	4042	1	1	0	1
781944	1115	3	2013-07-31	6209	421	1	1	0	1

1115 rows × 9 columns

In [7]:
```
without_promo_in_2013 = record.loc[(record["Promo"]==0) & (record["Date"]=="2013-01-01")]
without_promo_in_2013
```

Out[7]:

	Store	DayOfWeek	Date	Sales	Customers	Open	Promo	StateHoliday	SchoolHoliday
1016095	1	2	2013-01-01	0	0	0	0	a	1
1016096	2	2	2013-01-01	0	0	0	0	a	1
1016097	3	2	2013-01-01	0	0	0	0	a	1
1016098	4	2	2013-01-01	0	0	0	0	a	1
1016099	5	2	2013-01-01	0	0	0	0	a	1
...	...	...	...	...	...	...	...	...	...
1017204	1111	2	2013-01-01	0	0	0	0	a	1
1017205	1112	2	2013-01-01	0	0	0	0	a	1
1017206	1113	2	2013-01-01	0	0	0	0	a	1
1017207	1114	2	2013-01-01	0	0	0	0	a	1
1017208	1115	2	2013-01-01	0	0	0	0	a	1

1114 rows × 9 columns

In [8]:
```
promo_in_2015.groupby("Date").sum().iloc[:,2:4]
#Total sales and Total customers after promo in 2015
```

Out[8]:

Date	Sales	Customers
2015-07-31	10109742	968620

In [9]:
```
without_promo_in_2015.groupby("Date").sum().iloc[:,2:4]
#Total sales and Total customers after no promo in 2015
```

Out[9]:

Date	Sales	Customers
2015-01-01	130162	23049

In [22]:
```
promo_in_2013.groupby("Date").sum().iloc[:,2:4]
#Total sales and Total customers after promo in 2013
```

Out[22]:

Date	Sales	Customers
2013-07-31	10293490	1012667

In [10]:
```
without_promo_in_2013.groupby("Date").sum().iloc[:,2:4]
#Total sales and Total customers after promo in 2015
```

Out[10]:

Date	Sales	Customers
2013-01-01	97235	19491

In [13]:
```
a = promo_in_2015.groupby("Date").sum().iloc[:,2:4]
a.insert(2, "Date", "2015-07-31")
```

In [14]:
```
b = without_promo_in_2015.groupby("Date").sum().iloc[:,2:4]
b.insert(2, "Date", "2015-01-01")
```

In [15]:
```
c=promo_in_2013.groupby("Date").sum().iloc[:,2:4]
c.insert(2, "Date", "2013-07-31")
```

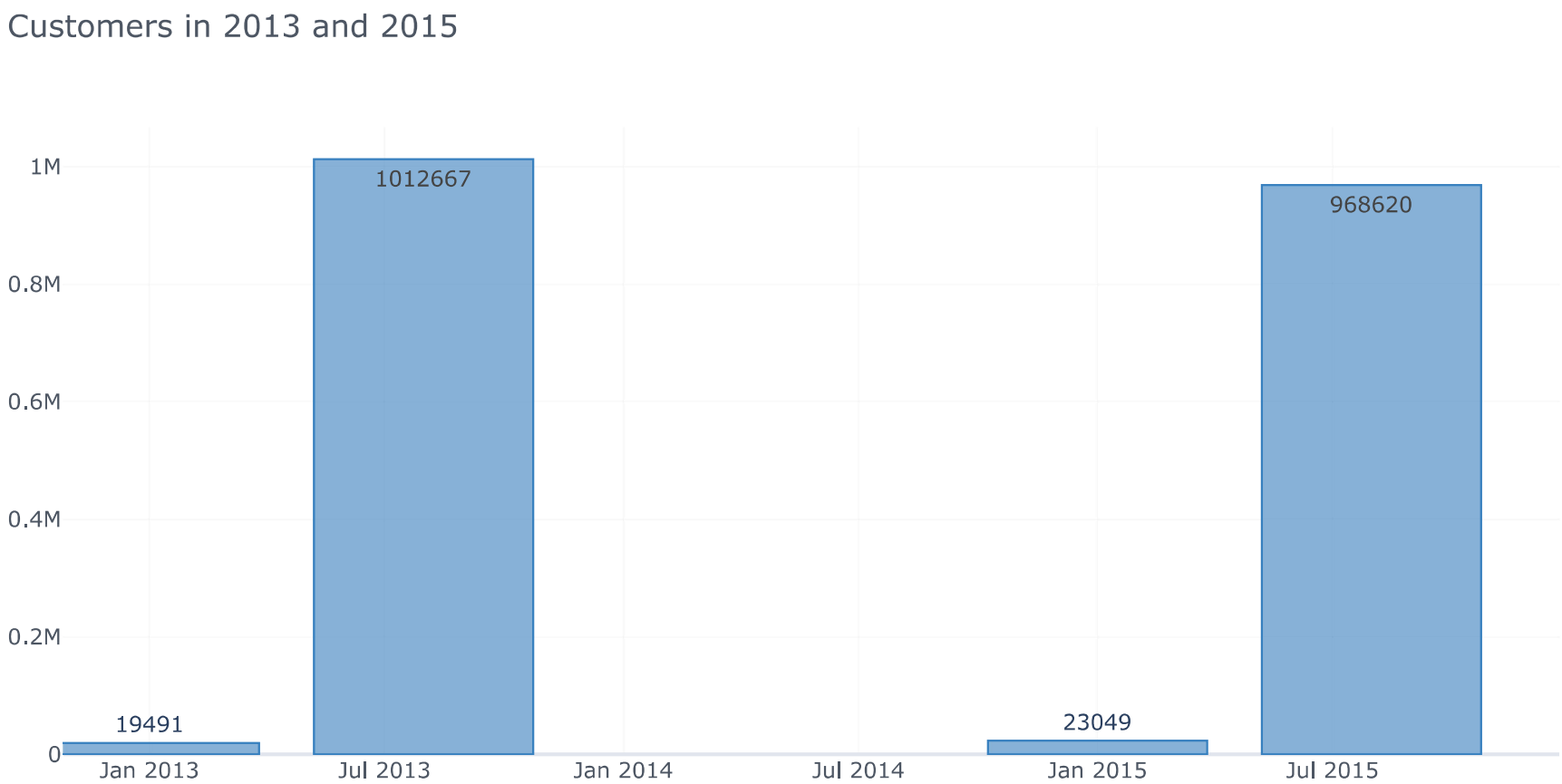
In [16]:
```
d=without_promo_in_2013.groupby("Date").sum().iloc[:,2:4]
d.insert(2, "Date", "2013-01-01")
```

In [17]:
```
integrated_grouped = pd.concat([a,b,c,d])
integrated_grouped
```

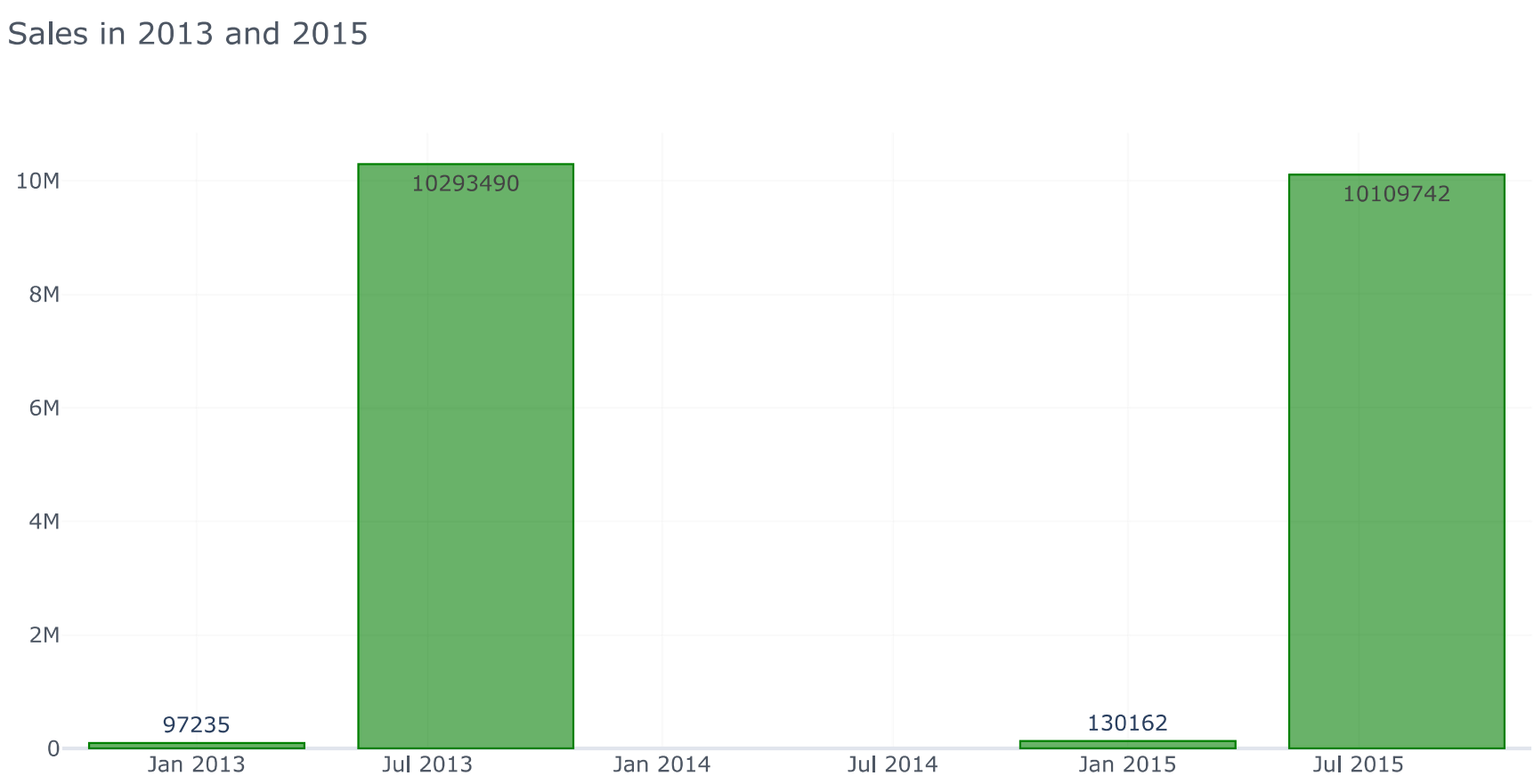
Out[17]:

Date	Sales	Customers	Date
2015-07-31	10109742	968620	2015-07-31
2015-01-01	130162	23049	2015-01-01
2013-07-31	10293490	1012667	2013-07-31
2013-01-01	97235	19491	2013-01-01

In [20]:
```
integrated_grouped.iplot(kind="bar", x="Date", y="Customers", text="Customers", title="Customers in 2013 and 2015",
#customers based on promotion in 2015(with promo 968620, withoutpromo 23049) & 2013(with promo 1012667,without promo
```



In [21]:
```
integrated_grouped.iplot(kind="bar", x="Date", y="Sales", text="Sales", title="Sales in 2013 and 2015", color="green
#sales based on promotion in 2015(withpromo 10109742, withoutpromo 130162) & 2013(withpromo 10293490,withoutpromo 97
```



In [ ]:
```

```