

# Explore\_bikeshare\_data

June 2, 2021

```
In [ ]: ### Explore Bike Share Data
```

For this project, your goal is to ask and answer three questions about the available bikeshare data.

You will be graded against the project [Rubric](<https://review.udacity.com/#!/rubrics/2563/view>)

```
In [4]: ny = read.csv('new_york_city.csv')
wash = read.csv('washington.csv')
chi = read.csv('chicago.csv')
```

```
In [3]: head(ny)
```

X	Start.Time	End.Time	Trip.Duration	Start.Station	End.Station
5688089	2017-06-11 14:55:05	2017-06-11 15:08:21	795	Suffolk St & Stanton St	W Broadway St
4096714	2017-05-11 15:30:11	2017-05-11 15:41:43	692	Lexington Ave & E 63 St	1 Ave & E 7 St
2173887	2017-03-29 13:26:26	2017-03-29 13:48:31	1325	1 Pl & Clinton St	Henry St & E 12 St
3945638	2017-05-08 19:47:18	2017-05-08 19:59:01	703	Barrow St & Hudson St	W 20 St & E 8 St
6208972	2017-06-21 07:49:16	2017-06-21 07:54:46	329	1 Ave & E 44 St	E 53 St & 3 Ave
1285652	2017-02-22 18:55:24	2017-02-22 19:12:03	998	State St & Smith St	Bond St & E 4 St

```
In [2]: head(wash)
```

X	Start.Time	End.Time	Trip.Duration	Start.Station
1621326	2017-06-21 08:36:34	2017-06-21 08:44:43	489.066	14th & Belmont St NW
482740	2017-03-11 10:40:00	2017-03-11 10:46:00	402.549	Yuma St & Tenley Circle NW
1330037	2017-05-30 01:02:59	2017-05-30 01:13:37	637.251	17th St & Massachusetts Ave NW
665458	2017-04-02 07:48:35	2017-04-02 08:19:03	1827.341	Constitution Ave & 2nd St NW/DOL
1481135	2017-06-10 08:36:28	2017-06-10 09:02:17	1549.427	Henry Bacon Dr & Lincoln Memorial
1148202	2017-05-14 07:18:18	2017-05-14 07:24:56	398.000	1st & K St SE

```
In [4]: head(chi)
```

X	Start.Time	End.Time	Trip.Duration	Start.Station	End.Station
1423854	2017-06-23 15:09:32	2017-06-23 15:14:53	321	Wood St & Hubbard St	Dan
955915	2017-05-25 18:19:03	2017-05-25 18:45:53	1610	Theater on the Lake	She
9031	2017-01-04 08:27:49	2017-01-04 08:34:45	416	May St & Taylor St	Wo
304487	2017-03-06 13:49:38	2017-03-06 13:55:28	350	Christiana Ave & Lawrence Ave	St.
45207	2017-01-17 14:53:07	2017-01-17 15:02:01	534	Clark St & Randolph St	De
1473887	2017-06-26 09:01:20	2017-06-26 09:11:06	586	Clinton St & Washington Blvd	Car

## 0.0.1 Question 1

Your question 1 goes here.

```
In [ ]: What 's the common month?
```

```
In [1]: library(readr);  
library(tidyr);  
library(plyr)  
library(dplyr);  
library(ggplot2)
```

Attaching package: dplyr

The following objects are masked from package:plyr:

```
arrange, count, desc, failwith, id, mutate, rename, summarise,  
summarize
```

The following objects are masked from package:stats:

```
filter, lag
```

The following objects are masked from package:base:

```
intersect, setdiff, setequal, union
```

```
In [2]: chi <- read.csv('chicago.csv')  
ny <- read.csv('new_york_city.csv')  
wash <- read.csv('washington.csv')
```

```
In [3]: chi$Start.Month <- format(as.Date(chi$Start.Time), '%m')  
ny$Start.Month <- format(as.Date(ny$Start.Time), '%m')  
wash$Start.Month <- format(as.Date(wash$Start.Time), '%m')
```

```
In [4]: month.sum <- function(x){  
  df <- x %>% group_by(month=x$Start.Month) %>% filter(!is.na(month)) %>% summarise(mo  
  return(df)  
}  
chi_month <- month.sum(chi)  
chi_month  
ny_month <- month.sum(ny)  
ny_month  
wash_month <- month.sum(wash)  
wash_month
```

month	mon.count
01	650
02	930
03	803
04	1526
05	1905
06	2816
month	mon.count
01	5745
02	6364
03	5820
04	10661
05	12180
06	14000
month	mon.count
01	8946
02	11563
03	12612
04	18522
05	17072
06	20335

```
In [5]: df <- data.frame(Month=chi_month$month, Chicago=chi_month$mon.count, New_York=ny_month$mon.count, Washington=wa_month$mon.count)
df
```

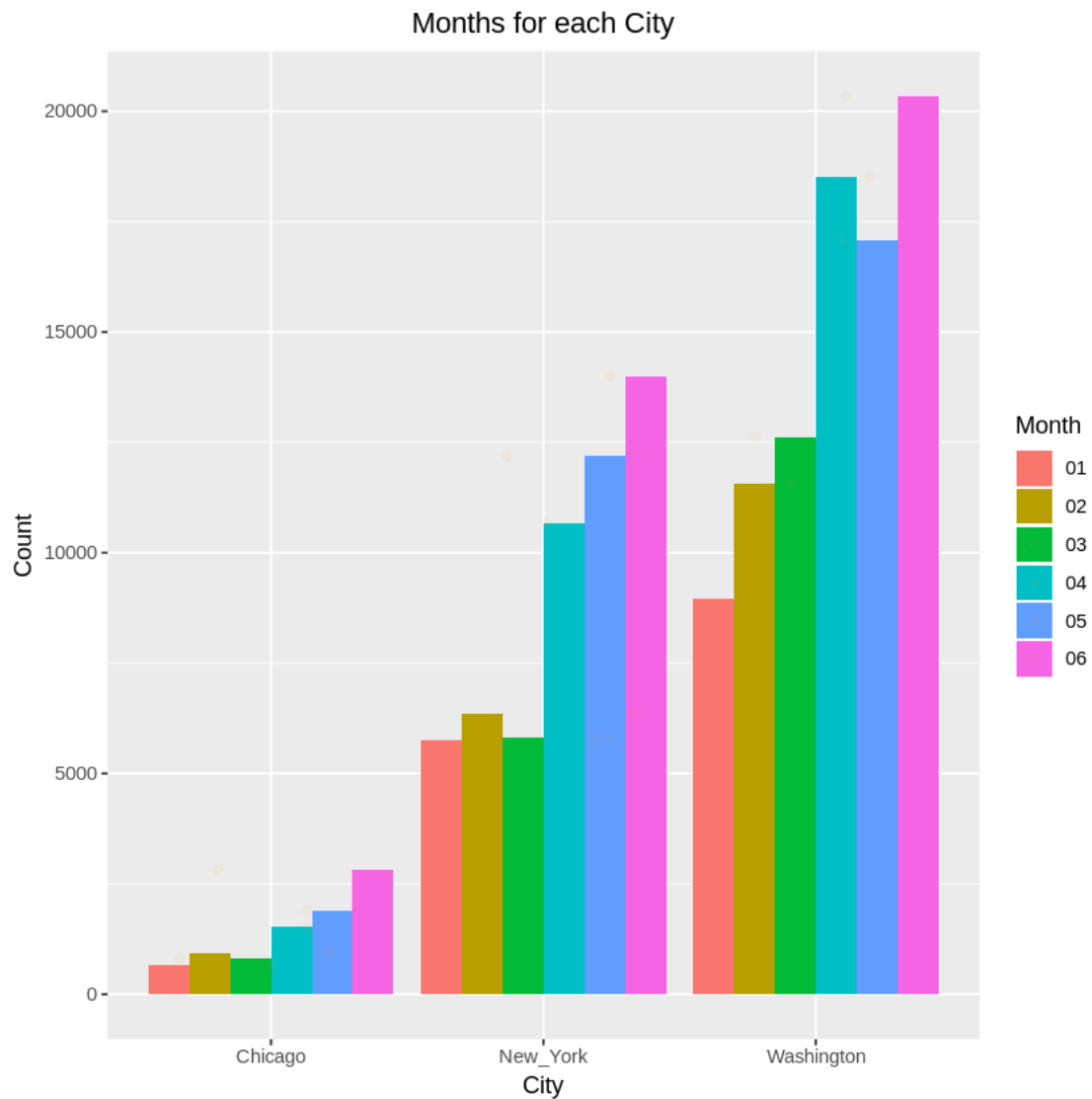
Month	Chicago	New_York	Washington
01	650	5745	8946
02	930	6364	11563
03	803	5820	12612
04	1526	10661	18522
05	1905	12180	17072
06	2816	14000	20335

```
In [6]: newdf <- df %>% gather(key=City, value=Count, Chicago, New_York, Washington)
newdf
```

Month	City	Count
01	Chicago	650
02	Chicago	930
03	Chicago	803
04	Chicago	1526
05	Chicago	1905
06	Chicago	2816
01	New_York	5745
02	New_York	6364
03	New_York	5820
04	New_York	10661
05	New_York	12180
06	New_York	14000
01	Washington	8946
02	Washington	11563
03	Washington	12612
04	Washington	18522
05	Washington	17072
06	Washington	20335

```
In [11]: common_month_plot <- ggplot(newdf, aes(City, Count, fill=Month)) + geom_col(position =
```

```
In [12]: common_month_plot
```



```
In [49]: mean_june <- mean(c(df[6,2],df[6,3],df[6,4]))
         mean_june
```

12383.6666666667

**Summary of your question 1 results goes here.**

In [ ]: From the above bar plot, we can conclude that the common month for all three cities is J

## 0.0.2 Question 2

**Your question 2 goes here.**

```
In [45]: What's the average travel time in all the three cities
```

```
Error in parse(text = x, srcfile = src): <text>:1:5: unexpected INCOMPLETE_STRING
1: What's the average travel time in all the three cities
  ^
```

Traceback:

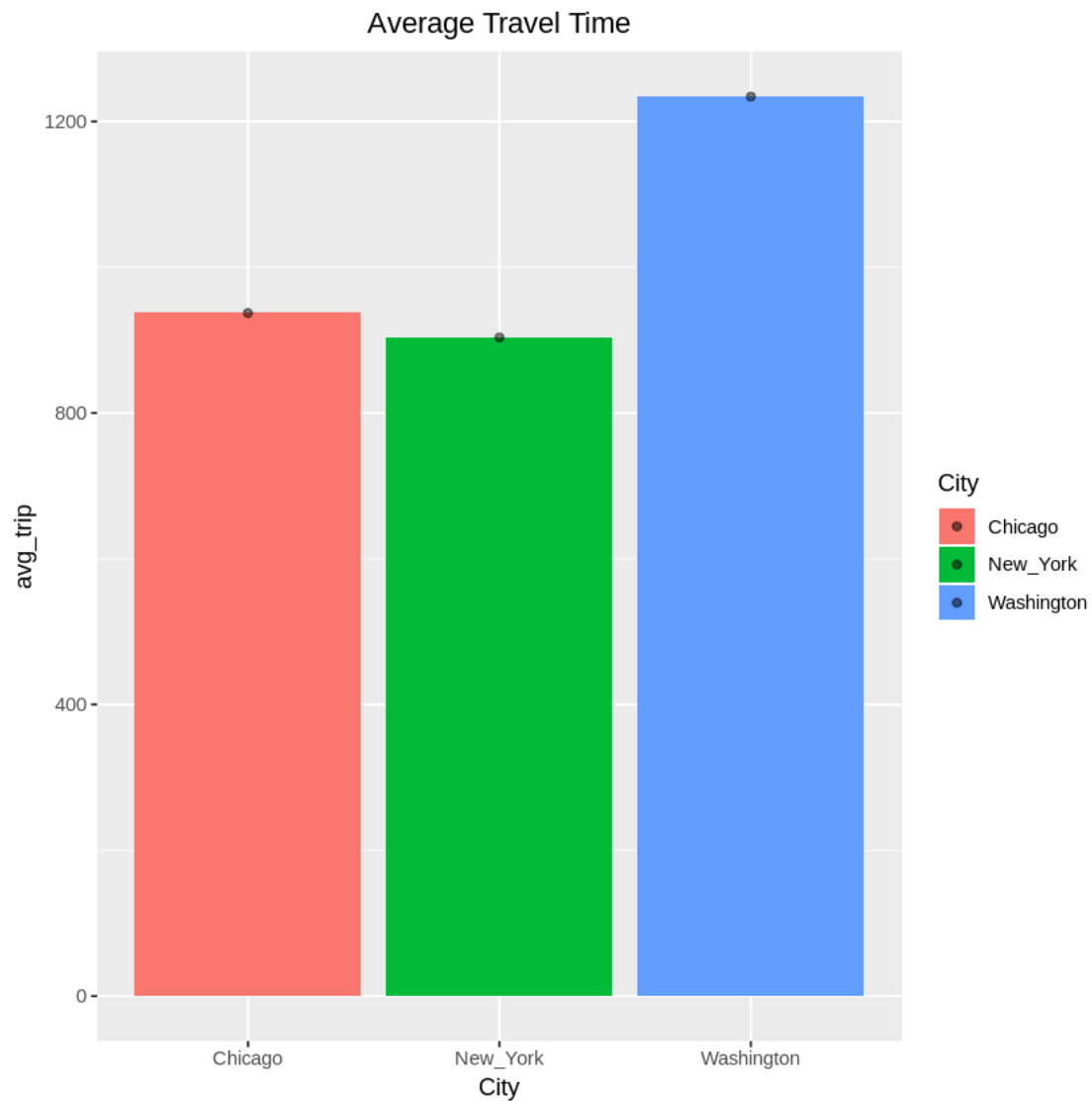
```
In [13]: ny = read.csv('new_york_city.csv')
        wash = read.csv('washington.csv')
        chi = read.csv('chicago.csv')

In [14]: library(readr);
        library(tidyr);
        library(plyr);
        library(dplyr);
        library(ggplot2)

In [15]: trip_dur <- list(chi$Trip.Duration, ny$Trip.Duration[!is.na(ny$Trip.Duration)], wash$Tr
        avg_trip <- lapply(trip_dur, mean)
        avg_trip

1. 937.172769409038
2. 903.614690792237
3. 1233.95333709152

In [16]: df <- data.frame(City=c('Chicago', 'New_York', 'Washington'), avg_trip=c(avg_trip[[1]],
In [20]: ggplot(df, aes(City, avg_trip, fill=City)) + geom_col(position='dodge') + geom_point(al
```



In [10]: In sum, average trip timr in Washington is the highest amoung all three cities.And mean

```
Error in parse(text = x, srcfile = src): <text>:1:4: unexpected symbol
1: In sum
  ^
Traceback:
```

Summary of your question 2 results goes here.

### 0.03 Question 3

Your question 3 goes here.

In [ ]: What are the counts of each `gender` (only available for NYC and Chicago)?

```
In [21]: ny <- read.csv('new_york_city.csv')
        chi <- read.csv('chicago.csv')
```

```
In [22]: library(readr);
        library(tidyr);
        library(dplyr);
        library(ggplot2)
```

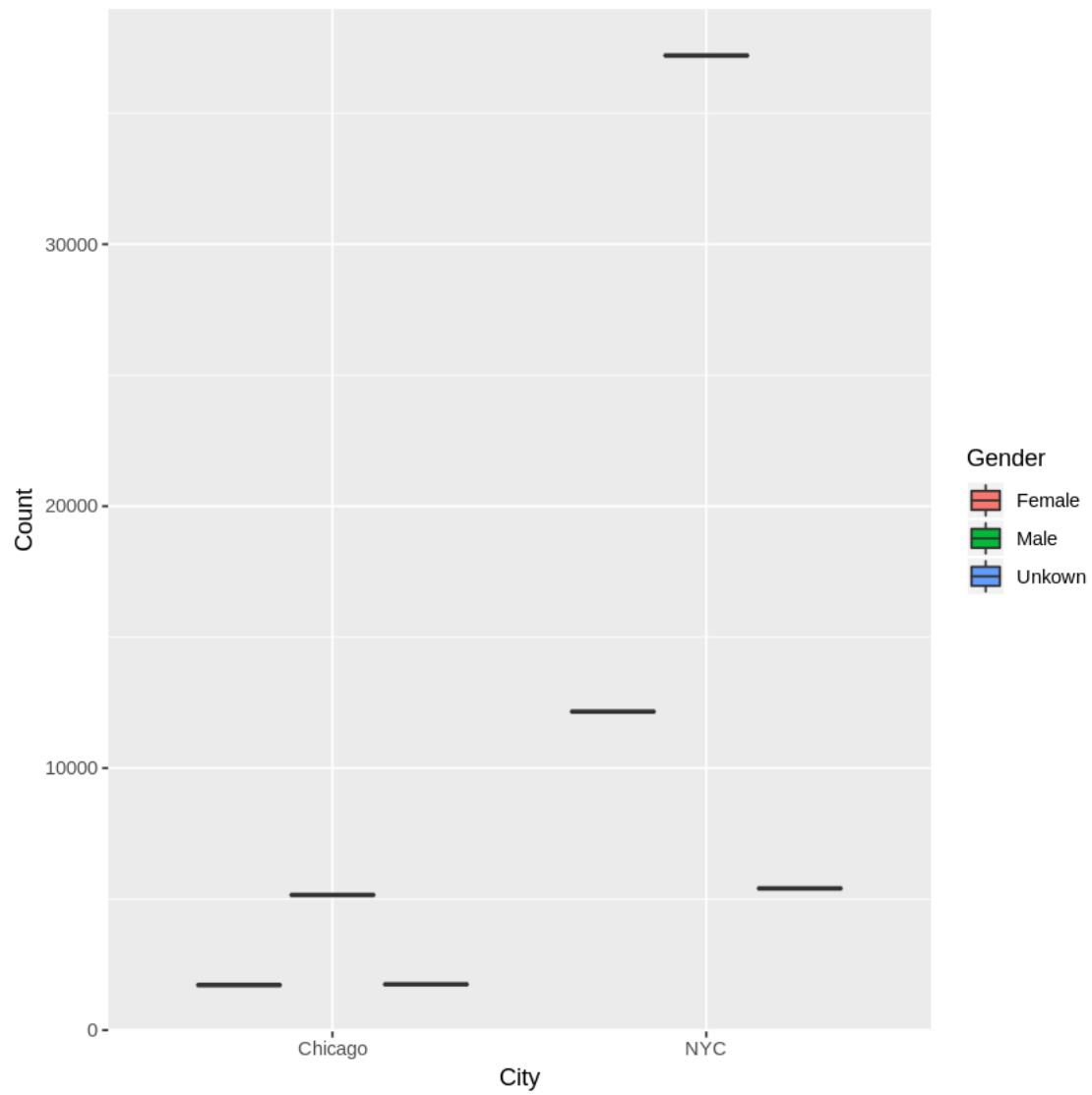
```
In [23]: gender_ct_chi <- chi %>% group_by(Gender) %>% summarize(chi.gender.ct=n())
        gender_ct_chi
        gender_ct_ny <- ny %>% group_by(Gender) %>% summarize(ny.gender.ct=n())
        gender_ct_ny
```

Gender	chi.gender.ct
	1748
Female	1723
Male	5159
Gender	ny.gender.ct
	5410
Female	12159
Male	37201

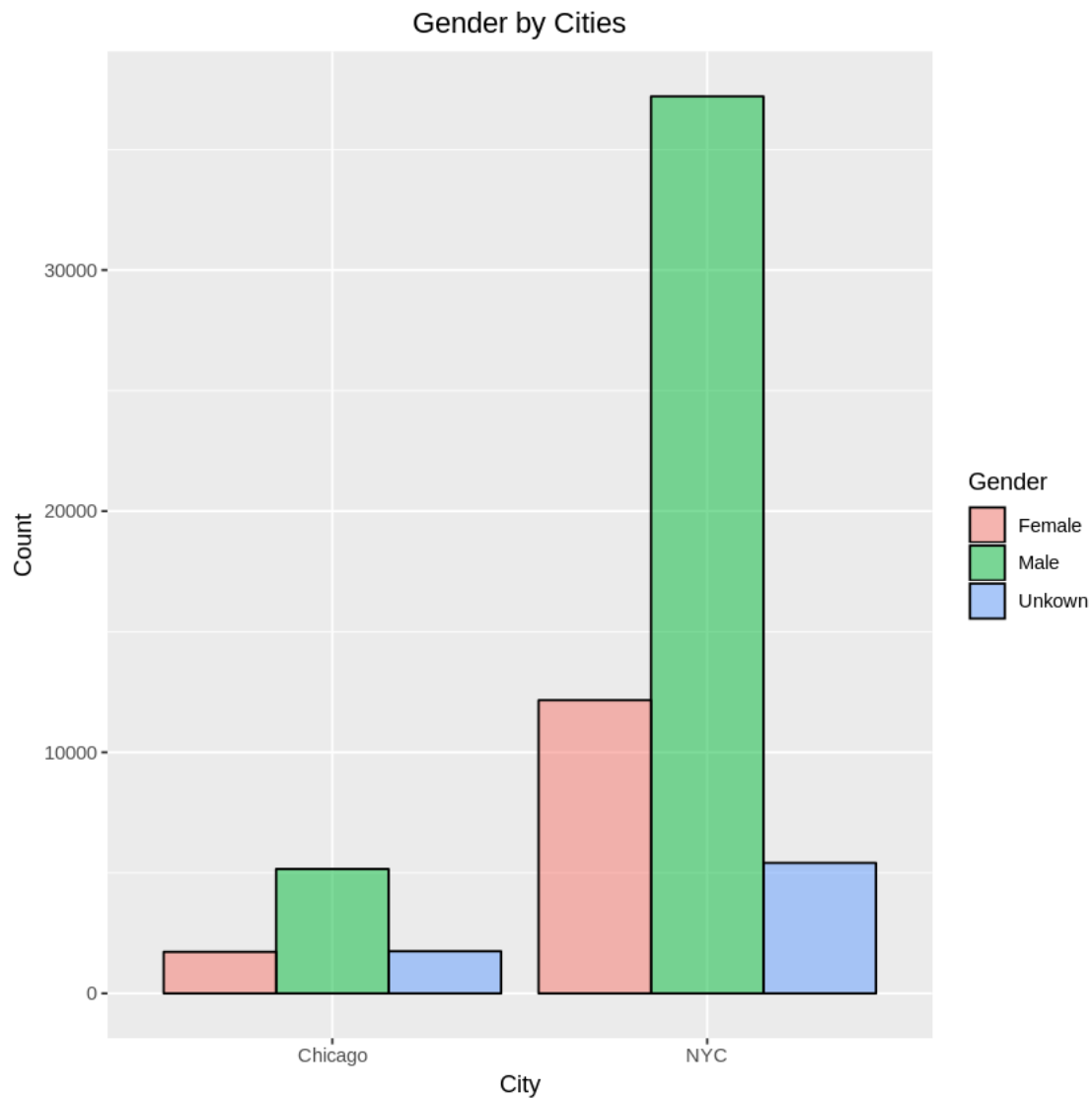
```
In [25]: df <- data.frame(Gender=c('Unkown', 'Female', 'Male'), NYC=gender_ct_ny$ny.gender.ct, C
        newdf <- df %>% gather(key=City, value=Count, NYC, Chicago)
```

```
In [28]:
```





```
In [31]: ggplot(newdf, aes(City, Count, fill=Gender)) + geom_col(position='dodge', alpha=0.5, co
```



In [27]: The result is obviously contrasting, the common feature in two cities is that the number

```
Error in parse(text = x, srcfile = src): <text>:1:5: unexpected symbol
1: The result
   ^
Traceback:
```

**Summary of your question 3 results goes here.**

## 0.1 Finishing Up

Congratulations! You have reached the end of the Explore Bikeshare Data Project. You should be very proud of all you have accomplished!

**Tip:** Once you are satisfied with your work here, check over your report to make sure that it satisfies all the areas of the [rubric](#).

## 0.2 Directions to Submit

Before you submit your project, you need to create a .html or .pdf version of this notebook in the workspace here. To do that, run the code cell below. If it worked correctly, you should get a return code of 0, and you should see the generated .html file in the workspace directory (click on the orange Jupyter icon in the upper left).

Alternatively, you can download this report as .html via the **File > Download as** sub-menu, and then manually upload it into the workspace directory by clicking on the orange Jupyter icon in the upper left, then using the Upload button.

Once you've done this, you can submit your project by clicking on the "Submit Project" button in the lower right here. This will create and submit a zip file with this .ipynb doc and the .html or .pdf version you created. Congratulations!

```
In [11]: system('python -m nbconvert Explore_bikeshare_data.ipynb')
```