

**Prakash Singh Sijwali**

Divvy Bicycles Converting Data to Discovery

**Executive Summary**

New age technologies, with analytics at core, are disrupting the business across verticals. Industries are reimagining their business with these technologies at faster pace. Businesses who realized the potential early are seeing exponential growth fuelled by enhanced and personalized customer engagement. At the centre (or Core) of this disruption, “Data” is the driving factor. Businesses now understand that with right utilization of “Data”, many useful insights can be uncovered resulting in paradigm shift.

“Divvy Bikes” is one of the leading players in bikes renting business, based out at USA. For any TTH (Travel, Transport and Hospitality) business, Customer Satisfaction is key. In context of Divvy Bikes, key factor for customer satisfaction is “availability of bikes” at customer disposal i.e. strengthening their forecasting and supply chain.

Using historical data, we have uncovered some interesting insights, which could be leveraged to create focused approach in attaining the key goals.

**Project Description**

Divvy bikes has system which records parameters of each and every trip (Bike rent) originated online or at renting station. As a result, Divvy Bikes has been able to create a rich historical data, which can be leveraged to identify customer renting patterns and most importantly forecast the Demand and innovative ways of maintaining the supply chain. Along with this, descriptive stat on historical data can uncover parameters like busiest rent station, what is customer base i.e. subscribers or pay as your go model etc.

We ingested 3 years of Rent data (~9.5 Million records – source: Kaggle), each data record having following vital information:

* Trip Year
* Trip Month
* Week of the year
* Weekday/Weekend
* Customer Type – Subscriber or Pay as you go
* Gender of customer
* Start timestamp of trip
* End timestamp of trip
* Trip Duration
* Temperature (in Fahrenheit) during Trip
* Weather during the trip
* Trip Start station ID and Name
* Trip End station ID and Name
* Trip Start Station Latitude and Longitude (Geo-codes)
* Trip Start Station Capacity
* Trip End Station Latitude and Longitude (Geo-codes)
* Trip End Station Capacity

Total size of data: 1.9 GB

Format of Data: CSV file.

**Business Solution Themes**

Using the above historical data, we have uncovered following business insights along with appropriate business recommendations:

1. **Descriptive analysis, uncovering following key statistics:**

* Which are top 10 stations generating highest amount of trips
* What is yearly trend of top 5 stations generating trips
* Top 10 stations with highest Average Trip time
* Is Busiest station generating highest Average tip?
* Top 5 most preferred routes by customer for bike trip
* Trend of Preferred routes on yearly basis
* Top 10 Bike capacity station – compare with highest trip generating stations?
* Time of the day (Hour) where most bike renting is happening – split by weekdays and weekends? Is it coinciding with office going times?
* Visualization

1. **Bike Demand and Supply**

* Identify key parameters which influence the pattern of Bike rent (count)
* Create a predictive model of “number of footfalls” based on the key parameters
* Ensure that capacity/supply of bikes is sufficient, based on the predictive model – Optimizing capacity of bikes across stations based on predictive model
* Business recommendations
* Visualization

1. **Opportunities of Business consolidation**

* Bottom 10 stations which are generating least amount of Bike rents
* Least preferred routes by customer
* Identify nearest stations which are doing good in terms of bike rents and exploring opportunities of moving the capacity of “worst performing stations” to nearby “good” ones
* Business recommendations
* Visualization

1. **Customer Base Analysis**

* Subscriber v/s pay-as-you-go customer
* Trend of subscribers – i.e. if more customer becoming subscriber on yearly basis?
* Top 5 and bottom 5 station IDs based on number of subscribers
* Business recommendations based on above
* Visualization

1. **Customer Bike riding pattern based on Gender**

* Who prefers Bike riding more – Male or Female?
* Any trend over years
* Which station IDs have more Gender specific customer
* Business recommendation
* Visualization

**Tools used**

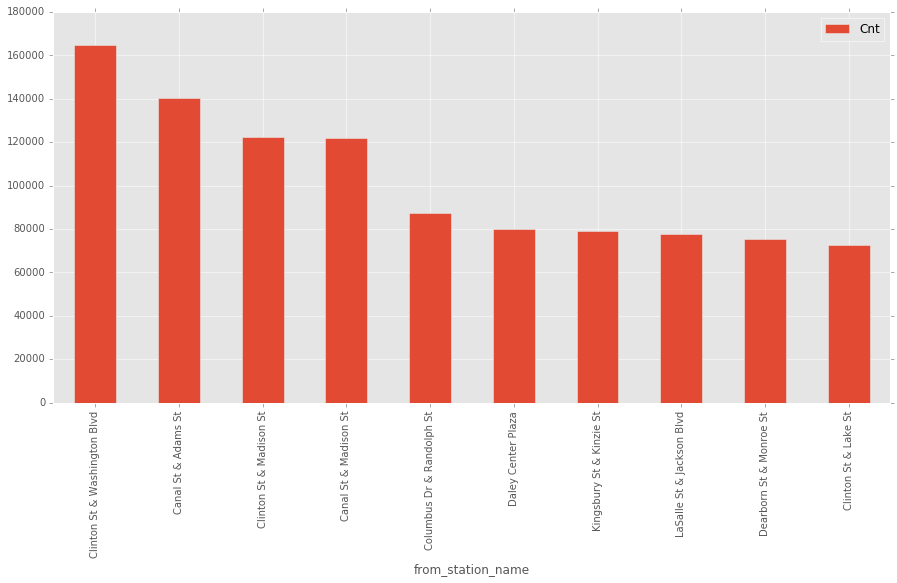
* Spark Core
* Spark SQL
* Spark Mlib
* Spark Graphframe
* Spark Streaming (to Capture Live tweets for Divvy Bikes)

Load the data in to SparkSQL generating auto schema, as depicted in the ipynb file in section 1. Then create view of the data which enable SQL statements to be executed.

**Descriptive Analysis**

1. **Identify top 10 busiest stations:**

**Approach:** On loaded data, selected stations and count of stations group by stations and ordered by descending order of count.

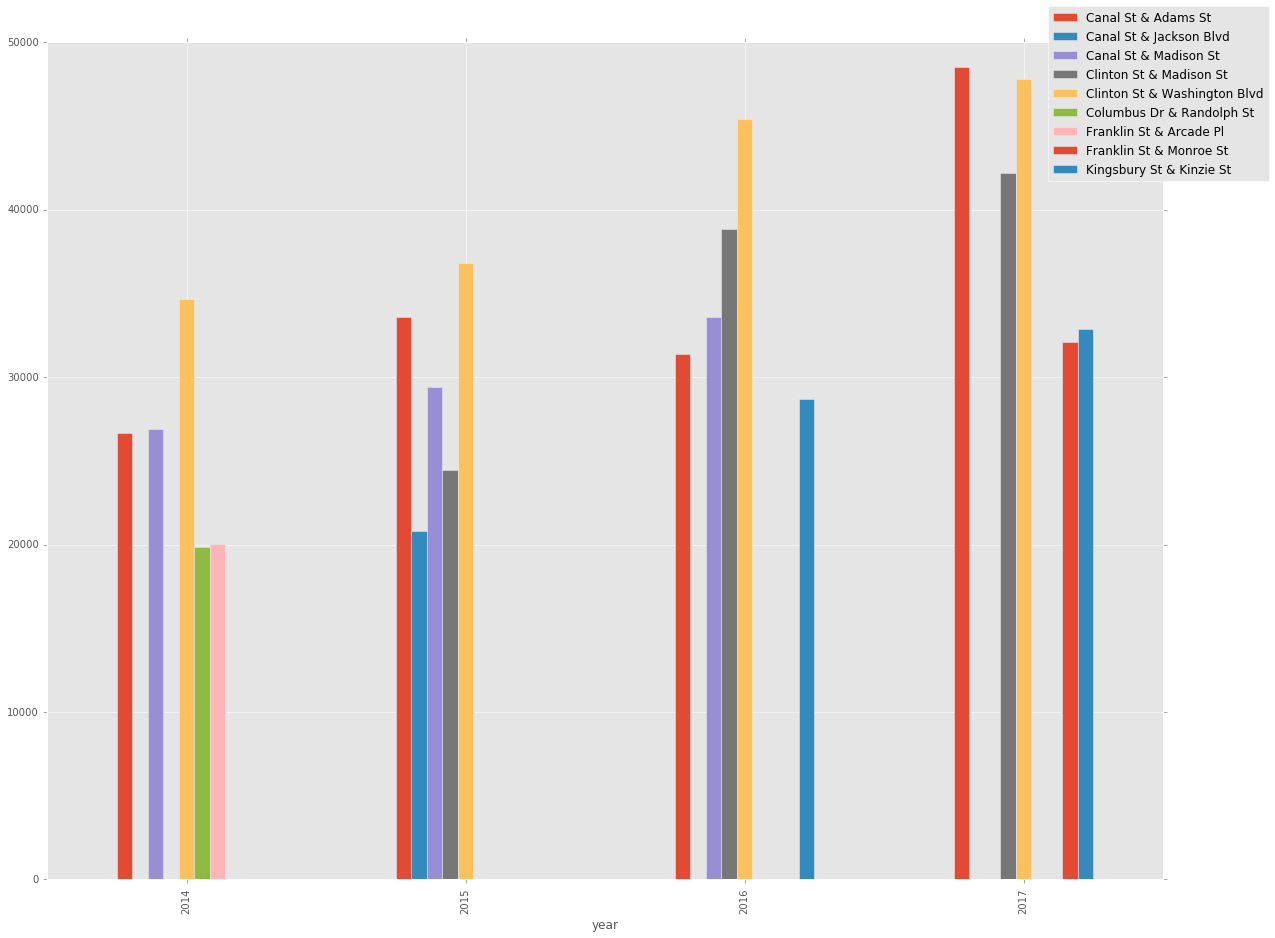


|  |  |
| --- | --- |
| **Station\_ID** | **Trip\_Counts** |
| **Clinton St & Washington Blvd** | **164669** |
| **Canal St & Adams St** | **140197** |
| **Clinton St & Madison St** | **122229** |
| **Canal St & Madison St** | **121728** |
| **Columbus Dr & Randolph St** | **87577** |
| Daley Center Plaza | 80139 |
| Kingsbury St & Kinzie St | 79267 |
| LaSalle St & Jackson Blvd | 77547 |
| Dearborn St & Monroe St | 75224 |
| Clinton St & Lake St | 72823 |

**Clinton St & Washington Blvd is historically busiest station**

1. **Yearly Trend of top 5 busiest station**

**Approach:** On loaded data, selected years stations and count of stations group by year and stations and ordered by descending order of count.



Based on above visualization, we can see that:

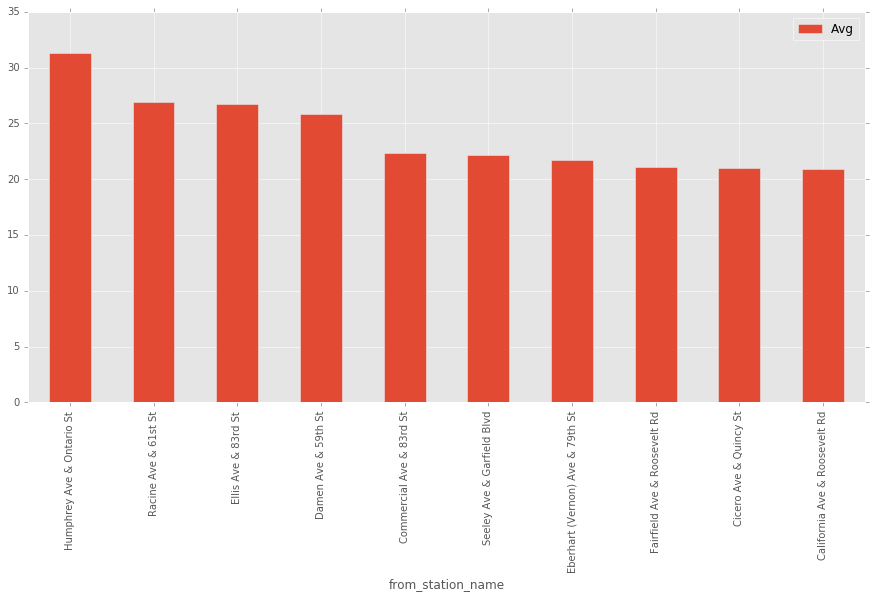
* Clinton St & Washington Blvd station was leading till 2016, however in 2017 there is tremendous growth is seen for Canal St & Adams St station due to which it moved to pole position in 2017.
* Columbus Dr & Randolph St station was in top 5 in 2014, but later moved out from top 5 chart.
* Clinton St & Madison St was in top 5 in 2016, but in 2017 moved out from top 5 chart.

**Business Questions:**

* Analyse growth journey of Canal St & Adams St (2016 v/s 2017) – what worked well, can it be replicated to other stations?
* What went wrong for Columbus Dr & Randolph St and Clinton St & Madison St, why there is decline? Analyse customer sentiments?

1. **Top 10 stations with highest Average Trip time**

**Approach:** On loaded data, selected stations and Average of trip duration group by stations and ordered by descending order of Average.



1. **Is Busiest station also generating more Duration rides?**

|  |  |  |  |
| --- | --- | --- | --- |
| **Top 10 stations by Average Trip Duration** | | **Top 10 Busiest Stations** | |
| **Station\_ID** | **Average Trip Duration** | **Station\_ID** | **Trip\_Counts** |
| Humphrey Ave & Ontario St | 31.31 | Clinton St & Washington Blvd | 164669 |
| Racine Ave & 61st St | 26.93 | Canal St & Adams St | 140197 |
| Ellis Ave & 83rd St | 26.77 | Clinton St & Madison St | 122229 |
| Damen Ave & 59th St | 25.80 | Canal St & Madison St | 121728 |
| Commercial Ave & 83rd St | 22.39 | Columbus Dr & Randolph St | 87577 |
| Seeley Ave & Garfield Blvd | 22.18 | Daley Center Plaza | 80139 |
| Eberhart (Vernon) Ave & 79th St | 21.71 | Kingsbury St & Kinzie St | 79267 |
| Fairfield Ave & Roosevelt Rd | 21.13 | LaSalle St & Jackson Blvd | 77547 |
| Cicero Ave & Quincy St | 21.05 | Dearborn St & Monroe St | 75224 |
| California Ave & Roosevelt Rd | 20.89 | Clinton St & Lake St | 72823 |

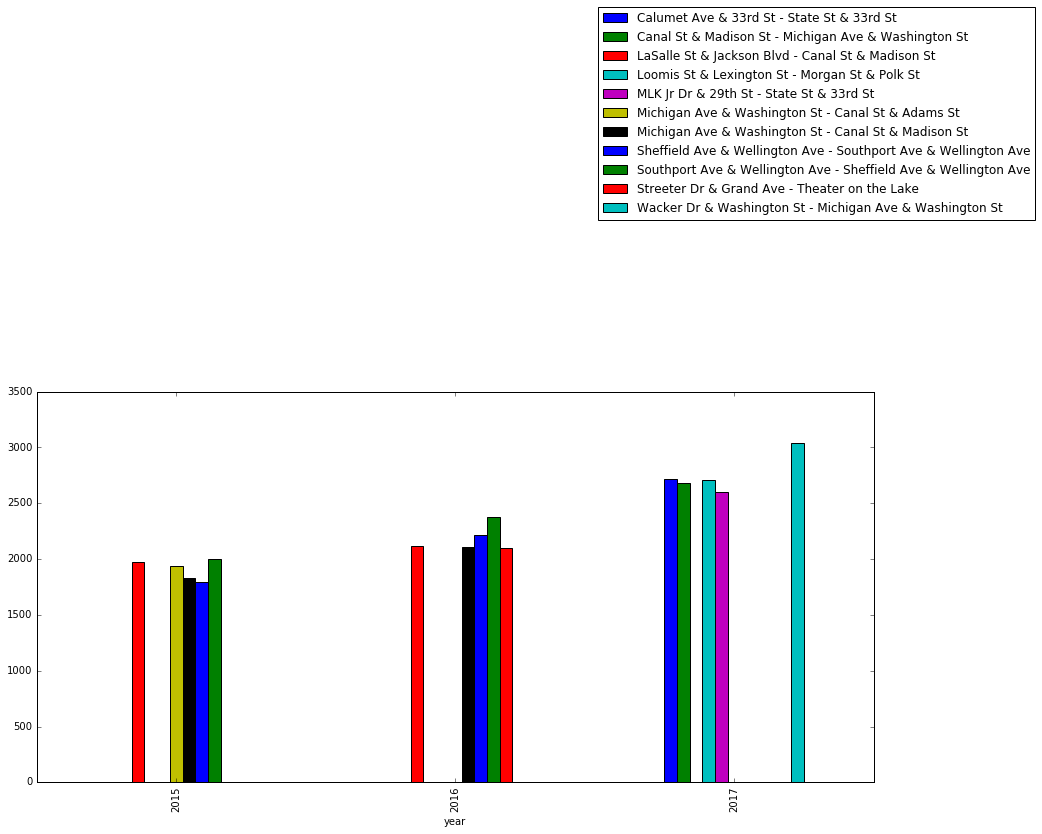
**As we can see from above table, that none of top 10 Busiest station are on Top 10 “Average Trip Duration” list.**

1. **Top 5 most preferred routes by customer for bike trip**

Approach: Queried the data to obtain stations routes (From station and To station) where we see the greatest number of trip counts. We observed that below station routes have the greatest number of trips:

|  |  |  |
| --- | --- | --- |
| From Station | To Station | Trip Counts |
| Southport Ave & Wellington Ave | Sheffield Ave & Wellington Ave | 6449 |
| Sheffield Ave & Wellington Ave | Southport Ave & Wellington Ave | 5804 |
| Columbus Dr & Randolph St | Clinton St & Washington Blvd | 5774 |
| Michigan Ave & Washington St | Canal St & Adams St | 5733 |
| Michigan Ave & Washington St | Canal St & Madison St | 5639 |

1. **Trend of Preferred routes on yearly basis**

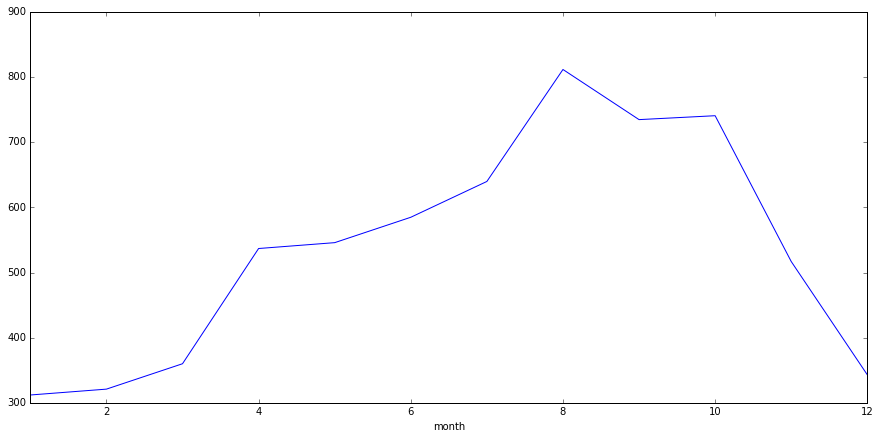


* + Routes - (***Wacker Dr & Washington St, Michigan Ave & Washington***) and (***Loomis St & Lexington St, Morgan St & Polk St***) jumped up under the top 5 most preferred route in year 2017.
  + Routes – **(Canal St & Madison St,Michigan Ave & Washington St),** **(Calumet Ave & 33rd St, State St & 33rd St)** have consistently been under the most preferred routes across year 2015, 2016 & 2017.
  + Route – (**Michigan Ave & Washington St, Canal St & Adams St**) was among the top 5 most preferred route in year 2015 but disappeared from the top list later on.

**Business Decisions:**

* Tie-up with third party on specific campaigning on these routes
* Any new stations in these routes?

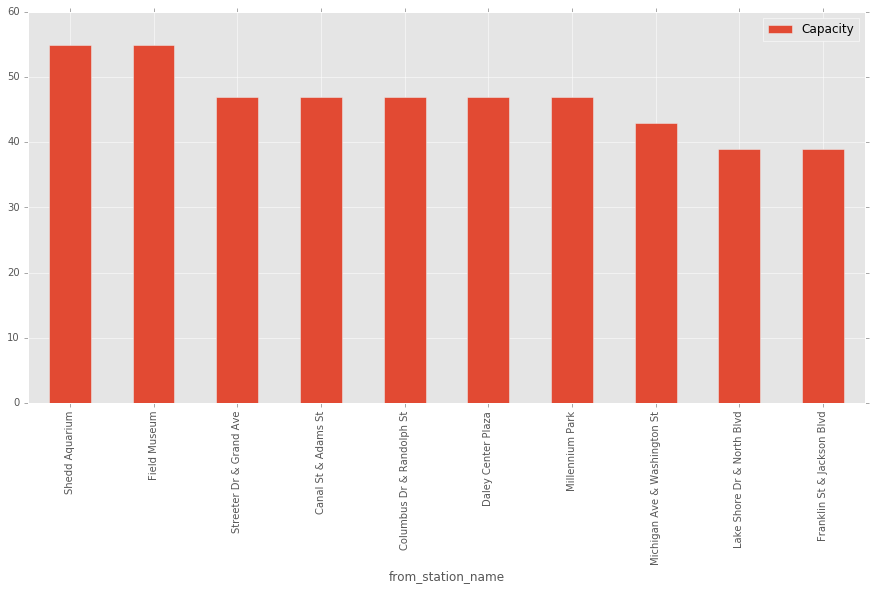
Month-wise tripcount trend for a particular most preferred route - 'Southport Ave & Wellington Ave' -'Sheffield Ave & Wellington Ave'



Above plot depicts that most of the bicycle trips have consistently been happening during **August- October period**.

1. **Top 10 highest Bike capacity station – is this in-line with highest trip generating stations?**

**Approach:** On loaded data, selected stations and Maximum of Capacity group by stations and ordered by descending order of Capacity.



Comparing with Top 10 List of capacity with busiest station:

|  |  |  |  |
| --- | --- | --- | --- |
| **Top 10 Station with highest Capacity** | | **Top 10 Busiest Stations** | |
| **Station\_ID** | **Capacity** | **Station\_ID** | **Trip\_Counts** |
| Shedd Aquarium | 55 | Clinton St & Washington Blvd | 164669 |
| Field Museum | 55 | **Canal St & Adams St** | 140197 |
| Streeter Dr & Grand Ave | 47 | Clinton St & Madison St | 122229 |
| **Canal St & Adams St** | 47 | Canal St & Madison St | 121728 |
| **Columbus Dr & Randolph St** | 47 | **Columbus Dr & Randolph St** | 87577 |
| **Daley Center Plaza** | 47 | **Daley Center Plaza** | 80139 |
| Millennium Park | 47 | Kingsbury St & Kinzie St | 79267 |
| Michigan Ave & Washington St | 43 | LaSalle St & Jackson Blvd | 77547 |
| Lake Shore Dr & North Blvd | 39 | Dearborn St & Monroe St | 75224 |
| Franklin St & Jackson Blvd | 39 | Clinton St & Lake St | 72823 |

From above we can see that there are 3 stations which are in top 10 of Busiest list and also have top capacity.

But what about the **top busiest** station, the capacity of this station is not in Top 10.

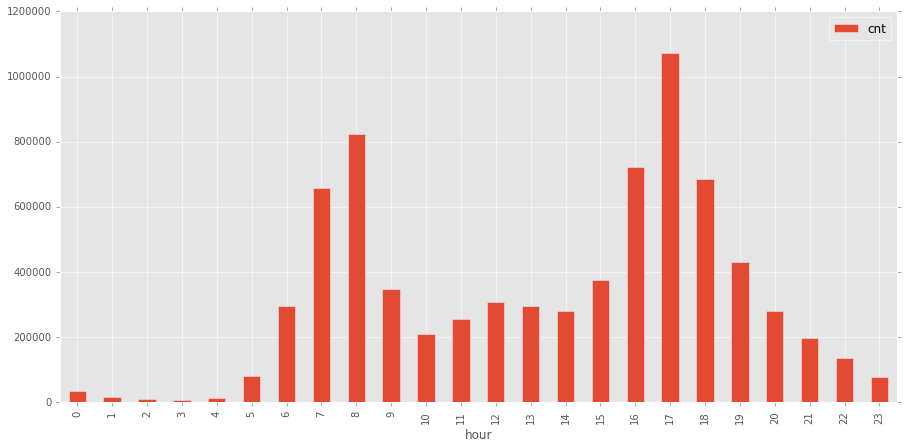
**Business Questions:**

* How are we coping with demand?
* Is there high waiting time for Bikes, which customer is experiencing?
* How about adding Capacity to this station?

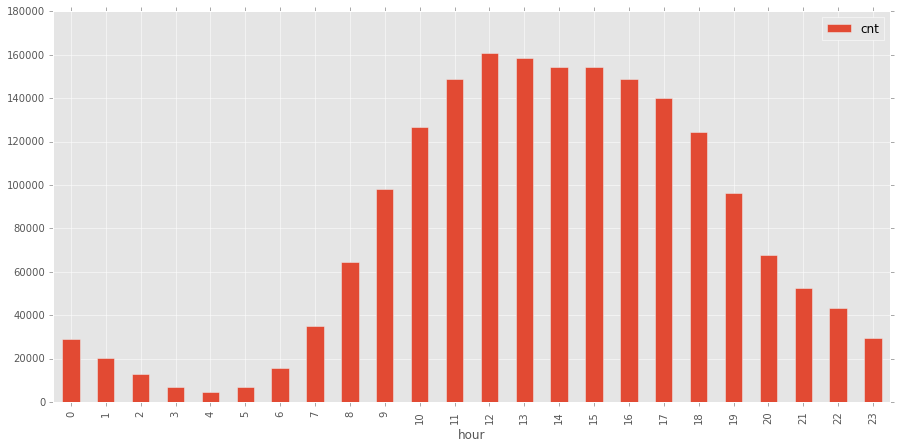
1. **Time of the day (Hour) where most bike renting is happening – split by weekdays and weekends? Is it coinciding with office going times?**

**Approach:** On loaded data, selected hours and trip count when it is weekday, group by hours and ordered by Ascending order of hours. Same step repeated with weekend.

* **Bike trip per hour distribution on Weekdays:**



* **Bike trip per hour distribution on Weekends:**



We can see 2 different distributions here, so we can infer that:

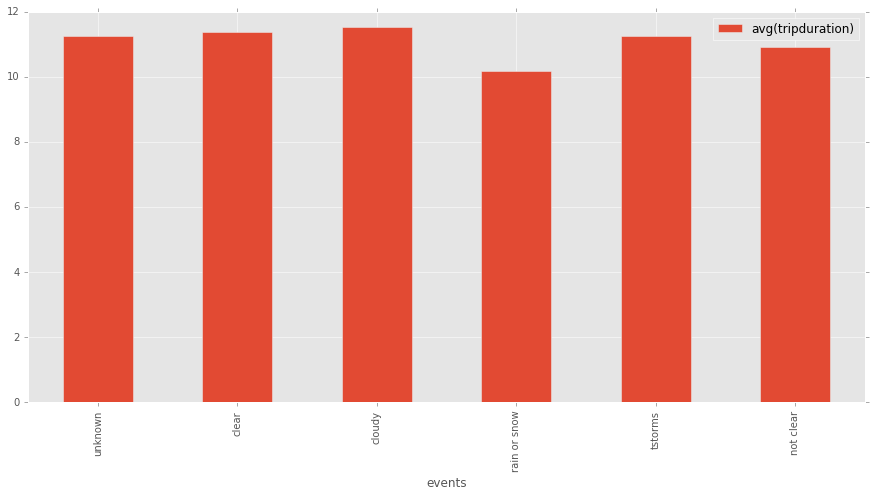
* On Weekdays major bike renting happens during **morning 7-8** and **evening 4-6**, which is making our **hypothesis strong** that people prefer to use Bikes to go workplace and return
* On Weekends, it follows more of a **normal distribution** where most of the bike renting happening during afternoon and very less in morning times.

**Key Business Decisions:**

* Optimum process deployment during morning and evening on weekdays to ensure Bike availability to office goers.
* Advertising Go-Green initiative to attract more office goers to use bike rather than personal vehicles.

1. **Is weather has an effect on Average trip duration?**

**Approach:** On loaded data, selected weather and average trip duration group by weather.



**We can safely assume that Average trip duration has very little variation on weather. In case of rain or snow, customers tend to stop their trip early compared to other type of weather.**

**Bike Demand and Supply**

Supply chain management is very critical for customer satisfaction, as we infer the importance in above descriptive visualizations. We should be able to effectively forecast/predict the demand and put process in place to fulfil.

The analysis is divided in 2 parts:

1. **Predicting Demand:**

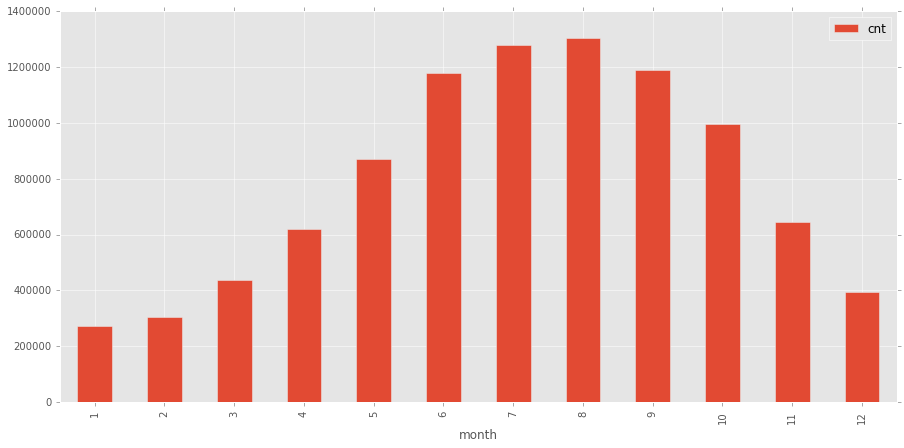
Approach:

* Identify key parameters which influence the pattern of Bike rent (count)
* Create a predictive model of “number of footfalls” in a station based on the key parameters

After analysing various fields of data, we tried to establish the relationship of # of trips v/s other data parameters:

* **# of Trip v/s month**

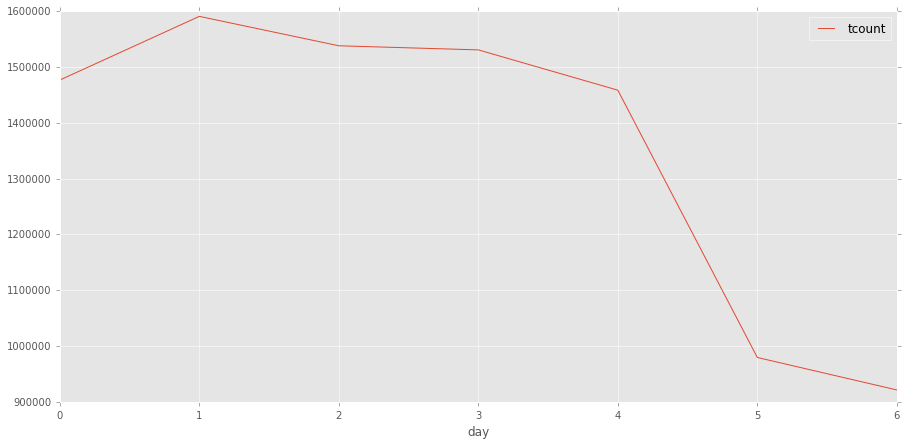
Select count of trips by month, group by month



**We can see that # of trips varies by month lowest in Jan, peak by August and then start decreasing till Dec.**

* **# of Trip v/s Weekday or Weekend**

Select count of trips by day of week (0 is Monday and 6 is Sunday), group by day of week

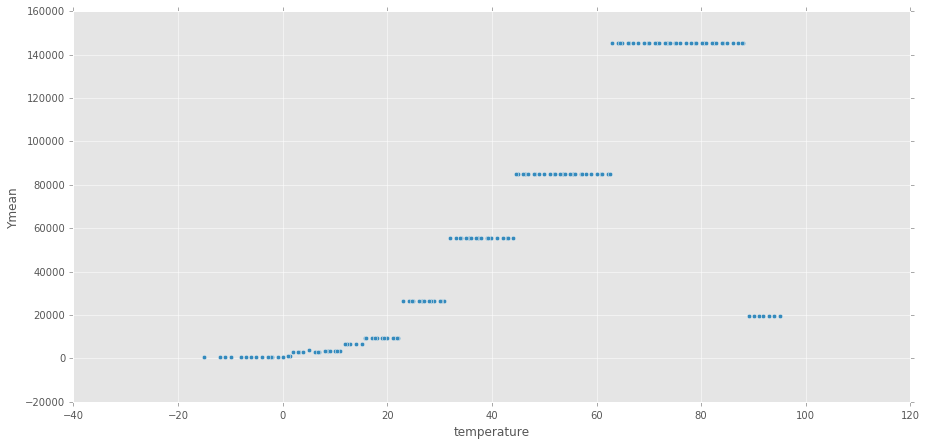


**We can see that # of trips varies by day of week: Steady during weekdays and fall sharply on weekends**

|  |  |
| --- | --- |
| **Day wise trip Analysis** | |
| **day** | **trip** |
| Monday (0) | 1230587 |
| Tuesday (1) | 1325667 |
| Wednesday (2) | 1271789 |
| Thursday (3) | 1266894 |
| Friday (4) | 1208680 |
| Saturday (5) | 810444 |
| Sunday (6) | 765491 |

* **# of Trip v/s temperature**

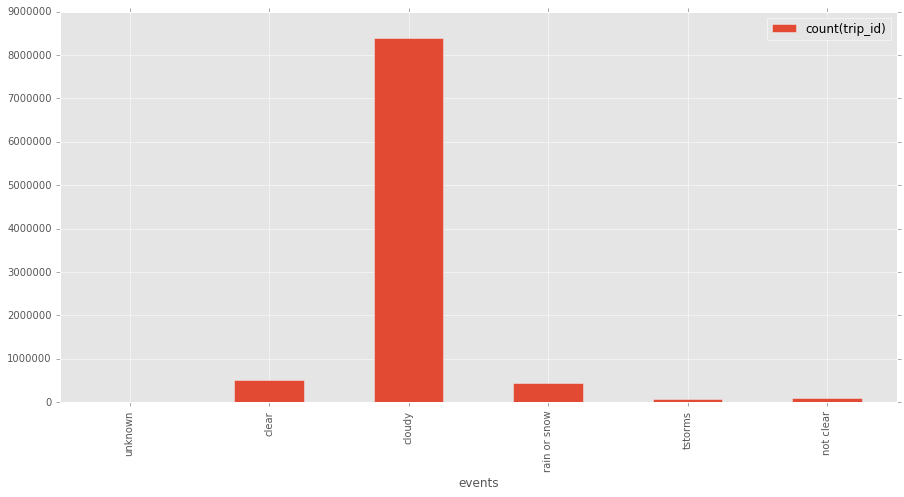
Select count of trips by temperature, group by temperature. Then put bins in temperature range



**We can see that # of trips varies by temperature: # of trips increases when temperature rises till 90 F, but decreases post that.**

* **# of Trip v/s Weather type**

Select count of trips by weather type, group by weather type.



We can see that # of trips varies by weather type**: Cloudy weather attracts more customers compared to Clear weather!!**

**Based on above findings we can concur that, # of trip on a day depends on:**

* Month of the Day
* Whether it is Weekday or weekend
* Weather of the day – clear/cloudy etc
* Temperature of the day

**Can we build a predictive model using Linear Regression? Let us look next section**

**Approach for Linear regression model:**

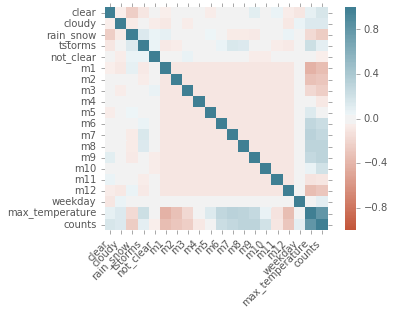
1. **Data Preparation Phase–**
2. Filter data for year 2015, 2016 & 2017 – using 3 years data to build the model
3. Feature Transformation – Our regressor variables, which got identified in above visualization exercise, have following categorical variables with multiple classes:
   * *Month – Values 1-12*
   * *Events – Clear day, Cloudy etc*
   * *Day* – weekday, having value 0 to 6

So, we transformed the categorical values/columns to binary dummies variables.

1. Modify column names - some of column names have spaces between words and therefore causing error while running query.
2. Create a new temporary view & use sql query to filter dataset to get relevant observations for most preferred station name (i.e. Southport Ave & Wellington Ave).
3. Drop original and irrelevant column-
4. Aggregate filtered data for most preferred station to get day basis trip records –

Group by day, week, month, year to retrieve trip records on day basis. Final prepared dataset has 1096 observations for 3 years.

1. Drop irrelevant columns and make data ready for Regression model – Drop **day, week, month, year** columns since we have derived dummies out of them for modelling.
2. Create co-relation matrix between x variables:



There is enough correlation (0.80) between max\_temperature (independent variable) and counts (dependent variable).

Other variables are however binary encoded and therefore evaluating correlation won't be much meaningful

1. **Modelling Phase -**
   1. Extracted features from the data and excluded target variable (i.e. Trip count per day)
   2. Used vector assembler class of pyspark.ml.features module to create a single vector column of features.
   3. Split the transformed data into traindata and test data with a proportion of 70:30 respectively.
   4. Fit model to train dataset (i.e. traindata)
   5. Print coefficients and intercept of trained model.
   6. Evaluate it on test dataset (i.e. testdata)
   7. Print mean absolute error, RMSE and R2 value.
      * Mean Absolute Error = 6.02
      * RMSE = 7.91
      * R2 value = 74.5

|  |  |  |
| --- | --- | --- |
| **Predictors** | **Co-efficient** | **Analysis** |
| clear | 1.9529 | Positive on Trip Counts - expected as per our Visualization |
| cloudy | 3.423 | Strong Positive on Trip Counts - expected as per our Visualization |
| rain\_snow | -3.7434 | Strong Negative on Trip Counts - expected as per our Visualization |
| tstorms | -4.7323 | Strong Negative on Trip Counts - expected as per our Visualization |
| not\_clear | 0.0669 | Very little Impact |
| m1 | -7.337 | Strong Negative on Trip Counts - expected as per our Visualization |
| m2 | -7.4239 | Strong Negative on Trip Counts - expected as per our Visualization |
| m3 | -8.4974 | Strong Negative on Trip Counts - expected as per our Visualization |
| m4 | -3.8466 | Strong Negative on Trip Counts - expected as per our Visualization |
| m5 | -0.8619 | Very little Impact |
| m6 | 4.3511 | Strong Positive on Trip Counts - expected as per our Visualization |
| m7 | 5.5146 | Strong Positive on Trip Counts - expected as per our Visualization |
| m8 | 7.775 | Strong Positive on Trip Counts - expected as per our Visualization |
| m9 | 8.2666 | Strong Positive on Trip Counts - expected as per our Visualization |
| m10 | 8.114 | Strong Positive on Trip Counts - expected as per our Visualization |
| m11 | -2.1082 | Strong Negative on Trip Counts - expected as per our Visualization |
| m12 | -5.7203 | Strong Negative on Trip Counts - expected as per our Visualization |
| weekday | 3.9995 | Strong Positive on Trip Counts - expected as per our Visualization |
| max\_temperature | 0.3483 | Not so strong impact on number of counts |

Using this Model, we can predict the demand/number of trips expected on particular station by providing following parameters:

* Month of the Day
* Whether it is weekday or weekend
* Weather
* Temperature

Now we have done the modelling, let us ponder on second part of Analysis.

1. **Supply Chain i.e. Demand fulfilment process**

What if projected demand for the day for a Station ID exceeds its Capacity?

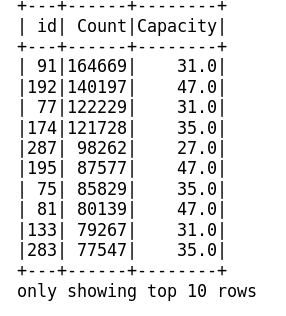
How we will increase the capacity, may be 2 options:

* Buy new bikes to fulfil the Demand
* Or check the spare capacity at nearby stations and loan the unused capacity from there?

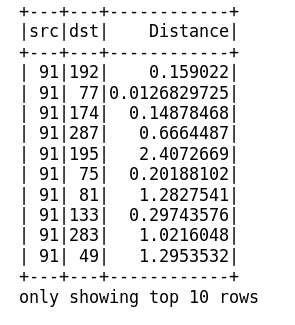
Second option sound more optimum and cost friendly.

**Approach for creating solution for second option using Graph Frame:**

* Created edges file containing each station ID as id, capacity of each station and count of trips originated from each station

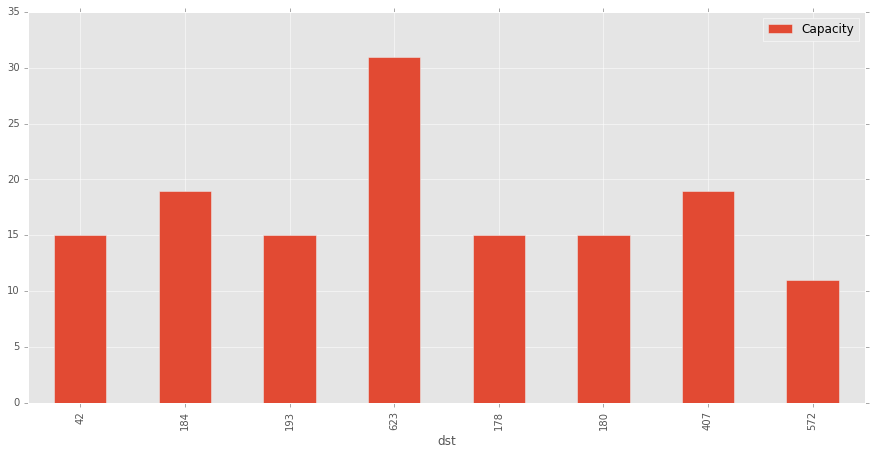


* Created a function which takes 2 sets of geo codes and provide the km distance between these geo-code
* Then we calculated distance between all permutations of stations (using geo codes and above function) for example if there are 3 stations, then distance between 1-2, 1-3 and 2-3. We have **586** stations, so total of **171405** permutations were captured
* Created Edge file, with source and destination as 2 stations and relationship is distance between them



* Created graph frame dataset using vertex and edge file
* Now we can get the list of stations which is x distance (user can specify) from my station and can see the capacity of these stations
* This way we can easily identify the spare capacity near-by for quick fulfilment.

For example, if we want to see which is the nearest station (within 400 metres) from station id 148 and what is the capacity, we can run the file and visualize as follows:



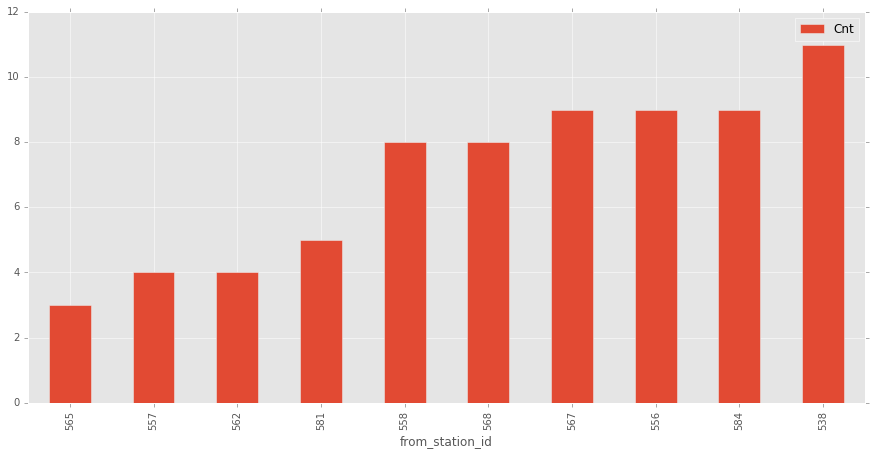
The above visualization is sorted based on nearest distance from station id 148, i.e. station 42 is nearest then 184 henceforth. And then for each station showing the capacity.

**Opportunities of Business consolidation**

Business need to see the consistent non-performing entity and plan strategy to either eliminate, consolidate or revamp. On same lines, we look at the data and see which station ids are consistently generating low trips and finding out opportunity to consolidate with other stations, using the capacity.

* Bottom 10 stations which are generating least amount of Bike rents:

Select station ID, count trip and group by station id order by count descending, check for latest year



* **Least preferred routes by customer**

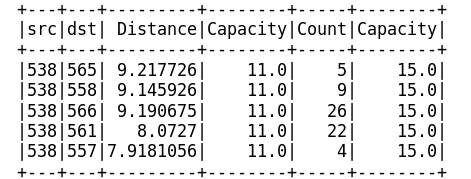
Approach: Queried the data to obtain stations routes (from station - to station) where we see least number of trip counts. We observed that below station routes have least number of trips

|  |  |  |
| --- | --- | --- |
| **From Station** | **To Station** | **Trip Counts** |
| **Racine Ave & 13th St** | **Ashland Ave & Division St** | **1** |
| **Halsted St & 63rd St** | **MLK Jr Dr & 56th St (\*)** | **1** |
| **Albany Ave & Bloomingdale Ave** | **California Ave & 21st St** | **1** |
| **Dearborn St & Adams St** | **Western Ave & Division St** | **1** |
| **Budlong Woods Library** | **Seeley Ave & Roscoe St** | **1** |

* **Identify nearest stations for consolidation:**

For this we used the **graph frame** which we created in supply chain use case.

* Taking example of 565 station ID, which has generated only 3 trips in 2017 (as per above visualization).
* Using Graphframe, find out nearest (<500 metre) stations having good footfall: **there is nothing**.
* Nearest station details as follows:



**With above output, we can interpret that:**

* Nearest station to 538 station is ~8 kms, but there also very sparse footfall
* In this case consolidation looks like not possible as distance between stations are high.
* We might consider to optimize the capacity; do we need 11 for all the above stations?

**Customer Base Analysis**

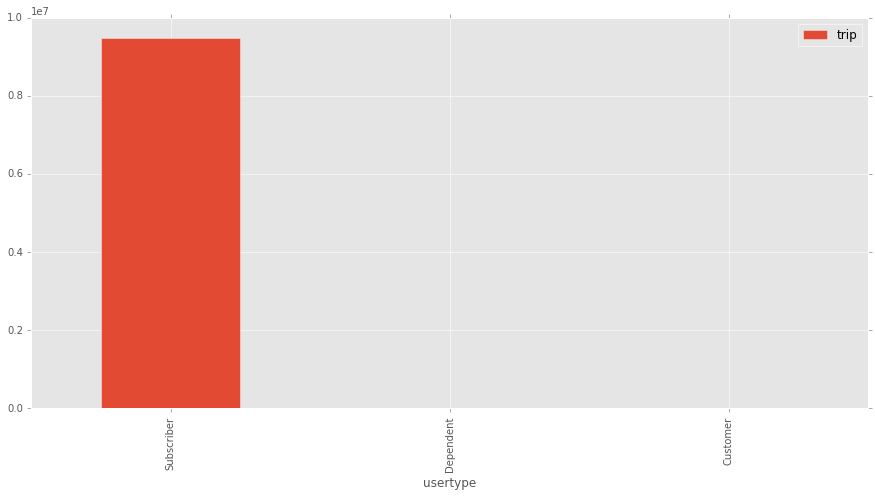
Divvy bikes has both “Subscriber” based business model and “Pay-as-you-go” business model. Which model is more popular?

* Subscriber v/s pay-as-you-go customer:

Approach:

* Used Spark SQL to write query on above data set and taken two variables user type and count of trips
* Usertype defines types of users as per dataset. On analysis we found there are three types of users
  1. Subscribers
  2. Dependent
  3. Customer
* Dependent and Customers consist only .017 % of the total user types and are too small to consider compared to Subscribers.

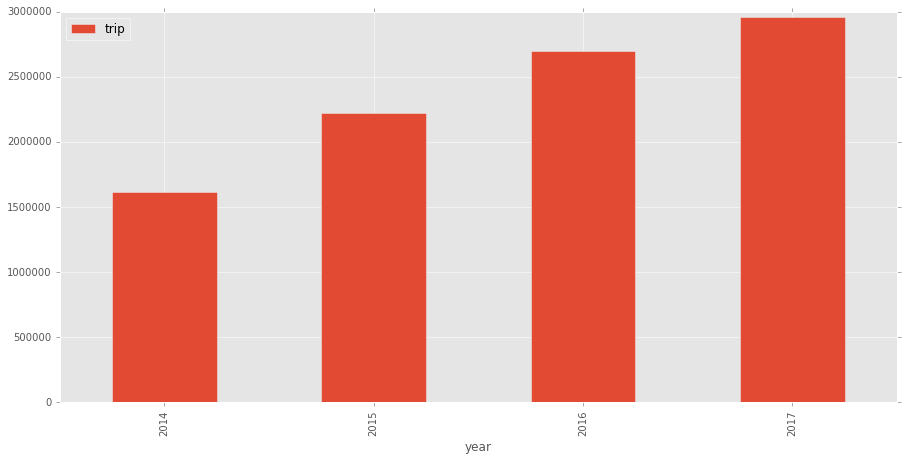
This denotes Subscribers are the major revenue contributors



“**Subscriber” based business model is clear winner**, compared to that other business model are very negligible.

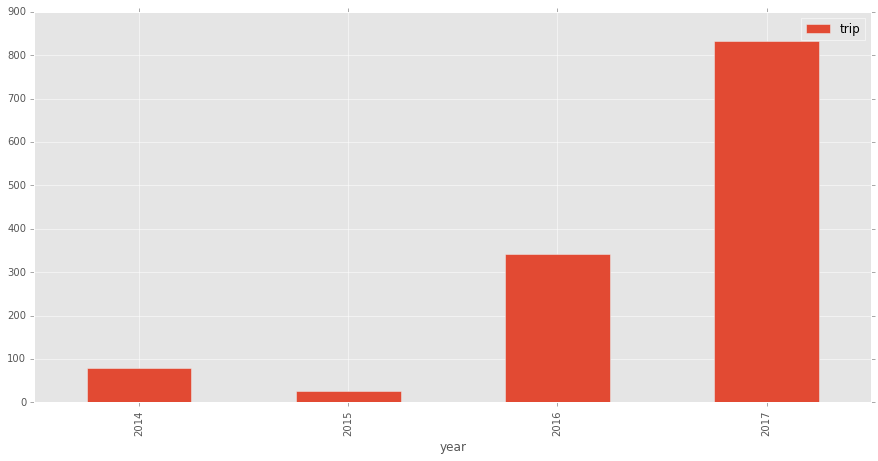
* **Trend of subscribers – i.e. if more customer becoming subscriber on yearly basis?**

Approach: Select count of trips based on usertype, group by year



And we can see that there is consistent growth in subscriber-based business model, year on year more and more people are becoming subscribers.

Also, there has been a growth in “pay-as-you-business” over years, however overall count is insignificant compared to subscribers

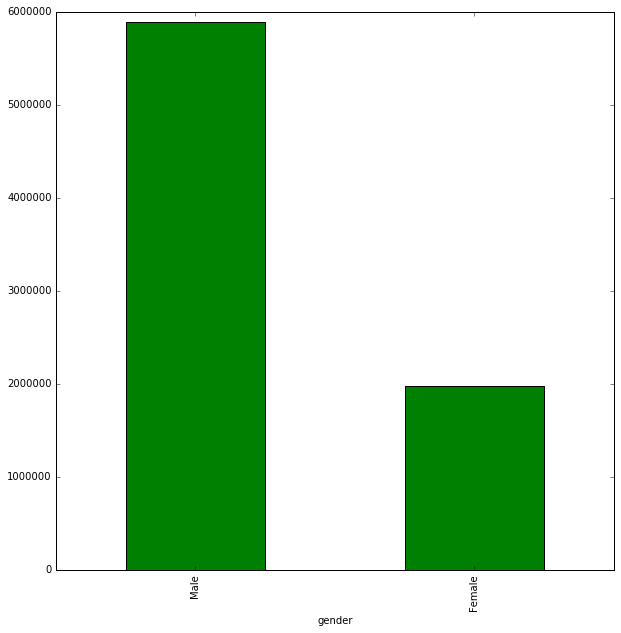


**Customer Bike riding pattern based on Gender**

* Who prefers Bike riding more – Male or Female?

**Approach:** Queried the filtered data for year 2015, 2016 and 2017 to get the gender, and count of observations across each gender.

Such observations give us the count of times a particular gender group took bicycle on rent for ride.

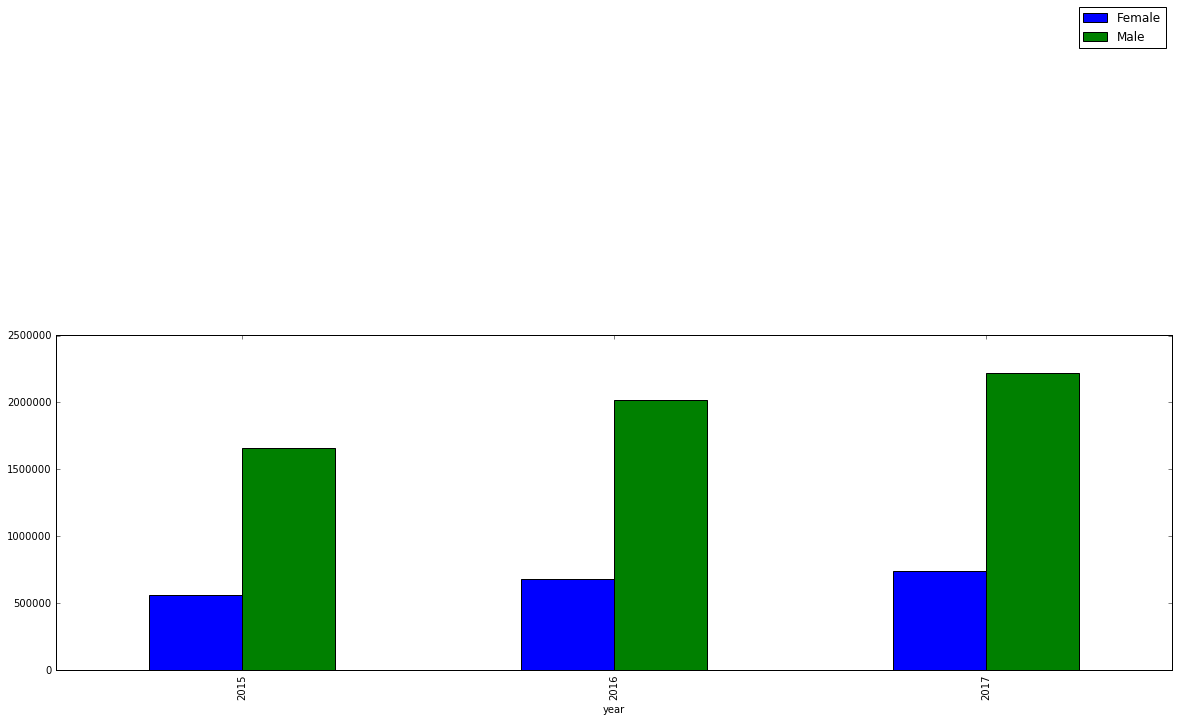


Number of **males** hiring and riding bicycles are much **higher** than **females**.

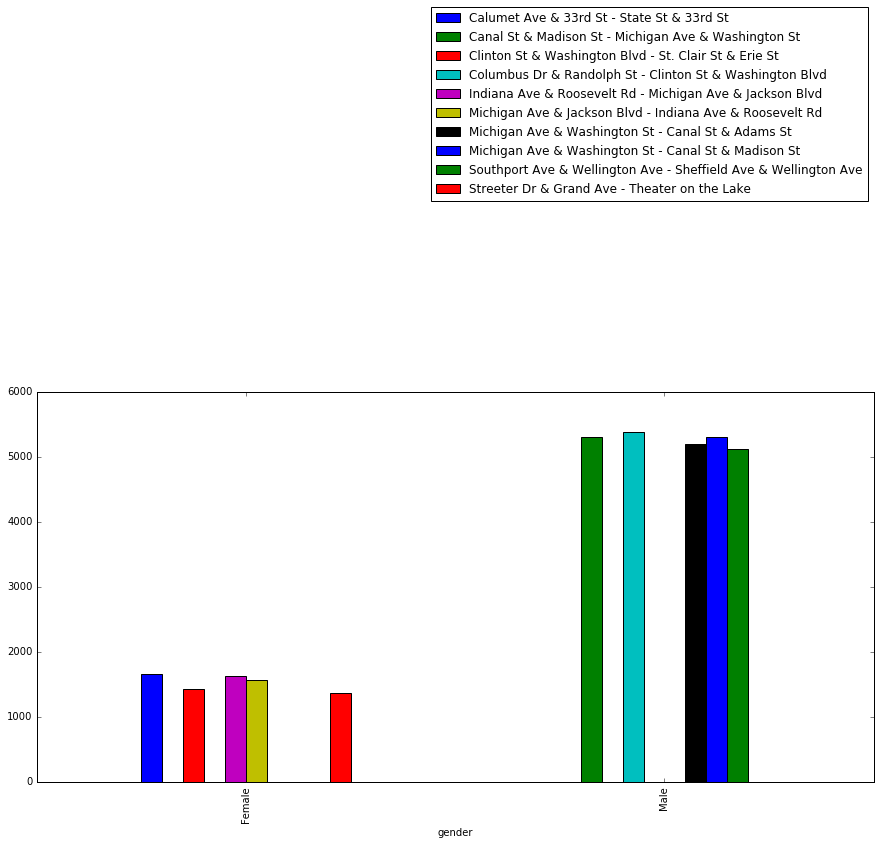
* **Any trend over years**

**Approach:**

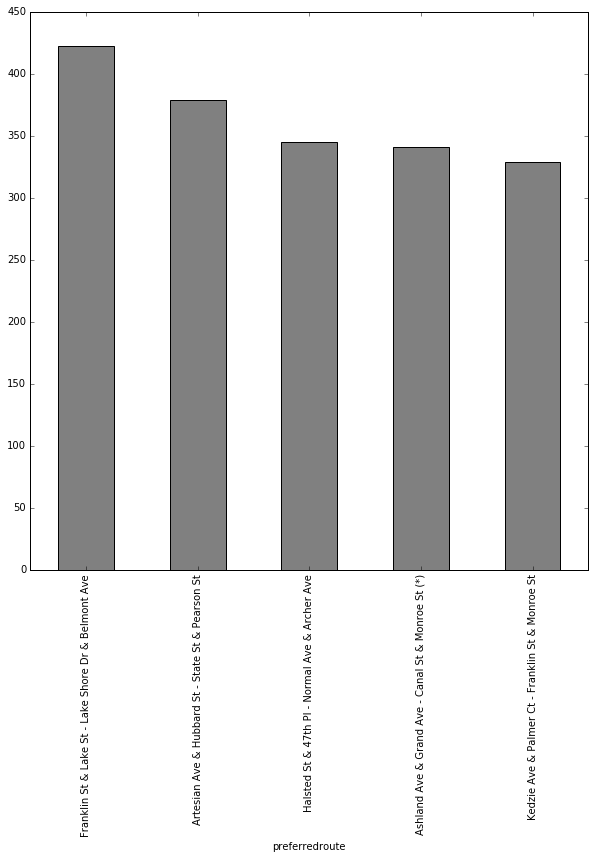
* + Queried the data to fetch all possible gender & year combination. We got 6 combinations (i.e. observations) wherein each row gives us number of times a particular gender group rented bicycle in a particular year.
  + Used Spark SQL - for querying the data and Spark Core - to convert Spark dataframe to pandas dataframe & plots.
  + Across year 2015, 2016 & 2017, males have shown a considerably more interest in rented bicycle for commuting as compared to female.



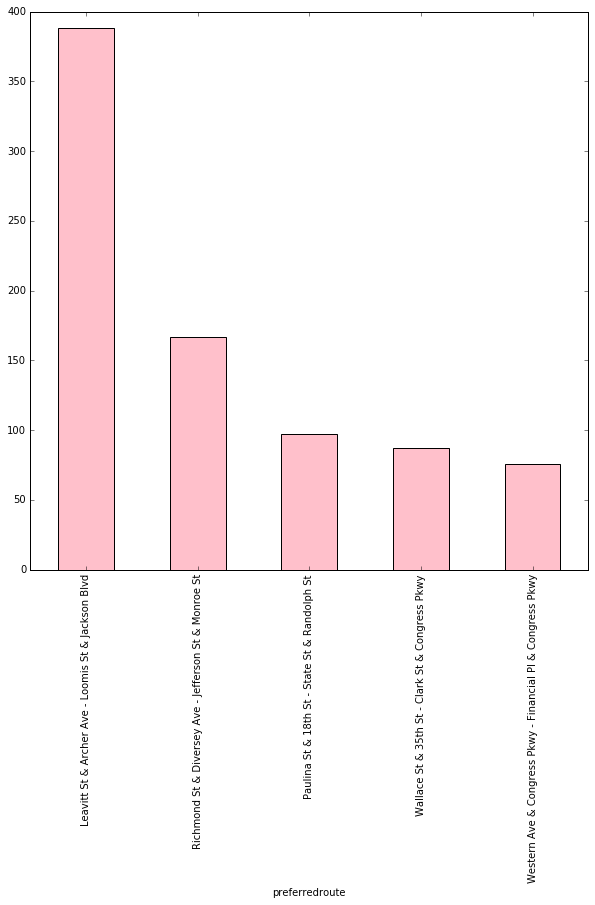
* **Gender-wise preferred routes (Include common routes opted by both gender**):



* + Queried data to fetch gender(male/female), preferred route and their count. Order in descending order of counts to get the most preferred routes for each gender group.
  + Filter data based on gender to get gender specific preferred route (Group by Gender, preferred route). After filter, we have two tables– each table specific to each gender.
  + Extract the preferred route information from each of the table and store them in separate variables.
  + Apply subtract() to get the observations(i.e. preferred routes) specific to a gender base.
  + For example – variable A has preferred routes by males information and variable B has preferred routes by females information. Now subtracting B from A i.e. A – B will give us preferred routes by males only.
  + Below **point** below gives us the routes preferred by males and females only :
* **Top 5 Routes explored by Gender – Male only –**



* + Above are the top 5 routes (from station, to station) which are explored by males only. Among them, route **- Franklin St & Lake St – Lake Shore Dr & Belmont Ave** is most explored.
* **Top 5 Routes explored by Gender – Female only –**



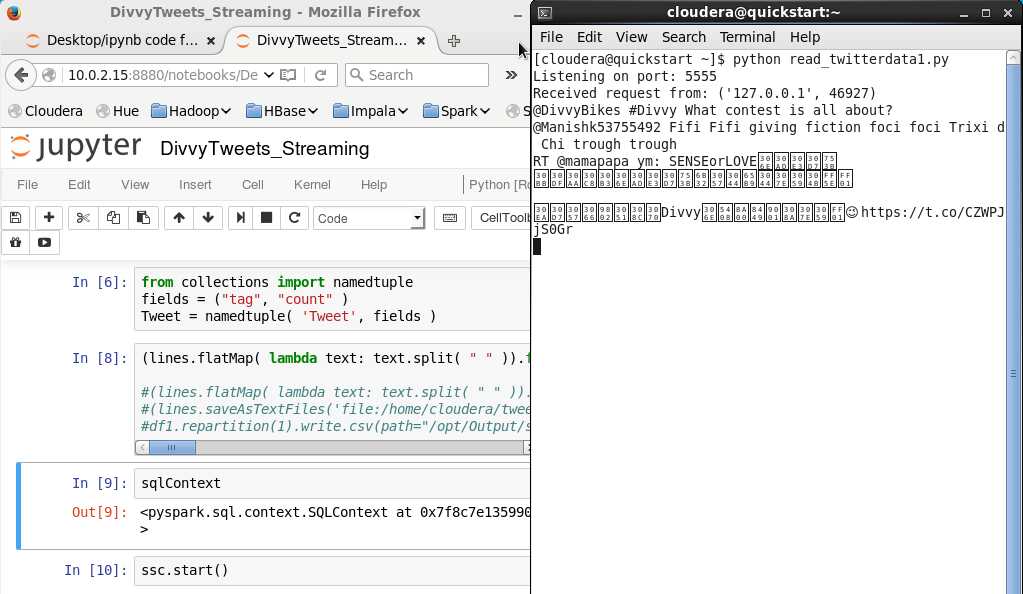
* + Above are the top 5 routes (from station, to station) which are explored by females only. Among them, route - **Leavitt St & Archer Ave– Loomis St & Jackson Blvd** is most explored.
* **Business recommendation:**
  + Gender based campaign to these routes by tying-up with third party
  + Gender specific cycle accessories as a cross-selling business proposition can be explored.

**Live Sentiment Analytics**

Divvy Bikes has been talked about in twitter and other social media channels. Hence text analytics becomes quite imperative, as it will help in mining out positive and negative sentiments and accordingly take the informed and calculated decisions.

Approach:

* Using Spark Streaming, connect with twitter and capture Live tweets filtered with “Divvy” keyword



* Capture in Dstream and apply text analytics on tweet
* Capture the sentiment and key words
* Take corrective actions (in case of negative sentiment)

**Conclusion:**

Right data used at right time can do wonders. With limited data points, we uncovered various dimensions on which “Divvy Bikes” business can venture out and look for answers, implications.

Be it Supply/Demand mystery or understanding customer behaviour or creating personalized campaign, Data holds the key. We used this key to unlock various unknowns and provided solution and right direction for business.

Using these insights and recommendations aided with Business knowledge and other data (like customer sentiments), a lot of dots can be connected and new solutions could be emerged which can be applied further to increase customer engagement/appreciation (our Key goal) and journey towards path of “exponential growth”.