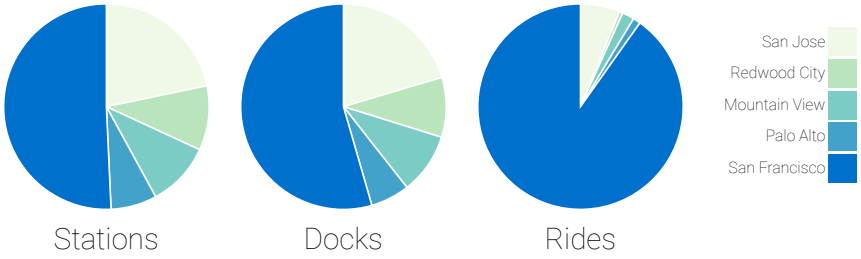
Bay Area Bike Share Data Challenge

Home Explore

What is the Bay Area Bike Share?

The Bay Area Bike Share system allows users to rent bicycles for short journeys between stations throughout the city. Users can be annual members or short term (1 or 3 days). The system is completely automated for users.

There are **69** stations across **5** cities in the Bike Share system, with an average of **17** docks per station.



About **50%** of the stations and docks are located in San Francisco, but it makes up **90%** of the system use.

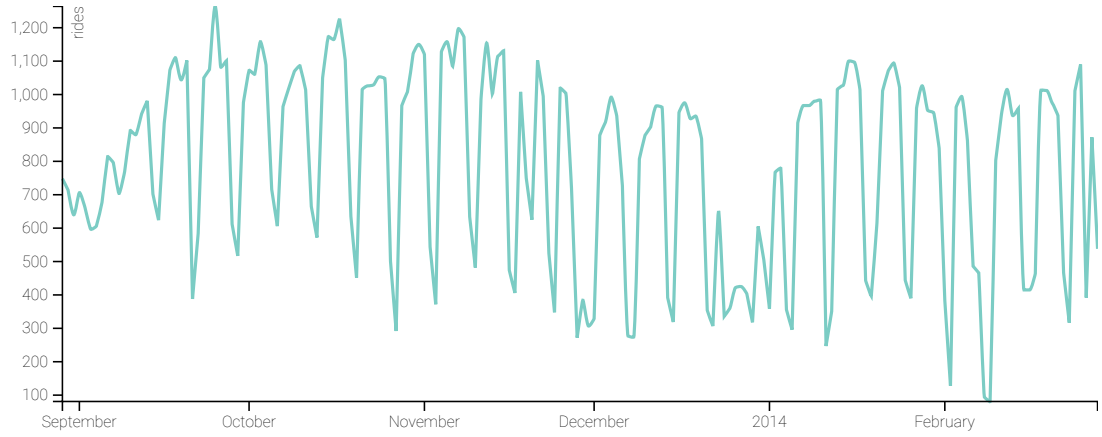
The data we analyze here came from rides between August 29, 2013 and February 28, 2014.

In those 185 days there were
144,015 rides, averaging about
20 minutes per trip.

Total riding time:
49,241½ hours, or **5 years, 32 weeks, 1 day, 12 hours, 27 minutes, and 4 seconds.**

How much is the Bay Area Bike Share used?

☐ Highlight Weekends



System-wide Rides per Day

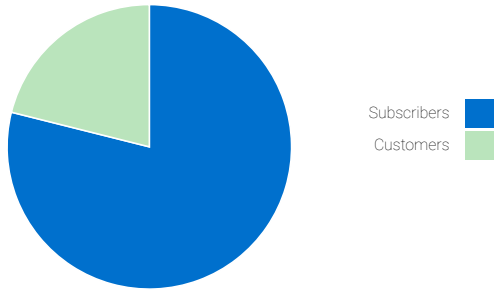
This chart makes it pretty clear that ridership drops on weekends. Check the box to highlight Saturdays and Sundays.

917 rides are made on average each weekday

442 rides are made on average each weekend

Busiest day	Calmest day
2013-09-25	2014-02-09
1,264 rides	81 rides

Who uses the Bay Area Bike Share?



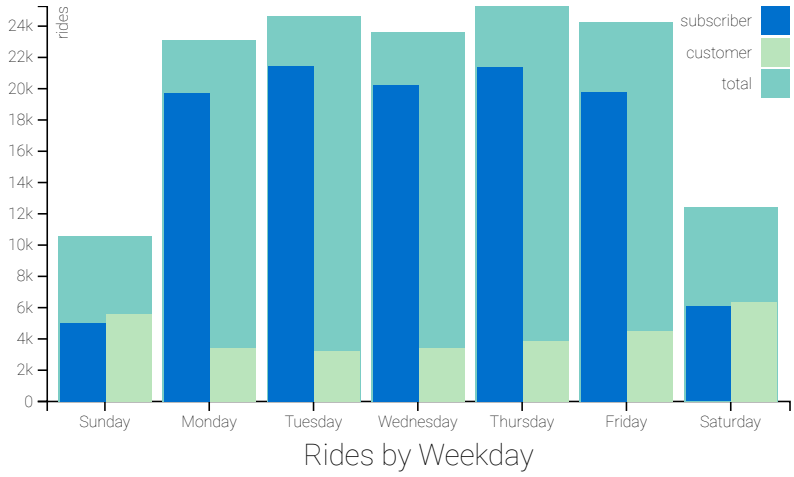
Riders by membership

Almost **80%** of riders have an annual subscription. The remaining **20%** of riders were customers who purchased a 24-hour or 3-day pass.

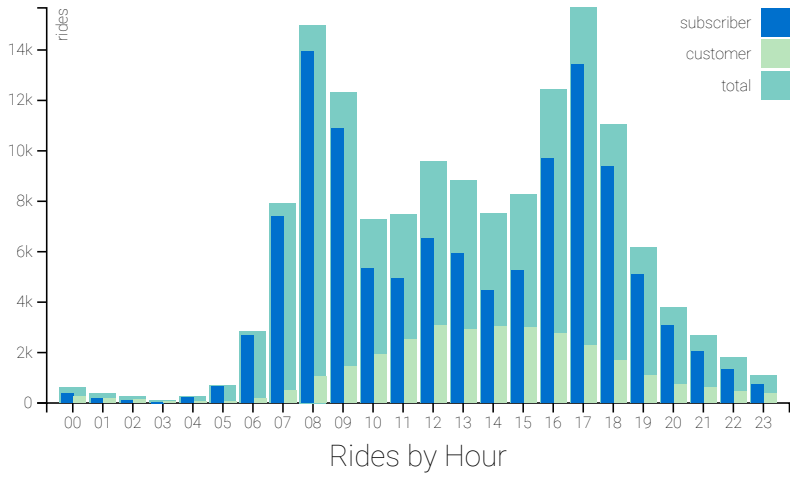
When is Bay Area Bike Share used?

We saw already that users are likely to be annual subscribers, and system use drops off on the weekends. Let's group the rides by weekday and see how subscribers' use compares to customers':

Weekday riders are overwhelmingly subscribers, and ridership among subscribers falls on weekends so that rides by customers just manage to outnumber them.



And now we'll group rides by time of day:



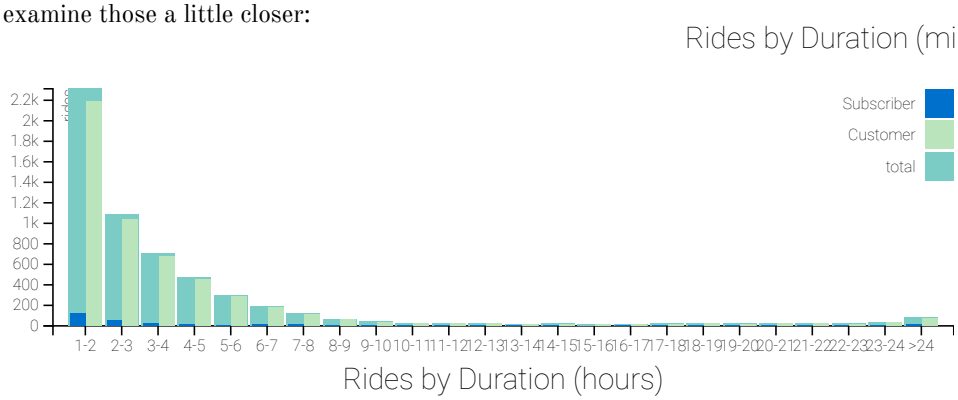
Among subscribers we see spikes in use at 8am and 5pm with another small bump at 12 noon. These users must be riding a bike to get to work, to go to lunch, and to head home.

Customers hourly usage seems to fall along a bell-shaped distribution peaking at two in the afternoon. There doesn't seem to be a lunchtime rise among customer use. These users must be riding around throughout the daytime at their leisure.

From these usage behaviors, it would be fair to characterize the two groups of subscribers and customers as **commuters** and **tourists**, respectively.

The Bike Share system is intended to be used for short rides: trips under a half hour do not incur any additional charges. Do riders use the system in the intended way?

Yes they do. The chart on the right shows that the most common ride length is **5 to 10 minutes**. Subscribers are clearly savvy to the price structuring. Very few rides longer than a half hour are taken by subscribers. For the most part, customers are savvy as well. Their trips last a little longer on average, but mostly less than 30 minutes. There's a bump in rides that last longer than an hour, let's examine those a little closer:



Rides lasting longer than an hour are most commonly 1 to 2 hours long, and almost always taken by customers. This could be due to confusion over the nature of the "24-hour pass" or one of many other factors including theft, forgetfulness, or getting lost.

By the current pricing structure, a customer purchasing a 24 hour pass and taking a ride that lasts 2 hours and 59 minutes would pay in total $41. A ride of 2 hours and 29 minutes would only be $34. There are several bike rental companies in San Francisco which offer 3-hour rentals for around $32. Customers taking trips less than three hours likely did not understand the bike share system and incurred unwanted overtime fees, or they decided the convenience of

the automated system was worth the premium paid over other rental offerings.

Where do people ride Bike Share?

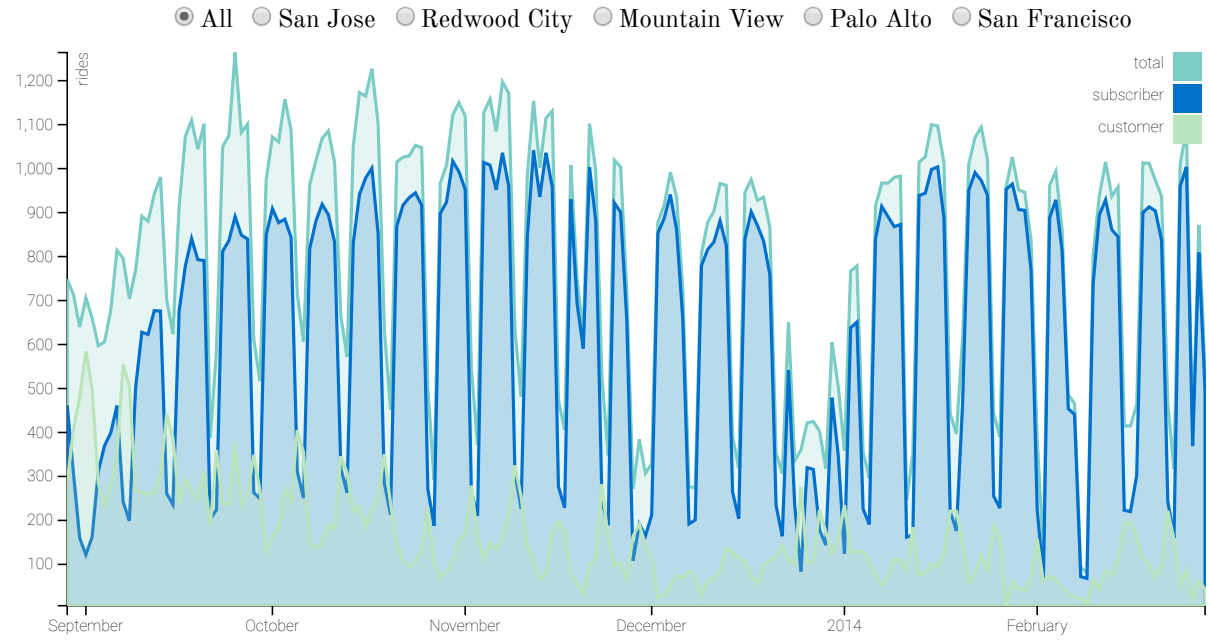
We saw earlier that 90% of Bike Share rides took place in San Francisco. Let's examine the average numbers for each city on a day-to-day basis:

Average Rides per Day

	System Wide	San Jose	Redwood City	Mountain View	Palo Alto	San Francisco
Total	778	48	4	15	9	702
Subscriber	614	39	3	12	5	555
Customer	164	9	1	3	4	147

Hmm. Redwood City and Palo Alto don't see much daily use, but not much interesting here. ~~Let's look at a chart.~~ Let's look at an interactive chart.

Here we see total daily rides for each category of user plotted across the entire timespan, similar to the chart at the top of the page. Use the radio buttons on top to examine different cities and the checkboxes below to highlight days with a factor that might have influenced riders.



Show Average Rides
Reset

Weekends
Holidays

Rain
Temp

49ers Game
Giants Game
Sharks Game

Americas Cup
BART Strike
Government Shutdown

Explore these factors closer on the next page →

If our generalization of users as commuters and tourists were true, we would predict use among subscribers to drop off when commuters don't go to work and increase when tourists are in town. With the overlays we can see this correlation holds true. Rides by subscribers in all cities decreased on weekends, over Thanksgiving, and between Christmas and New Years. Rides by customers increased noticeably on weekends, holidays, and during the America's Cup finals.

Rain seems to be a large deterrent to bike use; on rainy days system use falls almost as predictably as on weekends or holidays. Temperature, however does not seem to have a major influence on riders. Bay area temperatures being notoriously moderate, the lack of correlation is not surprising.

We can't draw any conclusions about the influence of home games by the 49ers, Giants, or Sharks. Heading to a Giants game would seem to be an ideal use of the Bike Share, but during this period the only Giants home games overlapped with Americas Cup races. There are a few dates where a Sharks game correlated with a slight rise above average usage, but such deviation from average is normal throughout this time period.

Interestingly, both the federal government shutdown of October and the two BART worker strikes of October and December did not seem to have a major influence. Rider numbers during the strikes remained similar to numbers from the weeks before and after the strikes. This poses an interesting question, do BART commuters use the Bike Share System? To begin looking for an answer, let's look at the numbers for rides to and from individual stations.

Most Popular Starting Stations

Station	Rides
San Francisco Caltrain (Townsend at 4th)	9,838
Harry Bridges Plaza (Ferry Building)	7,343
Embarcadero at Sansome	6,545
Market at Sansome	5,922
Temp. Transbay Term. (Howard at Beale)	5,113
Market at 4th	5,030
2nd at Townsend	4,987
San Francisco Caltrain 2 (330 Townsend)	4,976
Steuart at Market	4,913
Townsend at 7th	4,493

Most Popular Destinations

Station	Rides
San Francisco Caltrain (Townsend at 4th)	11,637
Embarcadero at Sansome	7,590
Harry Bridges Plaza (Ferry Building)	7,475
Market at Sansome	6,238
2nd at Townsend	5,655
San Francisco Caltrain 2 (330 Townsend)	5,112
Market at 4th	5,109
Steuart at Market	5,080
Townsend at 7th	5,073
2nd at South Park	4,431

Here we see the top 10 stations in the entire system to start or end a ride. Interestingly, the stations near Caltrain and the Ferry building top the list. Both Harry Bridges Plaza and Steuart at Market are close to the Ferry Building. San Francisco Caltrain 1 and 2 obviously serve the Caltrain line. The station at Embarcadero at Sansome is next to Pier 27, where the America's Cup Pavilion was located.

San Francisco Caltrain (Townsend at 4th)

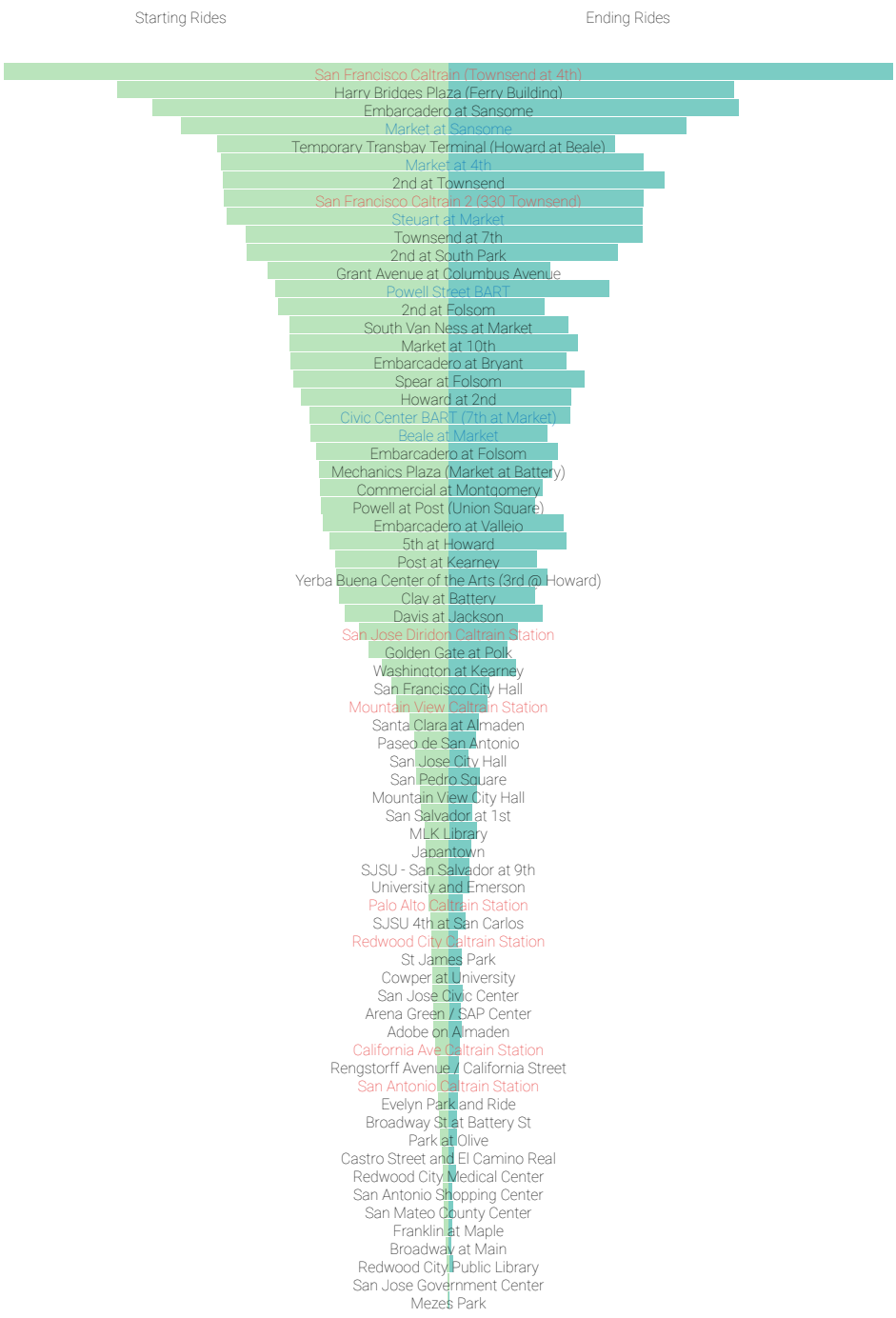
Most popular starting point	Most popular destination
9,838 rides	11,637 rides

The Bike Share stations on the top ten lists within walking distance of a BART stop are Market at Sansome, Market at 4th, and Steuart at Market. Keep in mind that all these stops are also within close proximity to many points of interest in downtown San Francisco.

Of course, the rest of the stations serving BART commuters could all be just outside the top ten. Let's take a look at the list of rides for all stations. BART stations are identified with blue text, Caltrain stations with red text.

Sort by:

◉ Starting station ○ Ending station



While the BART stations are indeed all among the busiest stations, they are all in San Francisco, and their numbers don't set them apart from any other station within San Francisco. These stations are centrally located to a number of attractions in downtown San Francisco, so we would expect them to be highly trafficked. Bike Share Riders docking at these stations are not necessarily transferring to or from BART. If there was a large contingency of BART commuters who used the Bike Share, we would expect to see rides to and from BART stations in numbers setting them above other destinations. With what we've seen here, it seems safe to conclude that BART commuters do not make up a large proportion of Bike Share riders.

Is there any more evidence to back up this assumption? Let's take a look at the rides at stations serving different forms of commuting to and from San Francisco:

Total rides at stations serving:

	BART	Caltrain	Ferry Building	Transbay Terminal
Starting	20,919	14,814	12,256	5,113
Ending	21,325	16,749	12,555	4,356

These numbers are somewhat misleading since there are four stations within walking distance of a BART station, but only two nearby the Caltrain and Ferry building. Instead, let's look at average rides to/from these stations:

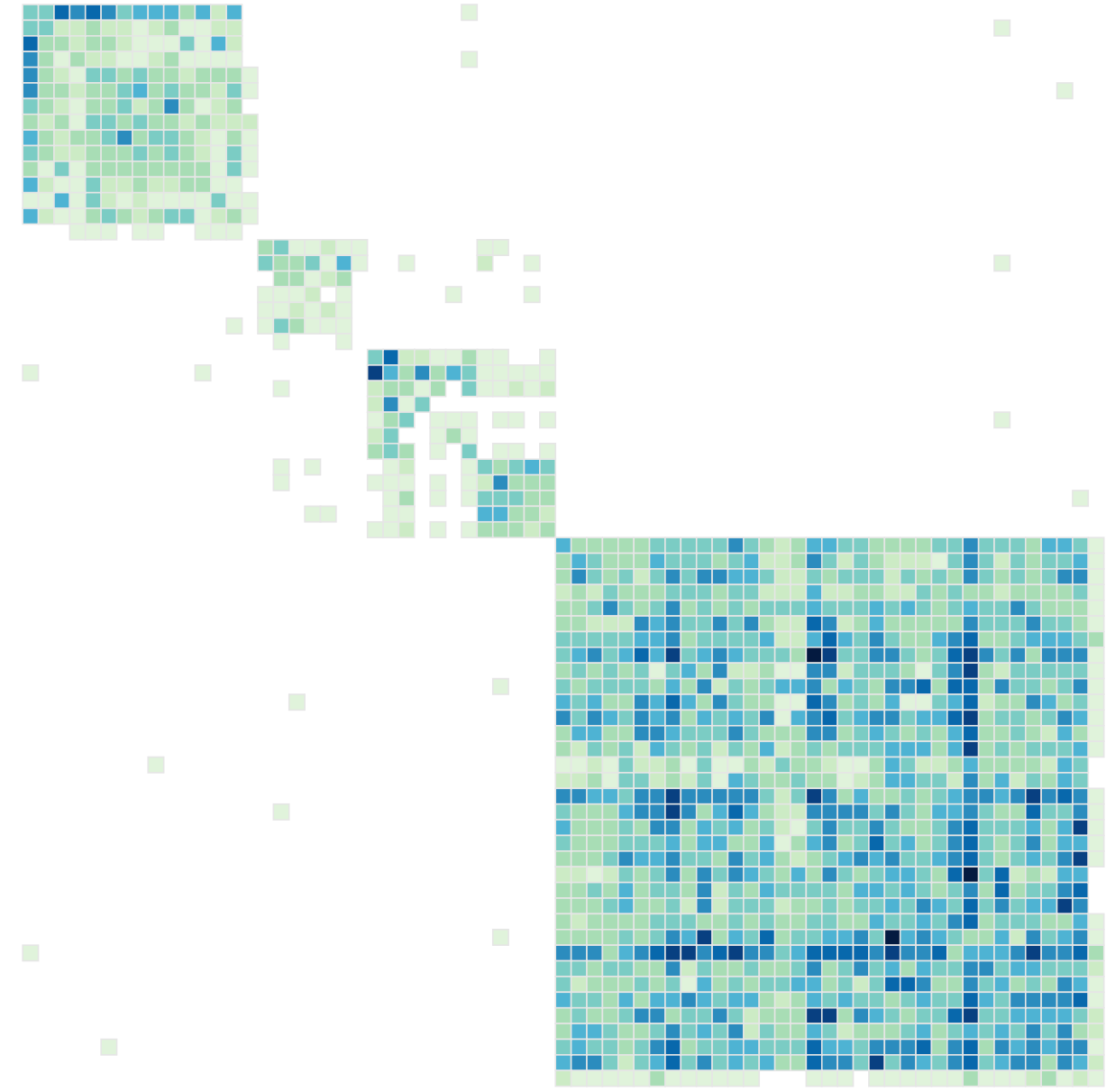
Average rides at stations serving:

	BART	Caltrain	Ferry Building	Transbay Terminal	All SF
Starting	4,305	7,407	6,128	5,113	3,710
Ending	4,401	8,374	6,278	4,356	3,710

Bike Share stations serving BART do not seem to be more trafficked than other stations in San Francisco. Of course, we would be making a massive mistake if we were to assume that all rides to or from these stations started or continued with trips on the related commuter system. The Ferry Building is a major tourism destination, every BART stop is within walking distance to numerous restaurants, museums, business, MUNI rail and bus stops. The Caltrain station, on the other hand, is only within walking distance of a handful of other points of interest. When we compare the numbers for the Caltrain station to the average for all stations in San Francisco, it becomes clear that Caltrain commuters are making great use of the bike share system.

Each trip within Bike Share is recorded by starting station and ending station, so in addition to examining the popularity of stations, we can examine the popularity of individual routes. In the heatmap below, starting stations are listed in rows and the ending stations in columns. The total number of rides from least to most along each route is indicated by the cell color, lightest to darkest.

◉ All ○ San Jose ○ Redwood City ○ Mountain View ○ Palo Alto ○ MV/PA ○ San Francisco



On the heatmap diagram of systemwide rides we see that activity tends to be grouped into squares. These are trips that took place within city boundaries, and we notice that not many riders go beyond their starting city. An exception to this are trips between Mountain View and Palo Alto, between which 191 rides went from one city to the other.

Looking at the station heatmap for San Francisco, we note that riders leaving from the most popular station, SF Caltrain, disperse throughout the system. Riders heading to SF Caltrain similarly tend to come from throughout the system.

Two popular routes which seem to make up a round trip are Townsend at 7th to SF Caltrain 2 and vice-versa. Near to Townsend at 7th is an Expo center and at handful of influential tech companies including Adobe, Heroku, Citrix, Advent, and Zynga. This could reflect a corporate membership program popular among commuting employees, a conference well-attended by peninsula-dwellers, or more simply, a dearth of other forms of transportation to the area.

The most ridden route, Harry Bridges Plaza (Ferry Building) to Embarcadero at Sansome, does not have a return route with as great numbers, indicating that riders tended to ride to Embarcadero at Sansome and continue their journey elsewhere rather than return to the Ferry Building. Embarcadero at Sansome is the northernmost station along the Embarcadero, closest to tourism-heavy Pier 39 and Fisherman's Wharf. The bike path heading north along the Embarcadero is also much more bicycle-friendly than the southern route.

Most Traveled Routes

From	To	Rides
Harry Bridges Plaza (Ferry Building)	Embarcadero at Sansome	1,330
Townsend at 7th	San Francisco Caltrain (Townsend at 4th)	1,322
San Francisco Caltrain 2 (330 Townsend)	Townsend at 7th	1,116
Market at Sansome	2nd at South Park	866
Embarcadero at Sansome	Steuart at Market	811
2nd at South Park	Market at Sansome	798
San Francisco Caltrain (Townsend at 4th)	Harry Bridges Plaza (Ferry Building)	782
2nd at Townsend	Harry Bridges Plaza (Ferry Building)	757
Steuart at Market	Embarcadero at Sansome	717
Harry Bridges Plaza (Ferry Building)	2nd at Townsend	710

In Conclusion

We saw here that Bike Share is used mostly in San Francisco, by commuters, when it isn't raining, for rides under 15 minutes. It looks like BART commuters don't use the system to the extent that Caltrain commuters do. On weekends and holidays, visitors to San Francisco use the system to ride around town.

One dimension not explored in this analysis was popularity of each station among subscribers vs customers. With what we saw of the behavior of customers vs subscribers we could identify stations more popular with tourists or with commuters and potentially identify areas of the city with demand for future stations.

One recommendation that can be made is for station growth into SOMA. Users are already shown to be commuters heading to work, and SOMA is one area of San Francisco with a high density of businesses, close proximity to the existing stations, and currently lacking any Bike Share stations.

Post-contest Addendum

After submitting this entry to the bike share contest, I kept thinking to myself, "It doesn't make sense that BART riders don't make up a big group of Bike Share users. How else could I show where the Subscribers come from?" It turns out that information was in the original dataset all along, so I made a map of subscribers by zip code.

Which shows that many Riders are Subscribers who live in the East Bay, near a BART station. Clearly, some more analysis is in order to figure out why the BART Bike Share stations don't stand out among other San Francisco stations.

I did not perform any actual statistical tests; with figuring out D3, I didn't have time to re-learn T-tests and whatnot. I would maybe be interested in doing that to see if my conjectures have any basis in math.
Made with D3 and Open Office. Data from the Bay Area Bike Share Data Challenge.