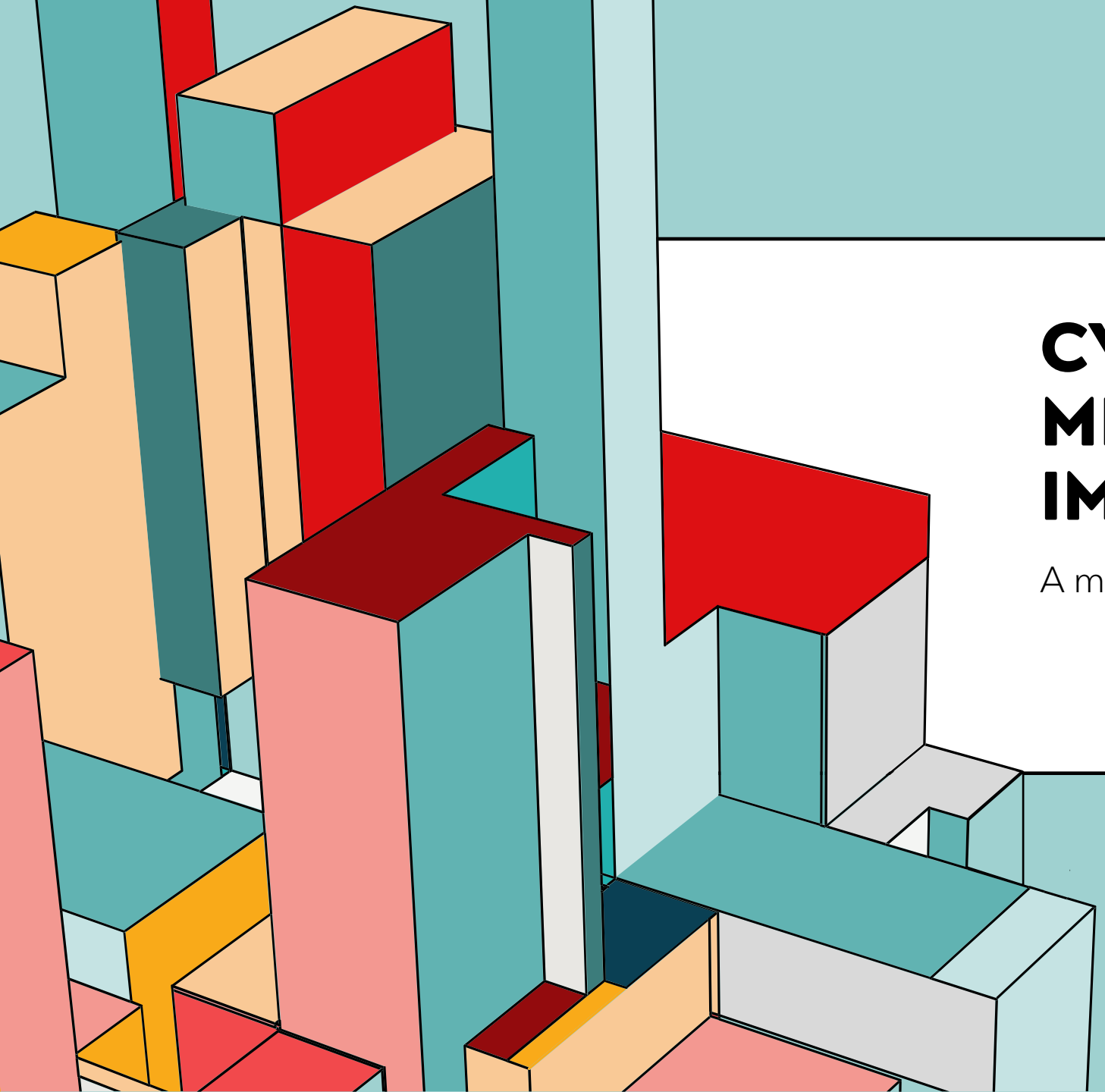


CYCLISTIC'S MEMBERSHIPS IMPROVEMENT

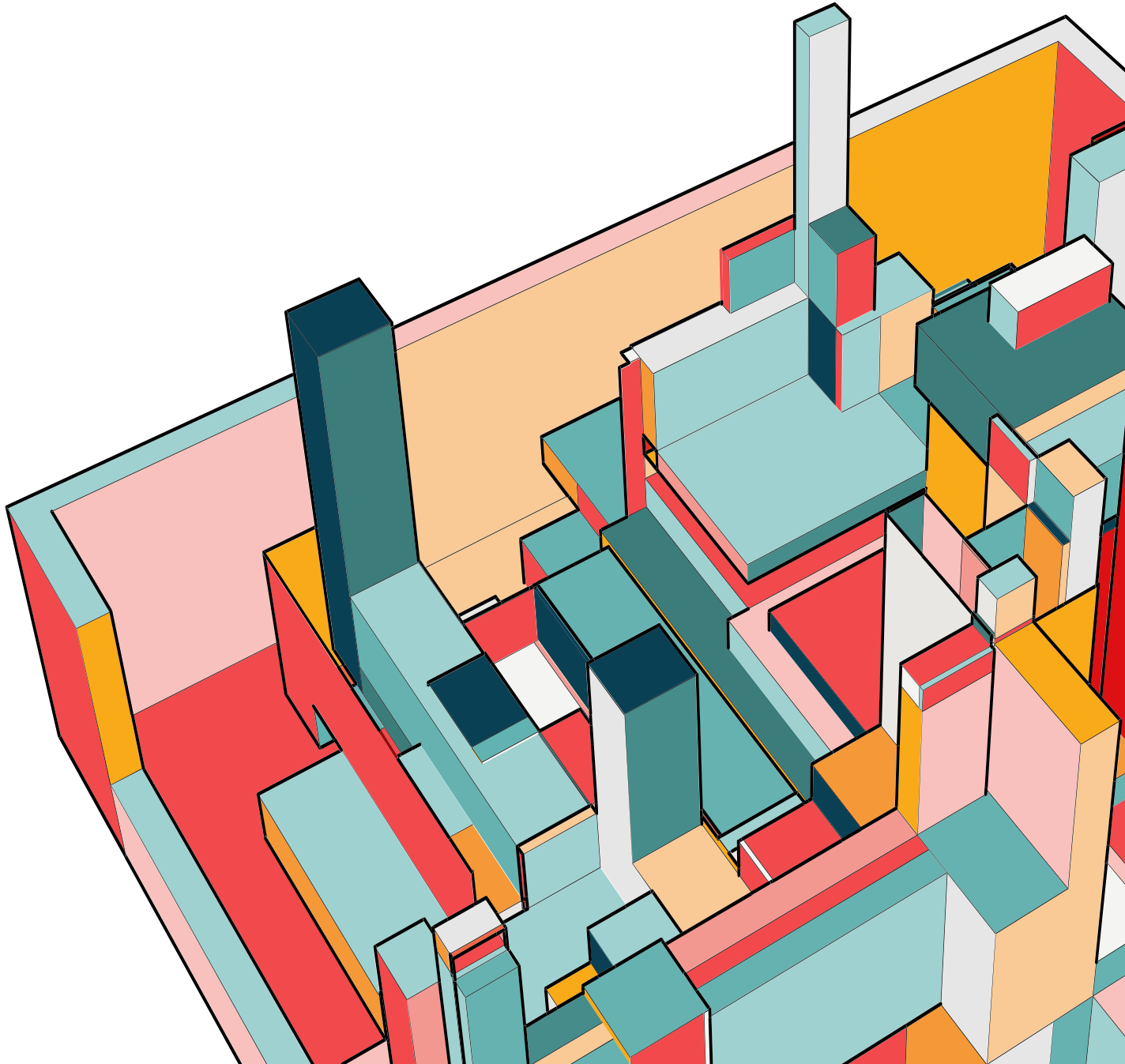
A marketing analysis for the future.

By: Marco Alegria – Nov 2023

QUICK OVERVIEW

Cyclistic's stakeholders require an improvement on the number of membership engagement for casual members. The final objective is to convert casual riders into members that use our bikes in general.

The data analysis and proposal will be performed thanks to the historical data recovered from Cyclistics data base. Based on this, the intention is to find trends on the use of different type of bikes, date and time and routes to propose the best marketing strategy.



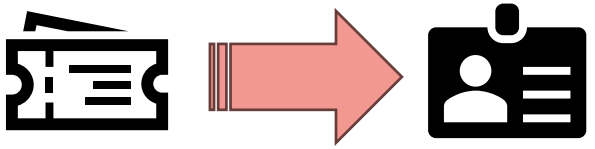
WHAT WILL WE FOCUS ON?

Members vs. Casuals use of bikes

Our first insight into the information, the first step of the data analysis, is going to be to find the differences of the use of the bicycles between our casual users and membership owners.

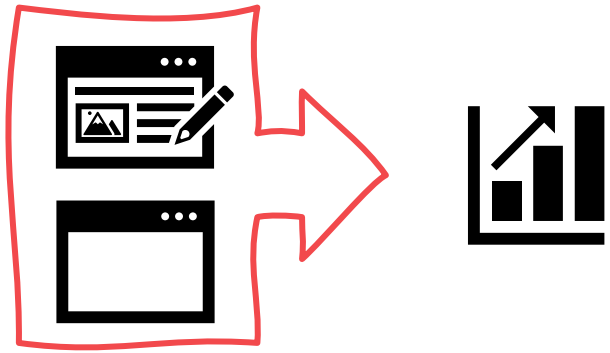
Attract casuals to a membership

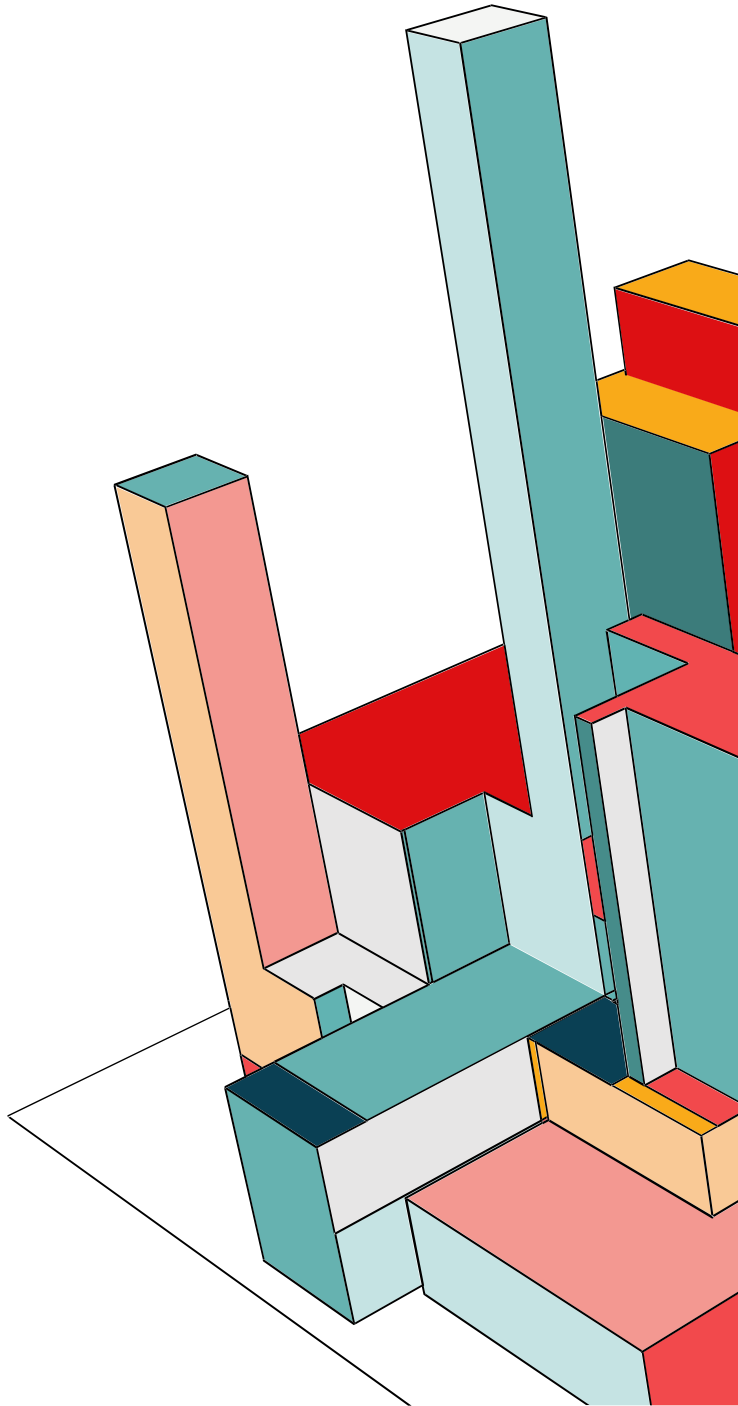
Can we find a way and reason to make our casual users to get them interested in becoming members?

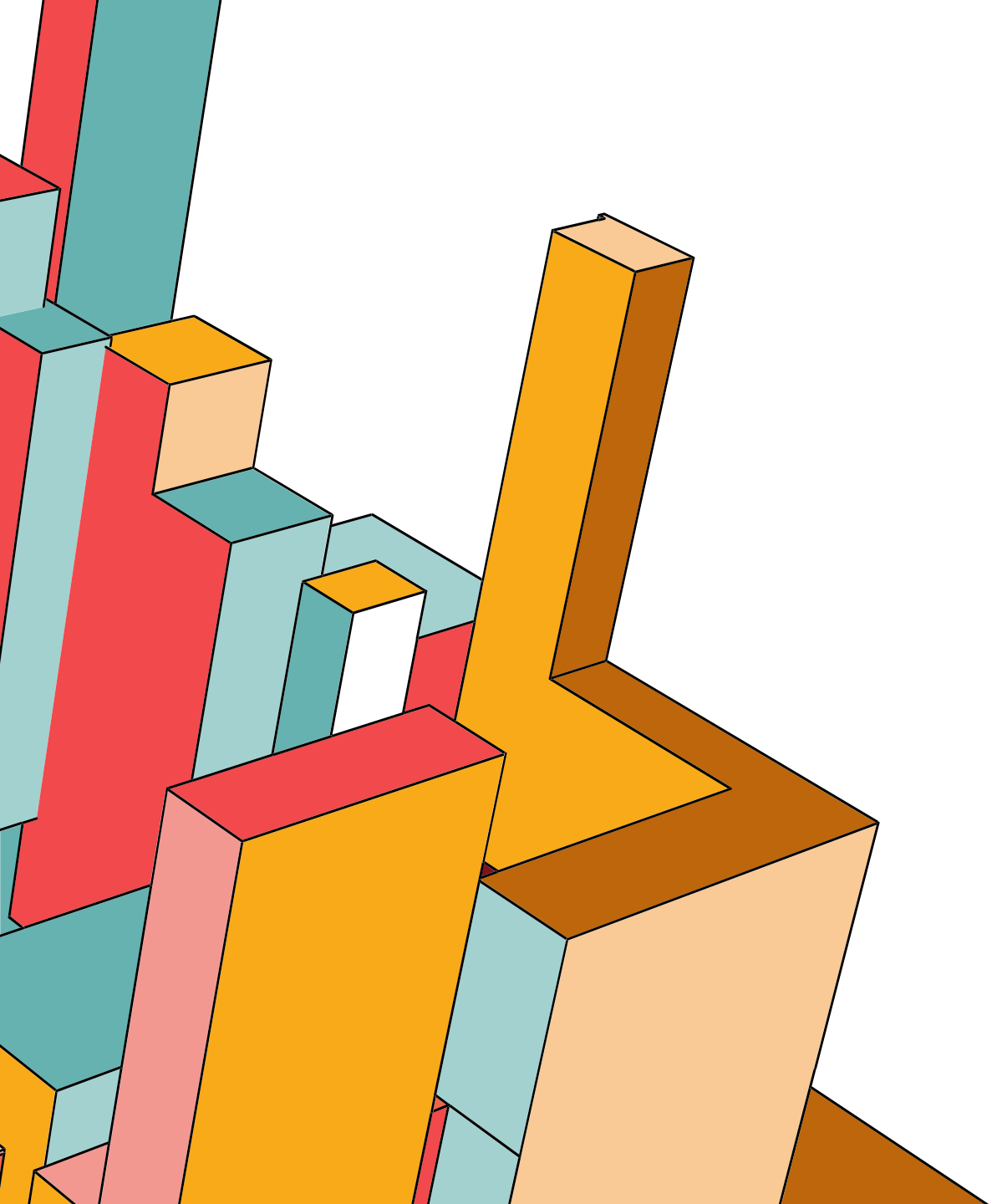


Digital media boost campaigns?

The final objective: Finding the best route to improve our engagement with our users. What would be the best marketing strategy to attract new users?







SOLUTION

TARGET AUDIENCE

Our main objective is to focus on our casual users. Focusing on the majority of type of bicycles used and the best zone to work in.

TARGET ZONES

Find the most populated zones where the users take our bikes and decide what is the best marketing strategy to engage this users.

MARKET REACH

Of our 3 types of bikes, what are the most used? What are the best locations and venues to publish the marketing campaigns? Does the casual member travel more / less distance with our bikes? Should we motivate the users with more distance traveled?

PREPARING OUR ANALYSIS

Data location and storage

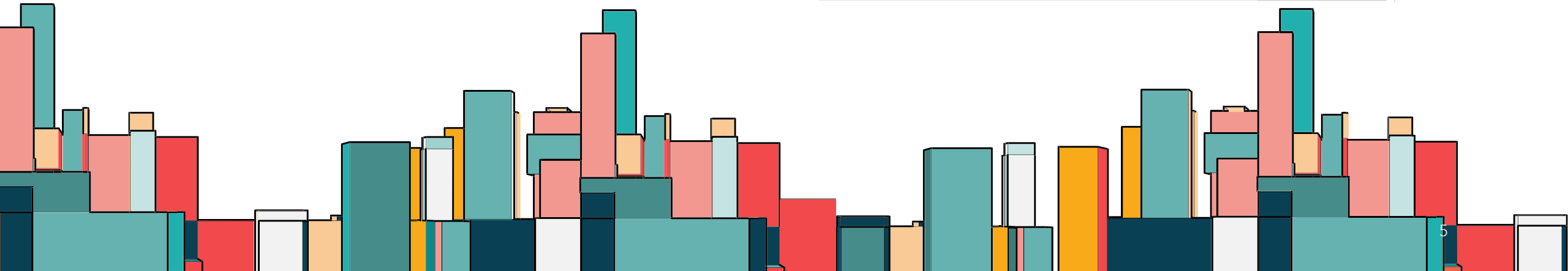
The data of the travels was provided to us in CSV format in a tabular form.

For making our analysis easier we will use R studio to build our datasets importing the CSV files into the R environment.

After this we are going to clean our data and make sure that only correct observations are including in our working data set.

Working data set.

rideable_type <chr>	started_at <chr>	ended_at <chr>	member_casual <chr>
classic_bike	2020-12-27 12:44:29	2020-12-27 12:55:06	member
electric_bike	2020-12-18 17:37:15	2020-12-18 17:44:19	member
electric_bike	2020-12-15 15:04:33	2020-12-15 15:11:28	member
electric_bike	2020-12-15 15:54:18	2020-12-15 16:00:11	member
electric_bike	2020-12-22 12:08:17	2020-12-22 12:10:59	member
electric_bike	2020-12-22 13:26:37	2020-12-22 13:34:50	member
electric_bike	2020-12-03 16:23:48	2020-12-03 16:33:39	member
electric_bike	2020-12-03 15:03:38	2020-12-03 15:12:39	member
electric_bike	2020-12-12 09:26:17	2020-12-12 09:26:35	member
electric_bike	2020-12-18 12:52:06	2020-12-18 12:52:23	member



UNDERSTANDING OUR DATA SET

As the data imported is not clean, we need to begin investigating what information is useful and meaningful for our analysis, for this we need to discard irrelevant information and empty observations.

As we can see, there is a lot of information about our riders, however, for improving our marketing campaigns we will focus on **locations**, **dates** and **types of bikes** used.

```
#What are the names of the columns on each observations?
names(bikeTrips)
```

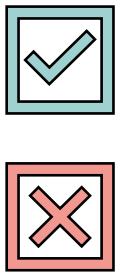
```
##  [1] "ride_id"            "rideable_type"      "started_at"
##  [4] "ended_at"           "start_station_name" "start_station_id"
##  [7] "end_station_name"   "end_station_id"     "start_lat"
## [10] "start_lng"          "end_lat"            "end_lng"
## [13] "member_casual"
```

```
#We know that there are 13 variables of each observation, that shows that each user has specifically information t
o its ride.
```

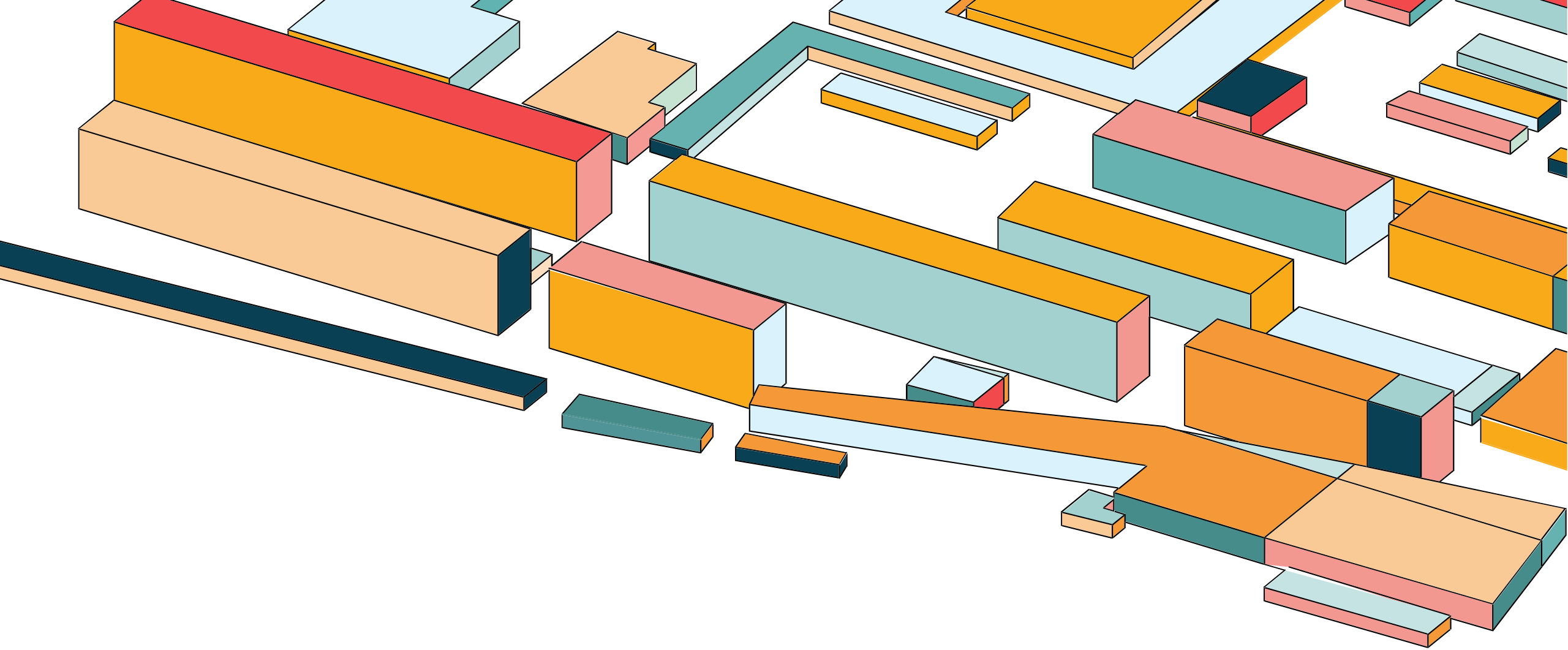
UNDERSTANDING OUR DATA SET

As expected, our raw information contains empty values on some categories, we only need data that is complete and correct, for this, we will create a new data set with clean data, and we will work with this on our analysis.

```
## Rows: 131,573
## Columns: 13
## $ ride_id           <chr> "70B6A9A437D4C30D", "158A465D4E74C54A", "5262016E0F…
## $ rideable_type     <chr> "classic_bike", "electric_bike", "electric_bike", "…
## $ started_at        <chr> "2020-12-27 12:44:29", "2020-12-18 17:37:15", "2020…
## $ ended_at          <chr> "2020-12-27 12:55:06", "2020-12-18 17:44:19", "2020…
## $ start_station_name <chr> "Aberdeen St & Jackson Blvd", "", "", "", "", "", "…
## $ start_station_id  <chr> "13157", "", "", "", "", "", "", "", "", "", "", ""…
## $ end_station_name  <chr> "Desplaines St & Kinzie St", "", "", "", "", "", ""…
## $ end_station_id    <chr> "TA1306000003", "", "", "", "", "", "", "", "", "",…
## $ start_lat         <dbl> 41.87773, 41.93000, 41.91000, 41.92000, 41.80000, 4…
## $ start_lng         <dbl> -87.65479, -87.70000, -87.69000, -87.70000, -87.590…
## $ end_lat           <dbl> 41.88872, 41.91000, 41.93000, 41.91000, 41.80000, 4…
## $ end_lng           <dbl> -87.64445, -87.70000, -87.70000, -87.70000, -87.590…
## $ member_casual     <chr> "member", "member", "member", "member", "member", "…
```



```
#We can identify by the glimpse of the data set that some columns show empty values, which we might have to ignore
when we work with the analysis.
```



THE NEW CLEAN DATA SET

LET'S FOCUS ON WHAT REALLY MATTERS

Using only the needed data

Yes, we have many information from our users, but reallisticaly if we are looking for a specific target and campaign to reach more engagement, we don't need to know the User ID of each one, or the Station IDs.
A clean data is easier and faster to analyze than a bunch of information of our users. And we will work with a resumed version of the important values of our dataset. As we can see here:

Bike_type <chr>	User_type <chr>	started_at <date>	ended_at <date>	start_lat <dbl>	start_lng <dbl>	end_lat <dbl>	end_lng <dbl>
classic_bike	casual	2020-12-19	2020-12-19	41.90	-87.63	41.90	-87.63
classic_bike	casual	2020-12-06	2020-12-06	41.92	-87.64	41.92	-87.64
classic_bike	casual	2020-12-22	2020-12-22	41.90	-87.62	41.90	-87.62
classic_bike	casual	2020-12-06	2020-12-06	41.95	-87.66	41.95	-87.66
classic_bike	casual	2020-12-04	2020-12-04	41.90	-87.62	41.90	-87.62
classic_bike	casual	2020-12-10	2020-12-10	41.90	-87.62	41.90	-87.62
classic_bike	casual	2020-12-20	2020-12-20	41.95	-87.66	41.95	-87.66
classic_bike	casual	2020-12-10	2020-12-10	41.92	-87.64	41.90	-87.63
classic_bike	casual	2020-12-28	2020-12-28	41.88	-87.65	41.89	-87.64
classic_bike	casual	2020-12-20	2020-12-20	41.92	-87.64	41.92	-87.64

1-10 of 10,000 rows Previous 1 2 3 4 5 6 ... 1000 Next

RELEVANT INFO INSIGHTS

Upon analyzing the clean data set, we've found that our users typically travel within a 2–3-kilometer radius.
This insight narrows down our target locations for digital campaigns. Next, we'll explore popular zones and seasons to fine-tune our strategy.

	maxDistance <dbl>	meanDistance <dbl>	minDistance <dbl>
distanceRides	29351.87	1992.315	0

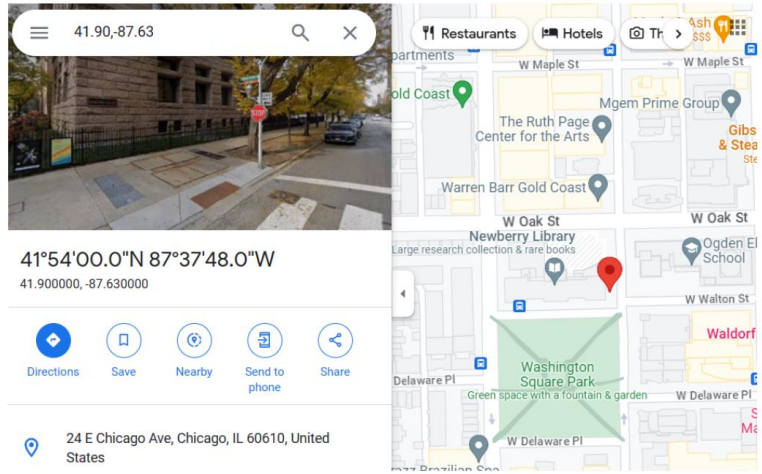
Target Zone:

popLng <dbl>	popLat <dbl>
-87.63128	41.90297

Target Season:

popDate <chr>
2020-12-18 15:55:33

Now we have one of the most important information, now we have the most popular place where we can focus the marketing campaign, we have now the Longitude and Latitude of the place where must of the people take and use our bikes. With a quick Google using this information, we know that the place is the zone nearby the Washington Square Park in Chicago, Illinois. Around this park we can start using new marketing campaigns to attract new members.

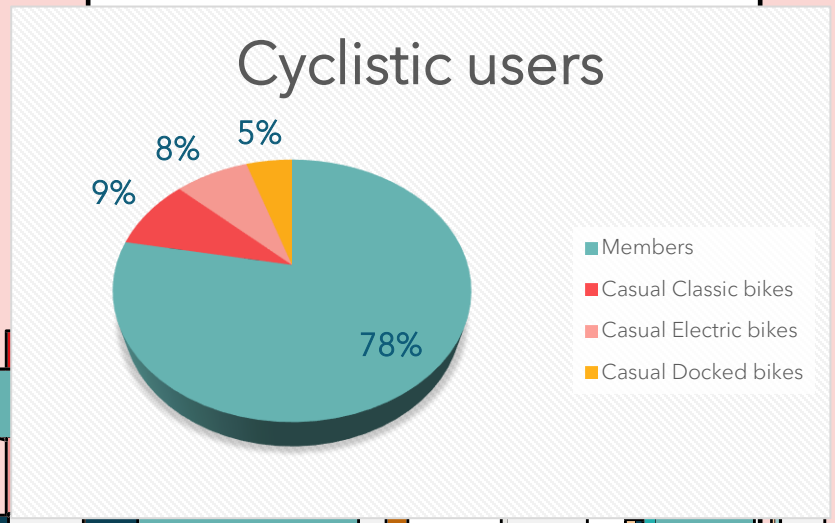


A simple search on Google Maps, brings us the relevant info we needed.
We may focus our campaigns **2Km around the Washington Square Park, Chicago**. And this will mainly be presented on **winter**.

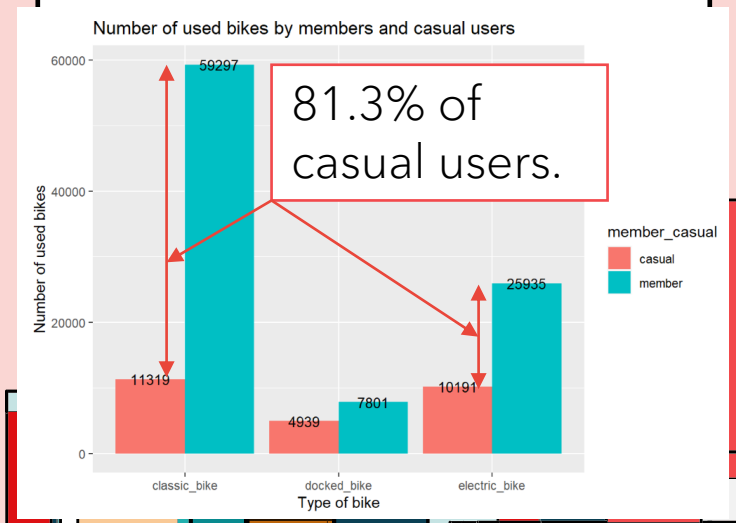
TECHNICAL INSIGHTS

The important numbers for deciding our targets.

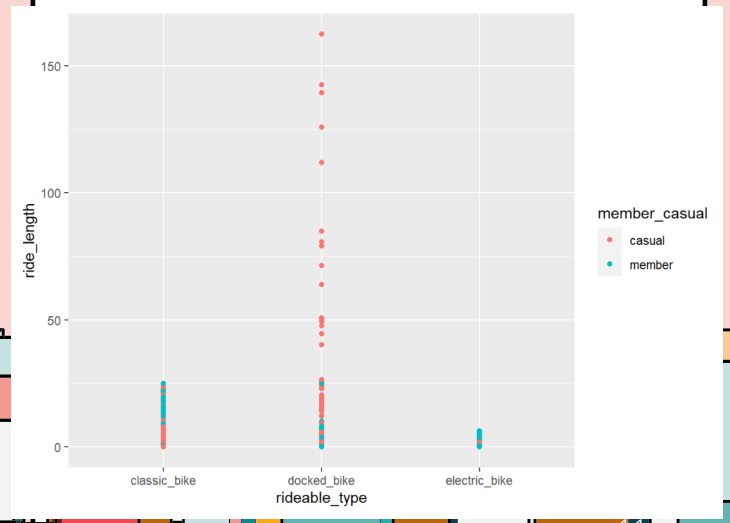
Out of all of our users, 22% of them are casual users. From this 22% the majority of them use more the **Classic (42.8%)** and **Electric (38.5%)** type bikes.



Thanks to the **Pareto principle**, we can be confident that focusing on the 80% of our casual members will gives us an advantage on the market.



Another good point to focus, is that the **docked users** are the ones that travel the **longest distance**. Another improvement can be placed on our **docked stations**.



THE FINAL VERDICT.

The objective

- Reaching a bigger market focusing our Casual users to convert them into Members

Different uses of bikes

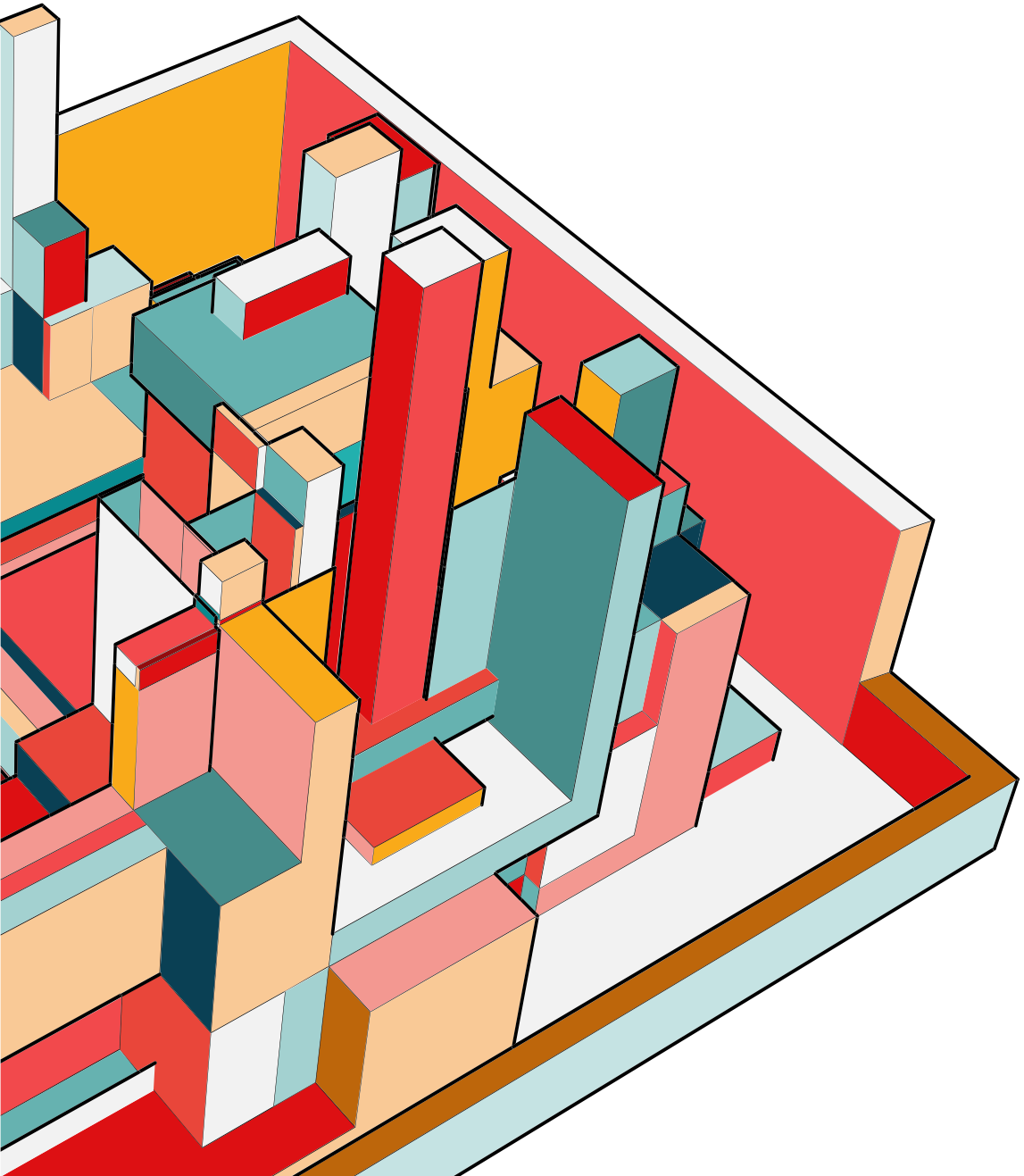
- Our casual members tend to use the **classic** and **electric** bikes more.
- And the **docked** users tend to travel longer distances.

Why change to a member?

- We can encourage our casual users to become members. Focusing on giving incentives according to travel distance.
- And coupons according to the times the service has been used.

The marketing strategy

- Show the benefits of becoming a member on advertisements near the **Washington Square Park**
- Improve the conditions of the campaign in Winter.
- Leave coupons on docking stations giving benefits according to travel distance.



SUMMARY

Our annual membership can be improved, by engaging our "**Pareto users",** thus, the Classic and Electric bike users. Using social networks, we can focus our marketing content with the users registered. And focusing the range of the campaigns around 2Km of the **Washington Square Park.**

Planning the marketing campaigns should be implemented on **Winter** to make a test phase and evaluate the impact on our market coverage.

THANK YOU

Marco Alegria

For more information on the Data Analysis process, the R

Notebook can be reviewed on the GitHub page.

www.contoso.com