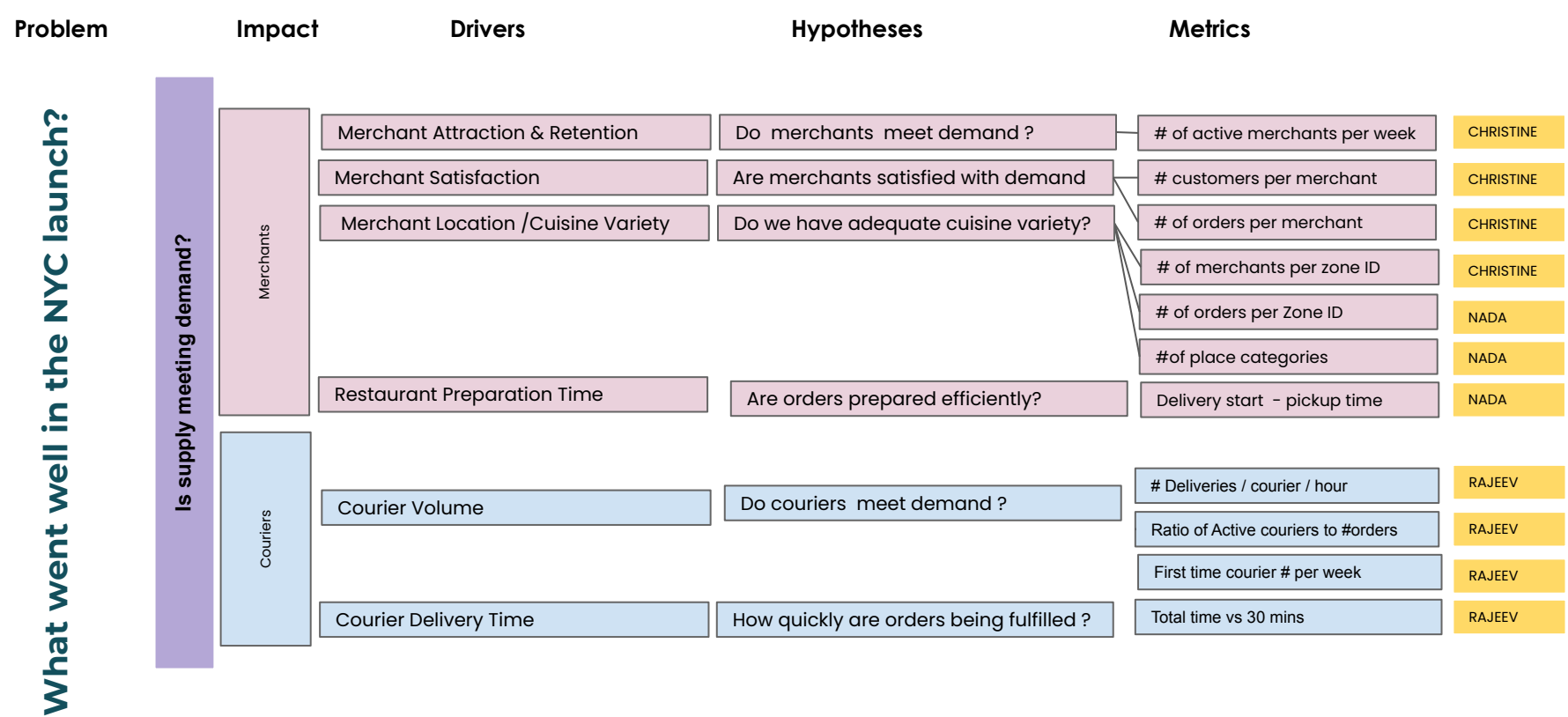
PROJECT JUNIPER

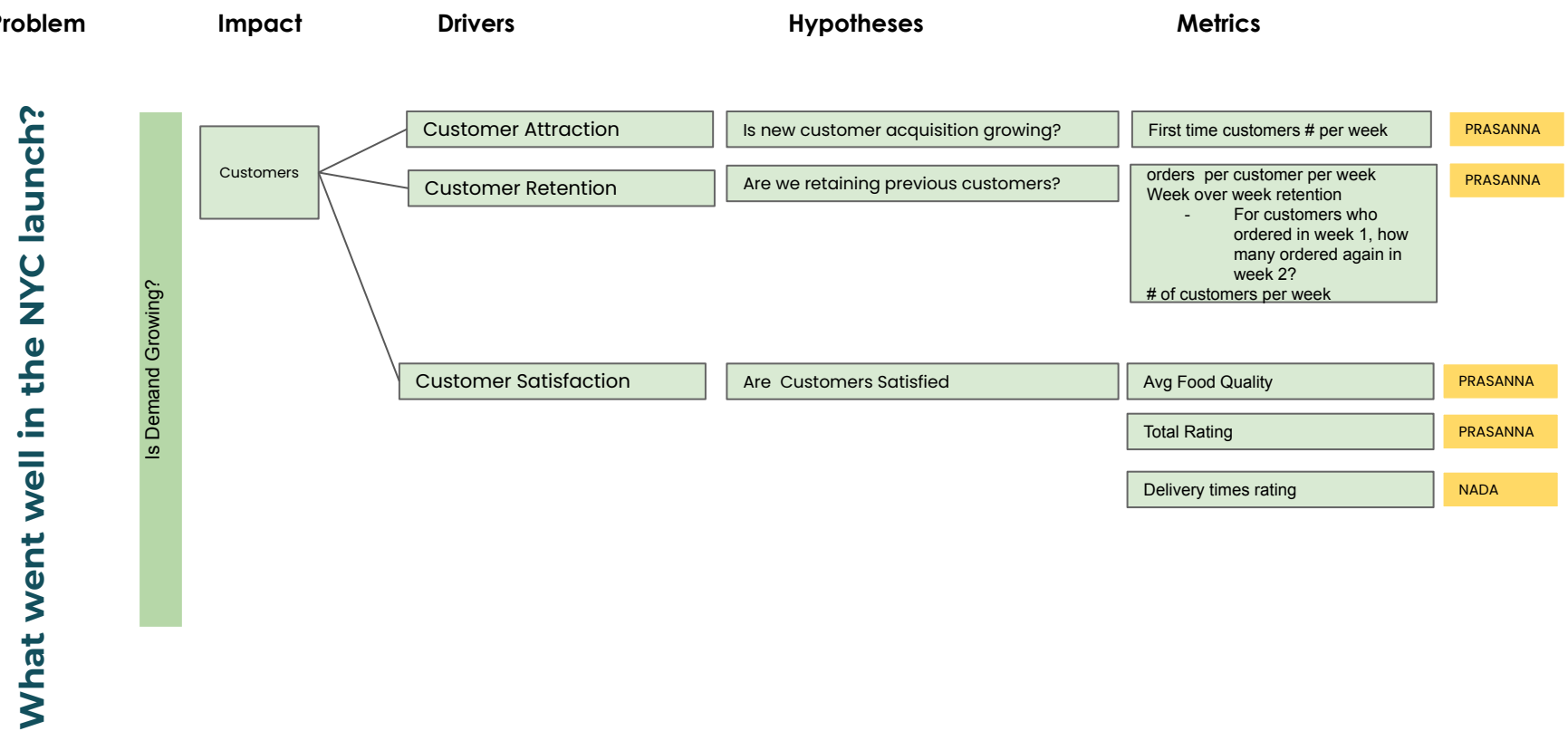
Issue Tree Development

PRASANNA MUPPIDI

Issue Tree - Project Juniper (Final)



Issue Tree - Project Juniper (Final)



```sql
SELECT
    timestamp_trunc(when_the_delivery_started, week) week,
    COUNT(DISTINCT delivery_id) volume
FROM
    Skillful_Data.Project_Data
GROUP BY 1
ORDER BY 1
```

Has our weekly delivery volume been increasing or decreasing?

```sql
SELECT
    customer_id,
    COUNT(DISTINCT delivery_id) as volume
FROM
    Skillful_Data.Project_Data
GROUP BY 1
HAVING volume >= 2
ORDER BY 1
```

customer_id	volume
1,311	2
5,139	2
6,987	2
7,922	2
9,666	6

How many customers have placed at least 2 orders?

```sql
SELECT
    place_category,
    COUNT(DISTINCT delivery_id) as volume
FROM
    Skillful_Data.Project_Data
WHERE place_category is not NULL
GROUP BY 1
ORDER BY 2 DESC
LIMIT 5
```

What are the top 5 most popular place categories?

place_category	volume
Italian	437
Burger	395
American	357
Japanese	335
Dessert	277

Which place category has the fastest delivery times?

```
SELECT
    place_category,
    AVG(TIMESTAMP_DIFF(when_the_Courier_arrived_at_dropoff, when_the_deliver
FROM
    Skillful_Data.Project_Data
WHERE place_category is not NULL
GROUP BY 1
ORDER BY 2
```

place_category	delivery_time
Grocery Store	77.27
Art Store	72.00
Specialty Store	68.20
Clothing	63.00
Department Store	59.44

```sql
SELECT
    place_category,
    COUNT(DISTINCT delivery_id) as volume
FROM
    Skillful_Data.Project_Data
WHERE place_category is not NULL
GROUP BY 1
ORDER BY 2 DESC
LIMIT 5
```

place_category	volume
Italian	437
Burger	395
American	357
Japanese	335
Dessert	277

What are the top 5 most popular place categories?

```sql
SELECT
    zi.borough,
    COUNT(DISTINCT pd.delivery_id) volume
FROM
    Skillful_Data.Project_Data pd
JOIN Skillful_Data.Zone_Information zi
ON CAST(pd.pickup_zone_id as integer) = CAST(zi.zone_id as integer)
GROUP BY 1
ORDER BY 2 DESC
```

Which pickup borough is doing the most orders?

borough	volume
Manhattan	5,027
Brooklyn	186
Queens	1

```sql
SELECT
    zi1.zone_id as pickup_zone_id,
    zi1.zone_name as pickup_zone,
    zi1.zone_id as dropoff_zone_id,
    zi2.zone_name as dropoff_zone,
    COUNT(DISTINCT pd.delivery_id) volume
FROM
    Skillful_Data.Project_Data pd
JOIN Skillful_Data.Zone_Information zi1
ON CAST(pd.pickup_zone_id as integer) = CAST(zi1.zone_id as integer)
JOIN Skillful_Data.Zone_Information zi2
ON CAST(pd.dropoff_zone_id as integer) = CAST(zi2.zone_id as integer)
GROUP BY 1,2,3,4
ORDER BY 5 DESC
```

What are the most common pickup-zone and dropoff-zone combinations?

pickup_zone_id	pickup_zone	dropoff_zone_id	dropoff_zone	volume
79	East Village	79	East Village	
79	East Village	79	Greenwich Village North	
234	Union Sq	234	Union Sq	
144	Little Italy/NoLiTa	144	TriBeCa/Civic Center	
79	East Village	79	TriBeCa/Civic Center	

```sql
SELECT
    pd.place_category,
    AVG(r.total_rating) as total_rating
FROM
    Skillful_Data.Project_Data pd
JOIN Skillful_Data.Ratings r
ON pd.delivery_id = r.delivery_id
GROUP BY 1
ORDER BY 2 DESC
```

Which place category has the best total rating?

place_category	total_rating
Russian	5.00
Coffee	4.75
Fast Food	4.46
Ice Cream	4.40
Caribbean	4.33

What is the busiest time of the day - Breakfast, Lunch or Dinner?

```
Select
     case when extract( hour from when_the_delivery_started) > 5 and extract( hour from when_the_delivery_started) < 12 then 'Breakfast'
          when extract( hour from when_the_delivery_started) >= 12 and extract( hour from when_the_delivery_started) < 17 then 'Lunch'
          else 'Dinner'  END as time_of_day
    , count(distinct delivery_id) as num_deliveries
from Skillful_Data.Project_Data
group by 1
```

time_of_day	num_deliveries
Dinner	3,277
Lunch	1,369
Breakfast	568

```sql
with first_week as (
Select
  customer_id
, min(DATETIME_TRUNC(when_the_delivery_started, week)) as cust_first_week
from Skillful_Data.Project_Data
group by 1
)

Select
   cust_first_week
  , count(distinct customer_id) as num_new_customers
from first_week
group by 1
order by 1
```

27/09/20 00:00	44
04/10/20 00:00	86
11/10/20 00:00	69
18/10/20 00:00	66
25/10/20 00:00	52

How many new customers were acquired each week? In which week were the most customers acquired?

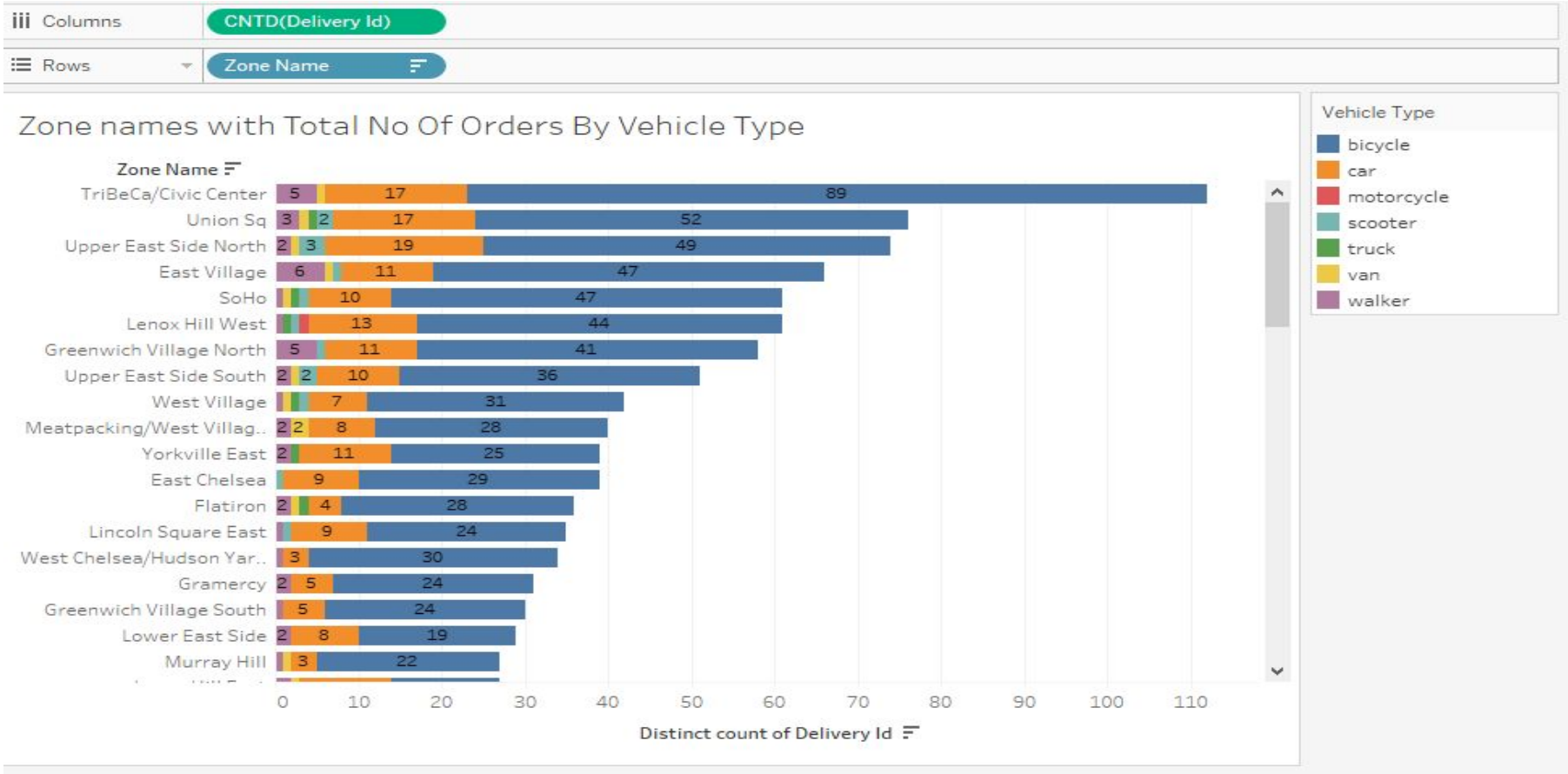
```
with delvery_value as (
 Select distinct
   delivery_id
 , sum(Total) as delivery_total
 from  Skillful_Data.Project_Data
 group by 1
 )

 Select
    case when delivery_total <10 then '<10'
         when delivery_total <20 then '10-20'
         when delivery_total <30 then '20-30'
         when delivery_total <40 then '30-40'
         when delivery_total <50 then '40-50'
         else '50+' end as value_group
  , count(delivery_id) as num_deliveries
 from delvery_value
 group by 1
```

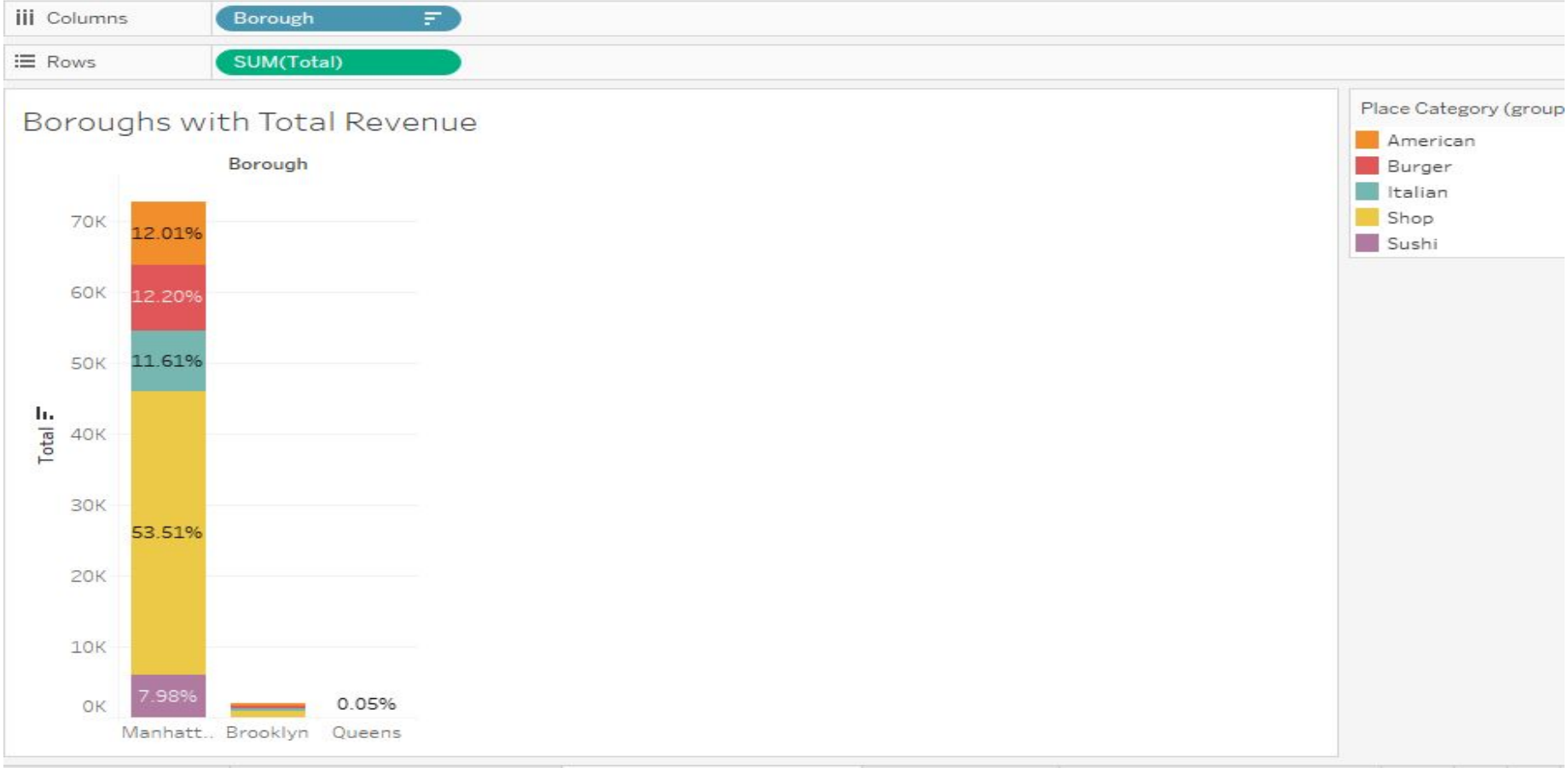
How many deliveries occurred in the denominations of $10, ranging from <$10 to $50+. (ie, <$10, 10-20,.... 50+)? What is the most popular range?

value_group	num_deliveries
<10	1,80
20-30	82
50+	36
40-50	23
10-20	1,64
30-40	35

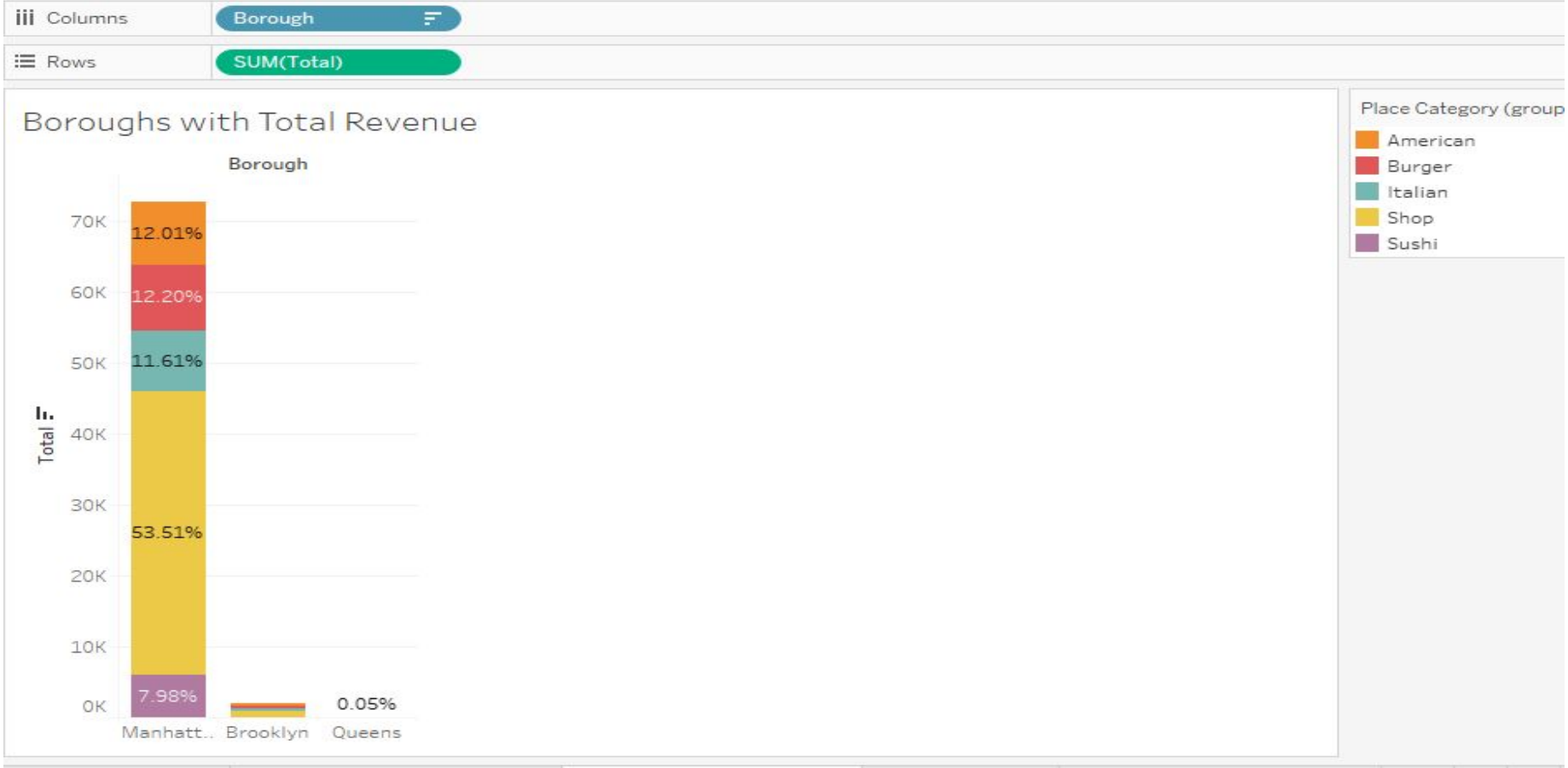
Recommendations



There seems to be a lot of bicycle deliveries, if we can switch to Car, MotorCycle , we can reduce the delivery time drastically and can attract new customers or retain customers

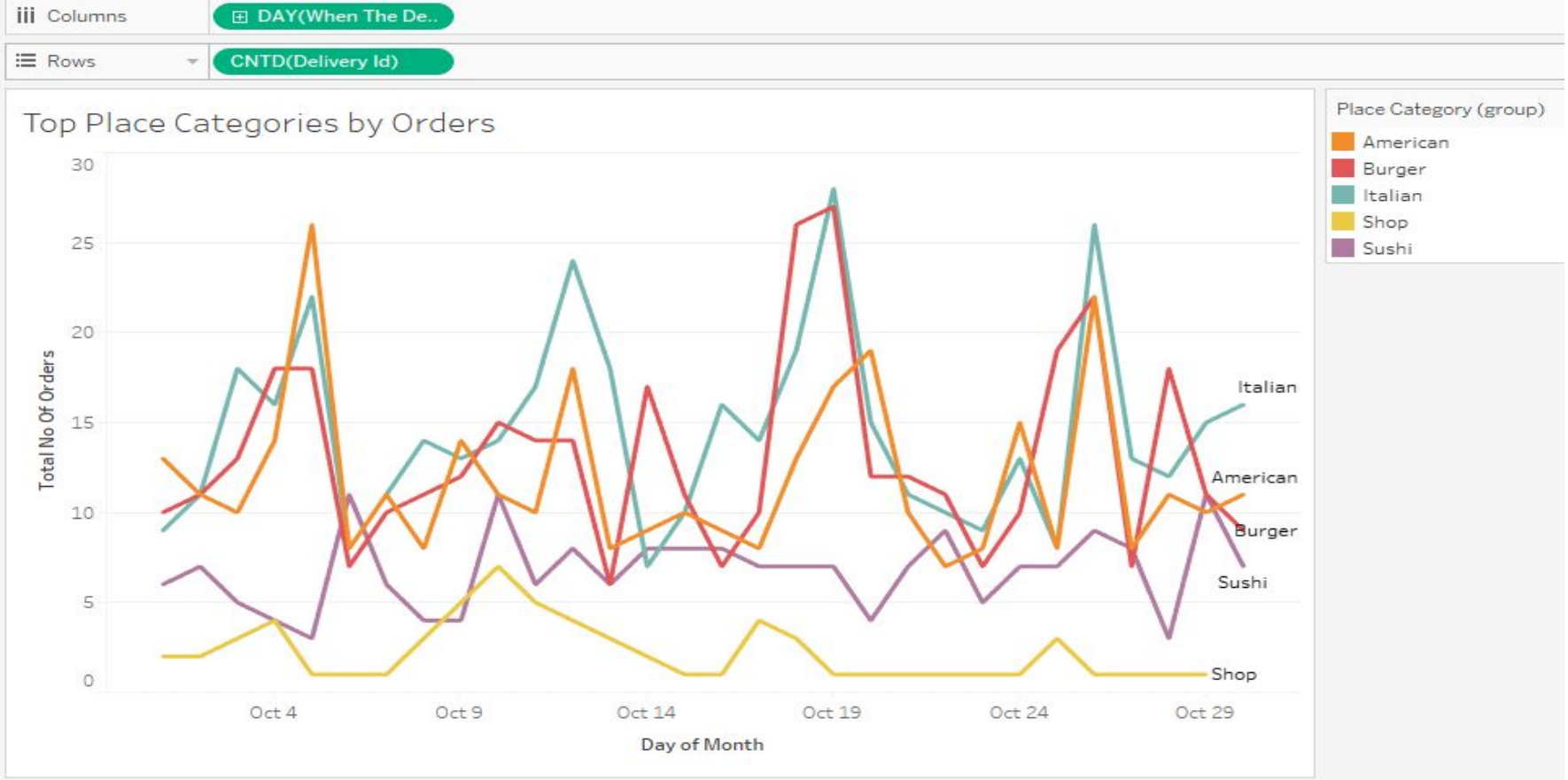


Manhattan is the borough with highest total revenue and we can see 3 Top Place categories are American, Burger and Italian

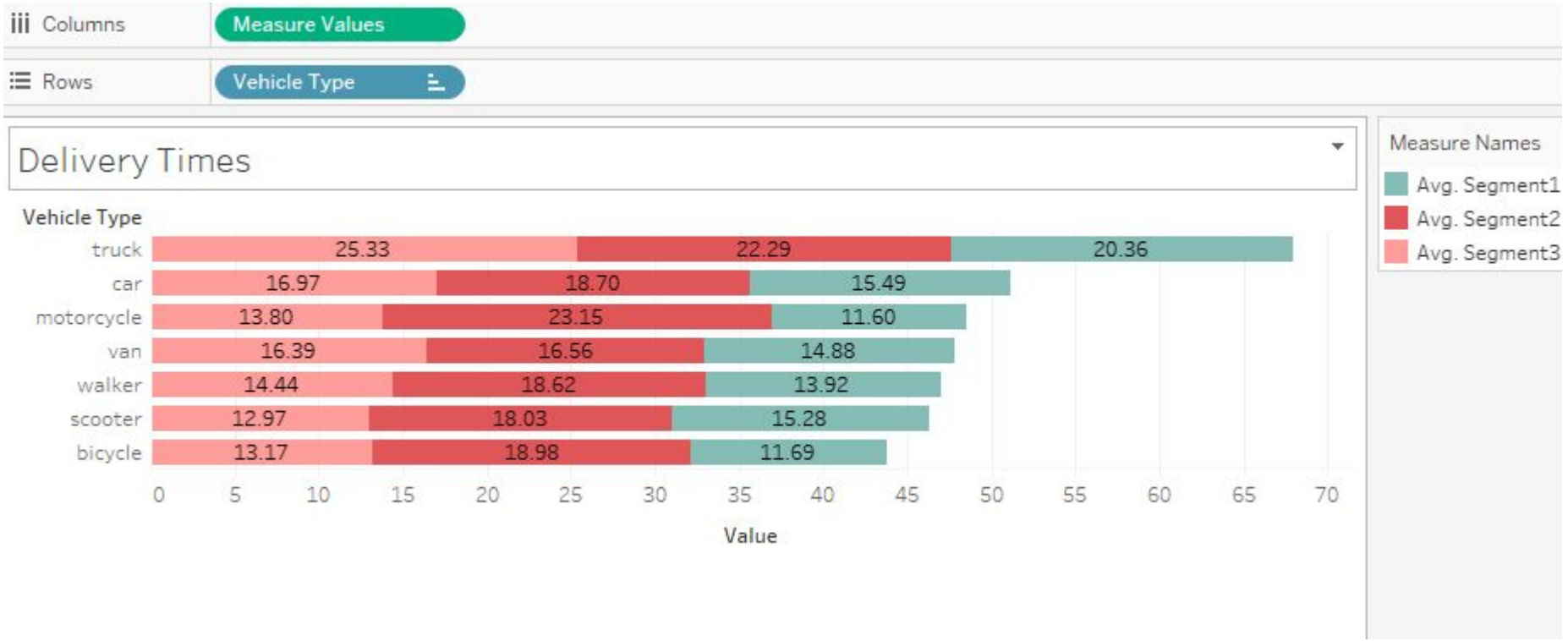


Manhattan is the borough with highest total revenue and we can see 3 Top Place categories are American, Burger and Italian

We need to look to First Quarter and second as, there's less no of Orders even we have reasonable no of restaurants available for delivery



We can focus more on promoting American, Burger and Italian, as these place categories are consistently in Top 3 Place categories by No Of Orders



We can see that the delivery agent is spending more in restaurant for pickup, if we can improve this, we can reduce delivery time by large fraction