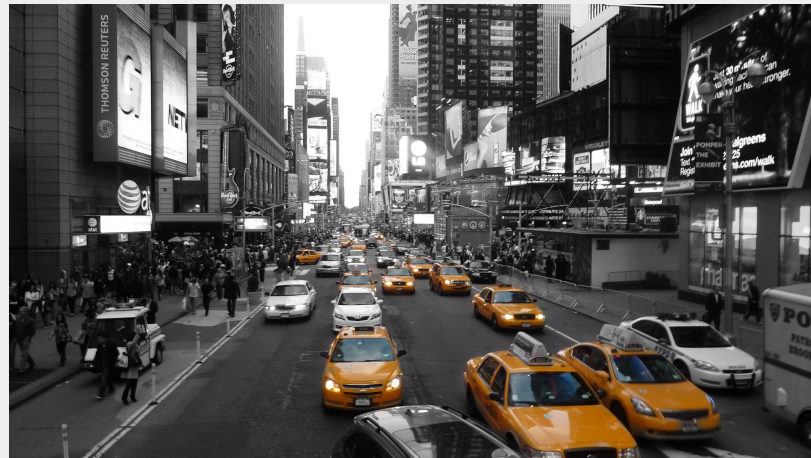# Vehicle collision analysis of New York City

Sunny Yoo
Haohan Zeng
Lynn Li

Background

1. We are analyzing the vehicle collision in New York City from 2012-2020
2. It includes the location, time, vehicle type, injury/ death statistics, contributing factor for each crash
3. Notice: some data has null/ unspecified value, also the year 2020 has incomplete data

Problem Statement

1. What time of the day police should be more concerned about drivers under influences? Is there also a seasonal pattern?
2. What percentage of incidents are preventable? Are divers more responsible in different borough?
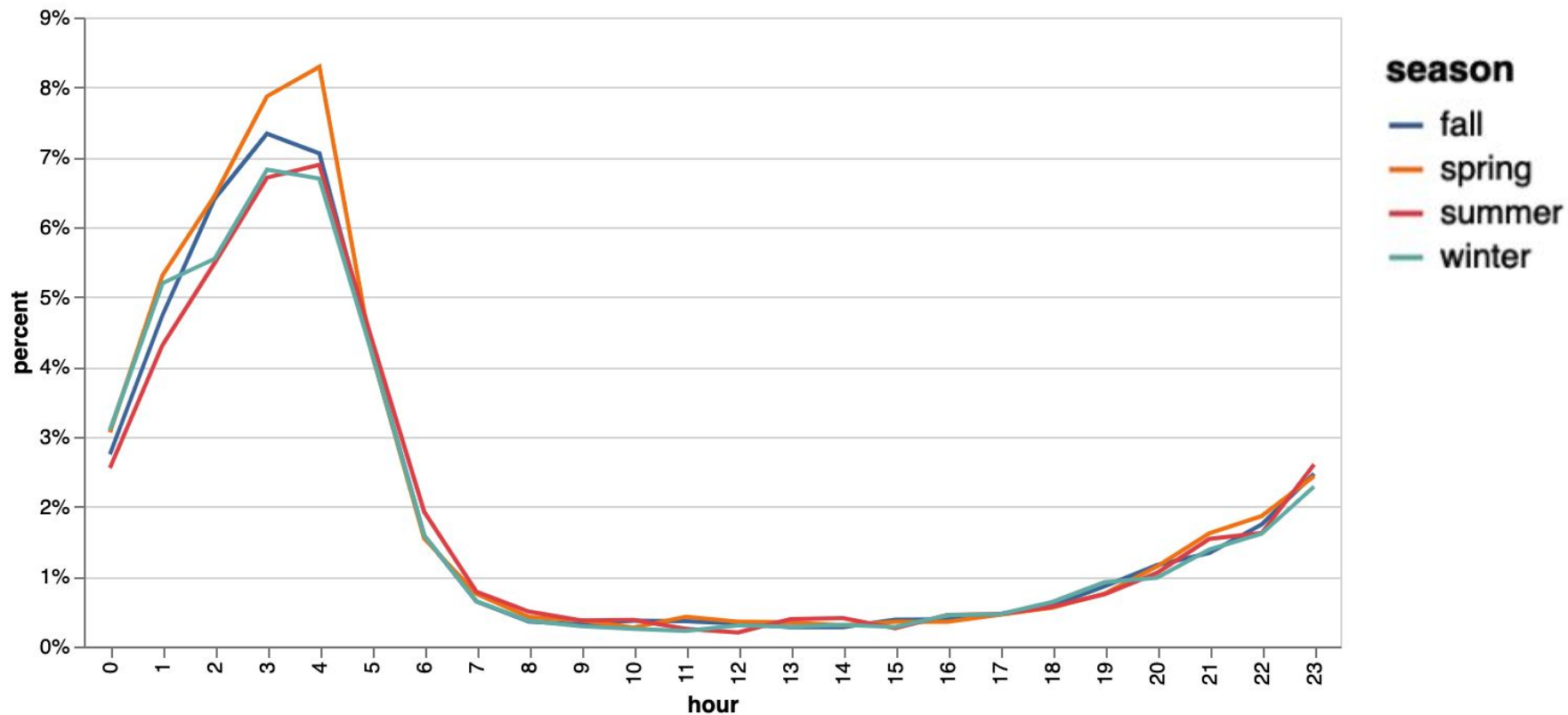3. What type of vehicle driver is more responsible?

Data Question 1

Is there any pattern for alcohol/ drugs incurred crash throughout the day for each season?
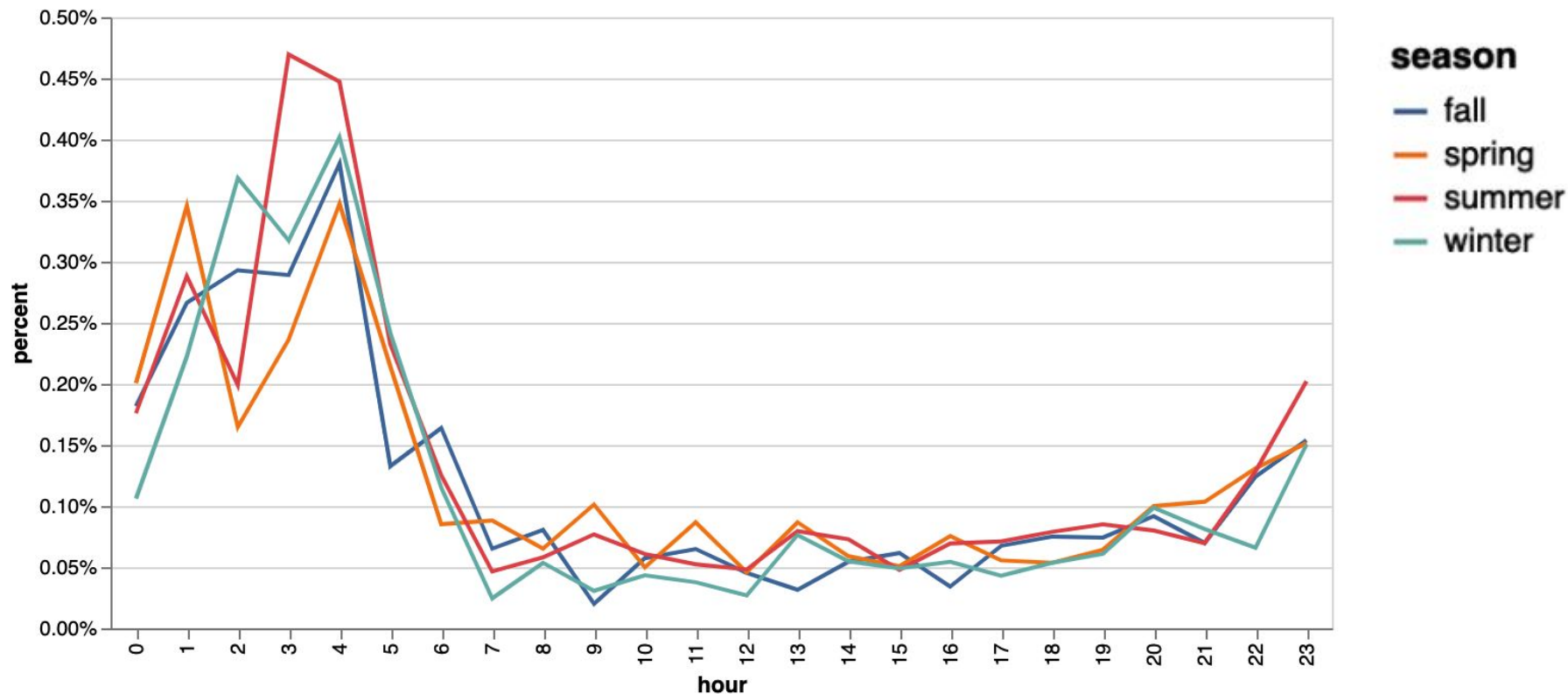
Daily pattern of total crashes by season
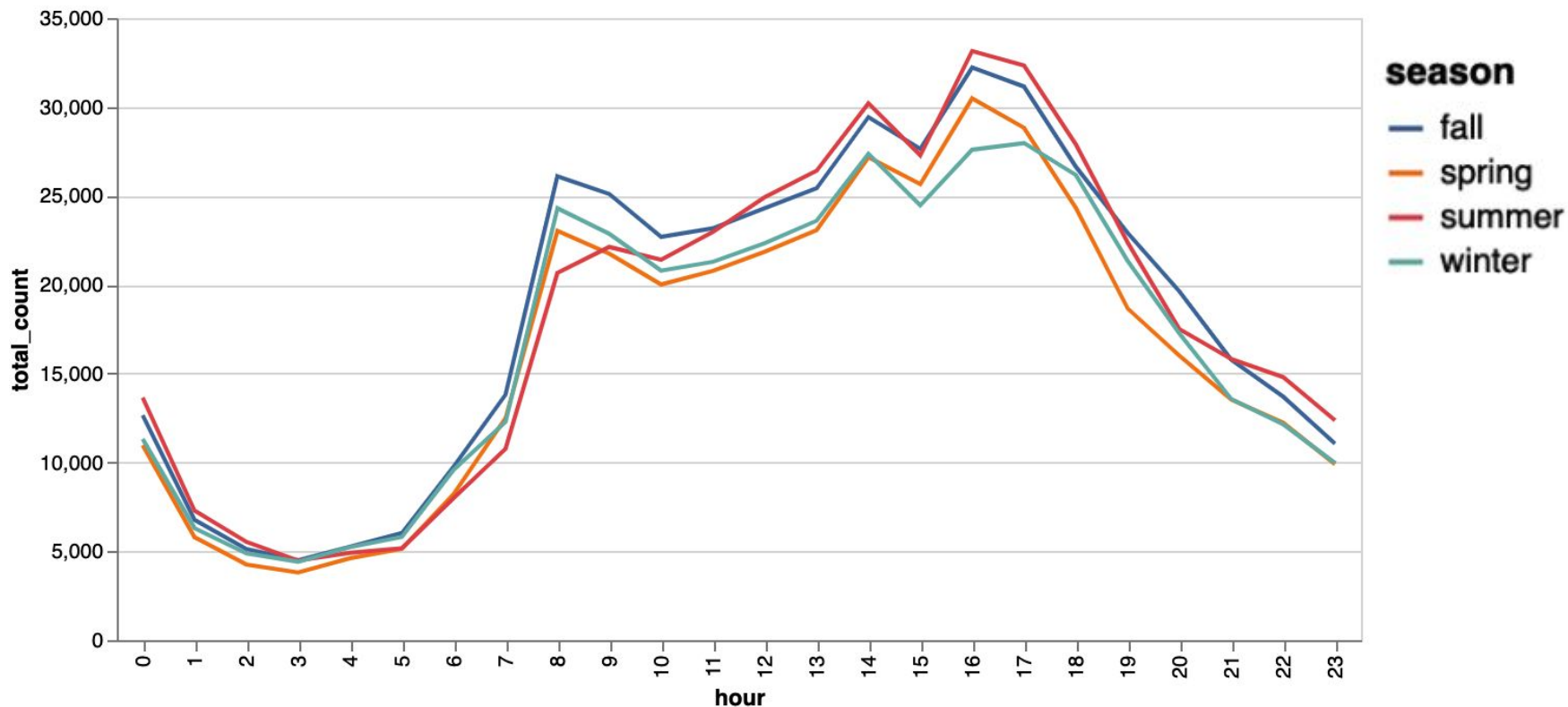
Daily pattern of influence-incurred crashes of vehicle 1 by season

Daily pattern of influence-incurred crashes of vehicle 2 by season

Data Question 2

Are there difference in collisions caused by Driver's responsibilities (using headphone, falling asleep) vs Non-Driver's responsibility (Brakes defective, animals action) in different Boroughs? Particularly, are there reasonable number difference of alcohol involvement?

How we define driver's cause vs passive cause

```python
driver_cause = ['Passing or Lane Usage Improper','Driver Inattention/Distraction','Turning Improperly',\
                'Following Too Closely','Failure to Yield Right-of-Way','Lost Consciousness',\
                'Unsafe Lane Changing','Unsafe Speed','Backing Unsafely', 'Passing Too Closely',\
                'Aggressive Driving/Road Rage','Fell Asleep','Failure to Keep Right','Cell Phone (hand-Held)',\
                'Texting','Other Electronic Device', 'Listening/Using Headphones','Fatigued/Drowsy',\
                'Traffic Control Disregarded','Driver Inexperience',\
                'Prescription Medication','Physical Disability',\
                'Drugs (illegal)','Illnes','Drugs (Illegal)','Illness','Eating or Drinking',\
                'Cell Phone (hands-free)','Cell Phone (hands-free)','Using On Board Navigation Device',\
                'Reaction to Other Uninvolved Vehicle','Cell Phone (hand-held)']
```

```python
passive_cause = ['Brakes Defective','Obstruction/Debris','Shoulders Defective/Improper','Pavement Slippery',\
                 'Traffic Control Device Improper/Non-Working','View Obstructed/Limited',\
                 'Glare','Lane Marking Improper/Inadequate','Other Lighting Defects', 'Pavement Defective',\
                 'Tire Failure/Inadequate','Headlights Defective','Animals Action','Accelerator Defective',\
                 'Tow Hitch Defective','Vehicle Vandalism', 'Windshield Inadequate', 'Other Vehicular',\
                 'Outside Car Distraction','Pedestrian/Bicyclist/Other Pedestrian Error/Confusion',\
                 'Driverless/Runaway Vehicle', 'Tinted Windows']
```
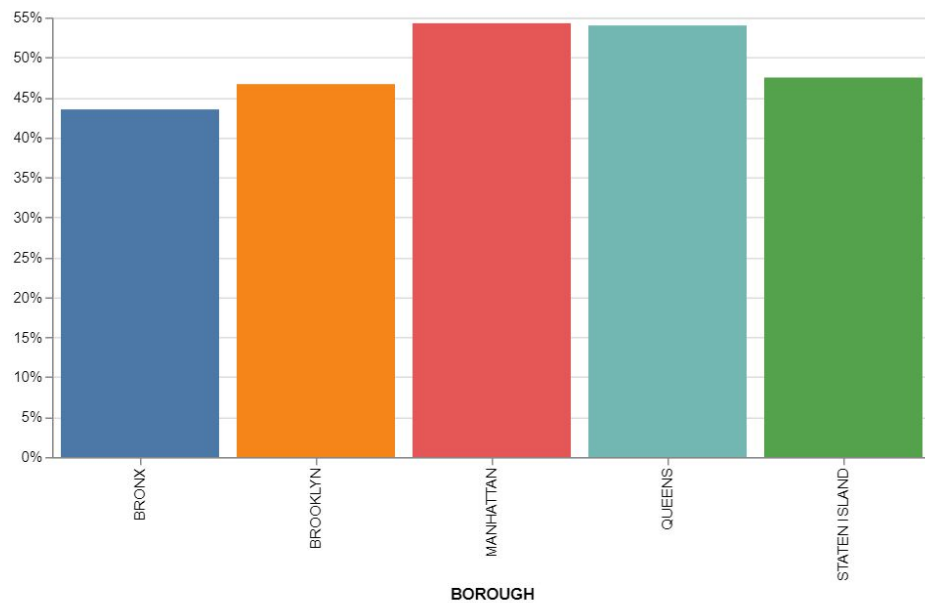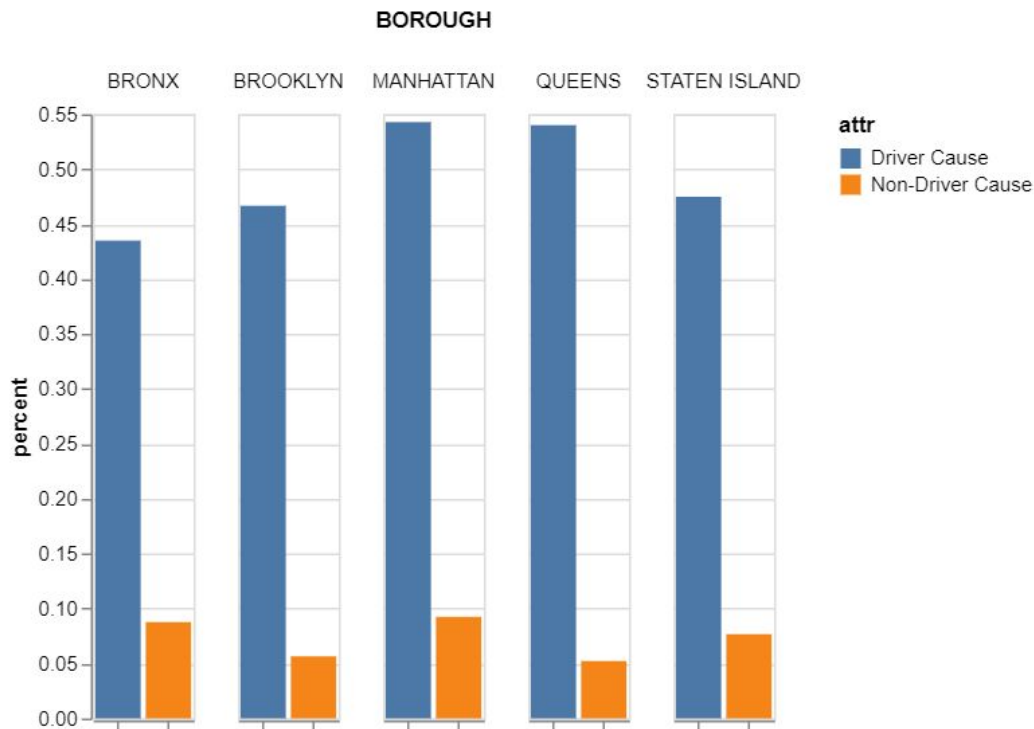
Collision caused by Driver's responsibilities

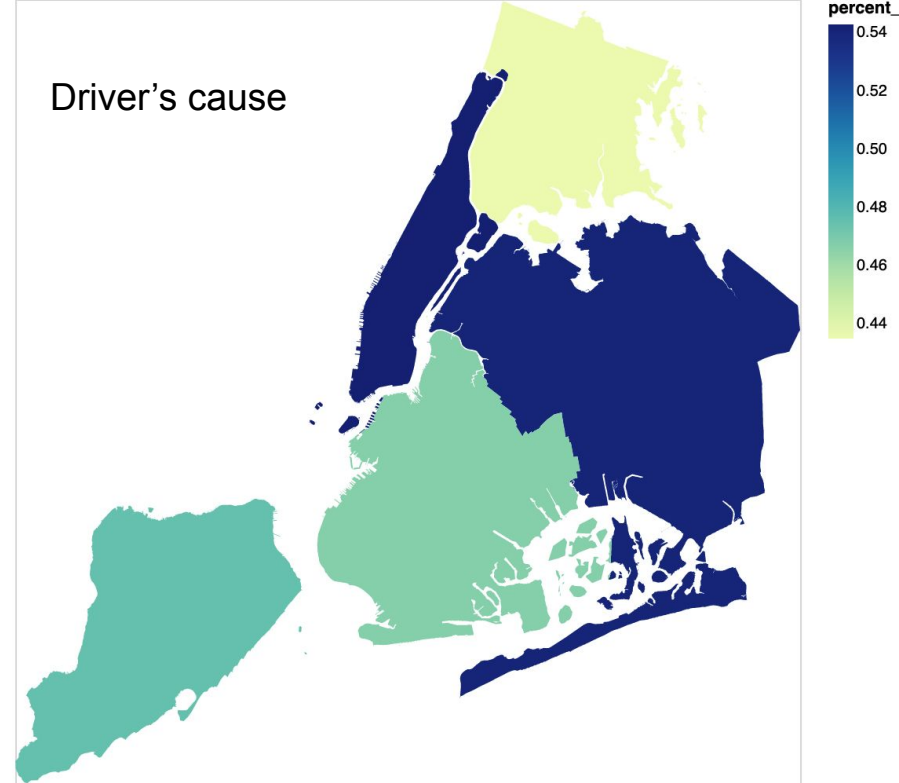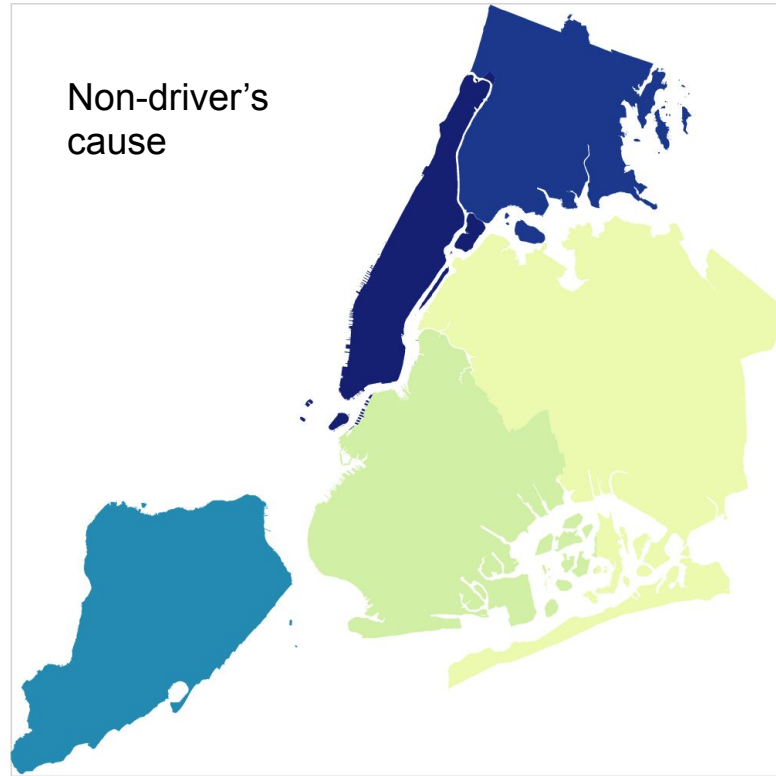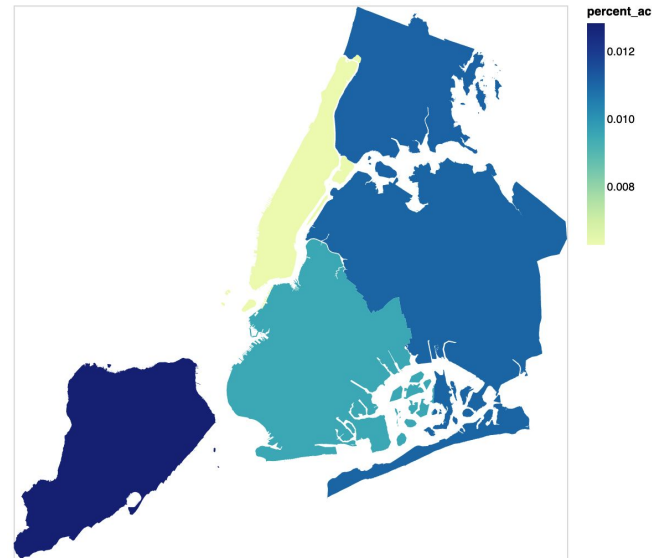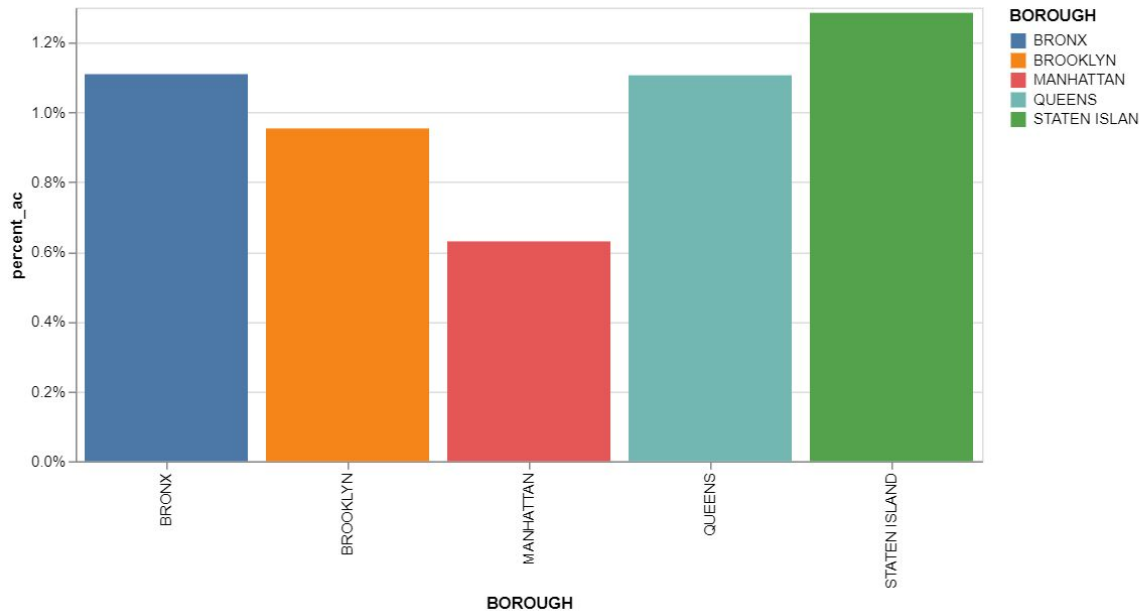Collision caused by non-driver's responsibilities
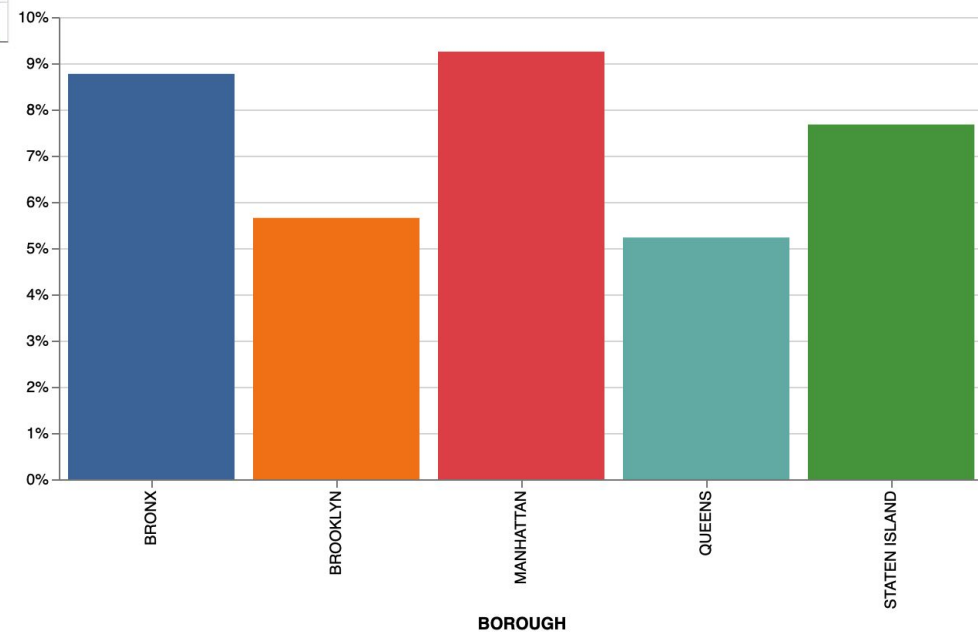
Comparison of collision caused by Driver vs Non-driver responsibilities

Comparison of collision caused by Driver vs Non-driver responsibilities

Non-driver's cause

Driver's cause

Alcohol involvement in different Boroughs

Data Question 3

Which types of vehicle are most likely involved in driver's caused crash versus non-driver's caused crash

Driver caused crash count by vehicle type

	VEHICLE TYPE CODE 1	count
271	PASSENGER VEHICLE	347944
327	SPORT UTILITY / STATION WAGON	164868
342	Sedan	119273
350	Station Wagon/Sport Utility Vehicle	101914
357	TAXI	28671
381	Taxi	13531
276	PICK-UP TRUCK	12343
411	VAN	11153
293	Pick-up Truck	8312
263	OTHER	7470
57	Box Truck	6466
322	SMALL COM VEH(4 TIRES)	6322
194	LARGE COM VEH(6 OR MORE TIRES)	6318
399	UNKNOWN	6025
202	LIVERY VEHICLE	5327
48	BUS	5036
61	Bus	4924
36	BICYCLE	3055
228	MOTORCYCLE	2935
389	Tractor Truck Diesel	2541

Non-driver caused crash count by vehicle type

	VEHICLE TYPE CODE 1	count
115	PASSENGER VEHICLE	48825
144	SPORT UTILITY / STATION WAGON	22657
151	Sedan	12084
154	Station Wagon/Sport Utility Vehicle	9738
157	TAXI	4509
185	VAN	2312
118	PICK-UP TRUCK	1760
112	OTHER	1680
85	LARGE COM VEH(6 OR MORE TIRES)	1362
168	Taxi	1312
141	SMALL COM VEH(4 TIRES)	1202
23	BUS	1178
87	LIVERY VEHICLE	954
178	UNKNOWN	879
18	BICYCLE	873
127	Pick-up Truck	823
97	MOTORCYCLE	751
26	Box Truck	695
25	Bike	683
29	Bus	583

Percentage of driver/ non-driver caused crash by vehicle type

	VEHICLE TYPE CODE 1	count	total	percent	attr
271	PASSENGER VEHICLE	347944	715236	0.486474	dirver cause
327	SPORT UTILITY / STATION WAGON	164868	313500	0.525895	dirver cause
342	Sedan	119273	183138	0.651274	dirver cause
350	Station Wagon/Sport Utility Vehicle	101914	150340	0.677890	dirver cause
357	TAXI	28671	50670	0.565838	dirver cause
115	PASSENGER VEHICLE	48825	715236	0.068264	non_dirver cause
144	SPORT UTILITY / STATION WAGON	22657	313500	0.072271	non_dirver cause
151	Sedan	12084	183138	0.065983	non_dirver cause
154	Station Wagon/Sport Utility Vehicle	9738	150340	0.064773	non_dirver cause
157	TAXI	4509	50670	0.088988	non_dirver cause

Percentage of driver/ non-driver caused crash by vehicle type

VEHICLE TYPE CODE 1

attr
- dirver cause
- non_dirver cause
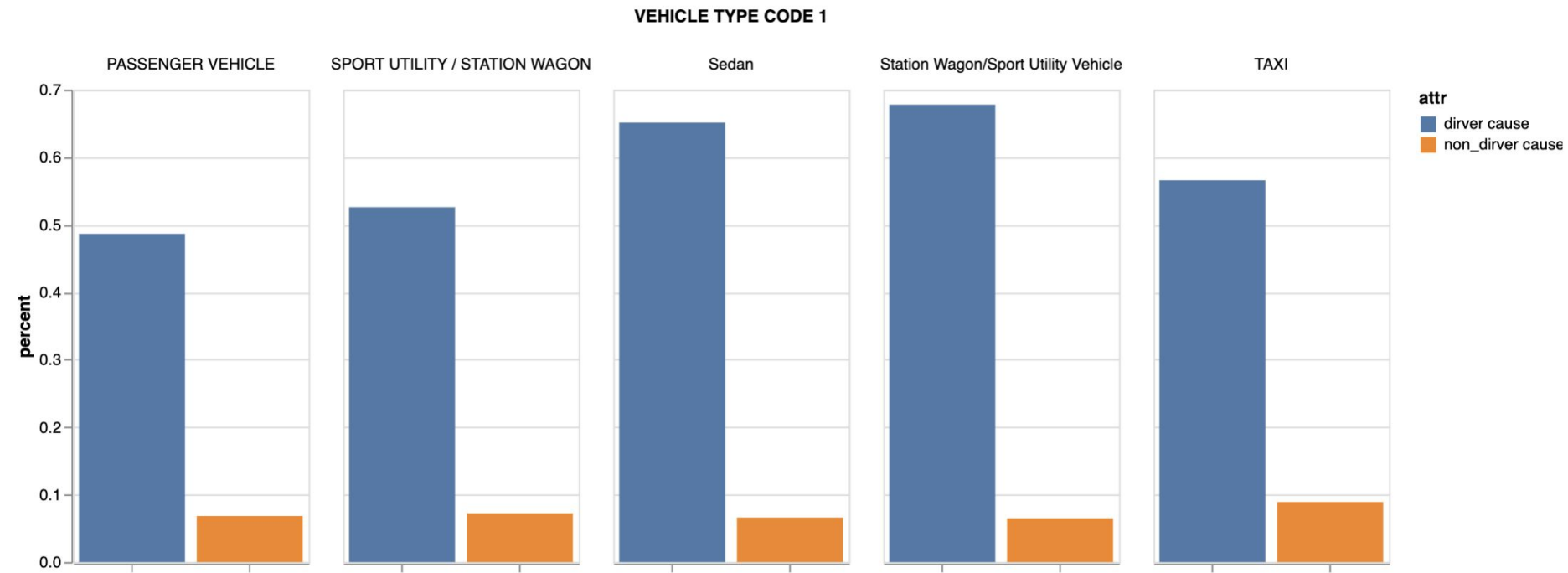
Summary

1. There's a clear peak of influence-incurred crashes at late night, especially between 3am-5am for both vehicle 1 and 2
2. Spring has a higher frequency of influence related late night crashes
3. Collisions caused by the responsibilities from driver are much higher than those caused by non-driver responsibilities
4. There isn't reasonable number of difference in alcohol involvement across different boroughs. Therefore, it is hard to conclude that collisions involving the use of alcohol happen more in particular borough.
5. The type of vehicles which are most common involved in crash are Passenger Vehicle, SUV, Sedan and Taxi. SUV has the highest percentage for driver's responsibility and Taxi has the highest percentage for Non-driver's responsibility

Implications

1. The police should look out for drunk drivers at late night between 3am-5am especially during the Spring.
2. Comparison between driver vs non-driver responsibility in different boroughs was an interesting analysis because it shows that passive causes happen more in the borough with less density (ex: Staten Island, Bronx) compared to higher density places like Brooklyn and Queens. Therefore, it'll be important for city to focus on fixing passive causes like Road defection to reduce the collision.
3. Taxi driver should be more care about non driver effect, and Station Wagon driver should be more care about driver effect.