**Data Description:**

The dataset used for analysis is obtained from NYC Open data program. It contains over 2 million rows with 29 columns. The dataset provided us multiple columns about the geospatial and temporal information about the accident using columns like data and time of accident, latitude and longitude, borough, and street name information etc. It also has collision\_id assigned which can be used to join with multiple other datasets which we have obtained from NYC Open data.

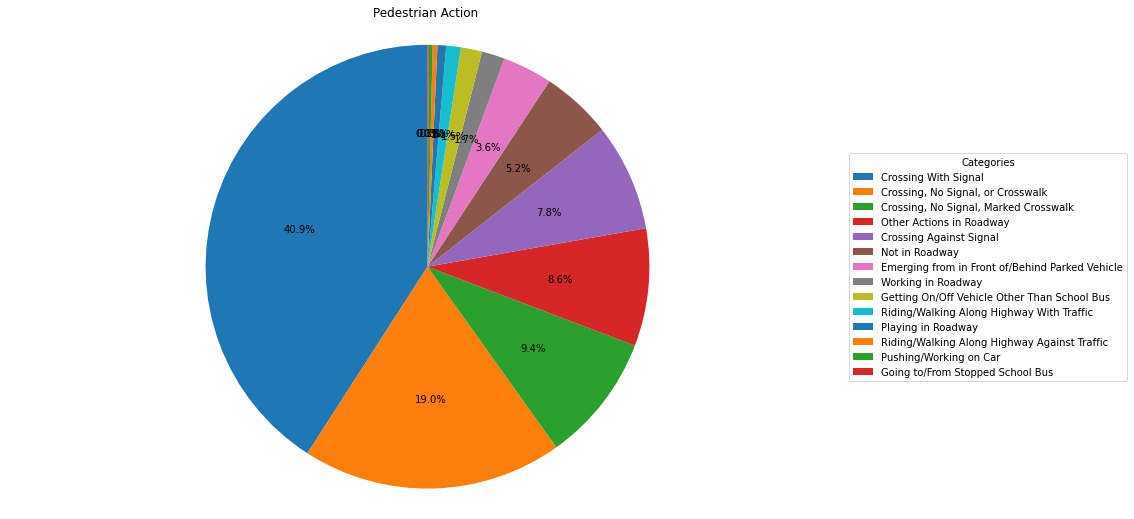
We used to other datasets namely Motor Vehicle Collisions – Vehicles, Motor Vehicle Collisions – Persons which give even more detailed information about the accident like the make and model of vehicle, its registration status, license status of the driver, the primary and other reasons of the accident. We used these datasets to join with the main dataset we obtained to understand the causality behind the accidents involving multiple groups such as pedestrians, motorists, and cyclists.

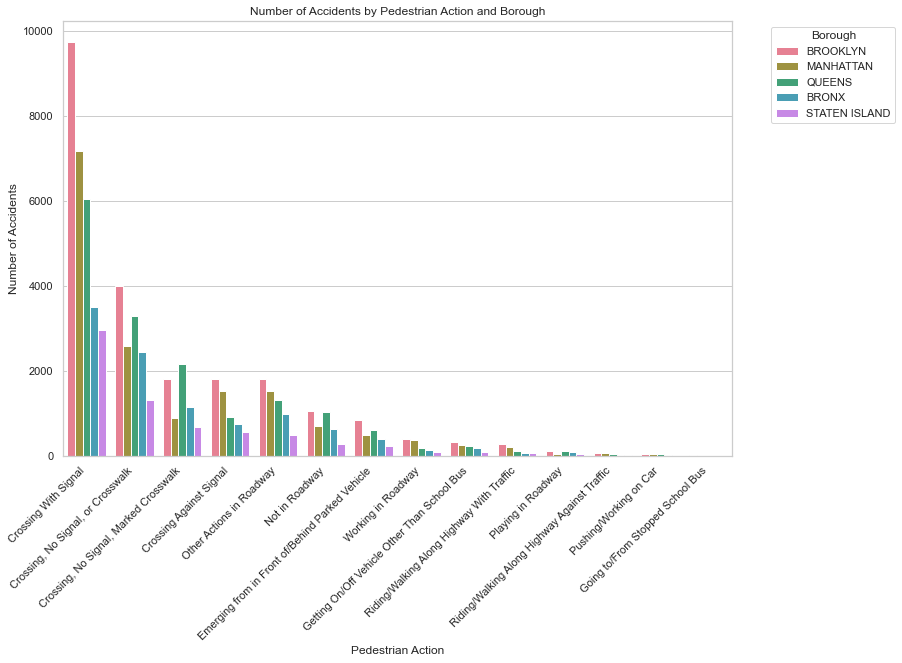
**Data Pre-Processing**

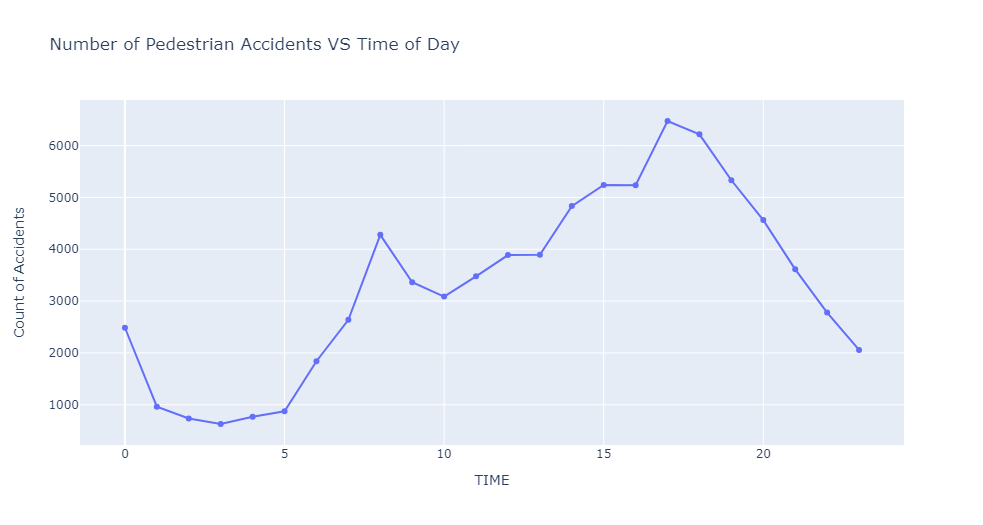
One of the important tasks involved is finding out the null values and categories involved in different attributes of the dataset. For example make of the vehicle has multiple vehicle models and we classified them into different sub-categories.

Since there are many columns and many categories of analyses that can be obtained, a row cannot be completely ignored based on one null value in a single column instead we will ignore when that row when it is considered in that specific analysis.   
For example, when analysing the total number of accidents happened at 2:00 AM in the morning, a column with null value in pedestrian\_action can be still considered, but when analysing the contributing factor count for the accident, the row cannot be considered if the value is null in that column. We also found out that there are 400,000 rows that had null value for borough while there is latitude and longitude information available for them. We calculated the borough values and populated those values.

PEDESTRIAN ACTION:







One of the main reasons why pedestrian accidents happen is “Crossing With Signal”.

Reason:

Given that the pedestrians are crossing with signal we can say that the error is from the side of vehicle driver. On searching for the reasons, we found out the following are the main reasons:

1. Most number of accidents are happening at the intersection points of roads.
2. Next contributing reason is the inattention/distraction of the driver, this is further expanded in the next sections where we have explained with numbers this reason.
3. Failure to yield right of way is another significant reason causing the accidents.

There is a confusion in the concept of right of way, in few states the pedestrians have the absolute authority in the right of way while in few states it is not. The right of way in NYC states that:

1. Drivers must yield to pedestrians legally using marked and unmarked crosswalks.
2. If stopped at a red light and the light turns green, drivers must wait if a pedestrian is still crossing.
3. Motorists must yield to pedestrians on the left or right when making a turn.
4. Drivers [must yield to both vehicles and pedestrians](https://dmv.ny.gov/about-dmv/chapter-5-intersections-and-turns) when entering the road from a driveway, alley, private road, or any other non-roadway.
5. A yield sign gives the right of way to pedestrians in a crosswalk. If a driver goes past a yield sign without stopping and hits a pedestrian (or another vehicle in the intersection), the accident will be deemed [prima facie (automatic) evidence](https://www.ny.gov/sites/ny.gov/files/atoms/files/Final_VT_Law.pdf)  of failure to yield the right of way.

The fifth rule is introduced as a part of VISION ZERO project, where the drivers had severe consequences if they hit a pedestrian without yielding right of the way.

As a supporting factor for our conclusion, we can see that the greatest number of pedestrian accidents are happening at the busiest hours in the city.

1. **Rush Hours:** The afternoon peak from 3 PM to 5 PM coincides with evening rush hours when people are getting off work or leaving school. Increased vehicular and pedestrian traffic during this time can contribute to a higher likelihood of accidents.
2. **Visibility:** In the evening hours, decreasing natural light as the day progresses could contribute to reduced visibility for both drivers and pedestrians. Diminished visibility might increase the risk of accidents.
3. **Distractions:** Afternoon hours are often associated with higher levels of traffic, and both pedestrians and drivers may be more likely to be distracted by mobile devices, conversations, or other factors, increasing the risk of accidents.
4. **Morning Commute:** The peak around 9 AM corresponds to the morning rush hour when people are commuting to work or school. Like the afternoon peak, increased activity during this time can contribute to a higher number of accidents.
5. **School Start/End Times:** The morning peak may also be influenced by school end times, as parents, students, and school buses contribute to increased pedestrian and vehicular activity in the vicinity of schools

Conclusion: More accidents are happening because of the confusion caused by right of way rule and driver’s inattention. Highest number of accidents are happening in the timeframe 3 PM – 5 PM.

More number of accidents are happening in the borough Brooklyn and this trend is also observed in all other accident patterns. This is because of the area and population of Brooklyn is far higher than all other boroughs. But one interesting point to note is that when we calculate the population to accident ratio it is highest in Queens Borough.

**Queens Boulevard** noted highest number of accidents, it has some of the most dangerous intersections in New York city causing these accidents not only of pedestrians but also for vehicles too. Also, there are few other dangerous intersections in Queens which are recorded as the hotspots for accidents.

Tricky intersections along with driver negligence and incompetence (several reasons are combined here) are resulting in a greater number of accidents.

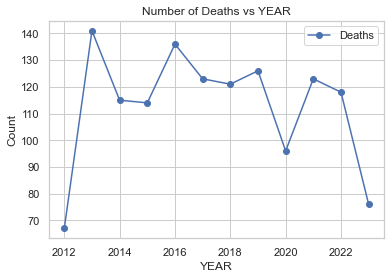
Also, we can see that the percentage increase in population in queens is less when compared to that of other boroughs. Even then the increase in accident percentage suggests that there is a need for special care in case of Queens.

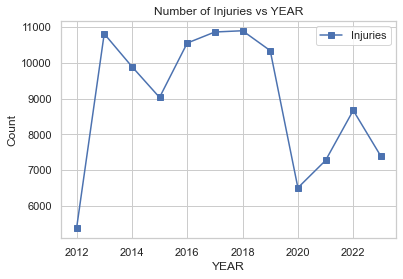
**Measures to be followed:**

More awareness programs must be conducted to explain the drivers how important pedestrian safety is and how even slight inattention while driving could cause a fatal accident.

Based on vision zero program, midblock crossways have been introduced in streets to ease the process of crossing the streets, this should be implemented rapidly.

Careful monitoring and speed restriction signs must be set up in the lanes where schools and hospitals are located.Pedestrian safe crosswalks needs to be the important goal all throughout NYC.   
More number of safety initiatives needs and awareness programs needs to introduced in Queens and Brooklyn borough





When we compare the increase in population with the increase in the fatality and injuries we can say that situation is improving but there needs to be additional care taken in few areas, there should be strict monitoring and penalties imposed in case of driver negligence as this is the most contributing reason for accidents happening. But we can see sudden increase in number of injuries and deaths rising from 2016.

<https://www.nyc.gov/content/visionzero/pages/legislation>

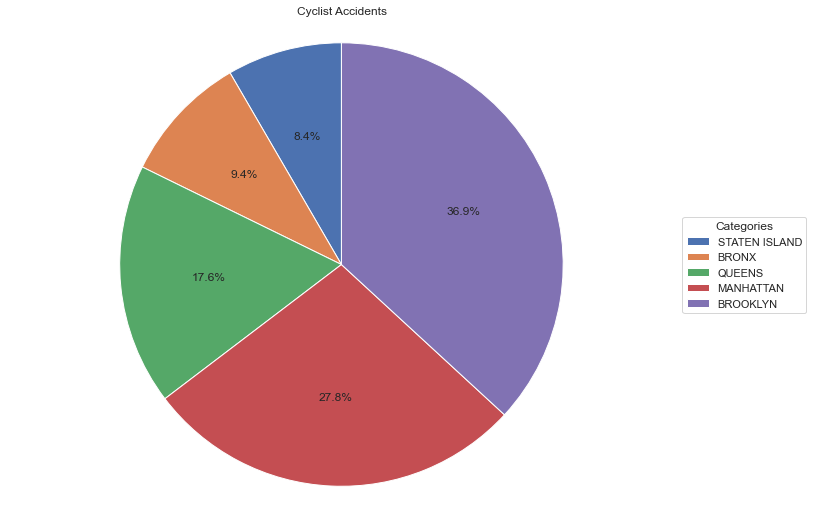
This is because in 2016 NYC Legislation declared that pedestrians have the right of way even when the flashing upraised hand signal or “countdown” is displayed. Eliminating confusion over right of way in this phase emphasizes that traffic must prioritize keeping pedestrians safe.

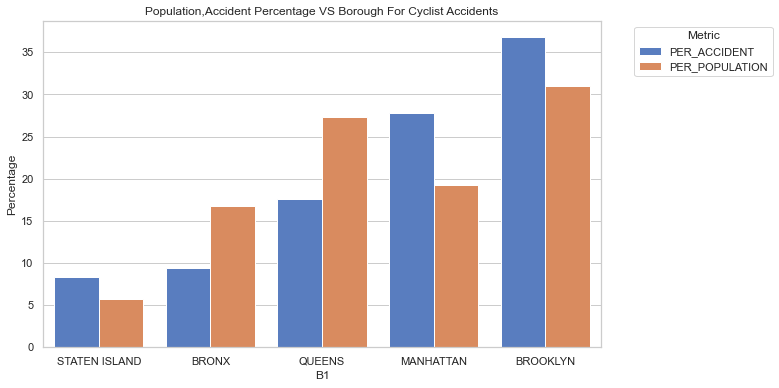
Now this caused a confusion among the pedestrians and motorists on how to cross resulting in increase of accidents.

<https://www.sciencedirect.com/science/article/pii/S221414052200233X#:~:text=A%20study%20by%20Tefft%20(2011,and%2090%25%20at%2058%20mph>.

According to the study the speed of the vehicle plays an important role in fatality of the accident. It states that the likelihood of pedestrian injury severity at high-speed limit zones is about 3.1 times higher for roadways. Considering this the speed limit introduced as a part of Vision Zero provided some advantage to the pedestrians.

**Cyclist Analysis**





Although the trend of total accidents for motorists seems like that of cyclists one important point to note here is we have normalised the percentage obtained with the percentage of population. Then we can see that highest disparity is found in the case of Staten Island and Manhattan.

Manhattan is one the dangerous boroughs for motorists in NYC. The highest motor vehicle dense area of NYC is Manhattan. Especially with rise in Uber and other cab services a greater number of vehicles are travelling across the city from Manhattan. On analysing the contributing factor, we found out that driver attention out-weighed all other reasons. Along with this we found out few other factors which are:

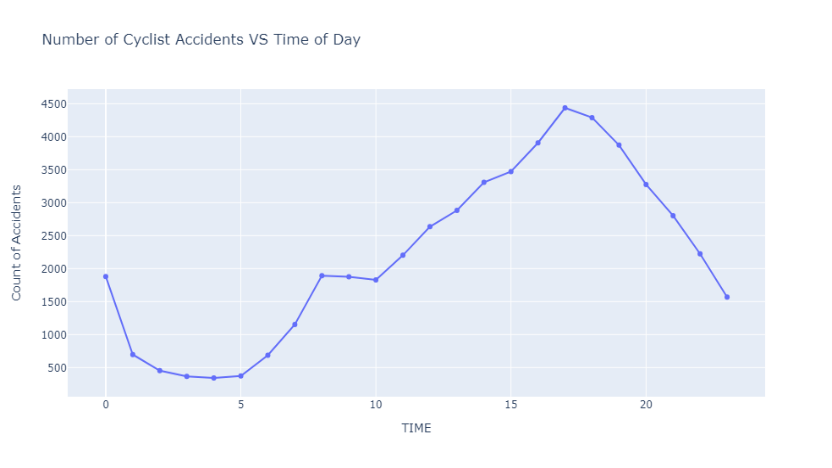
1. **Busy Traffic Structure:** Manhattan is known for its busy and congested streets, particularly in commercial and business districts. The increased traffic volume, combined with the presence of cyclists, can contribute to a higher risk of accidents.
2. **Complex Street Grid:** Manhattan's street grid is known for its complexity, with numerous intersections, one-way streets, and diverse traffic patterns. Cyclists navigating through this intricate network may face challenges, and the complexity can contribute to the risk of accidents.
3. **Tourist Presence:** Manhattan attracts a significant number of tourists, especially in popular areas such as Times Square and Central Park. Tourists may not be familiar with local traffic rules, leading to potential conflicts with cyclists.
4. **Limited Cycling Infrastructure:** While efforts have been made to improve cycling infrastructure in Manhattan, such as the addition of bike lanes, the existing infrastructure may still face challenges in accommodating the high demand for cycling. Limited or poorly designed bike lanes can contribute to accidents.
5. **Delivery and Service Vehicles:** Manhattan has a high volume of delivery and service vehicles. Interactions between cyclists and these vehicles, especially in loading and unloading zones, can contribute to accidents.
6. **Intersection Challenges:** Intersections are often high-risk areas for accidents. Manhattan's numerous intersections, along with turning vehicles and pedestrians, can create challenging conditions for cyclists

**Pedestrian countdown signals: Experience with an extensive pilot installation**

We can say that because of installation of countdown timers for pedestrian signals, the number of car crashes are increasing.

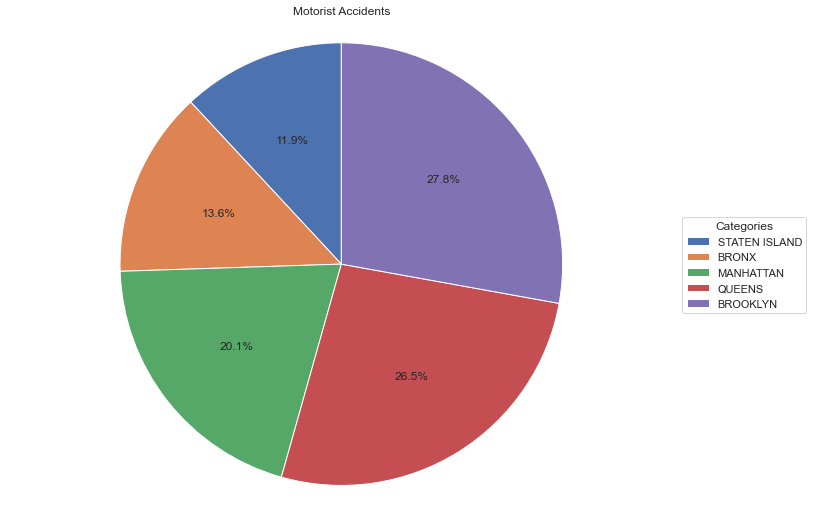
Crucially, accidents caused by drivers watching the countdown timer tended to be far more serious and were more likely to be fatal. That is because they were most often t-bone collisions, which tend to hit the driver or their passengers directly, without the padding of the crumple zone to cushion the impact.

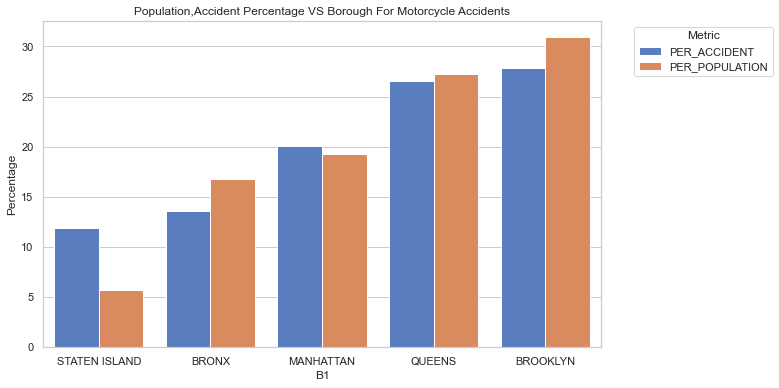
Another reason for increase is the role of countdown clocks introduced in 2010, this is because when there is a countdown, now there is pressure on driver either to start immediately after the timer goes out or to cross the road before the timer runs out



The trend that we observed wrt time of accidents is almost same as that of the pedestrians. Almost all the reasons mentioned for pedestrians can be applied here.

**Motorist Accident Analysis**





Unlike other boroughs we see highest difference between the population percentage and accident percentage in Staten Island.

The

.