

Car Accidents Severity

Introduction:

Urbanization is the major reason for increasing in number of vehicles on road. In this timely world everyone is in their own thought and hurry which are leading to road accidents. There are many factors leading to accidents. It depends on weather, road condition, vehicle condition, driver condition and many other factors. The main aim of our project is to reduce the car accidents.

Here we deal with the dataset which has been recorded by Seattle traffic management division since 2004 in Seattle. This data includes the types of collisions occurred, and all the necessary data.

We make a report on the factors leading to accidents and what can be done to reduce these accidents.

Dataset:

Seattle traffic management division has recorded the data of accidents which occurred since 2004, which includes many factors. We consider the following factors which determine the major accidents type.

Weather, road condition, vehicle condition, driver condition, Light condition, pedestrians, Junction are important to determine the accident in the dataset.

Firstly, to analyze the data we have imported data in CSV format into IBM Watson Studio.

```

In [10]: import itertools
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
from matplotlib.ticker import NullFormatter
import matplotlib.ticker as ticker
from sklearn import preprocessing
import matplotlib inline

In [12]: import types
import pandas as pd
from botocore.client import Config
import boto3

def __iter__(self): return 0

# @hidden_cell
# The following code accesses a file in your IBM Cloud Object Storage. It includes your credentials.
# You might want to remove those credentials before you share the notebook.
client_86d3d75d55f4ebda1fa8bcb298586 = boto3.client('s3',
    aws_access_key_id='2_82ghjXGNEHDBR85JUS7DvB8K4cduyQ4X_k02a86',
    aws_secret_access_key='...',
    endpoint_url='https://iam.cloud.ibm.com/oidc/token',
    config=Config(signature_version='auth'),
    endpoint_url='https://ss.eu-gb-objectstorage.service.networklayer.com')

body = client_86d3d75d55f4ebda1fa8bcb298586.get_object(Bucket='caraccidentseverity-donotdelete-pr-yukvsyuytazme',Key='data-collisions.csv')['Body']
# add missing __iter__ method, so pandas accepts body as file-like object
if not hasattr(body, "__iter__"): body.__iter__ = types.MethodType(__iter__, body)

df_vech = pd.read_csv(body)
df_vech.head()

```

SEVERITYCODE	X	Y	OBJECTID	INCKEY	COLDKEY	REPORTNO	STATUS	ADORTYPE	INTKEY	ROADCOND	LIGHTCOND	PEDEWNOTORNT	SDOTCOLNUM	SPEEDING	ST_COLCODE	ST_COLDESC	SEGLANEKEY	CROSSWALKKEY	HITPARKEDCAR		
0	2	-122.323148	47.703140	1	1307	1307	3502005	Matched	Intersection	37475.0	...	Wet	Daylight	NaN	NaN	10	Entering at angle	0	0	N	
1	1	-122.347204	47.647172	2	82200	82200	2607959	Matched	Block	NaN	...	Wet	Dark - Street Lights On	NaN	8354039.0	NaN	11	From same direction - both going straight - so...	0	0	N
2	1	-122.334540	47.607871	3	26700	26700	1462393	Matched	Block	NaN	...	Dry	Daylight	NaN	4223031.0	NaN	32	One parked-one moving	0	0	N
3	1	-122.334803	47.604803	4	1144	1144	3503937	Matched	Block	NaN	...	Dry	Daylight	NaN	NaN	23	From same direction - all others	0	0	N	
4	2	-122.306426	47.545739	5	17700	17700	1807429	Matched	Intersection	34387.0	...	Wet	Daylight	NaN	4028032.0	NaN	10	Entering at angle	0	0	N

5 rows x 38 columns

I. Methodology:

A. Data Processing:

- Firstly, I have checked the data types of all columns
- Then I've created a data frame of required columns.
- Converted the nan values to either 0 or 1.
- The time format is changed to default standard.

B. Exploratory Analysis:

- In total number of accidents in Seattle most of them are property damaged only collision are more than in injured.

```

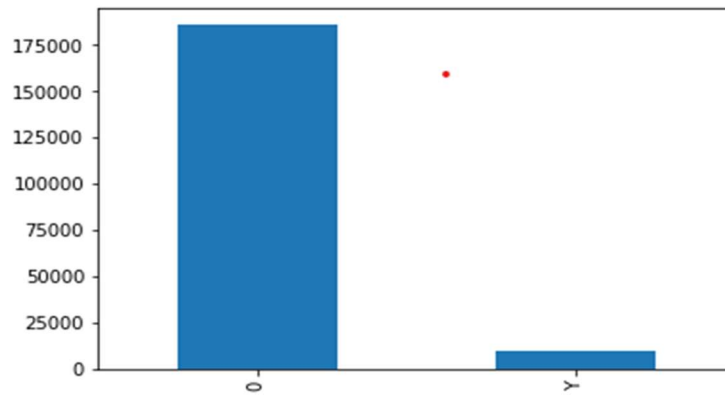
In [6]: df_vech['SEVERITYDESC'].value_counts()

Out[6]: Property Damage Only Collision    136485
        Injury Collision                  58188
        Name: SEVERITYDESC, dtype: int64

```

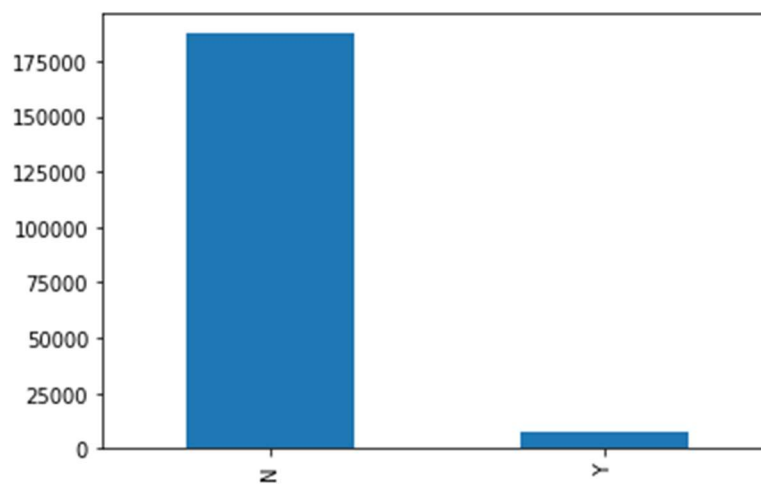
- Very of the accidents occurred due to Over Speeding.

```
In [8]: df_vech['SPEEDING'].replace(np.nan, 0, inplace=True)
df_vech['SPEEDING'].value_counts().plot(kind='bar')
Out[8]: <matplotlib.axes._subplots.AxesSubplot at 0x7f1aab533278>
```



- Car hitting a car at rest are very less.

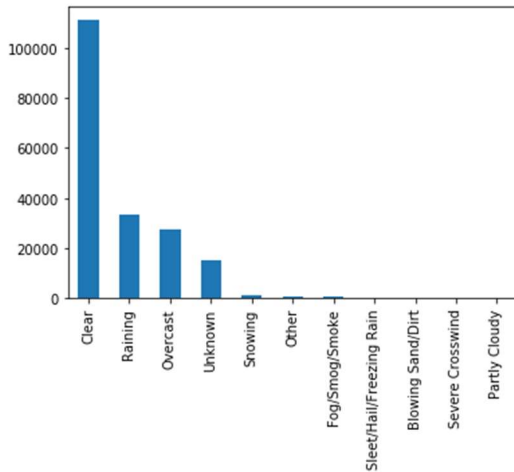
```
In [33]: df_vech['HITPARKEDCAR'].value_counts().plot(kind='bar')
Out[33]: <matplotlib.axes._subplots.AxesSubplot at 0x7f1aaa256e48>
```



- As we can see below most accidents occurred during a clear weather at daylight on a dry road.

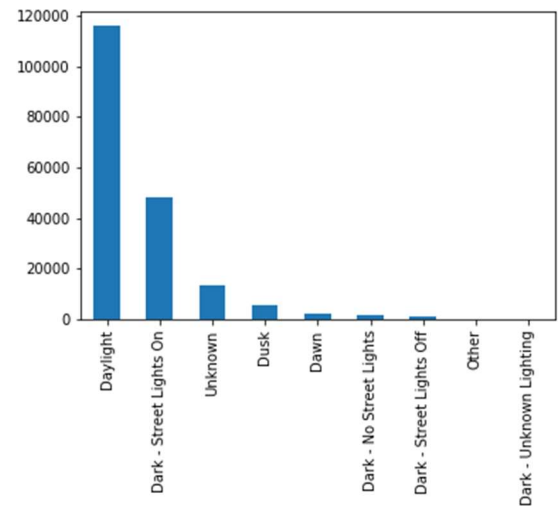
```
In [11]: df['WEATHER'].value_counts().plot(kind='bar')
```

```
Out[11]: <matplotlib.axes._subplots.AxesSubplot at 0x7f1aab490518>
```



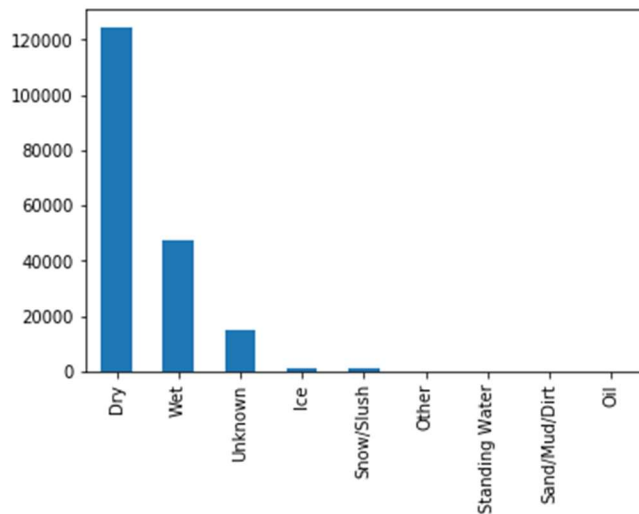
```
In [13]: df['LIGHTCOND'].value_counts().plot(kind='bar')
```

```
Out[13]: <matplotlib.axes._subplots.AxesSubplot at 0x7f1aab373ac8>
```



```
In [12]: df['ROADCOND'].value_counts().plot(kind='bar')
```

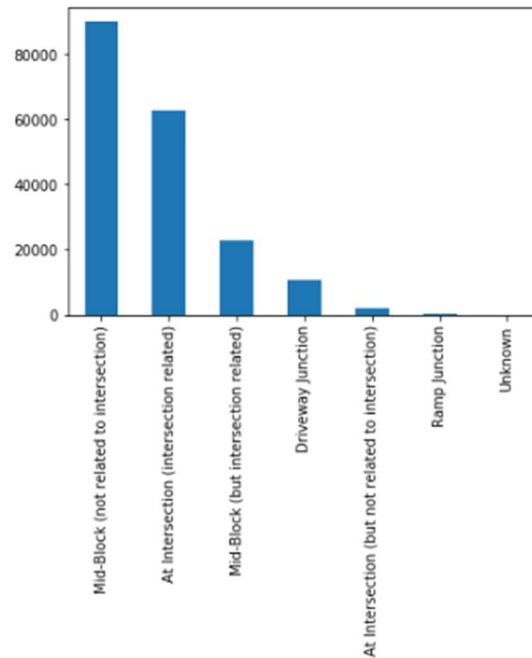
```
Out[12]: <matplotlib.axes._subplots.AxesSubplot at 0x7f1aab506ac8>
```



- Here below accidents did not occur at intersection.

```
In [14]: df['JUNCTIONTYPE'].value_counts().plot(kind='bar')
```

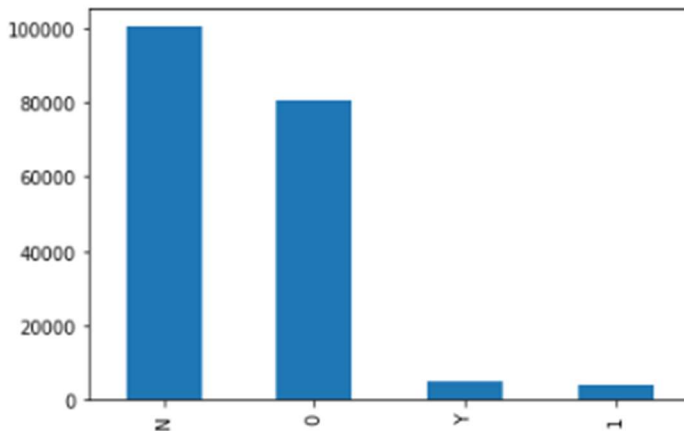
```
Out[14]: <matplotlib.axes._subplots.AxesSubplot at 0x7f1aab367438>
```



- Accidents due drivers under influence.

```
In [39]: df_vech['UNDERINFL'].value_counts().plot(kind='bar')
```

```
Out[39]: <matplotlib.axes._subplots.AxesSubplot at 0x7f1aa9eda390>
```



II. Results:

- As we can observe that most of the accidents occurred during daytime, on dry road surface, not even at intersection and in a clear weather.
- Most of them are not by over-speeding.
- The collisions which occurred under influence are also very less.

III. Discussions:

- As seen most of the accidents occurred on a dry road during daytime in clear weather. As these did occur under any influence. We can say that there might be some distractions on road as most of them did not occur at intersection.
- Those how got distracted ran into accident and other might not.
- Next to that, many occurred during raining or on a wet road, these can be expected as many accidents occur during these situations. For this people have to use proper tires with good grip.

IV. Conclusion:

- People have to be much cautious and alert during driving.
- People have to frequently check their tires as they get wear due to usage.