CAPSTONE PROJECT ON CAR ACCIDENT SEVERITY

BY
GVG

CONTENTS

- Background
- Data Understanding
- Methodology
- Results and Discussion
- Conclusion

Background

- Vehicles may be car, trucks, etc., .
- accidents may happen at different latitudes and longitudes.we have to find those based on city or province.
- Severity code and its description specifies the situation of the victim.
- victim vehicle and how many victims affected gives us how many are under that severity.

Data Understanding

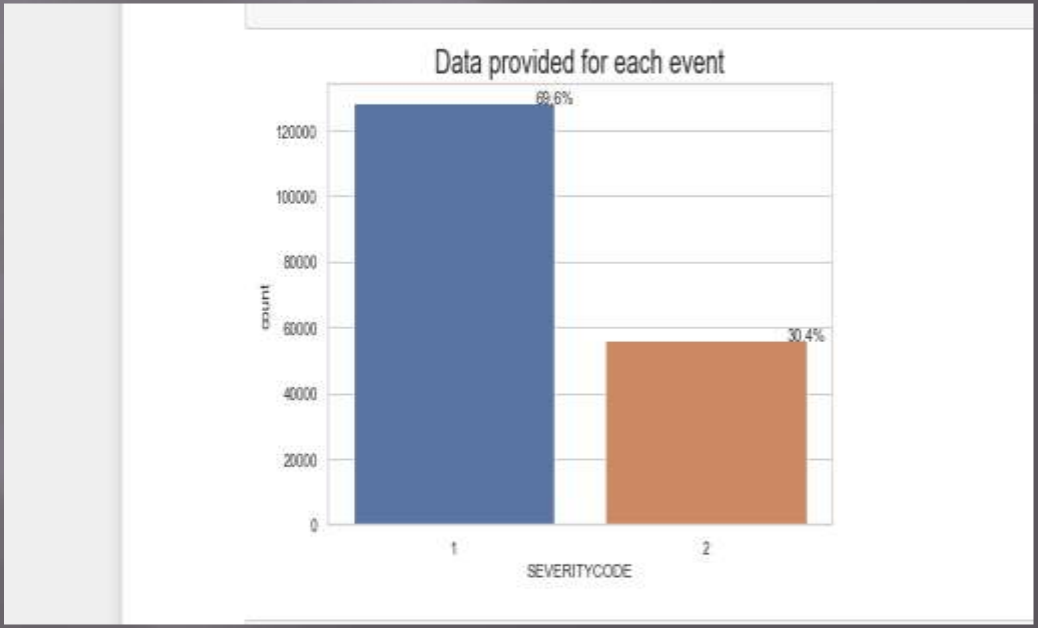
- Load file
- Display categorical features
- Inspect the features like Accident_Severity, Speed_Limit,etc.,
- Most accidents occur at low speed limits but they are more severe in proportion at higher speed limits (makes sense)
- Weather, road surface and light conditions
- Most accidents occur in normal weather, only Fog or High winds increase the proportion of severe accidents
- Junctions types

Methodology

- Initially, I thought that Our data is now ready to be fed into machine learning models.
- No algorithm is suitable to do predictions as described below.
- First we imported the data through read_csv(). I noticed that it had 194,674 rows and 38 columns. Therefore, we narrowed it down to 9 columns (‘Severity’, ‘X’, ‘Y’, ‘Location’, 'Vehcount’, ‘Weather’, ‘Roadcond’, ‘Lighdcond’ and ‘Hitparkedcar’) and delete the missing values, which made the final dataset with 184,167 observations and 9 variables.

Methodology

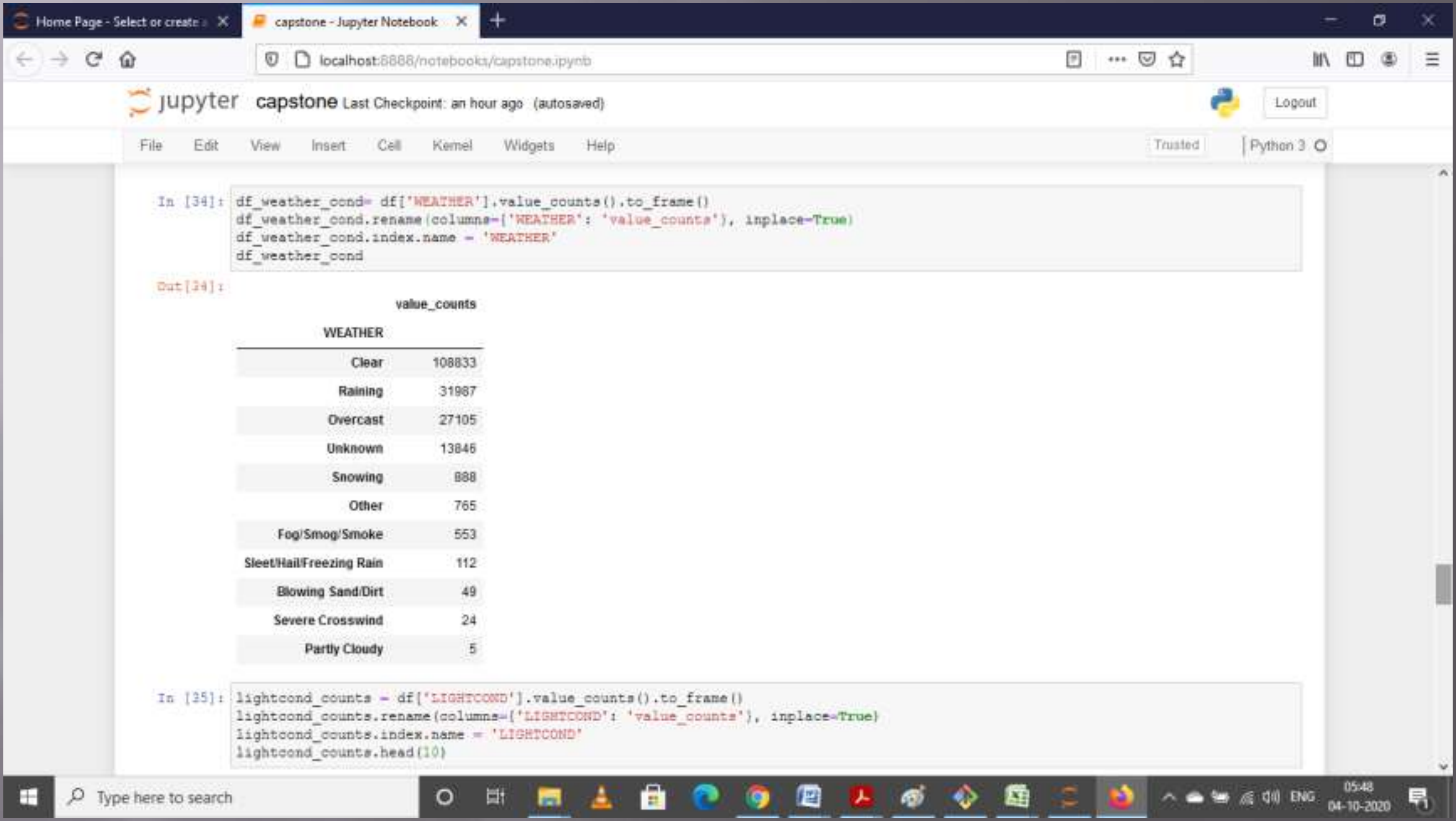
- Since most of the variable were categorical, it was hard to make the regression model. So, in this study, we focused more on the graphical data and the value count for different categories. There were around 69.6% (2/3) level 1 accidents and 30.4% (1/3) level 2 accidents.



Results and Discussion

- We generated the graphical information based on Seaborn library. The result showed that some locations did have more car accidents than the other places.
- After that, we checked about the weather, road, and light condition. We calculated the total number of car accidents under different situations. There was no significant evidence showed that they might be the reason for the accidents.

Results and Discussion



Home Page - Select or create a
capstone - Jupyter Notebook
localhost:8888/notebooks/capstone.ipynb
jupyter capstone Last Checkpoint: an hour ago (autosaved)
Logout
File Edit View Insert Cell Kernel Widgets Help
Trusted
Python 3
Type here to search
ENG
05:48
04-10-2020

```
In [34]: df_weather_cond= df['WEATHER'].value_counts().to_frame()
         df_weather_cond.rename(columns={'WEATHER': 'value_counts'}, inplace=True)
         df_weather_cond.index.name = 'WEATHER'
         df_weather_cond
```

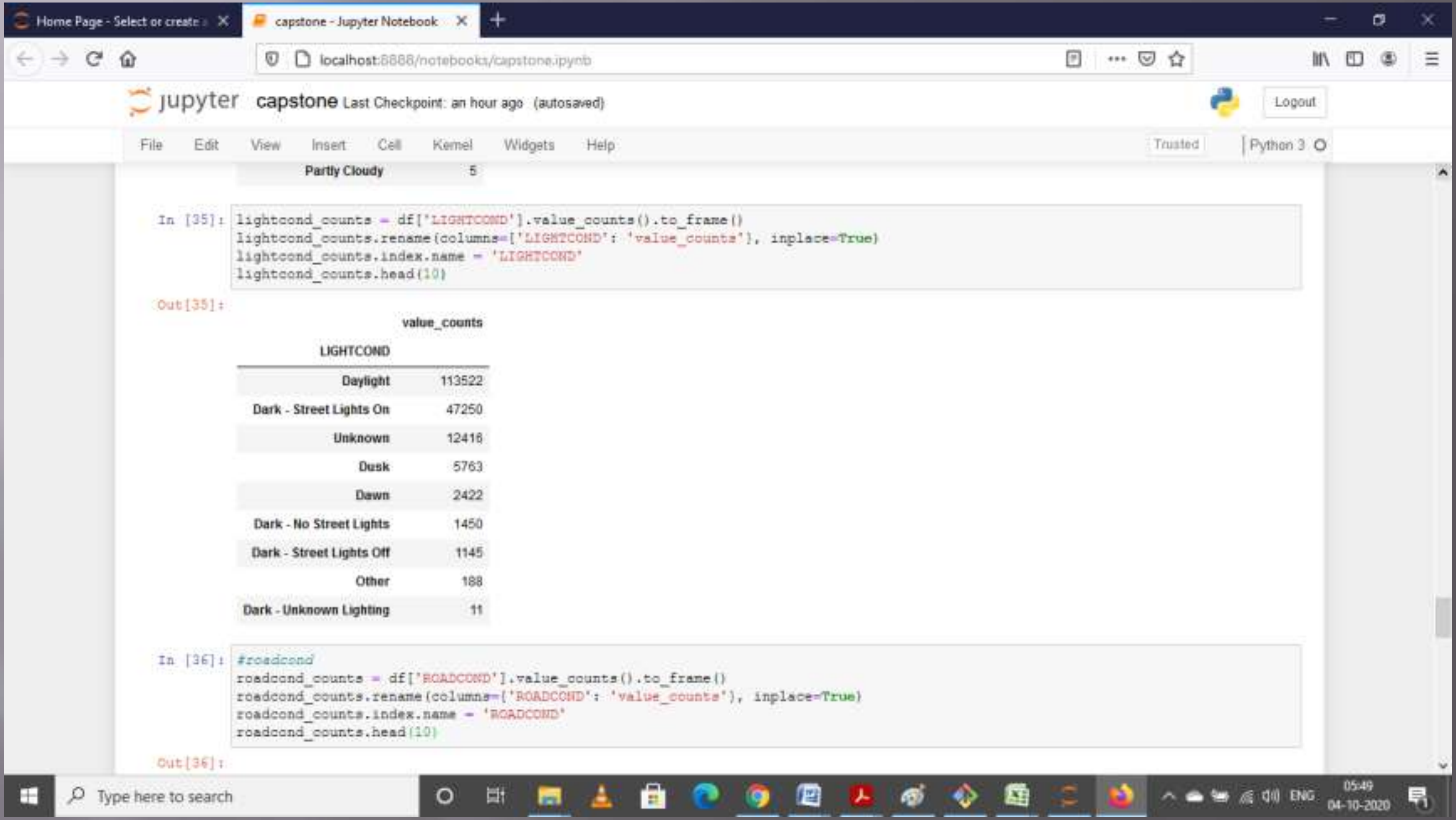
Out[34]:

WEATHER	value_counts
Clear	108833
Raining	31987
Overcast	27105
Unknown	13846
Snowing	888
Other	765
Fog/Smog/Smoke	553
Sleet/Hail/Freezing Rain	112
Blowing Sand/Dirt	49
Severe Crosswind	24
Partly Cloudy	5

```
In [35]: lightcond_counts = df['LIGHTCOND'].value_counts().to_frame()
         lightcond_counts.rename(columns={'LIGHTCOND': 'value_counts'}, inplace=True)
         lightcond_counts.index.name = 'LIGHTCOND'
         lightcond_counts.head(10)
```

Different Weather Conditions

Results and Discussion



Home Page - Select or create a
capstone - Jupyter Notebook
localhost:8888/notebooks/capstone.ipynb
jupyter
capstone Last Checkpoint: an hour ago (autosaved)
Logout
File
Edit
View
Insert
Cell
Kernel
Widgets
Help
Trusted
Python 3
Type here to search
ENG
05:49
04-10-2020

Partly Cloudy	5

```
In [35]: lightcond_counts = df['LIGHTCOND'].value_counts().to_frame()
         lightcond_counts.rename(columns={'LIGHTCOND': 'value_counts'}, inplace=True)
         lightcond_counts.index.name = 'LIGHTCOND'
         lightcond_counts.head(10)
```

Out[35]:

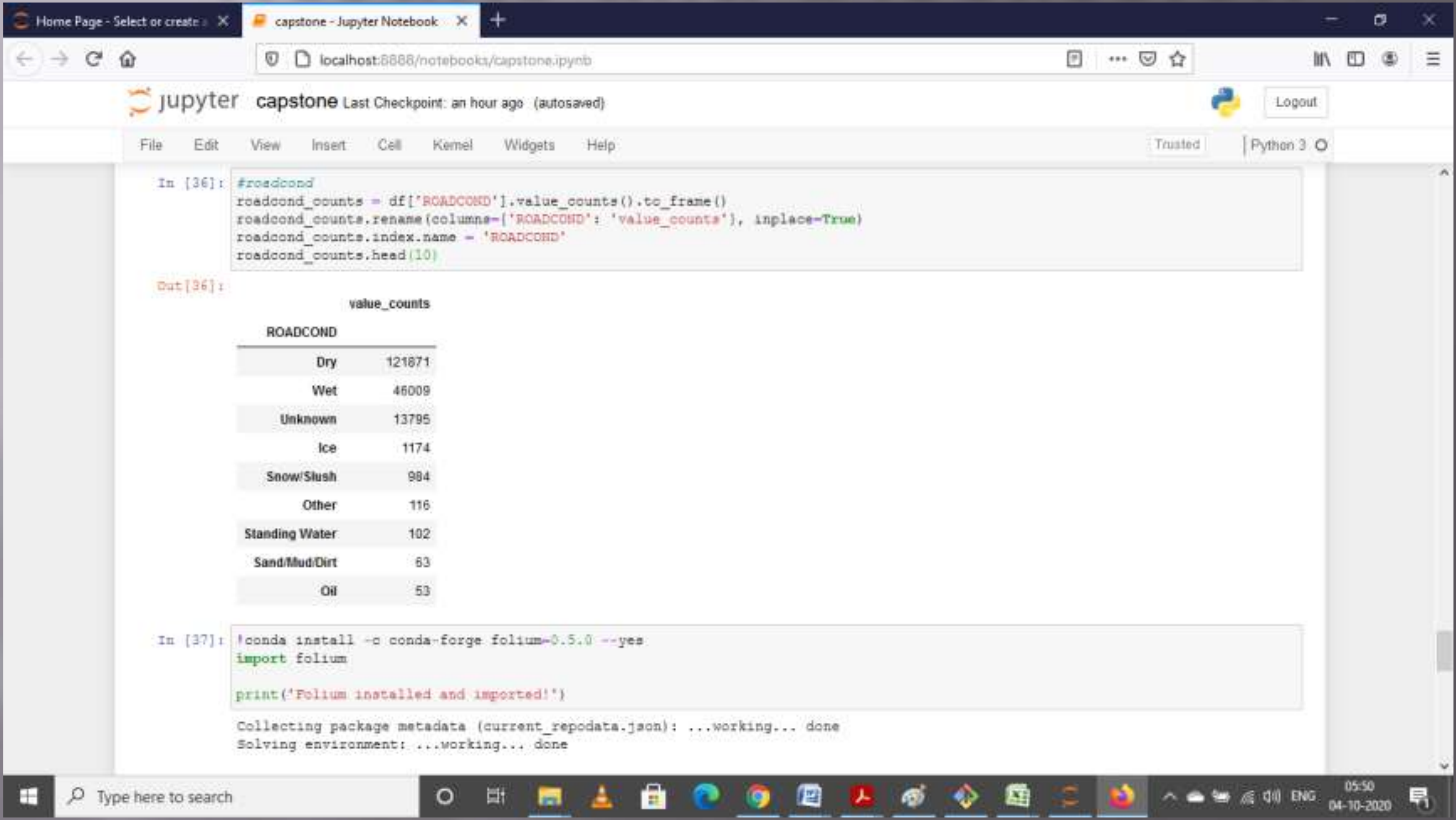
	value_counts
LIGHTCOND	
Daylight	113522
Dark - Street Lights On	47250
Unknown	12416
Dusk	5763
Dawn	2422
Dark - No Street Lights	1450
Dark - Street Lights Off	1145
Other	188
Dark - Unknown Lighting	11

```
In [36]: #roadcond
         roadcond_counts = df['ROADCOND'].value_counts().to_frame()
         roadcond_counts.rename(columns={'ROADCOND': 'value_counts'}, inplace=True)
         roadcond_counts.index.name = 'ROADCOND'
         roadcond_counts.head(10)
```

Out[36]:

Different lighting conditions

Results and Discussion


Home Page - Select or create a
capstone - Jupyter Notebook
localhost:8888/notebooks/capstone.ipynb
jupyter capstone Last Checkpoint: an hour ago (autosaved)
Logout
File Edit View Insert Cell Kernel Widgets Help
Trusted
Python 3
Type here to search
ENG
05:50
04-10-2020

In [36]:

```
#roadcond
roadcond_counts = df['ROADCOND'].value_counts().to_frame()
roadcond_counts.rename(columns={'ROADCOND': 'value_counts'}, inplace=True)
roadcond_counts.index.name = 'ROADCOND'
roadcond_counts.head(10)
```

Out[36]:

ROADCOND	value_counts
Dry	121871
Wet	46009
Unknown	13795
Ice	1174
Snow/Slush	984
Other	116
Standing Water	102
Sand/Mud/Dirt	63
Oil	53

In [37]:

```
!conda install -c conda-forge folium=0.5.0 --yes
import folium

print('Folium installed and imported!')
```

```
Collecting package metadata (current_repodata.json): ...working... done
Solving environment: ...working... done
```

Different road conditions

Conclusion

- Besides, there are a few spots which has more mishaps during the dull time(more information might be given). For those spots, including lights may be a decent answer for decrease the crashes. Additionally, when more vehicles engaged with the mishap, it appears to be that the degree of seriousness will increment. They may should be reacted quickly to spare life and wards.