

**MINI-PROJECT REPORT**

**ROAD ACCIDENT ANALYSIS**

*Submitted to the partial fulfillment of the requirement*

*for the* ***18CSE355T-Data Mining and Analytics*** *course and*

*for the award of the degree of*

**BACHELOR OF TECHNOLOGY**

**IN**

**COMPUTER SCIENE AND ENGINEERING**

***Submitted by-***

**Mr. MOKSH VIJ**

**[Registration No.RA1811003030124]**

**Mr. NIMISH DEEP**

**[Registration No.RA1811003030129]**

**Group ID: 2020DDDD**

*Under the guidance of*

**Mr. PARTHA SARATHI CHAKRABORTY**

*(Assistant Professor, Department of Computer Science and Engineering)*

**DEPARTMENT OF COMPUTER SCIENE AND ENGINEERING**

**SRM INSTITUTE OF SCIENCE AND TECHNOLOGY**

**(Deemed to be University under Section 3 of the UGC Act, 1956)**



NOVEMBER 2020

**BONAFIDE CERTIFICATE**

Certified that this project report titled   
“ROAD ACCIDENT ANALYSIS" is the bonafide work of   
Mr. MOKSH VIJ [Registration No. RA1811003030124] and Mr. NIMISH DEEP [Registration No. RA1811003030129], who carried out the project work under my supervision and submitted to the partial fulfillment of the requirement for the 18CSE355T Data Mining and Analytics course and for the award of the degree of Bachelor of Technology in Computer Science and Engineering of SRM Institute of Science and Technology.

|  |  |  |
| --- | --- | --- |
|  |  | ***Mr. Partha Sarathi Chakraborty***  *Supervisor,*  *Assistant Professor, Department of*  *Computer Science and Engineering* |

**ABSTARCT**

At present, India is in a period of steady development of highways. At the same time, traffic safety issues are becoming increasingly serious. There are many inventories in automobile industries to design and build safety measures for automobiles, but traffic accidents are unavoidable. There is a huge number of accidents prevailing in all urban and rural areas. Patterns involved with different circumstances can be detected by developing an accurate prediction models which will be capable of automatic separation of various accidental scenarios. Data mining technology is an effective method for analyzing traffic accidents. In-depth information mining of traffic accident data is conducive to accident prevention and traffic safety management. This cluster will be useful to prevent accidents and develop safety measures. We believe to acquire maximum possibilities of accident reduction using low budget resources by using some scientific measures. Based on the data of traffic accidents from 2001 to 2014, this study selected factors including time at which the accident occurred, the state in which the accident occurred and the increase in the number of the accidents in different states in different years.

**TABLE OF CONTENTS**

|  |  |  |
| --- | --- | --- |
| **S. No** | **Topic** | **Page No** |
|  | Introduction | 5 |
|  | Literature Survey | 6 |
|  | Dataset Description | 7 |
|  | Data Preprocessing | 9 |
|  | Data Analysis | 12 |
|  | Conclusion | 18 |
|  | References | 19 |

**INTRODUCTION**

There is a huge impact on the society due to traffic accidents where there is a great cost of fatalities and injuries. In recent years, there is increase in the researches attention to determine the significantly affect the severity of the driver’s injuries which is caused due to the road accidents. Accurate and comprehensive accident records are the basis of accident analysis. The effective use of accident records depends on some factors, like the accuracy of the data, record retention, and data analysis. There are many approaches applied to this scenario to study this problem. Road accidents constitute a major problem in our societies around the world. The World Health Organization (WHO) estimated that 1.25 million deaths were related to road traffic injuries in the year 2010.

A recent study illustrated that the residential and shopping sites are more hazardous than village areas. As might have been predicted, the frequencies of the casualties were higher near the zones of residence possibly because of the higher exposure. A study revealed that the casualty rates among the residential areas are classified as relatively deprived and significantly higher than those from relatively affluent areas. Patterns involved in serious accidents can be able to classify various traffic accidents with developing accurate prediction models. Identifying the primary road traffic accident factors will help to provide an appropriate solution to minimize the adverse effect of severity on human and property loss. Data mining algorithms can be used to help in preventing these accidents by providing a systematic analysis of the accidents around the country which will eventually help in creating some major measures to prevent these accidents.

**LITERATURE SURVEY**

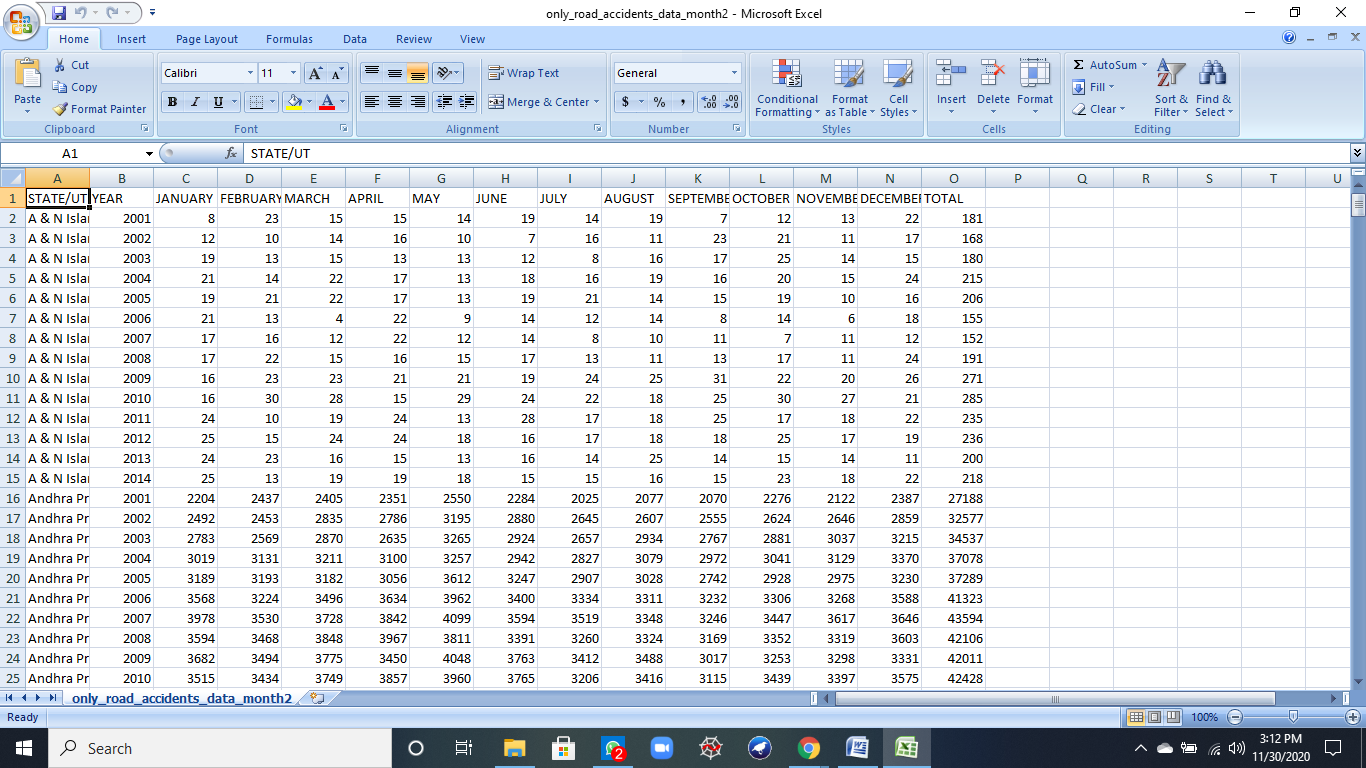
Data Mining is the process of discovering interesting patterns and knowledge from a large database and data classification is a form of data analysis that involves extracting models with important data classes.

This literature survey provides a comprehensive review about the data mining techniques applied to achieve this analysis report.

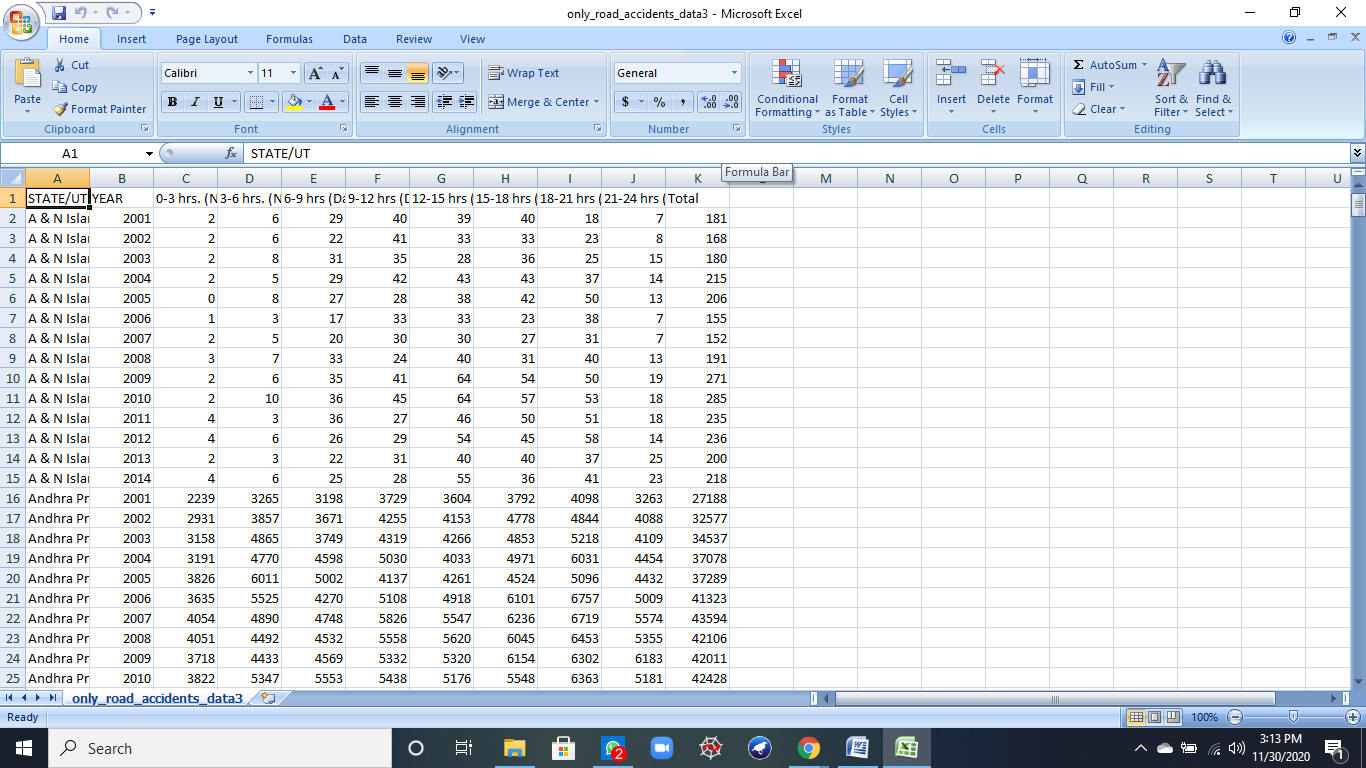
The given data is used to do a detailed analysis of road accidents across India using clustering which provide us with some unique patterns. These patterns help in classifying accidents across the country on the factors of accident-prone time, weather and place.

Identifying the primary road traffic accident factors will help us to reach an appropriate solution to minimize the adverse effects of accidents.

**DATASET DESCRIPTION**

* The dataset used here contains a detailed information of the accidents happening all across the country based on the factors of time, month and seasons.
* The first table of the dataset contains the accident count of different states of India between the period of 2001-2014 during the different months in a year. The attributes are-
  1. State/UT - This column contains all the states and union territories of India.
  2. Year - This column contains the year count (2001-14).
  3. Column C to N contains the count of accident each month in different states.
  4. Total - This column contains the total count of accidents each year in each state.
* The attributes of this table is also used to group months according to seasons and give a accident pattern on the basis of seasons i.e. Summer, Winter, Autumn, Spring.
* Note - All attributes are being used of the table.

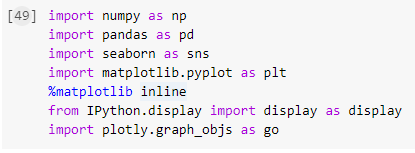
*Table 1*

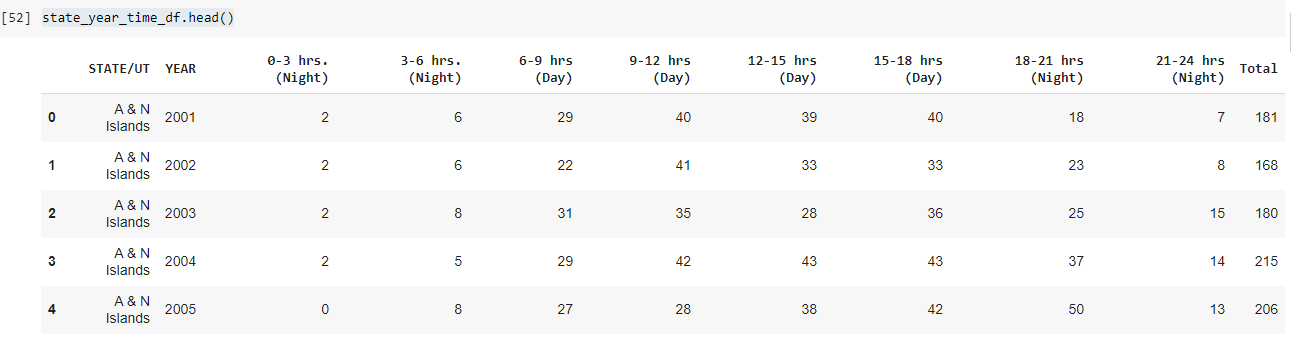
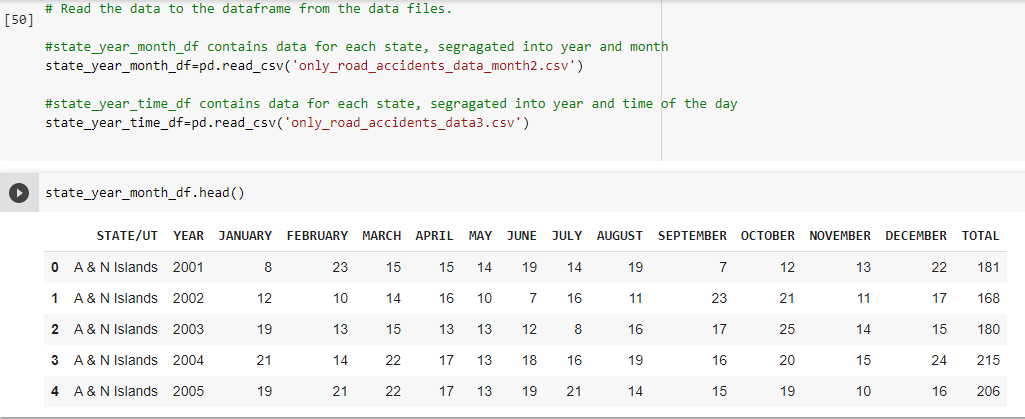
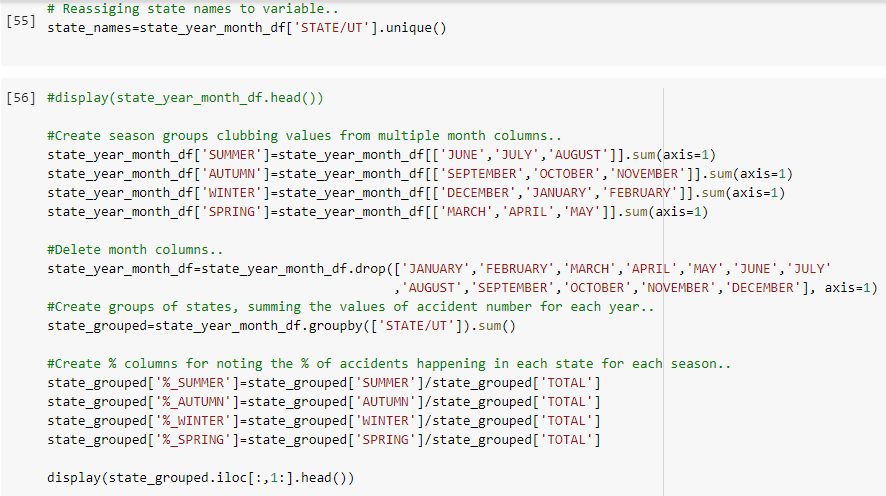
* The second table of the dataset contains the accident count of different states of India between the period of 2001-2014 during the different hours of the day in a period of 3 hours. The attributes are-
  1. State/UT - This column contains all the states and union territories of India.
  2. Year - This column contains the year count (2001-14).
  3. Column C to N contains the count of accident in every 3 hours during a single day in different states.
  4. Total - This column contains the total count of accidents each year in each state.
*  Note - All attributes are being used of the table.

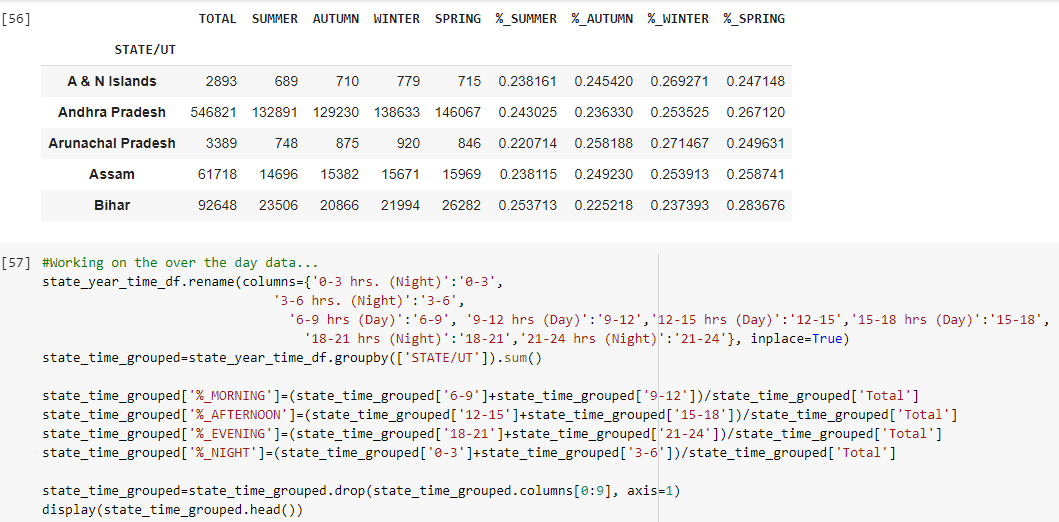
*Table 2*

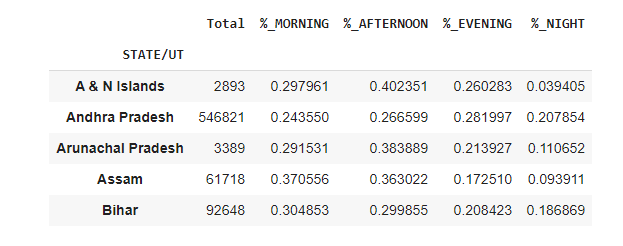
**DATASET PREPROCESSING**

1. **Importing Libraries:**



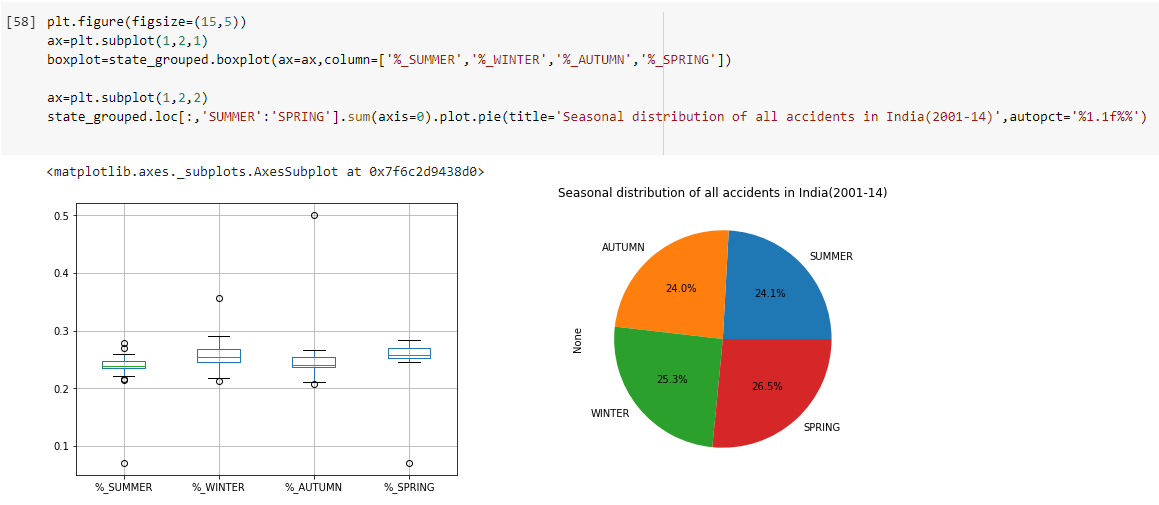
1. **Reading Files:**
2. **Grouping of Months into Seasons and Timeslots in a More Definite Way:**



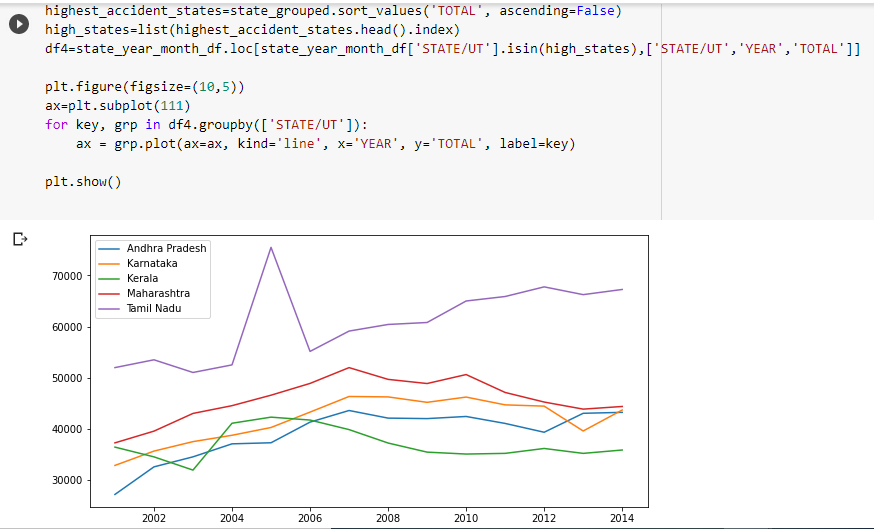
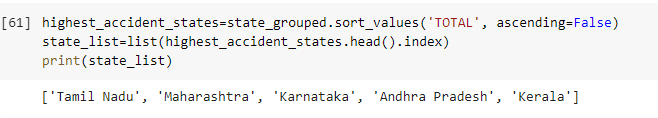


**DATA ANALYSIS**

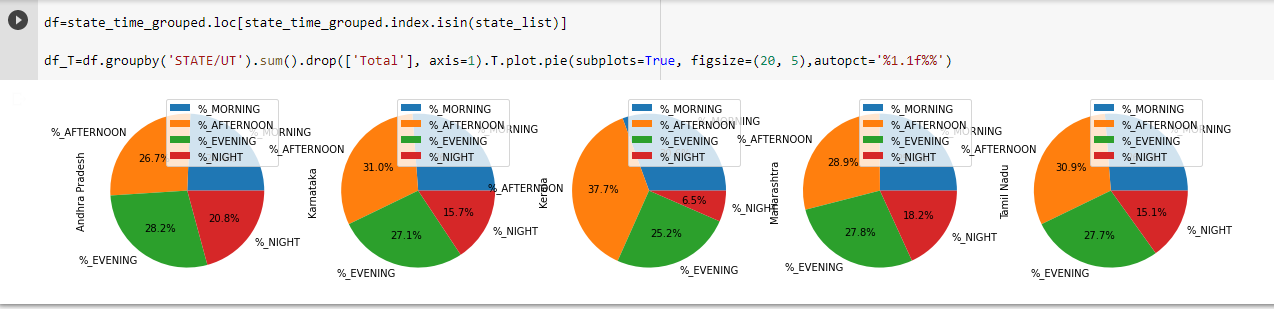
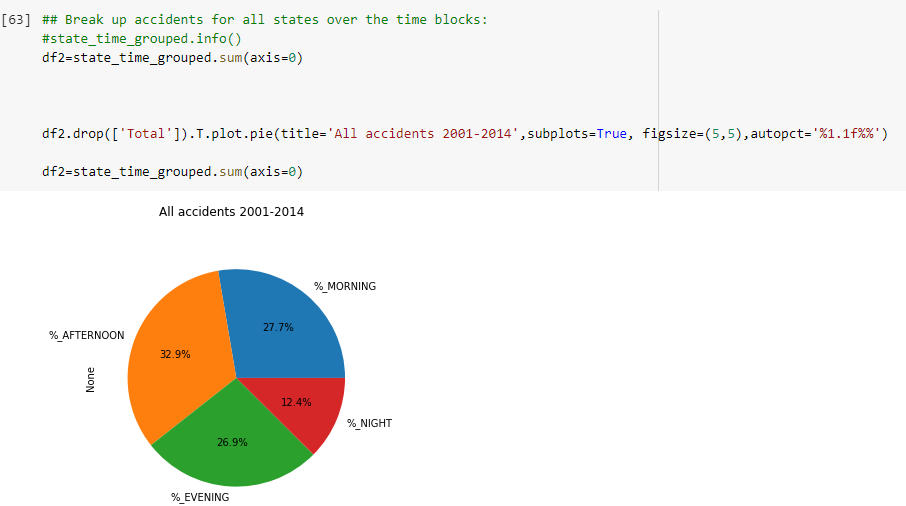
1. **Seasonal Distribution of Accident Count:**

* The pie chart and boxplot given below contains the data of the accidents occurring in the different seasons in a year across the country which will help us in predicting the most accident-prone season.

1. **States\UT Wise Distribution of Accident Count:**

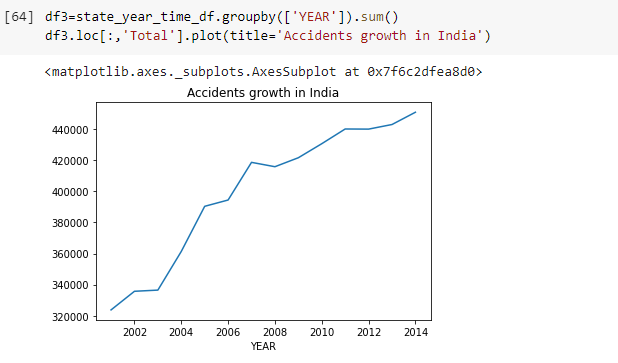
* The graph given below contains the data of the total number of accidents occurring in the different states\UT in a year across the country which will help us in predicting the most accident-prone state\UT.

1. **Time Wise Distribution of Accident Count:**

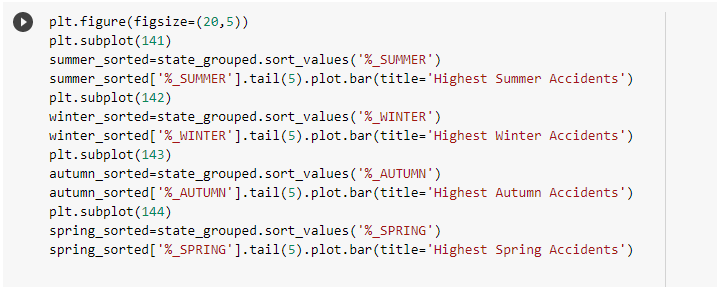
* The pie charts given below contains the data of the total number of accidents occurring in the different states\UT during different time period of a day which will help us in predicting the most accident-prone time period of the day in different states\UT.
* The pie chart given below contains the data of the total number of accidents occurring during different time period of a day in the country which will help us in predicting the most accident-prone time period of the day in the country.

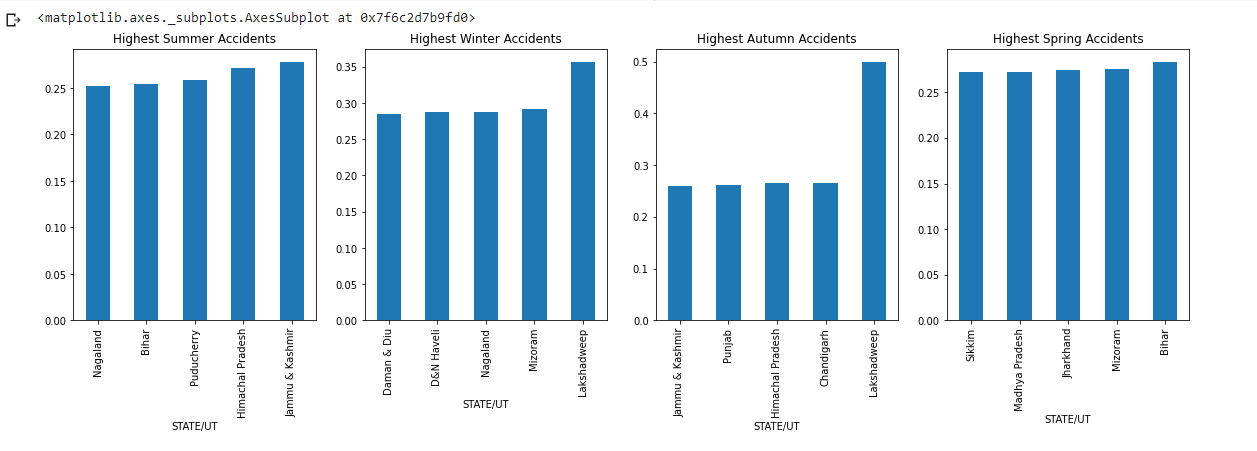
1. **Accident Growth in India:**

* The graph given below shows the increase in the accidents across the country by each passing year.



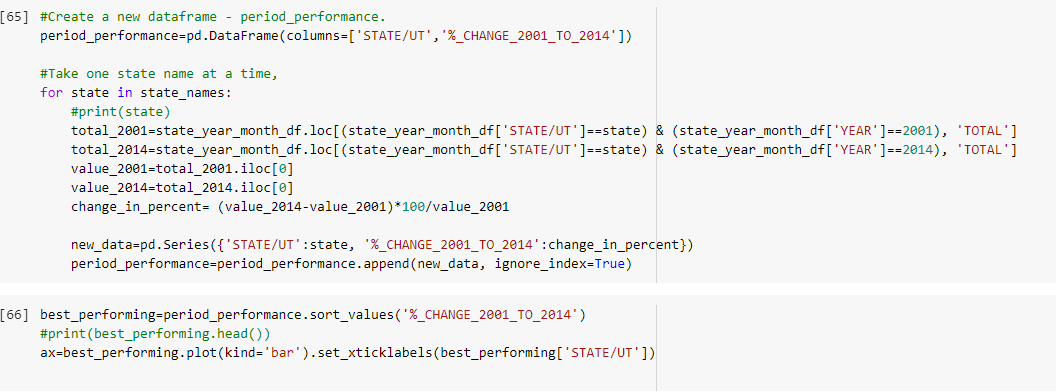
1. **States Wise Distribution of Accident Count in Different Seasons:**

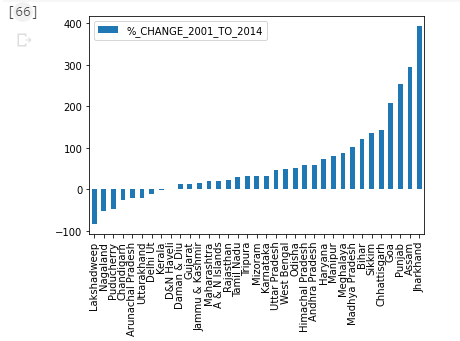
* The bar graphs given below contains the data of the total number of accidents occurring in the different states\UT during different seasons of a year across the country which will help us in predicting the most accident-prone seasons in different states\UT of India.



1. **Checking Performance of States\UT from 2001 to 2014:**

* The bar graphs given below tracks the performance of each state\UT of India during the period of 2001-2014. This will help us in tracing the state with highest increase in accident count between different states\UT of India.





**CONCLUSION**

In this project we did a detailed analysis on the road accidents across the country using python’s machine learning algorithms on the basis of various different attributes. The analysis helped us in predicting accident-prone states\UT, time period in a day, seasons in a year and months across the country and in individual states. The project also comments on the overall growth of the accident count in India during the time period of 2001-2014 and helps us in tracking the performance of individual states over the same time period. All the above gained information will help the central government as well as state government in preparing required action plans to tackle this problem by studying the unique patterns.

**REFERENCES**

* Dataset - <https://www.kaggle.com/manugupta/road-accidents-in-india?select=only_road_accidents_data_month2.csv>
* <https://morth.nic.in/road-accident-in-india>
* <https://www.who.int/entity/healthinfo/global_burden_disease/GHE_Deaths_2012_country.xls?ua=1>
* G. GURURAJ (2008), Road traffic deaths, injuries and disabilities in India: Current scenario, The National Medical Journal of India, volume 21, no 1, page 116
* <https://www.who.int/violence_injury_prevention/road_safety_status/country_profiles/india.pdf>