Work Plan

# Focus of Data Set

The dataset contains traffic violation information from all electronic traffic violations issued in Montgomery County. It is obtained from Montgomery County Maryland’s government data which is updated on a daily basis since January 1st, 2013.

# Research Questions

1. Are particular characteristics of a car (such as color & make) prone to higher incidents of traffic violations?
2. Is there an effect of the time of day/year (i.e. season/weather conditions, rush hours, weekends) on the number of traffic violations caused?
3. Does belonging to a particular race increase the probability of being issued a particular violation type?

# Timeline for project completion

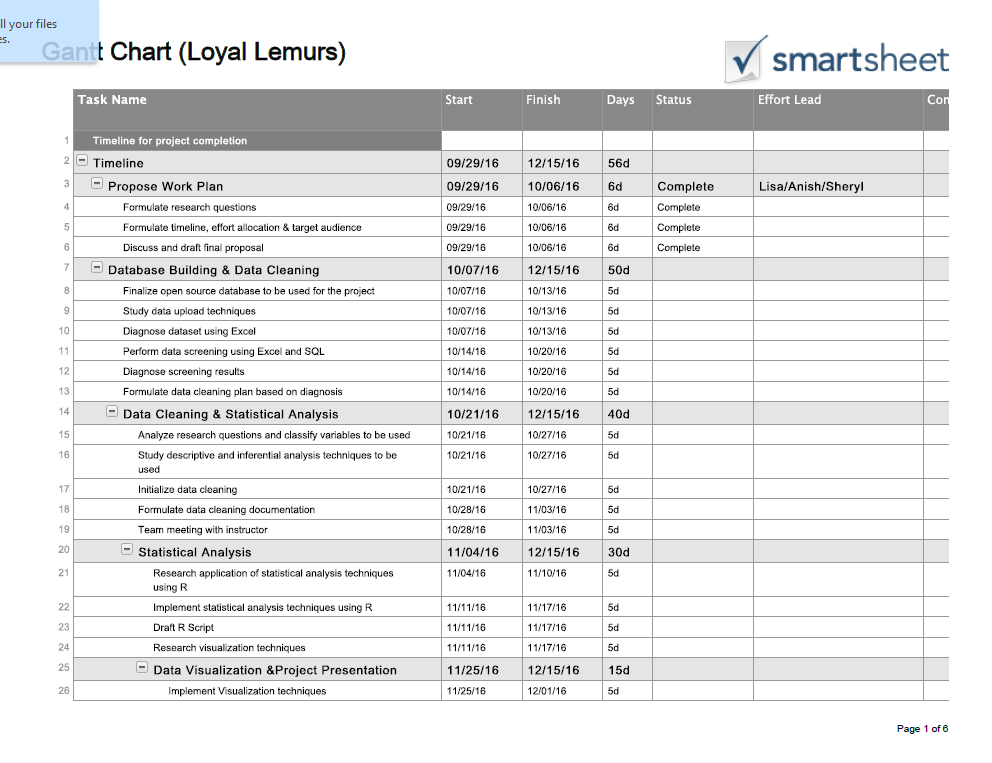
The timeline for project completion has been represented in the table below. The table contains the general task categories and sub-tasks to be performed within those categories. A target completion date has been assigned for each of the tasks.

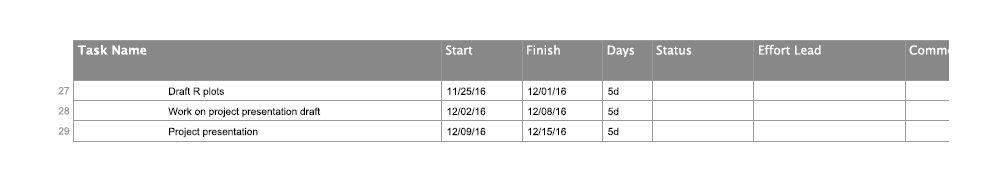
|  |  |  |
| --- | --- | --- |
| **TARGETTED DATE OF COMPLETION** | **TASKS** | **SUBTASKS** |
| October 6 | Propose work plan | * Formulate research questions * Formulate timeline, effort allocation and target audience * Discuss and draft final proposal |
| October 13 | Database Building & Data Cleaning | * Finalize open source database to be used for the project * Study data upload techniques * Diagnose dataset using Excel |
| October 20 | Database Building & Data Cleaning | * Perform data screening using Excel and SQL * Diagnose screening results * Formulate data cleaning plan based on diagnosis |
| October 27 | Data Cleaning & Statistical Analysis | * Analyze research questions and classify variables to be used * Study descriptive and inferential analysis techniques to be used * Initialize data cleaning |
| November 3 | Data Cleaning & Statistical Analysis | * Formulate data cleaning documentation * Team meeting with instructor |
| November 10 | Statistical Analysis | * Research application of statistical analysis techniques using R |
| November 17 | Statistical Analysis | * Implement statistical analysis techniques using R * Draft R Script * Research visualization techniques |
| December 1 | Data Visualization | * Implement Visualization techniques * Draft R plots |
| December 8 | Project Presentation | * Work on project presentation draft |
| December 15 | Project Presentation | * Project presentation |

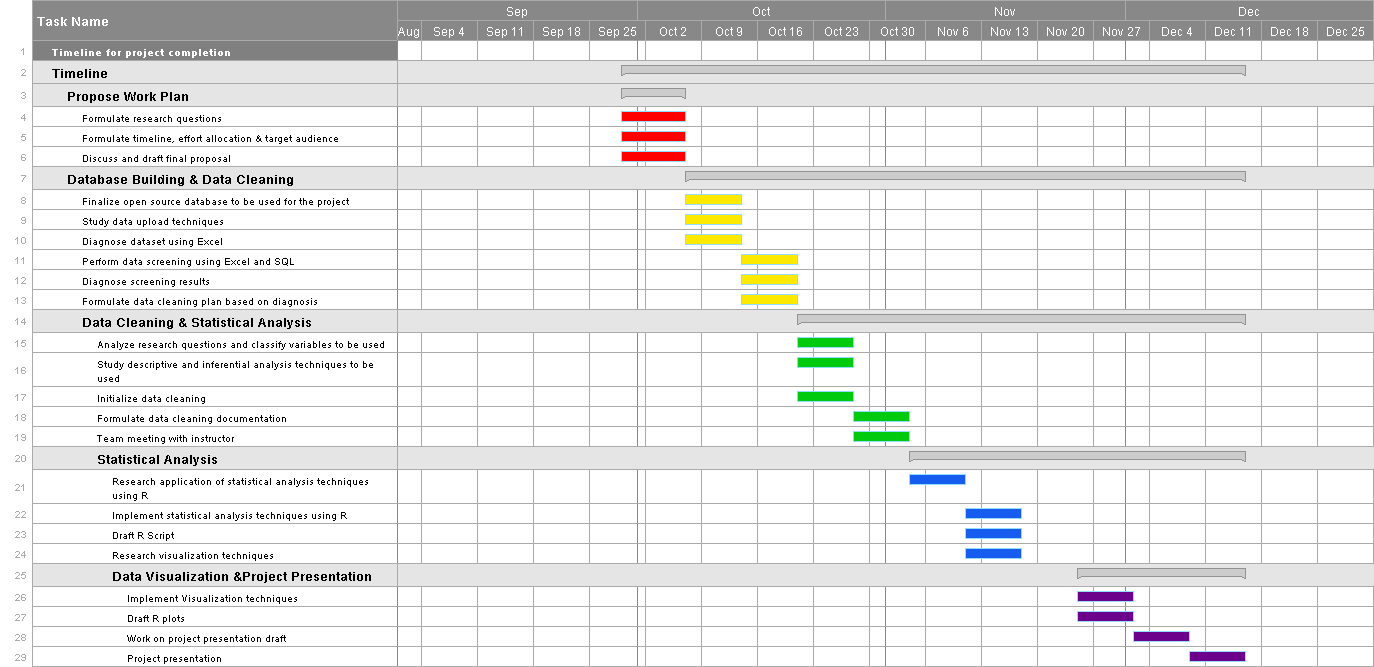
# Gantt chart

A Gantt chart depicting the project timeline has been shown below. Sub-tasks have been color coded depending on their parent task.

1. Propose Work Plan (Red)
2. Database Building & Data Cleaning (Yellow)
3. Data Cleaning & Statistical Analysis (Green)
4. Statistical Analysis (Blue)
5. Data Visualization & Project Presentation (Purple)







# Projected Effort Allocation

For this team project, all team members have decided on being equally involved in all stages of the project. A team meeting will be held at the start of every week to discuss the task and sub-task allocations. Based on prior knowledge and experience, one team member will be assigned as the head, to lead the efforts for a particular week, allocate responsibilities and guide the other team members. The Gantt chart will be updated with the status and details of the person who assumes primary responsibility as the project progresses. By adopting this method of effort allocation, each member will be exposed to every facet of the project be it data cleaning, analysis techniques, visualization techniques or project presentations. It provides a chance for increased learning and better co-operation and guidance. The project repository in GitHub will be regularly updated with the allotted tasks and sub-tasks and will be used for tracking as well.

# Target Audience

•   County Executives  
•    Division of Highway Services (MC DOT – Montgomery County Department of Transportation)   
•    Insurance Companies  
•    Police Department for deployment of officers/resources; training  
•    Civil Rights Agencies   
•    Mobile application developers

The effect of various factors such as season/weather conditions, rush hours, weekends on the frequency of violations issued can help the police department plan resource allocation and training of personnel. Policies can be framed to provide better training to people applying for motor vehicle licenses. The department of transportation and county executives can use this information to build better infrastructure to ensure safety. Mobile applications can be built to make drivers aware of the possible danger or high risk zones and time periods in the county. Insurance companies can use the information and trends extracted to frame policies and plans more suited to owners of a particular type of car. To some extent, civil rights agencies can make use of the information trends observed as well.

Word Count: 686