A title and summary for your project

* A description of the **goals** of the project - what are you aiming to show by your analysis
* A summary of the **datasets** that you will use in the analysis. Where were these obtained, what format are they in, what work (if any) will you need to do to make them usable
* The **techniques** you expect to use in the project (regression, clustering etc)
* A project plan - set at least two **milestones** for yourselves over the time that you have to complete the project.  For each milestone, describe what you should have achieved at that time and how you will evaluate whether you succeeded.

Project goal

1. ~~analyzing value of second hand cars~~
2. ~~self-driving car and data science~~
3. Electric car– energy distribution to different parts of vehicle
4. Car racing – analyzing car’s performance

Dataset

In this project, we will take the [data set] taken from [source] in [data format].

Through this study, we will develop a model that show the relationship between [different variables] and the outcome of [results].

1. Carsales.com.au to collect
   1. Maker, model, year manufactured, odometer
2. Electric vehicle trial data

Technique

Develop a predictive model that can;

1. Predict the sale price (buyer and seller) of used car based on different variables
2. Detect pedestrians on the road and avoid an accident
3. Identify the most (electric) energy consuming parts and improve on energy efficiency
4. Analyze the performance of racing car and make decision on optimal pit stop

Clustering data: Look for common group of properties

Predictive modelling:

* Linear regression on some variables within the group
* Logistic regression to establish a baseline classification performance
* Apply neural network to see if performance can be improved

Time Series modelling - simple, forward, backward

Milestones

1. Data preparation
   1. Collect and clean data set for analysis (ETL)
   2. Clustering algorithms to find structures within dataset
2. Develop a simple model to define a baseline performance
3. Apply a neural network for possible performance enhancement
   1. Supervised learning
      1. K-nearest neighbor algorithm
      2. Cross-validation to conduct the parameter tuning