**Abstract**

**Introduction**

In this project I will be analyzing the safety ratings of various brands of cars and how such a safety rating is related to other variables of the same vehicle. I will be using the automobile data set from the UC Irvine Machine Learning Repository, which consists of 205 observations and 26 variables. 24 of the variables account for a scalar attribute of the vehicle (fuel-type, weight, price, stroke, etc.). The other two variables are the safety rating (called symboling) and the make of the vehicle (Audi, BMW, etc.).

Symboling is represented as an integer that falls within the set [-3, 3]. A value of -3 indicates that the vehicle is considered safe, whereas a value of +3 indicates that the vehicle is “risky”. Symboling is calculated by… [NEED INFO].

Through understanding the attributes of automobiles that make them risky, we can alter these specific attributes to make automobiles generally safer.

**Goal**

The primary goal of this project is to find what attributes of a vehicle contribute most towards its safety rating.

**Methods**

**Results**

**Software and Packages Used**