Feasibility Study of Racing Car Club Project

Proposal

Abstract

This paper studies the statistical difference between states to provide insights for investments in the racing track business. To find the optimal location of investment, I studied exotic car ownership, weather, and demographics. Statistical tests and visualizations are used to aid data analysis in SAS.

Table of Contents

[1 Background 5](#_Toc37080002)

[2 Purpose 6](#_Toc37080003)

[3 Data 7](#_Toc37080004)

[3.1 vehicles.csv 7](#_Toc37080005)

[3.2 2017 Demographical Data by Counties 7](#_Toc37080006)

[3.3 Climate 8](#_Toc37080007)

[4 Analysis Plan 9](#_Toc37080008)

[5 Discussion 11](#_Toc37080009)

# Background

A company intends to build a racetrack for exotic cars in the USA. This racetrack is designed to target high-net worth individuals who are passionate about sports cars and car racing. The racetrack, after building, will partner with local exotic car dealerships and car clubs to host racing shows. Also, this racetrack will allow customers to become members so they can use this track to race too. This company will also build peripheral infrastructures that would provide both a luxury/convenient place to have fun, and enough support for car racing. This company wants to decide the potentially good regions to build this project. At this point this company only needs to know how big the market can be and where the market is, so factors like building cost, opportunity cost, regulations, and etc. are not of its concern.

# Purpose

In this paper, I will find the regions that have large high net worth population, high sports car parc, and good weather for car racing that provide big enough market. Then I will use demographical data including income and age to find out how many potential customers will this racetrack have

# Data

## vehicles.csv

<https://www.kaggle.com/austinreese/craigslist-carstrucks-data>

This dataset contains more than 500,000 records of used car sales on craigslist, containing all used car transactions ever happened on craigslist, collected in January 2020.

This dataset features 25 variables including craigslist region, entry price, entry year, model of vehicle, condition of vehicle, generic type of vehicle, state of listing, and etc. I will mainly use type and price to filter exotic sports cars out, and sort them by state and county to find which state has the most exotic cars sales, and therefore find the ideal region to start the car club. Car condition might be a useful factor, reflecting if the car is heavily used for racing and fast driving, or if it is carefully maintained as a collection or commute tool.

## 2017 Demographical Data by Counties

<https://www.kaggle.com/muonneutrino/us-census-demographic-data>

This data is collected from the 2017 census, it contains demographical information of 3220 counties. It features variables including population, ethnic population, employment status, professions, income per capita, and means of transportation. These are all useful information to help find out the number of potential customers in each county. It is obvious that the counties with high income per capita and high population drive to work are likely to have more potential customers, but whether other factors including ethnic, professions, and gender would have a relationship with tendency to become customers require further research. Unfortunately, I could not find any data about age group, but I will keep looking for it. The age group will be very useful in defining the number of potential customers.

## Climate

<https://www.usclimatedata.com/>

This data is from US Climate Data, featuring monthly average temperature of every state. Since car racing is an outdoor sport, states with very low average temperature will not be considered.

# Analysis Plan

Step One--Exploratory Research:

I will read articles and find study’s that address the relationship between factors such as income, means of commute, age, gender, profession and ethnic and tendency to own exotic cars and to participate in car racings.

Then, I will read studies and interview car racers and outdoor sport viewers to find out what temperature and weather are best for car racing and what weather will not permit car racing.

I will also try to find other important factors for operating a racetrack that I might have overlooked in structuring my analysis and make adjustments accordingly.

Step Two—Descriptive Research:

In “Vehicles” dataset, I will filter out the data that have prices above 40,000USD and have type as “coupe” or “convertible”. Also, to make the processing easier and to make sure the analysis reflects the current situation, I will only use data within the most recent 3 years. After filtering the data, I will count the numbers of records in each state, to locate 5 states that have high number of expensive sport cars.

In the demographical data set, I will find those states that have high population of our potential customers (defined using factors found in step one).

In the climate data, I will cross out 10 to 20 states that have most months of average temperature and other weather factors that are not fit for car racing.

Finally, I will find states that have the ideal number of exotic cars, enough population of potential customers, and ideal climate. It is likely that I will find multiple states that fulfil these criteria, then I choose the state that has the most exotic cars, because unlike the weather, which just needs to be good enough, more cars and people bring more income for the racetrack, and the number of cars is more clear/concrete than the number of potential customers.

Step Three--Hypothesis Testing:

I will use hypothesis testing to test that the selected state has statistically significant difference in number of exotic cars.

# 

# Discussion

My analysis gives this company a clear idea of which state is an idea location for building and operating a racetrack, in term of marketing. This information will be vital in deciding where to invest this racetrack and give a clear direction of further, more structured analysis.