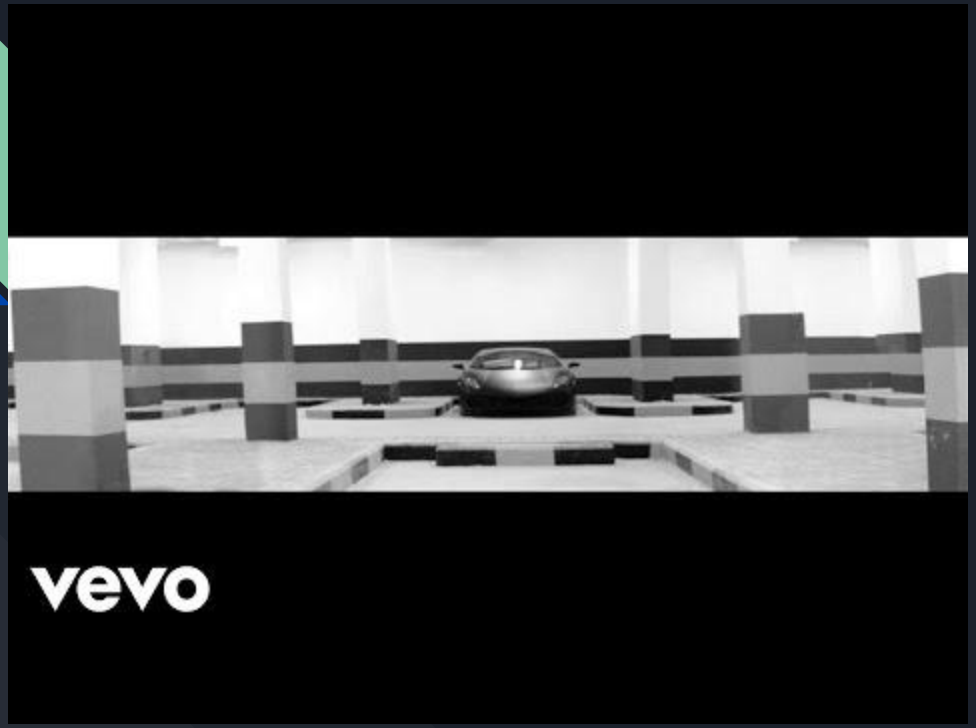
Horsepower

A Statistical Study by Walter Castillo, Bradley Harrison, Richy Peterson and Joanna Vasquez

"Horsepower, horsepower, all this Polo on, I got horsepower…"
-2 Chainz, "Mercy"


vevo

The Data

	mpg	cylinders	displacement	horsepower	weight	acceleration	model	origin	car_name
0	18.0	8	307.0	130.0	3504.0	12.0	70	1	chevrolet chevelle malibu
1	15.0	8	350.0	165.0	3693.0	11.5	70	1	buick skylark 320
2	18.0	8	318.0	150.0	3436.0	11.0	70	1	plymouth satellite
3	16.0	8	304.0	150.0	3433.0	12.0	70	1	amc rebel sst
4	17.0	8	302.0	140.0	3449.0	10.5	70	1	ford torino

- **398 rows (each representing a unique model), 9 columns:**
  - mpg: miles per gallon of vehicle (numerical)
  - cylinders: the number of cylinders in this vehicle (numerical)
  - horsepower: horsepower of vehicle (numerical)
  - weight: weight of vehicle in pounds (numerical)
  - acceleration: acceleration in feet per second squared (numerical)
  - model: Model year of the vehicle (categorical)
  - origin: Country of origin where 1 is the United States, 2 is Germany, and 3 is Japan (categorical)
  - car_name: Description of car (categorical)

The Data (cont.)

	mpg	cylinders	displacement	horsepower	weight	acceleration	model	origin	car_name
0	18.0	8	307.0	130.0	3504.0	12.0	70	1	chevrolet chevelle malibu
1	15.0	8	350.0	165.0	3693.0	11.5	70	1	buick skylark 320
2	18.0	8	318.0	150.0	3436.0	11.0	70	1	plymouth satellite
3	16.0	8	304.0	150.0	3433.0	12.0	70	1	amc rebel sst
4	17.0	8	302.0	140.0	3449.0	10.5	70	1	ford torino

We wanted to examine the horsepower for different car manufacturers, so we focused on the ‘horsepower’ and ‘car_name’ columns of the dataset.

Cleaning

- A few entries in the ‘horsepower’ and ‘car_name’ columns had ‘?’ instead of a value so we removed so we removed those entries
- We created a new row ‘car_makes’ that contains only the manufacturer name for each car
- Cars made by Chevrolet were represented by a few different names in the ‘car_makes’ column, so we found all of these names and changed them to ‘chevrolet’

Our Project Goal

“To See Whether There Was a Significant Difference Between the Horsepower of Ford and Chevrolet Cars Made from 1970-1982.”

Null Hypothesis: There is no significant difference between the horsepower of Ford and Chevrolet Cars.

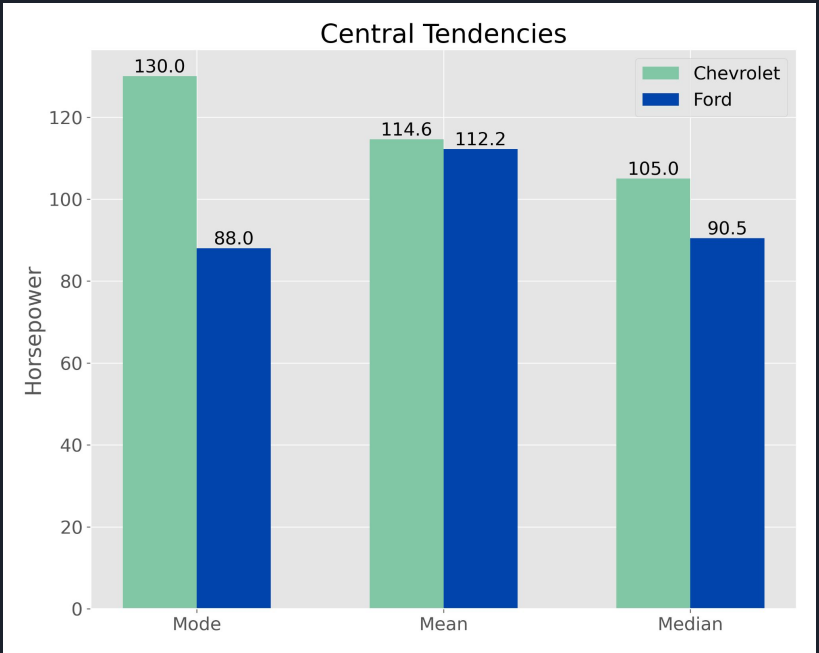
Alternate Hypothesis: Ford cars typically had less horsepower than Chevrolet Cars.

HORSEPOWER

"...the power an engine produces. It's calculated through the power needed to move 550 pounds one foot in one second or by the power needs to move 33,000 pounds one foot in one minute."

Exploratory Data Analysis

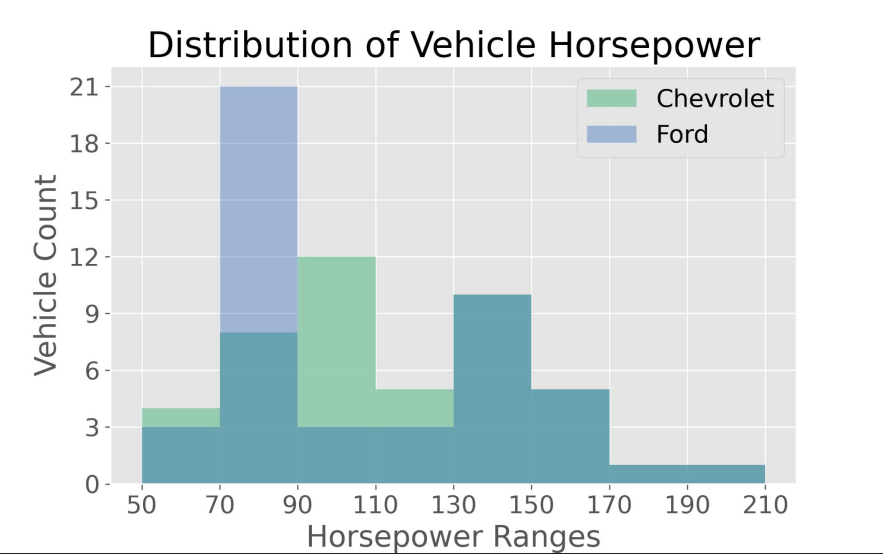
To get a general idea of how Chevrolet and Ford horsepowers differed, we looked at the mean, median and mode:



In all cases, it appears that Chevrolet surpasses Ford in the amount of average horsepower. The biggest discrepancy was noted in the most frequently seen horsepower, with Ford at 88 hp and Chevrolet at 130 hp.

Exploratory Data Analysis (cont.)

To better understand the distribution of our data, we created a histogram of the horsepower for Ford and Chevrolet:



We initially wanted to compare the average horsepower for the two manufacturers, but after looking at the distribution of the data, we realized that it was not normally distributed.

We decided to use **the Mann-Whitney Test** to compare the two populations.

Mann-Whitney U Test

Takes one datapoint from one dataset and compares it to all the data points in the second dataset and repeats the process until all data points from the first and second datasets have been compared.

HYPOTHESIS TESTING

Null Hypothesis: There is no significant difference between the horsepower of Ford and Chevrolet Cars.

Alternate Hypothesis: Ford cars typically had less horsepower than Chevrolet Cars.

Significance Level:
α = 0.05
Power = 0.80

Result:

After performing the Mann-Whitney U Test, we received a P-Value of 0.21, and because our P-Value was greater than our Significance Level (α), we **FAIL TO REJECT THE NULL HYPOTHESIS.**

Conclusions

Based on our test results, we cannot say whether there was a significant difference between the horsepower of the Ford and Chevrolet cars manufactured between 1970 and 1982.

Despite these findings, Chevrolet had higher mean, median and mode horsepower than Ford. This might be because the Chevrolet dataset included several trucks while the Ford dataset only contained a few.

Future Research:

Break data up by vehicle type (i.e. Trucks, Sports cars, etc.) and compare them between manufacturers.

Questions?