

ANALYSIS OF CARS

BDA 03



Akbota Kairatova, Zhannur Murat, Dayana Kassenova

CONTENT

- INTRODUCTION
- RESEARCH QUESTIONS
- WEB SCRAPING
- DATA PREPARATION
- EDA
- ENCODING AND PREPARING CATEGORICAL VARIABLES
- MODEL BUILDING & PREDICTION ACCURACY
- CONCLUSION



INTRODUCTION: CARS PRICE PREDICTION ANALYSIS

- THE PURPOSE OF OUR CARS PRICE PREDICTION ANALYSIS WAS TO DEMONSTRATE AND PREDICT PRICES OF CARS IN HAMILTON CITY. ACCORDING TO THIS ANALYSIS, PEOPLE CAN EASILY FIND THE FAVORITE AND SUITABLE CAR FOR FUTURE DRIVING IN THE CITY.
- FIRSTLY, WE HAVE SCRAPED THE DATA MORE THAN 1300 ROWS FROM THE WEBSITE KIJJI.COM, WHICH ARE GROUPED BY MODEL, BRAND, PRICE, COLOR, CONDITION, TRANSMISSION, FUEL TYPE AND MILEAGE.
- AFTER THAT WE HAVE DONE AN EXPLORATORY DATA ANALYSIS BY MAKING GRAPHS IN ORDER TO UNDERSTAND AND FIND OUT MORE CLEARLY THE DATASET. FURTHERMORE, WE HAVE MADE A DATA PREPARATION BY DROPPING MISSING VALUES, AND PREPROCESSING BY USING STANDARD SCALER LIBRARY. FINALLY, CREATED THE SUITABLE MODELS BY USING LINEAR REGRESSION AND KNN CLASSIFIER.



RESEARCH QUESTIONS



Analyze the which colored cars people are prefer to drive

Analyze the ratio of car transmissions

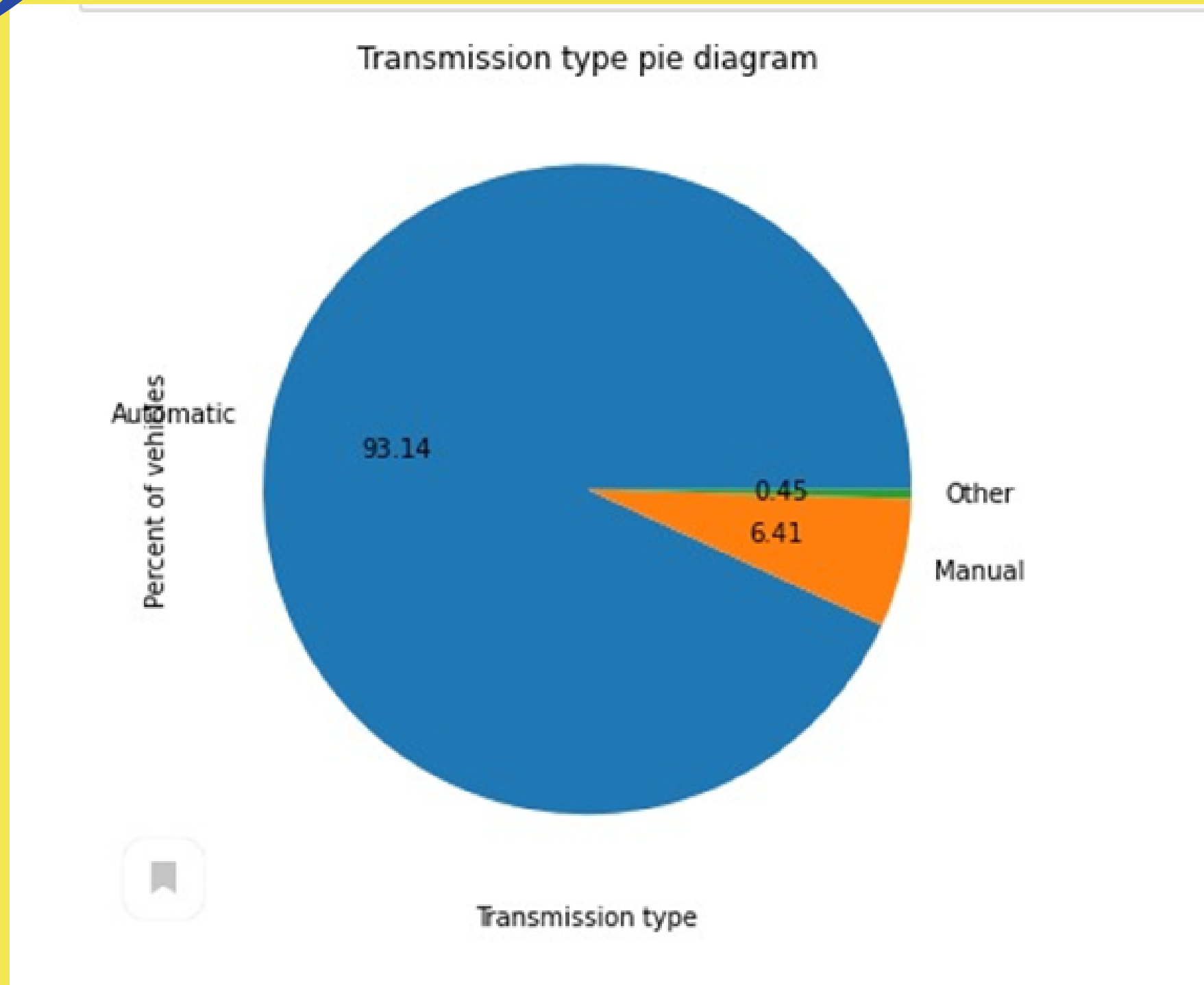
Analyze the time period when cars had high selling (sale)

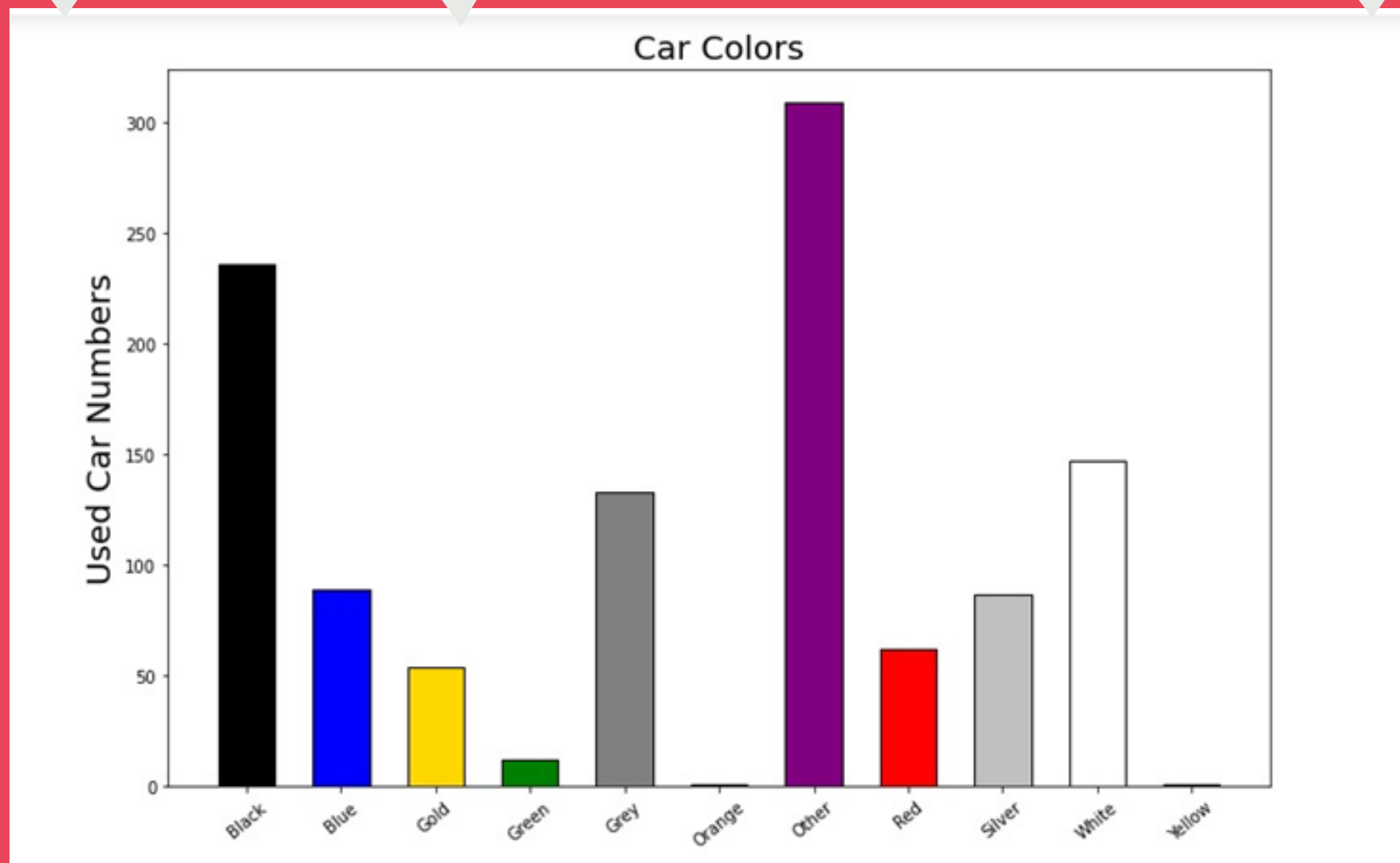
Analyze the correlation between fuel type and car transmission

Analyze the connection between the price and mileage category of each car transmission type

AS IT IS CLEAR

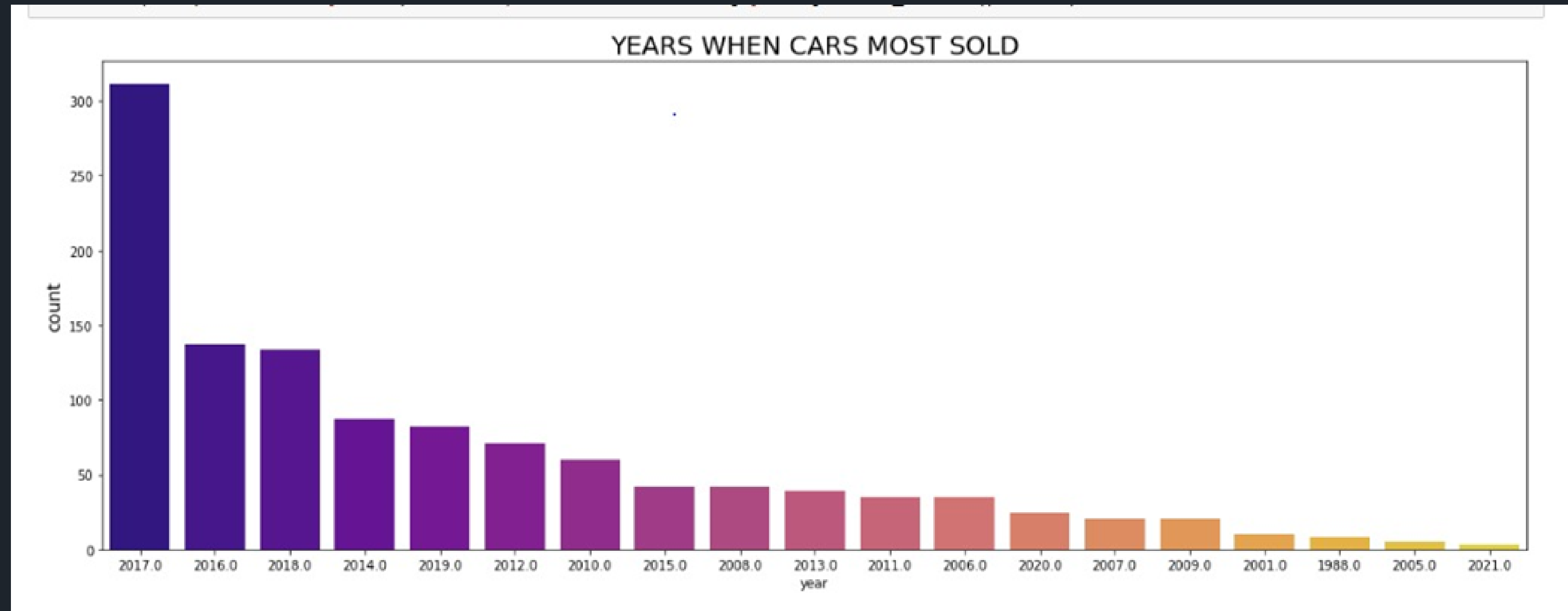
**IN THE PIE CHART
THE AUTOMATIC
TRANSMISSION
DOMINATES ON
MANUAL ALMOST
9 TIMES.**



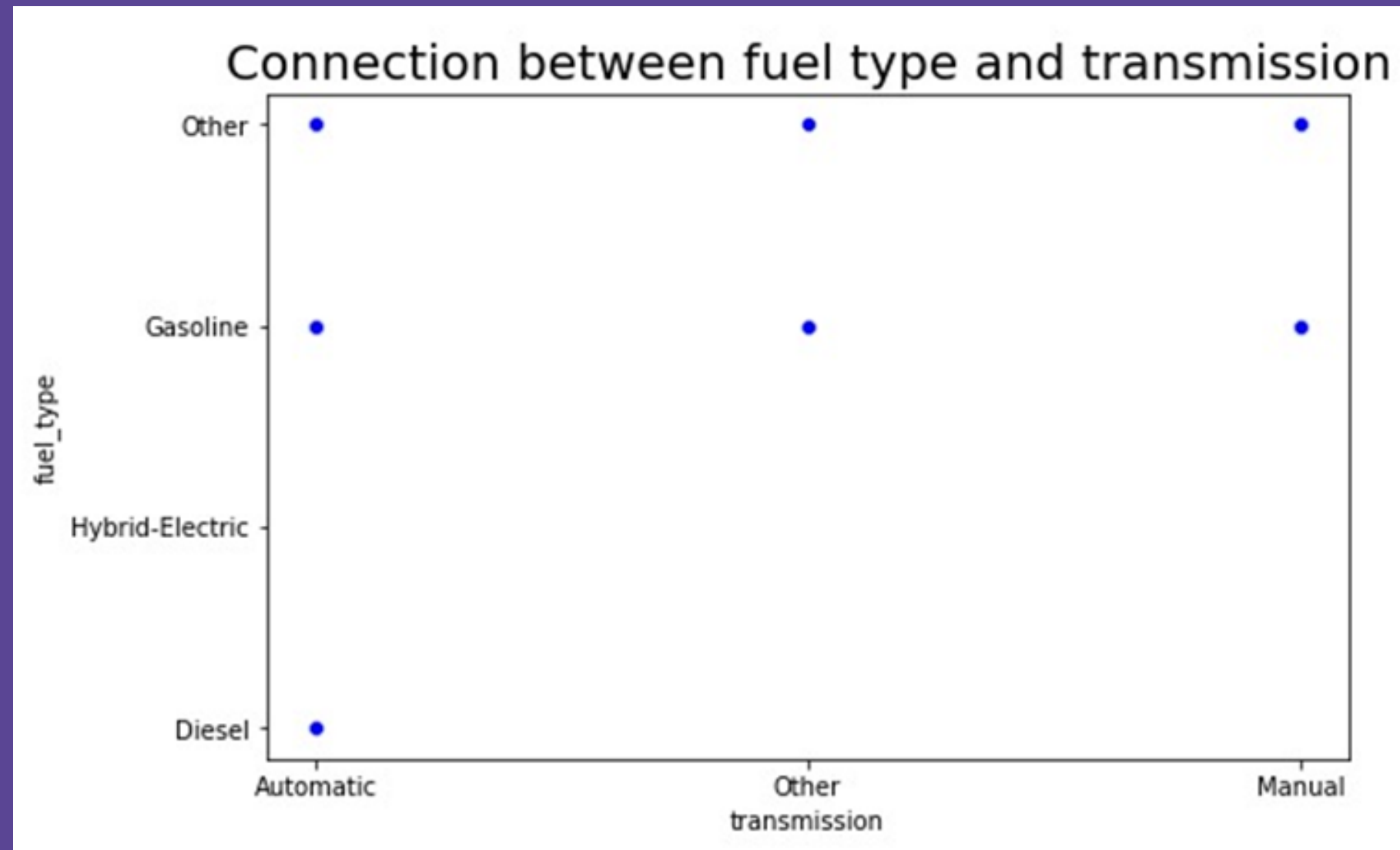


ACCORDING TO THE GRAPH, THE MOST USED COLOR WAS BLACK, WHILE THE LEAST YELLOW.

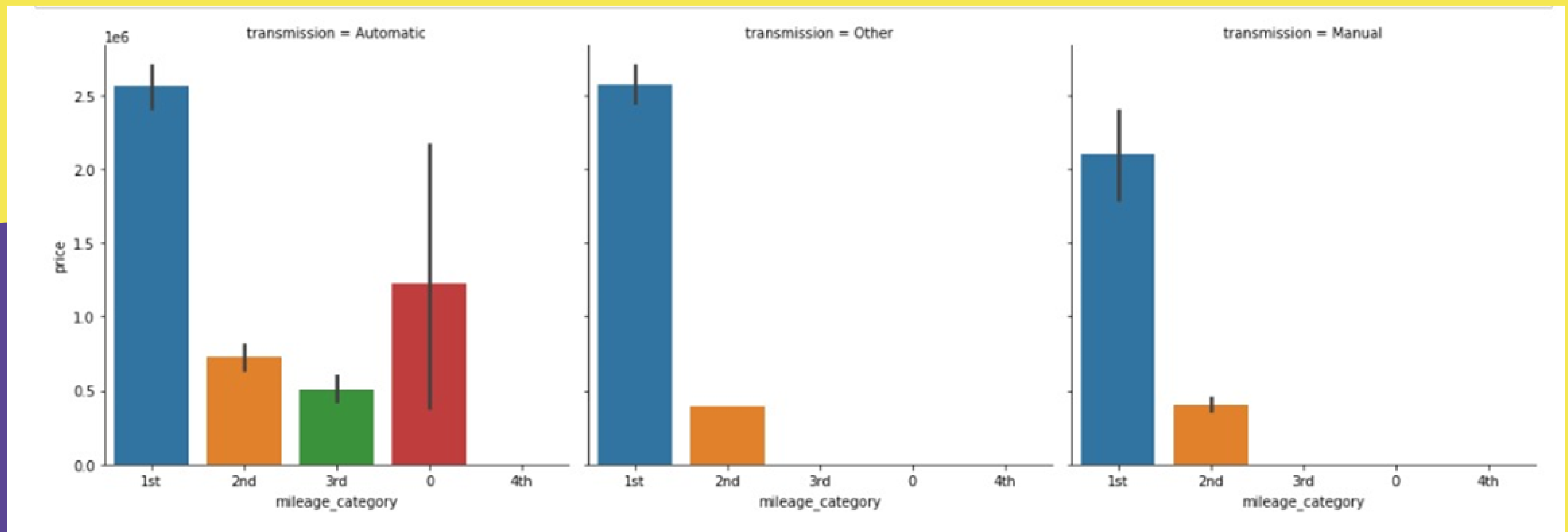
Cars that are produced 6-7 years ago still actual than the newly produced cars approximately 10 times more.



THE FOURTH CHART SHOWS THE RELATION BETWEEN FUEL TYPES AND CAR TRANSMISSION. SO AUTOMATIC TRANSMISSION REPRESENTS MOST CORRELATION WITH ALL FUEL TYPES, WHILE MANUAL ONLY WITH GASOLINE.



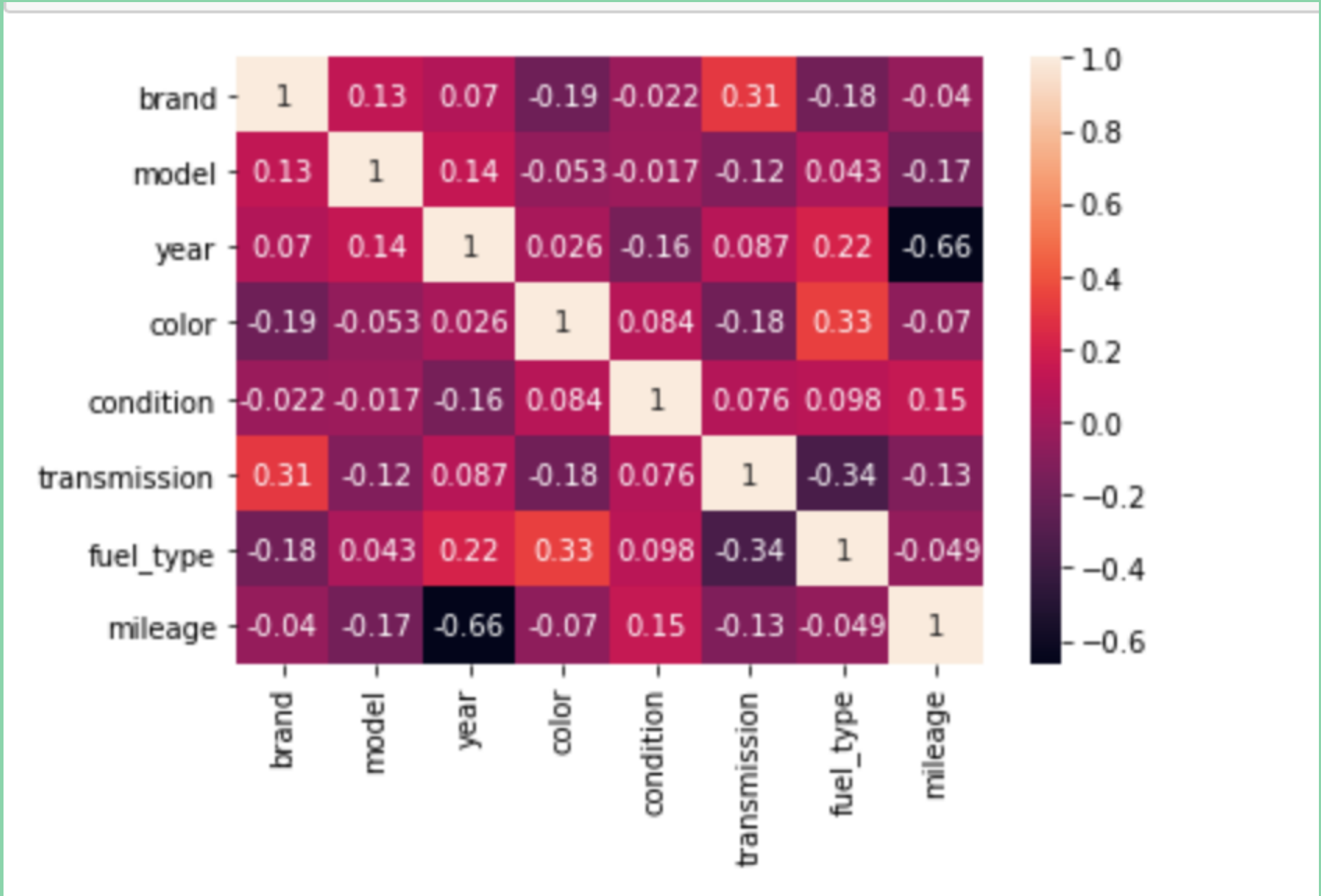
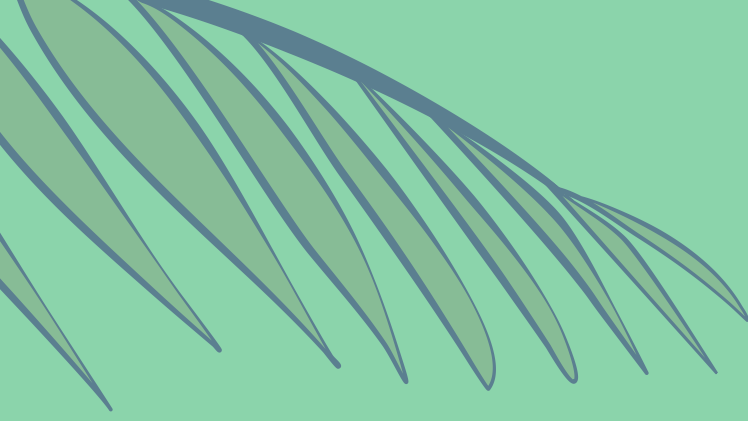
Finally, this illustrates the connection between the price and mileage category of each car transmission type.





CORRELATION MATRIX BEFORE ENCODING





AFTER LABEL ENCODING



REFERENCE LINKS:

1. [HTTPS://SCIKIT-LEARN.ORG/STABLE/MODULES/GENERATED/SKLEARN.NEIGHBORS.KNEIGHBORSCLASSIFIER.HTML](https://scikit-learn.org/stable/modules/generated/sklearn.neighbors.KNeighborsClassifier.html)

2. [HTTPS://WWW.KIJJI.CA/](https://www.kijiji.ca/)

3. [HTTPS://WWW.GEEKSFORGEEKS.ORG/LINEAR-REGRESSION-PYTHON-IMPLEMENTATION/](https://www.geeksforgeeks.org/linear-regression-python-implementation/)

4. [HTTPS://SCIKIT-LEARN.ORG/STABLE/MODULES/CLASSES.HTML?HIGHLIGHT=PREPROCESSING#MODULE-SKLEARN.PREPROCESSING](https://scikit-learn.org/stable/modules/classes.html?highlight=preprocessing#module-sklearn.preprocessing)

5. [HTTPS://STACKABUSE.COM/LINEAR-REGRESSION-IN-PYTHON-WITH-SCIKIT-LEARN/](https://stackabuse.com/linear-regression-in-python-with-scikit-learn/)



*Thank
you!*

