Group 9 Final Project

By: Nana Amoah, Ethan Hisle, Joseph Rae, Quintin Sargent, and Noah Donnelly

Teamwork Contribution

Nana- worked on the Project Proposal and finding the dataset, worked on creating questions and structuring the dataset, worked on Data Analysis I, worked on the project report.

Ethan- worked on the Project Proposal and finding the dataset, worked on creating questions and structuring the dataset, worked on Data Analysis I, worked on the project report.

Joe- worked on the Project Proposal and finding the dataset, worked on creating questions and structuring the dataset, worked on Data Analysis I, worked on Data Analysis II and Predictive Modeling.

Quintin- worked on the Project Proposal and finding the dataset, worked on creating questions and structuring the dataset, worked on Data Analysis I, worked on Data Analysis II and Predictive Modeling.

Noah- worked on the Project Proposal and finding the dataset, worked on creating questions and structuring the dataset, worked on Data Analysis I, worked on the presentation.

Background

Background: Over the past 12 years in Baltimore there were more than 265,000 instances of towed vehicles. Baltimore kept detailed records of each towing instance.

Baltimore has struggled with negative press coverage regarding potential corruption in its towing practices with officers and councilpersons being indicted for theft, fraud, and extortion for their involvement in corruption scandals

Business Application Scenario

Scenario: The City of Baltimore is performing an audit to ensure that their towing practices are unbiased. They are interested in having some potential questions answered about their towing practices in order to create a more transparent, effective, and profitable process. They hired us to analyze their data and present the outcome as a neutral party.

Project Questions

Questions to answer:

- **Is the city of Baltimore unfairly favoring one towing company over the other competitors?**
- **Is the city of Baltimore towing specific types of vehicles more than others?**
- **Is the city of Baltimore towing visiting and traveling visitors disportionately to local residents?**
- **Can we predict if a vehicle will be auctioned based on the data provided in the dataset?**

Data Collection

The dataset is available as a csv file provided publicly by the city of Baltimore.

- [ArcGIS Table]
- [Table Description]
- [Raw Data]

It has 39 columns and 265,816 rows. Large Dataset.

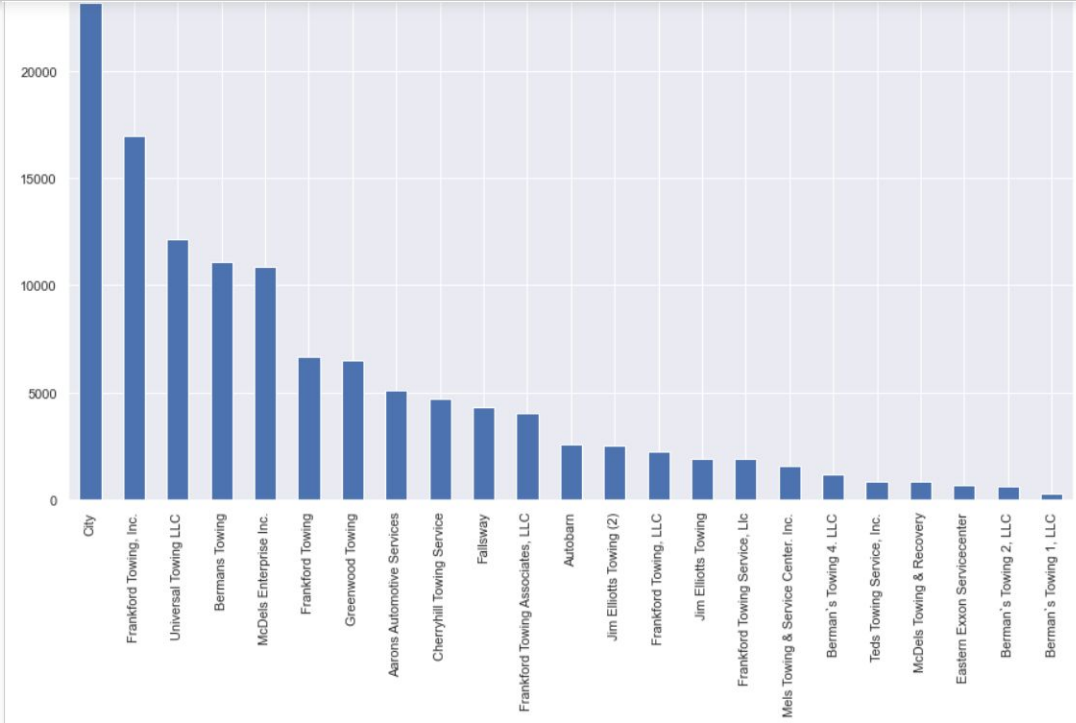
The data is not clean. This dataset contains many missing values and there are multiple time variables which will need to be reformatted into integers for ease of analysis.

Of the 265,000 records, nearly 253,000 had at least 1 null value - after dropping a few inconsequential and repetitive columns we were able to retain ~120,000 records

The dataset has not been analyzed by anyone else. The dataset appears to be untouched.

Q1) Is the city of Baltimore unfairly favoring one towing company over the other competitors?

When looking into the first question, we found that there were multiple entrances for the same towing company in the dataset. We suspect this is because there are multiple people entering in data



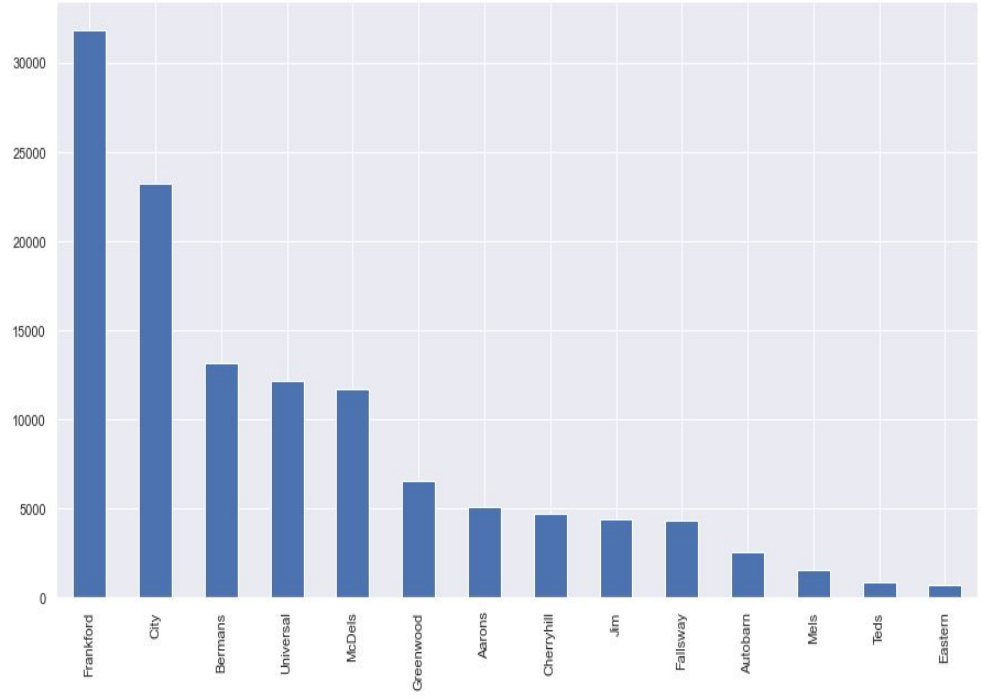
Question 1 (continued)

When combining the companies together, we find that Frankford is the most popular towing company followed by City and Bermans.

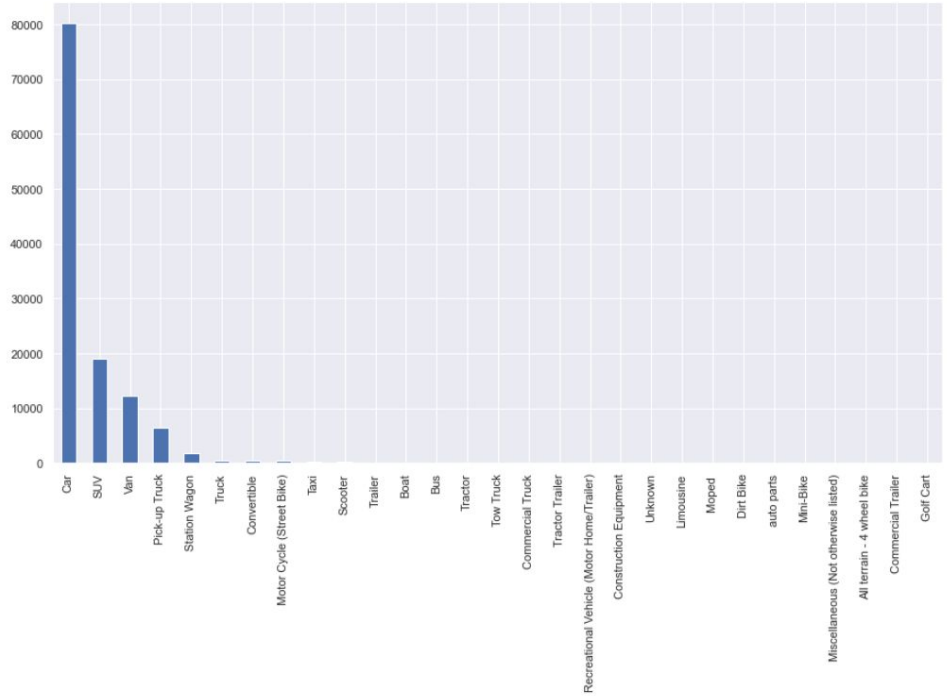
Although the towing choices are a little spread out among the companies, Frankford has 32,812 choices, City has 23,180 choices, and the next company only has 13,123 choices. Maybe these towing companies have contracts with the city? Could be because of multiple locations/bigger business.

Q1) Is the city of Baltimore unfairly favoring one towing company over the other competitors?

- **Not necessarily but could be possible (maybe look into further)**

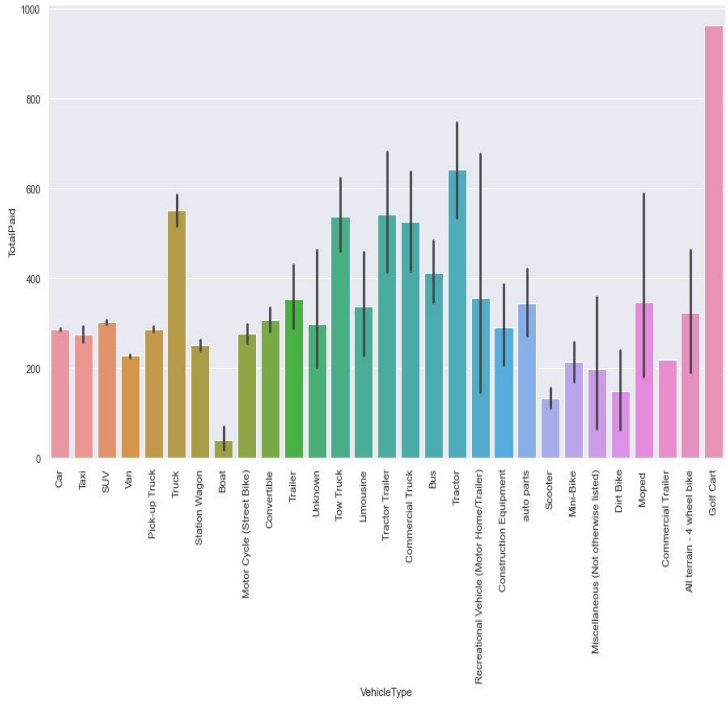


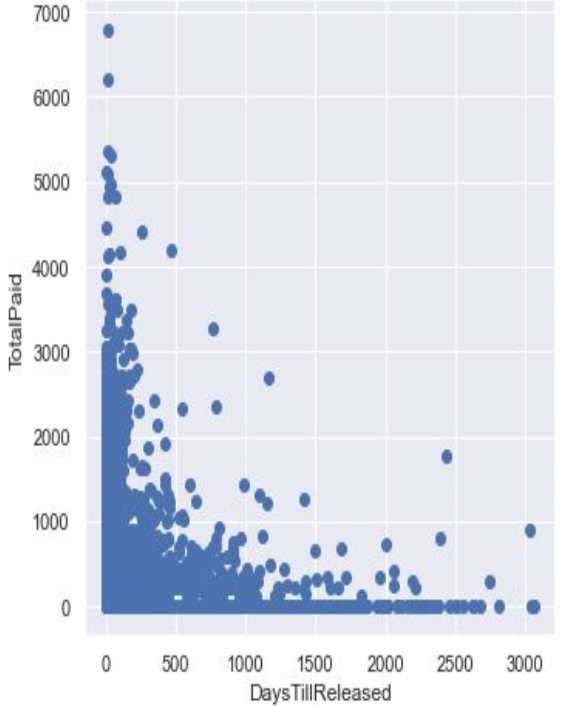
Q2) Is the city of Baltimore towing specific types of vehicles more than others?



When looking at the second question, we can see that cars get towed way more by far than any other type of vehicle. Although cars may seem like a clear choice, we wanted to also do analysis to see if other types of vehicles were being unfavorably treated even if they are not as prevalent as cars.

Question 2 (continued)





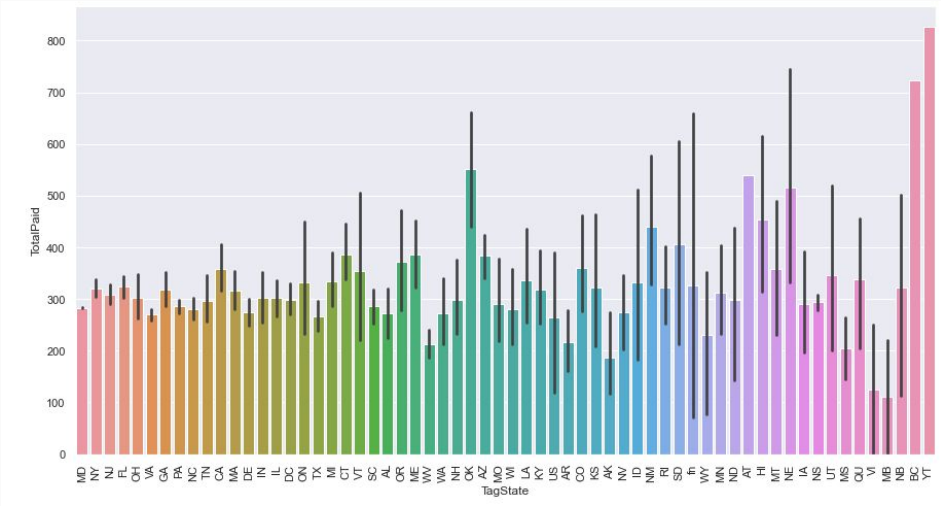
We can see that smaller types of vehicles have smaller charges, and people tend to not pay big charges as more time is spent in the impound. If the city was targeting a specific type, they would(should) go for cars of higher price.

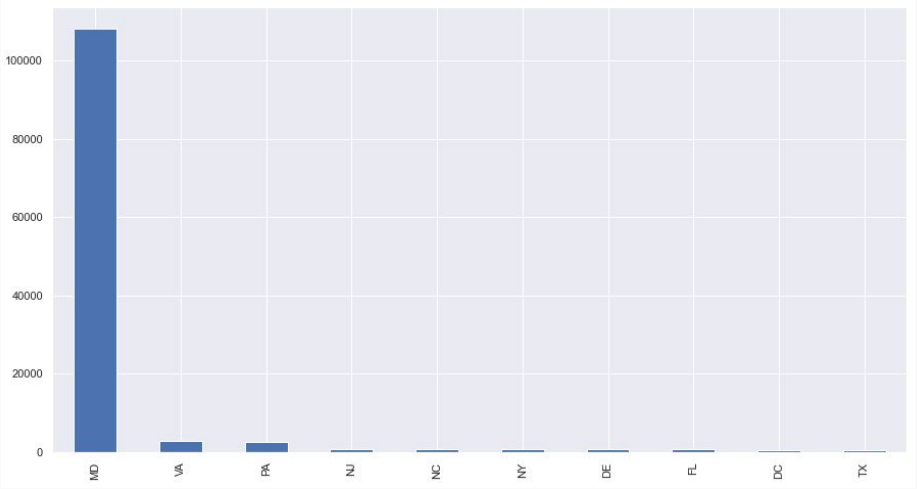
Q2) Is the city of Baltimore towing specific types of vehicles more than others?

- **No**

Q3. Is the city targeting visitors disproportionately?

In the barchart of total tows by state tags, we can see that Maryland has the vast majority of occurrences, with neighboring states being towed much less frequently





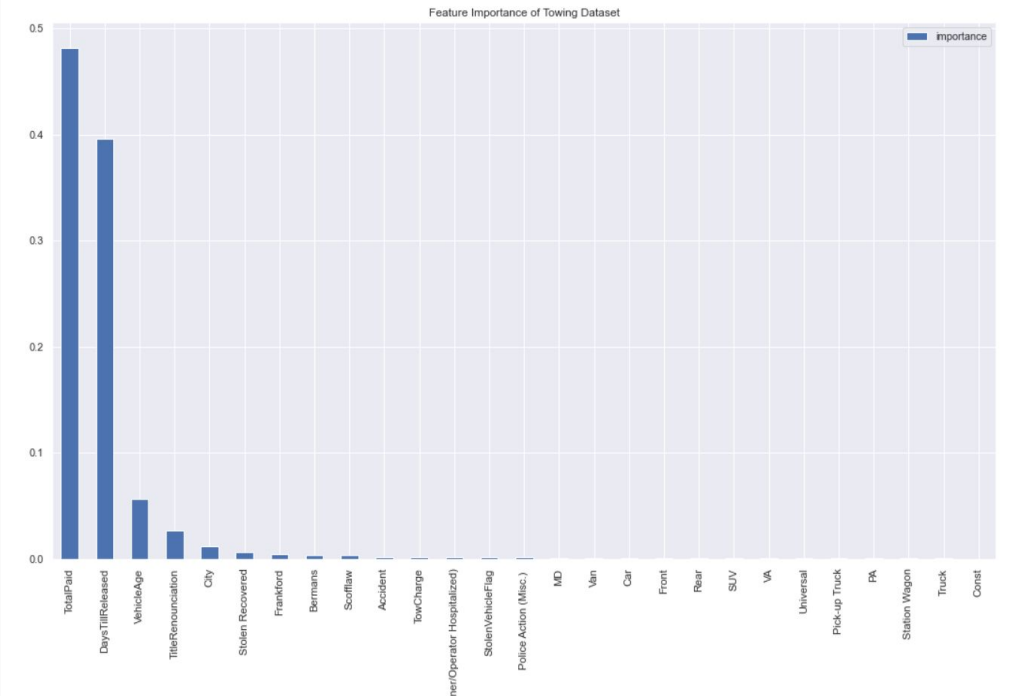
No, It is evident that Baltimore is not targeting or towing visitors disproportionately

Q4) Can we predict Baltimore's auctioning trends?

We conducted 6 different regression analysis and predictive modeling methods to find which method will work best for this dataset. Artificial NN was found to be the best model with an AUC score of 99.51%

	Accuracy	Precision	Recall	F1 Score	AUC
k-NN (k=5)	0.975385	0.953419	0.902439	0.927229	0.946583
Gaussian Naive Bayes	0.920083	0.807805	0.708724	0.755028	0.836630
Bernoulli Naive Bayes	0.988630	0.961120	0.973968	0.967501	0.982841
Decision Tree	0.991768	0.978333	0.974203	0.976263	0.984832
Random Forest	0.993928	0.978600	0.986632	0.982600	0.991047
Artificial NN	0.994865	0.975414	0.995544	0.985376	0.995133

Question 4: Impact of variables



We found that the totalPaid and daysTillReleased variables had the greatest impact on if a vehicle was auctioned.

Q4. Can we predict Baltimore's auctioning trends?

YES, using our artificial neural network, we could correctly predict if a car would be auctioned off or not.

Conclusion 1

Questions to answer:

- Is the city of Baltimore unfairly favoring one towing company over the other competitors?
  - **Possibly** - further analysis is required. It is evident that Frankford Towing and the city's towing branch are the most commonly used. However, it is unknown if this is based on availability/effectiveness or if it is caused by nepotism or corruption.

Conclusion 2

- Is the city of Baltimore towing specific types of vehicles more than others?
  - **NO** - cars are the most commonly towed vehicles, and yet have a lower price, so there does not seem to be a profit motive in their towing practices. Further analysis could involve determining the distribution of total amounts of vehicle type ownership throughout the city. It is possible that many of the residents of the city prefer to drive cars - skewing the results

Conclusion 3

- Is the city of Baltimore towing visiting and traveling visitors disportionately to local residents?
  - **NO** - Cars with Maryland tags are the most commonly towed vehicles, and the total amount paid is fairly consistent across states. However, we do see that impound charges are the highest in the first few days after towing and decrease with time - possibly indicating that the city charges high rates for people to retrieve their vehicles quickly

Conclusion 4

- Can we predict if a vehicle will be auctioned based on the data provided in the dataset?
  - **YES** - using our model, basing it mostly off of the totalPaid and DaysTillReleased variables, we could correctly predict the auction outcome of a towed vehicle.

Final Conclusion

- It is unlikely that the city of Baltimore is corrupt in their towing practices
- Further analysis is needed on their choice of towing companies, but with the data available it is not evident

Q & A

Any Questions?