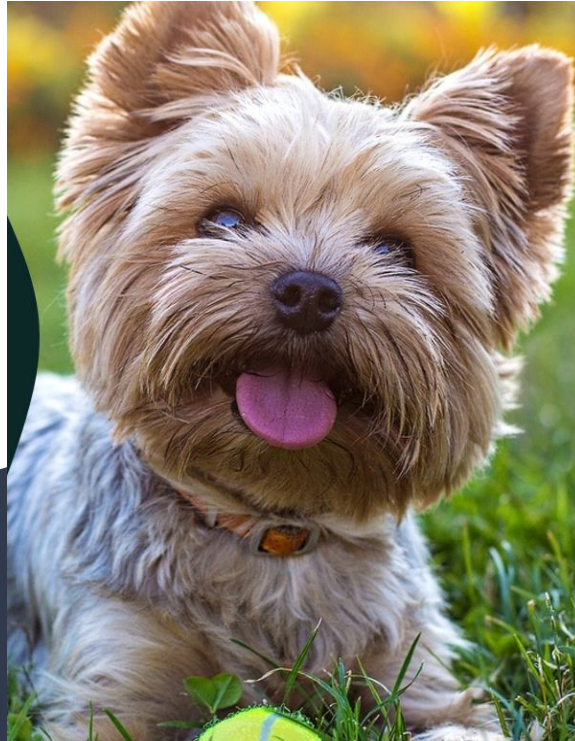


Dog Biting Trends in New York

Presentation Team:

Alex Moore
Thripura Pakala
Joy Weishan



Introduction

PURPOSE:

- The dog bite dataset was chosen because we found a common interest of dogs and all own dogs.
- Demonstrate data cleaning, data analysis, data visualizations and data presentation skills
- Show our team has grasped the concepts provided to date in the course
- Display through data visualizations what potentially increases their chances of biting and hope this educates our larger audience as well.
- Answer Data Question through Data Visualization
 - What type of Breeds are more likely to bite?
 - How do Seasons and Days of Week impact biting?
 - Does the age of a dog show biting trends?
 - Are dog bites increasing or decreasing

Data Cleaning

-Breed:

- Over 1600 breeds
- Misspelling, added characters, mixes
- Dropped unknown/rare breeds

-Age:

- The age data had decimals, alpha as well as timestamps in it.
- Examples of some of the data prior to data cleaning: 1W, 3WKS, 4WK, 2M, 2MTH, 2MTHS, 5Y, 4 YRS, Timestamps. before and after the alpha characters, There was a dog with and age of 41 that was deleted with breed processing.
- Dropped Null Ages

-Date/Season:

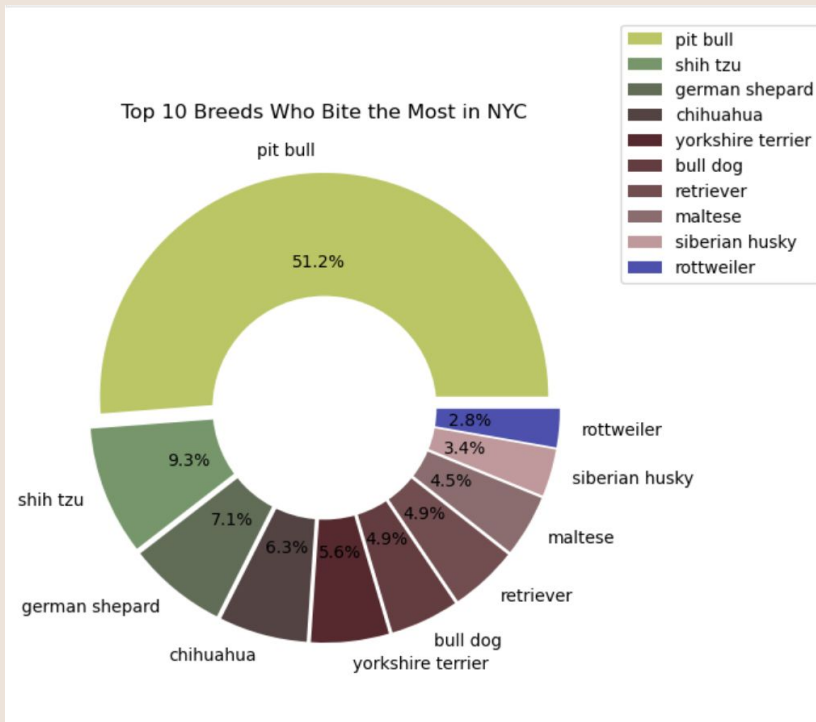
- The DateOfBite had Dates in text format, converted into datetime, and split into year, month, day, based on the month, day, created seasons, week of days cloumns. Set as week starts on Monday(Mon=0)
- Dropped the column Species.(Dog)

Leading Breeds in Bites

- Leading the pack are Pit Bulls
- Small dog BIG bite
- Dogs leading in most bites are all common types of breeds
- NYC most common dogs, bull dog, poodle, retrievers, and pit bulls



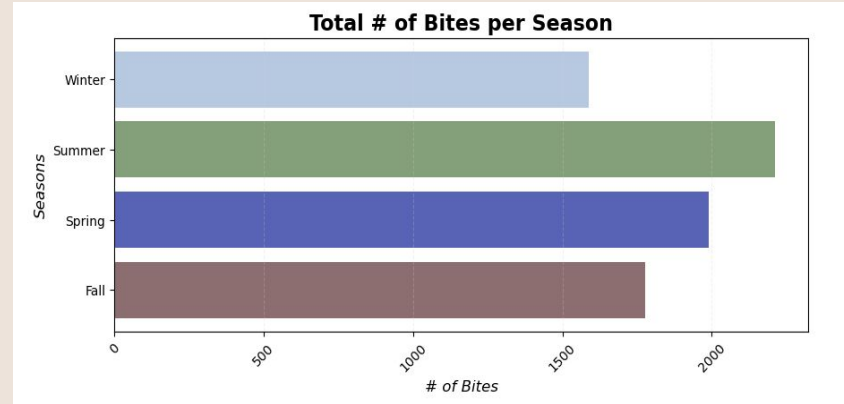
AMERICAN
KENNEL CLUB®



Seasons/Day of Week

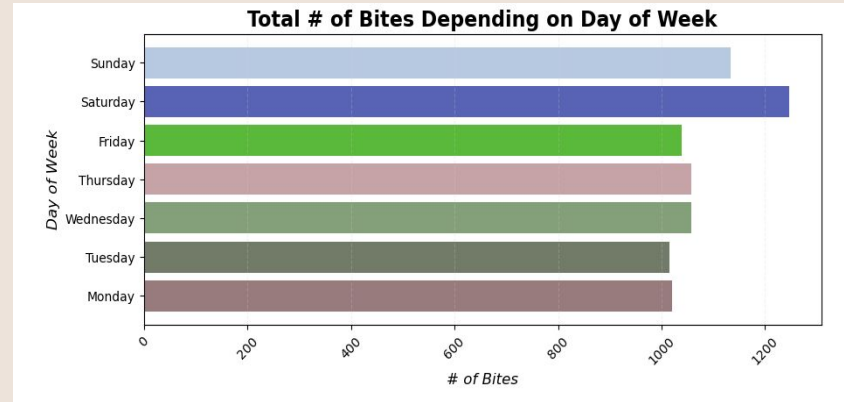
BAR CHART- Seasons

- The length of the bar indicates the number of dog bites reported during that Season
- There are more dog bites reported during Summer, followed by Spring
- dog safety measures and Awareness (Summer)



BAR CHART- Day of Week

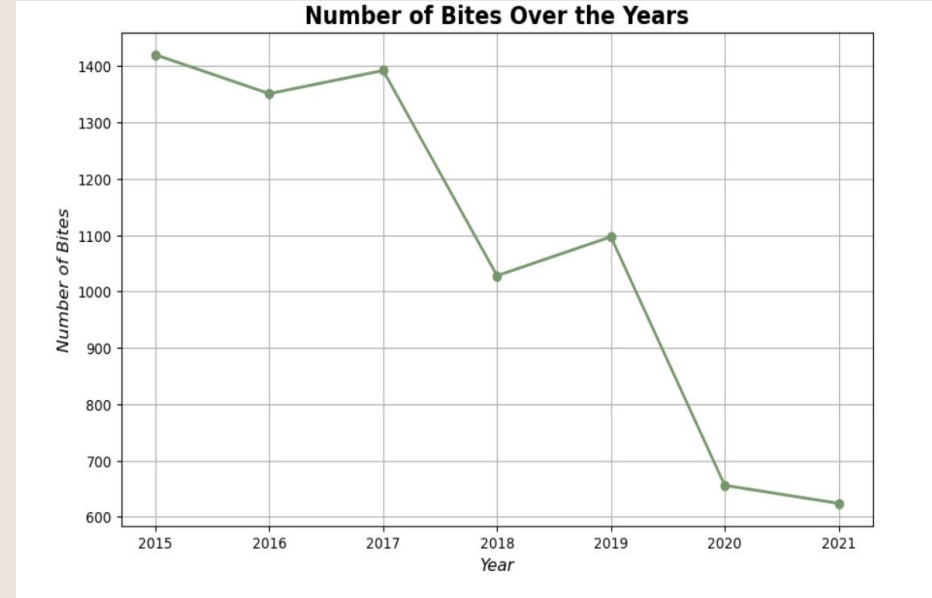
- The length of the bar indicates the number of dog bites reported on that particular day
- There are more dog bites reported on weekends than weekdays
- dog safety measures and Awareness(Weekends)



Number of Bites over the Years

Line Chart:

- Data points(markers) on the chart represents Number of Dog bites each year
- Line - points the trend, which shows the Number of Dog bites decreasing or Increasing over the years (Decreasing)
- Dog bites has been generally dropping since 2017-2021
- The year 2021 has lowest Dog bites, because of the Covid impact (2019-2021) people spend more time with their dogs and social distancing, which reduces the number of dog bites, hopefully the trend continues to decrease (current data not available in our dataset)



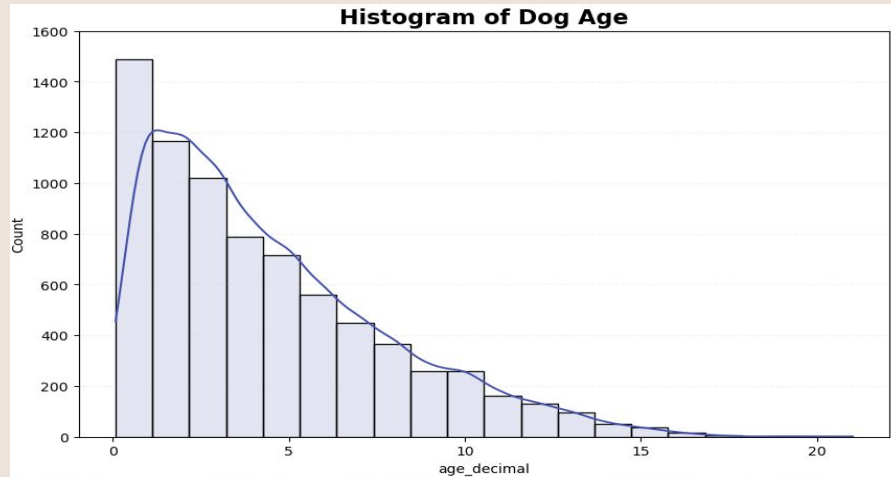
Age

HISTOGRAM

- Histogram shape of the data distribution
- Vertical bite count (type) vs Horizontal (age)
- Younger dogs vs Older dogs bite count
- right skewed; shows mean(avg) > median, positive skew, peak (mode) is on the left
- Manually entered age data; how accurate

Violinplot

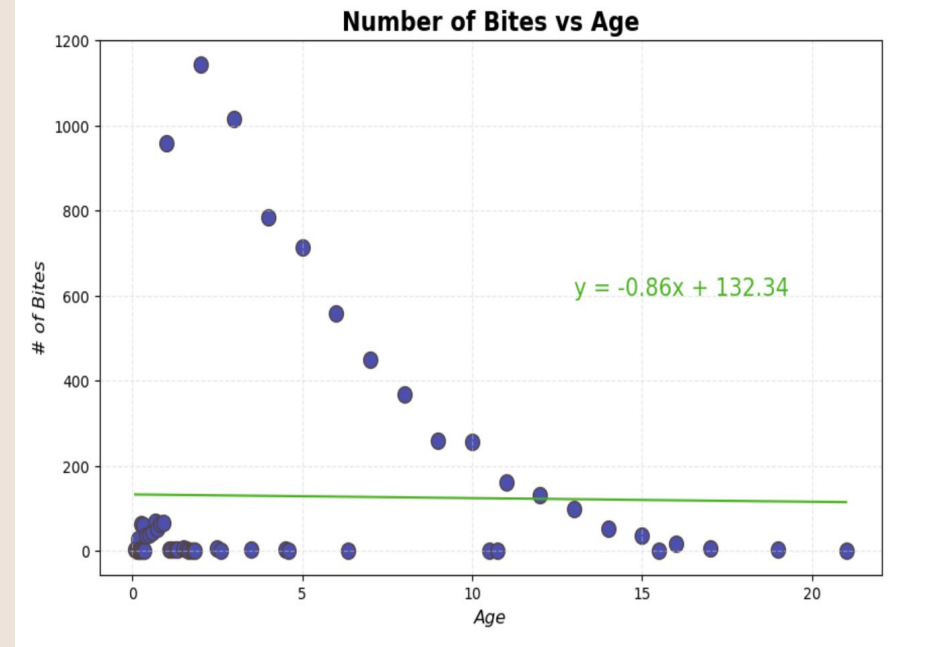
- Density Curve
- Min, 1st and 3rd quartile ranges, median, max
- Useful for bimodal data visualization; this is not the case for this data which is normal
- Younger dogs vs Older dogs bite count
- Oldest dog Dalmation (21) least likely
- Skewed



Regression Number of Bites VS Age

Scatter Plot with Regression

- Correlation between age and # of Bites
- The r-squared is: 0.00032908540749924215
- Regression explains very little about the variation. Why? Similar Age? Young dogs?
- Is Age regression useful in many situations where the large population of data contains a similar age? Can it be ignored?
- Can regression be improved by removing outliers?



Call to Action

- We should add regulations around pit bulls and leashes
- Dogs under the age of 5 need to be on leashes
- Mandatory dog classes once they reach 10 weeks
 - You typically adopt dogs when they turn 8 weeks
 - Most receptive during this time
- We recommend dog owners take their dog to the vet before the summer to check for any allergies or other symptoms



Data Analytics Data Biases

Data Bias

- Data is limited to only New York City and its boroughs
- We need more data to understand the dog population in New York. Are there more young dogs which could skew the data? Does this dataset have proper data across age? Are elderly dogs represented properly?
- What is the definition of a dog bite?
- Dataset had many manual entry errors, especially with Age and Breed columns as well as null values
- Project Time constraints restricted efforts to clean all breed data which resulted in removal of large portions of data
- Spay/Neuter was not verified data and therefore may contain false information
- Age of person bitten could also be beneficial to tell a better story

Future Work

- More data cleaning is needed under breed column.
- Addressing data bias (More Years, New York only, Manually entered incorrect data for breed
- Review a larger dataset that expands to the U.S. and not just NYC with more consistent data
- Borough
 - This data set is limited to these places Brooklyn, Bronx, Staten Island, Queens, Manhattan from New York in the future we can collect data from more places to compare and make a good decision.
- Gender
- Dog Bite Incidents dataset has data from 2015-2021, having only 7 years of data, future work would be collecting data for more years and ensure that we can make fair and ethical decision on this dataset.

Sources

"Most Popular Dog Breeds by City in 2021." *American Kennel Club*, 01 March 2024,

<https://www.akc.org/expert-advice/news/most-popular-dog-breeds-city-2021/>

"Dog Bite Incidents Dataset." *Kaggle*, n.d., <https://www.kaggle.com/datasets/michaelbryantds/dog-bite-incidents/data>

"Coolors." *Coolors.co*, n.d.,

<https://coolors.co/visualizer/bdc667-4bd61d-626d58-544343-4f282c-653c40-734e51-737171-8c6d6f-7885e7>

"What's the Right Age to Start Puppy Training?" *Dogsee Chew Blog*, 22 March 2022,

<https://www.dogseechew.in/blog/whats-the-right-age-to-start-puppy-training#:~:text=Having%20your%20pup%20socialise%20and,%2C%20stand%2C%20stay%20and%20come>

"Microsoft Template Designs" *Microsoft*

<https://create.microsoft.com/en-us/template/blob-design-62339b2c-5cf2-47ff-90b5-2207f1f6a677>

Q&A