DOG BITING PROJECT PROPOSAL

Project 1 – Group 13

OVERVIEW

1. Project background and description

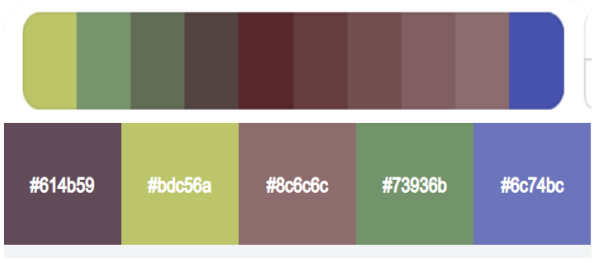
- *This Project is to accomplish demonstrating data cleaning, analysis and data presentation skills which will show the team has grasped the concepts provided the past 6 weeks in the UW Data Analytics Certification Course. We chose the Dog Biting dataset because we found our common interest of dogs and all own dogs. We are hoping to gain a better understanding of our dogs regarding what increases their chances of biting as well as educate others.*

2. Project scope

- *The Scope is limited to analysis on a CSV dataset with no more than 10MB of data with only two groups in the class allowed to use the same dataset. A minimum of three research questions will be answered including inspiration as we build upon previous topics. The dataset can come from any source, however, Kaggle.com has a wealth of data and previous analysis. Five visualizations will be included, extra visualizations out of scope, yet will be considered. The visualizations will include Bar charts, line graphs, pie charts (donut) as well as other useful visualizations to accommodate proper presentation of the data. A color pallet will be included for consistency across the graphs. Roles and Responsibilities will be initially defined and presented in the project proposal.*

3. High-level requirements

The new system must include the following:

- Project Proposal
- Github Link to Project
    - https://github.com/mooreforlesss/project_1_group_13
- Dataset and Inspiration Links
    - Kaggle page for Dataset information
      https://www.kaggle.com/datasets/michaelbryantds/dog-bite-incidents/data
    - Kaggle page for inspiration – EDA
      https://www.kaggle.com/code/yeonseokcho/eda-for-dog-bite-incidents
- Three research questions
- Regression Analysis
- A minimum of 5 Visualizations
- Color Pallet Consistency across the chart
    - 

- Analysis and Presentation of the data

4. Deliverables

- Project Proposal
- Data Cleaning, merging and preparation for further presentations of the data
- The Most Likely Dog Breed to Bite
- How Age impacts dog biting. Is an old dog more likely to bite. Age is defined by 10 years or older, 5 – 9 years, under 5 years.
- Which Gender is most likely to bite
- Review the inspiration, EDA for Dog Bite Incidents as we build upon it and do the analysis
- Bar chart for breed, spay/neuter donut chart (pie chart), line graph bite per year (possibly use months for seasonal biting), Does a warmer Season impact dog biting?
- Bar chart by Zip, possibly Borough as well
- Regression to include number of bites for future per year

- Color Pallet Consistency across the chart
- Analysis and Presentation of the data

5. Work Responsibilities

Alex, Thripura, Joy will prepare the project proposal, Data cleaning will be done as a group, Alex, Joy and Thripura will each answer one question and prepare two data visualizations per person, Marin will work on the regression. Alex created the group git hub project, the team will share responsibilities of organizing the daily standups, Joy will do the initial work on the project proposal with review and updates provided by the team. The team will share responsibilities in doing quality reviews for each other, Alex will merge the project data and other general administrative work for the git hub project. Thripura can manage the communication for the daily standups.

Assistance from TAs as needed, Guidance from professor as needed

6. Affected business processes or systems

This is a standalone system with no other impact to other business processes or systems.

7. Specific exclusions from scope

Scope exclusions will be defined throughout the project as well as project risk and mitigation

8. Implementation plan

Implantation and Project Presentation will be defined after work has been completed. Each person will be responsible for their own works presentation, not to exceed 20 minutes total leaving approximately 5 minutes per person for the presentation.

9. High-level timeline/schedule

- *Project Proposal due by noon 05/30/2024*
- *Data Cleaning complete by 10PM CDT 06/03/2024*
- *Research, Visualizations, Regression Complete by noon 06/06/2024*
- *Presentation Complete by 10PM CDT 6/08/2024*
- *Presentation 5:30PM CDT 06/10/2024*
- *Final Write-up 5:30PM CDT 06/17/24*

APPROVAL AND AUTHORITY TO PROCEED

We approve the project as described above, and authorize the team to proceed.

Name	Title	Date

Approved By	Date	**Approved By**	Date