

# GEOSPATIAL DATA ANALYSIS ON SQUIRRELS IN NYC'S CENTRAL PARK

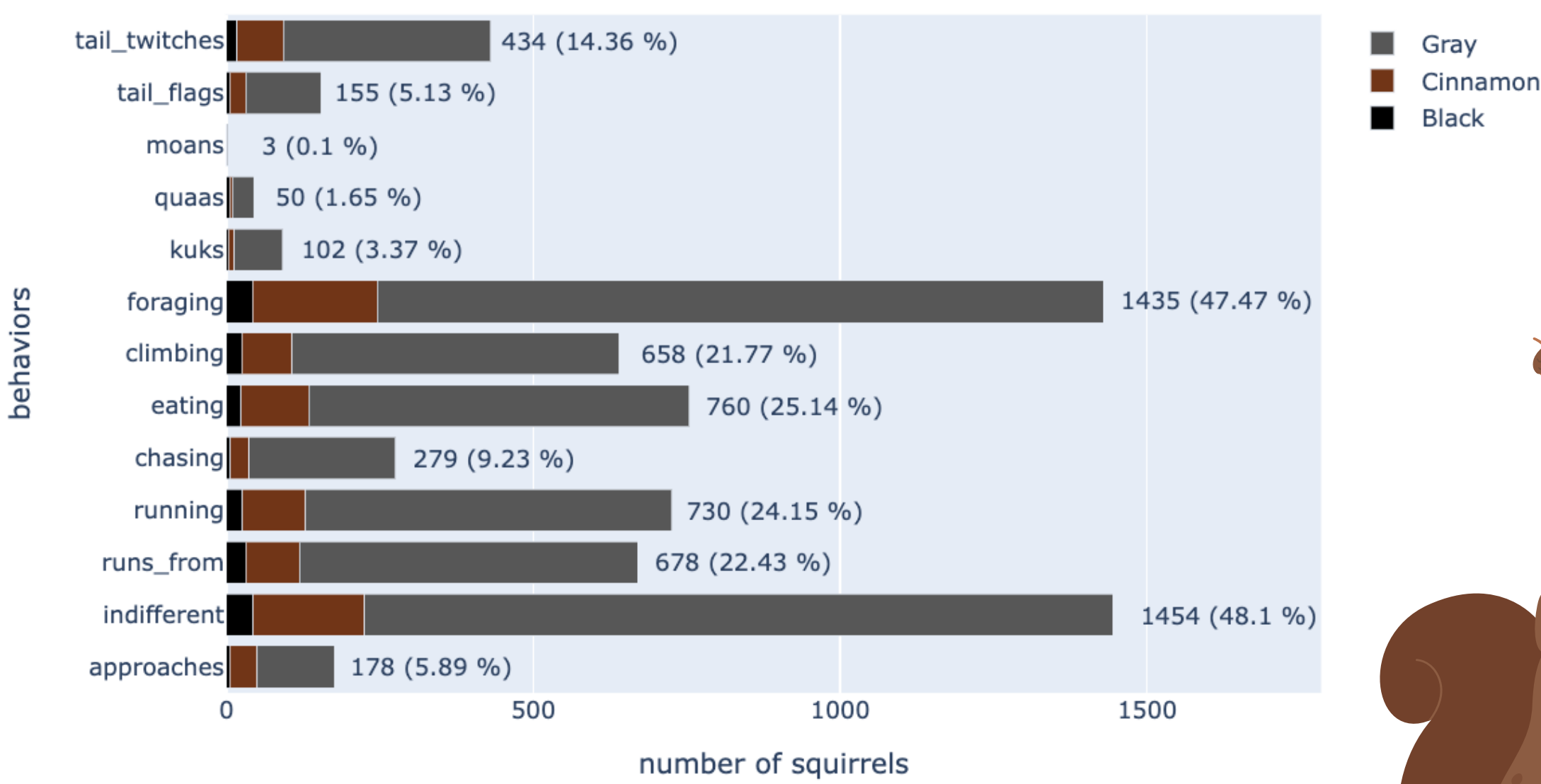
## Introduction

The 2018 Central Park Squirrel Census collected data about squirrels' fur colors, behaviors, and geographic coordinates. Using this dataset, we sought to find out how different aspects of Central Park (bodies of water, forests, buildings, and paved pathways) affect squirrels' behaviors (such as eating or foraging).

## EDA

We explored the fur colors and behaviors of the squirrels and found that there were not enough data for black squirrels and squirrels' communication behaviors like kuks, quaas, and moans.

Primary Fur Colors and Behaviors



## Key Findings

- 18 significant combinations out of 105 total combinations (~17.1%) of behavior and two distinct features
- 8 out of the 18 significant combinations (~44.4%) had "indifferent" as the behavior
- top 3 features of the significant combinations are: woods (~55.6%), grass (50%), and pedestrian (~44.4%)
- "foraging" behavior did not have any significant results
- combinations with the "indifferent" behavior had the most significant results (i.e., the lowest p-values)

## Process

1. Cleaned the raw data by converting to correct types and transforming coordinates to projection systems
2. Obtained data on Central Park features from the OpenStreetMap database in geojson format and clipped features only inside the park
3. Performed buffer analysis to create a dataframe of a squirrel's geographic relationship with Central Park features
4. Conducted permutation tests for every combination of a behavior and two buffers using simulation based on our null hypothesis:

*There is no difference between the proportion of squirrels exhibiting a specific behavior near one feature of Central Park than the proportion of squirrels exhibiting the same behavior near a distinct, different feature.*

## Future

- Predict squirrel behavior based on location using ML
- Fix the problem of squirrels being near multiple features by choosing the "closest" feature
- Conduct a one-sided permutation test to determine which feature has a higher proportion of squirrels doing a certain behavior

P-Value Heatmap of Permutation Tests



## Implications

- General application to studies on animal behaviors in the urban landscape
- Insights for city planners on how human-made environments affect the natural world living with us

Squirrel Distribution in Central Park's Features



For interactive maps and more details, check out our [website!](#)



(Desktop recommended for best user experience)

## Key

- building
- garden
- grass
- pedestrian
- water
- woods
- squirrel