

A wide-angle photograph of a large stadium, likely an NFL arena, during a winter storm. The field is covered in a thick layer of snow. Two green tractors with front loaders are actively clearing the snow from the field. Numerous people, some in dark clothing and others in bright yellow safety vests, are scattered across the field and the sidelines. The stadium's seating areas are visible in the background, and the bright lights of the stadium are on, creating a hazy atmosphere. In the upper right corner, there are several logos, including the NFL shield and the Georgia Tech 'G' logo.

Can NFL game winners be predicted with
machine learning based on weather?

Jacob Garrison
Ike Hannah
Frank White
Aaron Wolfe

Selected Topic

- Machine learning model to predict outcomes of NFL games based on weather data.
- Includes use of data measurements, such as:
 - Teams playing
 - Winner/loser
 - Location of stadium
 - Susceptibility of stadium to elements
 - Temperature
 - Humidity
 - Wind speed



Reason topic was selected



- Collective interest among team members
- Large amounts of data available
- Several variables to measure
- Able to assess validity using historical data

Description of the source of data

- Kaggle
- <https://www.kaggle.com/datasets/tobycrabtree/nfl-scores-and-betting-data>
- NFL scores from 1966 to 2022
- Prior to cleaning, original dataset provided information on 13,232 NFL games

kaggle

Questions the team hopes to answer with the data

- Does weather affect the outcome of a given NFL game?
- If so, what effect does it have?
- Are the machine learning findings consistent with predictions from historical data?



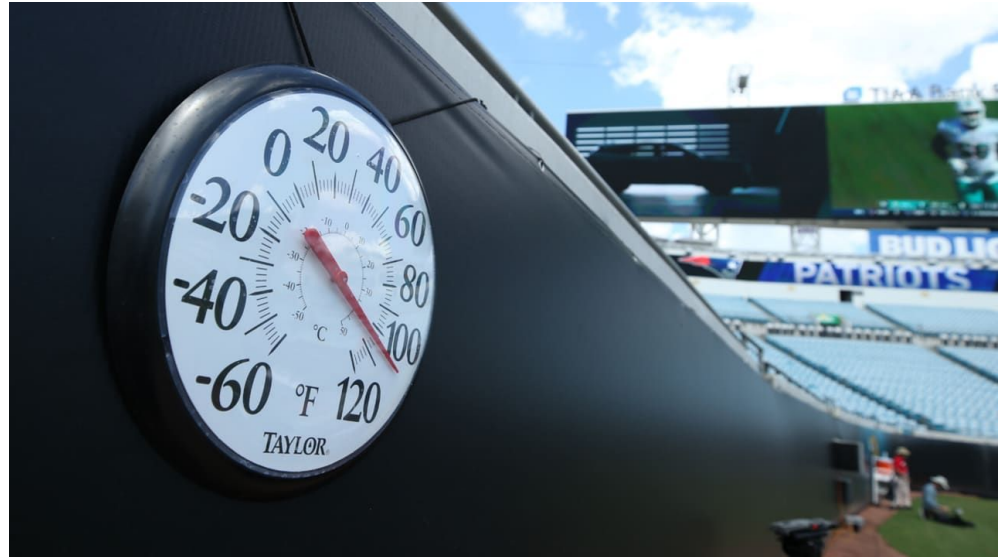
Description of the data exploration phase of the project

- Dropping columns/excluding data
- Assigning new values to domes/covered stadiums
- Rain-or-not: dealing with vague variables
- Eliminating null values
- Creating a key to combine tables
- Verified matching variables between data sets (stadium names)



Description of the analysis phase of the project

- Linear regression
- Dashboard visualizations
- Relationships we expected to find but didn't



Technologies, languages, tools, and algorithms used throughout the project

- Jupyter Notebook
- Postgres
- Python
- Tableau

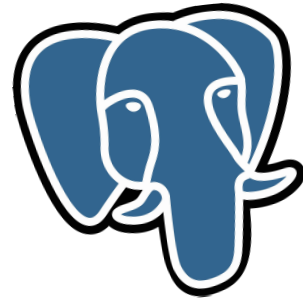
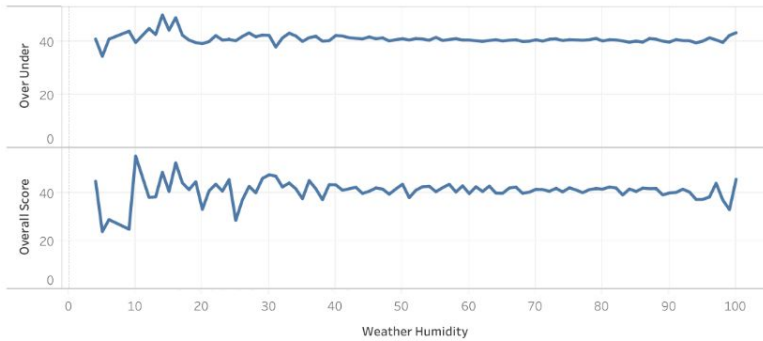
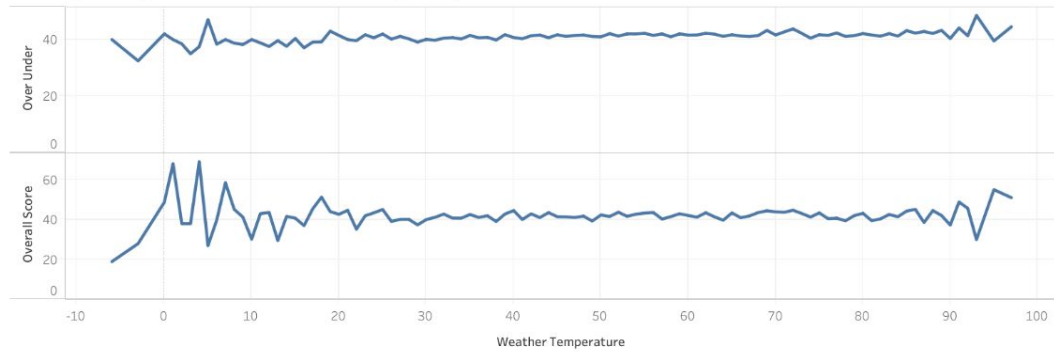


Tableau Dashboard

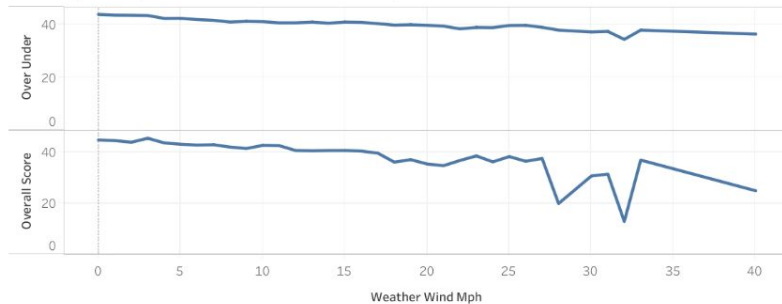
Average Over/Under and overall score by Humidity



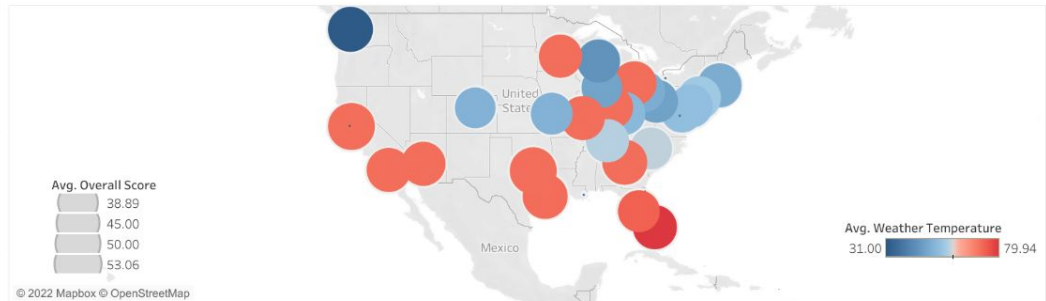
Average Over/Under and overall score by Temperature



Average Over/Under and overall score by Wind MPH



NFL Stadiums - Avg overall score by Temperature



Thank you!