Michael Zaychikov, Ethan Ha, Rocky Tran, Sri Paruchuri

Dataset gameplan:

Seattle and Pittsburgh comparison of results data

About: This dataset is a set of arena distances from Seattle and Pittsburgh as well as their results from the 2022-2023 NHL season. The purpose of the dataset was to see where either Pittsburgh or Seattle played, if they won, and how far they had to travel in miles to the arena where the game was.

Data Creation Range: October 2022- Present

Created By:

NHL scorekeepers, distance columns added on by Michael Zaychikov (used google maps as a source for the distances)

Source:

https://www.google.com/maps/dir/Climate+Pledge+Arena,+1st+Avenue+North,+Seattle,+WA/PPG+Paints+Arena,+1001+Fifth+Ave,+Pittsburgh,+PA+15219/@40.4393092,-79.9998526,15z/data=!3m1!5s0x8834f15e8b78d0fb:0xd182066140b3ea85!4m13!4m12!1m5!1m1!1s0x5490154471be8ed3:0xde04af6753ca2e27!2m2!1d-122.3540714!2d47.6220713!1m5!1m1!1s0x8834f15e8eaaaaab:0x1c1b1fdbf269c5fe!2m2!1d-79.9895529!2d40.4393097?entry=ttu, https://fixturedownload.com/results/nhl-2022/pittsburgh-penguins, https://fixturedownload.com/results/nhl-2022/seattle-kraken

Use Cases: Real world applications of the dataset

1. How many wins/losses did Pittsburgh and Seattle get in the 2022-2023 season
2. How much less travel does Pittsburgh do compared to Seattle
3. Is there a pattern of which teams lose to which teams
4. Understanding schedules based on distances between teams
5. Determines if travel the only factor between winning and losing

Gameplan:

Bias:

The data is simply raw numbers of scores and distance therefore there is no bias

Misrepresentation:

There was an assumption that for every game the team traveled from their home arena. This may not be true if they played a team close to where the played the game before. This affects total miles traveled per year but not flight distance from the home arena.

Period one
- Collect data for both teams
- Verify that the results are legitimate

Period two
- Understand misrepresentations
- Interpret how this affects the data

Period three
- Compare results of East and West games for both teams
- Compare the travel for both teams