Project 2 Report

John Silberstein and Kristina Mulholland

Description:

For this project we will be examining running and weather data. Performance will be measured by Average Heart Rate and minutes per mile. How does weather affect performance?

Original Data Sources and Formatting:

John's Strava Data
CSV file
strava_data.df

https://openweathermap.org/api
Weather API: Open Call 3.0 (subscription, 1000 calls free per day)
JSON
running_weather_data[]

Data Cleaning and Transformation:

- strava_data.df
  - Loaded csv dataframe and filtered to show only columns with desired comparison data
  - Checked data types of each column
  - Deleted rows with missing data
  - Format "Activity Date" into format to use alongside Open Weather Map API call

- running_weather_data[]
  - Converted Philadelphia latitude and longitude from degrees to decimal to fit API formatting using: https://www.fcc.gov/media/radio/dms-decimal (ensuring that longitude includes "-")
  - Make the API call to Open Weather Map One Call for weather by date and location (temperature, humidity, uv index, wind speed)
  - Format the JSON into a dataframe

Final Database Structure:

- Load the two tables into a SQL relational database
- Use the primary key of datetime for this database

Ideas for Future Improvement:

- Using multiple Strava data reports to increase the scope of performance to weather analysis
- Including elevation in data for more accurate understanding of pacing
- Identifying additional datasets that provide location and weather