

Project 2

John Silberstein and Kristina Mulholland

Description:

For this project examined running data and corresponding weather data with the intention of preparing the data for Performance analysis. Performance is measured by two primary metrics, Average Heart Rate and minutes per mile. The question to be answered is how does weather affect performance?

Original Data Sources and Formatting:

Garmin Activity Data (John Silberstein's) downloaded in a CSV file.

Weather data downloaded using an API from <https://openweathermap.org/api>.

Weather API: Open Call 3.0 (subscription, 1000 calls free per day) Data received in JSON format.

Data Cleaning and Transformation:

- running data was initially loaded into a data frame called, `garmin_data_df`. The initial data included several different types of activities, running, biking, walking and swimming. None running activities were removed as well as running activities with missing data. Also, the working data frame had extraneous columns removed so that only data we had interest in using was included.
- The activity data was going to be used as the key index, but needed to be converted to a timestamp that Open Weather Map could use.

- All of the activities took place in Philadelphia, so those GPS coordinates were used in the API calls. The actual coordinates were determined using the <https://www.fcc.gov/media/radio/dms-decimal>. This data also had to be cleaned as it included a decimal point that was extraneous.
- Due to the limitations of the weather resource, we limited the API call to 10 rows of data, `running_test_df`.

`clean_running_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.df` ○ 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, 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