Health Data Tracking Web App Documentation

Overview

The Health Data Tracking Web App is a comprehensive health monitoring and analysis application designed for individuals to manage and gain insights into their health-related data. The app provides a user-friendly interface for tracking information sourced from various devices, as well as manual entries.

The primary focus of the program is to empower users in understanding and visualizing their health metrics over time. It facilitates comparisons, statistical analysis, and correlation studies between different health variables, aiding users in making informed decisions about their well-being.

OS:

The application is platform-independent and can be run on any operating system supporting Python. It has been tested on Windows and Linux environments.

Language:

The app is developed using the Python programming language.

Libraries and External Components:

- Bokeh Library: Used for creating interactive and dynamic visualizations in the web app.
- AirTable Integration: The app interfaces with AirTable to fetch and store user health data.
- NumPy and SciPy: Utilized for numerical and scientific computations in data analysis.
- Others: Various standard Python libraries and tools for web development.

Development Environment:

The development environment consists of a standard Python development setup, along with the necessary libraries installed. Developers can use preferred code editors or integrated development environments (IDEs) for coding and testing.

Data Source Integration:

The app integrates with AirTable to collect user data. Users store information from multiple sources, such as smartwatches and manual entries.

Configuration:

To use the app, users need to configure their AirTable credentials in the locals/config.py file.

Features of the App

The app offers several features to help users track their health data. Each feature is represented by a separate tab of the program. So far, the app utilizes 6 tabs:

1. Comparison Tab

The 'Comparison' tab is a powerful feature allowing users to compare two health variables with various settings:

1.1 Variable selection

1.1.1 Category: Represents the data source.

Users can choose the category of the variable, representing the source or nature of the data. This helps in organizing and identifying the type of information associated with the selected variable.

1.1.2 Variable Name:

Selecting the specific name of the variable within the chosen category is essential. This corresponds to a column name in the AirTable record and defines the dataset to be analyzed.

1.2 Filtering Options

1.2.1 Weighted Average:

For each variable, users can apply different weighted average filtering settings. The available options include 'none,' 'Gauss,' 'PastGauss,' or 'FutureGauss.' This feature allows users to smooth out fluctuations and focus on underlying trends in the data.

1.2.2 Sigma Value:

The sigma value, by default set to 0.2, can be adjusted by the user. This parameter influences the extent of the weighting applied during filtering. Users have the flexibility to fine-tune this value based on their data characteristics.

1.3 Shifting Settings

1.3.1 Shifting Direction:

The shifting setting determines the directional relationship between the two variables. By default, it is set to 'no shift,' meaning a direct comparison over time. Users can also choose 'variable2 -> variable1' or 'variable1 -> variable2' to analyze how changes in one variable relate to the other.

1.4 Plot Visualization

Corresponding to the settings set by the user, the 'Comparison Tab' generates several insightful plots to aid in the analysis of paired data.

Time Series Plot:

This plot visualizes the time series data for the two selected variables over time. It allows users to observe trends, patterns, and fluctuations in the raw and filtered data.

Filtered Line Plots:

Filtered Line Plots provide a detailed view of the filtered values for each variable separately. This helps users understand the impact of the applied filters on individual variables.

Correlation Plot:

The Correlation Plot is a scatter plot of one variable against another, displaying statistical information such as mean, SEM, and correlation coefficients. It aids in understanding the relationship between the two variables.

Cross-correlation Plot:

The Cross-correlation Plot is a line plot showing cross-correlation coefficients over time. It visualizes how changes in one variable relate to changes in the other with a time lag.

◊ Each plot serves a distinct purpose in analyzing and comparing paired data, providing insights into relationships, trends, and statistical significance. The interactive nature of the plots allows users to explore data by selecting specific ranges or categories - all plots in the 'Comparison' tab are interactive. Users can explore the data further by selecting specific ranges or categories of interest, providing a dynamic and user-friendly experience.

2. Event-Based Analysis Tab

The 'Event-Based Analysis' tab in the Health Data Tracking Web App is designed to enable users to comprehensively explore the relationship between two health variables in the context of events. Users can choose and analyze these variables with the same settings as in the 'Comparison' tab.

2.1 Plot Visualization

Corresponding to the settings set by the user, the ‘Event-Based Analysis' tab generates 2 different plots to aid in the analysis of paired data.

Time Series Plot:

Type: Line Plot

This plot visualizes the event-triggered average over time, providing a dynamic view of how variable 1 and variable 2 change around specific events.

• Time series plot focuses on the continuous time series aspect, offering insights into the nuanced relationship between the two variables during events.

Grouped Bar Plot:

Type: Grouped Bar Plot

This plot visualizes the mean values of variable 1 and variable 2 across different time windows, presenting a snapshot of their behavior before and after events. It also utilizes grouped bars to clearly depict mean values, allowing users to compare the variables in distinct time intervals.
The plot highlights significant differences between the variables using p-values and distinctive colors for enhanced clarity. It also connects corresponding points with lines, emphasizing the temporal evolution of mean differences between variable 1 and variable 2.

• Grouped bar plot provides a summary view that emphasizes mean disparities for discrete time windows and their statistical significance.

In summary: The combination of these plots equips users with a great toolkit for gaining insights into the relationship between the chosen variables over time, with a specific focus on events and varied time windows.

By concurrently exploring continuous trends and discrete changes in the data, users can deepen their understanding of health variables and their interplay in the context of specific events.

3. Correlations Tab

The 'Correlations' tab within the app offers an interactive dashboard dedicated to correlation analysis. This feature enables users to explore relationships between variables, apply filters, and seamlessly switch between different views, providing a comprehensive toolkit for a thorough correlation exploration.

The tab can be configured through various settings:

3.1. User Actions:

3.1.1 User Can:

- Select the start and end dates for the correlation analysis.
- Choose a category of variables from the dropdown.
- Select a variable within the chosen category from the dropdown.
- Click the "Recalculate" button to initiate the correlation analysis.

3.1.2 Data Viewing Configuration:

A user can also modify how the data is displayed:

- View All Correlations: Default view displaying all correlations.
- View Picked Correlations: Pick specific correlations.
- Hide/Show Correlations: Choose to display or hide correlations on ignore or black lists.
- Show Shift Options: Choose the shift type to display in the correlation results (variable 1 -> variable 2, No Shift, variable 2 -> variable 1).

3.1.3 Filtering/Selection for detailed information:

- Filter by p-value: Set a maximum p-value threshold to filter correlations.
- Users can also select variables for detailed inspection.

3.1.4 List Management:

- Add to Ignore List: Add selected correlations to the ignore list to exclude selected correlations from future analyses.
- Add to Black List: Add selected correlations to the black list to exclude selected correlations from future analyses, marking them as significant issues.

After configuring the settings, the app displays the DataTable-like widget with all the generated information.

4. Relationships Tab

The 'Relationships' tab enables users to manage relationships between variables efficiently. Users can navigate, analyze, and control the exploration of correlations and relationships within the dataset. This code integrates these functionalities into an interactive web-based application, creating a user-friendly environment for data exploration and analysis.

4.1 User Contributions:

4.1.1 View Ignored and Blacklisted Relationships:

Users can display relationships marked as ignored or blacklisted.

4.1.2 Relationships Management:

Users can select one or more variables and remove the corresponding relationships from the ignored list or the blacklist.

4.2 Relationships Table:

According to the way user configures the settings, i.e. what variables are added/removed from the blacklist and the ignored list, the table is displayed as follows:

Variable 1	Variable 2	IgnoreList	BlackList

◊ To aid the process of configuring the relationships, the console output displays messages like "Starting filtering" and "Finished filtering" during the filtering process.

5. Blood Tests Tab

The 'BloodTests' tab is designed to handle and analyze blood test data provided by the user. This tab plays a crucial role in offering a comprehensive visual representation of various biomarkers over time, providing insights into the user's health progression. The detailed functionalities, visualization strategies, and user-friendly features make the 'BloodTests' tab an essential component for users keen on tracking and understanding their health metrics.

Features of the 'BloodTests' tab:

5.1 Biomarker Table Representation:

5.1.1 Columns:

Represents a structured compilation of data pertaining to each biomarker derived from blood test results, with the first column depicting corresponding dates.

Each non-empty entry within the table is thoughtfully color-coded, providing an immediate visual indication of the biomarker's health status.

5.1.2 Color Codes:

Red: Signifies biomarkers that fall outside the specified range.
Yellow: Indicates biomarkers within the normal range.
Green: Represents biomarkers performing optimally within the user-defined optimal range.

5.2 Dropdown Widget for Different Views

Users have the flexibility to select different views corresponding to various blood test records effortlessly.

Functionality:

The 'BloodTests' tab utilizes metadata, using user-provided information to dynamically categorize biomarker values into optimal, normal, and out-of-range classifications.

The color-coding system adapts dynamically based on the user-defined optimal and normal ranges, providing a personalized and accurate visual representation.

Code Structure:

The relevant code creates an instance of the 'BloodTests' class, extending the 'AnalysisPanel' class. It includes functions to update widgets, data, and plots based on user interactions. The 'classify_range' and 'color_mapper' functions play a crucial role in dynamically categorizing biomarker values and mapping them to appropriate colors for display.

6. Blood Tests Correlation Tab

The 'BloodTests Correlations' tab in the app is a crucial feature designed to demonstrate relationships between blood test results and user-specific lifestyle metrics. This tab stands out by offering a detailed exploration of how a secondary variable impacts blood test biomarkers over extended periods. Unlike the 'Comparison' tab, which examines two variables at specific time points, the 'BloodTests Correlations' tab employs a unique approach of segmenting secondary variable data based on blood test dates and calculating averages for corresponding time frames. This strategy provides users with valuable insights into the dynamic interplay between lifestyle factors and blood test results.

6.1 Functionality:

6.1.1 Variable Segmentation:

The 'BloodTests Correlations' tab slices the secondary variable's data into segments based on blood test dates.

Users can adjust the segmentation period to explore correlations over specific time frames, capturing both short-term and long-term effects.

6.1.2 Interactive Widgets:

The tab features dynamic widgets enabling users to choose the primary and secondary variables for analysis, such as blood test biomarkers and lifestyle metrics.

A segmentation slider provides users with control over the analysis, allowing them to focus on specific time intervals.

6.2 Plot Visualization

6.2.1 Time Series Plot:

The first plot serves as a time series visualization, illustrating the trends of both variables over time. Color-coded circles represent each variable, providing an immediate visual correlation between the primary and secondary variables.

6.2.2 Correlation Plot:

The second plot delves into detailed statistical correlation analysis between the selected variables. Data points are plotted, and a trend line, along with shaded areas indicating standard error of the mean (SEM), aids in interpreting the strength and direction of the correlation.

A Slope annotation on the correlation plot visually illustrates the correlation line, offering an intuitive understanding of the relationship direction.

6.3 Statistical Information:

The tab incorporates statistical analysis, including correlation coefficients and p-values, offering users quantitative insights into the strength and significance of observed correlations.

6.4 Implementation

The code creates a dynamic interface with widgets for user interaction, facilitating the selection of variables and adjustment of segmentation parameters. The tab employs a segmentation and averaging strategy to align the time frames of the two variables for meaningful correlation analysis.

Both the time series plot and the correlation plot are designed to provide users with a comprehensive view of the data.