Apple Watch Exploratory Data Analysis

Raul Eulogio

RforDataScience Community Talk

About

With the advent of wearable technologies, daily health tracking has become a popular trend for users to gain insight on day to day aspects.

These tools have been important because the data collection requires little overhead; vasts amounts of data are being collected without too much interaction between the user and the technology.

Utilizing this very personalized data with other outside sources could be beneficial to...

- Understanding a person’s daily exercise routines
- Being able understand trends across different time frame measurements
- Understanding the relationship between this data and other available data

Packages used in this project

- tidyverse
- lubridate
- here
- reticulate
- Rmarkdown
- packrat

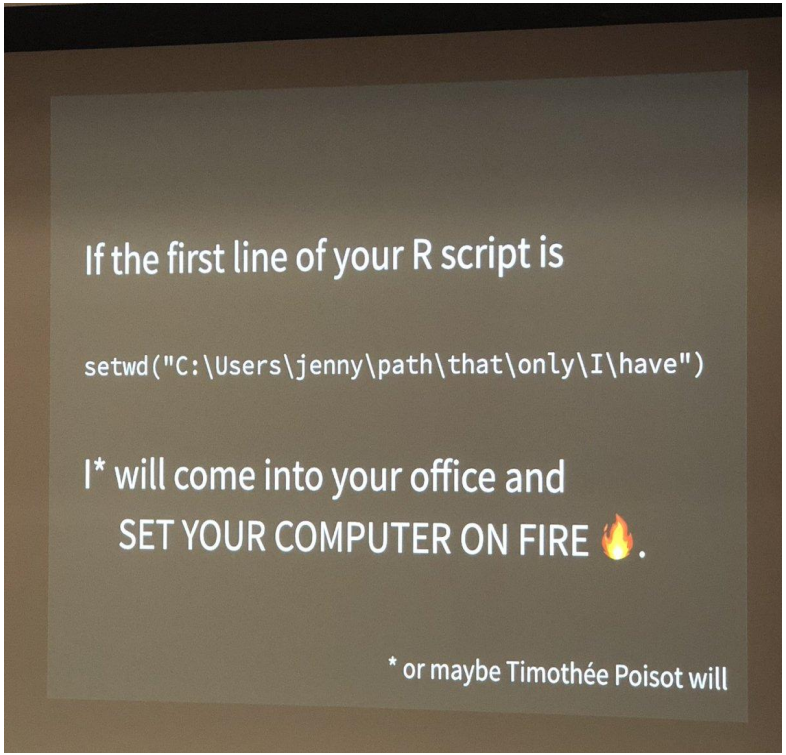
Tidyverse

Majority of packages utilized in this analysis rely on tidyverse

- ggplot2 - Data Visualization (Chapter 3 in r4ds)
- dplyr - Data Manipulation (Chapter 5 in r4ds)
- lubridate - Datetime data Structures(Chapter 16 in r4ds)

here package

- Utilized to help with navigating through the directory tree for your project.
- Effective workflow especially for projects that you want to be reproducible
- Find more information from Jenny Bryan here



Directory Structure

- Inspired and heavily borrowed from Driven Data's Cookie Cutter Data Science Structure

Includes:

- .Rproj file - utilized with here package to ensure cross compatibility for project
- .gitignore file - utilized to ignore certain files in our directory that don't serve a purpose within the public repo
- Packrat directory - uses packrat package to abstract dependencies (Much like virtualenv or anaconda for Python)

Collect Data

- Data was collected on Apple Watch used by friend (kept anonymous).
- Extracted from xml file using Python script by Nicholas J. Radcliffe from Test-Driven Data Analysis
- Focusing only on one aspect of the data collected (Active Energy Burned), many of the files contain similar schema so most of this analysis can be translated to the other aspects. More in near future.

Before we begin

Resources:

- Github repo will be made available at the release of the video
- inertia7 write up
- Any questions or concerns reach out to me at rauleulogio3@gmail.com

Let’s Begin!