

# CSE 550 Software Engineering

*Term Project – Visualizing and analysing Wearable Sensor Data*





## Term Project

- Design and develop an interactive user interface to use visualization techniques and visual analysis to explore multimodal sensor signal data and to tell how and what you can discover interesting information and pattern from the visualization.



# The data

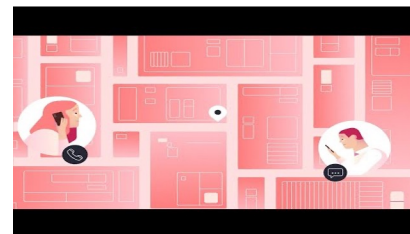
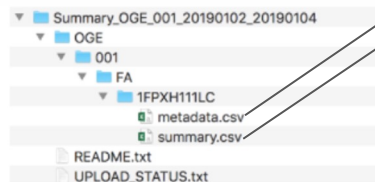
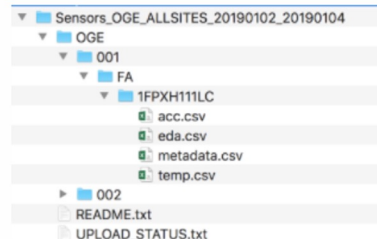
Datetime (UTC)	Timezone (minutes)	Unix Timestamp (UTC)	Acc magnitude avg	Eda avg	Temp avg	Movement intensity	Steps count	Rest
2020-01-17T23:48:00Z	-300	1.5793E+12	1.063262	0.541921	30.155257	0	0	0
2020-01-17T23:49:00Z	-300	1.5793E+12	1.005967	0.539037	29.9799	0	0	0
2020-01-17T23:50:00Z	-300	1.5793E+12	1.045804	0.535254	29.713417	0	0	0
2020-01-17T23:51:00Z	-300	1.5793E+12	1.017389	0.532977	29.416833	0	0	0
2020-01-17T23:52:00Z	-300	1.5793E+12	1.03043	0.532688	29.2752	0	0	0
2020-01-17T23:53:00Z	-300	1.5793E+12	0.995176	0.541594	29.4671	0	0	0
2020-01-17T23:54:00Z	-300	1.5793E+12	1.066957	0.548665	29.84755	0	0	0
2020-01-17T23:55:00Z	-300	1.5793E+12	1.076301	0.547429	30.069133	0	0	0
2020-01-17T23:56:00Z	-300	1.5793E+12	1.507012	0.546359	30.031567	0	0	0

Datetime (UTC)	Timezone	Embrace firmware version	App	App version	Mobile OS	Mobile OS version	GTCS algorithm version
2020-01-17T23:48:28Z	-300	1.3.5602	Mate	2.1.0	iOS	Version	0.6.4724

## Raw Data - Archive Structure

The data will download into an archive with a 3-level hierarchy. It is divided into sites, enrolled subjects, and devices as shown in the examples below:

The left image shows an example of the folder structure for **Sensors Data** of an entire study (e.g. with identifier OGE), named here as **Sensors\_OGE\_ALLSITES\_20190102\_20190104**. On the right is an example of the folder structure for **Summary Data** for one site (e.g. with identifier 001) of the study, named here as **Summary\_OGE\_001\_20190102\_20190104**.



[more details](#)

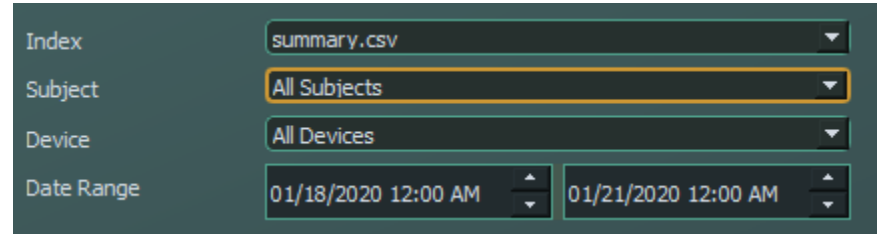


# Basic requirements

1. A Data Loader to load sensor data for visualization and exploration

## Basic requirements

1. A Data Loader to load sensor data for visualization and exploration
  - a. **Allow users to select and visualize data between different participants.**



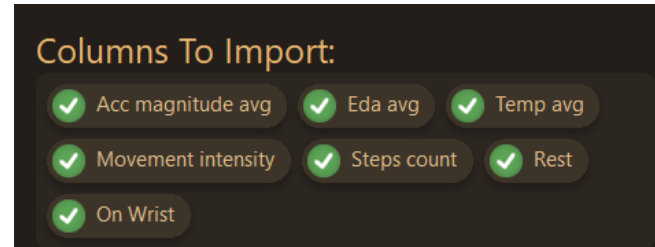
A screenshot of a data loader interface with a dark teal background. It features four filter controls: 'Index' with a dropdown menu showing 'summary.csv'; 'Subject' with a dropdown menu showing 'All Subjects' (highlighted with a yellow border); 'Device' with a dropdown menu showing 'All Devices'; and 'Date Range' with two date pickers showing '01/18/2020 12:00 AM' and '01/21/2020 12:00 AM' respectively.

**example**



## Basic requirements

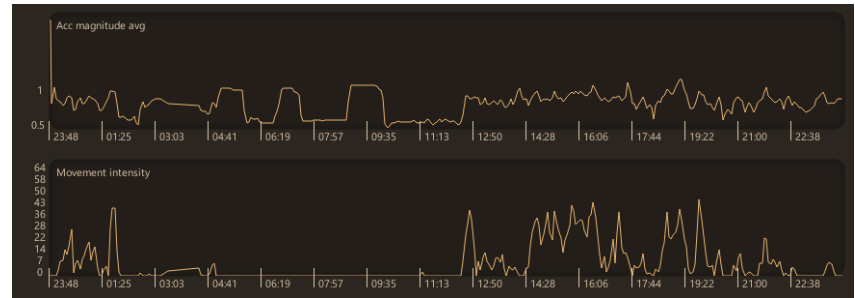
1. A Data Loader to load sensor data for visualization and exploration
  - a. Allow users to select and visualize data between different participants.
  - b. **Allow users to specify the interested data streams to be imported.**



**example**

# Basic requirements

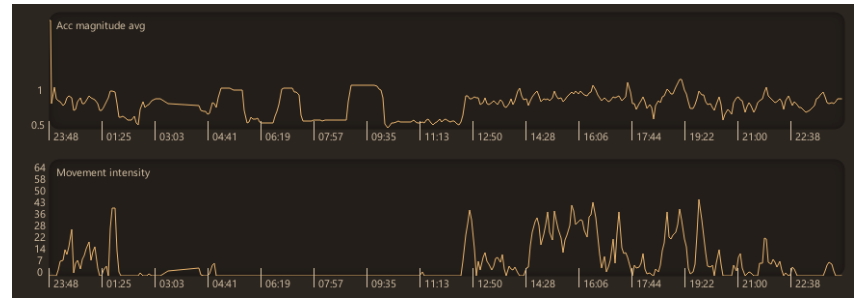
1. A Data Loader to load sensor data for visualization and exploration
  - a. Allow users to select and visualize data between different participants.
  - b. Allow users to specify the interested data streams to be imported.
2. A visual panel to show the multiple time series



**example**

# Basic requirements

1. A Data Loader to load sensor data for visualization and exploration
  - a. Allow users to select and visualize data between different participants.
  - b. Allow users to specify the interested data streams to be imported.
2. A visual panel to show the multiple time series
  - a. Synchronized across series



**example**

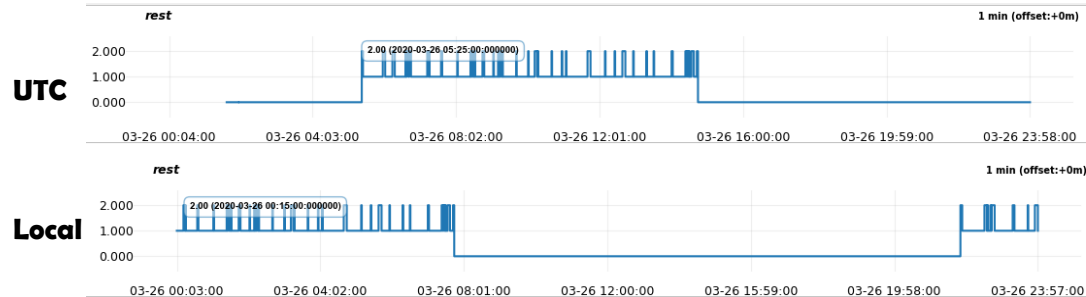


# Basic requirements

1. A Data Loader to load sensor data for visualization and exploration
  - a. Allow users to select and visualize data between different participants.
  - b. Allow users to specify the interested data streams to be imported.
2. A visual panel to show the multiple time series
  - a. Synchronized across series
  - b. Allow to choose the appropriate type of chart

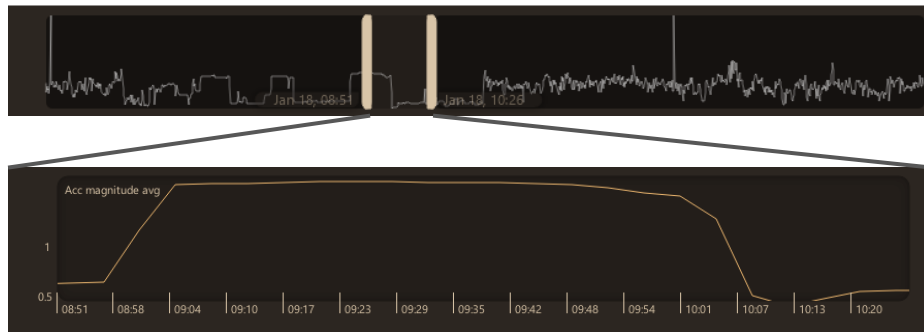


# Basic requirements



1. A Data Loader to load sensor data for visualization and exploration
  - a. Allow users to select and visualize data between different participants.
  - b. Allow users to specify the interested data streams to be imported.
2. A visual panel to show the multiple time series
  - a. Synchronized across series
  - b. Allow to choose the appropriate type of chart
3. Switch between UTC / local time using time zone information
  - a. Coordinated Universal Time or UTC is the primary time standard by which the world regulates clocks and time.
  - b. Local time is the time observed in a specific locality, which can be derived by the UTC and time zone information.

# Basic requirements



**example**

1. A Data Loader to load sensor data for visualization and exploration
  - a. Allow users to select and visualize data between different participants.
  - b. Allow users to specify the interested data streams to be imported.
2. A visual panel to show the multiple time series
  - a. Synchronized across series
  - b. Allow to choose the appropriate type of chart
3. Switch between UTC / local time using time zone information
  - a. Coordinated Universal Time or UTC is the primary time standard by which the world regulates clocks and time.
  - b. Local time is the time observed in a specific locality, which can be derived by the UTC and time zone information.
4. Data navigation
  - a. **A widget required to be designed, e.g. a time box, to explore different regions of data with different granularities.**

# Basic requirements



1. A Data Loader to load sensor data for visualization and exploration
  - a. Allow users to select and visualize data between different participants.
  - b. Allow users to specify the interested data streams to be imported.
2. A visual panel to show the multiple time series
  - a. Synchronized across series
  - b. Allow to choose the appropriate type of chart
3. Switch between UTC / local time using time zone information
  - a. Coordinated Universal Time or UTC is the primary time standard by which the world regulates clocks and time.
  - b. Local time is the time observed in a specific locality, which can be derived by the UTC and time zone information.
4. Data navigation
  - a. A widget required to be designed, e.g. a time box, to explore different regions of data and with different granularities.
  - b. **Aggregation.** Data *can be* gathered and expressed in a summary form for analysis.

# Basic requirements

1. A Data Loader to load sensor data for visualization and exploration
  - a. Allow users to select and visualize data between different participants.
  - b. Allow users to specify the interested data streams to be imported.
2. A visual panel to show the multiple time series
  - a. Synchronized across series
  - b. Allow to choose the appropriate type of chart
3. Switch between UTC / local time using time zone information
  - a. Coordinated Universal Time or UTC is the primary time standard by which the world regulates clocks and time.
  - b. Local time is the time observed in a specific locality, which can be derived by the UTC and time zone information.
4. Data navigation
  - a. A widget required to be designed, e.g. a time box, to explore different regions of data and with different granularities.
  - b. Aggregation. Data *can be* gathered and expressed in a summary form for analysis.
  - c. **Basic statistical analysis.**

Acc magnitude avg	Point count: 1404.00	Mean: 1.03	SD: 0.07	Min: 0.90	25%: 0.97	Median: 1.04	75%: 1.08	max: 1.51
Movement intensity	Point count: 1404.00	Mean: 9.27	SD: 16.85	Min: 0.00	25%: 0.00	Median: 0.00	75%: 14.00	max: 87.00
Eda avg	Point count: 1404.00	Mean: 1.46	SD: 0.63	Min: 0.02	25%: 0.85	Median: 1.68	75%: 1.94	max: 2.62
Steps count	Point count: 1404.00	Mean: 12.20	SD: 19.29	Min: 0.00	25%: 0.00	Median: 0.00	75%: 36.00	max: 92.00
Temp avg	Point count: 1404.00	Mean: 30.50	SD: 1.13	Min: 26.36	25%: 29.75	Median: 30.38	75%: 31.54	max: 32.65
Rest	Point count: 1404.00	Mean: 0.38	SD: 0.58	Min: 0.00	25%: 0.00	Median: 0.00	75%: 1.00	max: 2.00

example



# Basic requirements

1. A Data Loader to load sensor data for visualization and exploration
  - a. Allow users to select and visualize data between different participants.
  - b. Allow users to specify the interested data streams to be imported.
2. A visual panel to show the multiple time series
  - a. Synchronized across series
  - b. Allow to choose the appropriate type of chart
3. Switch between UTC / local time using time zone information
  - a. Coordinated Universal Time or UTC is the primary time standard by which the world regulates clocks and time.
  - b. Local time is the time observed in a specific locality, which can be derived by the UTC and time zone information.
4. Data navigation
  - a. A widget required to be designed, e.g. a time box, to explore different regions of data and with different granularities.
  - b. Aggregation. Data *can be* gathered and expressed in a summary form for analysis.
  - c. Basic statistical analysis.
5. Data query



# Basic requirements

1. A Data Loader to load sensor data for visualization and exploration
  - a. Allow users to select and visualize data between different participants.
  - b. Allow users to specify the interested data streams to be imported.
2. A visual panel to show the multiple time series
  - a. Synchronized across series
  - b. Allow to choose the appropriate type of chart
3. Switch between UTC / local time using time zone information
  - a. Coordinated Universal Time or UTC is the primary time standard by which the world regulates clocks and time.
  - b. Local time is the time observed in a specific locality, which can be derived by the UTC and time zone information.
4. Data navigation
  - a. A widget required to be designed, e.g. a time box, to explore different regions of data and with different granularities.
  - b. Aggregation. Data *can be* gathered and expressed in a summary form for analysis.
  - c. Basic statistical analysis.
5. **Data query**
  - a. Allow users to filter and visualize data under certain conditions.



# Basic requirements

1. A Data Loader to load sensor data for visualization and exploration
  - a. Allow users to select and visualize data between different participants.
  - b. Allow users to specify the interested data streams to be imported.
2. A visual panel to show the multiple time series
  - a. Synchronized across series
  - b. Allow to choose the appropriate type of chart
3. Switch between UTC / local time using time zone information
  - a. Coordinated Universal Time or UTC is the primary time standard by which the world regulates clocks and time.
  - b. Local time is the time observed in a specific locality, which can be derived by the UTC and time zone information.
4. Data navigation
  - a. A widget required to be designed, e.g. a time box, to explore different regions of data and with different granularities.
  - b. Aggregation. Data *can be* gathered and expressed in a summary form for analysis.
  - c. Basic statistical analysis.
5. Data query
  - a. Allow users to filter and visualize data under certain conditions.
6. UI design





# Principles of User Interface Design

- **Minimize actions** – Minimize action means steps per screen. The tasks and actions are streamlined so that they can be done in as few steps as possible. The interface should be designed keeping in mind to maintain the steps as few as possible for performing any tasks.
- **Consistent** – make sure elements in a user interface are uniform. They'll look and behave the same way. This helps constantly prove a user's assumptions about the user interface right, creating a sense of control, familiarity, and reliability.
- **Providing useful feedback** – The user should be provided with feedback for every action. This keeps the user informed and helps them to know whether some action was successful or not.
- **Clarity** – Content should provide the user with clarity. There should not be anything which confuses the user, as it becomes an obstacle for the user in interacting with the application.

# Demos





## Bonus points

1. A database for **easy accessing, managing and updating** the raw data, the aggregations of data and immediate analytic results.
2. Advanced analysis
  - a. Correlation analysis for different variables.
  - b. Visual analysis or pattern mining. For example, the baseline estimation for studying the normal behavior and abnormal events for these sensor signals. Or using statistical hypothesis testing methods to analyze certain proposed assumption, like how is the eda signal looking like in sleep versus daytime and if exists significant difference between them.

...

# Our technology



- Open source
- Syntax is intuitive and easy to learn and use
- ***Interpreted*** language, rather than a ***compiled*** language

```
jupyter wikidata (autosaved)
File Edit View Insert Cell Kernel Widgets Help
In [1]: endpoint = "http://query.wikidata.org/sparql"
display table
show all

SELECT DISTINCT ?item ?itemLabel ?formula WHERE {
  {
    SELECT DISTINCT ?item ?formula WHERE {
      { ?item [(wd:p1)*]/wd:p279 wd:p266. } UNION { ?item wd:p2534 ?formula. }
      FILTER NOT EXISTS { ?item wd:p13 ?image. }
      FILTER NOT EXISTS { ?item wd:p1 wd:p166546. }
      FILTER NOT EXISTS { ?item wd:p73 ?category. }
    }
    LIMIT 5
  }
  SERVICE wikibase:label { bd:serviceParam wikibase:language "[AUTO_LANGUAGE],en". }
}
ORDER BY ASC(?item)

Endpoint set to: http://query.wikidata.org/sparql
Display table
Result maximum size: unlimited

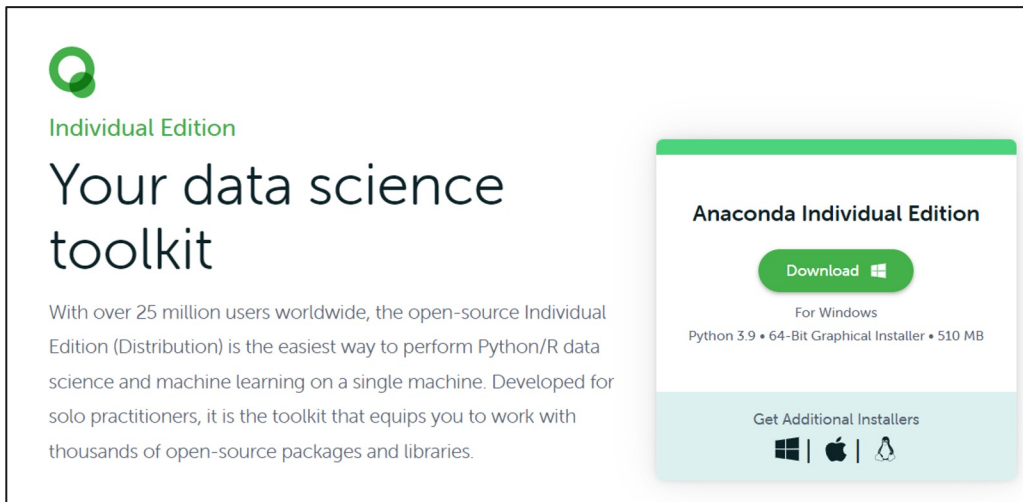


| Item                                                                                        | ItemLabel                        | Formula |
|---------------------------------------------------------------------------------------------|----------------------------------|---------|
| <a href="http://www.wikidata.org/entity/Q118078">http://www.wikidata.org/entity/Q118078</a> | CORDIC                           |         |
| <a href="http://www.wikidata.org/entity/Q130762">http://www.wikidata.org/entity/Q130762</a> | multiplication algorithm         |         |
| <a href="http://www.wikidata.org/entity/Q140710">http://www.wikidata.org/entity/Q140710</a> | General number field sieve       |         |
| <a href="http://www.wikidata.org/entity/Q217146">http://www.wikidata.org/entity/Q217146</a> | Teuchterberg system              |         |
| <a href="http://www.wikidata.org/entity/Q283883">http://www.wikidata.org/entity/Q283883</a> | common subexpression elimination |         |



Total: 5, Shown: 5
```

## 1. Install Anaconda

<https://www.anaconda.com/products/individual>




The image shows a screenshot of the Anaconda Individual Edition product page. On the left, there is a green Anaconda logo and the text "Individual Edition" in green. Below this, the heading "Your data science toolkit" is displayed in a large, dark font. A paragraph of text follows: "With over 25 million users worldwide, the open-source Individual Edition (Distribution) is the easiest way to perform Python/R data science and machine learning on a single machine. Developed for solo practitioners, it is the toolkit that equips you to work with thousands of open-source packages and libraries." On the right side of the page, there is a white box with a green header that says "Anaconda Individual Edition". Inside this box, there is a green "Download" button with a Windows logo. Below the button, it says "For Windows" and "Python 3.9 • 64-Bit Graphical Installer • 510 MB". At the bottom of the white box, there is a light blue section with the text "Get Additional Installers" and icons for Windows, Apple, and Linux.

 Individual Edition

### Your data science toolkit

With over 25 million users worldwide, the open-source Individual Edition (Distribution) is the easiest way to perform Python/R data science and machine learning on a single machine. Developed for solo practitioners, it is the toolkit that equips you to work with thousands of open-source packages and libraries.




#### Anaconda Individual Edition

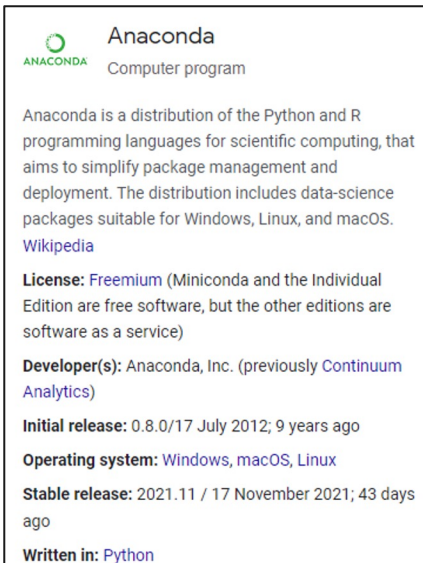
[Download](#) 

For Windows


Python 3.9 • 64-Bit Graphical Installer • 510 MB

Get Additional Installers



The image shows a screenshot of the Anaconda product information page. At the top, there is a green Anaconda logo and the text "Anaconda" in a large, dark font, followed by "Computer program" in a smaller font. Below this, there is a paragraph of text: "Anaconda is a distribution of the Python and R programming languages for scientific computing, that aims to simplify package management and deployment. The distribution includes data-science packages suitable for Windows, Linux, and macOS." followed by a link to "Wikipedia". Below this, there is a section for "License: Freemium (Miniconda and the Individual Edition are free software, but the other editions are software as a service)". Below that, there is a section for "Developer(s): Anaconda, Inc. (previously Continuum Analytics)". Below that, there is a section for "Initial release: 0.8.0/17 July 2012; 9 years ago". Below that, there is a section for "Operating system: Windows, macOS, Linux". Below that, there is a section for "Stable release: 2021.11 / 17 November 2021; 43 days ago". At the bottom, there is a section for "Written in: Python".

 **Anaconda**  
Computer program

Anaconda is a distribution of the Python and R programming languages for scientific computing, that aims to simplify package management and deployment. The distribution includes data-science packages suitable for Windows, Linux, and macOS. [Wikipedia](#)

**License:** Freemium (Miniconda and the Individual Edition are free software, but the other editions are software as a service)


**Developer(s):** Anaconda, Inc. (previously [Continuum Analytics](#))

**Initial release:** 0.8.0/17 July 2012; 9 years ago

**Operating system:** [Windows](#), [macOS](#), [Linux](#)

**Stable release:** 2021.11 / 17 November 2021; 43 days ago

**Written in:** [Python](#)

- 
1. Install Anaconda
  2. Open Jupyter notebook

