In the reading “Prosody of Text Communication” (week 5), you have seen how data-science can be used to judge the efficacy of different types of text-chat interface.

In this coursework, you carried out a task (maze game), while using two different interfaces. One interface was the typical Turn-By-Turn (TBT) Interface that is used in most instant messaging programs. The other interface shows each character as it is typed. These interfaces are called What You See Is What You Get (WYSIWYG) Interfaces.  
  
Your task as a datascientist is to determine which interface is better. To this end, your task is to develop some quantitative measures of the interaction, and then compare both interfaces, e.g. in a graph or table. You do not need to do inferential statistics (i.e. statistical tests). It is enough to do summary statistics.

**The data**

The data are stored in a spreadsheet

***TurnID:***Contains a unique identifier of the turn. Use this ID if you want to ask any questions about particular turns.

***DyadID:***A unique number that identifies each pair of participants

***Experimentid:***There were two different experimental conditions:

WYSIWYGTBT: Participants played 10 games using the WYSIWYG Interface. On the 11th and 12th games, participants swapped to the TBT Interface.

TBTWYSIWYG: Participants played 10 games using the TBT Interface. On the 11th and 12th games, participants swapped to the WYSIWYG Interface.

***Interfaceused:***

WYSIWYG: Participants were using the WYSIWYG Interface

TBTWYSIWYG: Participants were using the TBT Interface

***TURNTYPE:***  
**Data:** This contains information about each game. It records whether the maze was solved or if there was a timeout.

**WYSIWYG:** These are turns that are calculated by analyzing participants’ keypresses. For example, suppose P1 and P2 type the following keypresses in this order:   
  
P1: Hello how are you? N i c e  
P2: I am great a n d you?

This would be analyzed as:

P1: Hello how are you?  
P2: I am great  
P1: N  
P2: a  
P1: i  
P2: n  
P1: c  
P2: d  
P1: e  
P2: you?

This is because when using the WYSIWYG Interface there is no clear notion of turns (like there is in the TBT Interface)

The dataset contains these types of turns for text that is typed using the WYSIWYG Interface AND for text that is typed using the TBT Interface! (this allows direct comparison of their typing behaviour)

**Normalturn**

This contains the text that was sent “normally” using the TBT Interface.

**Sender:**

The participant who produced the turn.

**Text:**

The text that was sent in the turn.

**Timestamp:**

The time when the data were produced. The time is formatted using UNIX epochs.(The number of milliseconds since the 1st of January 1970. <https://www.epochconverter.com/>)

Look up the time you participated and try to find your dialogue (Note – a couple of dialogues were removed because of software crash).

**Duration:**

The time (in milliseconds) it took to type the turn.

**GameNo:**

Ranges between 0 and 11 (participants played 12 games)

**Your task is to develop and use measures to compare both interfaces:**

**Q1.** Which interface is the most efficient? Define a measure of efficiency and compare both interfaces in a suitable graph  
**Q2.** Which interface has better turn-taking? Define a measure and compare both interfaces in a suitable graph  
**Q3.** In which interface do they copy each other’s language more? Define a measure and compare both interfaces in a suitable graph  
**Q4.** Which is the happier interface? Define a measure and compare both interfaces in a suitable graph  
**Q5.** Which interface has more emotional synchrony? Define a measure and compare both interfaces in a suitable graph  
**Q6:** Can you think of any other interesting measure(s)? Define them and compare both interfaces in a suitable graph.

**Deliverables:**  
Submit a zip filethat contains   
- a folder containing the data you used for analysis  
- a folder for each analysis that contains  
 - a python script that runs out of the box  
 - R code (or other code) for generating the graph  
 - a pdf containing a short description of the measure,   
 the algorithm used, including any special steps/preprocessing/cleaning of the data,   
 the graph  
 two or three sentences summarizing the graph.

**N.B.**  
You can work in pairs or individually.  
It is very easy to write very coarse measures  
It requires a lot of thought to create more accurate measures.  
Make sure you know what you are comparing.  
Most of the measures will require some pre-processing.   
You will be graded on the elegance of your measures.