EXERCISE: Annotating Data for Interpersonal Emotion Regulation Strategies & Calculating Inter-Annotator Agreement

In this practical session, we will label a conversation with a predefined set of interpersonal emotion regulation strategies. The resulting annotations will be used in a Jupyter Notebook to calculate agreement between the annotators of your group. This notebook provides a basic introduction on different metrics that can be used to evaluate categorical annotation schemes. In the final part of this session, we will discuss how agreements scores should be interpreted and what methods could be used to increase agreement between annotators.

—

During the first 30 minutes:

Open the '1_annotation' folder and inspect the annotation scheme (see pdf). Once you have read through the labels and their meaning, open the 'annotation_excercise' file to start the exercise. The file is made available in two different extensions: an Excel-version (.xlsx) and a csv-version (.csv). If your computer has Excel installed, we advise you to use Excel, since you can choose labels from a drop-downlist, thus avoiding typos. If this is not the case, you can also use the csv-file.

If you need additional context for the annotation exercise, the full conversation is available in the file 'conversation_covid.pdf'.

Once you have annotated all instances, you can save the in the folder '2_iaa_scores'. VERY IMPORTANT: the name of your file in the '2_iaa_scores' folder has to correspond to this format:
'labels_' + <name_annotator> + <file_extension> (either '.csv' or '.xlsx')
→ e.g., 'labels_sofie.xlsx'

We will reuse your annotations there for the second part of this exercise. If you're making this exercise in a group, please share your annotated file with the other members of the group. Each member should place all annotated files in the folder '2_iaa_scores'.

—

During the next 25 minutes:

In this second part of the exercise, we will analyze the annotations of your group in terms of agreement. The folder '2_iaa_scores' already contains a Jupyter Notebook and a requirements file with the relevant Python packages that have to be installed. The folder should also contain the labeled files of the previous sessions (see instructions above), which you have to add yourself.

We advise to run the Jupyter Notebook in a virtual environment (you should have venv installed). To do so, you should execute the following lines of code:

1. Open terminal in '2_iaa_scores folder' or navigate to this directory in your terminal.
```
cd /path/to/your/project
```

2. Create a virtual environment by typing 'python3 -m venv <your_venv_name>'. In the code snippet below, we called the venv 'venv_tutorial'.
```
python3 -m venv venv_tutorial
```

3. Activate the virtual environment
    a. On macOS and Linux:
    ```
    source venv_tutorial/bin/activate
    ```

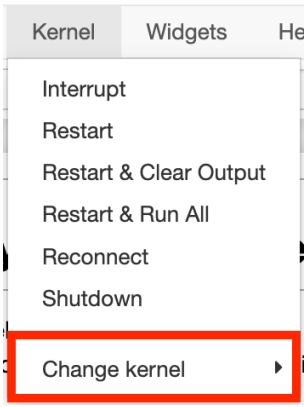
    b. On Windows:
    ```
    venv_tutorial\Scripts\activate
    ```

4. Install the packages listed in 'requirements.txt'
```
pip install -r requirements.txt
```

5. Add your virtual environment to the notebook:
```
python -m ipykernel install --user --name=venv_tutorial
```

6. Launch Jupyter Notebook
```
jupyter notebook
```

7. Navigate to 'calc_iaa.ipnyb' & check if the kernel in the toolbar is set to your virtual env (venv_tutorial in the example)



—

During the next 5 minutes:

Share the findings of your group in the general discussion. What is your group's opinion on the following questions:

- How do you interpret the inter-annotator scores you obtained? What could be possible causes of low agreement scores?

- Which labels were easy to annotate? Which labels were often confused between annotators or by yourself?

- How would you improve agreement? Which strategies would you implement to accomplish this?