

## Instructions for coding the attendees in excel

The meeting data is messy and needs to be cleaned as well as be connected to several features of the groups that have met with MEPs

The data-cleaning consists of the following things

1. Harmonizing the spelling of the attendees
2. Classifying attendees as being NGO, business, academia or other
3. Classifying attendees in terms of whether they are a single actor, an EU umbrella group or a national umbrella group.

## Steps for the classification of attendees in the files

1. Download your group's file with attendees from GitHub. The link can be found in the list of groups and legislative files on Moodle and in the Announcement that was sent last week.
2. This file has just a single column with the names of attendees. Create 4 columns next to the attendees column:
3. Column labels are in **bold**, column entries are next to it:
  - **Fixed\_names**: contains the standardized name for an organization (further instructions below)
  - **Class**: NGO, business, government, academia, other
  - **Structure**: single\_actor, European\_umbrella\_group or National\_umbrella\_group
  - **Comment**: remarks for groupmates or future notes

\*Do not forget underscores( \_ ) when typing the class and structure variables in excel. Otherwise you might run into trouble later. A screenshot can be found below.

4. Copy the attendees entries into the fixed\_names column.  
Before you can begin correcting the spellings of the entrees, read the following notes:
  - **DO NOT** change entries in the attendees column. We need this column to stay intact if we want to successfully merge the correct spellings into the larger data later on. You can standardise the spelling by changing the entries in the fixed\_name column.
  - Sort the names in alphabetical order. Some entries are very similar. Standardising their spelling works best if you replace the varying names in the fixed\_name column by copy+pasting the most suitable name into the other rows.

- It is preferable to have attendee names that include a transparency register number, if listed.
- Be aware that some attendees have different names in different languages. Thus, you will have to do some research to discover which organisation they are part of. The same goes for attendee entries that solely consist of a number.
- The process works the fastest if you also fill in the other columns (class, structure, comment) as you fix the spelling of each attendee.

Here is a screenshot of how the file is going to look (note that the Umbrella\_group score by now is differentiated in European\_umbrella\_group and national\_umbrella\_group respectively:

	A	B	C	D	E
	attendees	fixed_names	class	structure	comment
1					
2	Amnesty International Limited	Amnesty International Limited	NGO	Single_actor	
3	APEAL	APEAL - Transparency Register Number: 7003690	Business	Umbrella_group	
4	Apeal	APEAL - Transparency Register Number: 7003690	Business	Umbrella_group	
5	APEAL - Transparency Register Number: 70036906501-10	APEAL - Transparency Register Number: 7003690	Business	Umbrella_group	
6	Human Rights Watch	Human Rights Watch	NGO	Single_actor	
7	International Commission of Jurists	International Commission of Jurists	NGO	Single_actor	
8	McDonald's	McDonald's	Business	Single_actor	
9	US Ambassador	US Ambassador	Government	Single_actor	
10	Vinces Consulting	Vinces Consulting	Other	Single_actor	Consultant
11					

- In order to overcome language differences and discover the class/structure of attendees, you will need to use a variety of tools. Here's a list with some suggestions:
  - Google translate.
  - Transparency Register: [Zoeken in het register - Europese Unie \(europa.eu\)](https://zoekregister.europa.eu/)
  - Lobbyfacts: [Search | lobbyfacts](https://lobbyfacts.eu/)
  - Websites of the attendees.
  - ChatGPT or AI tools can be useful if you are really unsure about the class or structure of a given attendee. However, make sure it identifies the same attendee as the one in the data.
- Make sure to work in your shared online file and ensure automatic save.** Also be sure to give the modified version of the file a different name, otherwise your progress will be lost if you re-run the R code.
- After the table has been filled in, we can then import it back into R and join it with our file data.