

To do list

- write about the context of the study : Michal Raszka, Rita, Roberts, Alina (RK)
- SD for data (TG : Calculate it & find an example on GGplot)
- group graphs by sections so that we have each major section of the survey on one page. (RK)
- Improve esthetism of the graphs. Do all plots with *ggplot2* as I know it supports colour palettes / themes very nicely (RK)
- Percentages in the mergins with likert figures (RK)
- Replace not Selected by No (TG) ** http://kbroman.org/knitr_knutshell/pages/figs_tables.html **
<https://www.qwant.com/?q=ggplot&t=images>
 - just to mention - in ggplot version I have put all the questions in one - producing a comparable barchart; perhaps it is not necessary to account for “Not selected” - this way the output is more compact. This can be seen in **Fig. 6** and **Fig. 9** (RK)

On-going

- Get the French Psychotherapy study (marion.azoulay13@gmail.com contacted, waiting the answer)
- Sweden data get them (Michal Raszka)

Deadline April 28

- Write an abstract for EABCT April 28 – Poster abstracts due
- Make a poster for EABCT April 28 – Poster abstracts due
 - Is it the whole poster that is needed by then, or just the abstract with selected graphs? (RK)

Questions :

- In which country psychotherapy training is the best? (TG)
- In which country there are more hours of supervision ? (TG)

Free text responses to be analysed:

There is a number of freeform text questions that provide interesting material. Freeform text questions were generally not mandatory - people gave answers if they felt like doing so. In providing summaries, I am sure we could quote some of the most interesting or prevalent answers/thoughts. We have freeform answers to analyse for:

- **q_0005_other** - Other kind of psychotherapy person is training in. We also should reflect it in the existing graph (Fig. 6) - by adding one more column with count how many non-NA responses are there (*will be done by RK*). However we should at least provide some kind of grouping of other therapies that have been mentioned and an overview of what kind of responses people have provided under **other**. 24 responses here.

- **q_0006** - this is list of therapies people are training in. Only people who said they were training in psychotherapy were presented with this question. Hence we need to classify the responses in some sensible manner and provide count of responses in each class (e.g. CBT, Psychodynamic, etc). 329 responses.
- **q_0009_other** - this is very similar to q_0005_other - but it gives an 85 items long list of other therapies that people are *interested in*. This also needs classification and counting within classes.
- **q_0010** - this is a very important and more difficult question. It asked people for their comments regarding their preferences in psychotherapy and we have 165 textual answers which we need to classify and make an overview of. I think this is a challenging, but particularly interesting task.
- **q_0012_other** - provides **other** answers for question - Are you qualified to practice psychotherapy after your psychiatry training? This is very short - it only gives 7 answers.
- **q_0013** - another interesting classification/summary - what therapies people have seen available for patients at the place they are receiving their training. That is rather straightforward, but people have given 510 responses, so will require some effort. It is very interesting though.
- **q_0015_other** - describes **other** kinds of supervision that people receive. 21 responses.
- **q_0023** is the final text question - it asks for general feedback after the survey and it contains 55 responses. Contains some of the most interesting thoughts.

How to extract text responses from the data

Generally it is quite simple. We can print non-NA responses using the command to omit NA `na.omit(data_Europe_clean$q_0023)` - which gives us numbered rows of answers, but requires a bit of scrolling in R. If we want to make it into a dataframe for easy export to Excel etc. we can do `df.0023 <- na.omit(data_Europe_clean$q_0023)`.

Done

- add a column with name
- merge all the data : analysis in the whole europe (TG) <http://www.r-bloggers.com/string-concatenation-in-r/>
- number of trainee in each country (TG) principle - list directories read that list into a vector then for each element of the vector (i.e. country name) and all the syntax files for that reason also start with the country name so if country name is a variable we simply substitute that variable into the script when we loop over the list (i.e. the vector) containing names of directories
- Europe map: R actually has good mapping capabilities. We should be able to produce plots of Europe with each country in colour depending on the particular variable in that country. A heatmap of sorts, where the word “map” is for real (TG) Done
- Ask help for statistical inference ? MOOC : <https://www.coursera.org/learn/statistical-inference/home/welcome>
- NB : change labels in the analysis file not in syntax file
- UEMS recommendation abline <http://stat.ethz.ch/R-manual/R-devel/library/graphics/html/abline.html> Done - * ? just a thought: what data we are trying to show by it? Perhaps we best show percentage differences between awareness and implementation awareness?
- Mean European abline

```
q9df <- data[,16:21]
```

To use git (after installation)

- cd to the appropriate repo !!!
- git status
- git pull origin master
- modify stuff
- git status
- git add .
- git commit -m "my message"
- git push origin master