**zoomGroupStats Tutorial: Supplementary Guide**

This guide accompanies the tutorial and demonstration of zoomGroupStats. [You can access the tutorial code](http://apknight.org/zoomGroupStats_tutorial_code.R) here and [zoomGroupStats here](http://apknight.org/zoomGroupStats.R).

*Please feel free to insert comments throughout the document if you have questions or run into trouble with any of the functionality described here.*

**Guidance for Segment on Setting Up Zoom & Downloading Files to Use in Analyses**

Recommendation: Maximize degrees of freedom in recording settings

* If using cloud-based recording, select all possible recording options (of different views). This gives you the ability to make selective decisions after-the-fact.
* Select options to enhance the recording for 3rd party video editing.
* Select option to have Zoom produce an audio transcript
* Make other option selections in a manner consistent with your research goals (e.g., having names on videos, having video time stamped).

Recommendation: Develop a standardized protocol for yourself + assistants to follow

* [Sample of a guide given to those charged with recording meetings](https://docs.google.com/presentation/d/1B9Cdc-tdB4mKYjIXQ7R-RgF5-HEazU_Daiik_1GH8WY/edit?usp=sharing)
* [Sample video guide for how to set up Zoom recording features](https://youtu.be/Y82nf9lfeQU)
* [Sample video guide for recording the meeting itself](https://youtu.be/HbbKcmbaLYI)

Recommendation: When possible, require users to be registered

* A major challenge in collecting data at scale is a lack of a persistent individual identifier for participants who join a session.
* When your research design allows for it, require users to be registered in Zoom. This will at least give you an email address that you can use to identify participants and will be linked to their initial screen name.

Downloading Files

* Download the Usage Report for the focal meeting:
  + From the “Reports” page in Zoom, click “Usage Reports”.
  + Scroll to the “Participants” column for your focal meeting. Click the linked # of participants.
  + Checkbox “export with meeting data” and “show unique users”
  + Click Export
* Download selected files from the Recordings page:
  + Chosen video files (e.g., gallery, active speaker, shared screen)
  + Audio file
  + Transcript
  + Chat

**Guidance for Segment on Analyzing Language**

Load the zoomGroupStats Functions

* Use source(“<http://apknight.org/zoomGroupStats.R>”)
* Can also pull from and/or fork on github (<https://github.com/andrewpknight/zoomGroupStats>)
* You will need to have the following packages installed on your machine:
  + reshape2
  + stringr
  + paws
  + magick
  + data.table
  + jsonlite
* Additional requirements are listed below (i.e., AWS account with credentials, ffmpeg) for specific functionality

Parsing the Transcript File

* One challenge in parsing the transcript file and/or the chat file is the lack of persistent individual identifiers. The name applied is the person’s screen name. This can both change and be duplicated with someone else in the meeting. The best solution is to manage this through requiring registrations and making sure people fix their screen names (as described above).
* If you do not manage this using registrations, you will have tco do manual clean up and linkages between people’s screen names and their identities in your other data sources.
* NOTE: Zoom timestamps the chat file anchored on the start of the session itself.

Parsing the Chat File

* Same issue as above in people’s screen names not being a unique identifier
* Sometimes there are some weird things that happen in the chat file as a function of people inserting tabs or special characters. The parsing is still pretty crude. So, it is worth taking a look through your chat files + output to see if everything is correct.
* NOTE: Zoom timestamps the transcript file anchored on the start of the recording. The issue here is that you may not start the recording until a period of time after you have launched the session. This means that you could have a transcript file out of sync with the chat file. Resolving this issue requires careful records of when the recording was started *or* setting up your session to automatically record.

Analyzing the Conversation Dynamics

* The basic conversation analysis provides an output at the meeting level and an output at the speaker level.
* The turn-taking analysis is preliminary (alpha). But, it provides output at the dyad-level and the individual level.
* The sentiment analysis requires that you have an AWS account and that you have properly configured your credentials. [This is a useful guide for doing so](https://github.com/paws-r/paws/blob/master/docs/credentials.md).
* The windowed analysis is also currently in a preliminary stage. It is only available for transcripts right now. But, will be extended to the chat soon.

**Guidance for Segment on Analyzing Visuals**

* Using these features requires that you have ffmpeg installed on your machine. This is used to break the video up into still frames. On a Mac, the easiest way to install is:
  + [Install Homebrew](https://brew.sh/)
  + [Then install ffmpeg](https://formulae.brew.sh/formula/ffmpeg)
* You also have to have an AWS account with correctly configured credentials. Again, t[his is a useful guide for doing so](https://github.com/paws-r/paws/blob/master/docs/credentials.md).
* If you want the software to detect identities in the photos, you need to have already created a repository on AWS that contains the identified reference images.
* The output of this is at the personXframe-level. So, you will need to decide how you want to aggregate this information up to the meeting, person, or time level of analysis.