

Information sheet

Working with sensible data on the LAV Postgres server

Some of the datasets from our server you will be working with are sensitive, for one of two reasons:

- a) the data are proprietary, i.e. they belong to companies or individuals, who shared this information with us under the condition that we keep it secret.
- b) the data contain personal information subject to data protection laws.

Contracts between ETHZ-LAV and the respective owners of the data regulate their usage. This is a very serious matter! For one, a breach can have severe legal repercussions. And even if not, it would still be disastrous to ETHZ-LAV: we depend on data; if people don't trust us to keep their information safe, we're finished.

To grant you access to the data, we had to extend the data protection and usage contracts to you; that is why your supervisor asked you to sign the "project worker agreement". You are now subject to the contract(s) the agreement references; you can view them at any point during your project — just ask your supervisor.

So, also in your own interest, please handle the data with care. Here are a couple of guidelines:

1. Generally: use common sense; when in doubt, ask your supervisor.
2. Do not use the data for anything other than the purposes of your project. Generally: as long as your supervisor knows what you are doing and why, you are fine.
3. Do not make copies of the data; the only exception are temporary copies for processing; but we recommend working directly from the database, without temporary storage on the filesystem. If you made any copies for processing (including intermediary results), you must destroy them irreversibly at the latest at the end of your project.
4. Do not share the data or intermediary results with anybody except your supervisor. This counts for the dataset as a whole, but also individual records. Next to giving somebody physical or digital access, sharing also includes showing someone, e.g. on a screen or print-out, or telling them. Asking someone for technical support may fall under that definition.
5. Analyses must be statistical in nature, i.e. you may only evaluate the records as part of a population, not as individuals on their own account. Beware: because of this, some contracts have special stipulations about what can be put in a graph.
6. The data in our database is anonymized; do not circumvent this. In particular, do not try to determine which record(s) relate to a given individual.
7. Be sure to use a secure password. Do not share it with anybody, do not write it down and do not put it verbatim in your code or any files accessible to other users. To change your password, ask your supervisor.
8. As per the ETHZ "Acceptable Use Policy for Telematics Resources", do not store code or data on non-ETHZ cloud services, such as *Dropbox*, *OneDrive* or *iCloud* — *polybox* is OK.
9. Do not put data on USB-sticks; they are easily lost.