Task 2

GitHub

To properly conduct version controlling as well as storing the proper documents via github its required to create a repository and store the initial documents in it. The following images provides an overview of the gitcommands used to open, add and commit the initial files to a master branch. In order to have multiple versions available and the ability to share across devices or to other people the need for communication rises, as well as properly conducted uploading once the work is done.

Firstly its important to log on via the git config command, locating the files via cd documents- into the correct folder before git add \*, commit m- message and git remote add origin before pushing the files onto the server. Also remembering to close the programs before committing any files.

Et bilde som inneholder skjermbilde, innendørs, overvåke, datamaskin

Automatisk generert beskrivelseEt bilde som inneholder skjermbilde, overvåke, datamaskin, bærbar PC

Automatisk generert beskrivelse

Considering the possible involvement of different people continuously working on the files one might deem it necessary to define several crucial steps to the strategy. One might start with the data subject; the original data collection would gather it from somewhere or someone. Which brings us to us, data controller is the person(s) in charge of monitoring and store the data. Security is crucial when focusing on sensitive data, luckily that is not the case here. Data processor, which github proves useful to depend on when regarding letting other entities gain access to the files. The data controller is responsible for encryption and must take this in consideration before sharing. Furthermore, the strategy itself relies upon the completion of git commands in respect to the server. By using the add, commit, and push command it enables the uploading to the server, from here it is possible for participants to use the clone command if they are provided with the repository code. Allowing participants in the work to clone, add and commit changes to the master file on the server site. It is important to note that for this to work the main data processor need to make use of the pull command. The committing of files via push commands uploads to the server, but the pull command enables the handler to get access to the changes made.

Github also gives the ability to branch the work, if we were to conduct a split research approach, meaning someone performs tasks conducting analysis of one aspect like marketing. A purposeful approach would be to branch the files into separate departments, marketing, sales or other respective departments of a company. It is though important to note that the branches cannot involve themselves in other departments files, because of the merge back into master branch. If this were to happen the collaboration might include elements from other departments and make statements, data or analysis redundant or repetitive.

This might draw upon principles from data cataloging, where principles are implemented to sort or divide data into respective compartments. This would of course rely on large datasets with many variables and the ability to compartmentalize them, the project currently conducting does not hold the depth which would justify detailed compartmentalizing. It is through mention worthy to adopt such an approach if possible or departmental collaboration, also with regards to security and sensitive data.

The workflow itself relies upon the continuous committing of changes or files from day to day, as said previously the possibility to branch is helpful, if needed. >This project does not make use of it and therefore makes use of one master branch alone. Updates made from day to day stores themselves in the server and people responsible for changes would be required to use the git config; email and password in order to keep track of the different contributors, which in turns help to produce a purposeful overview of data processors.