# Course:

URL: <https://class.coursera.org/datascitoolbox-015/lecture>

# Week 1:

Data Scientist:



How to ask questions:

<http://www.catb.org/~esr/faqs/smart-questions.html>

Drew Conway:

<http://drewconway.com/>

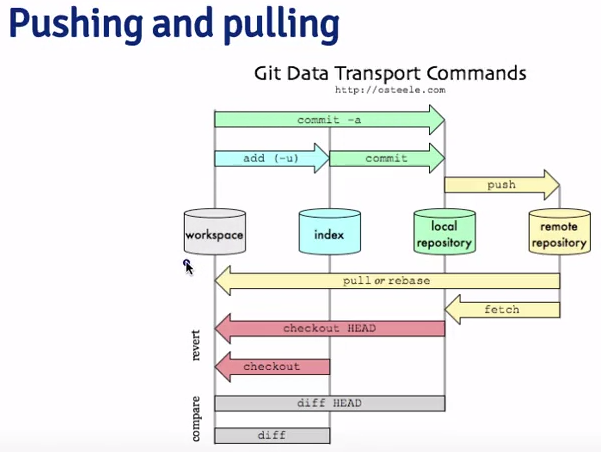
# Week 2

Command Line Interface (CLI) – Git Bash:



Path working directory, clear to home, list folders, change directory, make directory, create empty file, etc.

Git architecture:



Git commands:

* Git add .
* Git add –u
* Git add –a
* Pwd – path to the current working directory
* Cd – change directory
* Cd .. – move up in directorys
* Mkdir ~name – create a directory called “name”
* Git clone <https://github.com/sacooper74/xxxx.git> - create a clone of a directory
* Git commit –m “insert comment” – commits changes to the local repository
* Git push – after the commit, pushes changes to the remote
* Touch xx.md – creates a markup file (next you have to add, commit and push)

Setting up Git Bash:

git config --global user.name "sacooper74"

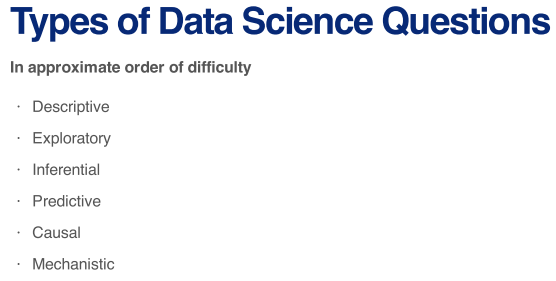
git config --global user.email "steve.a.cooper@gmail.com"

cd /c/Users/steve.a.cooper/'Google Drive'/Training/datasciencedata/assignment

Git clone https://github.com/sacooper74/data\_assignment.git

git init

# Week 3



Data are values of qualitative (gender) or quantitative (height) variables (measurement or characteristic), belonging to a set of items (the population).

Control for confounding variables by (a) fixing, (b) stratifying or (c) randomizing.

