Github | Docker | Actions

Automatisation of R scripts

Statisticcafe 2022, 16.November
Michael Kahle
(michael.kahle@geographie.uni-freiburg.de)
Physical Geography, University Freiburg

UNI FREIBURG

Outline

- Version Control
  - git
  - github
- Docker
  - Ubuntu
  - R
  - Libraries
- Github-Actions
  - Submit, Push
  - Cron
  - Manuell
  - Install additional libraries

Version Control

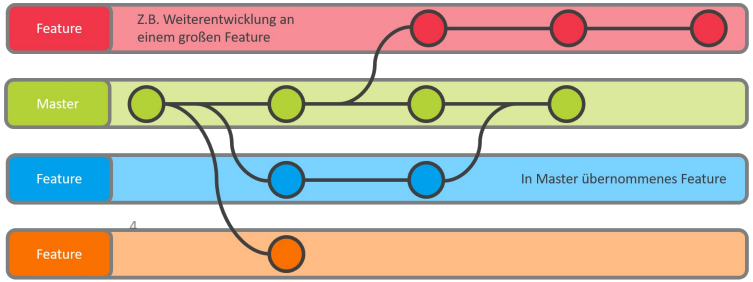
- Lost Code
  - Print out
  - Backup
- New Versions
  - filename_v3_final.R
  - code_2022_11_16.R
- Different people
  - stat_mk_cd.R
- Better
  - Version Control
  - Central repository
  - See History

Hamilton in 1969, standing next to listings of the software she and her MIT team produced for the Apollo project

git

- Free and open source (GPL)
- Linus Torvald for developing Linux
- Linux, macOS, Windows
  - download
  - Documentation
  - Some GUI available as well

- Most important commands
  - **Clone**: Get remote repository (once)
  - **Pull**: Update from remote
  - **Status**: Info on changes
  - **Add**: Add local changes
  - **Commit**: Prepare local changes
  - **Push**: Send local changes to remote repository



github.com

- Hosted Software Platform
  - Free and paid plans available
- Bought by Microsoft
- Git and much more…
- Own CLI interface
  - download
  - Documentation
  - GUI is the website
- Integration with Visual Studio Code
- Integrates with zenodo (DOI for code)
- Shows markdown (README.md)

Docker

- (Container-)Virtualisation
  - Defined by Code
- Based on Linux
  - Windows & MacOS via VirtualBox
- Defines
  - OS (i.e Ubuntu)
  - Installed Software (i.e. R & packages)
  - Volumes (~ hard- or usb-drive)
  - (ports, memory, …)
- Own defined containers can be hosted via hub.docker.com -> available for all, everywhere

Docker example

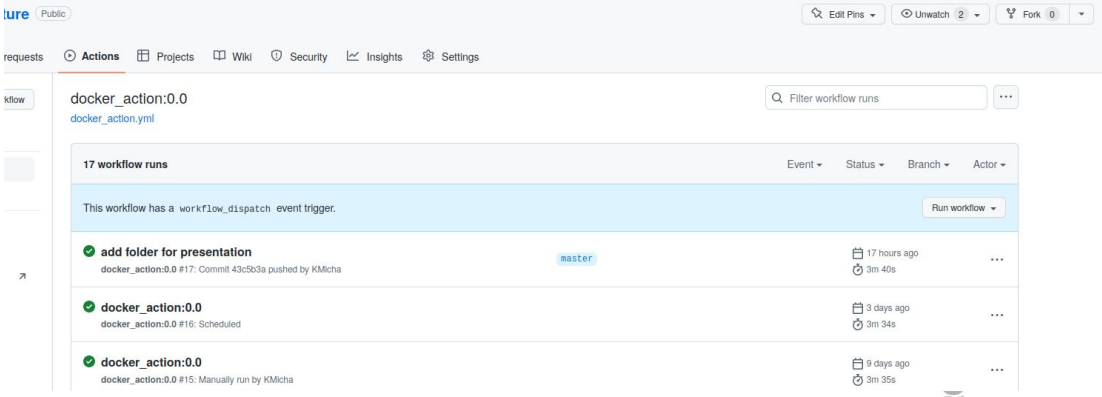
- Defined container for using R
  - Usable on every computer without version hell
- creCoding (part of CRE/VRE infrastructure)
  - @dockerhub
  - Windows & MacOS via VirtualBox

- `FROM tamboraorg/creubuntu:2020.0`
  - Based on Ubuntu 20.04
- `RUN apt-get -y --no-install-recommends install r-base r-recommended …`
  - Install R (and many more packages)
- `WORKDIR /cre/code`
  - Shares directory for code

Github Actions

Example see [climdata/dwdTemperature]

- Quite new concept
- Idea: Run program code with/for repository
- Trigger on
  - Repository change (submit/push)
  - Periodically in time-intervals
  - Manually by UI
- Simply needs a file:
  [.github/workflows/docker_action.yml]
  (Name is choosable; .yml needed)



Github Actions: action.yml

Example see [climdata/dwdTemperature]

Structure:

- # Comment (anywhere)
- name: doThis:0.0 (name of action, version)
- on: (several options to trigger action, next slide)
- jobs:
  - docker-run-action:
    - runs-on: ubuntu-latest (OS to use on github side)
    - container: (which container to use)
    - steps: (commands to execute)

Github Actions: action.yml

Example see climdata/dwdTemperature

Trigger: (one or more)

- on:
  - Push:
    - Branches:
      - - ‘master’ (new&default: ‘main’ !)
  - Workflow_dispatch:
  - Schedule:
    - # ‘m h d/m m d/w’ : weekly,Sunday,10:00
    - - cron: ‘00 10 * * 0’ (CRON syntax)

Github Actions: action.yml

Example see [climdata/dwdTemperature]

jobs/docker-run-action:

- jobs:
  - docker-run-action:
    - Runs-on: ubuntu-latest
    - Container:
      - **image: tamboraorg/crecoding:2020.0** (R environment)
      - Env:
        - NODE_ENV: development (not needed here)
      - Ports:
        - - 80 (not needed here)
      - **volumes:**
        - **- ${{ github.workspace }}:/cre/R** (binds to repository)

Github Actions: action.yml

Example see [climdata/dwdTemperature]

jobs/steps:

- jobs:
  - steps:
    - - uses: actions/checkout@v3 (Copies repository into action workspace)
    - -name: Set Ownership
    - Run: chown -R $(id -u):$(id -g) $PWD () (fix needed for older git versions)
    - -name: Install pandoc (install ubuntu packages)
    - Run: (apt-get -y --no-install-recommends install pandoc)
    - -name: Install markdown (install R packages)
    - Run: (R -e "install.packages('rmarkdown')")
    - -name: Run Knitr (Finally run the knitr - or any other R file)
    - Run: (R -e "rmarkdown::render('/cre/R/README.Rmd')")
    - -name: Submit Changes
    - Run: EndBug/add-and-commit@v9 ()

UNI FREIBURG

Github Actions:

Other languages can be used as well (i.e. python, julia)
Action Cascade: climdata (T,P,TI,SPI,drought)

How to start:

- Existing R repository on github.com
  - Simply create an action.yml file below .github/workflows
    Start with copy from here
  - Activate Actions (by clicking on tab)
- No account on github.com
  - Join on github.com
  - Fork climdata/dwdTemperature
- Run docker locally
  - sudo docker run --volume $(pwd):/cre/R tamboraorg/crecoding:2020.0
    R -e "rmarkdown::render('/cre/R/README.Rmd')"