

Get going with git

Mark Jones

2023-11-02

Table of contents

Introduction	5
Pre-requisites	6
Background reading	6
Command line interfaces	6
Windows	6
macOS	7
Extra credit	7
Operating system management - environment variables	7
Windows	7
macOS	7
Operating system management	7
GitHub	8
I Part 1 - Logistical elements	9
1 What is (this thing) called revision control	11
1.1 The big picture	11
1.2 Why commandline	13
2 Git install	14
2.1 RStudio	14
2.2 Install git	14
2.2.1 Mac OSX	14
2.2.2 Homebrew	15
2.2.3 Windows	17
3 Git setup	19
3.1 Configuration for git	19
4 Github setup	20
4.1 GitHub account	20
4.1.1 Personal access token	20
4.2 Git Credential manager	20
4.2.1 GCM install	21

4.2.2	GCM demo	21
4.2.3	GCM configuration (advanced only)	25
4.3	GitHub CLI	25
4.3.1	GitHub CLI authentication	26
II	Part 2 - Fundamentals	29
III	Aside.	31
5	Repositories	33
5.1	Git repositories	33
5.1.1	Initialisation	33
5.1.2	Repository components	33
5.1.3	Adding project files	33
6	Commit	34
6.1	Adding files to projects	34
6.2	Commit process	34
6.3	Tracking status	34
6.4	Reviewing commit history	34
7	Reviewing differences	35
8	Branch	36
8.1	Why branches?	36
8.2	What is a branch?	36
8.3	Special branch concepts	36
8.4	Creating a branch	36
8.5	Switching branches	36
8.6	Working on branches	36
9	Merge	37
9.1	Merge concepts	37
9.2	Merge processes	37
10	Tags	38
11	Reviewing history	39

IV Part 3 - Collaboration	40
About	42
Repository status	42

Introduction

When you are getting into the nitty gritty, version control is a large topic. The git, github, github cli and github pages ... technologies can be daunting to set up and use. However, once you are setup properly and understand how things fit together at a high level, users can go a long way with a minimal set of commands.

I know, it is true that a little knowledge can be a dangerous thing, but it probably isn't quite as bad as complete ignorance.

Get going with git - aims to provide a minimal knowledge base, sufficient to get you started without having to wade through reams of documentation distributed all over the internet. It can be browsed online or downloaded as a word document for your reference.

What is version control? Setting up git Time to commit: working with a local repository Push and pull: working with a remote repository Making your first repository on GitHub Day 2: Collaborating with Others

Dealing with (git) conflicts Branching for sanity Creating pull requests Collaborating with GitHub tools Day 3: Dealing with Complications

Undoing changes Learning good repository organization principles Ignoring things (in git) Working with large files Using the README Creating GitHub templates Extending git and GitHub with other tools

Pre-requisites

Background reading

There is no point reinventing the wheel, Jenny Bryant motivates the use of version control in the following paper: [Excuse Me, Do You Have a Moment to Talk About Version Control?](#)

Command line interfaces

A CLI is a software mechanism that you use to interact with your operating system via your keyboard rather than a mouse. You enter in commands as text and the system will do something, e.g. delete a file.

CLIs are software that are supplied with the operating system. Software that implements such a text interface is often called a command-line interpreter, command processor or a shell.

Nearly universally, if I use the word *terminal* I am referring to the operating system command line interface.

Windows refers to its CLI as the *command prompt*, and in macOS we have the *terminal*.

If you do not know how to operate your operating system CLI then you need to address that.

Windows

You can do these tutorials to familiarise yourselves:

- [How to use Windows 10's Command Prompt](#)
- [Learning Windows Terminal](#)

The following link needs admin rights (TKI people probably won't have nor will be able to obtain these rights) and is not strictly necessary here, but it gives a lot of useful commands and an opportunity to gain a bit more familiarity with the commandline.

- [40 Windows Commands you NEED to know](#)

macOS

- [How To Use Terminal On Your Mac](#)
- [What Is the Mac Terminal?](#)
- [Absolute BEGINNER Guide to the Mac OS Terminal](#)

Extra credit

- [Learn the command line](#)

Operating system management - environment variables

You will need to have some minimal technical competence in driving your computer. At a minimum, you need to know how to set an environment variable under your operating system of choice. If you do not know how to set an environment variable, then you need to address that.

Other items that are important to be familiar with are file and directory concepts. For example, if you do not know what the command `tree` does, then it would be useful to find out. Similarly, if you do not know what file permissions are then, again, it would be useful to find out.

Windows

You can do these tutorials to familiarise yourselves:

- [Environment Variables : Windows 10](#)
- [How to Set Environment Variables in Windows 11](#)

you should be able to set a user variable even if you do not have admin priviledges.

macOS

- [How to Set Environment Variables in MacOS](#)
- [PATH Variable \(Mac\)](#)
- [Use environment variables in Terminal on Mac](#)

Operating system management

Create a directory on your machine where we will store all the files for this workshop.

GitHub

Follow part 1 of the instructions provided by [Getting started with your GitHub account](#) to create and configure your account.

Warning

The part on **configuring 2-factor authentication** is absolutely mandatory, the rest of the 2-factor content can be skimmed. See [Configuring two-factor authentication](#).

To use the USyd GitHub Enterprise Server, you will need a unikey. If you have a unikey, you should have access. Go [here](#) and confirm that you can login.

Part I

Part 1 - Logistical elements

In this part we will get everything set up. This may well be the hardest part.

1 What is (this thing) called revision control

1.1 The big picture

First, let's briefly introduce some minimal terminology and context.

i Note

TODO - Timelines

A **repository** is the mechanism that is used to implement version control by git.

i Note

The repository is implemented by a hidden directory called `.git` that exists within the project directory and contains all the data on the changes that have been made to the files in the project. You should never touch this directory nor its contents.



There are two types of repositories, *local* and *remote*.

The local repositories reside on your machine. Remote repositories are hosted by service providers, the most common being GitHub, GitLab and Bitbucket. We only deal with GitHub here. GitHub comes in a few varieties:

- GitHub Enterprise is hosted by the company called GitHub, see github.com. It is a commercial platform, but parts of it are made freely available.
- GitHub Enterprise Server is self-hosted; this is what USyd provides via <https://github.sydney.edu.au/>

In a nutshell, git provides a set of commands that allow you to manage the files that are retained in these local and remote repositories.

Again, I cannot sum it up better than Jenny, so please take the time to read it.

[Excuse Me, Do You Have a Moment to Talk About Version Control?](#)

1.2 Why commandline

Because it is the best way.

2 Git install

Inevitably there are some installation tasks that we need to take care of before we proceed.

Warning

The following steps can be a bit of a pain. Don't be disheartened, it gets less tiresome.

2.1 RStudio

Developers - is your instance of RStudio up to date? If not, update it. Ditto for R. Keep them both updated.

2.2 Install git

I am going to break this down into Mac and Windows because they are the two systems that most of AHI seem to use and the installation is somewhat different for each. If you are using Linux, you probably have no need to be reading this.

2.2.1 Mac OSX

First, do you have git installed already? Launch the **terminal** app (see the pre-requisites on the landing page if you do not know how to do this). In the terminal, type:

```
which git
```

which should show the location of the version of git in use:

```
## /opt/homebrew/bin/git
```

If you have homebrew (see below) installed, you can just type:

```
brew install git
```

and git will be installed, otherwise, follow the instructions below and then come back here.

Once git is installed run the `which git` command again and then run `git --version` which is shown (along with the output) below:

```
git --version
## git version 2.42.0
```

If you got here, then you have git installed. You can close down terminal, open it up again and then run the `git --version` command again to make certain that everything is ok.

2.2.2 Homebrew

In the previous section, you can see that the path output from the `which git` command includes `homebrew`. For macOS, `homebrew` is a package manager. This basically just lets you install and manage packages (applications) on your mac.

To use homebrew, you need to install it first. To do that, go [here](#), then follow the instructions, which amount to going to the terminal and running the commands listed below.

Please go and read the landing page for homebrew before you proceed any further.

The first command ensures that pre-requisites are met, see [here](#):

```
xcode-select --install
```

if this has already been done you will get an error, or be asked to run Software Update. Generally, you can just move on to the next command:

```
/bin/bash -c "$(curl -fsSL https://raw.githubusercontent.com/Homebrew/install/HEAD/install.sh)"
```

`curl` is a tool for transferring data from a url. It is usually stored under `/usr/bin` but if you are having issues then exporting the following might assist:

```
export HOMEBREW_FORCE_BREWED_CURL=1
```

which basically tells homebrew to use its own version of curl. After the above is complete, homebrew should be installed. Now you can run

```
brew install git
```

to install git for you.

For reference, here are a minimal set of commands for using homebrew (additional information can be found in the homebrew man pages).

Basic information on homebrew:

```
# Display the version of Homebrew.
$ brew --version
# Print Help Information
$ brew help
# Print Help Info for a brew command
$ brew help <sub-command>
# Check system for potential problems.
$ brew doctor
```

Keep your homebrew applications up to date:

```
# Fetch latest version of homebrew and formula
$ brew update
# Show formulae with an updated version available
$ brew outdated
# Upgrade all outdated and unpinned brews
$ brew upgrade
# Upgrade only the specified brew
$ brew upgrade <formula>
# Prevent the specified formulae from being upgraded
$ brew pin <formula>
# Allow the specified formulae to be upgraded.
$ brew unpin <formula>
```

The core commands for managing commandline applications are:

```
# List all the installed formulae.
$ brew list
# Display all locally available formulae for brewing.
$ brew search
# Perform a substring search of formulae names for brewing.
$ brew search <text>
# Display information about the formula.
```



```
$ brew info <formula>
# Install the formula.
$ brew install <formula>
# Uninstall the formula.
$ brew uninstall <formula>
# Remove older versions of installed formulae.
$ brew cleanup
```

Homebrew casks allow you to install GUI applications. Unless you are an advanced user, you will rarely need to use these, but for completeness:

```
# Tap the Cask repository from Github.
$ brew tap homebrew/cask
# List all the installed casks .
$ brew cask list
# Search all known casks based on the substring text.
$ brew search <text>
# Install the given cask.
$ brew cask install <cask>
# Reinstalls the given Cask
$ brew cask reinstall <cask>
# Uninstall the given cask.
$ brew cask uninstall <cask>
```

2.2.3 Windows

The official site for the git windows binary download is <https://git-scm.com/download/win>.

Download the 64-bit standalone installer, run it, agree to the conditions and license, choose the default location.

Ensure that the following install components are chosen:

- windows explorer integration
- large file support

and accept any other defaults.

With the exception of the following, for any of the other prompts, just accept the defaults.

1. You will need to nominate a text file editor for editing commit messages and so on. Unless, you know what you are doing, I would advise just select the Windows Notepad application, you can reconfigure this later if you want to.

2. You should select to override the default branch name as `main`. The reason to do this is so that git aligns with github (which uses `main` as its default branch).
3. For adjusting the `PATH` environment variable, ensure that you select `Git from the command line and also from 3rd-party software` which is the default.
4. Ensure that line ending conversion is set to `Checkout as-is, commit as-is`.
5. For the terminal emulator, select `Use Windows default console window`. This has some limitations but it is ok for an introduction.
6. Ensure that `Git Credential Manager Core` is selected when prompted.

We will run through this install for someone in the group.

To keep git up to date, you will need to go to the above site and download and reinstall git.

Open the command prompt and type:

```
git --version
## git version 2.42.0
```

3 Git setup

3.1 Configuration for git

Per the sentiment of Fred Basset, you are now up but not quite running.



Figure 3.1: Fred Basset

One of the first things we need to do is to set a username and email address:

```
git config --global user.name "Fred"
git config --global user.email "fred.basset@comic-land.com"
```

You can list your configuration with

```
git config --global --list
```

We will get into the why later, but basically any interaction you have with git will be tied to your username and email address. This has obvious benefits if we want to be able to figure out who has done what, when and why.

4 Github setup

4.1 GitHub account

As noted in the pre-requisites for using this knowledge base, you have to have GitHub account.

While there are multiple ways to interact with GitHub from your local machine, here we will use the commandline. There are two protocols that can be used, HTTPS and SSH. We will use HTTPS.

First we need to set up a Personal access token.

4.1.1 Personal access token

GitHub introduced personal access tokens a short while ago. Personal access tokens are basically a password with some bells and whistles.

1. Login to your GitHub account.
2. Open [Creating a personal access token \(classic\)](#) in a new tab in your browser and follow the instructions.
3. Set the expiry to at least several months into the future.

4.2 Git Credential manager

The [GCM](#) is a platform agnostic credential manager (in English, that translates loosely to a *password manager*). Once it's installed and configured, Git Credential Manager is called by git and you shouldn't need to do anything special.

The next time you clone an HTTPS URL that requires authentication, Git will prompt you to log in using a browser window. You may first be asked to authorize an OAuth app. If your account or organization requires two-factor auth, you'll also need to complete the 2FA challenge.

Once you've authenticated successfully, your credentials are stored in the macOS keychain and will be used every time you clone an HTTPS URL. Git will not require you to type your credentials in the command line again unless you change your credentials.

4.2.1 GCM install

For Windows users it can be installed by selecting this option during the installation wizard, see Section 2.2.3, step 6.

For macOS, use homebrew again, specifically:

```
brew install --cask git-credential-manager
## ==> Downloading https://formulae.brew.sh/api/cask.jws.json
## #####
## ==> Downloading https://github.com/git-ecosystem/git-credential-manager/releases/download,
## ==> Downloading from https://objects.githubusercontent.com/github-production-release-asse
## #####
## ==> Installing Cask git-credential-manager
## ==> Running installer for git-credential-manager with sudo; the password may be necessary
## Password:
## installer: Package name is Git Credential Manager
## installer: Installing at base path /
## installer: The install was successful.
## git-credential-manager was successfully installed!
```

4.2.2 GCM demo

Below I demo the process by cloning a private repository from my GitHub account.

```
192-168-1-100:tmp mark$ git clone https://github.com/maj-biostat/wisca_2.git
Cloning into 'wisca_2'...
info: please complete authentication in your browser...
```

at this point the following window is launch by GCM:

selecting **Sign in with your browser** the following will launch in your default browser (Chrome, Safari, etc)

selecting **Authorize git-ecosystem** will result in

at which point you use the 2-factor authenticator tool (I use google authenticator) to respond with an authentication code.

Looking back at the terminal, the following output can be observed, which details the repository being cloned.

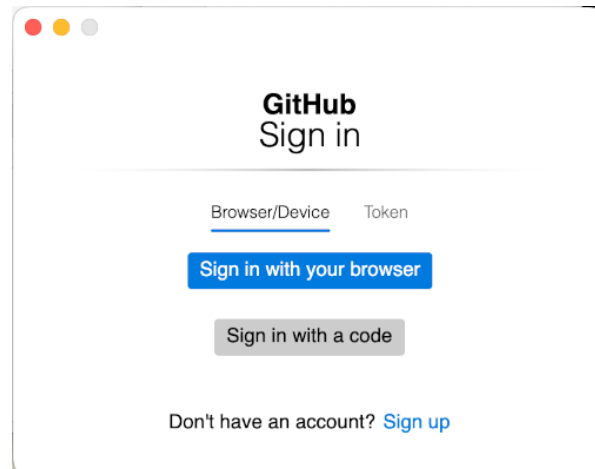


Figure 4.1: GCM

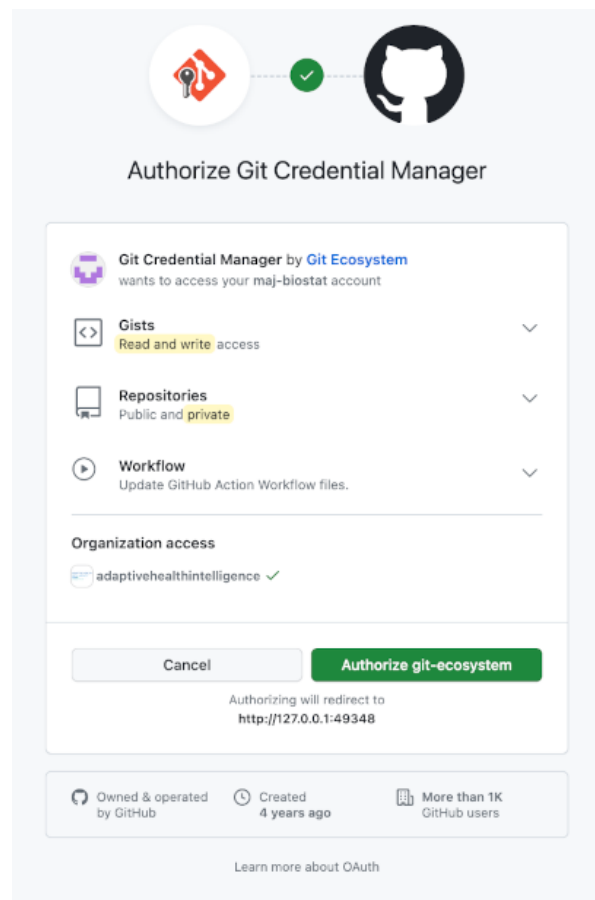


Figure 4.2: Sign in with browser



Confirm access



Signed in as @maj-biostat



Authentication code

Verify

Open your two-factor authenticator (TOTP) app or browser extension to view your authentication code.

Having problems?

- [Use your password](#)

Tip: You are entering [sudo mode](#). After you've performed a sudo-protected action, you'll only be asked to re-authenticate again after a few hours of inactivity.

[Terms](#) [Privacy](#) [Docs](#) [Contact GitHub Support](#)

Figure 4.3: Sign in with browser

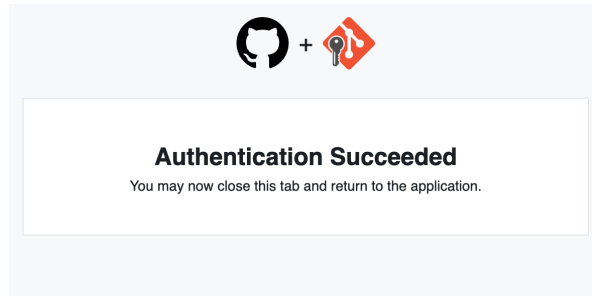


Figure 4.4: Auth success

```
remote: Enumerating objects: 297, done.
remote: Counting objects: 100% (297/297), done.
remote: Compressing objects: 100% (156/156), done.
remote: Total 297 (delta 148), reused 284 (delta 137), pack-reused 0
Receiving objects: 100% (297/297), 7.85 MiB | 2.13 MiB/s, done.
Resolving deltas: 100% (148/148), done.
```

Finally, you will receive an email of this sort:

Hey maj-biostat!

A first-party GitHub OAuth application (Git Credential Manager) with gist, repo, and workflow permissions. Visit <https://github.com/settings/connections/applications/0120e057bd645470c1ed> for more info.

To see this and other security events for your account, visit <https://github.com/settings/security>.

If you run into problems, please contact support by visiting <https://github.com/contact>.

Thanks,
The GitHub Team

On repeating this process a second time, all the authentication works in the background and there will be no need to go through various authentication handshakes again.

The same process applies irrespective of whether you are using [GitHub.com](https://github.com) or the [USyd GitHub Enterprise Server](#). However, it is advisable to get this working in GitHub first and then work on getting it to work in the USyd GitHub Enterprise Server.

The transition from the old authentication approach has (so far) proved completely seamless for macOS. It will be interesting to see what happens for the Windows platform.

4.2.3 GCM configuration (advanced only)

You can view the current credential manager by running the following commands:

```
git config --local credential.helper
git config --global credential.helper
# /usr/local/share/gcm-core/git-credential-manager
git config --system credential.helper
```

Of the local, global and system, the first one checks the local repository config, the second is your ~/.gitconfig, and the third is based on where git is installed. Note that only one credential help is configured in the above example.

In some circumstances you may need to reconfigure things. If you have to start from scratch, the following may be useful:

```
git config --local --unset credential.helper
git config --global --unset credential.helper
git config --system --unset credential.helper
```

For windows users check the contents of the credential manager. This can be accessed via Control Panel » All Control Panel Items » Credential Manager or by simply typing Credential Manager in the Windows task bar. Under generic credentials you should see the git entries.

4.3 GitHub CLI

In the day to day grind, having to deal with GitHub through its Web interface can be a little cumbersome. You can obviate having to interact with GitHub through the browser by using the [GitHub CLI](#). This tooling allows you to review, create and manage your repositories from the comfort of your commandline. You can think of it as an extension of git that allows you to invoke the GitHub specific functionality.

The extremely term gh CLI manual can be found [here](#).

For Windows users, you can pick up the latest Signed MSI executables from the [release page](#).

For macOS, use `homebrew`:

```

brew install gh
## ==> Downloading https://formulae.brew.sh/api/formula.jws.json
## #####
## ==> Downloading https://formulae.brew.sh/api/cask.jws.json
## #####
## ==> Downloading https://ghcr.io/v2/homebrew/core/gh/manifests/2.37.0
## Already downloaded: /Users/mark/Library/Caches/Homebrew/downloads/331c0b76fd34aa97342efa0
## ==> Fetching gh
## ==> Downloading https://ghcr.io/v2/homebrew/core/gh/blobs/sha256:a8c21e08d77963c2d12102ae
## Already downloaded: /Users/mark/Library/Caches/Homebrew/downloads/d0e6a3f8f7a4b138b36484e
## ==> Pouring gh--2.37.0.arm64_ventura.bottle.tar.gz
## ==> Caveats
## Bash completion has been installed to:
##   /opt/homebrew/etc/bash_completion.d
## ==> Summary
##   /opt/homebrew/Cellar/gh/2.37.0: 191 files, 44.2MB
## ==> Running `brew cleanup gh`...
## Disable this behaviour by setting HOMEBREW_NO_INSTALL_CLEANUP.
## Hide these hints with HOMEBREW_NO_ENV_HINTS (see `man brew`).

```

4.3.1 GitHub CLI authentication

In order to make use of `gh` we need to go through another round of authentication setup. To do this, go to the terminal and run:

```

gh auth login
## ? What account do you want to log into? GitHub Enterprise Server
## ? GHE hostname: github.sydney.edu.au
## ? What is your preferred protocol for Git operations? HTTPS
## ? Authenticate Git with your GitHub credentials? Yes
## ? How would you like to authenticate GitHub CLI? Login with a web browser

```

For additional information, see [gh auth --help](#).

In order to use `gh` with `github.com` directly you need to authenticate for that platform too. Repeat the above, but now the responses look like this:

```

gh auth login
## ? What account do you want to log into? GitHub.com
## ? What is your preferred protocol for Git operations? HTTPS
## ? Authenticate Git with your GitHub credentials? Yes
## ? How would you like to authenticate GitHub CLI? Login with a web browser

```

You are nearly set. You can verify that what you have configured worked via:

```
gh auth status
## github.sydney.edu.au
##   Logged in to github.sydney.edu.au as mjon7053 (keyring)
##   Git operations for github.sydney.edu.au configured to use https protocol.
##   Token: gho_*****
##   Token scopes: gist, read:org, repo, workflow
##
## github.com
##   Logged in to github.com as maj-biostat (keyring)
##   Git operations for github.com configured to use https protocol.
##   Token: gho_*****
##   Token scopes: gist, read:org, repo, workflow
```

However, for `gh` to work with the desired host you need to set an environment variable to tell `gh` which platform to use. On macOS, you can set this up easily with the following entries in the `.profile` shell initialisation script (or `.bash_profile` for those inclined).

```
gh-ent() {
    export GH_HOST=github.sydney.edu.au
}

gh-std() {
    export GH_HOST=github.com
}
```

On Windows, I have no idea how you are supposed to do the above in an easy manner. You may just have to resort to running

```
set GH_HOST=github.sydney.edu.au
```

or

```
set GH_HOST=github.com
```

each time you want to switch.

Now (on macOS) when you want to interrogate github.com repositories we can use the following commands.

Note

Do not worry about the meaning of the commands, this is just to establish that we have configured things correctly.

```
gh-std
gh repo list
```

```
## Showing 30 of 185 repositories in @maj-biostat
##
## maj-biostat/misc-notes           info for manjaro/arch linux setup
## maj-biostat/wisca_2             Revised approach to antibiogram
## maj-biostat/motc.run            Simulation for motivate c trial
## maj-biostat/motc.sim            Stan models for motc
## maj-biostat/motc.stan           Demo using Quarto to render to word document
## maj-biostat/quarto_demos_basic  Dose response models in stan
## maj-biostat/BayesDRM
## maj-biostat/motc.modproto
```

and for the USyd Enterprise GitHub Server, use:

```
gh-ent
gh repo list
```

```
Showing 12 of 12 repositories in @mjon7053
```

```
mjon7053/motc-mgt           Monitoring statistics for Motivate-C study
mjon7053/fluvid.analysises   Analyses for fluvid coadministration study (COVID19 + FLU) vaccine
mjon7053/motc.sap           Motivate-C simulation report
mjon7053/motc-sim-report     Notes relating to the ROADMAP project.
mjon7053/roadmap-notes
mjon7053/mjon7053.github.io
```

Part II

Part 2 - Fundamentals

Now we will make a start with git. Initially we will focus on using git within the confines of your local machine. Yes, that's right, for the moment, we won't be using github at all. The point of this is to give you a chance to get to grips with the basic ideas. After the main concepts are bedded in, we will move to thinking about github, which is a whole new beast.

Part III

Aside.

```
git remote add origin https://github.com/career-karma-tutorials/ck-git gh repo create my-newrepo -public -source=. -remote=upstream -push
```

The first part of the one liner: `gh repo create my-newrepo` creates and names a repo in your account (note: ‘my-newrepo’ should be replaced by the repo name of your choice) The `-public` flag makes sure the repo is public (swap this for `-private` if necessary) The `-source=.` flag specifies the source directory to be pushed Finally, the `-remote=upstream` flag specifies the remote repository to which the local repository is going to be compared with when pushing i.e the ‘upstream’ default.

5 Repositories

5.1 Git repositories

A repository is the most basic component of git. It is where you store your files and each files history. Repositories can be public or private, can have multiple collaborators and can be stored locally (on your personal computer) or remotely (in the cloud hosted by a service provider like github).

Using git, you can create and configure repositories, add or remove files and review history of the files in the repository.

5.1.1 Initialisation

There are two basic ways of creating a repository on your local machine

1. Initialise a new repository
2. Clone an existing repository

We will look at initialising a new repository, cloning will come later. Run the following on your machine

```
cd ~
```

5.1.2 Repository components

5.1.3 Adding project files

6 Commit

6.1 Adding files to projects

6.2 Commit process

6.3 Tracking status

6.4 Reviewing commit history

7 Reviewing differences

7.1

8 Branch

8.1 Why branches?

8.2 What is a branch?

8.3 Special branch concepts

8.4 Creating a branch

8.5 Switching branches

8.6 Working on branches

9 Merge

9.1 Merge concepts

9.2 Merge processes

10 Tags

10.1

11 Reviewing history

11.1

Part IV

Part 3 - Collaboration

Remotes

Hosting git

About

Repository status

Details on github repository files, tags, commits follow:

```
Local:    main /Users/mark/Documents/project/misc-stats/get-going-with-git
Remote:   main @ origin (https://github.com/maj-biostat/get-going-with-git)
Head:     [803a376] 2023-11-02: reconfig to new naming
```

```
Branches:      2
Tags:          0
Commits:       23
Contributors:  1
Stashes:       0
Ignored files: 4
Untracked files: 20
Unstaged files: 0
Staged files:  0
```

Latest commits:

```
[803a376] 2023-11-02: reconfig to new naming
[3ee74ba] 2023-11-02: Start of git repo section
[2fbe117] 2023-11-01: Minor edits
[3557833] 2023-11-01: Spelling
[c5321da] 2023-11-01: extended clean and made about unnumbered
```