Mini Project

Course 1: Data Scientist's Toolbox - Code Summary

---

1. Setting Up R, RStudio, and Git

Install R and RStudio (Manual Step, No Code)

- Download R from https://cran.r-project.org/
- Download RStudio from https://posit.co/download/rstudio-desktop/

Install Git on macOS

xcode-select –install # Installs Git via Xcode command-line tools

Or use Homebrew (if installed):

brew install git

Verify installation:

git –version

Configure Git (Set Username and Email)

git config –global user.name "Yuemin2025" # Replace with your name git config –global user.email "your.email@example.com" # Replace with your email

Check settings:

git config –global –list

Check settings:

git config –global –list

2. Creating a New Project in RStudio

New R Project (e.g., datasciencecoursera)

In RStudio: File > New Project > New Directory > New Project Name: datasciencecoursera Location: ~/Desktop/datasciencecoursera Click "Create Project" ### Or Terminal (Optional): *cd ~/Desktop mkdir datasciencecoursera cd datasciencecoursera*

3. Version Control with Git (Initializing and Managing Repositories)

Initialize a Git Repository

git init ### Output: "Initialized empty Git repository in . . . /.git/" with "master" branch warning

Rename Branch to main (Optional, Recommended)

git branch -m main

Check Git Status

git status ### Shows untracked or staged files (e.g., HelloWorld.md)
git add # e.g., git add HelloWorld.md ### Or add all files: *git add .*

Commit Files

git commit -m "Descriptive message" # e.g., git commit -m "Add HelloWorld.md"

Set Remote (Link to GitHub)

git remote add origin https://github.com/Yuemin2025/datasciencecoursera.git

Verify:

git remote -v

Push to GitHub

git push -u origin main ### If PAT needed, generate it in GitHub (Settings > Developer settings > PAT)

Pull (If Conflicts)

git pull origin main –allow-unrelated-histories

4. Creating and Managing Markdown Files

Create HelloWorld.md

echo "## This is a markdown file" > HelloWorld.md ### Or in RStudio: File > New File > Text File, add "## This is a markdown file", save as HelloWorld.md

Add to Git and Push (Use Version Control Commands Above)

Example: *git add HelloWorld.md git commit -m "Add HelloWorld.md" git push origin main*

Create Practice.md (Additional Practice)

echo "## Testing Git" > Practice.md ### Add, commit, push as above

5. Forking a Repository (From datasharing)

Fork on GitHub (Manual Step, No Code) Go to https://github.com/jtleek/datasharing Click "Fork," select your account (Yuemin2025) Result: https://github.com/Yuemin2025/datasharing

Get Fork Link (For Submission)

https://github.com/Yuemin2025/datasharing Optional Clone Locally (If Editing Later) *git clone https://github.com/Yuemin2025/datasharing.git cd datasharing*

6. GitHub Authentication (PAT Setup in R, Optional)

Generate and Store Personal Access Token (PAT)

install.packages("usethis") usethis::create_github_token() # Opens GitHub, generate PAT with repo scope install.packages("gitcreds") *gitcreds::gitcreds_set()* Enter username and PAT

7. Additional Practice (From Waiting Period Suggestions)

Create Simple R Script (Practice.R)

Practice.R

x <- 1:10 mean_x <- mean(x) print(paste("Average:", mean_x)) Save in ~/Desktop/datasciencecoursera, add to Git, push.

R Markdown Mini-Project (MiniProject.Rmd)

Simple Analysis

```r
data(mtcars)
mean_mpg <- mean(mtcars$mpg)
print(paste("Average MPG:", mean_mpg))
```

```
## [1] "Average MPG: 20.090625"
```