

CPP 528 – Group Project Instructions

The project for this course is designed as an opportunity to practice project management skills. It is broken into six steps, leaving one week at the end of the semester for a final round of revisions, documenting the process, and house-cleaning in your GitHub repo.

- Week 1: Neighborhood Revitalization Background
- Week 2: Construct Measures of Neighborhood Stability
- Week 3: Descriptive Analysis
- Week 4: Predictive Analysis
- Week 5: Merge Program Data
- Week 6: Estimate Program Impact
- Week 7: Finalize Project Repository and Report

Each week you will work on one step of the analysis as a lab. These labs should be completed and submitted individually (you are still allowed to collaborate on labs but submit your labs individually). Your team can then synthesize a final chapter from components of team member files.

The group grade will primarily focus on the integration of content into a clean final report format that makes all of the steps in the process transparent and easily reproducible.

Evaluation Report: Table of Contents

You can think about the final deliverable as sections in a report. The table of contents will be as follows:

1. Executive Summary (2 pages)
 - a. Overview / Research Question
 - b. Program Details
 - c. Data
 - d. Methods
 - e. Results
2. Part 1 – Neighborhood Change
 - a. Metrics
 - i. Data sources
 - ii. Median Home Value
 - iii. Neighborhood Health
 - iv. Gentrification
 - b. Descriptive Analysis of Neighborhood Change
 - i. Average change in MHV from 2000 to 2010
 - ii. Average change in neighborhood metrics
 - c. Predicting Change Based on 2000 Neighborhood characteristics
 - i. Neighborhood demographics
 - ii. Neighborhood health metrics
 - iii. Spatial characteristics (population density, adjacent tracts)
3. Part II – Evaluation of Tax Credits
 - a. Overview of Programs
 - i. New Market Tax Credits
 - ii. Low Income Housing Tax Credits
 - iii. Data sources
 - b. Descriptive Statistics
 - i. Dollars given out
 - ii. Characteristics of neighborhoods that received them
 - iii. Characteristics of those that did not
 - c. Predictive Analysis
 - i. Aggregate credits given between 2000 and 2010
 - ii. Update models from 2-C adding tax credit amounts
4. Part IV – Results and Conclusion

Note that there are some sections like “Overview of Programs” that require a succinct narrative for context. But most of the sections will focus on describing the methodology to the audience and walking them through the process of generating results. These chapters will have more of a tutorial or code-through feel than a final report tone. The goal of those sections is to make the methodology as transparent as possible and to make it easy for others to reproduce the work and extend it.

Grading Rubric

Your grade will largely be based upon your demonstration of mastery of project management principles and your ability to implement the project steps.

Your deliverable will be a report packaged as a Jekyll website hosted on GitHub, and all of the components necessary to replicate results (data, code, documentation of packages, etc).

You will be allocated points based upon your performance on the following criteria:

FINAL DELIVERABLES [50 points]

Project Website in Jekyll [15 points]

You will create an *index.md* file to serve as the landing page for the report, and link to report chapters from that document. Individual chapters should be stored as separate RMD files (rendered as HTML files for the report). You can link to rendered HTML files by referencing them in the proper subdirectories:

- <http://site-url.com> (*will load the index file*)
- <http://site-url/analysis/file-name.html> (*file in analysis subdirectory*)

- Website is active and live
 - All links work properly
- GitHub pages template is clean and effective
 - Custom CSS used for consistent report style
- Landing page includes:
 - a table of contents
 - a link to files on the GitHub repo with description of content
 - replication instructions (software needed, how to access files, etc)
 - license info
- About Us page

Data [10 points]

- Original data sources are documented and raw data used in the project has been archived
- Data steps clearly walk through the process of creating new variables, cleaning data, and joining tables.
 - When appropriate create an RMD describing the process of joining and cleaning datasets that can be referenced in the report
- README in the data folder that explains the organization of the files

Analysis [15 points]

- Quality of models and presentation of the steps to produce results:
 - Chapter on metrics
 - Chapter on descriptive analysis of neighborhood change
 - Chapter predicting change with neighborhood characteristics
 - Chapter describing tax credit programs (NMTC & LIHTC)
 - Predictive models after adding tax credit programs
- Data manifest:
 - A table that reports the initial sample size in the raw dataset, the sample size in the final model, and explain where and why each observation was dropped
- All results are easily reproducible
 - Packages reported
 - Clear description of which data is used for each model
 - Appropriate references to detailed data step files when helpful
 - Functions and arguments are explained when they are not clear
 - Specialized packages or functions
 - Custom functions

Portability [10 points]

- Can anyone run project code on their computer without having to change working directories or other settings?
- Note that Windows and Mac use different styles for paths (forward vs backwards slash) so functions like `here()` can help avoid those issues.
- Report the R environment and package versions used to create the analysis
 - Consider adding a private package library with packrat

TEAM PROCESS [10 points]

Utilization of Kanban Boards

When used correctly, a manager should never have to ask for a progress report. The Kanban board should show exactly what people are working on at the current moment, and progress thus far on the project.

- Project broken into appropriate-sized steps (cards)
- Steps broken into task lists
- Each card is assigned to at least one team member
- Card status is updated weekly

GitHub Commits

- Commits have useful names and clear descriptions

Code Review

- Each member has managed the code review process at least once

Git Ignore Files and Secret Passwords

- Add file names and file types to the gitignore file to prevent them from syncing to the team repository
 - Add the *.Rhistory* file to the list
- Never store passwords on GitHub!
 - Place passwords (or API keys) into a file called “password.R” and at the point in the program where you need to password you can type: `source("password.R")`
 - Add the file “password.R” to gitignore, and it won’t be synced. That means you can all keep separate passwords in a file with the same name on your own computers, and the code will run fine on each person’s machine

PROJECT ORGANIZATION [15 points]

The ideal directory structure makes it possible for a random stranger to find files easily, even if they have never seen your project before. The “stranger” is most likely you six months from now when you are trying to find code that you want to re-use or update your project and you have no recollection which files are the final working versions and which of the 100 datasets you created are used in the final program.

Folder names:

- data
 - raw-data
 - wrangling (data steps used to construct the final dataset)
 - rodeo (data used in the final analysis)
- analysis
- report-templates
- assets
 - images
 - css
- r-package (optional if you are creating a custom package)

Additionally:

- NEVER capitalize unless necessary or helpful for emphasis
- Use a dash instead of a space in file and folder names since spaces can cause a lot of problems in paths and are usually replaced with arbitrary characters
- Avoid locating your website files in a docs folder because you then cannot link to other HTML files outside of docs
- Add a README.md file inside of each folder, and use it to make notes on what the folder contains
- If files need to run in a specific order, or they represent things like chapters of a book, consider naming them something like step-01, step-02, etc.

File names:

- Consistent naming throughout project, including rules for capitalization and dates
- When order matters (for example steps in analysis or chapters of a report) the file order matches order in which they should be run
 - Effective use of leading zeros to maintain proper file order (09, 10, 11, ...)
- Clear system to keep track of the current version and archive revisions

It is helpful to have only the current version of a file in the project directory, and move all old versions to an “old-files” or “archive” folder if you need to reference them. If you have multiple versions of the same file in your project folder there is a high probability that one team member accidentally uses the wrong version (especially if they have not synced their project directory). It also prevents you from reading old versions of the file into a script.

DOCUMENTATION [25 points]

README Files [10 points]

Main README.md in the repository

- Overview of the project and the code
- Links to appropriate pages in the report
- Be kind and include convenient code to install all packages you use in the report
- If you write a custom package, include the command to install the package from GitHub so it can be easily copied and pasted
- Contact information for people that will be maintaining the project
- If you include a shiny app, include the launch from GitHub command in the README file so that it is accessible

Sub-folder README files

- Should explain the purpose and organization of the directory
- Include instructions for using the files in the folder

Readable Code [5 points]

- Everyone follows the same code style
 - Clean and consistent feel
 - Indents used to group code appropriately
- Objects follow consistent naming conventions
 - Data sets are nouns, functions are verbs
- Code is commented appropriately throughout the files

Citations [5 points]

- All sources are appropriately cited
- Bibtex citation database is created for the project
- R Markdown generates a bibliography automatically
 - Include a citation list at the end of each chapter

License [5 points]

- Select an appropriate open-source license for the project and include it in the repository.
- Add a LICENSE.md file on GitHub and include the appropriate licensing info

Citation Integration in RMD:

https://rmarkdown.rstudio.com/authoring_bibliographies_and_citations.html

[HTML] [One-click science marketing](#)

[HTML] [nature.com](#)

[M.Fenner](#) - Nature materials, 2012 - nature.com

When supply exceeds demand, marketing becomes vitally important. Marketing—a "set of processes for creating, communicating, and delivering value to customers" 1—helps companies distinguish themselves from their competitors. It is usually seen as a business ...

☆ 99 Cited by 8 Related articles All 6 versions

[book] [Advances in electronic marketing](#)

[L.Clarke](#), [T.Flaherty](#) - 2005 - igi-global.com

... a Free IGI Global Library Account to Receive a 25% Discount on All Purchases. Exclusive benefits include one-click shopping, flexible ... of **Marketing Education**, **Marketing Education**, **Journal of Internet Commerce**, and the **Journal of the Academy of Marketing Science**

☆ 99 Cited by 70 Related articles All 3 versions

[Online advertising and promotion: Modern technologies for](#)

[P.Hanafizadeh](#), [M.Behboudi](#) - 2012 - igi-global.com

... an Additional 5% Discount on All Purchases. Exclusive benefits include one-click shopping, flexible ... students who study **marketing**, **advertising**, **business management**, **information systems**, and **decision sciences** at **Arizona State University** is an Associate Professor of Information Systems and Decision Sciences at **Arizona State University**

☆ 99 Cited by 59 Related articles All 3 versions

[Mobile marketing management: marketing objectives, types, and implementation techniques](#)

[K.Pousttchi](#), [DG.Wiedemann](#) - ... of **Research on Mobile Marketing** ..., 2010 - igi-global.com

... for the first, second, and third editions with the purchase of the **Encyclopedia of Information Systems and Decision Sciences** ... Library Account to Receive a 25% Discount on All Purchases. Exclusive benefits include one-click shopping, flexible ... "Mobile Marketing Management: Marketing Objectives, Types, and Implementation Techniques"

☆ 99 Cited by 15 Related articles All 3 versions

[e-marketing: Towards a Conceptualisation of a New Marketing Philosophy](#)

X Cite

MLA Fenner, Martin. "One-click science marketing." *Nature materials* 11.4 (2012): 261-263.

APA Fenner, M. (2012). One-click science marketing. *Nature materials*, 11(4), 261-263.

Chicago Fenner, Martin. "One-click science marketing." *Nature materials* 11, no. 4 (2012): 261-263.

Harvard Fenner, M., 2012. One-click science marketing. *Nature materials*, 11(4), pp.261-263.

Vancouver Fenner M. One-click science marketing. *Nature materials*. 2012 Apr;11(4):261-3.

BibTeX EndNote RefMan RefWorks

--- In file called: bibliography.bib

```
@article{fenner2012one,  
  title={One-click science marketing},  
  author={Fenner, Martin},  
  journal={Nature materials},  
  volume={11},  
  number={4},  
  pages={261--263},  
  year={2012},  
  publisher={Nature Publishing Group}  
}
```

```
---  
title: "Sample Document"  
output: html_document  
bibliography: bibliography.bib  
---
```