Software Design for Data Science

Technology Reviews

Melissa Winstanley
University of Washington
February 3, 2025

W

A Project Situation

- You need to write code to do <X>
- It's really hard to write the code for <X> from scratch
- You search online for existing Python packages
- You find that you could use packages <A>, <B>, or <C>
- Which do you use?

What are technology reviews?

- Package research
- Package evaluation for deployment in a project
- Contents
  - Background
  - Choices comparison
  - Decision & why
  - Drawbacks

Why technology reviews?

- Helps you make a good choice between different package options!
    - Solves the use case
    - Avoids pitfalls

Your Technology Review Format: Markdown Report + Demo

Markdown Section 1: Background & Use Case

- What's your application?
- What use case are you trying to solve?
- What about it needs a Python library?

Markdown Section 2: Python Package Choices

- Find 2-3 Python libraries that potentially address your use case
- Describe the libraries
  - Name
  - Author (if known - eg a particular company)
  - Brief summary of library's purpose (check out the README!)

Markdown Section 3: Package Comparison

- Side-by-side comparisons of the technologies (pros/cons)
  - *This will require that you actually install and use the technologies!*
- Things to think about, as a starting point:
  - Addresses all requirements of your project
  - Compatibility with other elements of the project
  - Ease of use
  - Computational efficiency for the scale of data you'll be using
  - No software bugs that will impair your use case
  - Availability of relevant examples
- Look at open issues on GitHub

Markdown Section 4: Your Choice

- What did you choose to use?
- Why did you choose it?

Markdown Section 5: Drawbacks/Remaining Concerns

- Is there anything in the "con" list to be concerned about?
- Will you have to mitigate any drawbacks?
- What will you watch out for?

Package Demonstration

- Screen recording or you using the package
- Show us how the package might work for you
- Could be:
  - Input => run code => output
    - Processing libraries
    - ML libraries
  - Visual demonstration
    - Map
    - Web app
    - Graphing/visualizations
    - Command line interaction
  - etc

Technology Review: Summary

Markdown document + screen recording demo, due by 2/24,
with in-class demonstration for instructors

But I STRONGLY encourage you to do the work sooner!

1. Background & Use Case
2. Python Package Choices
3. Package Comparison
4. Your Choice
5. Drawbacks/Remaining Concerns
6. Demo of the Package

Pick an interesting use case - NOT something that requires pandas

More details and instructions on the course website

https://uwdata515.github.io/projects.html#technology-review