# SoftDes Midterm Proposal Spring 2025

In the project proposal, we’ll ask you and your team to provide key information about your project. Your answers to the proposal questions serve several important purposes:

* Telling us what your project is about.
* Justifying why you and your team are well-suited to carry out this project.
* Showing us that you’ve done a few tasks to set up and plan your project.

## Project Overview

In one sentence, what is your project about?

(replace this with your answer)

What are the research questions you hope to answer by doing this project?

(replace this with your answer)

## Repository

Please make sure that you have created a new repository from the template repository, https://github.com/olincollege/python-project-template. Make sure that the repository fulfills the following criteria:

* The repository is hosted in olincollege and not in a personal account.
* The name of the repository is properly formatted: it does not contain underscores (\_), the name of the year/semester, team members, instructors/CAs, course, or “midterm project”.
* All team members are collaborators on the repository.
* All team members have cloned the repository to their machines.
* The softdes-2025-01-assistant and softdes-2025-01-teaching teams have permissions on your repository.

## Data Retrieval

Data overview: In a few sentences, what data are you collecting? (This should be understandable by someone who doesn’t know Python.)

(Replace this line with your answer)

Justification: How will this data help you answer the questions above?

(Replace this line with your answer)

Data Source: Where are you getting your data from?

(Replace this line with your answer)

Data retrieval method: Is your team retrieving data through API calls, Web scraping, or a Python library? (Then, answer the relevant detailed questions below, depending on your method)

(Replace this line with your answer)

### API Calls

API key preparation: Please be prepared to show your API key (we’ll review this with you at the check-in; **do not include it here or commit this to your GitHub repository under any circumstances**).

Sample code: Please be prepared to show a Python function that uses that key to download some of your desired data (we’ll review this with you at the check-in)

Description: Please provide samples of **all** data you plan to retrieve (at a minimum, sample output from every API endpoint you plan to use), **OR** data types of **all** data you plan to retrieve, by field/column (e.g., saying “JSON” or “data frame” is not specific enough)

(Replace this line with your answer)

### Web Scraping

Pages: Please provide a list of the webpages that you plan to scrape from (if you are going to replace one part of the URL with different values, like page=1, page=2, etc., then feel free to describe those - as long as we know what the actual URL for one of those pages would be).

(Replace this line with your answer)

Webpage sample: Please provide a sample webpage downloaded with wget or curl: you can do this with wget <URL> or curl <URL> --output <file>, with <URL> being replaced by the URL you want to scrape and <file> being replaced by the file you want to download the webpage to (wget will choose the file name for you automatically).

(Replace this line with your answer)

Data sample: Please provide a sample of the data you would extract from the webpage and where it is located in the page.

(Replace this line with your answer)

### Python Library

Library: What library do you plan to use?

(Replace this line with your answer)

API key preparation: If the library uses an API key, please be prepared to show your API key (we’ll review this with you at the check-in; **do not include it here or commit this to your GitHub repository under any circumstances**).

Sample code: Please be prepared to show a Python function that uses that key to download some of your desired data (we’ll review this with you at the check-in)

Description: Please provide samples of **all** data you plan to retrieve (at a minimum, sample output from every API endpoint you plan to use), **OR** data types of **all** data you plan to retrieve, by field/column (e.g., saying “JSON” or “data frame” is not specific enough)

(Replace this line with your answer)

## Analysis and Visualization

Data processing: what computations will you do with your data that will NOT appear in visualizations? (e.g., taking the average of something, filtering out invalid values or outliers)

(Replace this line with your answer)

Visualization descriptions: What data, or things computed from your data will you show in visualizations?

(Replace this line with your answer)

Sample visualizations: Please provide examples of your three visualizations below.

(Replace this line with your answer)

Narrative: What points will these visuals make?

(Replace this line with your answer)

Justification: How will this visualization help answer your research questions in ways that the data by itself will not?

(Replace this line with your answer)

## Planning

Collaboration style: How are you dividing the work for this project? (Will you write code together, or split the work up and merge your parts later?)

(Replace this line with your answer)

Task allocation: Who is doing what between now and the next deadline?

(Replace this line with your answer)

Backup plans: How will you change your plan for the project if you are not able to do all that you had planned?

(Replace this line with your answer)