

Phase 1: Project Familiarization and Setup (2 weeks)

1. Task: Understand PubChem API and Python Basics

- **Duration:** 1 week
- **Goal:** Familiarize with the PubChem API documentation and basic Python programming, focusing on handling JSON data.
- **Resources:**
 - [PubChem API Documentation](#)
 - [Python for Beginners](#)
 - [JSON in Python](#)

2. Task: Install and Configure Python Environment

- **Duration:** 1 week
- **Goal:** Set up Python environment with necessary tools (install Python, pip, requests library, and a code editor).
- **Output:** A working Python environment with the `requests` library installed.
- **Resources:**
 - [Installing Python](#)
 - [Installing Packages with pip](#)
 - [Python Requests Library](#)

Phase 2: Script Development and Data Retrieval (6 weeks)

1. Task: Design Script Structure

- **Duration:** 1 week
- **Goal:** Outline the structure of the script, identifying key functions and data flow.
- **Output:** A flowchart or diagram of the script's structure, including function descriptions.
- **Resources:**
 - [Python Script Structuring](#)
 - [Diagramming Tools](#)

2. Task: Implement Basic API Request Function

- **Duration:** 1 week

- **Goal:** Develop the function that takes a SMILES string as input and sends a request to PubChem.
- **Output:** A Python function that successfully sends a request and retrieves data.
- **Resources:**
 - [Handling API Requests with Requests Library](#)
 - [SMILES Strings and Chemical Information](#)

3. Task: Parse and Handle JSON Data

- **Duration:** 1 week
- **Goal:** Implement functionality to parse the JSON response from PubChem and extract relevant information.
- **Output:** A Python script that outputs parsed JSON data in a readable format.
- **Resources:**
 - [Parsing JSON in Python](#)
 - [Working with JSON Responses](#)

4. Task: Retrieve Specific Data Types from PubChem

- **Duration:** 3 weeks
- **Goal:** Retrieve the following specific data from PubChem using the API:
 - **Subtask 1:** Full description of the molecule.
 - **Subtask 2:** List of all names, identifiers, and synonyms.
 - **Subtask 3:** Commercial availability and suppliers.
 - **Subtask 4:** Computed properties (e.g., molecular weight, solubility).
 - **Subtask 5:** Toxicity or safety information.
- **Output:** A Python script that retrieves and displays each type of data.
- **Resources:**
 - [PubChem Data Descriptions](#)
 - [Retrieving Names and Identifiers](#)
 - [Computed Properties Guide](#)

Phase 3: Error Handling, Advanced Features, and Review (4 weeks)

1. Task: Handle Errors and Edge Cases

- **Duration:** 1 week
- **Goal:** Ensure the script handles invalid SMILES strings, network errors, and unexpected API

responses.

- **Output:** A robust Python script with error-handling mechanisms.
- **Resources:**
 - [Error Handling in Python](#)

2. Task: Implement Additional Features

- **Duration:** 2 weeks
- **Goal:** Add features like saving JSON output to a file, or retrieving multiple types of data (e.g., bioactivity, patents).
- **Output:** An enhanced Python script with additional features.
- **Resources:**
 - [Saving JSON Data to a File](#)
 - [Exploring PubChem Data Types](#)

3. Task: Write Detailed Documentation and Final Review

- **Duration:** 1 week
- **Goal:** Document the entire project, including how to use the script, code comments, and explanations of functions. Perform final testing and debugging to ensure all features work correctly.
- **Output:** A comprehensive user guide, code documentation, and a fully functional, bug-free Python script.
- **Resources:**
 - [Documenting Python Code](#)
 - [Effective Software Testing](#)

Phase 4: Final Submission and Wrap-up (2 weeks)

1. Task: Prepare Final Deliverables

- **Duration:** 1 week
- **Goal:** Compile the final script, documentation, and any other project materials.
- **Output:** A project package ready for submission.

2. Task: Submit Final Project and Receive Feedback

- **Duration:** 1 week
- **Goal:** Submit the project for review and implement any final feedback.
- **Output:** Finalized project, with feedback integrated (if necessary).

