Softsinc Internship Program 2025

Week 1 – Advanced Python & Data Handling

Topic 1: Python Programming – Deep Dive

Advanced Concepts:

- List/dictionary comprehensions
- Lambda functions & `map()`, `filter()`, `reduce()`
- Exception handling
- File and directory operations using `os` and `shutil`

Tasks:

1. Rebuild your CLI app to include:
    - Error handling (e.g., invalid input)
    - Modular structure (`functions.py`, `main.py`)
2. Build a mini **unit conversion CLI tool** using lambda and `map()`.
3. Create a Python script to:
    - Read a directory
    - List all `.txt` or `.csv` files
    - Copy them to a new folder

Challenge:
Build a **modular CLI toolkit** that includes 3 features: calculator, file organizer, and a password generator.

◆ Topic 2: OOP with Real-World Design

Concepts:

- Inheritance and polymorphism
- Magic methods (`__str__`, `__len__`, etc.)
- Encapsulation and abstraction

Tasks:

1. Create a `User` → `Intern` → `Mentor` class hierarchy.
2. Use inheritance to implement shared and unique behavior.
3. Add `__str__()` and validation logic.

🎯 **Challenge:**
Simulate a **chat app structure** using OOP: Users, Managers, and Messages (no real-time, just logic flow).

◆ Topic 3: Advanced File & CSV Handling

📚 **Concepts:**

- Handling CSV with Python (manually and via `csv`/`pandas`)
- Reading structured logs
- Writing formatted reports

 Tasks:

1. Parse a CSV file and:
    - Find rows with missing values
    - Count unique entries per column
2. Write a report to a `.txt` or `.csv` file with:
    - Summary stats
    - Timestamp of execution

🎯 **Challenge:**
Automate a weekly **log reader** that scans a directory for `.log` or `.txt` files, summarizes errors, and writes a report.

◆ Topic 4: Data Handling with Pandas – Extended

📚 **Concepts:**

- Advanced data filtering
- Grouping, sorting
- Handling missing and duplicate data
- Data type conversions

 Tasks:

1. Use `groupby()` to analyze a dataset (e.g., avg survival rate per gender in Titanic).
2. Filter a dataset using multiple conditions.
3. Drop missing and duplicate rows and explain the impact.

🎯 **Challenge:**
Create a **Jupyter Notebook** that:

- Loads any public dataset
- Cleans it thoroughly
- Explores 5+ different insights
- Includes plots (optional)

◆ Topic 5: Git & GitHub – Pro Workflow

📚 **Concepts:**

- Branching and merging
- GitHub Actions (basic intro)
- Markdown for README files

 Tasks:

1. Create at least 2 branches: `main` and `dev`.
2. Commit to `dev` daily, and merge to `main` at week-end.
3. Write a professional README with:
    - Project description
    - Task list
    - Demo screenshots

🎯 **Challenge:**
Configure a GitHub Action that runs `flake8` or any code linter (intro level CI/CD).