Midterm Project Report

Advanced Computer Programming

Weather Web Scraping with Python

Student Name : Reinhart Osfred

Student ID : 112021228

Teacher : DINH-TRUNG VU

2024-04

Chapter 1 Introduction

1.1 Github

1) Personal Github Account : Reinhartos

2) Group Github Account : https://github.com/advcomp-earthquake

3) Group Project Repository : https://github.com/advcomp-earthquake/earthquake-monitor

4) List of submitted files :

- weatherscraper
    - _pycache_
    - Spiders
        - _pycache
        - _init_py
        - example.py
        - weatherSpider.py
    - _init_py
    - items.py
    - middlewares.py
    - pipelines.py
    - settings.py
- scrapy.cfg
-

1.2 Topic

The topic of our group is Earthquake and my personal topic is scraping with Weather Data about earthquake

1.3 Project Overview

I have used the data class, pprint, and scrapy in my program. My program have extracted date, high temperature, low temperature, and weather prediction from website “the weather channel” on the page https://weather.com/weather/tenday/l/51aa95c773902d7614fdbac99637315ed8f99a0a0ec624f9b5b493b53a4d8dbc?unit=m .

Chapter 2 Implementation

2.1 Data Class Weather

Data class weather for know about date, high temperature, low temperature, and weather description.

2.2 Class Spider

Class weatherSpider it's a custom class name, so its exact structure and functionalities would depend on the specific requirements of the web scraping task it's designed for.

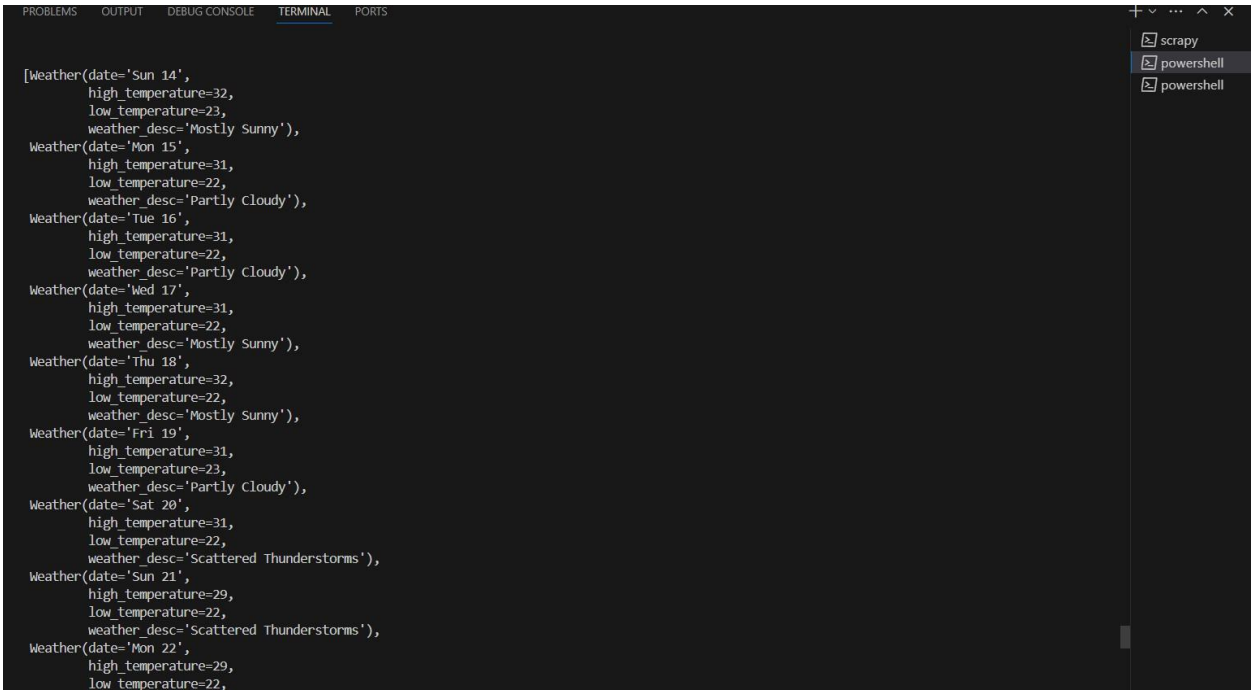
2.2.1 Method Parse

Parse refers to the process of extracting structured data from the raw HTML or XML content of a web page. Use for summary list, weather list, and weather item.

pprint for formatting complex data structures in a more readable and visually appealing way when printed to the console or output file.

Chapter 3 Results

3.1 Result



Chapter 4 Conclusions

After working on the midterm and learning the materials from Tronclass, I have learned how to do web scraping on a web page properly and how to use data class and pprint to improve my application.

The shortcomings that I have in my program, I don't use beautiful soap, pattern matching, etc. In the future, I will use beautiful soup and others to carry out my projects and become better at coding that I will create. Lastly, i would like to thank you to the professor who taught me this.