



CarrotDB and Analysis

[\[Project GitHub \]](#)

2023年/10月 - 2024年/12月

—

이건희 (イ・ゴニ / LEE KUNHEE)

sacredcrawler@gmail.com (business)

knutvikings@gmail.com (private-not business)

Seoul, Korea

Overview

평소 데이터의 수집, 수집된 데이터의 흐름과 관리에 관심이 많다. 더불어 데이터를 어떻게 활용하면 좋을 지에 대해서도 고민하는 것을 좋아한다. 그래서 이 모두를 실행해 볼 수 있는 재미있을 프로젝트를 진행하게 되었다.

Milestones

1. 실제 온라인 환경에서 떠다니는 정보를 수집한다.
2. 수집된 정보를 유효화된 데이터로 정제한다.
3. 정제된 데이터를 프로젝트 목적에 맞게 활용한다.

Description

Python, C++, SQL 을 메인 언어로 채택했으며, **Python & C++** 는 데이터 추출, 정제, 분석에 사용했고 **SQL** 은 데이터베이스 관리에 사용했다. 개발 환경으로는 **Google CoLab, Microsoft Visual Studio, SQLite** 를 이용했다.

Details and Goal

한국 내에서 가장 규모가 큰 중고 거래 시장 플랫폼은 '당근마켓'이다. 이 플랫폼에 매일 흐르고 있는 전국의 물건 현황 데이터를 자동화된 웹 스크래퍼를 통해 매일 수집한다. 수집된 데이터는 정제 과정을 거쳐 데이터베이스 파일에 저장되고, 해당 데이터베이스 파일은 **SQLite** 로 관리됨과 동시에 **Python & C++** 로 분석이 이루어진다.

이 프로젝트의 목표는, 실제 세계의 데이터를 직접 추출, 적재, 변형, 활용하는 경험을 쌓는 것과, 전체 과정 중 문제가 발생하면 어떤 방향으로 해결해 가면 좋을 지 숙고하는 능력을 기르는 것이 주(主)이다. 그 안에서의 목적, 즉 분석의 목적은 중고 거래 시장 플랫폼의 동향과 사회적 관계에 따른 구조적 특징을 파악하는 것이다.

Project Schema

I. Web Scraper

여기에 텍스트를 입력하세요 여기에 텍스트를 입력하세요 여기에 텍스트를 입력하세요
여기에 텍스트를 입력하세요 여기에 텍스트를 입력하세요.

II. Database

여기에 텍스트를 입력하세요 여기에 텍스트를 입력하세요 여기에 텍스트를 입력하세요
여기에 텍스트를 입력하세요 여기에 텍스트를 입력하세요.

III. Analysis

여기에 텍스트를 입력하세요 여기에 텍스트를 입력하세요 여기에 텍스트를 입력하세요
여기에 텍스트를 입력하세요 여기에 텍스트를 입력하세요.

IV. Management

여기에 텍스트를 입력하세요 여기에 텍스트를 입력하세요 여기에 텍스트를 입력하세요
여기에 텍스트를 입력하세요 여기에 텍스트를 입력하세요.