

서비스산업 데이터 분석가 - 서브 프로젝트

Index

1. 개요

- 프로젝트를 하는 이유
- 목적

2. 방향성

- 주제
- 데이터 관리

3. 평가지표

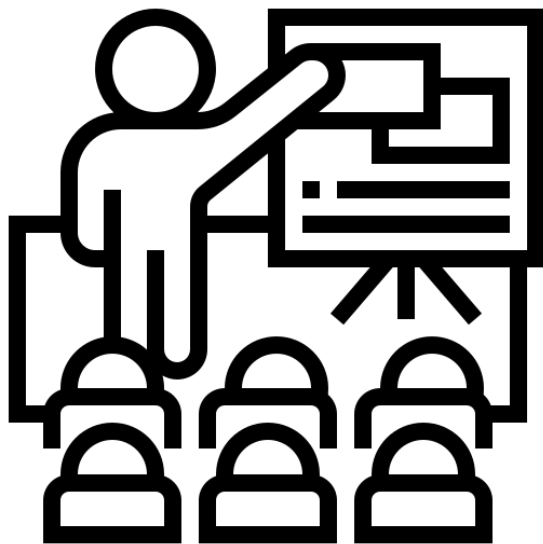
- 지표
- 방법

4. 발표

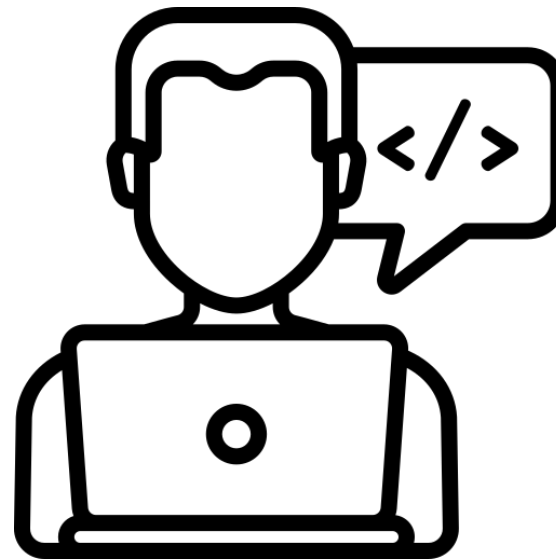
Overview

개요

개발자가 성장하는 방법



강의로 10 성장한다면,



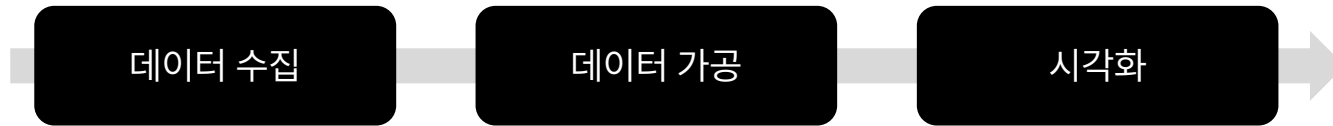
프로젝트로 90 성장할 수 있음

목적

- 데이터 수집, 전처리, 시각화 프로세스 이해
- 본 프로젝트의 프로세스는 데이터 분석 프로세스의 일부분
- 분석 프로젝트 : 데이터 수집/가공 → EDA (시각화) → 데이터 분석
- 데이터 수집 : ETL (Extract, Transform, Load)



- 본 프로젝트 : 데이터 수집 → 데이터 처리 → 시각화



앞으로 진행할 **데이터 분석 프로젝트**도 직접 **구축**하고, **구현**할 수 있도록 역량을 끌어올리는 것이 **최종 목표**

Direction

방향성

주제

- 관심있는 분야의 데이터 분석

프로젝트 방향성

- 크롤링한 데이터를 전처리 한 뒤, 시각화하여 인사이트를 도출
- 데이터 분석 README를 잘 작성하는 것도 나중에 포트폴리오로 활용 시 도움됨

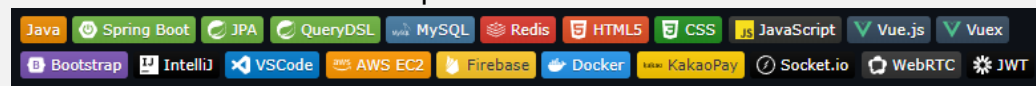
내부 방향성

- 지금 같이 교육받는 동료들 경쟁자라고 느끼지 말 것
(선의를 경쟁은 OK, 하지만 정보공유를 막는 순간 성장하기 힘들)
- 프로젝트 초반에 내용에 대해서 얘기해보며 좀 더 발전시킬 곳은 없는지 확인
- 프로젝트 범위를 넓게하기보단 깊게 할 것

README 구조 예시

프로젝트명

사용한 기술스택 (non-required)



성과

개요 - 프로젝트 설명, 주관사

분석 결과, 예측결과

인사이트 (너무 길지 않게 2~3문장)

분석 프로세스

1. 크롤링
 - 수집한 사이트, 목적
2. 전처리
 - 특별히 처리한 것
3. 시각화
 - 사용한 라이브러리, 시각화 내용

데이터 관리

- 파일 관리 시스템 : 하드디스크, 플로피디스켓, 외장하드, USB
- 데이터베이스 : 데이터베이스 서버
- 데이터의 volume을 보고 판단

코드의 모듈화

- 함수화
 - 특정 기능을 하는 코드들을 묶음
- 클래스화
 - 특정 속성, 메서드를 가지는 객체를 생성
- 지금 당장 잘하기보단 잘해지기 위해 계속해서 노력하는 것이 중요 !
- 한번 작성한 코드를 리팩토링하는 것도 실력 향상에 큰 도움이 됨 !

Metrics

평가지표

4가지 프로젝트 평가지표



목적성

활용 목적에 따라 적절한 주제를 선정하였는가?



데이터

수집한 데이터가 적절한가? 충분한가?



전처리

전처리 방법과 그 근거가 명확한가?



시각화

결과물이 이해하기 쉽고 인사이트를 얻는데 적합한가?

평가방법

- 강사 40% / 상호 60%
- 상호 평가는 미리 준비한 엑셀 파일 전달 예정

Presentation

발표

구성

- 주제
 - 데이터 수집 방법 & 수집 데이터 구조
 - 진행한 데이터 전처리
 - 시각화 결과 및 인사이트
- 발표 시간은 7분 이내