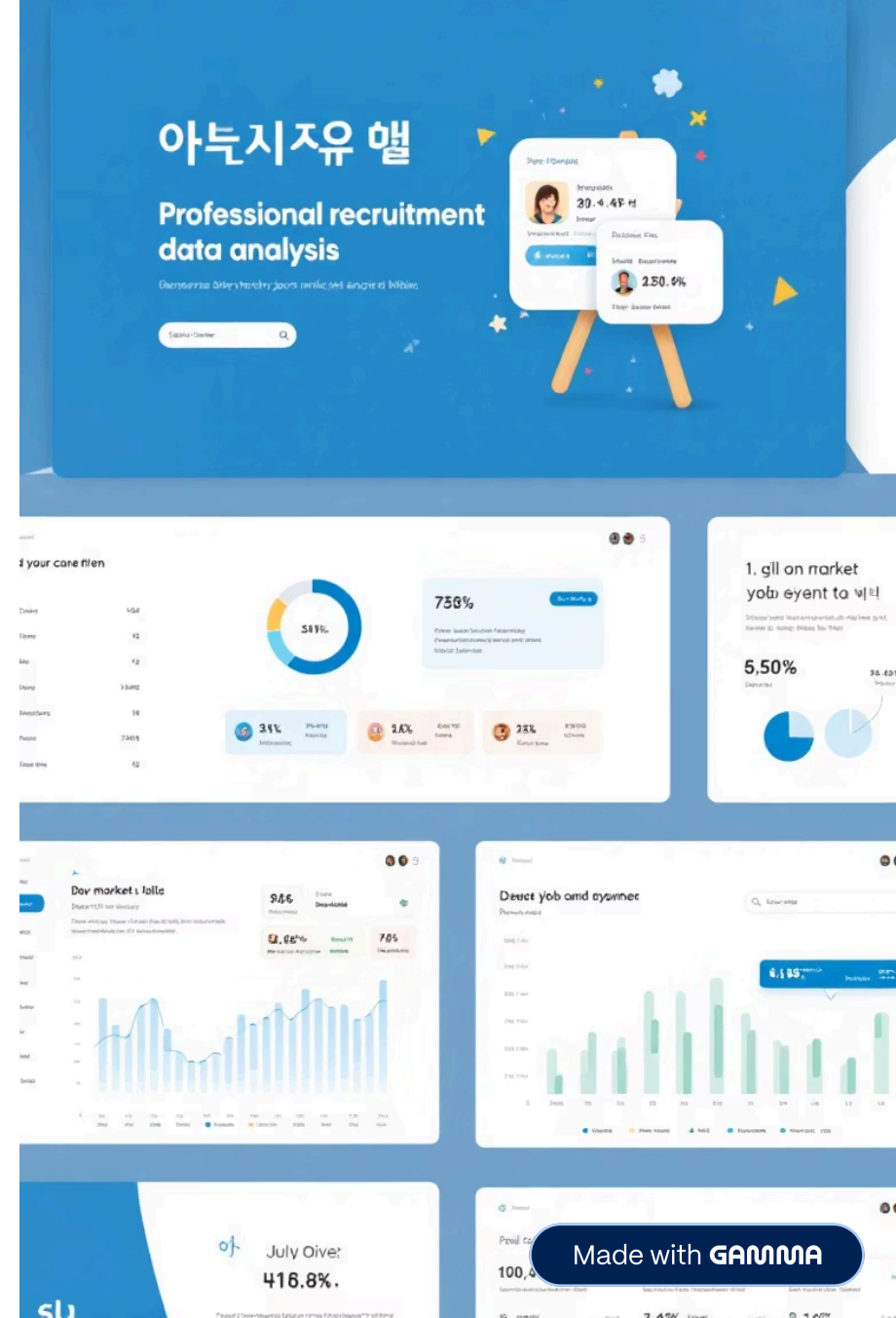


채용 정보 분석 플랫폼 기획서

프로젝트명: 채용 정보 분석 플랫폼

영준 작성자: 영준 홍



프로젝트 개요



안태경: 데이터 수집 및 전처리 담당

데이터 정제 파이프라인 구축 (중복/결측치 처리)

채용 공고 크롤링 및 API 연동

데이터베이스 설계 및 관리



홍영준: 데이터 분석 및 시각화 담당

EDA(탐색적 데이터 분석) 및 통계 분석

기술 키워드 NLP 처리 (KoNLPy, Soynlp)

자료 수집, test 진행, 개선안 제시



한명규: Streamlit 웹 개발 담당

Streamlit 기반 대시보드 UI/UX 기획 및 구현

인터랙티브 시각화 구현 (Plotly, Altair)



프로젝트 목표



"채용 시장 분석과 커리어 의사결정 지원 플랫폼" 구축을 목표로 합니다. 현재 기업과 회사의 직종은 많지만, 사용자가 직접 분석할 수 있는 부분이 라던가, 본인이 구인하는 모습은 회사가 아니라 본인이 직접 결정하는 부분이어서 추천 보다는 의사결정 존중 참고용 자료로 사용하려고 합니다.

데이터 수집 계획

수집 대상 데이터

- 수집 대상 사이트: 사람인(회사정보 all), 잡플래닛(평점)
- 기본 정보: 직무명, 회사명, 근무 지역, 고용 형태, 회사링크
- 분석 정보: 경력 요구 수준, 학력 및 자격요건 기준 분석
- 공고 일정 & 분석용 데이터: 공고 게시일 및 마감일

데이터 출처

- 사람인
 - URL: https://www.saramin.co.kr/zf_user/
 - 수집 방법: 웹 크롤링 (BeautifulSoup)
 - 대상 데이터: 포지션명, 회사명, 지역, 연봉, 경력 조건, 기술 키워드(본문 기반 추출), 공고 등록일/마감일
- 잡플래닛
 - URL: <https://www.jobplanet.co.kr/job>
 - 수집 방법: 웹 스크래핑
 - 대상 데이터: 기업 평점

Data analysis METHODOLOGY

EDA

Pre processing para analysis reports and collation for data 70% in how far as we're is really in internet gloves but the learning are offertile, an flicking test and access from the montal analyst. EDA, all people that heva colientity of pavemure.

Correlation and

Let the cordal in data co thatacually cementoce angrapee to be mating re and beclireling and on so

Let a ondy is and call the aerally porvarting are a pryoct sent us dille in s and ang or entserment fo using plak in colling lines

The machine learning multal evolves totk or m caldk in completions nen

Pre processing ends in data table.

MACHINE LEARNING MODELS

Neural N

Machine learn Clane is dltol one dlt modalely cenger ref for lce exollera for cor main cases con maratling in the stagers.

데이터 분석 방법론

전처리 목적

목표: 중복 제거, 노이즈 정제, 형식 통일

- 결측치 처리: 컬럼(링크, 회사링크, 제목, 회사, 직무 또는 중복 값 등) null 값, error 값 삭제 및 제외
- 이상치 처리: "해외", "경력" 튜는 숫자 값 제외
- 데이터 정규화: 등록일자, 기간, 직무 등 통일화 필요

탐색적 데이터 분석 (EDA)

- 지역별 채용 건수 분석
- 경력 요건별 분포 분석
- 학력 요건별 분포 분석
- 공고 마감 기간 분석
- 기업 배지(인증) 여부 분석
- 계약종류별 공고 수 분포 분석

상관관계 분석

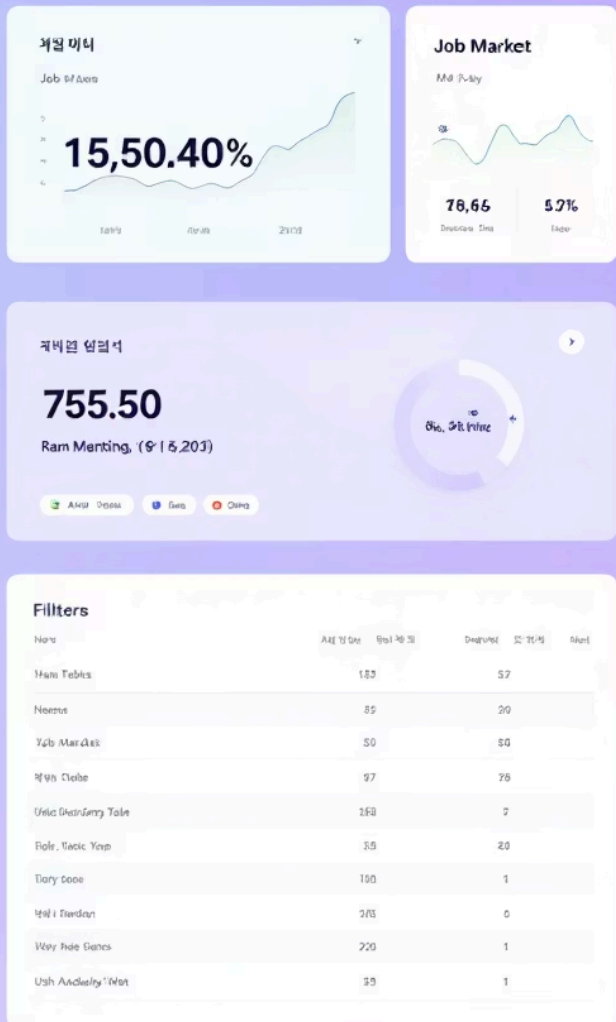
- 연봉과 직무의 상관 관계
- 기술 스택과 직무의 상관 관계
- 경력 수준과 직무의 상관 관계
- 연봉과 경력 수준의 상관 관계

머신러닝 모델 개발

- 연봉 예측 모델
- 직무 분류 모델
- 구직자-직무 매칭 시스템
- 채용 시장 분석 및 트렌드 예측
- 이상치 탐지 모델

웹 애플리케이션 구현

Job market analysis



메인 대시보드

- 지역별 채용 공고 수 시각화
- 지역별 채용 공고 인터랙티브 차트
- 기업 채용 공고 순위
- 인기 직무/기술 스택 추이

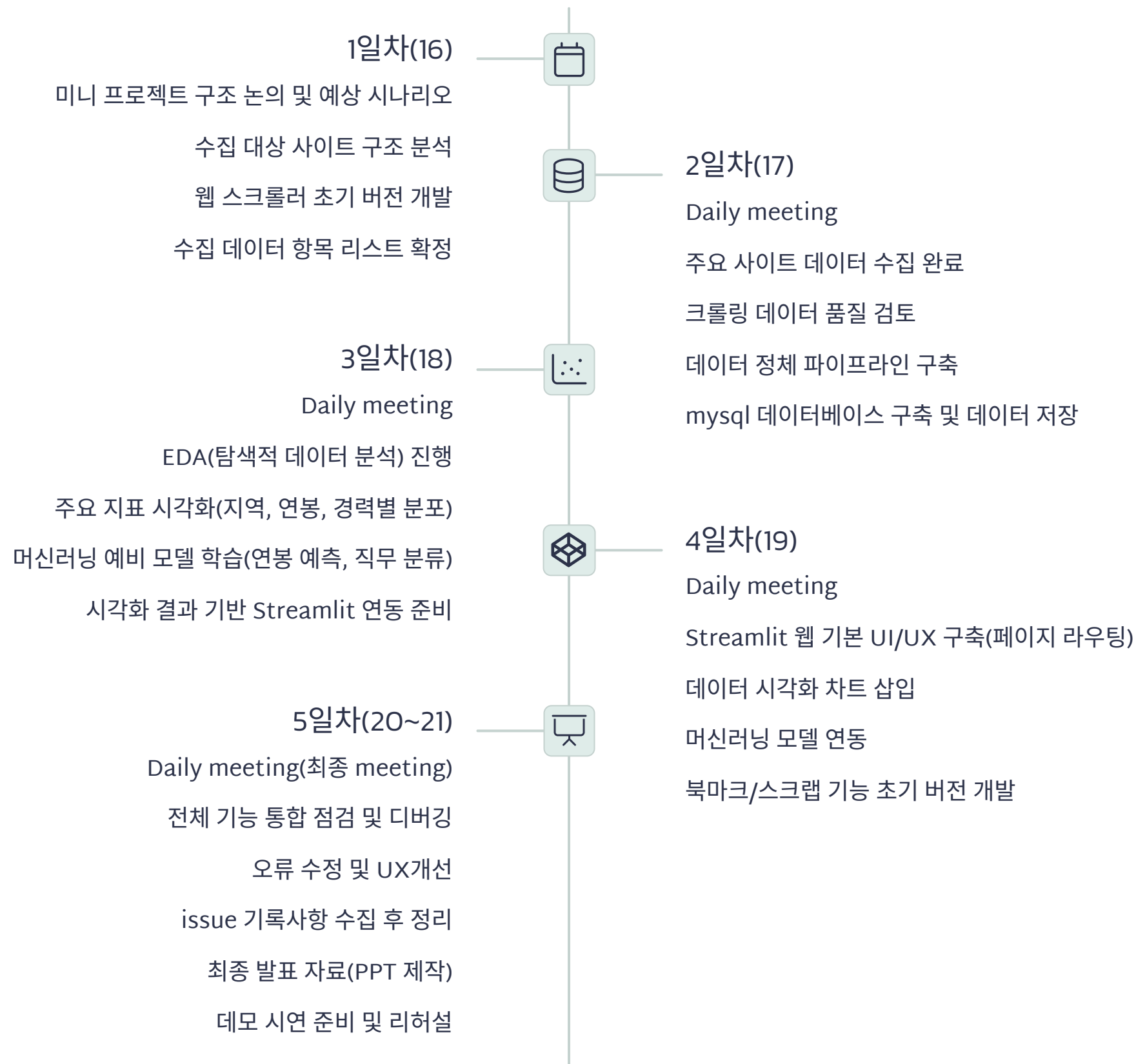
기업의 상세 분석 페이지

- 상세 정보 및 매출액 추이 제공
- 복지 정보
- 유사 업계 차이 비교
- 과거 채용 트렌드 분석
- 채용 경쟁력 점수화
- 구직자 평판/리뷰 분석

데이터 분석 결과 시뮬레이터

- 기업정보 클릭 시 상세 정보 제공
- 머신러닝 모델 기반 매출액 추이 예측
- 기업 특징 및 아쉬운점 제안
- 채용 포지션 별 경쟁력 점수 제공
- 구직자 지원 이력 분석 및 맞춤형 조언

일정 계획



종합 기대 효과 및 발전 방향

구직자 측면

- 구직자의 효율적인 자의 의사 결정 지원
- 실시간 지원 현황 분석 리포트 제공
- 개인 맞춤형 경력 성장 경로 제시
- 구직 활동 스트레스 감소
- 채용시장 투명성 향상

기업 측면

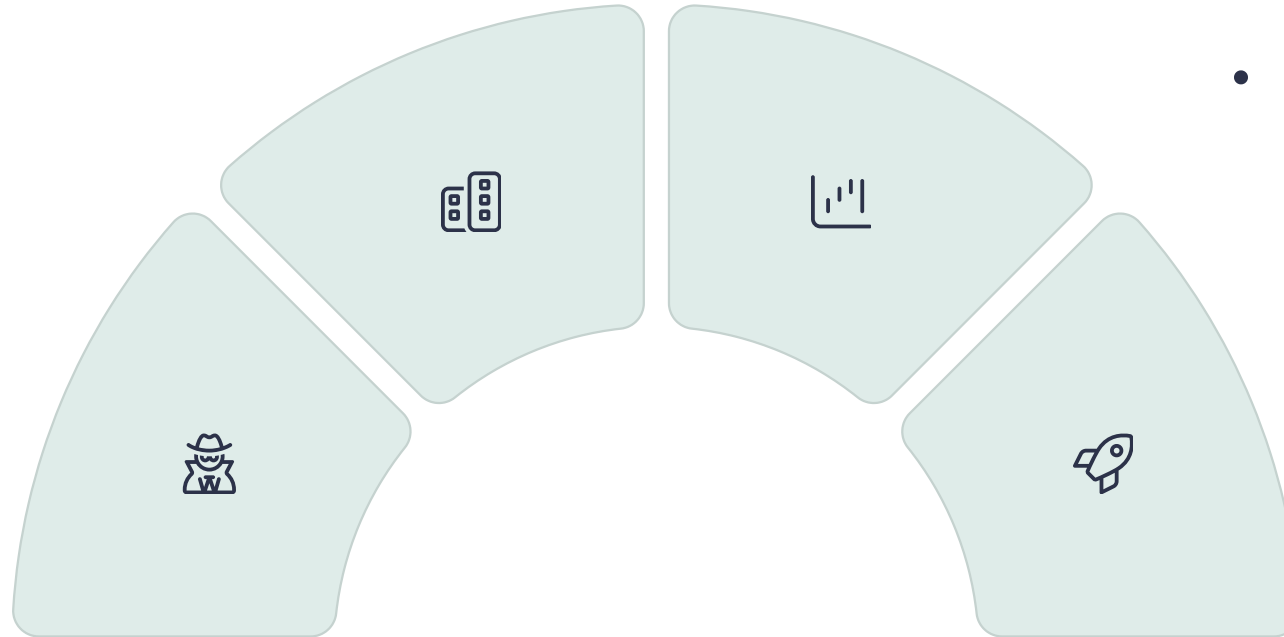
- 브랜드 인지도 강화
- 지원자 행동 데이터 확보 및 채용 광고 최적화
- 경쟁사 분석
- 우수 인재 확보 전략 수립

플랫폼 측면

- 데이터 기반 서비스 고도화
- 광고 및 유료 기능 확장 기반 마련
- 시장 분석 기반 비즈니스 모델 다변화

확장 가능성

- 기술적 확장: AI 고도화, 블록 체인 적용
- 서비스 영역 확장: Post-Hiring, 글로벌 크로스보더
- 데이터 활용 확장: 실시간 산업 리포트, 구직자 Lifetime Value 관리
- 수익 모델 다변화: B2B2E, Web3 연계



프로젝트 수행 후기

어려웠던 점 & 해결 방법

안태경: 프로젝트를 진행하면서 가장 어려웠던 점은 초기 아이디어 선정 과정이었습니다. 다양한 아이디어를 제시하긴 했지만, 관련 자료를 찾는 데 큰 어려움을 겪었습니다. 이를 해결하고자 'Gemini - Deep Research'를 활용해 자료를 찾아보기도 했고, 이를 통해 일반적인 웹 서핑으로는 찾기 어려운 몇몇 사이트를 발견하기도 했습니다. 그러나 여전히 자료의 신뢰도나 적절성이 부족하다고 판단되어, 결국 해당 아이디어는 폐기하게 되었습니다. 이후 팀원들의 적극적인 아이디어 제안과 끈질긴 자료 조사 끝에 지금의 주제와 데이터셋을 구축할 수 있었습니다. 아이디어가 정해진 후에도 팀원들과 지속적인 브레인스토밍을 통해 다양한 기능과 방향성을 논의했고, 기존 수업에서 배운 여러 코드들을 응용해 구현해 나갔습니다. 그러나 구현 과정에서 코드가 지나치게 복잡해지고 스파게티처럼 꼬여 있는 점이 문제로 드러났습니다. 이를 해결하기 위해 ChatGPT를 활용하여 코드 리팩토링을 진행했고, 덕분에 더욱 깔끔하고 유지보수가 쉬운 구조로 개선할 수 있었습니다.

한명규: 프로젝트를 진행에 있어 웹크롤링과 웹 스크래핑 방식을 사용하여 사이트를 구현 하였습니다, 하던중 크롤링 부분에서 어려움이 많아 ai의 도움을 받기도 하였지만 잘 구현이 되지 않았습니다, 하지만 팀원들과 함께 머리를 싸매고 잘 모르는 부분은 팀원들에게 질문을 해가며 도움을 받다보니 사이트 구현에 성공하였습니다.

홍영준: 미니 프로젝트를 진행 해보면서, 코드와 코드를 합치는 과정이 어려웠고, test진행 해보면서 부족한 부분도 많이 보였고, 혼자 해결이 안되는 부분은 태경님과 명규님에게 도움요청을 통하여 진행하여 프로젝트의 구현을 성공 하였습니다.

배운 점 및 느낀 점

안태경: 확실히 아이디어가 중요한 시대라고 느꼈습니다.

한명규: streamlit에 엄청 많은 기능이 있다는 것을 알게 되었습니다. 덕분에 원하는 방향의 사이트를 비교적 쉽게 구현 할 수 있어 편리하였습니다.

홍영준: 제가 생각했던 것 보다 훨씬 복잡했습니다. 처음 접해보는 프로젝트여서 많이 걱정되었습니다. 태경님과 명규님께서 많이 이끌어 주신 결과 완벽하게는 아니어도 잘 따라갈 수 있었습니다. 프로젝트를 진행 하면서 두 분께서 많이 가르쳐주시고 알려주셔서 많이 배울점이 많아 도움이 되었습니다. 같이 프로젝트를 할 수 있어 영광 이었고, 두 분께 감사를 표합니다.

개선하고 싶은 점(Next Step)

현재 계획에 있는 부분을 많이 실행을 하지 못 하였지만 한반도를 시각화 하여 전국구 채용 공고수를 확인 or 최다 공고수 지역 등 비교분석할 수 있는 데이터를 출력하는것이 제일 개선하고 싶은 점 입니다.