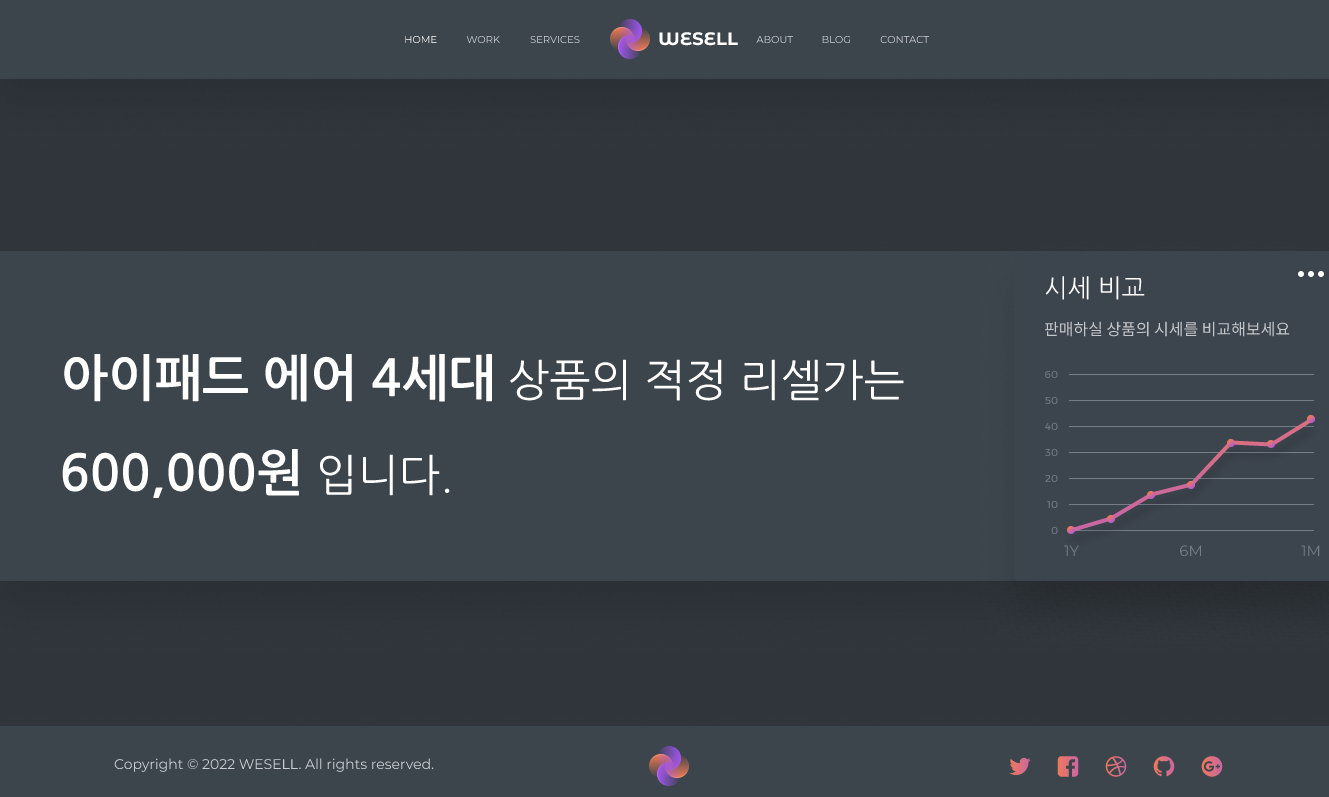
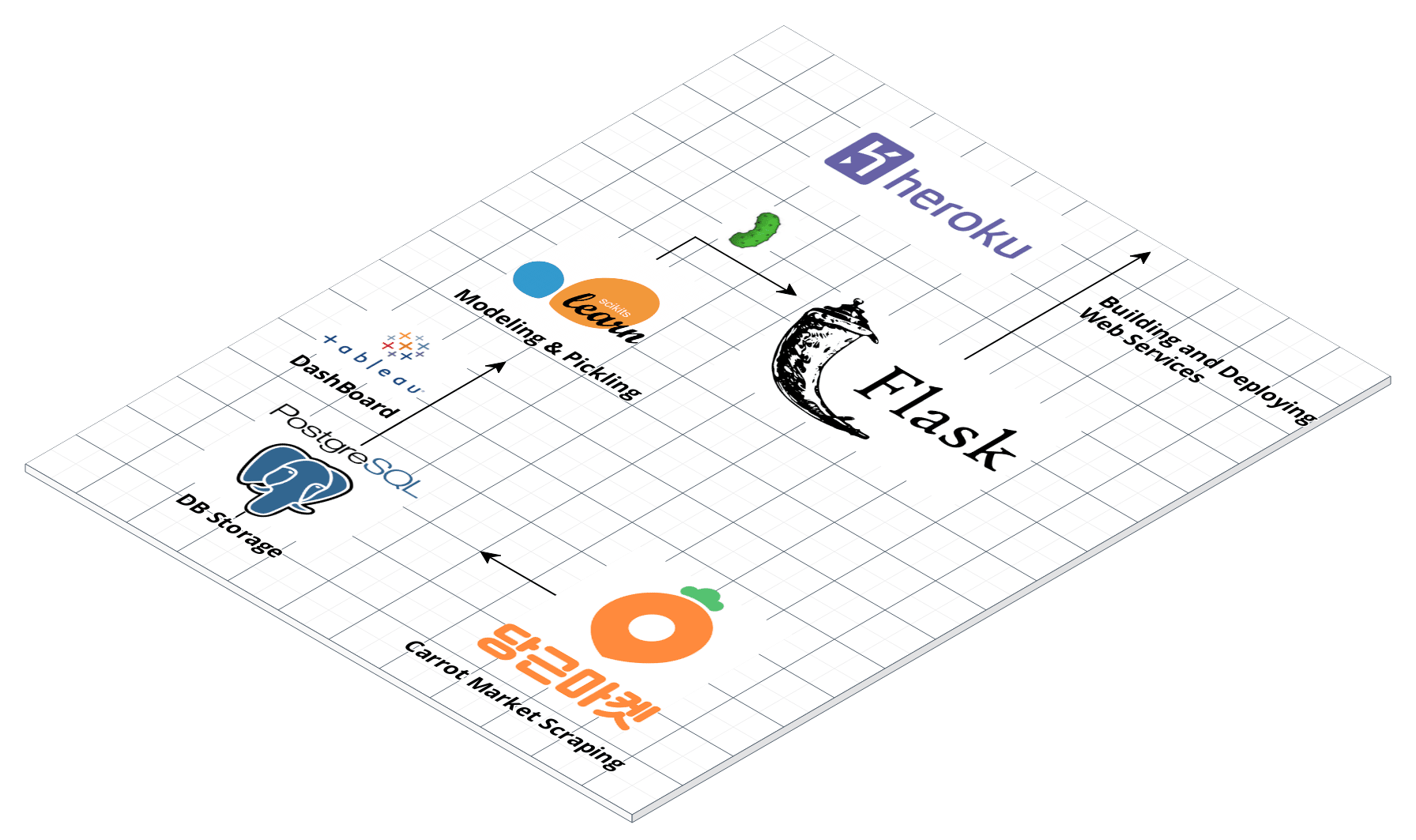


**Code States AIB(AI Bootcamp)   
수강생 프로젝트 기획서**

| 수강생 이름 : | 강병우, 이유진, 이창수, 임재민, 이다정 |
| --- | --- |
| 시작일 : | 2022. 08. 30 ~ |
| 프로젝트 : | < WESELL: > |

| **1. 프로젝트 주제** |
| --- |
| 최근 경기 침체가 장기화되면서 중고거래가 일상으로 자리잡고 있습니다. 저희 팀은 소비자들의 현명한 중고거래를 돕는, 중고 거래가 추천 웹 어플리케이션을 개발하고자 합니다. 당근마켓 등 거래가 많은 중고거래 플랫폼의 중고가격과 이용자들의 관심 수, 채팅 수 등을 스크래핑해 가격에 따른 이용자 관심도를 예측하는 ML 모델을 만든 뒤 ,이를 웹으로 이용가능하도록 플랫폼을 구축할 계획입니다. |
| **2. 프로젝트 목적** |
| 이번 프로젝트의 목적은 크게 두 가지입니다. 첫번째는 단순 프로젝트용 개발로 그치는 것이 아닌, 현실에서 사용가능할 정도로 실용적이고 유용한 앱 서비스를 개발하는 것입니다. 두번째는 개발 프로젝트를 통해 팀원들이 다수간 협업 프로세스를 경험하고 이에 익숙해지는 것입니다. 다수의 구성원이 모여 서비스를 개발하는 프로젝트는 팀원 모두 처음인 만큼 여러 측면에서 작지 않은 성장을 이룰 수 있을 것이라 기대합니다. |
| **3. 프로젝트 구현 기술 및 담당자 선정** |
| *프로젝트 구현에 필요한 기술과 담당자를 구분하여 작성해주세요.*  *모든 기술 구현을 팀원 전체가 수행하는 경우 또한 아래 양식에 맞춰 작성하여 주세요*   * 스크레이핑 : Python, Selenium, bs4   + 담당자 : 이창수, 강병우, 이유진, 임재민, 이다정 * 데이터베이스 설계 및 분석 : postgreSQL, DBeaver   + 담당자 : 임재민, 이창수 * 전처리 : Python pandas   + 담당자 : 강병우, 이유진, 이다정 * 모델링 : ML 예측 모델, Scikit-Learn, Tableau   + 담당자 : 강병우, 이유진, 이다정 * 앱 서비스 구현 : Flask, Heroku   + 담당자 : 임재민, 이창수 |
| **4. 프로젝트에 필요한 강의 컨텐츠** |
| *프로젝트에 필요한 강의 컨텐츠(책, 강의 영상, 강의 노트 등) 을 작성하여 주세요.*   * 강병우 - 해당 모델 관련 자료 혹은 강의, Tableau : 실무 밀착 데이터 시각화(Fastcampus) * 이창수 - 셀레니움, bs4, postgreSQL, Flask 로그인/회원가입 관련 강의 * 이유진 - Tableau 10: Training on How to Use Tableau For Data Science (Udemy) * 이다정 - 데이터 전처리, 분석 , 초간단 초스피드 데이터수집(파이썬크롤링) * 임재민 - 데이터 분석 SQL Fundamentals(인프런) 및 다양한 사례로 익히는 SQL 데이터 분석 |
| **5. 프로젝트 디자인** |
|  |





아이디어를 기반한 Mock-Up 제작 ([피그마](https://www.figma.com/proto/QP5iK71fiocKeYHRawtCtC/CP1-MOCKUP?node-id=3%3A335&scaling=scale-down&page-id=0%3A1&starting-point-node-id=3%3A239))(링크 클릭)

데이터 파이프 라인

| **6. 프로젝트 개발 일정** |
| --- |
| * 8월 31일 수요일   스크레이핑   * + 강병우 : 데이터 파이프라인 구성, 당근 마켓 스크래핑   + 이창수 : 파이썬 웹 스크래퍼 개발   + 이유진 : 데이터 스크래핑   + 이다정 : 데이터 탐색, 계획   + 임재민 : 당근 마켓 스크래핑 * 9월 1일 목요일   + 강병우 : 당근 마켓 스크래핑   + 이창수 : 파이썬 웹 스크래퍼 개발 및 테스트   + 이유진 : 데이터 스크래핑   + 이다정 : 데이터 탐색, 번개장터 데이터 스크래핑   + 임재민 : 당근 마켓 스크래핑, 데이터 베이스 방향 설계 * 9월 2일 금요일   + 강병우 : 당근 마켓 스크래핑, 데이터 전처리   + 이창수 : 당근 마켓/번개장터 스크래핑 테스트   + 이유진 : 데이터 스크래핑, 전처리   + 이다정 : 번개장터 데이터 스크래핑, 데이터 전처리 시작   + 임재민 : 당근 마켓 스크래핑, 데이터 베이스 방향 설계 * 9월 5일 월요일   웹구축 시작   * + 강병우 : 당근 마켓 스크래핑, 데이터 전처리   + 이창수 : 데이터 수집 마무리   + 이유진 : 데이터 스크래핑, 데이터 전처리, 가공, 모델명 분류   + 이다정 : 데이터 EDA 데이터 분석   + 임재민 : 당근 마켓 스크래핑, 데이터 베이스 설계 * 9월 6일 화요일   + 강병우 : 데이터 전처리, EDA   + 이창수 : 플라스크 로그인/회원가입/인증 메일 기능 개발   + 이유진 : 데이터 전처리, EDA   + 이다정 : 데이터 분석(KoNLPy를 이용한 언어분석과 wordcloud 시각화)   + 임재민 : 데이터 베이스 설계, 웹 서비스 모델 구현 * 9월 7일 수요일   + 강병우: 데이터 전처리, XGboost, Randomforest 모델 설계   + 이창수 : 플라스크 회원가입 기능 개발 마무리 및 User DB 구축   + 이유진 : 가설 검정, 데이터 모델링   + 이다정 : 데이터 모델링   + 임재민 : 데이터 베이스 설계, 웹 서비스 모델 구현 * 9월 8일 목요일   + 강병우: 데이터 전처리, XGboost, Randomforest 모델링 및 각종 인코더 사용   + 이창수 : 플라스크 구글 Oauth2 로그인 기능 개발   + 이유진 : CatBoost 모델 설계   + 이다정 : 데이터 모델링 Light GBM ,모델 시각화   + 임재민 : HTML 및 CSS 처리 * 9월 9일 금요일   + 강병우: 모델 튜닝, Tableau를 이용한 모델 시각화   + 이창수 : 플라스크 구글 Oauth2 로그인 기능 개발   + 이유진 : CatBoost 모델 튜닝, Tableau를 이용한 모델 시각화   + 이다정 : 데이터 전처리Tableau를 이용한 모델 시각화   + 임재민 : HTML 및 CSS 처리 * 9월 13일 화요일   + 강병우 : 모델 튜닝, 발표자료 작성   + 이창수 : 발표자료 제작 및 발표영상 촬영   + 이유진 : CatBoost 모델 튜닝, 발표자료 작성   + 이다정 : 발표자료 제작 제출 자료 검토   + 임재민 : 웹 서비스 배포 * 9월 14일 수요일   + 강병우 : 제출 자료 검토, 발표 영상 제작   + 이창수 : 최종 영상 편집 및 제출자료 검토   + 이유진 : 제출자료 검토, 발표 영상 제작   + 이다정 : 제출자료 검토, 발표 영상 제작   + 임재민 : 제출 자료 작성 |

| **7. 프로젝트 활용 데이터** |
| --- |
| *해당 프로젝트에 활용된 데이터에 대한 설명을 최대한 자세하게 작성하여 주세요.*   * 중고거래 물품 가격 및 좋아요, 조회수 등   + 데이터 설명 : 온라인 중고거래 사이트에 올라온 상품들의 가격과 좋아요 등 정보.   + 데이터 수집 방법 : Selenium, Bs4를 활용한 파이썬 스크래핑   + 데이터 타입(종류) : CSV   + 데이터 크기(예정) : 약 3만여개   + 데이터 활용 방안(예정) : 중고거래 평균가 시각화 및 가격 추천 ML 모델링 |

| **8. 팀 및 개인 Github 링크 작성** |
| --- |
| *팀 전체 깃 허브 혹은 팀 원 분들의 Github 링크를 작성하여 주세요.*  *해당 Github 의 기록은 Codestates Project 참여 확인 용도를 위해 참고할 예정입니다.*   * 이창수 : <https://github.com/grewtree> * 이유진 : <https://github.com/lyj7518> * 강병우 : <https://github.com/vkekd25> * 임재민 : <https://github.com/woals182> * 이다정 : <https://github.com/Daj21> |