|  |  |  |  |
| --- | --- | --- | --- |
| **[2차 프로젝트 수행 일지 ]** | | | |
| **서비스 산업 데이터를 활용한 빅데이터 분석 실무** | | | |
| **프로젝트 타이틀** | 지역별 초등 돌봄교실 수요 시각화를 통한 초등돌봄교실 우선지역 배치 | | |
| **프로젝트 팀명** | 일조 | **프로젝트 팀원** | (팀장) 김도훈  (팀원) 김지은, 정혜린 |

|  |  |
| --- | --- |
| **1. 2차 프로젝트 수행 계획 및 달성도** | |
| **이름** | **수행 계획** |
| 김도훈 | 1. 데이터 가공 및 정제, 추가데이터 수집(데이터 준비) 2. 변수 후보 탐색, 변수간 관계 분석(데이터 탐색) 3. 분석 가설 수립, 탐색적 데이터 분석(데이터 분석 가설 수립) 4. WBS 작성 |
| 김지은 | 1. 분석 가설 수립, 탐색적 데이터 분석(데이터 분석 가설 수립) 2. 모델링 성능평가, 통계적 유의성 검토(모델링) 3. 신규데이터 모델 적용, 모델링 성능평가(검증&테스트) 4. WBS 작성 |
| 정혜린 | 1. 데이터 가공 및 정제, 추가 데이터 수집(데이터 준비) 2. 분석 가설 수립, 탐색적 데이터 분석(데이터 분석 가설 수립) 3. 자료조사(이슈 및 문제점 조사) & 서비스 전략 수립&산출물 정리&발표자료 준비 4. WBS 작성 |

|  |  |  |
| --- | --- | --- |
| **2. 2차 프로젝트 수행 내용 및 문제점** | | |
| **이름** | **수행 내용** | **문제점** |
| 김도훈 | 1. 데이터 가공 및 정제, 추가데이터 수집: 완료 2. 변수 후보 탐색, 변수간 관계 분석: 완료 3. 분석 가설 수립, 탐색적 데이터 분석 | * 구별 초등 비만학생 BMI 수치 등 특정 데이터 수집에 어려움을 겪음 * 구별 초등 5-6학년 BMI 데이터를 찾았고, 이를 모집단으로 한 데이터를 가공함. |
| 김지은 | 1. 분석 가설 수립, 탐색적 데이터 분석 2. 모델링 성능평가, 통계적 유의성 검토 3. 신규데이터 모델 적용, 모델링 성능평가 | * ‘한부모가정’이 많은 지역구 일수록 돌봄교실의 수가 높게 나타나는 상관성을 띌 것으로 가설을 수립하였으나, 실제 분석 결과 상관성이 낮게 나타남. |
| 정혜린 | 1. 데이터 가공 및 정제, 추가 데이터 수집:완료 2. 주제와 관련한 이슈 및 문제점 조사 및 정리 3. 자료조사, 서비스 전략수립,산출물정리, 발표자료 준비 | * 논리구축 자료를 찾는데 어려움을 겪었으나, 국내외 논문과 현재 정부가 추진하고 있는 정책을 기반으로 자료를 수집함. |

|  |  |
| --- | --- |
| **3. 3차 프로젝트 수행 계획(차주)** | |
| **이름** | **수행 계획** |
| 김도훈 | 1. 서울 자치구별 결과값을 토대로 광역시 별 회귀분석을 진행한다. 2. 통계성 유의성을 검토한다. |
| 김지은 | 1. 서울 자치구별 결과값을 토대로 광역시 별 회귀분석을 진행한다. 2. 모델링 성능평가를 진행한다. |
| 정혜린 | 1. 분석된 데이터를 활용하여 향후 발전가능한 서비스 활용에 대해 구체화 한다. 2. 분석 결과를 기반으로 한 논리 구축 자료를 수집한다. |

|  |  |
| --- | --- |
| **4. 강사님 피드백** | |
| **금주 프로젝트 수행 내용** | **차주 프로젝트 수행 계획** |
| - | - |
| **피드백 일시** 년 월 일  **강사명** (서명) | |