|  |  |
| --- | --- |
| 교육 제목 | **AI 프로젝트 기반 S/W 전문가 양성과정 3일차** |
| 교육 일시 | 21년 10월 27일 |
| 교육 장소 | 영우글로벌러닝평생교원 |
| **교육 내용** | |
| 오전 | \*10월 26일 교육한 git 복습  : github 생성, github 서버와 작업자간 충돌이 있을 때 대처방법, branch 활용  \*기초통계학  : 자료의 종류 – 표본집단이 빅데이터로 가면 표본집단이 모집단에 수렴함  : 범주형 자료의 요약 – 많은 데이터를 요약 하려면 직관련이 중요하다 – 그림이 직관력에 좋다  : 연속형 자료의 요약 – 돗수분포표  히스토그램 – 데이터분포도  Boxplot  Violin plot – 최근에 쓰이기 시작, 데이터 직관력이 뛰  남, 분포상태에 관심이 있다  요점 : 직관력이 중요 |
| 오후 | \*수치를 통한 연속형 자료의 요약  : 중심위치의 측도 – 표본평균(이상치 영향O), 중앙값(이상치 영향X), 최빈값(거의 안 쓰임)  : 평균 = 무게중심  : 이상치(outlier)  \*퍼짐의 정도  : 분산과 표준편차  : 사분위수범위 – outlier 를 구하려면, Q3 + 1.5 X IQR , Q1 + 1.5 X IQR  1.5 계수는 전문가의 특별한 의견이 있진 않는 이상 거의 고정값  \*두 변수간의 자료 요약 – 데이터상관성을 이해하는 것이 가장 중요하다  : 상관분석 – 인과관계 <> 상관계수  \*그림을 통한 상관관계 요약  : 산점도  : 피어슨 상관관계 – 정규분포를 따를 때, 곡선형 그래프 불가  : 스피어만 상관계수 – 정규분포를 따르지 않아도 된다 |
|  |  |