**데이터 시각화 프로젝트 기획안**

기획안 작성일자 : 2022년 01월 13일

| 팀 명 | 4조 : 집값은왜오르조 |
| --- | --- |
| 프로젝트 주제 및 개요 | 주제 : 전국 집값 상승에 따른 여러 분야와의 상관관계    개요 : 전국 도, 광역시 별 집값 상승률과 경제, 인구, 교통에 여러 데이터들을 정제 하여, 집값과 각각의 데이터들을 비교하여 어떤 상관 관계가 있는지 파악하여 집값 상승에 어떤 요인들이 비슷한 영향을 미칠지 파악하고자 한다. |
| 프로젝트 수행 방향 | 1. 데이터 수집  a. 년도별 전국 집값 데이터 수집  b. 경제, 인구, 교통, 기타 등에 관련한 년도별 데이터 수집  - 경제 : 물가상승률, 근로자 평균 임금, 금리, 일자리(취업률) 등..  - 인구 : 출산율(산부인과), 혼인율, 무주택 가구 수, 인구이동 등..  - 교통 : 차량등록현황, 교통사고 발생률 등..    2. 데이터 전처리  - 지역별, 년도별 집값 데이터 : 나머지 분야별 데이터들로 dataframe화    2-1. 데이터 전처리 과정  - 데이터 출처 표기!  - 비교하고자 하는 데이터와 집값 상승과 어떤 관계가 있을지?  - 작성한 코드가 어떤 처리를 하는 코드인지 주석 달아주기  - 지역 명칭은 풀네임으로 통일(서울특별시, 경기도, 강원도, ...)  - row데이터 목록  전국, 지방, 수도권 + 각 광역시, 도 별 형태  - col 데이터 목록  년/월 형태의 데이터는 지수 또는 지표는 평균, 수량 값은 합계로 년도 데이터로 새로운 col 생성    3. 데이터 시각화  - 전처리한 데이터를 통해 집값과 각각의 데이터들의 상관 관계를 분석하고 시각화하여 집값 상승의 요인에 어떤 데이터가 영향을 미치는지 확인. |
| 프로젝트 조직  ㅓㅏㅣ(구성원 및 역할) | ● 역할분담  팀장: 김용현  데이터 전처리 - 차량등록현황, 교통사고 발생률    팀원1: 김주희  데이터 전처리 - 소비자 물가상승률, 소비자물가지수, 예금은행대출금액    팀원2: 윤성준  데이터 전처리 - 미분양주택현황, 삶의 만족도    팀원3: 이경희  데이터 전처리 - 거주지역별\_주택소유\_및\_무주택\_가구수(전국),  순이동인구\_시도\_시\_군\_구(전국)    팀원4: 이현호  데이터 전처리 - 시도별\_평균초혼연령, 연도 및 지역별 출생건수, 혼인건수, 월/분기/연간\_인구동향\_출생\_사망\_혼인\_이혼(출생아수) |
| 프로젝트 추진 일정 | -● 일정  1/12~ 1/13: 주제 선정 및 일정 수립, 데이터 수집  1/14 ~ 1/17: 데이터 전처리  1/17 ~ 1/18: 데이터 전처리 정리 및 보완  1/19 ~ 1/20: 데이터 시각화  1/21 ~ 1/24: 발표자료 준비 |