




Capstone Design 중간보고서						
1. 신청과제						
수행기간	2024학년도 2학기		교과목명	데이터분석모델링캡스톤디자인		
과제명	경기도 성남시 부동산 가격 예측모델					
팀명	맛동산		신청예산	총 1,600,000원		
지도교수(학과)	데이터사이언스학부		지도교수(성명)	안재준		
2. 참여학생(최소 2인 이상)						
구분	역할	성명	전공	학년	학번	
팀장	데이터분석 총괄	김보운	정보통계학	4	2019251153	
팀원1	데이터수집 및 문헌분석	이중강	정보통계학	4	2019251045	
팀원2	데이터분석 및 모델링	류원우	경제학	4	2018231181	
팀원3	데이터분석 및 모델링	박준영	정보통계학	4	2019251046	
팀원4						
팀원5						
3. 과제타입(택1)						
<input type="checkbox"/> 일반형		<input checked="" type="checkbox"/> 융합형		<input type="checkbox"/> 기업연계형		
				<input type="checkbox"/> 연세글로벌형		
4. 결과물 종류(택1)						
유형	<input type="checkbox"/> 시작품		<input type="checkbox"/> 하드웨어		<input type="checkbox"/> 시제품 및 결과모형	
	<input type="checkbox"/> 기타 ( )					
무형	<input type="checkbox"/> 인쇄물 및 영상		<input type="checkbox"/> 설계도면		<input checked="" type="checkbox"/> 보고서(조사, 분석결과 등)	
	<input type="checkbox"/> 소프트웨어(어플리케이션, 웹페이지 등)			<input type="checkbox"/> 기타 ( )		
5. 참여업체(기업연계형만)						
업체명		실무담당자		전화(H.P)		
				E-mail		
<p>위와 같이 Capstone Design 과제 중간보고서를 제출합니다.</p> <p>2024년 10월 25일</p> <p>(대표학생) 김보운 (인 )</p> <p>(지도교수) 안재준 (인 )</p>						
<div>  <b>연세대학교 미래캠퍼스 현장실습지원센터 귀하</b> </div> <div> <b>현장실습지원센터</b> </div>						

# Capstone Design 과제요약서

## 1. 예산집행현황 (홈페이지 확인 후 작성)

지원금 집행내역 (단위 : 원)				
지출항목	신청 예산	집행 액	잔액	비고
물 품 · 제 작 비	880,000	0	880,000	
회의비	320,000	107,900	212,100	
교통비	400,000	0	400,000	
합 계	1,600,000원	107,900원	1,492,100원	

## 2. 과제 진행현황

### 2.1. 개발 진행 사항

#### - 1~2주차 개발 진행 사항

초기에 정했던 ‘별집순환모형을 통한 서울 부동산 시장 변화’ 라는 주제로 진행하고자 하였지만 방대한 서울 데이터 규모로 인해 현재 관심이 많아지는 경기도 성남시에 대한 부동산 시장으로 범위를 줄여 진행하게 되었고, 정한 주제에 걸맞는 문헌 및 예시자료들을 수집 후 검토하여 구체적인 분석 목표와 분석과정을 정하게 되었다.

#### - 3~4주차 개발 진행 사항

부동산 시장 변화를 예측하기 위해서는 부동산시장가격과 연관이 있는 요소들을 찾아야 했고, 공공데이터포털, KOSIS 등 공공데이터를 활용하여 성남시의 각종 데이터들을 수집 할 수 있었다. 또한 분석의 용이성을 위해 범주형 데이터를 제외한 모두 수치형 데이터로 수집하였다.

#### - 5~6주차 개발 진행 사항

수집한 데이터를 바탕으로 하나의 데이터셋을 만들기위해 데이터 전처리와 결합을 하였고, 조원들의 각자 역할에 맞게 데이터 전처리조와 분석변수 추가수집조 로 나누어서 진행하였다. 결론적으로 총 17 개의 변수를 수집하여 하나의 데이터셋 파일로 만들게 되었고, 수집과정에서 월 별 데이터가 존재하지 않던 데이터는 12개월을 같은 데이터로 통일시켜 결측치를 처리하였다. 이 과정을 통해 12년, 즉 132 개의 월 별 샘플데이터를 추출하였다.

#### - 7~8주차 개발 진행 사항

분석목적인 다음달의 아파트 매매가격을 예측하기 위해 아파트매매가격 데이터를 1월씩 밀어 다음달의 아파트 매매가격 데이터를 추가해주었고, 최종 18개의 변수중에 다음달아파트매매가격을 종속변수로 두고 분석을 진행하기로 결정했다. 분석에 들어가기 전에 데이터에 존재하는 변수들을 MinMaxScaler를 이용하여 0과 1사이의 범위로 정규화를 해 주었고, 정규화된 데이터를 바탕으로 모델링에 사용할 변수를 추출하는 유의성 검정을 진행하기 시작함.

## 2.2. 향후 추진계획

만들어진 데이터셋을 이용해 상관분석을 하여 나온 상관계수를 이용해 분석에 필요없는 데이터를 소거하고, 추가로 변수중요도를 파악하여 데이터를 소각하는 과정 두가지로 모델링을 진행 할 것이고, 시계열데이터만큼 적절한 모델링 기법을 사용해서 RMSE값을 바탕으로 예측모형을 평가할 것이다.