

PORTFOLIO



금융 데이터 분석가
김혜인 포트폴리오

PORTFOLIO

금융 데이터 분석가

김혜인 포트폴리오



“일단 하고 생각하자”

김혜인 (Kim Hyein)

2000.3.9

Education

2019.03 ~ 성신여자대학교 통계학과 입학

2020.03 ~ 성신여자대학교 정보시스템공학과 복수전공

Internship

2021.09~12 양주시청 기획재정부 정보통신과 빅데이터팀 인턴

2023.01~02 기업은행 정보보호부 개인신용정보관리팀 계장

Certificate

2022.06 사회조사분석사 2급

2022.03 컴퓨터 활용능력 1급

2021.04 데이터 준전문가(ADSP)

Club Activity

2019.03~12 성신여대 봉사동아리 반짝반짝

2020.03~07 Smarter 20기 미디어홍보팀

2021.05~08 한국정보진흥원 공공빅데이터 4기

2022.01~02 성신여대 인재개발팀 행정인턴

2022.04~05 금융사관학교 50기(서포터즈)

2022.06~08 한국핀테크인턴십(개발인턴)

2022.09~12 통계학과 실습실 조교(인턴십근무)

2023.03~ 05 kb IT 아카데미 3기



금융데이터분석가라는
꿈을 위해 오늘도
한걸음씩 도전하는
김혜인입니다

“일단 하고 생각하자”

김혜인 (Kim Hyein)

2000.3.9

TEL: 010-9652-5997

EMAIL: haein0758@naver.com

Project

2021.2	신용카드 거래 데이터 분석
2021.5	기상청 날씨 빅데이터 콘테스트
2021.8	서울 열선 도선 우선 설치지 분석
2021.9	어린이 교통사고 취약지 분석
2021.10	폐렴구균 미접종자 데이터 분석
2021.11	양주시 민원데이터 분석
2021.12	양주시 코로나18 전후 매출변화 분석(지역화폐 및 신용카드)
2022.1	조달청 공모전
2022.3	통계청 텍스트 분류 분석
2022.7	Compas cctv 최적의 설치 위치 _ LH
2023.1	KOSPI 지수 시각화

Skill

Python	<div></div>	상	C++	<div></div>	중	SAS	<div></div>	중하
Excel	<div></div>	상	Java	<div></div>	중	Tabelu	<div></div>	중하
R	<div></div>	중상	SQL	<div></div>	중			
QGIS	<div></div>	중상	Spss	<div></div>	중			

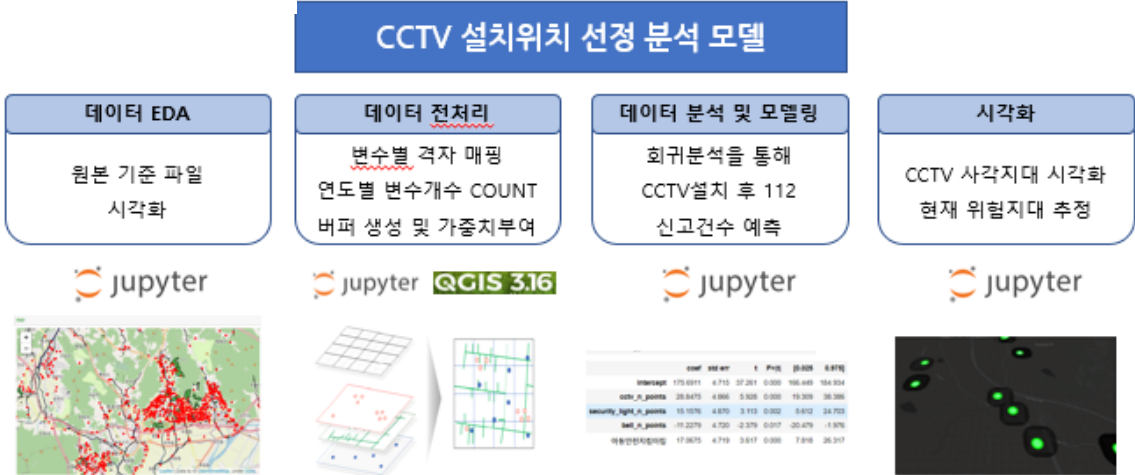
김해시 시민안전 사각지대 해소를 위한 CCTV 최적의 설치 위치 선정 _ COMPAS

김해시 시민안전을 위한 CCTV 설치현황, 112 신고이력등을 기반으로
CCTV 사각지대 분석

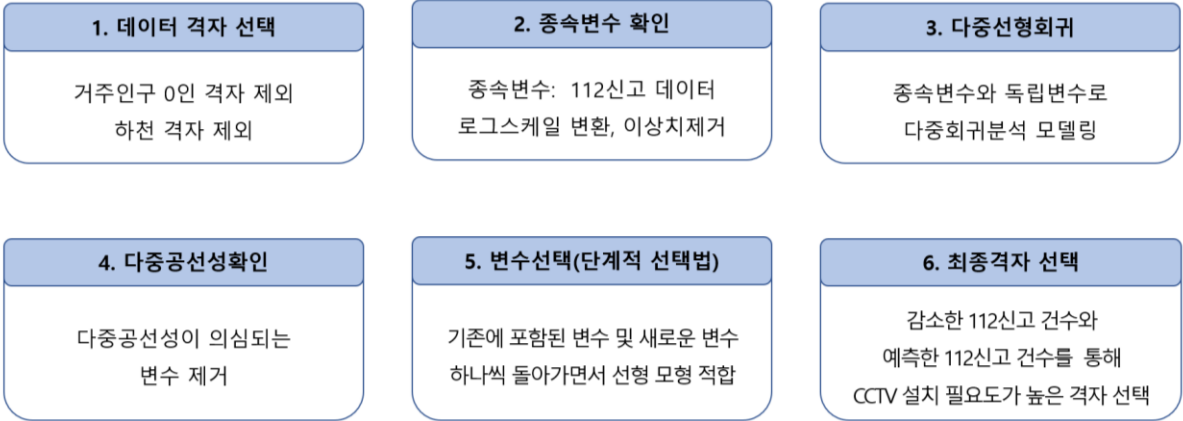
주체기관 COMPAS(LH)
분석기간 2022.7.15 ~ 2022.08.18

기획&보고서 80
모델링 90
데이터 시각화 50

분석 프로세스



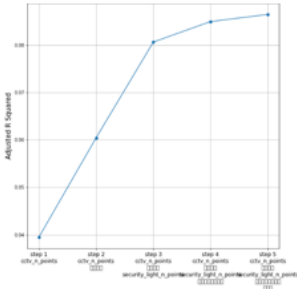
분석 모델링 - 회귀분석



행	변수명	VIF
0	cctv_n_points	1.333022
1	security_light_n_points	1.441408
2	bell_n_points	1.019555
3	cctv_buffer	1.702054
4	어린이집	1.112910
5	유치원	1.192928
6	초등학교	1.179435
7	중학교	inf
8	고등학교	inf
9	고등학교	NaN
10	아동안전지킴이집	1.035082

다중공산성이 보이는
초등학교, 중학교,
고등학교 변수 제거

단계적 선택법: 아무 변수도 선택되지 않은 Null model에서
출발하여 변수를 하나씩 추가하며 회귀식을 적당하는 방법



최종적으로 격자내 cctv 개수,
어린이집수, 보안등,
아동안전지킴이집을
독립변수로 사용하였다.

분석 결과



KOSPI와 외부 데이터와의 상관관계 분석 및 시각화

KOSPI 데이터를 기반으로 한 주식시장에 영향을 미치는 데이터 선정 후
회귀분석을 통한 주가 움직임 이해

데이터 수집	<div><div></div></div>	60
모델링	<div><div></div></div>	90

분석 활용 데이터 선정_데이터 수집

data1. 해외지수
data2. 해외통화

미국(Dow Jones, Nasdaq), eurostoxx,
독일(dax), 일본(Nikkei), 홍콩恒生지수

달러, 유로, 독일마스크, 일본엔
싱가폴 달러, 중국위안

data1. 물가지수

생산자물가지수, 소비자물가지수,
수출/수입물가지수, 주택매매가격지수

금융통계 수업에서 한국실물경제에 영향을 미치
는 변수로 민간소비수요, 민간투자수요,
수입/수출물가지수, 정부 수요로 배워서
한국경제시스템에서 기간별로 데이터 선정

data3. 유가 데이터

OIL

최근 급격한 유가 상승으로 코스피 지수가
큰 폭으로 변동했다는 기사를 읽고
필연적 관계가 있다고 판단함.

데이터 전처리 진행

데이터 타입 / 데이터 날짜 형식 맞추기 / 데이터 결합 / 년.월.일 칼럼 생성

분석 모델링 - 회귀분석

상관관계 분석

```
# 경고 뜨지 않게..
import warnings
warnings.filterwarnings('ignore')

# 그래프를 노트북 안에 그리기 위해 설정
%matplotlib inline

# 그래프에서 마이너스 폰트 깨지는 문제에 대한 대처
mpl.rcParams['axes.unicode_minus'] = False

# 그래프 설정
plt.rc('font', family='NanumBarunGothic')
plt.rcParams['font.size'] = 16
plt.rcParams['figure.figsize'] = 20, 10
plt.rcParams['axes.unicode_minus'] = False

data_heat = df.corr()
mask = np.zeros_like(data_heat, dtype=np.bool)
mask[np.triu_indices_from(mask)] = True
sns.heatmap(data_heat, mask=mask, annot=True, linewidths=4, fwt='.2f',
            plt.xticks(rotation=35))
plt.title('코스피 상관관계 열별')
plt.show()
```

다중회귀 분석

```
df.info()

<class 'pandas.core.frame.DataFrame'>
MultiIndex: 250 entries, (2000, 1) to (2022, 8)
Data columns (total 28 columns):
 #   Column                Non-Null Count  Dtype
---  -
col = ['원/미국달러(매매기준율)', '원/유로', '원/일본엔(100엔)', '원/싱가포르달러', '원/중국위안',
        '미국(Dow Jones)', '미국(NASDAQ)', 'Euro STOXX', '독일(DAX)', '일본(NIKKEI)', '싱가포르(STI)', '홍콩(항셍)',
        '중국(SHCOMP)', '생산자물가지수', '소비자물가지수', '수출물가지수', '수입물가지수', '주택매매가격지수', 'oilprice']

## 결측값 있는 열은 그냥 빼고 진행함.
X = df_reg_1[col]
y = df_reg_1['Close']

model = sm.OLS(y,X)
res = model.fit()
res.summary()
```

결측값 확인

결측값 열 제외 한 후 다중회귀 분석

PORTFOLIO

김혜인 (Kim Hyein)

Education

2019.3 ~ 성신여자대학교 통계학과 입학
2020.3 ~ 성신여자대학교 정보시스템공학과 복수전공

Internship

2021.9 ~ 양주시청 기획재정부 정보통신과 빅데이터부 인턴

Certificate

2022.6 사회조사분석사 2급
2022.3 컴퓨터 활용능력 1급
2021.4 데이터 준전문가(ADSP)

Club Activity

2019.3~12 성신여대 봉사동아리 반짝반짝
2020.3~7 Smarter 20기 미디어홍보팀
2021.2~ Daja Vu 빅데이터 분석 동아리
2021.5~8 한국정보진흥원 공공빅데이터 4기
2022.4~5 금융사관학교 50기,서포터즈
2022.6~ 한국핀테크인턴십

Project

2021.2 신용카드 거래 데이터 분석
2021.5 기상청 날씨 빅데이터 콘테스트
2021.8 서울 열선 도선 우선 설치지 분석
2021.9 어린이 교통사고 취약지 분석
2021.10 폐렴구균 미접종자 데이터 분석
2021.11 양주시 민원데이터 분석
2021.12 양주시 코로나18 전후 매출변화 분석
2022.1 조달청 공보전
2022.3 통계청 산업분류모델 자동화분석

Skill

Python	<div></div> 상	C++	<div></div> 중
Excel	<div></div> 상	JavaScript	<div></div> 중
R	<div></div> 중상	SQL	<div></div> 중
QGIS	<div></div> 중상	Spss	<div></div> 중



“일단 하고 생각하자”

김혜인 (Kim Hyein)

2000.3.9

TEL: 010-9652-5997

EMAIL: haein0758@naver.com

금융데이터분석가라는
꿈을 위해 오늘도
한걸음씩 도전하는
김혜인입니다

Education

2019.2 서울수락고등학교 졸업
2019.3 성신여자대학교 통계학과 입학
2020.3 정보시스템공학과 복수전공

Internship

2021.9 ~ 양주시청 기획재정부
정보통신과 빅데이터 인턴
2022.1~2 성신여대 일자리센터 행정인턴

Certificate

2022.6 사회조사분석사 2급
2022.3 컴퓨터 활용능력 1급
2021.4 데이터 준전문가(ADSP)

Skill

Python	<div></div>	상	C++	<div></div>	중
Excel	<div></div>	상	Java	<div></div>	중
R	<div></div>	중상	SQL	<div></div>	중
QGIS	<div></div>	중상	Spss	<div></div>	중

Club activity

2019.3~12 성신여대 봉사동아리 반짝반짝
2020.3~7 Smarter 20기 미디어홍보팀
2021.2~ Daja Vu 빅데이터 분석 동아리
2021.5~8 한국정보진흥원 공공빅데이터 4기
2022.4~5 금융사관학교 50기,서포터즈
2022.6~ 한국핀테크인턴십

Project

2021.2 신용카드 거래 데이터 분석
2021.5 기상청 날씨 빅데이터 콘테스트
2021.8 서울 열선 도선 우선 설치지 분석
2021.9 어린이 교통사고 취약지 분석
2021.10 폐렴구균 미접종자 데이터 분석
2021.11 양주시 민원데이터 분석
2021.12 양주시 코로나18 전후 매출변화 분석
(지역화폐 및 신용카드)
2022.1 조달청 공보전
2022.3 통계청 텍스트 분류 분석
2022.5~ 기상청 날씨 빅데이터 콘테스트
2022.7~ Compas cctv 최적의 설치 위치