
빅데이터 인력양성 사업(경기대)

수주 및 매출 분석을 통한 수요처별 맞춤형 마케팅 모형

김상진 교수님 홍석환(팀) 박창현 장태수 최지나

CONTENTS

프로젝트 목표

- 01 분석 배경
- 02 주제 선정 및 목표

데이터 분석 및 결과

- 01 활용 데이터 설명
- 02 데이터 처리과정
- 03 데이터 분석 기법 활용
- 04 데이터 분석 결과

결론

- 01 영업 전략 제안
- 02 출처

01. 분석배경

UPS – 무정전 전원 장치

- ✓ Uninterruptible Power Supply의 약자로, 무정전 전원 공급장치
- ✓ 공급되는 전원에 문제 발생 시, 중요 부하에 **안정적인 전원 공급**해주는 시설물
- ✓ 대표적 전원 문제: 정전, 순간전압강하, 순간전압상승, 과전압, 전선 노이즈, 주파수/스위치 변화, 고조파 일그러짐

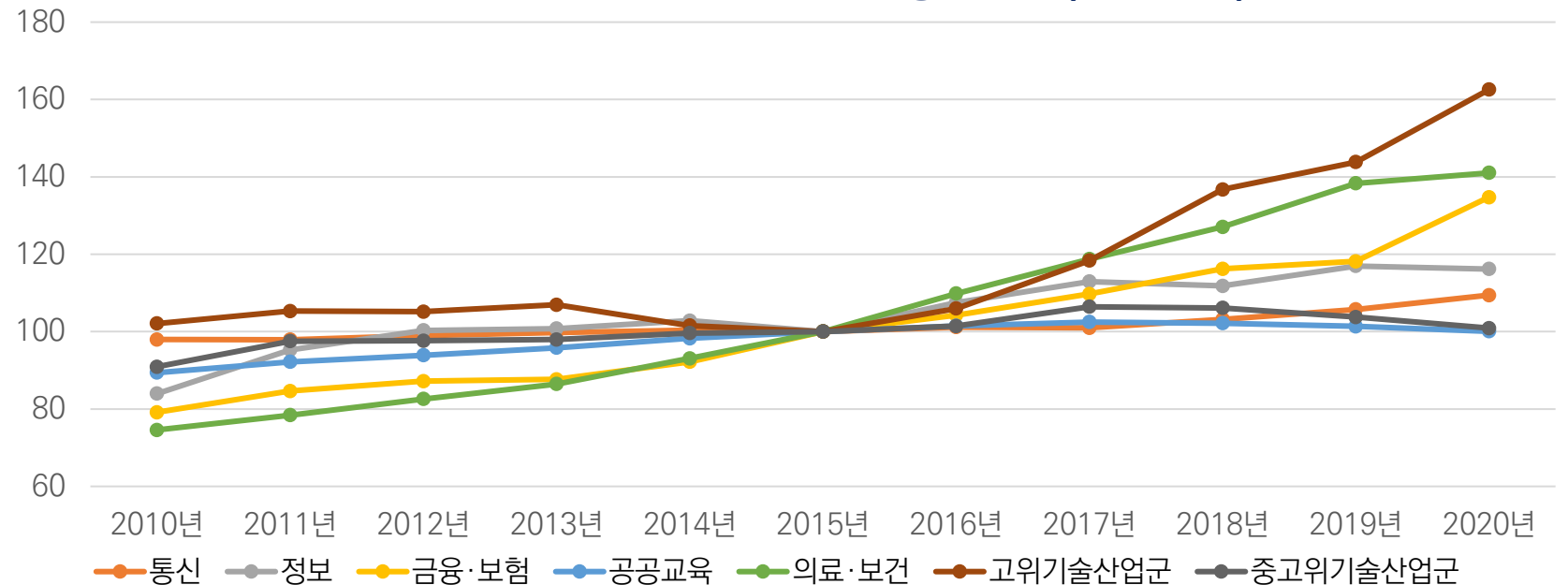
UPS 시장 성장 동력과 산업 동향

전력 부족현상으로 인한
전력 품질 저하 및
정전 가능성에 대한 대비 필요

산업용 전력 Back-up 수요 증가
IT산업체(통신, 정보, 금융/보험)
고위기술산업군(의약, 반도체등)
중고위기술산업군
(석유화학, 전기기기등)

녹색에너지 저장기술 사용 증가
전기차 폐배터리의 UPS화 로의
개발/연구 수행

주요 제조업 및 서비스업 생산지수(불변지수)



단위 : 2015년=100 기준
출처 : 산업통계 분석시스템(INSTANS)

02.주제 선정 및 목표

기존 문제점

- ✓ (주)성신전기공업의 경우, 선주문 후제작 기업으로 과거 판매량 데이터 만을 이용하여 상품 수요예측이 가능했음.
- ✓ 주문과 판매사이의 기간은 최소 30일 이상으로, 그 사이의 기간동안 UPS의 사양과 개수의 변화가 생길 수 있음.
- ✓ 따라서 각 산업체의 시기와 사양, 개수에 대한 수요예측 모델 제시 필요

분석 주제

1. 수요처별 상세구분

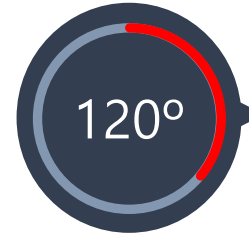
2. 수요처별 구매 시기 분석

3. 수요처별 선호 용량 및 사양 분석

4. 수요처별 수주 변화 추이 분석

5. 수주에서 매출까지의 소요기간

목표



대략적으로 알고 있었던 것에 대한 수치화 및 시각화 자료 제시



분석 주제들의 결과를 이용하여 연간 영업 전략 제안



의외의 결과에 대한 인사이트 발굴 및 제언

01.활용데이터 설명

DATA

수주 DATA

(2011.01 ~ 2020.12)



매출 DATA

(2011.01 ~ 2020.12)



Raw DATA

속성명	속성값	속성설명	속성유형
날짜	2020.01, 2020.12 ...	날짜 데이터	범주형,명목형
납품처	OO고등학교,OO전기, OO공업 ...	납품처 명	범주형,명목형
품목	UPS,부품,축전지,수리 ...	품목명	범주형,명목형
용량	1KVA, 2KVA, 5KVA, 10KVA ...	용량	범주형,명목형
수량	1,10,16 ...	수량	수치형

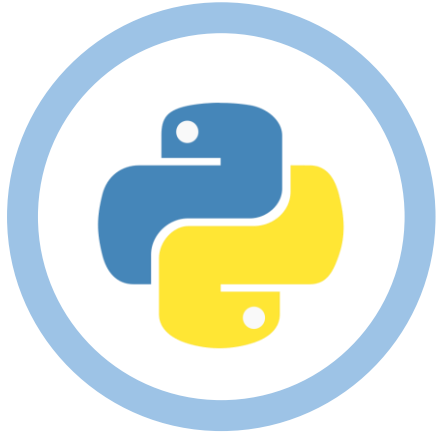


약 17,000여개의 데이터, 5개의 컬럼



동일 회사, 용량이지만 표기의 방법이 다름

02. 활용 도구



Python



R



Tableau



Excel

02.데이터 처리 과정

STEP 01 변수 정리

- 용량 단위 단일화로 인한 변수 내용 변경
- 날짜변수를 연도와 월로 별도 구분
- 회사명 단일화
- 이상치 처리
 - ⇒ 여러 지점에 대한 내용이 묶인 경우
 - ⇒ '설치처-주문처' 형태의 분리
 - ⇒ 품목에 대해 이중 개념이 부딪힌 경우 분리
(ex. 판매와 수리의 개념이 부딪힌 경우)
- 품목, 카테고리 분류 기준 수립

STEP 02 변수 생성

- 납품처명과 카카오맵 API를 이용해
산업 카테고리 추출
- 미분류 및 오분류에 대한 확인
 - ⇒ 카카오맵 재검색 또는 네이버 지도 이용
 - ⇒ 오타 가능성을 염두, 비슷한 회사명 검색

STEP 03 최종 데이터

- 품목, 카테고리 분류 기준에 따른 적용
- 단위, 날짜, 회사명의 형태 단일화
- 데이터의 출처 구분(kind)
 - ⇒ 데이터 분석을 위한 기본 데이터 완성

02.데이터 처리 과정

STEP 01 변수 정리

	name_ori	name	volumn_ori	volumn	date_ori	year	month
5598	성신전기공업	성신전기공업	7.5K (BL3000)	7.5	2011.01	2011	1
5599	성신전기공업	성신전기공업	80K (BL3000)	80.0	2011.01	2011	1
5600	성신전기공업	성신전기공업	CM FILTER	NaN	2011.01	2011	1
5601	성신전기공업	성신전기공업	NaN	NaN	2011.01	2011	1
5602	성신전기공업	성신전기공업	NaN	NaN	2011.02	2011	2
...
15436	성신전기공업	성신전기공업	20K/10K	20.0	2019.02	2019	2
15437	성신전기공업	성신전기공업	20K/10K	10.0	2019.02	2019	2
15438	성신전기공업 대전	성신전기공업	5K/3K/7.5K	5.0	2020.10	2020	10
15439	성신전기공업 대전	성신전기공업	5K/3K/7.5K	3.0	2020.10	2020	10
15440	성신전기공업 대전	성신전기공업	5K/3K/7.5K	7.5	2020.10	2020	10



각 납품처명 단일화

이름은 같으나 지역명,지점명이 들어가 있거나
오타가 존재하기에 각 기업명을 하나로 정리



용량 단위 단일화로 인한 변수 변경

단위 변환 및 volumn이 아닌
오표기의 경우 Nan값으로 수정



연도와 월의 열 구분

‘날짜’ 변수의 ‘연도.월’ 형태를 열로 구분

02.데이터 처리 과정

변수 정리 中 분류 기준 수립

카테고리 분류 기준

대분류	중분류
공공기관	공사,공단
	기관
	관공서
	기간산업
	군부대
교육기관	학교1
	학교2
	교육시설
금융기관	금융기관
민간기업	건축및부동산
	제조업
	기타서비스(C2C)
의료기관	관할의료기관
	병의원
정보통신	정보통신
대리점	대리점

품목 분류 기준

대분류	중분류
장비	UPS
	UPS 외
배터리	배터리
점검/수리	점검/유지보수
	수리
	설치
교육	교육
	교육
부품	부품

변수 정리 中 이상치 처리

대다수의 데이터가 하나의 기업명을 갖고 있거나 지점별로 기록되었기에 결합데이터에 대한 분리가 필요하였다. 따라서 '설치처-주문처' 형태와 여러 지점에 대한 내용이 묶인 경우에 대하여 동일 내용을 행을 추가하였다. 용량 또한 한 행에 여러 용량이 기록된 경우가 있었기에 분리하고 수량에 대해선 (기존 수량)/(기존 행+추가된 행) 으로 계산하였다.

판매와 교체는 새로운 제품이 고객에게 전달되는 것이고 점검 및 수리의 경우 사후 서비스라고 판단하였기에 품목 기준에 따른 개념적 충돌이 일어나 이중 작업의 분리가 필요하였다. 그래서 점검 및 수리의 경우 개수가 1로 집계되는 경우가 98.4%였기에 개수를 1로 입력하고 1이 아닌 숫자에 대해서 다른 내용에 넣음으로써 데이터 손실을 방지하였다.

Ori_item	count
충전지교체및수리	40

Ori_item	item	count
충전지교체	배터리	40
및수리	점검/수리	1

02.데이터 처리 과정

STEP 02 변수 생성



카카오 API + 파이썬



산업 분류 추출

기업명 정리와 산업 분류의 방법으로 카카오 API를 이용해 내용 추출
군부대의 경우 '논산 육군훈련소'로 검색하여 병무청으로 출력



카카오 맵 및 네이버 지도
이용한 미·오분류 수정



미분류 및 오분류에 대한 확인

미 검색건에 대해서 오타 가능성을 염두하여 비슷한 회사명으로 검색하거나
또는 카카오맵에 나오지는 않지만 네이버 지도에 검색되는 경우가 있어 재검색
산업 분류를 기준으로 잘못 분류된 건에 대하여 하위 검색 내용 참고 후 수정

기업명	카카오 API 산업분류	미분류 및 오분류 수정 후
성신전기공업	서비스,산업 > 제조업 > 전기,전자	서비스,산업 > 제조업 > 전기,전자
OO경찰서	사회,공공기관 > 행정기관 > 경찰서	사회,공공기관 > 행정기관 > 경찰서
OO부대	서비스,산업 > 음식점	공공기관 > 군부대 > 병무청

02.데이터 처리 과정

STEP 03 최종 데이터

date_ori	name_ori	item_ori	volumn_ori	count_ori	kind	agent	category
2013.01	AA부대	UPS	10KVA	1	s		사회,공공기관 > 행정기관 > 병무청 > 신병교육대
2013.03	00제조업	UPS	10KVA	1	m		서비스,산업 > 제조업 > 전기,전자 > 전기자재,부품
2016.12	00통신-agent1	이전설치		1	m	agent1	서비스,산업 > 건설,건축 > 시공업체 > 통신시공
2019.05	00전기공업	UPS	10KVA	1	m		대리점 > 대리점 > 대리점



year	month	name	item1	item2	count	volumn	kind	category1	category2	category3
2013	1	AA부대	장비	UPS	1	10	s	공공기관	군부대	병무청
2013	3	BB제조업	장비	UPS	1	10	m	민간기업	제조업	기기
2016	12	00통신	점검/수리	설치	1	NaN	m	민간기업	건축및부동산	시공
2019	5	00전기공업	장비	UPS	1	10	m	대리점	대리점	대리점

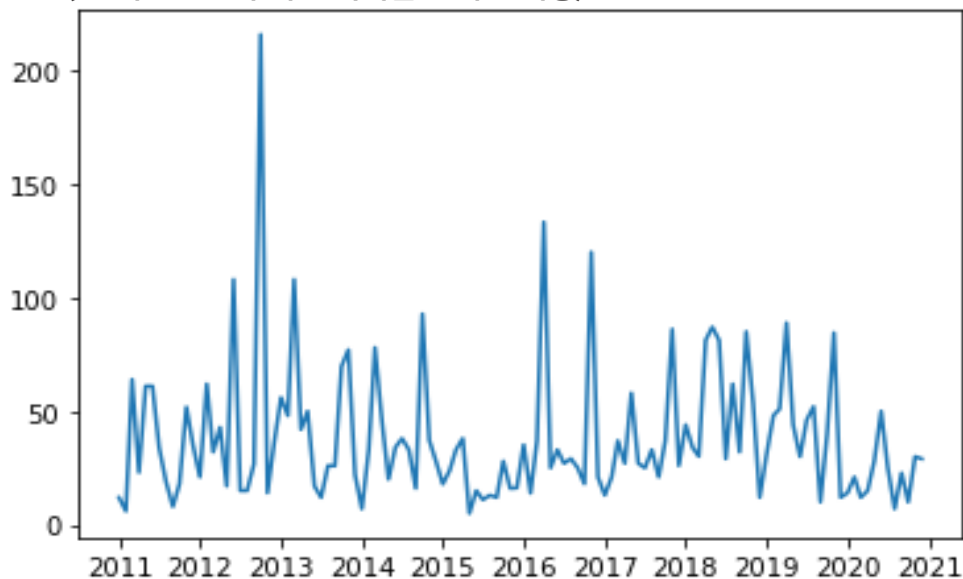
03.데이터 분석 기법 활용

SARIMA

데이터셋 준비 - Raw data에서 item2='UPS'+ 'UPS외' → 1로 변경, 'date_ori'와 'category1'기준으로 묶음

	date_ori	item2	category1	count		date_ori	category1	item2	count
0	2013.01	1	공공기관	1.0	0	2011.01	공공기관	6	12.0
9	2013.05	1	공공기관	1.0	1	2011.01	교육기관	3	3.0
12	2014.12	1	공공기관	1.0	2	2011.01	대리점	3	4.0
13	2019.08	1	공공기관	1.0	3	2011.01	민간기업	9	52.0
15	2014.02	1	민간기업	1.0	4	2011.01	정보통신	3	4.0
18	2014.12	1	공공기관	3.0	5	2011.02	공공기관	6	6.0

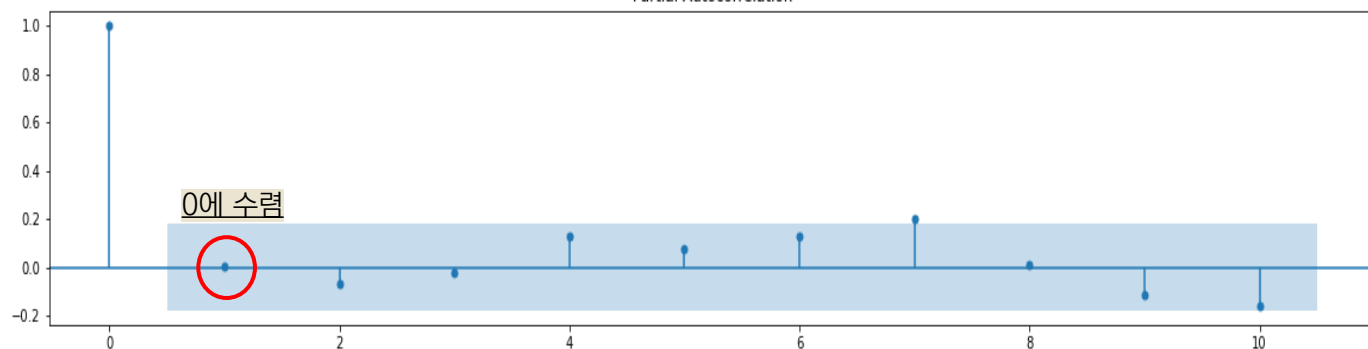
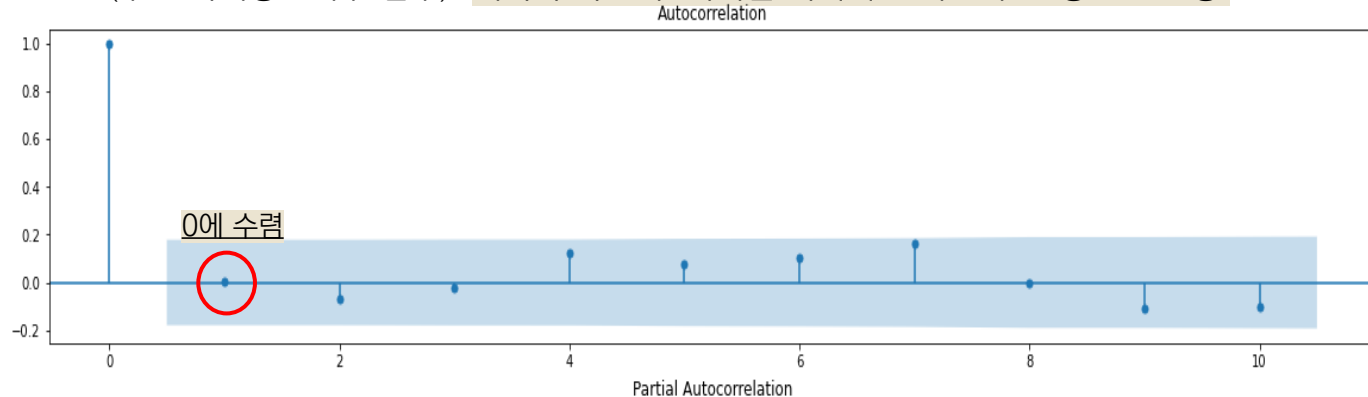
< 전체 DATA에 대한 시계열 그래프 개형 >



모델 선정: ACF, PACF로 확인 후, ARIMA 모델 선정

ACF(자기상관계수 함수): 시차에 따른 일련의 자기상관을 의미
상대적으로 빠르게 0에 수렴하는 모형을 보이면 정상 시계열로 판단.

PACF(부분 자기상관계수 함수): 시차가 다른 두 시계열 데이터 간의 순수한 상호 연관성 판단.



```
ARIMA(0,0,1)(0,0,0)[0] AIC=1097.596, Time=0.02 sec
ARIMA(0,0,2)(0,0,0)[0] AIC=1084.599, Time=0.03 sec
ARIMA(3,0,0)(0,0,0)[0] AIC=1047.217, Time=0.03 sec
```

(0,0,0) = (o, o, o),(o, o, o)[12]
계절성 확인 → SARIMA

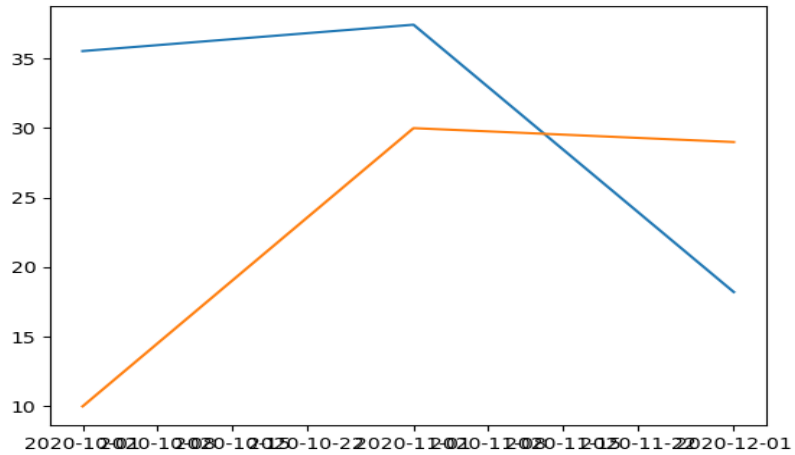
AIC 가 가장 적은 모델을 찾아서 최적의 모형 찾음

Best model: ARIMA(3,0,0)(0,0,0)[0]

03.데이터 분석 기법 활용

SARIMA

📄 SARIMA(1,0,3)(1,0,3,12) 결과 MSE: 274



〈공공기관, 3달치 예측〉



📄 설명

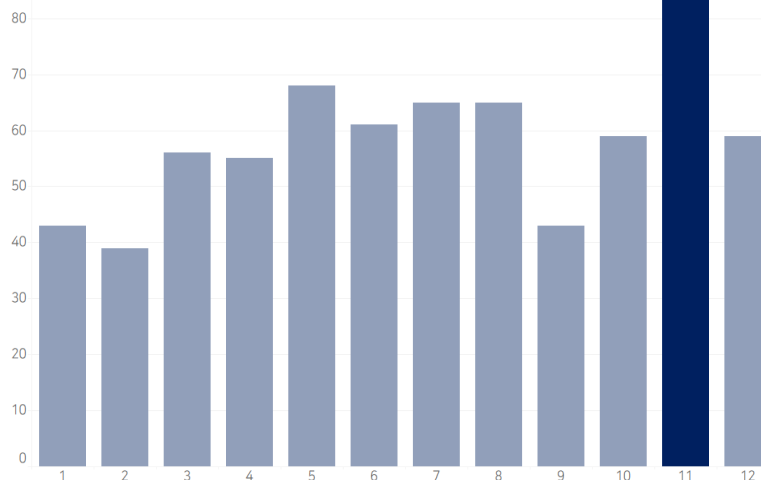
위 그래프에서 단순히 'MSE 낮다 = 모형 좋다' 라고 하기 힘들어 보인다. 결국 27.5로 수렴하는 모습을 보이고 있기 때문

기술 통계, SARIMA

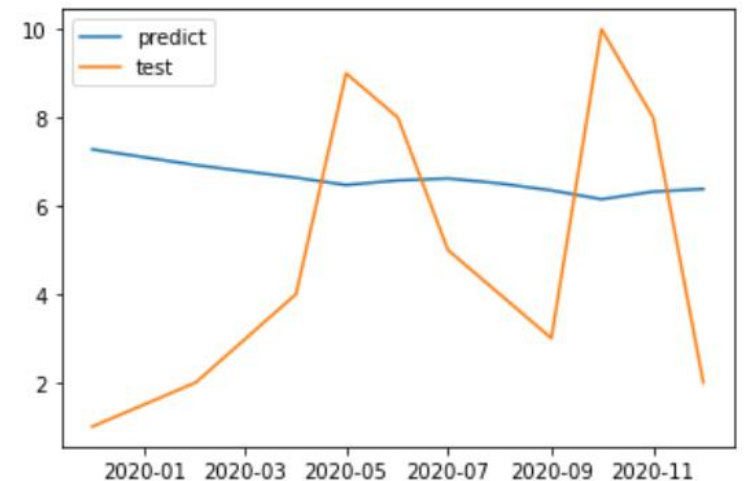
📄 설명

대리점의 경우 2분기와 4분기에 수요가 가장 많음을 보이고 있다. 또한 꾸준한 수주를 통해 안정적인 수업을 만들어내는 것을 볼 수 있다.

〈대리점〉



📄 SARIMA(3,0,0)(3,0,0,12) – 대리점

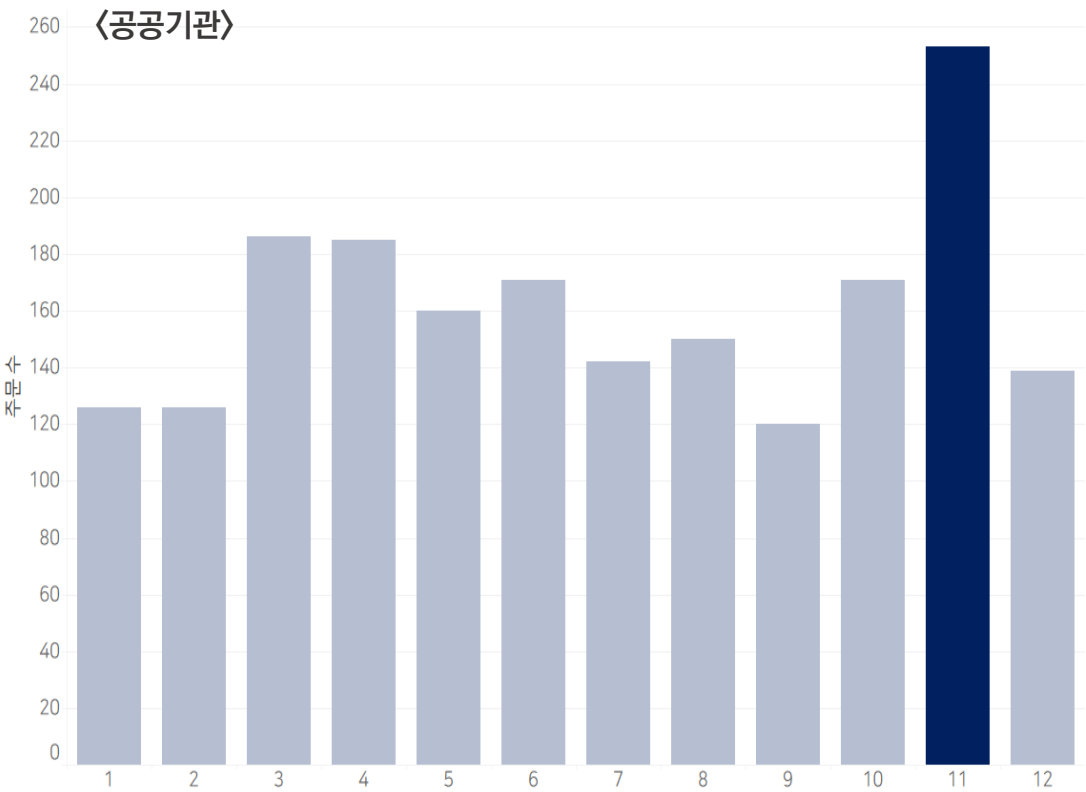


03.데이터 분석 기법 활용

〈정보통신〉

기술 통계, SARIMA

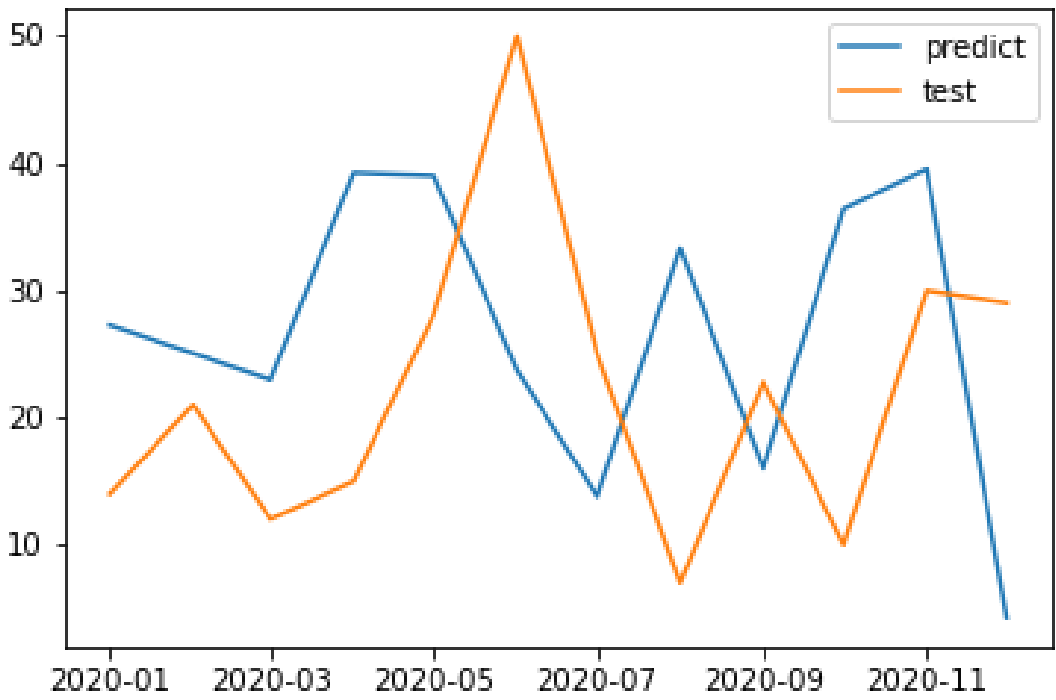
기술통계로 확인한 UPS 수주시기



알림

금융기관, 의료기관, 정보통신의 경우 수가 최고 많은 수주건수가 20건이 되지 않기 때문에 APPENDIX에 넣었음.

SARIMA(3,0,0)(3,0,0,12) – 공공기관



설명

공공기관, 교육기관과 같이 국가에서 직간접적으로 관련이 있는 경우, 11월에 높은 수주건수를 기록하고 있음을 확인.

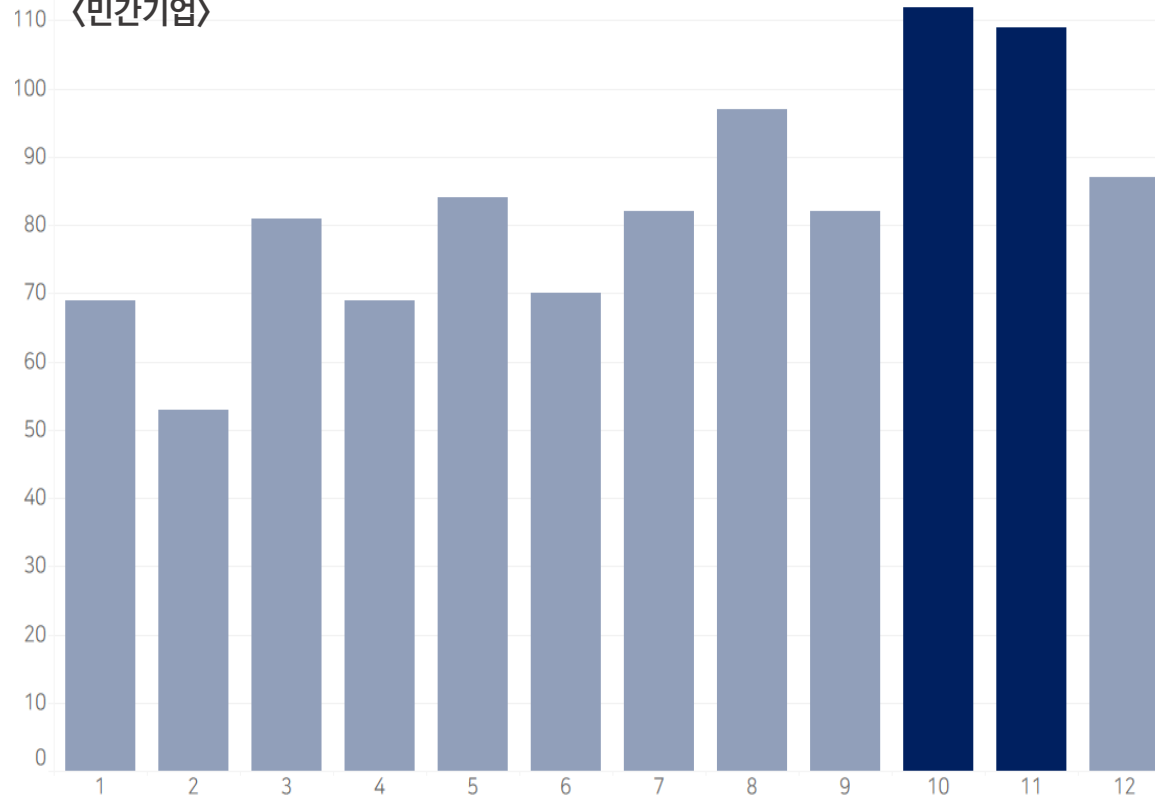
03.데이터 분석 기법 활용

〈정보통신〉

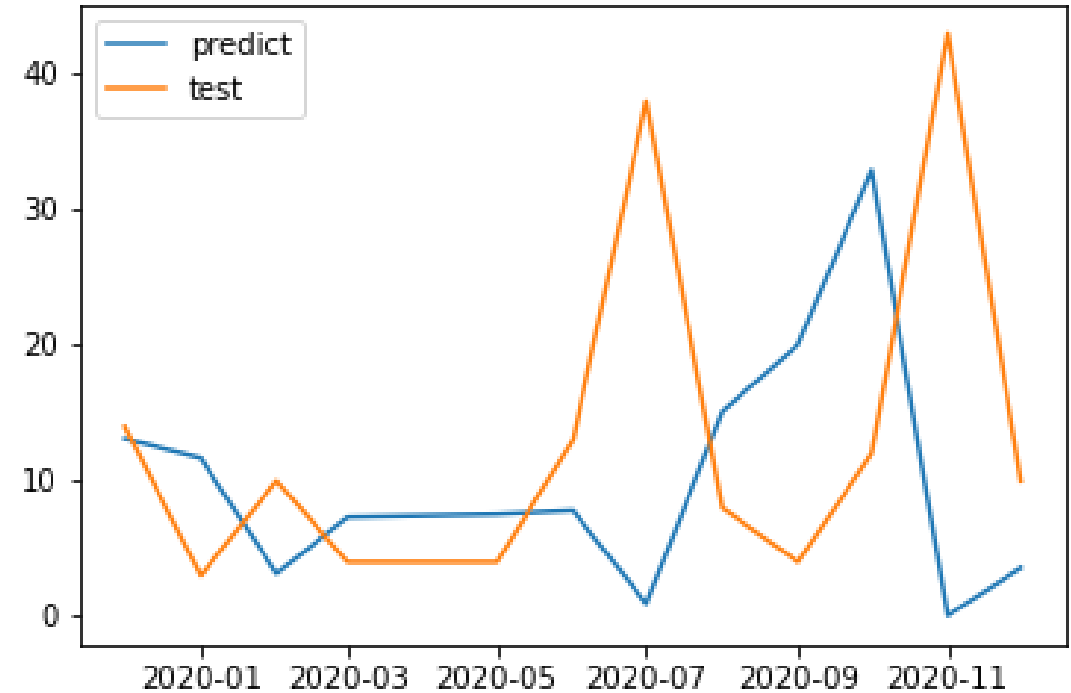
기술 통계, SARIMA

기술통계로 확인한 UPS 수주시기

〈민간기업〉



SARIMA(3,0,0)(3,0,0,12) – 민간기업



설명

민간기업의 경우 4분기에 수주건수가 많은 것을 모델이 잘 예측하고 있다.
기술통계를 확인했을 때 10월, 11월이 가장 수주건수가 많았음을 알 수 있다.

03.데이터 분석 기법 활용

KNN,RNN,LSTM

📖 KNN 알고리즘

K Nearest Neighbor(KNN) 으로 기계 학습 분야에서 가장 간단한 알고리즘 분류하고자 하는 데이터와 가장 가까운 k개의 이웃을 선택하여 분류함.

선택 이유: 동일한 산업일 경우, 선호하는 용량이 비슷할 것이라고 판단.

독립변수: 월(月), 수량, 대분류 **종속변수:** 용량

📖 RNN 알고리즘

한 hidden layer 내의 노드가 순차적으로 연결된 구조이기에 입력 집합의 값들이 시간 등의 순서에 따라 순차적으로 영향을 받을 때 사용.

선택 이유: 10년치의 데이터이기 때문에 오래된 순서대로 입력한다면 순서가 유의미할 것이라고 판단하여 시계열 데이터를 다룰 수 있는 RNN 선택

독립변수: 대분류, 중분류 **종속변수:** 용량

📖 LSTM 알고리즘

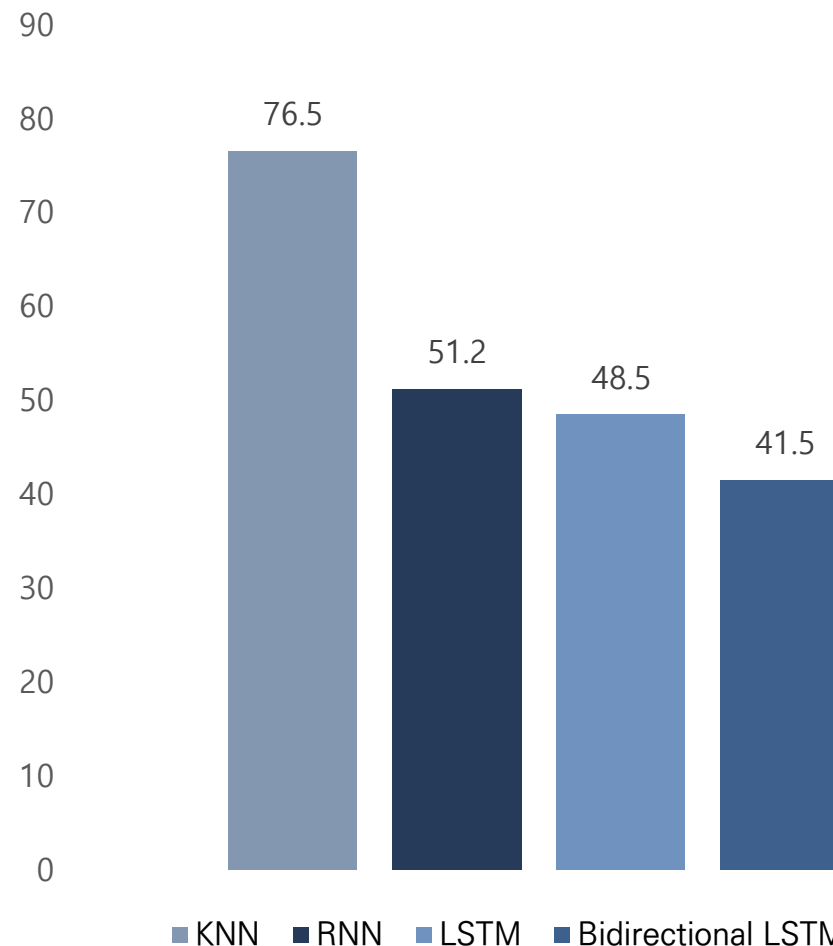
Long Short-Term Memory로 RNN의 단점인 오래 된 정보도 기억할 수 있는 장기간에 걸친 시간의존성 (long-tern dependency)이 있음.

선택 이유: 경사 소실로 인해 시계열 데이터의 초기에 있는 값들은 예측값에 미치는 영향이 현저히 줄어들게 되기에 장기간에 걸친 시간의존성 (long-tern dependency) 은 학습할 수 없기에 보완하고자 LSTM 선택

독립변수: 대분류, 중분류 **종속변수:** 용량

모델 별 정확도

단위: %



03.데이터 분석 기법 활용

KNN

📄 데이터셋 준비 - Raw data에서 kind=s(수주), item2='UPS'로 조건을 걸어 데이터 준비

📄 One-hot-encoding으로 category1 처리 & volume별 category 설정

📄 train, test dataset 분리

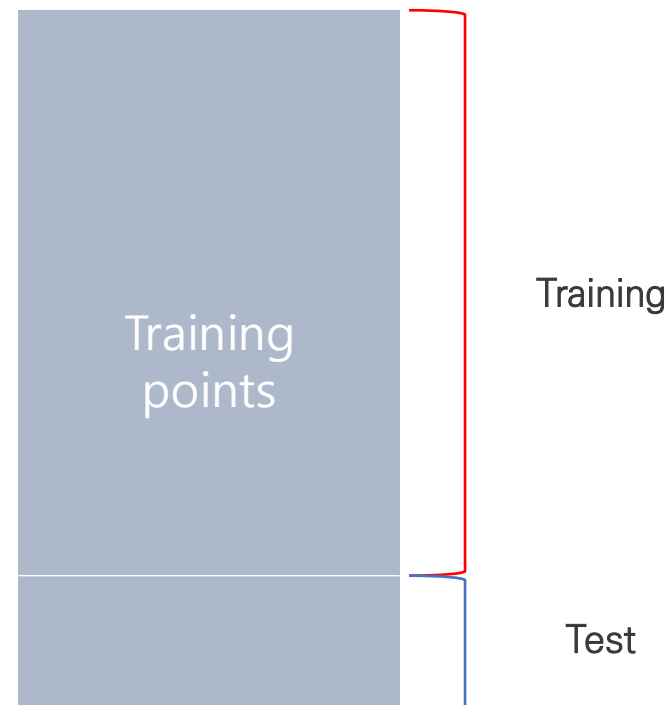
x_train, x_test, y_train, y_test
= train_test_split(x, y, test_size=0.2, random_state=4)

1	2	3	4	5	6
0~30	31~60	61~120	121~200	201~500	500~

〈Volume category〉

Training points										Labels
	month	count_normal	공공기관	교육기관	금융기관	대리점	민간기업	의료기관	정보통신	vol_category
0	1	-0.241797	1	0	0	0	0	0	0	1
9	5	-0.241797	1	0	0	0	0	0	0	1
12	12	-0.241797	1	0	0	0	0	0	0	1
13	8	-0.241797	1	0	0	0	0	0	0	1
15	2	-0.241797	0	0	0	0	1	0	0	1

〈Labels & Training points〉



03.데이터 분석 기법 활용

KNN

📄KNN결과

Accuracy_score: 0.765

Average_recall_score: 0.162

```
#model
classifier = KNeighborsClassifier(n_neighbors = CFG.neighbors)
classifier.fit(x_train, y_train)
guesses = classifier.predict(x_test)
print(confusion_matrix(y_test, guesses))
print("Accuracy_score:", accuracy_score(y_test, guesses))
print("Average_Recall_score:", recall_score(y_test, guesses, average='macro'))
```

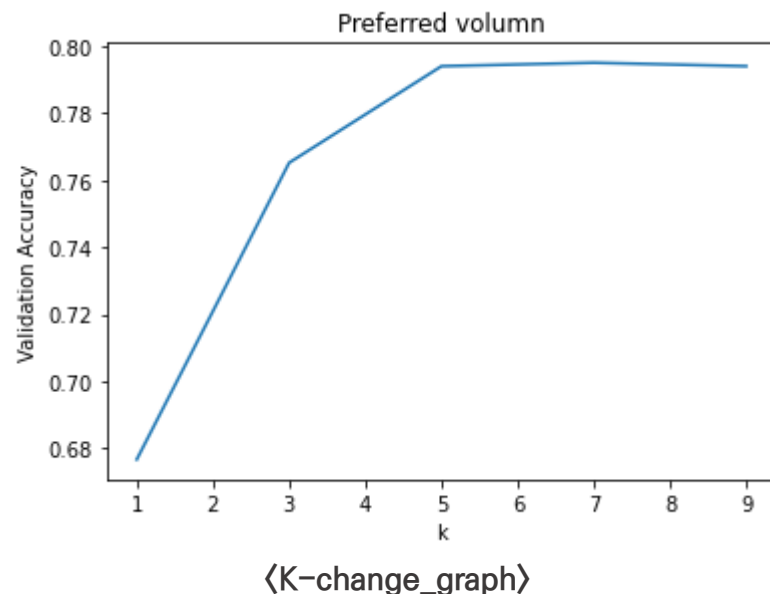
```
[[742  6  3 19  2  0]
 [ 88  0  0  2  0  0]
 [ 58  1  1  2  1  0]
 [ 27  0  0  0  0  0]
 [ 10  0  1  0  0  0]
 [  8  0  0  0  0  0]]
```

Accuracy_score: 0.7651905252317199

Average_Recall_score: 0.16283548537434547

〈KNN-code〉

📄K값의 변화에 따른 Accuracy_score의 변화



📄결론

값이 대부분 1값으로 예측됨에 따라

Accuracy_score는 특정 k이후에 변동폭이 0에 수렴하며,

Recall_score와의 차이가 크기 때문에 좋은 모델이라 보기 힘들다.

03.데이터 분석 기법 활용

RNN, LSTM

📄 데이터셋 준비 - Raw data에서 kind=s(수주), item2='UPS'로 조건을 걸어 데이터 준비

📄 Tokenizer화를 위한 데이터 조정으로 '대분류-중분류'로 변수 생성

〈Volume category〉

1	2	3	4
0~10	11~50	51~100	101~

📄 One-hot-encoding으로 volume별 category 설정
One-hot-encoding의 경우,
범주가 많으면 오히려 성능이 떨어질 수 있기에 4개로 한정

📄 train, test dataset 분리
x_train, x_test, y_train, y_test
= train_test_split(x, y, test_size=0.2, random_state=10)

📄 Tokenizer화를 하는 이유: 산업 분류가 목적이기에 **밀집 벡터**를 이용해
분류하는 성능을 높이는 것으로 목적 함수를 설정

〈Tokenizer 결과〉

```
#tokenizer
from tensorflow.keras.preprocessing.text import Tokenizer
np.random.seed(7)
tf.random.set_seed(7)
tokenizer = Tokenizer(num_words=CFG.max_words) # 빈도가 높은 2000개의 단어를 선택하도록 객체 생성
tokenizer.fit_on_texts(joinlist) #단어 인덱스 구축

X = tokenizer.texts_to_sequences(joinlist)
word_index = tokenizer.word_index
word_docs = tokenizer.word_counts

print('num of tokens:', len(tokenizer.word_index)) #토큰 갯수
for index, docs in enumerate([word_index, word_docs]):
    print(index, docs)
```

```
num of tokens: 22
0 {'공공기관': 1, '대리점': 2, '기관': 3, '민간기업': 4, '공사': 5, '공단': 6, '제조업': 7, '교육기관': 8, '학교1': 9, '건축및부동산': 10, '기타서비스': 11, 'c2c': 12, '정보통신': 13, '관공서': 14, '군부대': 15, '교육시설': 16, '의료기관': 17, '학교2': 18, '금융기관': 19, '병원': 20, '관할의료기관': 21, '기간산업': 22}
1 OrderedDict([('민간기업', 1089), ('건축및부동산', 279), ('공공기관', 2206), ('공사', 713), ('공단', 713), ('정보통신', 228), ('기관', 1169), ('교육기관', 547), ('교육시설', 122), ('기타서비스', 240), ('c2c', 240), ('학교2', 91), ('군부대', 140), ('제조업', 570), ('대리점', 1500), ('학교1', 334), ('관공서', 166), ('의료기관', 110), ('병원', 66), ('금융기관', 78), ('관할의료기관', 44), ('기간산업', 18)])
```

03.데이터 분석 기법 활용

RNN, LSTM

RNN결과

Accuracy_score: 0.512

Average_recall_score: 0.312

```
31/31 [=====] - 0s 6ms/step - loss: 1.2269 - acc: 0.5118
#Test accuracy:0.512
[[348 220 15 11]
 [108 141 14 4]
 [ 22 39 6 2]
 [ 20 16 3 2]]
0.5118434603501545
0.31242137076076093
```

Bidirectional-LSTM 결과

Accuracy_score: 0.415

Average_recall_score: 0.317

```
31/31 [=====] - 1s 6ms/step - loss: 1.2634 - acc: 0.4150
#Test accuracy:0.415
-----
[[328 79 11 176]
 [ 92 51 6 118]
 [ 19 13 2 35]
 [ 11 6 2 22]]
0.41503604531410915
0.32719266531091096
```

LSTM 결과

Accuracy_score: 0.485

Average_recall_score: 0.339

```
31/31 [=====] - 0s 4ms/step - loss: 1.2591 - acc: 0.4851
#Test accuracy:0.485
-----
[[328 206 11 49]
 [ 92 129 6 40]
 [ 19 35 2 13]
 [ 11 16 2 12]]
0.4850669412976313
0.3392507634199819
```

결론

값이 1~4 사이로 고루 찍히지만 정확도가 겨우 50%를 넘으며,
Recall_score와 앞서 knn보다 차이가 적지만 0에 수렴하므로
좋은 모델이라고 보기 힘들다.

03.데이터 분석 기법 활용

기술 통계

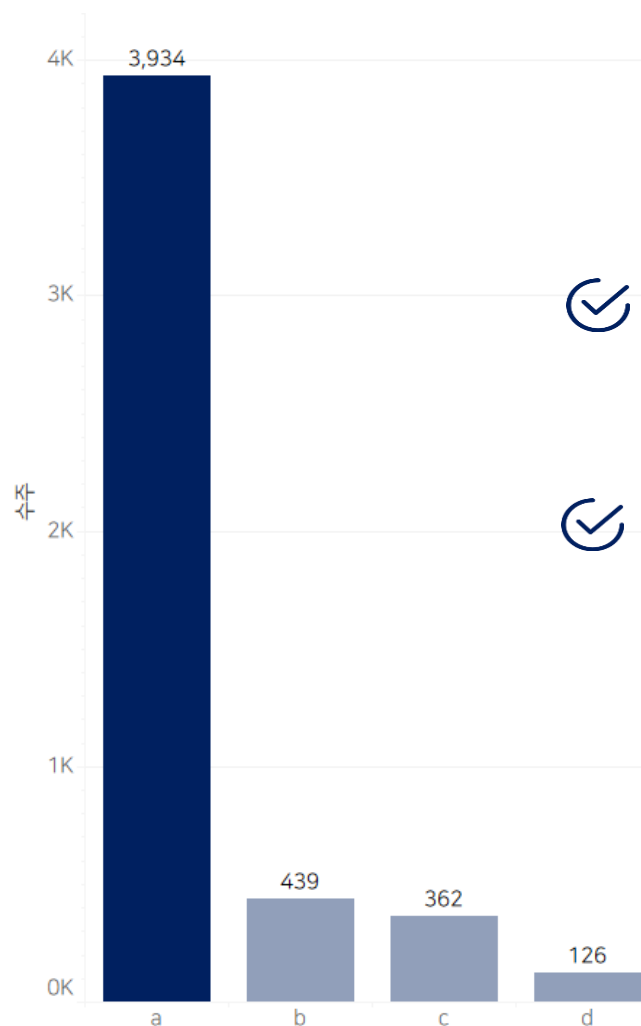
📖 용량 분류 기준(용량별 외함 크기 기준)

Label	용량(kva)
a	0 ~ 30 이하
b	30 ~ 60 이하
c	60 ~ 150 이하
d	150 초과

📖 카테고리 분류

대분류	중분류
공공기관	공사·공단, 기관, 관공서, 기간산업, 군부대
교육기관	학교1, 학교2, 교육시설
금융기관	금융기관
민간기업	건축및부동산, 제조업, 기타서비스(C2C)
의료기관	관할 의료기관, 병 · 의원
정보통신	정보통신
대리점	대리점

📖 용량 전체 분포도



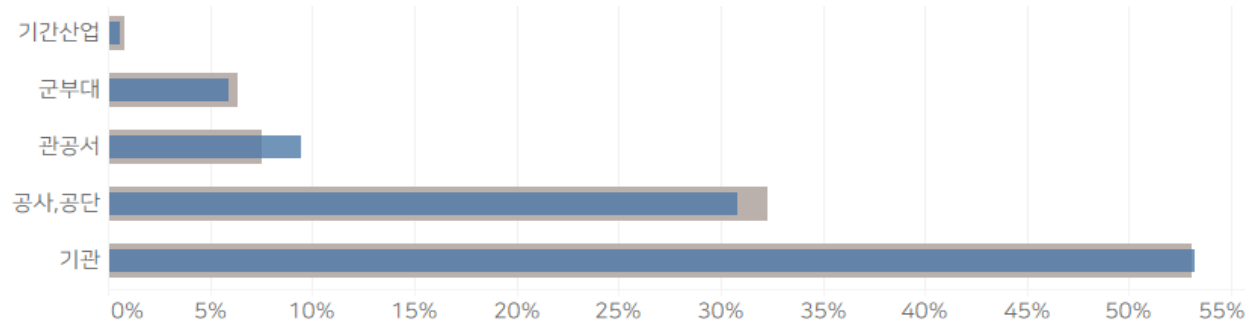
✓ 앞서 모델에서 살펴봤듯이, 적은 용량의 UPS를 선호하고 있음.

✓ 금융기관과 정보통신의 경우 표본 수가 적고, 대리점의 경우 용량 전체의 분포도와 유사한 형태를 보이기에 생략.

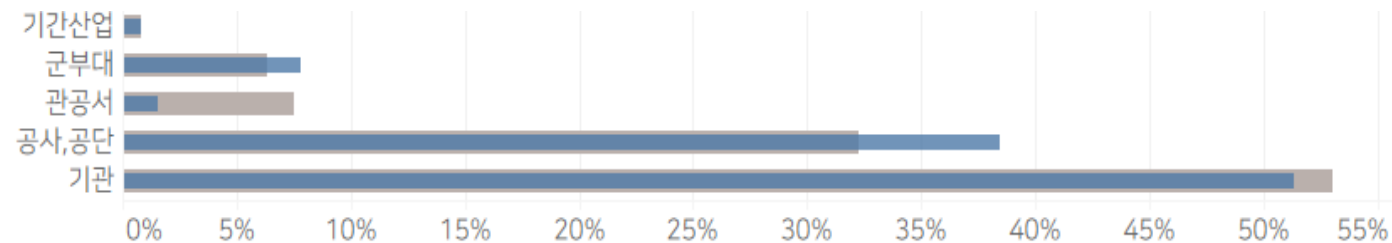
03.데이터 분석 기법 활용

기술 통계 中 공공기관

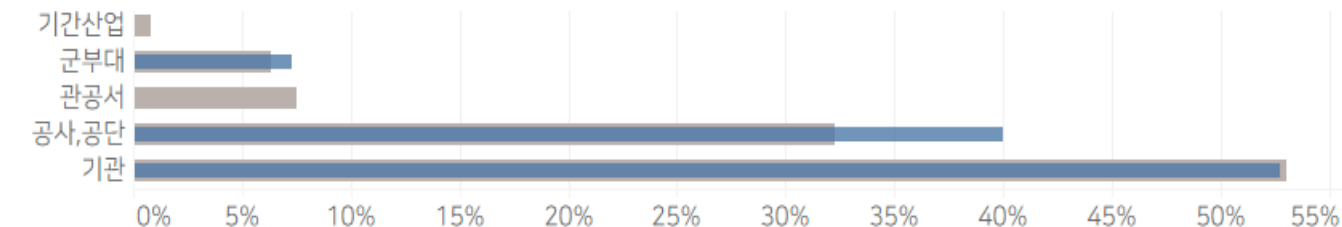
공공기관 中 'a'의 비율



공공기관 中 'b'의 비율



공공기관 中 'd'의 비율



회색: 공공기관에서의 비율
파랑: 공공기관에서의 용량 Label 비율
회색 < 파랑 : 전체 수주에서 잦은 조건의 용량

관공서: a(0~30kva) 용량 선호
군부대: b(30~60kva), d(150kva~) 선호
기간산업: 대체적으로 적은 용량인 a,b 선호
공사,공단: b 이상
기관: 다양한 용량 선호, 대체적으로 a 선호

관공서와 기간 산업의 경우 150kva 초과되는 대용량 UPS 주문이 존재하지만 니즈의 강세가 보이지 않고 있음.

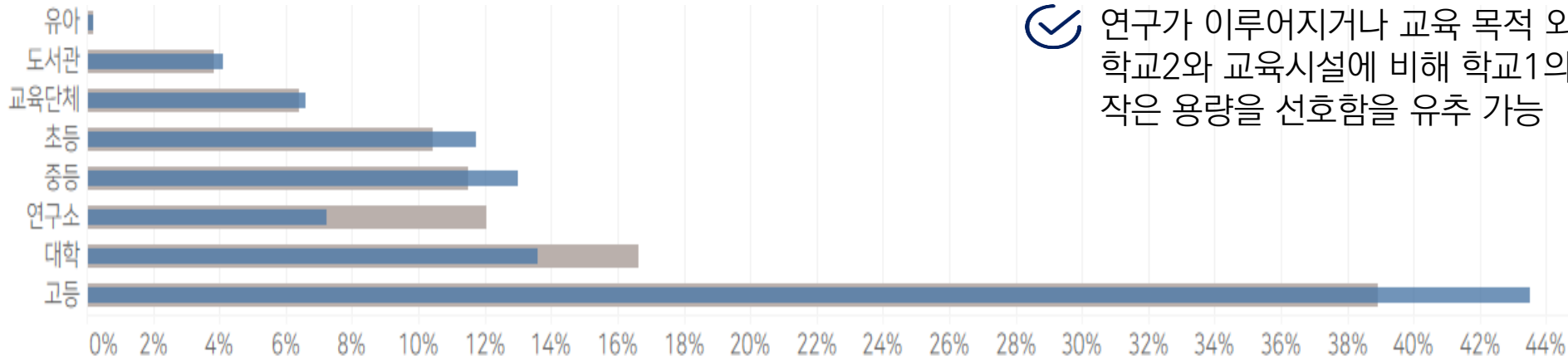
03.데이터 분석 기법 활용

선택 용량

- * 학교1: 유,초,중,고등학교
- * 회색: 교육기관에서의 비율
- 학교2: 대학교
- 파랑: 교육기관에서의 용량 Label 비율
- 교육시설: 교육단체, 연구소, 도서관
- 회색 < 파랑 : 전체 수주에서 찾은 조건의 용량

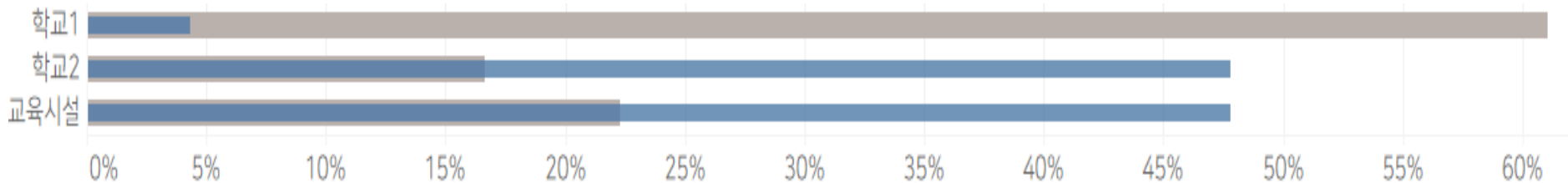
기술 통계 中 교육기관

교육기관 中 'a'의 비율



연구가 이루어지거나 교육 목적 외의 시설을 갖고 있는 학교2와 교육시설에 비해 학교1의 경우, 단순한 업무 중심으로 작은 용량을 선호함을 유추 가능

교육기관 中 'b'의 비율



고등학교의 경우, UPS 수주 양이 많은 이유는 듣기 평가를 포함한 국가 시험이 이루어지는 장소로 전기 공급 관리 중요 ⇒ UPS 수요 ↑

03.데이터 분석 기법 활용

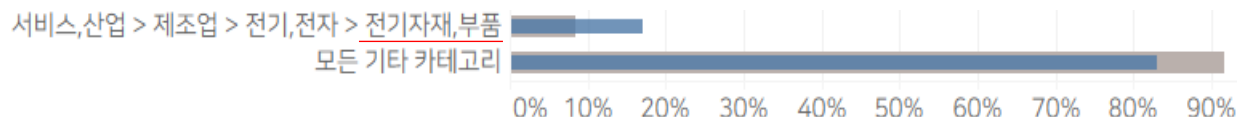
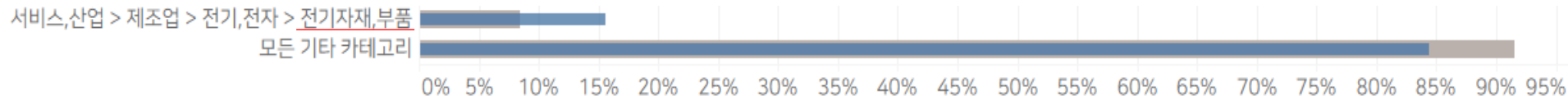
선호 용량

* 건축및부동산 : 건축, 시공
제조업 : 화학, 기기
기타서비스(C2C) : 관리,문화,판매

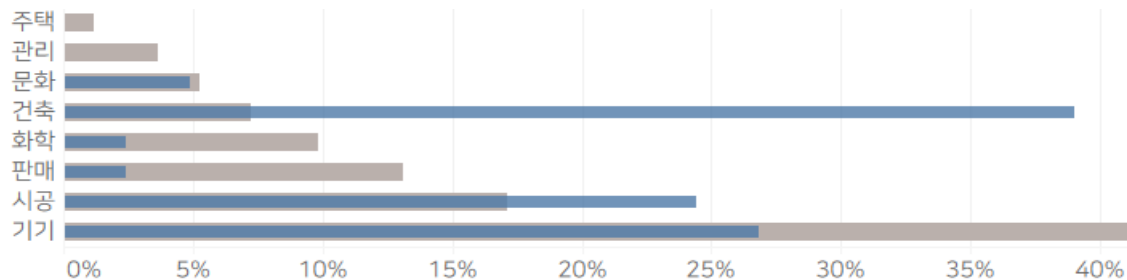
* 회색: 민간기업에서의 비율
파랑: 민간기업에서의 용량 Label 비율
회색 < 파랑 : 전체 수주에서 찾은 조건의 용량

기술 통계 中 민간기업

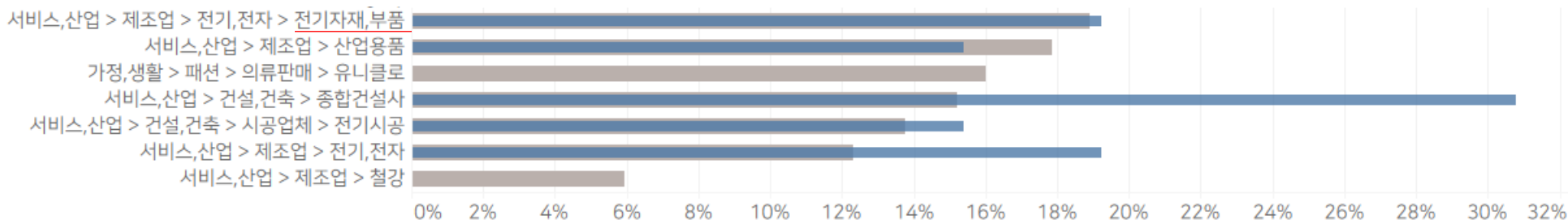
민간기업 기준 'b'의 비율(위), 'd'의 비율(아래)



민간기업의 소분류 기준 'd'의 비율



민간기업 中 'c'의 비율



전기자재, 부품을 생산하는 제조업체는 b,c,d 선호 중 b와 d에서 뚜렷한 니즈를 보이고 있음.

새로 생성한 소분류 기준에서 d는 건축과 시공업에서 확실한 니즈를 보이고 있고 c의 경우 종합건설사에서도 강한 니즈를 확인할 수 있다.
따라서 전기 시공과 전기, 전자를 포함한 전기 관련 산업이 아닌 민간기업에서는 건설/건축업에 주목해야 함.

03.데이터 분석 기법 활용

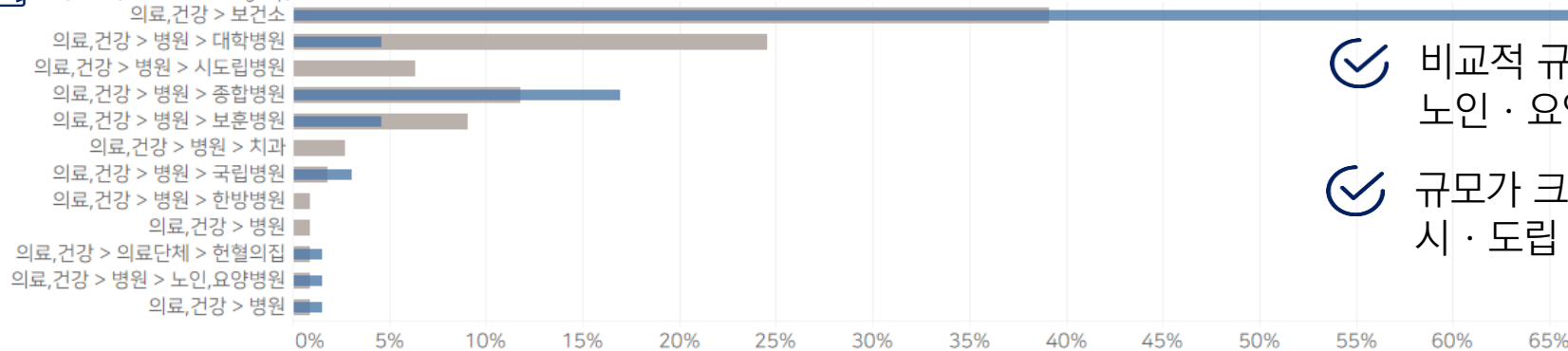
선호 용량

* 관할의료기관: 병·의원 이외 모든 것
병·의원: '병원'으로 끝나는 분류명

* 회색: 의료기관에서의 비율
파랑: 의료기관에서의 용량 Label 비율
회색 < 파랑: 전체 수주에서 찾은 조건의 용량

기술 통계 中 의료기관

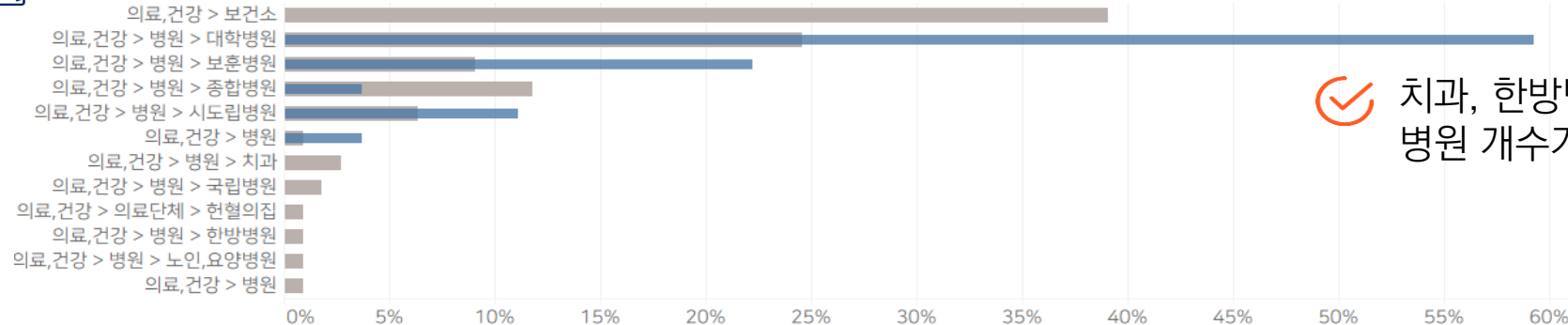
의료기관 中 'a'의 비율



비교적 규모가 작다고 판단되는 헌혈의 집, 노인·요양병원, 보건소에서는 a(0~30kva) 선호

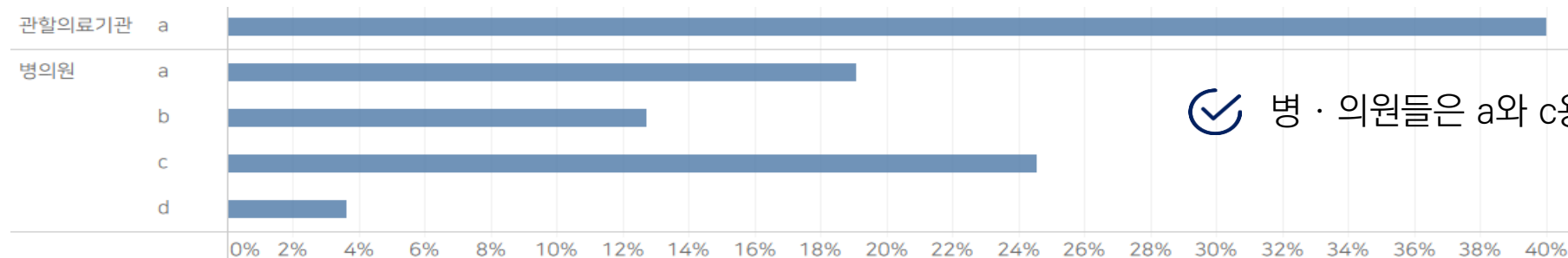
규모가 크다고 판단되는 대학병원, 보훈병원, 시·도립 병원의 경우 c(60~150kva) 선호

의료기관 中 'c'의 비율



치과, 한방병원의 경우, 다른 의료기관에 비해 수주량과 병원 개수가 적기에 뚜렷한 양상을 보이고 있지 않음.

의료기관 중분류 기준 용량 분포



병·의원들은 a와 c용량을 선호하고 있음을 확인 가능

03.데이터 분석 기법 활용

Python 기본 문법 이용

수주~매출 소요기간 분석

기준: name_ori가 같으면 매출 '년월'에서 수주 '년월'을 제거

우려사항: 수주데이터는 있는데 매출데이터가 없는 경우 기간이 크게 잡힐 수 있다.

한계점 : name_ori(납품처명) 기록할 때, 수주와 매출을 다르게 했으면 분석 불가능



해결책: '조달정보개방포털' 에서 공공기관 UPS 조달 내역 데이터 사용

Df = Raw data(제공받은 데이터)

	name_ori	date_ori	item_ori	item1	item2
0		2013.01	UPS	장비	UPS
1		2013.03	UPS	장비	UPS
3		2019.05	UPS	장비	UPS
9		2013.05	UPS	장비	UPS
10		2013.07	UPS	장비	UPS



New_Df = 물품 계약 상세내역(출처: 조달정보개방포털)

	수요기관명	최초계약일자	최대납품기한	업체명
0		20030819	20040920	
1		20030825	20041231	
2		20030901	20031130	
3		20030924	20031030	
4		20031013	20040111	

03.데이터 분석 기법 활용

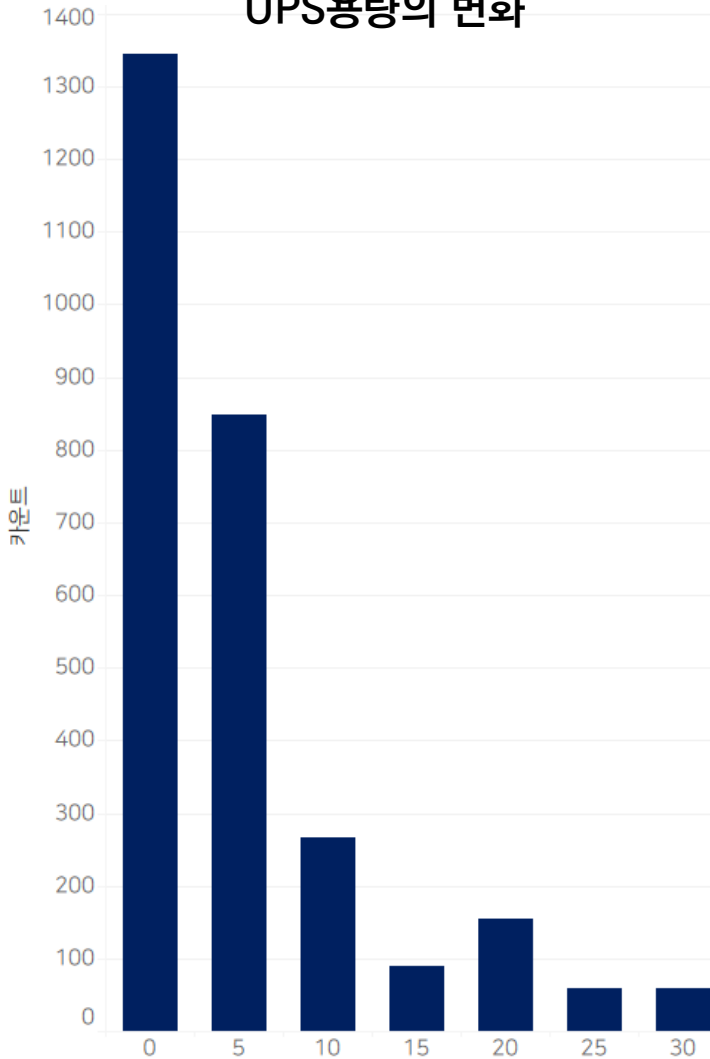
Python 기본 문법 이용

📄 UPS 용량의 변화를 보았을 때
0~5KVA 사이의 용량변화가 가장 많음.

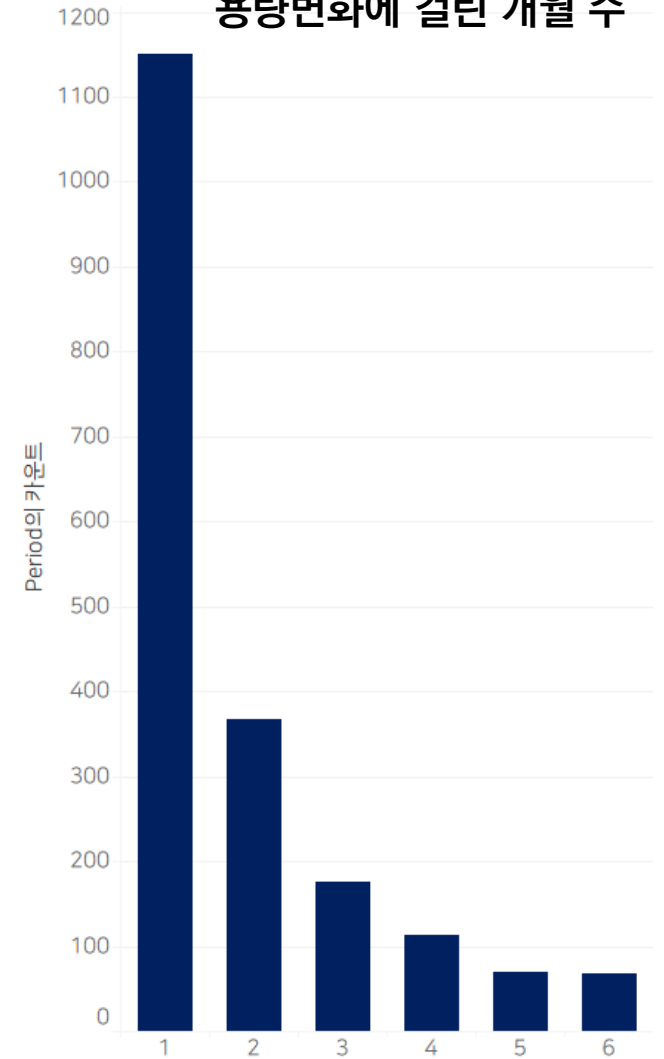
📄 용량변화가 가능했던 기간이 6개월 미만의
경우에서 나타났으며 1개월의 경우에 약
60%를 차지하며 두드러짐.

📄 UPS 수주 ~ 매출 기간의
대부분(87%)은 3개월 이하이다.

UPS용량의 변화



용량변화에 걸린 개월 수



結.영업 전략 제안

용량

의료기관: 0 ~ 30kva , 60kva ~ 150kva

학교1, 관공서, 기관: 0 ~ 30kva

기간산업: 0 ~ 60kva

군부대: 30kva ~ 60kva 이상 또는 150kva 초과

학교2, 연구시설, 공사및공단: 30kva이상

2분기, 4분기

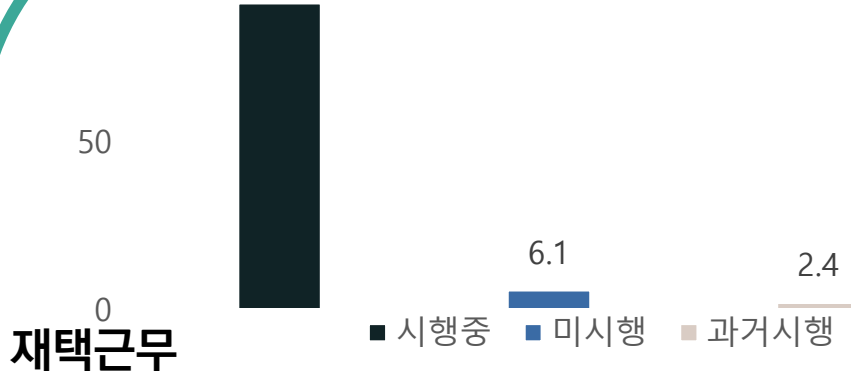
산업 : 금융기관, 의료기관 (5월)

정보통신, 교육기관(10월)

공공기관, 대리점(11월)

민간기업(10월, 11월)

매출 100대 기업 재택근무 시행 현황 단위: %



코로나19 전에는 가정에서의 UPS의 쓸모가 없었지만, 코로나 19가 터지고서부터 작고 감성적인 색으로 공략한다면 새로운 시장을 개척할 수 있을 것이라 판단

UPS의 가정화

結.영업 전략 제안

규제 완화



1. 한강변 첫 주동 높이 규제 해제
→ 여의도가 그 대상

[김정환, 여의도 아파트 15층 규제 없앤다...성남갑 대신 다채로운 초고층 아파트, 매일경제, 2021-07-27](#)

2. 종합 건설사 UPS 주요 용량은 고용량
3. 수익 증가 가능성 ↑

신도시 증가

수도권 신규 공공택지 위치



종합건설사

* 주동: 한강변에서 가장 가까운 아파트 동

감사합니다

01 Q & A

02 APPENDIX

APPENDIX – 산업 카테고리 소분류 포함

대분류	중분류	소분류
공공기관	공사,공단 기관 관공서 기간산업 군부대	공사, 공단 부,처/도·시·군·구청,단체 및 협회, 상하수도, 행정 경찰서,세무서,소방서,우체국 방송, 에너지, 항공, 해양, 도로 병무청
교육기관	학교1 학교2 교육시설	유치원, 초등학교, 중학교, 고등학교 대학 교육단체(문화센터 포함), 도서관, 연구소
금융기관	금융기관	금융(증권 포함), 은행
민간기업	건축및부동산 제조업 기타서비스(C2C)	건축,시공 기기,화학 관리,문화,판매
의료기관	관할의료기관 병의원	보건소, 관할의료기관 국공립, 사립병원
정보통신	정보통신	IT 기업, 정보통신
대리점	대리점	대리점

APPENDIX – 산업 카테고리 소분류 포함

대분류	중분류	내용
장비	UPS	UPS,UPS 임대
	UPS 외	INV(인버터) 충전기, 주파수변환기, 정류기, AVR, SNMP, 모드버스통신
배터리	배터리	축전지, 축전지교체
점검/수리	점검/유지보수	UPS점검,시운전, 유지보수, 오버홀
	수리	UPS수리, A/S, 부품수리, FAN교체,부품교체
	설치	추가작업,전기공사,이전설치
	교육	교육
부품	부품	회계용어, 부품, 기타 장치

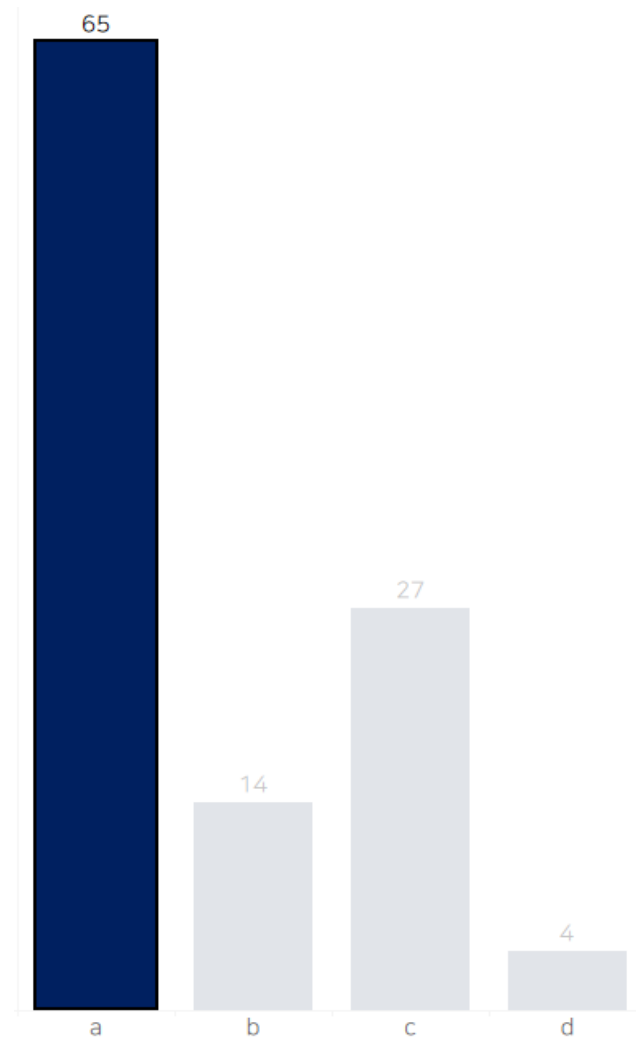
APPENDIX – 선호 용량

카테고리 전체, 수주건수

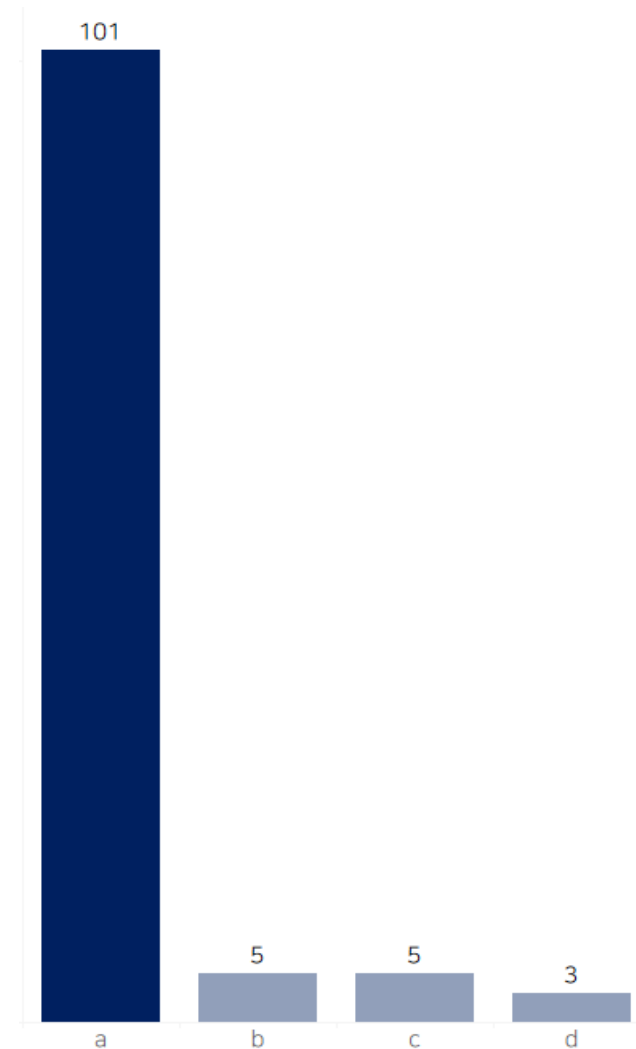
교육기관



의료기관



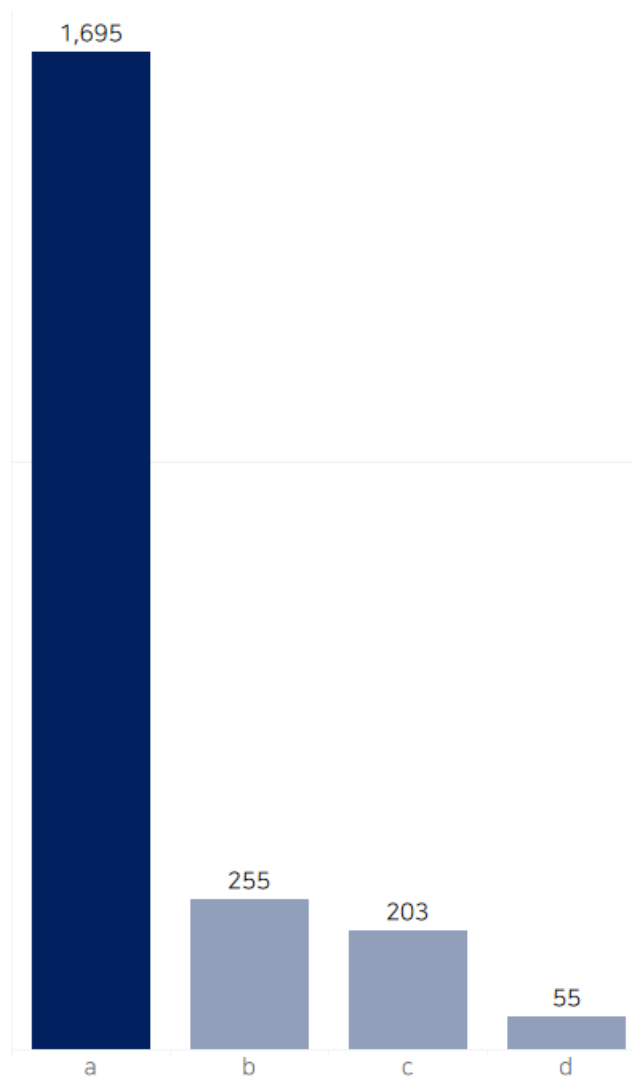
정보통신



APPENDIX – 선호 용량

카테고리 전체, 수주건수

공공기관



금융기관



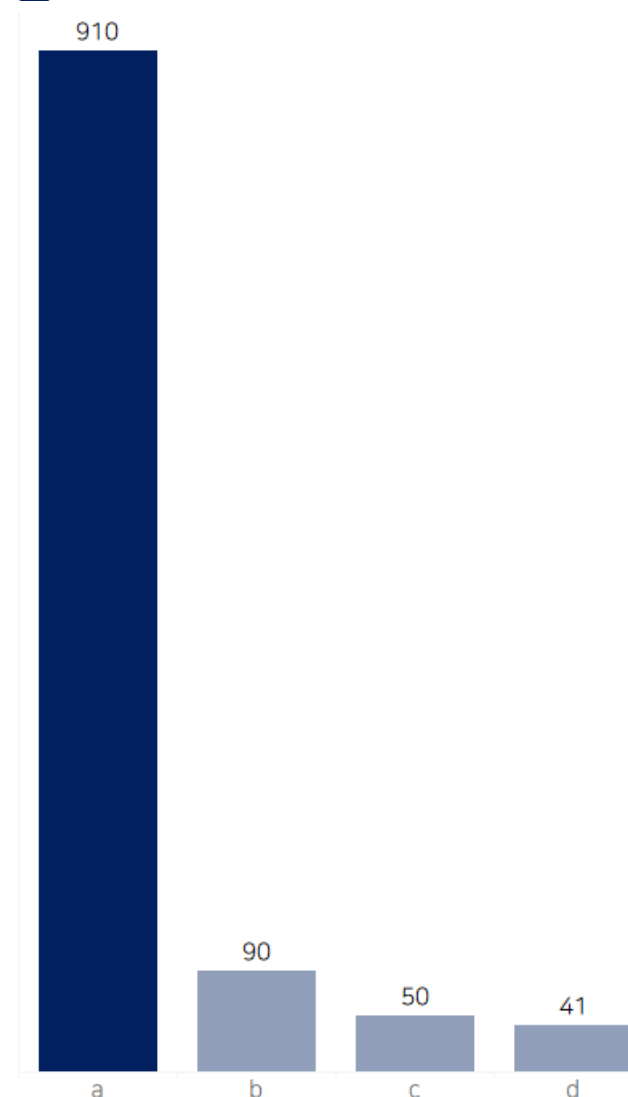
APPENDIX – 선호 용량

카테고리 전체, 수주건수

대리점



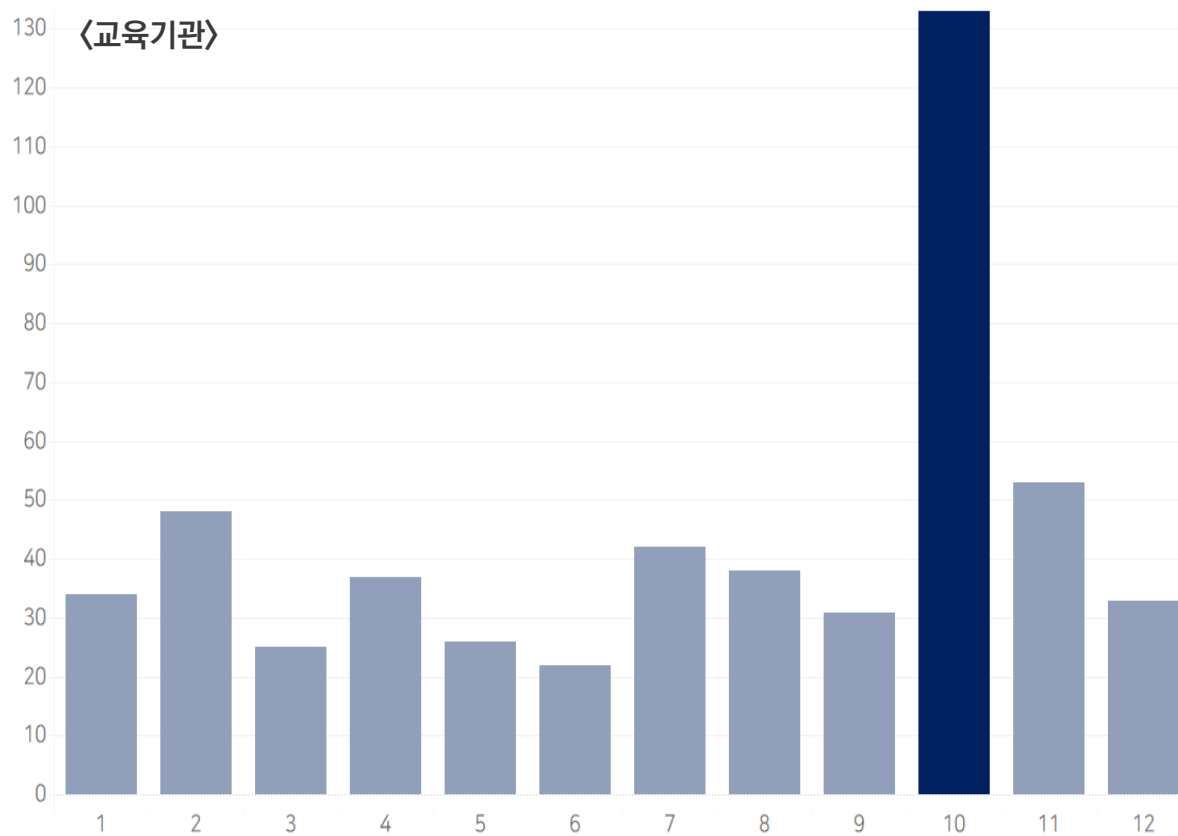
민간기업



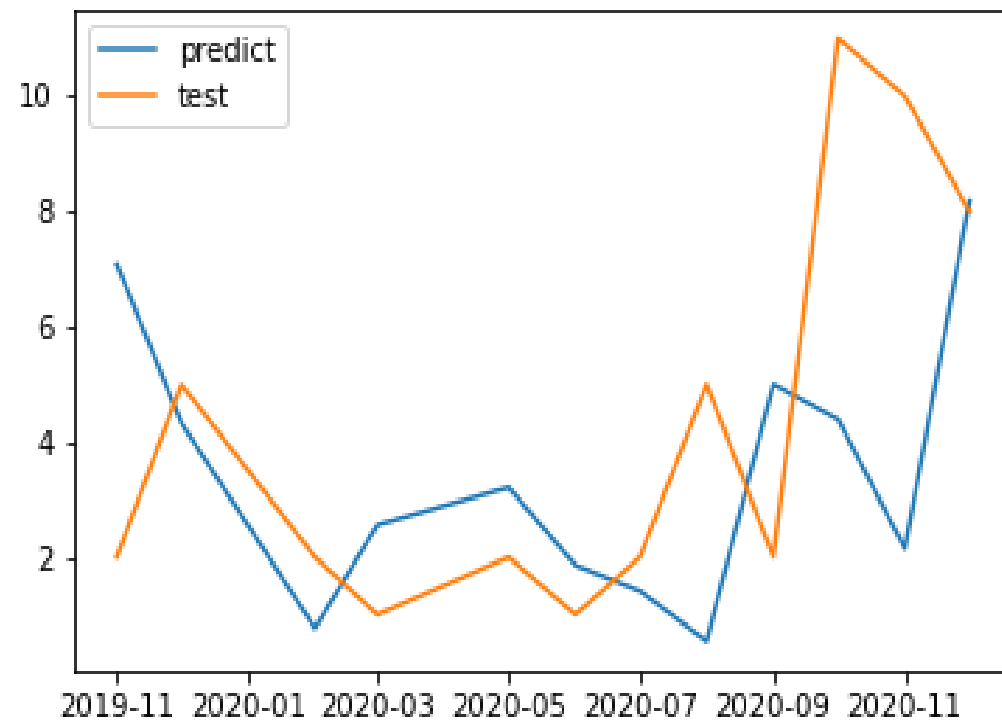
APPENDIX – 교육기관 UPS 수주시기

기술통계

기술통계로 확인한 UPS 수주시기



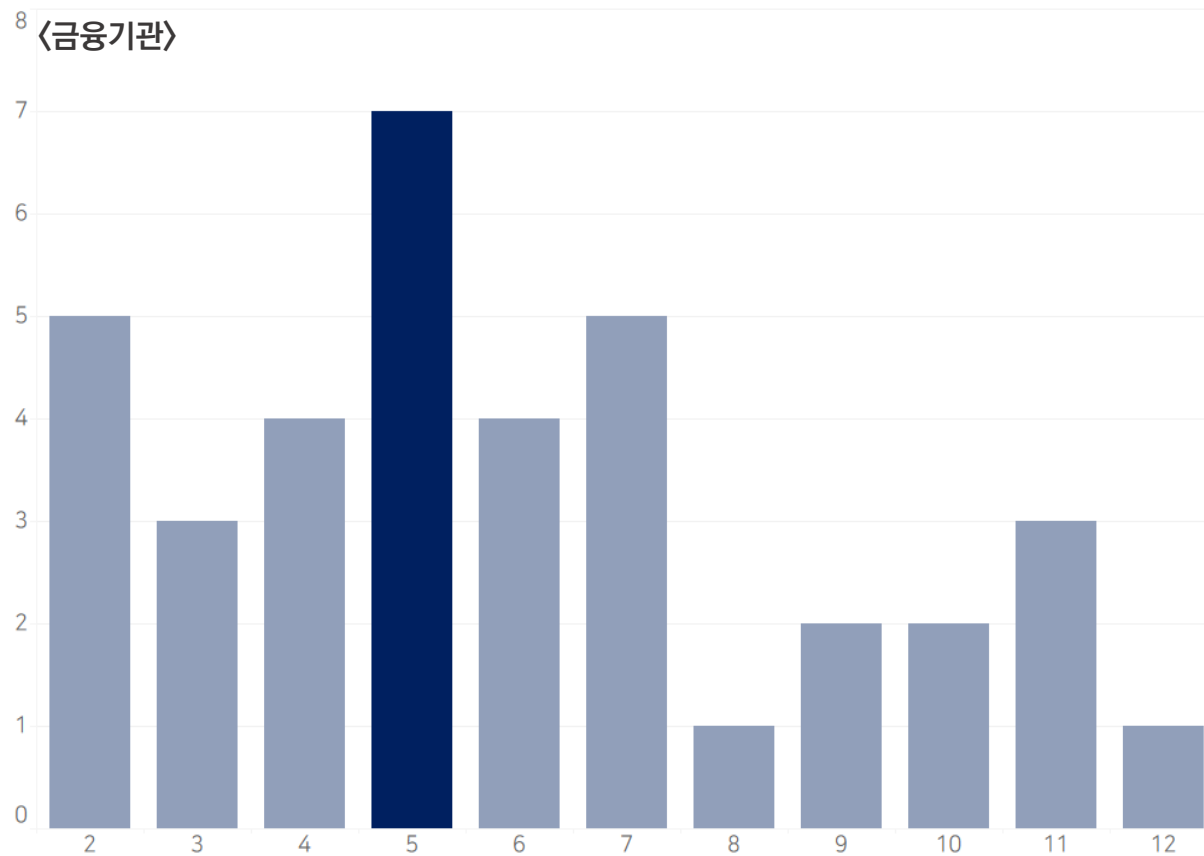
SARIMA(3,0,0)(3,0,0,12) – 교육기관



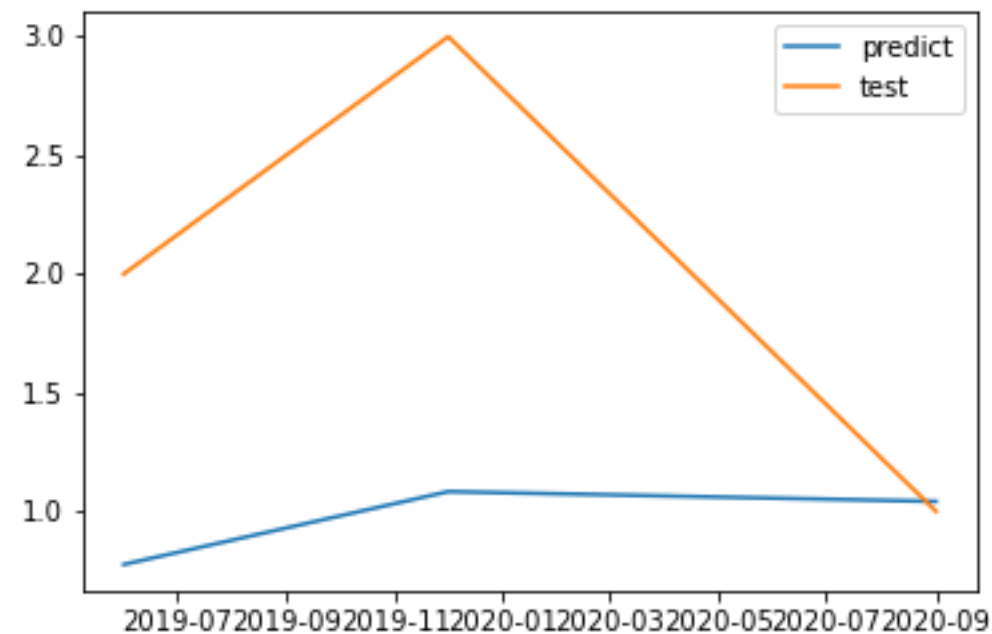
APPENDIX – 금융기관 UPS 수주시기

기술통계

기술통계로 확인한 UPS 수주시기



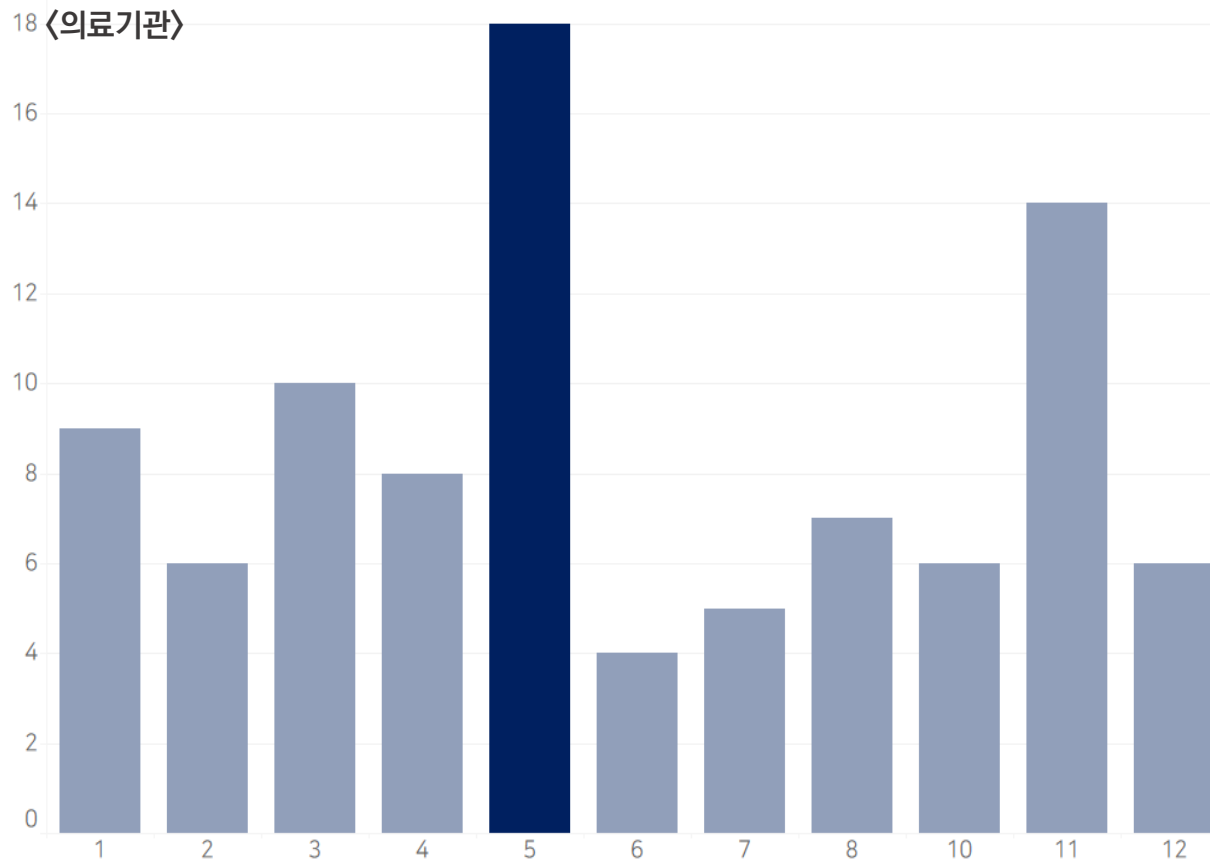
SARIMA(3,0,0)(3,0,0,12) – 금융기관



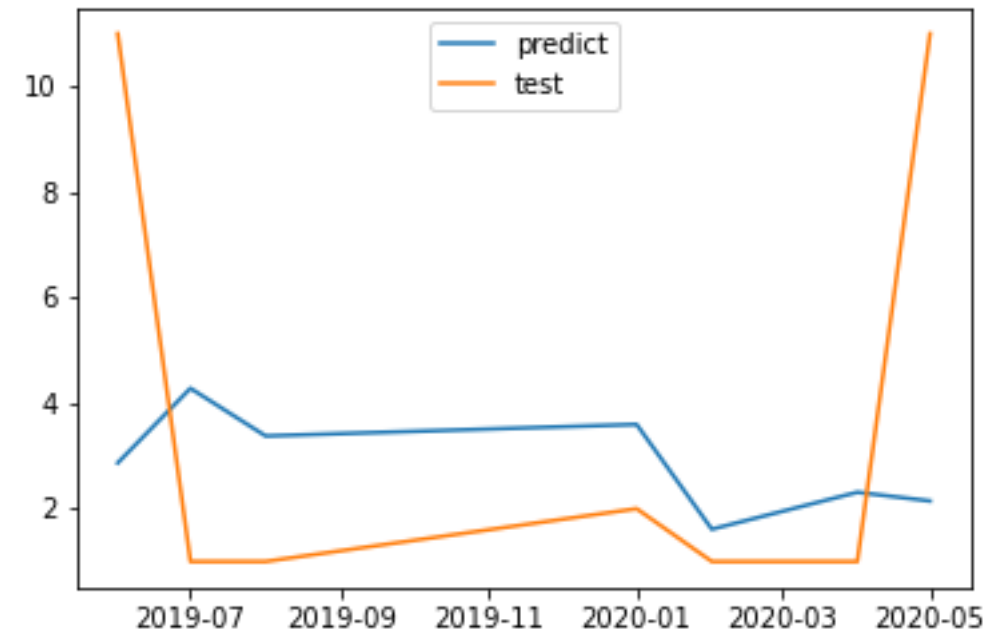
APPENDIX – 의료기관 UPS 수주시기

기술통계

기술통계로 확인한 UPS 수주시기



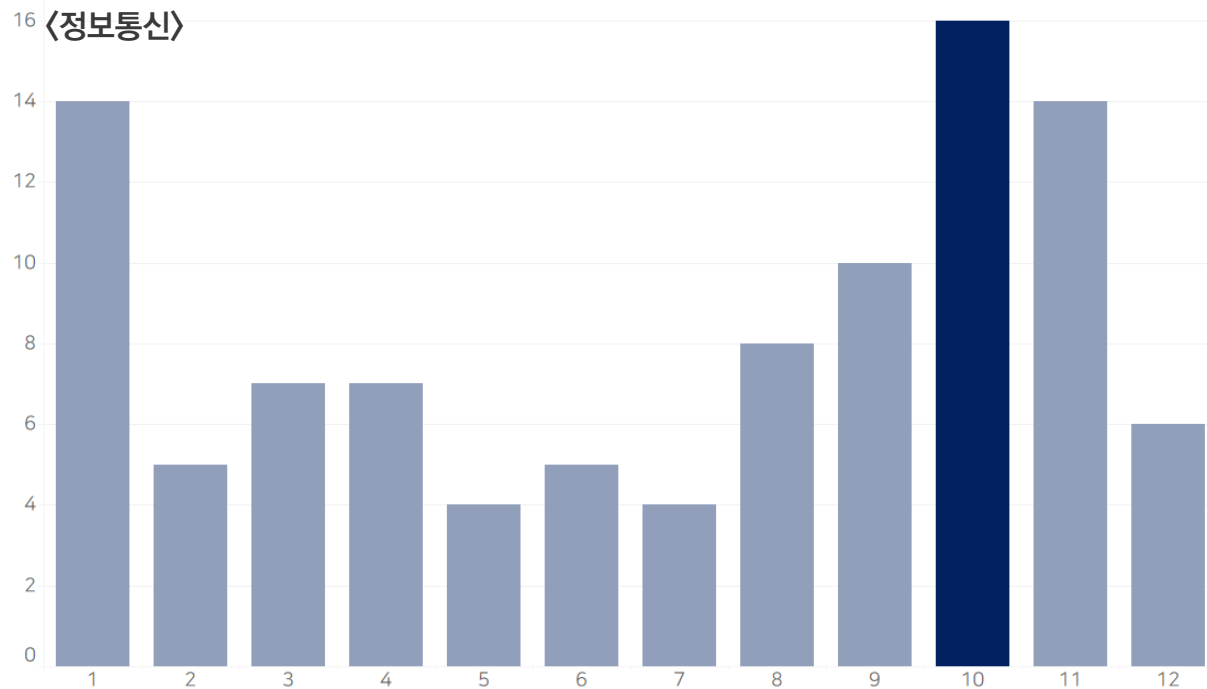
SARIMA(3,0,0)(3,0,0,12) – 의료기관



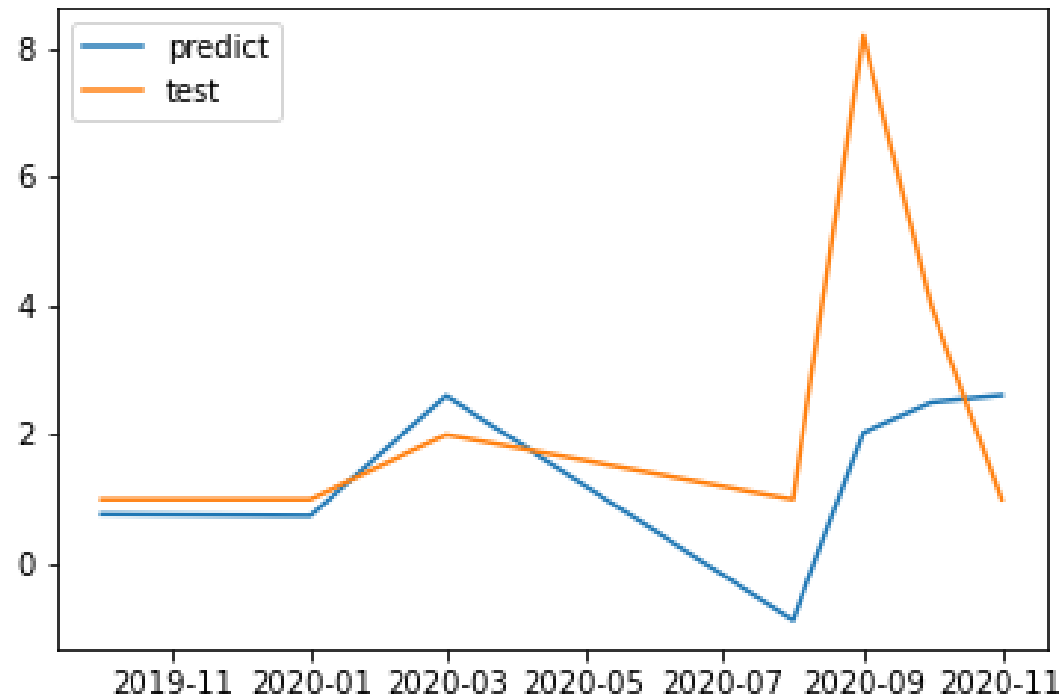
APPENDIX – 정보통신 UPS 수주시기

기술통계

기술통계로 확인한 UPS 수주시기



SARIMA(3,0,0)(3,0,0,12) – 정보통신



출처



카카오맵 API : <https://apis.map.kakao.com>

조달정보개방포털 – 무정전전원장치 조달 내역 : <https://data.g2b.go.kr:1443/pt/main/index.do>

ISTANS – 산업분류연계표 : <https://www.istans.or.kr/wh/whCodeT.do?chn=2>

주요 제조업 및 서비스업 생산지수 – 산업동향지수\생산지수(원지수) : <https://www.istans.or.kr/su/newSuTab.do?scode=S350>

반복 구매제품의 재구매시기 예측을 위한 다층퍼셉트론(MLP) 모형과 순환신경망(RNN) 모형의 성능비교, 송희석, 한국데이터베이스학회, 2017.03, 111-128page

[단독] 여의도 아파트 15층 규제 없앤다... 성냥갑 대신 다채로운 초고층 아파트, 매일경제, 2021.08.18,

<https://www.mk.co.kr/news/realstate/view/2021/08/800844/>