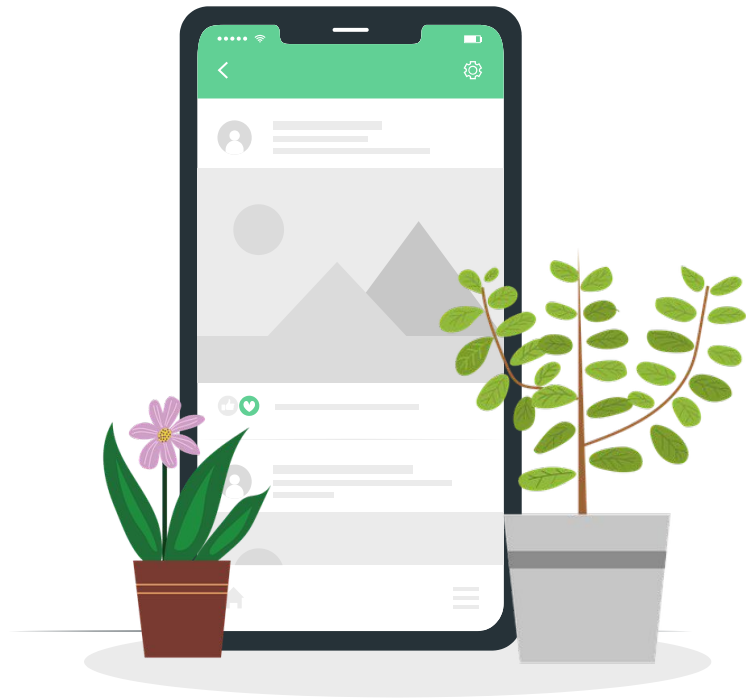


# 식물 가격 예측 서비스

식이랑 서비스기획팀 임지현



# Table of Contents

**1**

## Introduction

서비스 기획 배경과  
프로젝트의 목적

**2**

## EDA & 모델학습

가격 예측을 위한  
머신러닝 모델 생성 및 학습

**3**

## 모델 평가

모델의 평가와 해석

**4**

## 결론

예측 모델에 대한 결과와  
비즈니스 인사이트

1

# Introduction

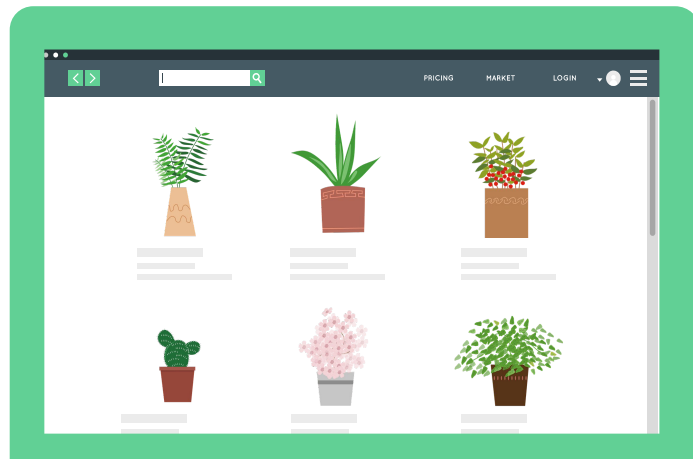
서비스 기획 배경과  
프로젝트의 목적



# 서비스 기획 배경

## 식물 플랫폼 식이량의 서비스

- 식물 관련 제품 판매(화분, 영양제 등)
- 콘텐츠 제공(식물 관련 정보 등)
- 사용자 위치 기반 근처 식물 병원 찾기



# 서비스 기획 배경

## 희귀종 '식테크'... 이 작은 게 500만원입니다

600명 모인 희귀식물마켓 가보니

이정구 기자

입력 2022.07.12 03:00



지난 10일 서울 용산구 아이파크몰에서 열린 '희귀식물마켓'에는 '식물 집사'로 불리는 희귀식물 애호가들이 몰려 정보를 나누고 유튜브 영상도 제작했다. /이정구 기자

## 1200만원짜리 식물이 있다고?... '식테크족' 불러온 희귀식물의 세계

2021.11.26 14:27 입력

장희정 기자

최근 한 온라인 식물오픈마켓에 1200만원짜리 식물이 올라왔다. 실수로 0이 더 붙은 것이 아니다. '최상품종'이란 타이틀이 붙은 이 식물의 이름은 '몬스테라 보르시지아나 알보 바리에가타'. 흔히 '알보몬'이라고 부르는 '희귀식물'이다. 식물업계에서 이른바 희귀식물이 인기다. 큰 잎에 구멍이 나거나 갈라져 이국적인 정취를 풍기는 열대성 관엽식물인 몬스테라의 일반종은 같은 마켓에서 5000원 미만에 판매되고 있다.



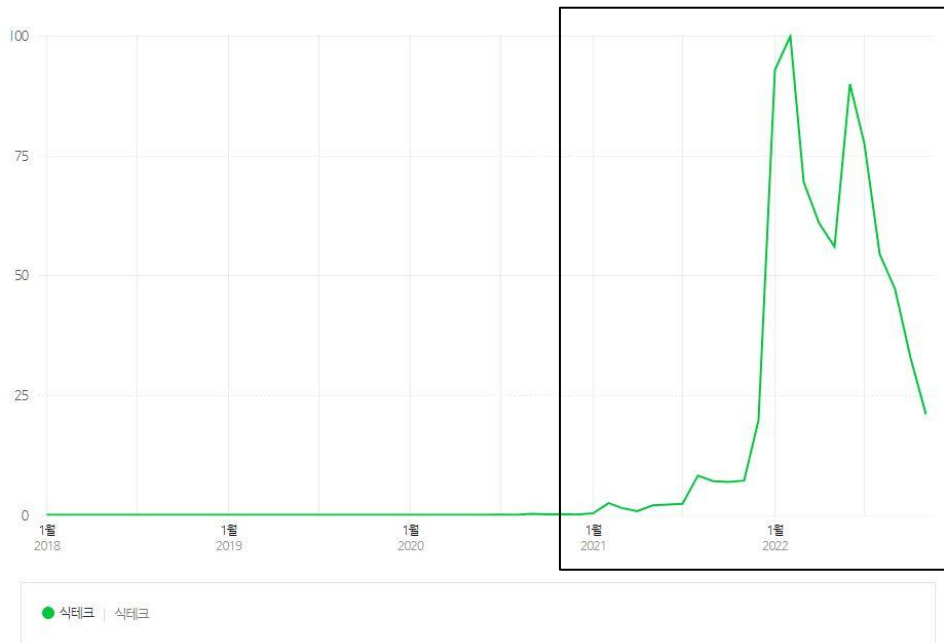
자연이 빚어낸 세상에 하나 뿐인 무늬로 '희소성'이라는 매력을 장착한 희귀식물들. 경기 파주의 조인폴리아에서 희귀식물의 다양함을 자랑하는 '조인폴리아'의 모습이다.

# 서비스 기획 배경

식테크의 성장



판매자와 소비자 사이의  
가격 정보의 불균형



2018년 ~ 2022년 사이의 식테크 검색량 추이

# 프로젝트 목적

**1**

식물 가격 예측을 통한  
정보의 불균형 해소

**2**

가격 예측 서비스를 통한  
구독 비즈니스 모델 구축

# 문제 정의 및 프로젝트 가설

## 구매를 원하는 식물의 평균 가격은 얼마인가?(회귀모델)

- 1 식물 화훼 구분(관엽, 난, 춘란 등)에 따라 가격 차이가 발생할 것이다.
- 2 식물이 비닐하우스에서 경작되더라도 계절에 따라 가격이 달라질 것이다.  
>> 계절에 따라 식물의 가격이 달라지면 날씨가 식물의 가격에 미치는 영향이 클 것이다.
- 3 가격을 결정하는데, 식물의 품종(장미 등)이 가장 중요할 것이다.



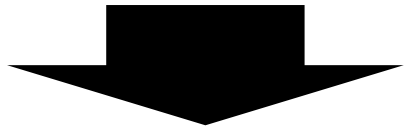
# 가설 정의의 목적

1

식물이 가장 저렴한 시기를 예측해 원하는 식물을 낮은 가격으로 구매할 수 있도록 도와준다.

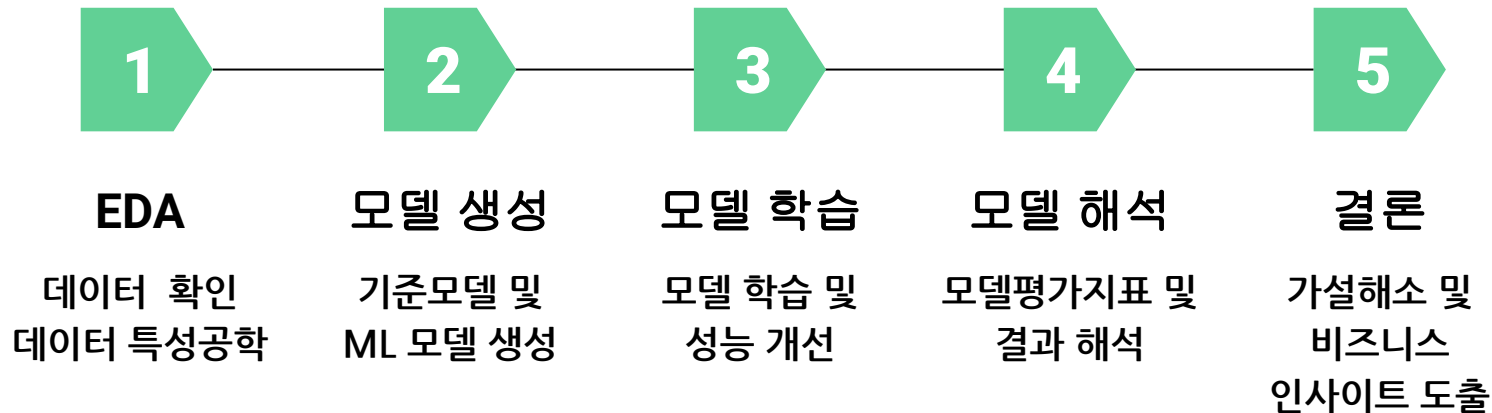
2

가격에 영향을 미치는 요인과 해당 요인의 영향력을 파악하여 식물의 적정 가격을 예측할 수 있도록 도와준다.



**정확한 가격 예측을 통한 서비스 이용 고객을 늘리기**

# 프로젝트 진행 순서



2

# EDA & 모델 학습

가격 예측을 위한  
머신러닝 모델 생성 및 학습

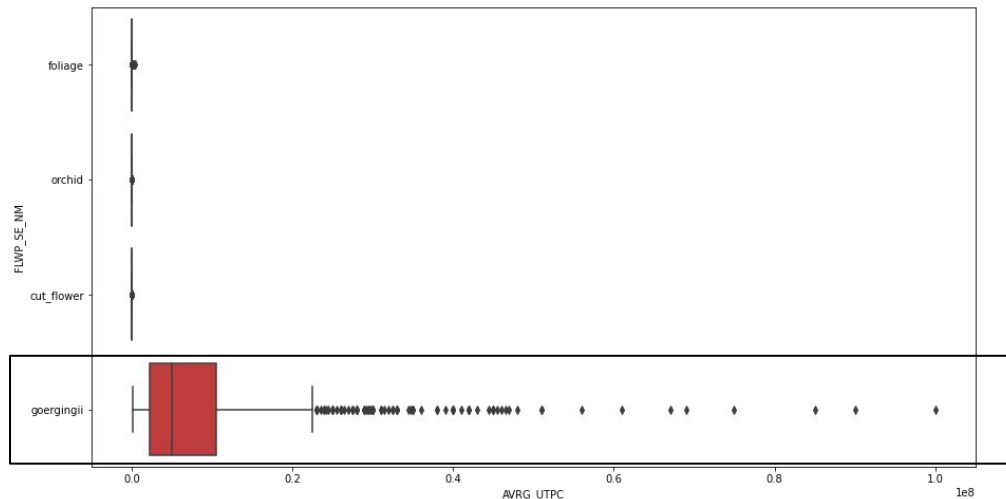


# 데이터 탐색

## 화훼 구분에 따른 가격 분포

- 춘란(goeringii)의 평균가격이 다른 품종에 비해 비싼 것을 확인 가능

➡ 정확한 가격 예측을 위해 춘란을 분리한 다음, 모델 학습을 진행



# 데이터 탐색

## 시간에 따른 평균 가격 추이

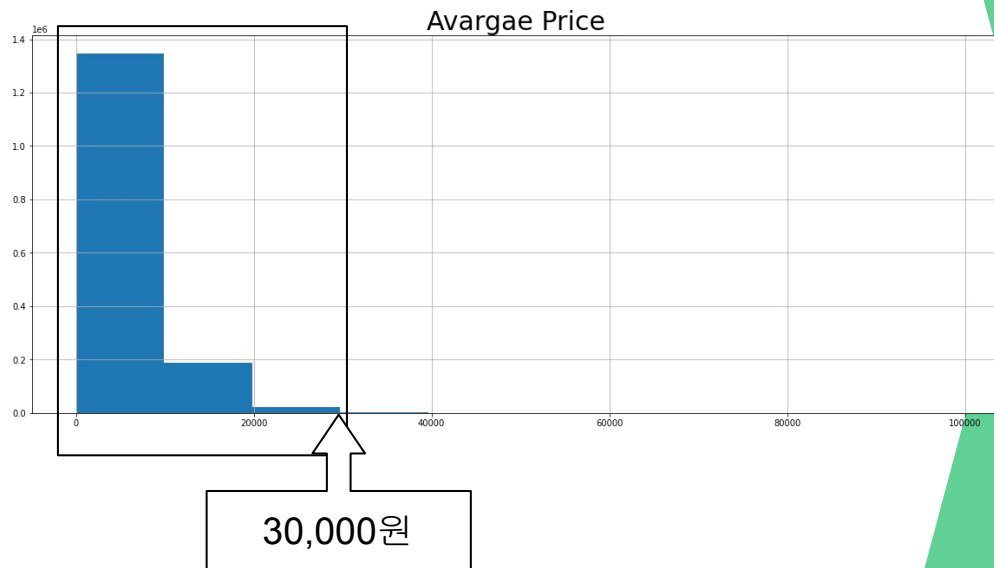
- 겨울에는 가격이 상승하고  
여름에는 가격이 하락



# 데이터 탐색

## 이상치 확인

- 평균 가격 데이터가 3만원 이하에 몰려있는 것을 확인 가능
- 이상치의 영향을 줄일 수 있는 모델 사용



# 데이터 특성 공학



상품 등급(특1, 특2 등)  
데이터 값 변환



식물 품종 등급  
데이터 추가



결측치 처리



데이터 누수 방지



불필요한  
데이터 제거



타겟값  
로그변환

# 머신러닝 모델 생성

## 기준모델

- 검증데이터의 평균

## 머신러닝 모델

- 날짜에 따라 훈련, 검증, 테스트 데이터로 분리
- RandomForest 회귀 모델 사용



# RandomForest 회귀 모델

## 특징

- 여러개의 모델을 생성하여 각 모델의 결과를 종합하여 최종 결과를 예측하는 모델
  - 이상치의 영향을 덜 받음
  - 과적합 방지
  - 카테고리가 많은 데이터에 용이

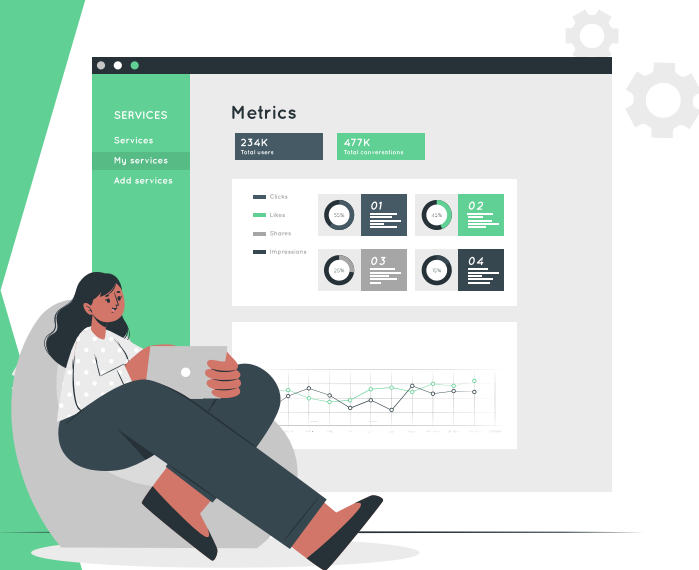
## 모델 성능 향상

- 하이퍼파라미터 튜닝
  - max\_depth
  - max\_features
  - n\_estimators

3

# 모델 평가

모델의 평가와 해석



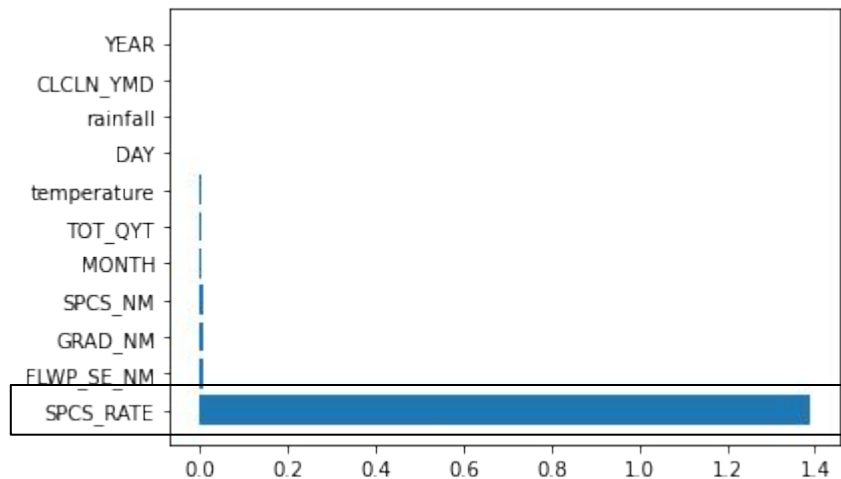
# 머신러닝 모델 평가지표

	기준모델	검증 데이터	테스트 데이터
MAE	0.63	0.26	0.27
MSE	0.64	0.12	0.13
R2 Score	0.00	0.82	0.80

# 모델 해석

## 특성중요도

- 식물 가격 예측 모델에서  
가장 영향력이 큰 특성은  
“식물품종등급(SPCS\_RATE)”

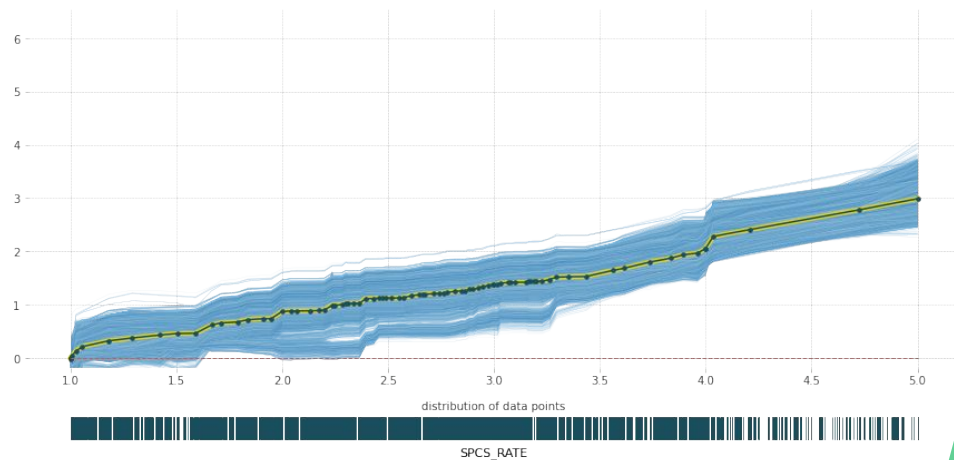


# 모델 해석

## 1 특성 PDP 그래프

- “식물품종등급(SPCS\_RATE)”이 5점에 가까울수록 식물의 평균 가격 증가

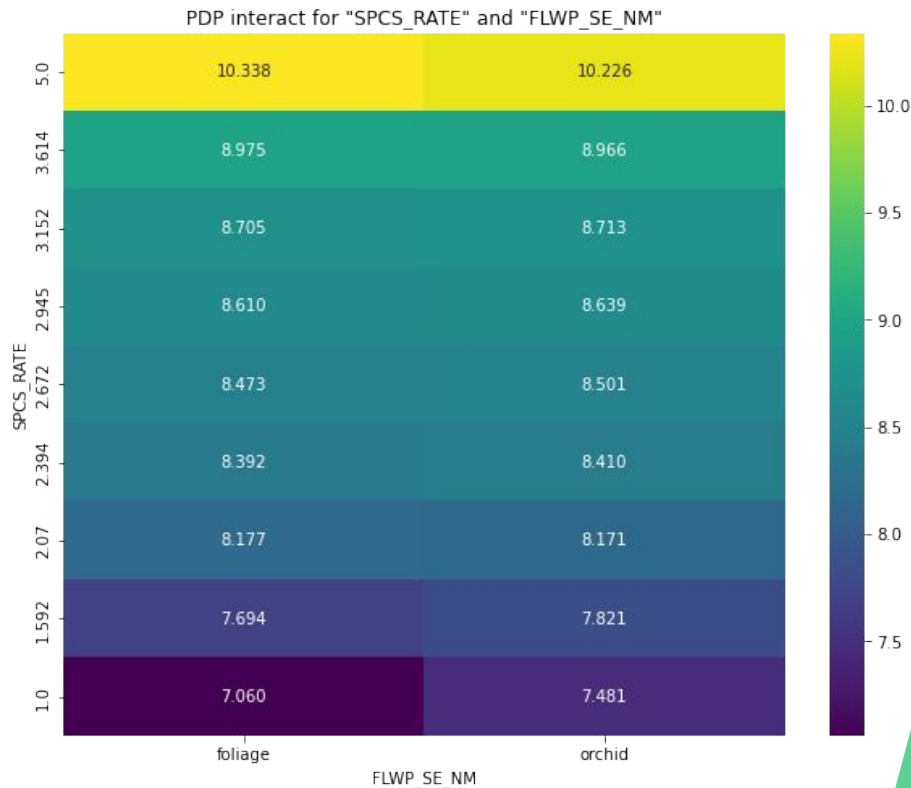
PDP for feature "SPCS\_RATE"  
Number of unique grid points: 83



# 모델 해석

## 2특성 PDP 그래프

- “식물품종등급(SPCS\_RATE)”이 5점에 가까울수록 식물의 평균 가격 증가
- “화훼 구분(FLWP\_SE\_NM)”은 식물의 가격에 큰 영향을 미치지 않음



# 4

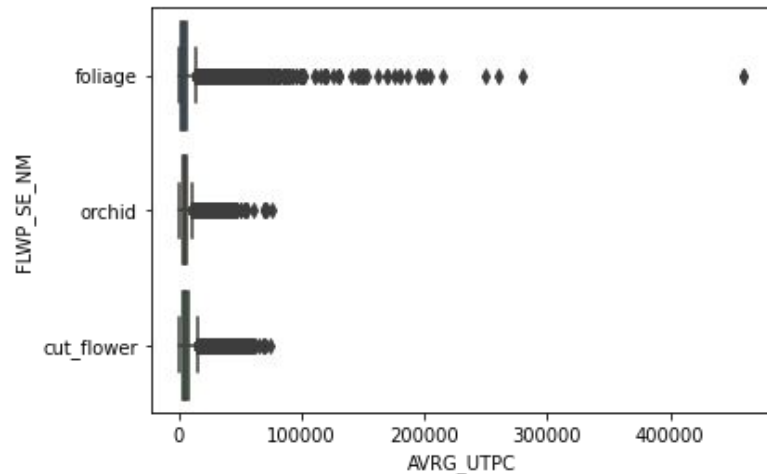
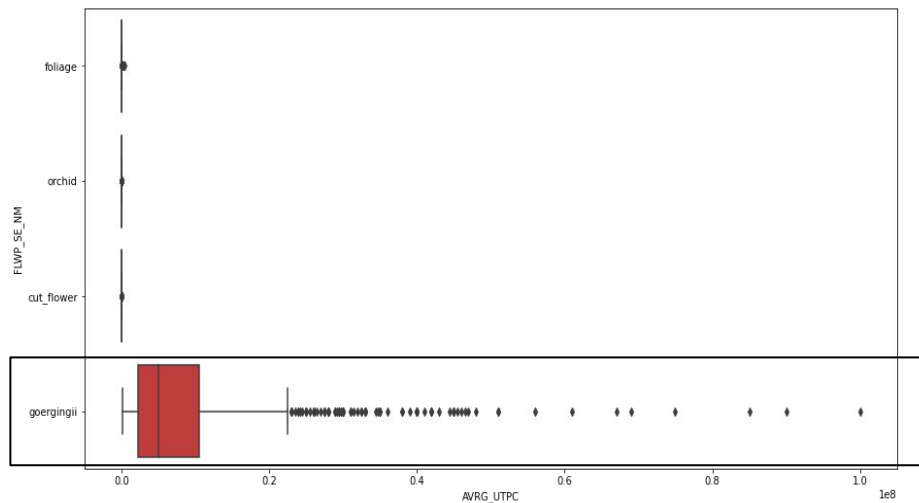
## 결론

예측 모델의 결과와  
비즈니스 인사이트



# 가설 해소

1 식물의 화훼 구분(관엽, 난, 춘란 등)에 따라 가격 차이가 발생할 것이다.



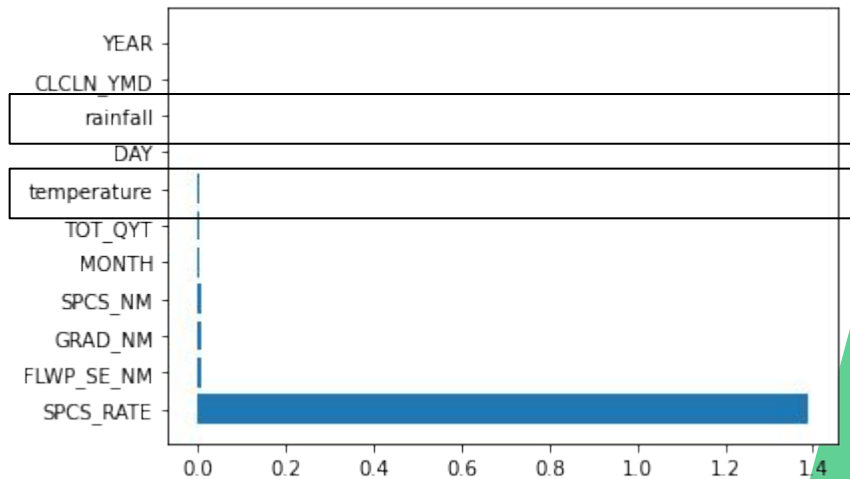
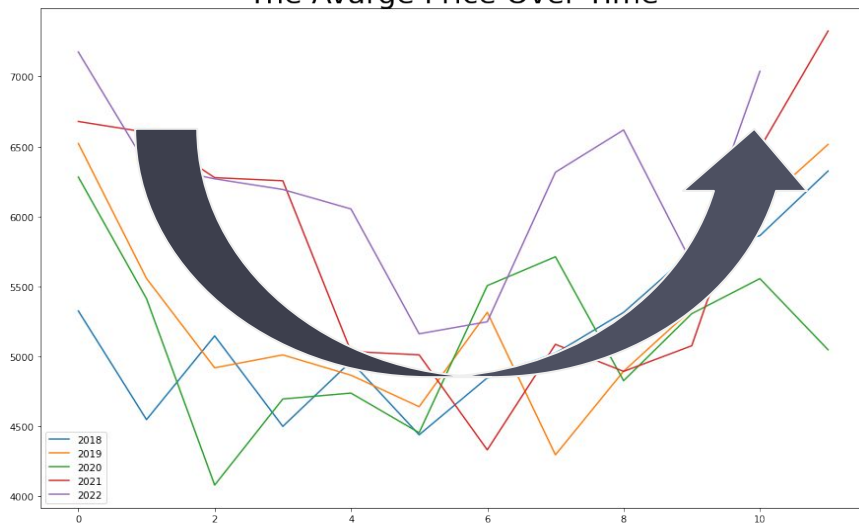


# 가설 해소

2

식물이 비닐하우스에서 경작되더라도 계절에 따라 가격이 달라질 것이다.  
>> 계절에 따라 식물의 가격이 달라지면 날씨가 식물의 가격에 미치는 영향이 클 것이다.

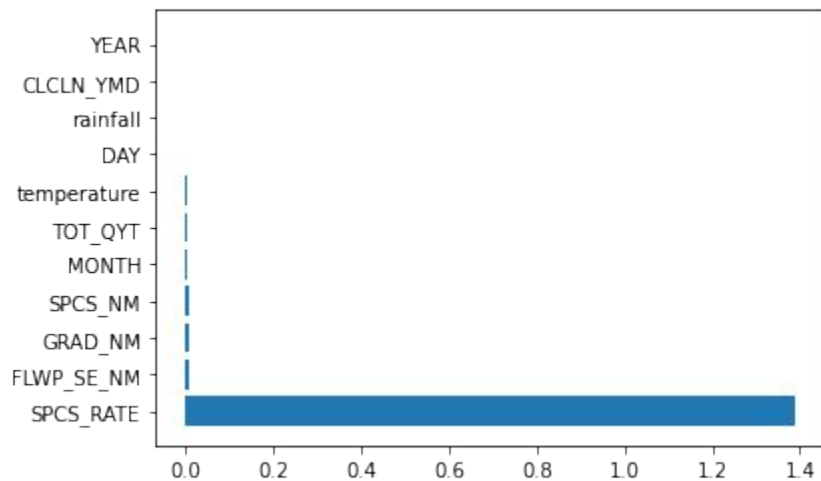
The Average Price Over Time



# 가설 해소

3

가격을 결정하는데, 식물의 품종(장미 등)이 가장 중요할 것이다.



모든 식물 등급을 통해  
식물의 가격을 예측, 제공

# 개선사항

## 식물 가격 예측 서비스



프리미엄 고객을 위한  
추관데이터 추가



식물 등급의  
지속적인 업데이트  
필요



타겟 소비자층 확대를  
위한 절화 데이터 추가

# 비즈니스 인사이트



# Resources

## DATA

- 요일별 화훼경매 시세데이터  
(2018년 ~ 2022년 11월)
- 요일별 온도 및 강수량 데이터  
(2018년 ~ 2022년 11월)
- 식테크 키워드 검색량  
(2018년 ~ 2022년 11월)

## DATA 출처

- 한국농수산물유통공사
- 기상청
- 네이버 트렌드 데이터랩

# Thanks!

This presentation template was created by **Slidesgo**, including icons by **Flaticon**, and infographics & images by **Freepik** and illustrations by **Stories**