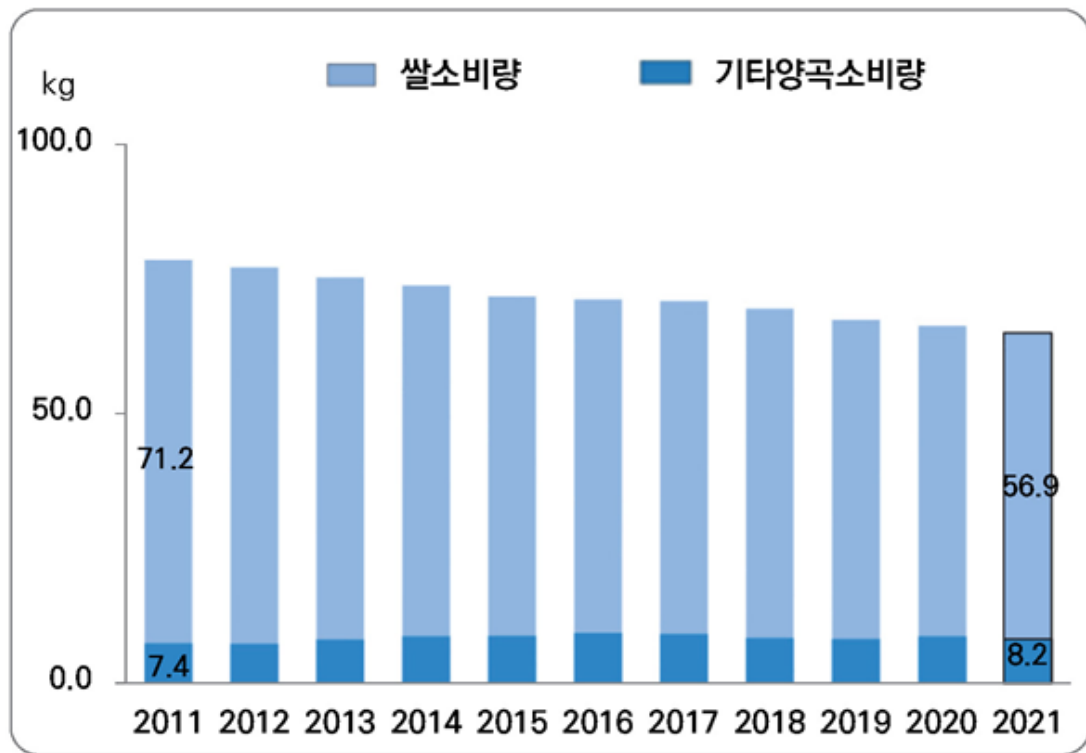


# 쌀 과잉공급으로 인한 정부와 농민의 갈등 해소 - 단수 예측 모형 구축을 통하여 -

6조

신민건 201827526  
권도윤 202027501  
김정현 201827511  
이지우 202027534  
추하연 202027545  
홍정환 201827550

# 제안발표 요약



출처) 2021년 통계청 양곡소비량조사 결과

## 1) 공공비축미 매입

매년 11월 즈음 다음 해 예측소비량의 17-18% 수준 매입

## 2) 쌀 시장격리 정책 시행

수요 대비  
3% 초과 생산

5% 가격 하락

# 제안발표 요약

< 정부 >



대한민국정부

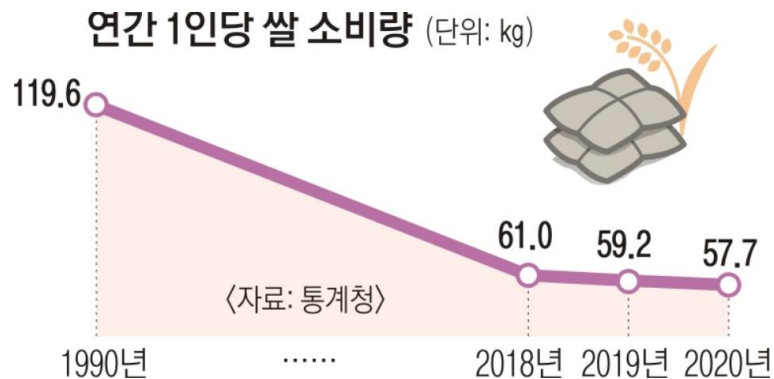
편성된 예산의 한계로,  
가격이 최대한 떨어졌을 때  
과잉공급량을 매입하려 함

< 농민 >



청와대 앞에 모인 쌀 생산 농가들...“쌀 30만톤 즉시 시장  
격리 하라” - 한국농업신문 (newsfarm.co.kr)

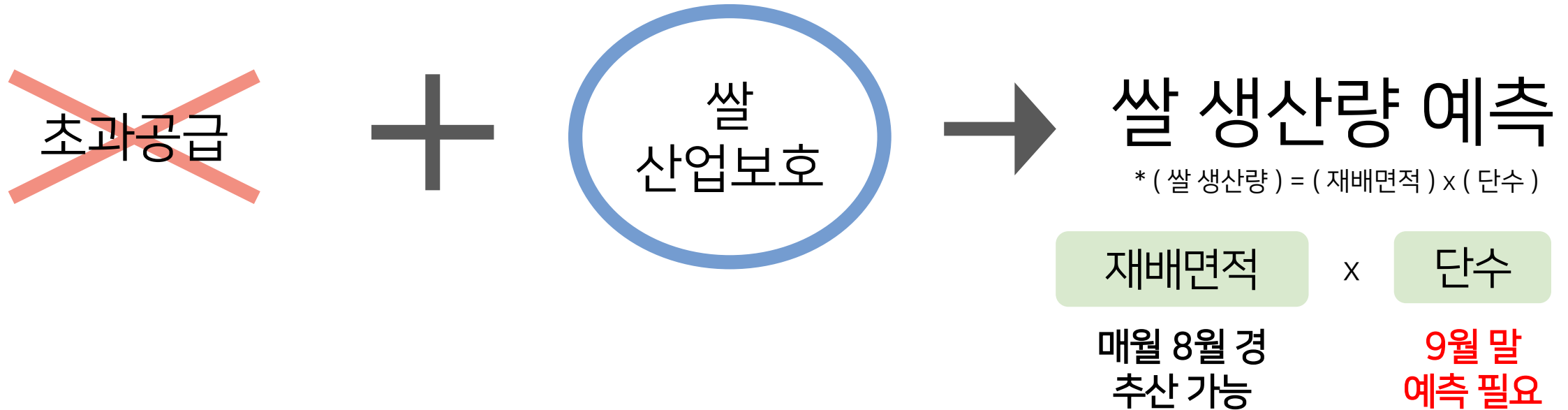
최대한가격이 높을 때  
과잉공급량을 해소하고 싶음



과잉공급량 ↑

# 제안발표 요약

- 문제 해결을 위한 접근



- 텀프로젝트의 진행 방향 및 선행연구와의 차이점
  - 예측 모형 구축 : 기상 데이터 (전통적인 변수) + 농업 관련 데이터 (기술적 요소)
  - 예측 모형의 구체적인 정책 활용 방안 고민

## 분석 결과-전처리(결측치처리)

- 변수 특성을 반영하여 진행

농업 관련 데이터	Linear interpolate	
기상 데이터	1) 강수량	2) 일조시간
	- 비가 오지 않은 날 : Fillna(0) - 비가 온 날 : Interpolate	- 해가 떠 있는 시간 : Interpolate - 그 외의 시간 : Fillna(0)
	3) 적산온도	4) 기온, 일교차
	- 결측치 X	- KNN

# 분석 결과-전처리 (최종 데이터프레임)

year	지역	5월 평균 기온	6월 평균 기온	7월 평균 기온	8월 평균 기온	9월 평균 기온	5월 평균 일교차	6월 평균 일교차	7월 평균 일교차	8월 평균 일교차	9월 평균 일교차	5월 평균 강수량	6월 평균 강수량	7월 평균 강수량	8월 평균 강수량	9월 평균 강수량	누적 일조 시간	5월 적 산온도	6월 적 산온도	7월 적 산온도	8월 적 산온도	9월 적 산온도	농기계 대수	농가인구	농업 용수	단수 (10a당 생산량)
1996	부산	17.9419 4	20.5047 2	23.8469 1	26.2028 2	22.5155 6	7.71290 3	5.00666 7	5.70967 7	6.20322 6	7.21333 3	0.06303 8	0.45430 6	0.38279 6	0.18871 0.03708 3	312.4	5802.3	15679. 3	28187. 95	20826. 9		13138	18896	132,63 8.20	493	

## - 독립변수

1	5월 평균기온
2	6월 평균기온
3	7월 평균기온
4	8월 평균기온
5	9월 평균기온
6	5월 일교차
7	6월 일교차
8	7월 일교차
9	8월 일교차
10	9월 일교차
11	5월 평균강수량
12	6월 평균강수량

13	7월 평균강수량
14	8월 평균강수량
15	9월 평균강수량
16	누적일조시간
17	5월 적산온도
18	6월 적산온도
19	7월 적산온도
20	8월 적산온도
21	농기계대수
22	농가인구
23	농업용수

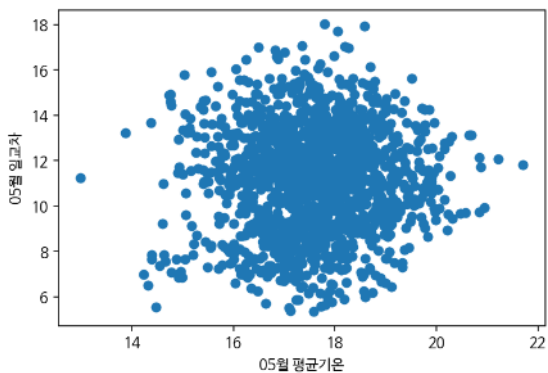
## - 종속변수

1	단수(kg/10a)
---	------------

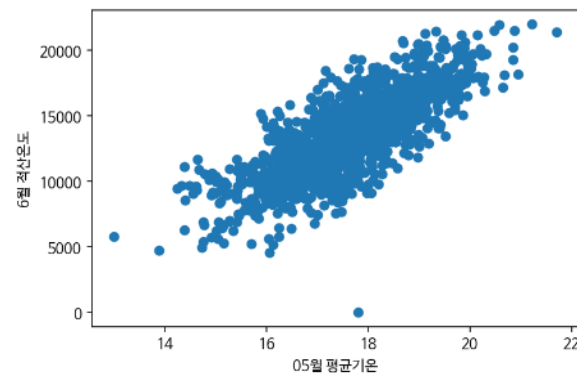
- 농업관련 데이터  
→ 종속변수의 특성을 고려해  
경지면적으로 나누어줌

- 경지면적 컬럼 삭제

상관관계를 보이지 않는 두 변수



강한 상관관계를 보이는 두 변수



→ 다중선형회귀모형 사용 보류

비선형회귀

트리기반 회귀모형

Decision Tree

Random forest

Xgboost

LightGBM

모델	특성 (이점 및 단점)
Decision Tree	• 트리기반 회귀 모형의 가장 기본 모형
Random Forest	• 여러 Decision Tree를 bagging 방식으로 앙상블한 모델
Xgboost	• 여러 Decision Tree를 boosting 방식으로 앙상블한 모델 (균형 트리)
LightGBM	• 여러 Decision Tree를 boosting 방식으로 앙상블한 모델 (불균형 트리)

# 분석 결과-모델 선정

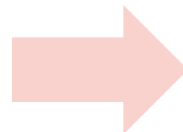
Min-max  
normalization



GridSearchCV를 사용해 모델별 최적 파라미터 선정

- 지정된 파라미터의 경우의 수마다 예측 성능을 평가하며 최적의 하이퍼 파라미터를 찾는 것

model	교차검증 RMSE
Decision Tree	37.82
Random Forest	37.33
Xgboost	40.43
LightGBM	37.04



가장 성능이 좋은  
LightGBM으로 모델 결정

<LightGBM Parameter>

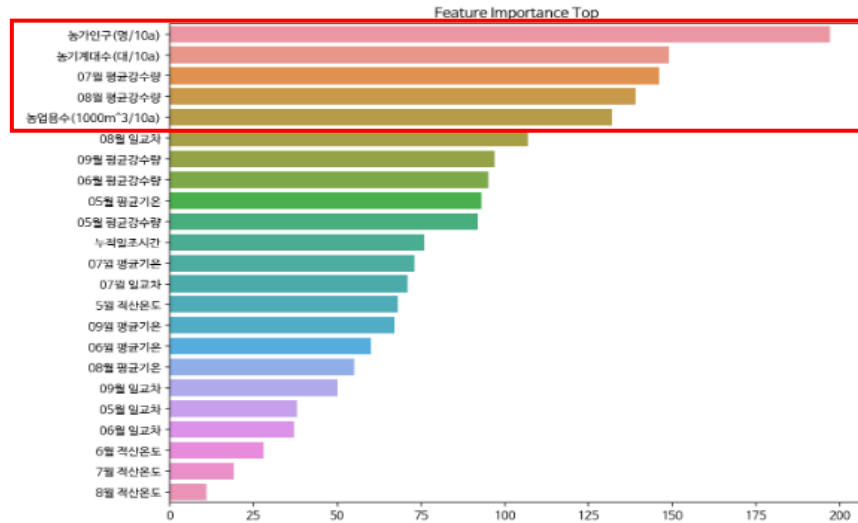
learning\_rate=0.2,  
max\_depth=7,  
min\_child\_samples=15,  
num\_leaves=20

Train data	Test data	모델 성능
1996년 ~ 2020년 data	2021년 data	MAE: 28.804 MAPE: 0.054 MSE: 1186.403 RMSE: 34.444



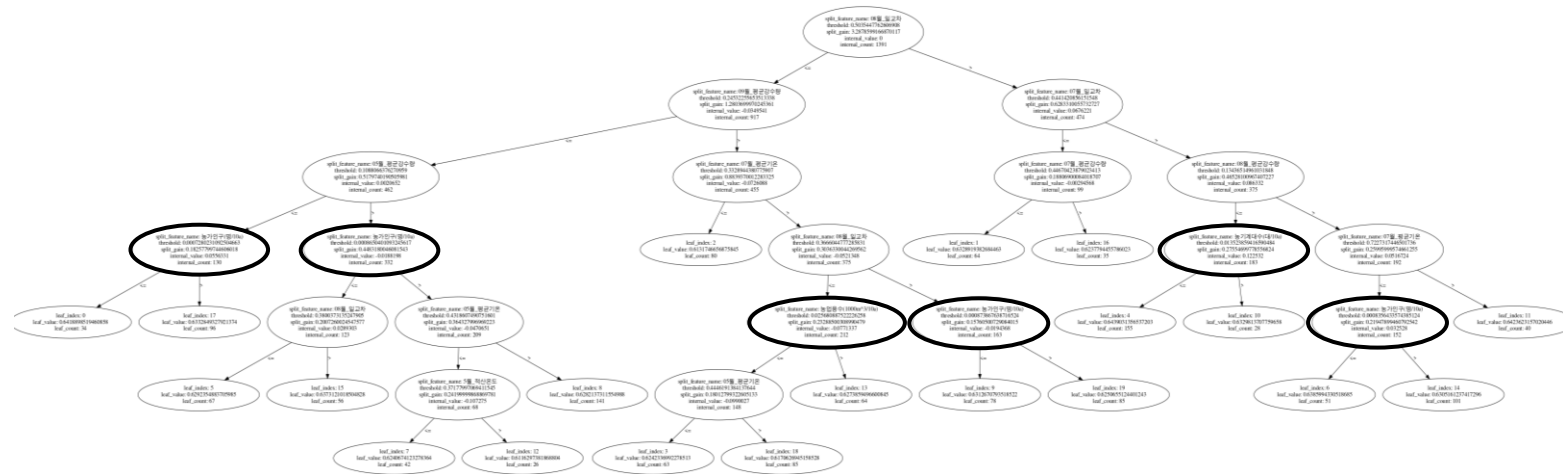
# 분석 결과-모델분석

<LightGBM feature importance>

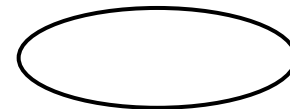


농업관련 변수의 중요도가 높음을  
확인할 수 있음

<LightGBM 내의 tree 100개 中 첫 번째 tree>



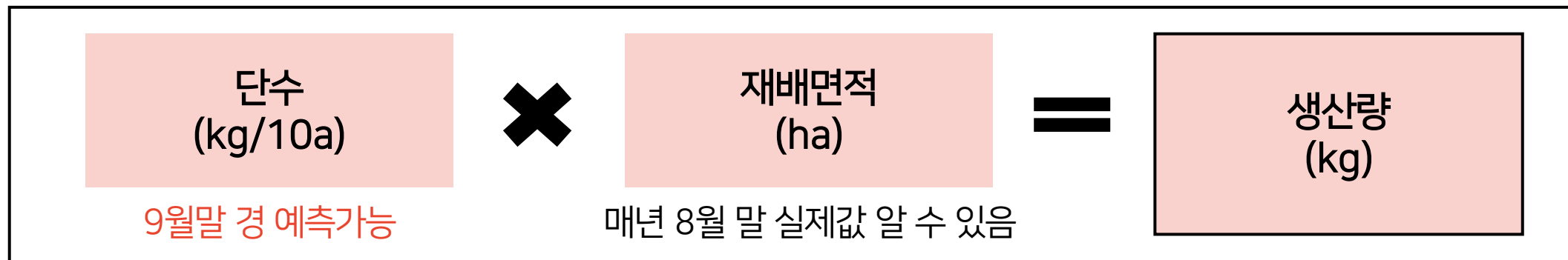
...



는 농업 관련 변수

# 문제 해결 방안 - 개요

## 1. 올해 생산량 예측 : lightGBM model 사용



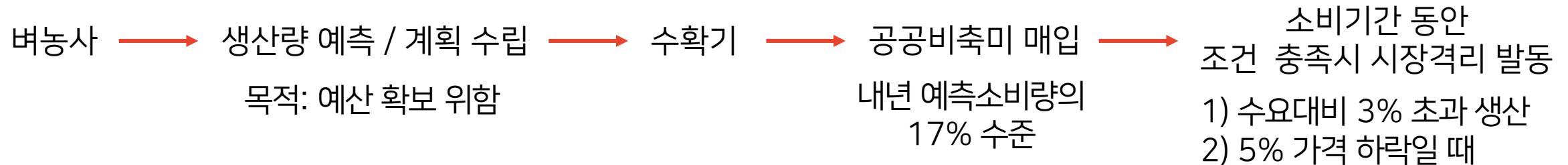
## 2. 내년 소비량 예측 : 시계열 예측모형 사용 → 지역별 권장 생산량 계산 (분배기준: 지역별 경지면적 비율)

## 3. 과잉공급량 예측 : 지역별 예측 생산량 - 지역별 권장 생산량

## 4. 문제 해결 방안

- ① 과잉공급량(예측생산량-권장생산량) 수준에 따른 지역별 차등 incentive 지급을 통한 과잉공급량 감축 촉구
- ② 타 작물 경작 장려 제도 시행 : 기존 과잉공급량 매입 위해 사용한 예산 이용

# 문제 해결 방안 - 현재 정부의 쌀 관련 정책 프로세스



## 문제 해결 방안

예시 : 2021년 말 부산 지역의 쌀 과잉공급정도 평가

$$[ \text{과잉공급량} = \text{예상생산량} - \text{권장생산량} ]$$

**권장생산량** 2021년 11월 ~ 2022년 10월 전국소비량 \* 해당 지역 경지면적비율 \* 1.17 (공공비축미 산정 비율 기준)

$$= \sum_{2021\text{년 } 11\text{월} \sim 2022\text{년 } 10\text{월}} (\text{월별 1인당 1일 평균 소비량}) * (\text{해당 월의 일수}) * (\text{해당 월의 인구})$$

- 월별 1인당 1일 평균 소비량(g): 자기상관 확인 → 시계열모형 필요
- Prophet 모형 이용: 페이스북이 만든 시계열 예측 라이브러리



시점	월별 1인 예측소비량(g)	월별 인구수(명)	월별 쌀 예측소비량(kg)
2021-11-01	156.722433	51662290	242,899,200
2021-12-01	154.366296	51662290	247,222,400
2022-01-01	156.292498	51662290	250,307,300
...			
2022-08-01	146.399791	51662290	234,463,800
2022-09-01	151.093298	51662290	234,174,800
2022-10-01	149.180646	51662290	238,917,400

시계열 예측

2021년 10월 인구  
유지된다고 가정

X 1.17

3,373,697,886.7

전국 권장생산량

## 문제 해결 방안

예시 : 2021년 말 부산 지역의 쌀 과잉공급정도 평가

$$[ \text{과잉공급량} = \text{예상생산량} - \text{권장생산량} ]$$

예상생산량

- 2021년 예측 단수 \* 2021년 재배면적 (8월 산정)

지점명	예측단수(kg/10a)	재배면적(10a)	예상생산량(kg)
부산	475.1699718	21,690	10,306,436.69

$$[ \text{과잉공급량} = \text{예상생산량} - \text{권장생산량} ]$$

$$\begin{aligned}
 &= 10,306,436.69 - 3,373,697,886.7 * (\text{부산 경지면적 비율}) \\
 &= 10,306,436.69 - 10,006,155.27 \\
 &= 300,281.41(\text{kg})
 \end{aligned}$$

- 예상생산량 中 0.29% 정도 초과생산한 양일 것으로 예측 가능
- 지역간 과잉공급비율 비교 후 incentive 차등 지급

## 전체적인 흐름

\* 특정 지역의 예상 과잉공급량 = 지역 예상 생산량 - 지역 권장 생산량

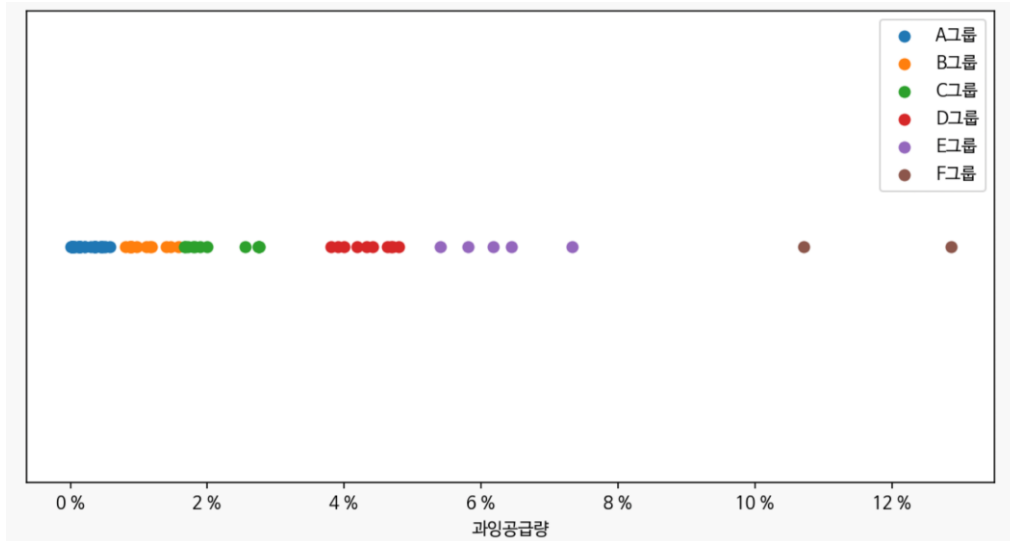
< 시장격리기간 예산 편성 >

- ① 부산 지역의 생산량 예측
- ② 과잉공급량 예측



## ① 정부의 쌀 매입 예산 감소

- 지역간 incentive 차등 지급을 시행할 시, 농민의 자발적인 생산량 조절 기대
- 정부가 쌀 매입하는데 쓰는 예산 감소



## ② 타 곡물자급률 향상에 기여

- 쌀매입 예산을 다른 곡물 지원에 분배하여 농민 지원  
보조금 없이 벼 재배면적 줄인다고?...농가 '코웃음'

입력 : 2022-04-08 00:00 | 수정 : 2022-04-10 00:36



보조금 없이 벼 재배면적 줄인다고?...농가 '코웃음' - 농민신문 (nongmin.com)

- 타 곡물자급률 향상을 통한 전체 곡물자급률 향상