

# 기업연계 3조

...

POSITION

박재형, 최기욱, 송찬범, 권혁돈, 이동찬, 이제형

# 목차

1. FOSITION 소개
2. 프로젝트 목표
3. Why? Coffee?
4. 프리미엄 커피농장
5. 비즈니스 모델
6. 스마트 팜 구조
7. AI 모델 설명
8. FOSITION 서비스
9. 스마트 팜 시연
10. 질문

저희는 초반 비즈니스 모델에서 많은 설명이 있을 예정입니다. 그리고 저희 농장의 구조와 기술에 대해서 설명 드리겠습니다.

# FOSITION



저희 FOSITION 농장의 브랜드 의미는 어느지역에서나 최적화된 시스템을 통하여 사람들에게 똑같은 품질의 서비스를 제공할수있다는 점에서 위치라는 명사인 position을 본따 만들었습니다. 그리고 모든곳에 작물을 만들수 있다는 것을 어필하기 위해 Anywhere Everywhere라는 모토를 사용하였습니다.

# FOSITION의 목표



- 차별화된 농장 시스템 구축
  - 어느 지역에서도 체계화된 스마트 팜 시스템으로 좋은 농작물을 선사
- 농작물의 자동 생산
  - 21세기에 맞추어 농민들에게 편리한 인터페이스 제공
- 커피 농장의 브랜드화
  - 한우와 같은 국산 파워브랜드로 성장 기대
- 농장의 시스템화로 양질의 데이터 제공
  - 농장의 환경데이터와 제어데이터를 수집하여 분석가들에게 편리한 데이터 제공

# Why Coffee?



# 한국인은 커피를 하루에 얼마나 마실까?



1인당 커피 하루 평균 2잔



한국인은 일주일에 얼마나 먹을까?

1 커피 12.2회



2 배추김치 11.9회



3 설탕 9.7회



4 잡곡밥 9.6회



커피의 국내 생산량  
63% 증가



커피 국내 생산액  
92% 성장



커피전문점 확대로  
커피믹스 시장 감소



현재 한국은 잡곡밥, 설탕, 배추김치보다 더 커피를 자주 마시고 있습니다.

# 국내 커피시장 규모

단위: 원, 2023년은 전망치

10조

5조

0

5조9000억

6조8000억

8조6000억

2016년

2018년

2023년

자료: 현대경제연구원

The JoongAng

커피 시장은 매서운 속도로 성장하고 있으며

2년을 주기로 약 1조원의 성장을 하고 있습니다.

23년 예측치를 보시면 총 8조원의 성장을 기대하고 있습니다.

때문에 저희는 커피농장을 저희의 작물로 선정하였습니다.

# 프리미엄 커피 농장



저희 포지션은 현재 국내에서는 존재하지 않는 프리미엄 커피 농장의 국내 생산을 목표로 프리미엄 커피 농가 시스템을 추구하고 있습니다.



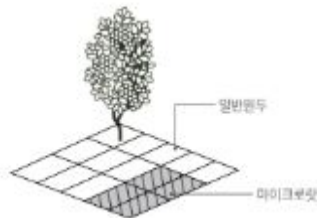
# FOSITION 커피농장 비즈니스 모델

마이크로랏 모델 사용하여 고급화 전략 사용

- 마이크로 랫이란 대규모 농장이 아닌 특정한 구역을 구분하여 특별관리된 소량의 커피를 생산하는것을 의미
- 저희는 소량의 특별한 커피를 생산하기에, 상품의 균일한 품질과 질에 많은 노력을 기울일수 있습니다.
- 일반적으로 마이크로랏의 커피의 경우 농장의 모든것이 투명하게 관리 되기에 사람들이 안심하고 질 좋고 건강에 좋은 커피를 제공할수 있는점이 장점입니다.
- 여기에 AI 기술까지 접목하여 스페셜티 등급까지 받는것을 목표로 두가지의 프리미엄을 받는것이 저희의 최종 목표입니다.

## 마이크로랏 원두

작은 규모의 농장에서 특별 관리하여 수확된 "마이크로 랫 커피"는 보통의 수확량보다 아주 적은 양을 생산하여 균일한 품질과 생산 지역의 특색이 잘 나타나는 특징이 있다.

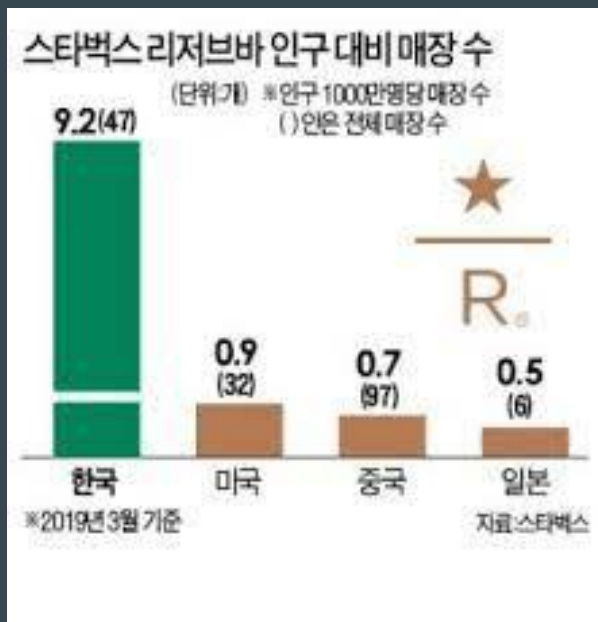


## 스페셜티 등급 원두 사용

미국 스페셜티 커피협회(SCAA)에서 원두를 평가하여 80점 이상의 결정도가 거의 없고 커피가 가지고 있는 산지 특성과 풍미를 지닌 원두를 스페셜티라고 한다.



# 프리미엄 커피의 성장성

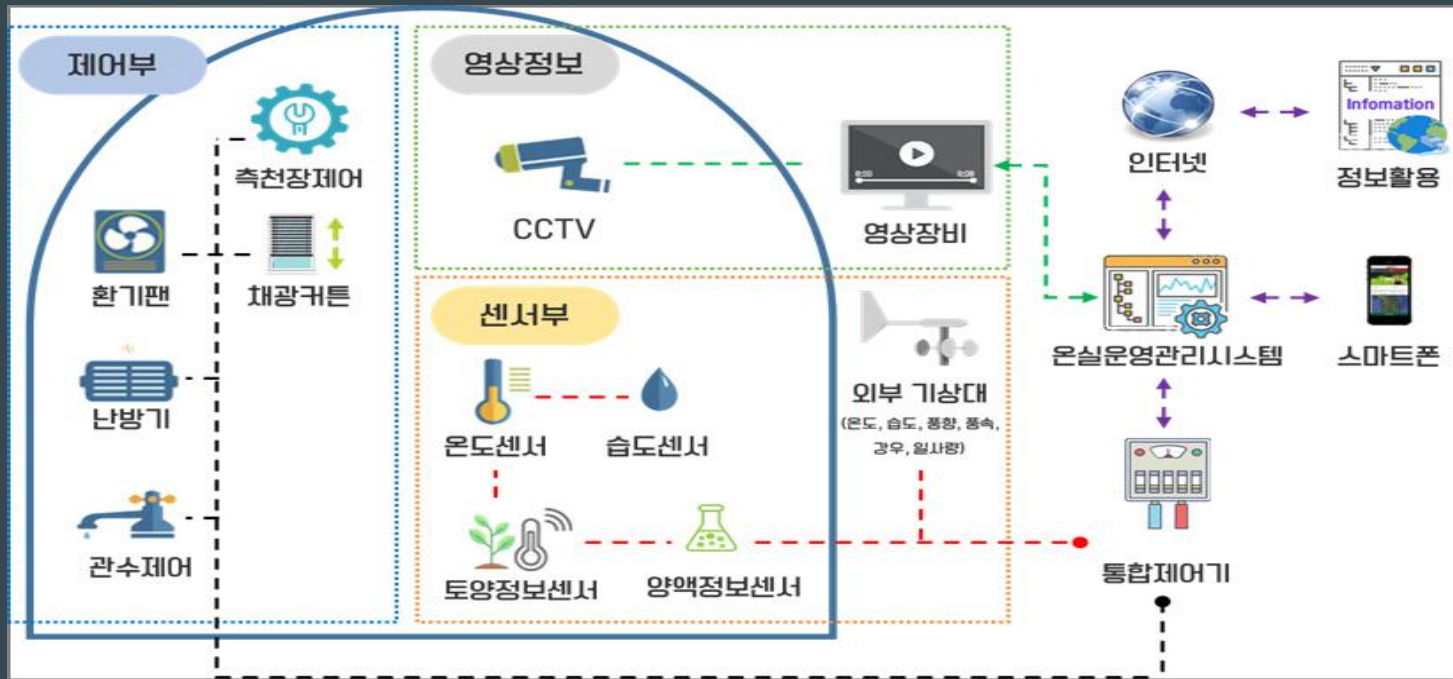


한국에서의 프리미엄 커피의 성장 가능성은 열려 있습니다. 고급 커피 브랜드인 스타벅스 리저브의 판매량과 매장수를 확인해보면 매년 프리미엄의 가치는 올라가고 있는 것으로 보여집니다.

# 스마트 팜 프로젝트

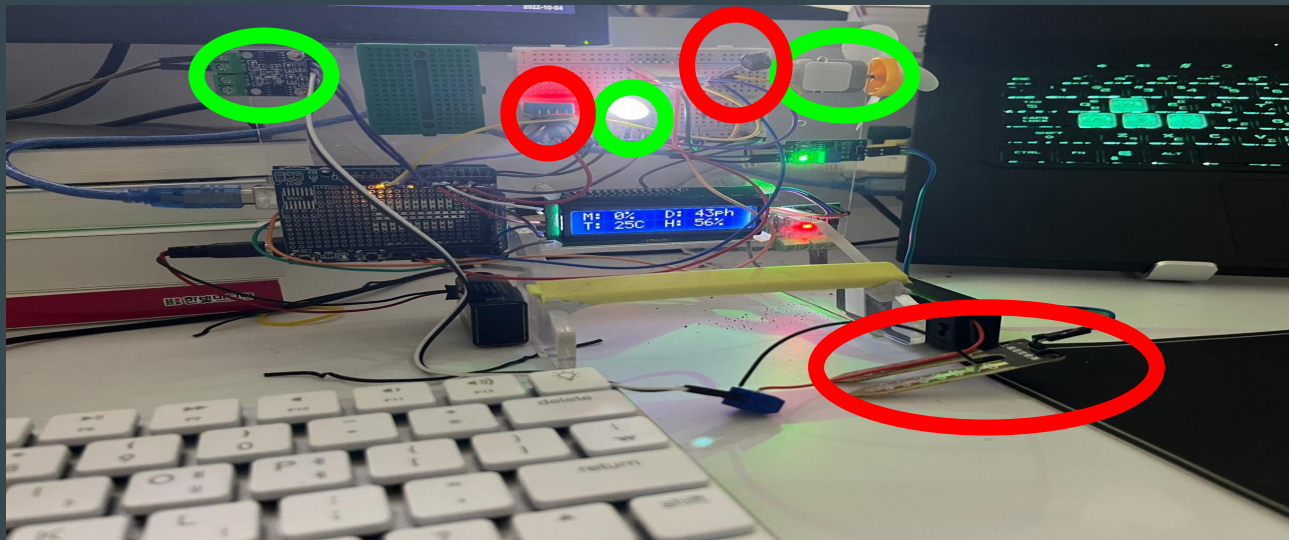


# FOSITION 농장 기본 구조



제어부 센서부가 있습니다. 이곳에서 데이터를 아두이노를 통하여 데이터를 습득하고 습득한 데이터를 DB에 보내고 WEB서버에서도 확인이 가능한 구조를 만들었습니다. 저희는 구조적으로 어떻게 저희가 작업을 했는지 설명하겠습니다.

# 데이터 수집(아두이노)



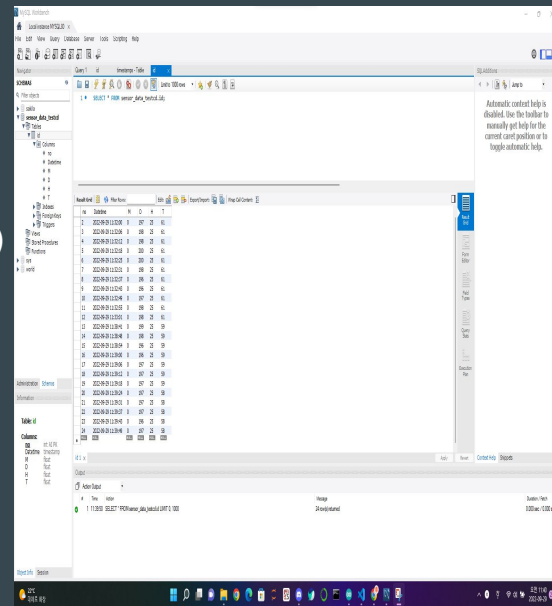
저희는 아두이노를 통하여 3개 수집센서와 3개 제어장치를 가지고 작물의 데이터를 추출하는것에 성공하였으며 이러한 데이터를 pysrial통신을 통해 데이터 베이스로 옮기는 작업까지 성공하였습니다.

# 데이터베이스 - MySQL



저희는 원활한 데이터 네트워크 통신을 위해서 SQLite대신 MySQL를 사용

1. 아두이노에 습도/조도/온도/센서/지습센서에서 나오는 값들을 시리얼통신을 이용하여 MySQL에 적재
  - a. 아두이노에서 출력되는 아스키코드를 유니코드로 전환
  - b. 불필요한 특수문자 및 데이터 정제
  - c. 시간단위로 데이터를 Mysql workbench에 입력
2. 타 농가의 환경 데이터를 SQL에 적재 (12개 농가데이터)
  - a. 20%의 높은 생산성과 수확률을 가진 우수 농가데이터 선별



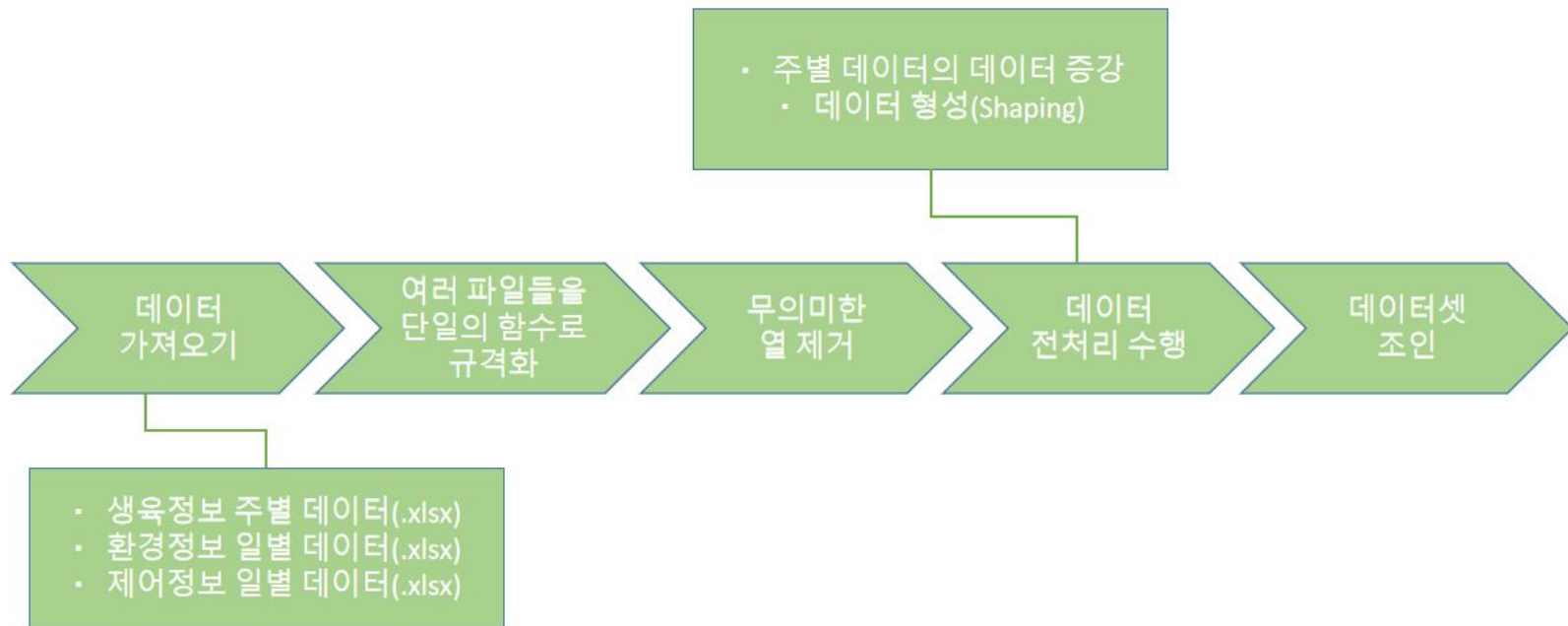


# AI 데이터



저희가 원래는 커피의 생장 데이터를 사용할 예정이었으나 현재 온라인상으로는 커피의 생장데이터 및 제어데이터를 찾는것에 어려움을 가지게 되었습니다. 그래서 저희는 일단은 온라인 방울 토마토 데이터를 사용하여 토마토 농가들의 환경 데이터, 제어 데이터, 생육 데이터 수집하여 MVP(최소검증제품) 검증을 통해서 모델의 가능성을 먼저 확인하도록 하였습니다.

# AI 데이터 과정





# AI 데이터 가공

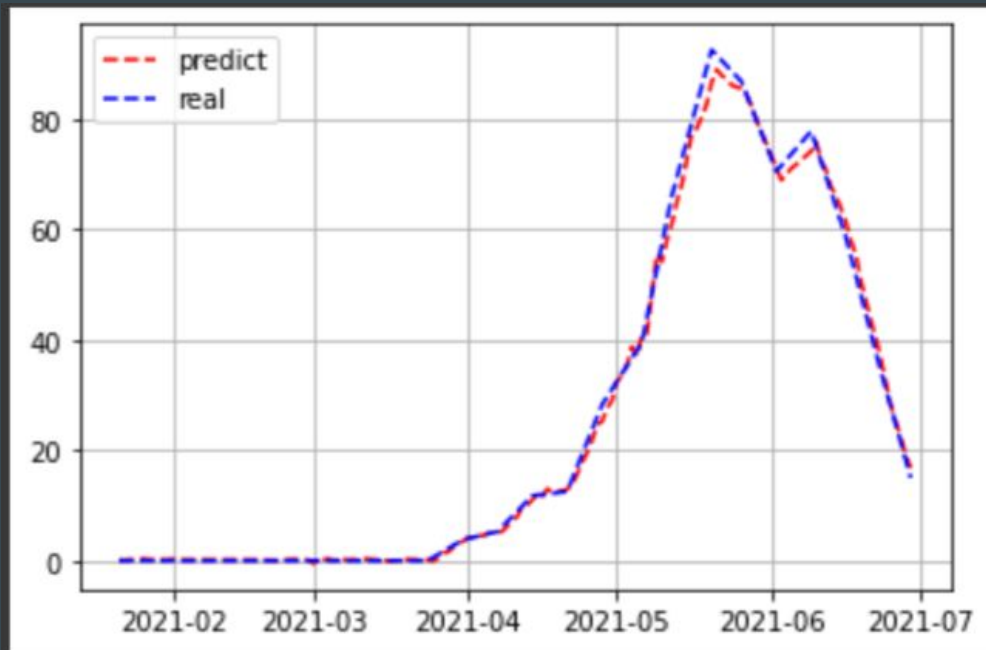
수집일	내부 CO2(ppm)	이슬점 온도(도)	(양액)공급 EY(ds/h)	토양 EY(ds/h)	(양액)배액 EY(ds/h)	외부습도(도)	내부습도(도)	지습(%)	(양액)공급배액(%)	...	(양액)배액PH	강우강지	(양액)총공급(t)	(양액)총배액(t)	일사량 (W/m <sup>2</sup> ·s)	외부온도(도)	내부온도(도)	지온(도)	외부풍량(도)	외부풍속(m/s)
1 2021-01-12	635.18	NaV	NaV	NaV	NaV	NaV	81.0	NaV	NaV	..	NaV	0.34	NaV	NaV	110.15	-3.13	14.74	NaV	NaV	0.2
2 2021-01-13	552.26	NaV	NaV	NaV	NaV	NaV	84.38	NaV	NaV	..	NaV	0.0	NaV	NaV	63.94	1.82	17.46	NaV	NaV	0.05
3 2021-01-14	532.32	NaV	NaV	NaV	NaV	NaV	82.04	NaV	NaV	..	NaV	0.0	NaV	NaV	173.33	1.74	17.91	NaV	NaV	0.35
4 2021-01-15	455.24	NaV	NaV	NaV	NaV	NaV	87.31	NaV	NaV	..	NaV	0.0	NaV	NaV	82.31	8.69	17.3	NaV	NaV	1.84
5 2021-01-16	530.14	NaV	NaV	NaV	NaV	NaV	83.2	NaV	NaV	..	NaV	0.0	NaV	NaV	129.99	-0.93	17.63	NaV	NaV	0.93



수집일	내부CO2(ppm)	내부습도(%)	강우강지	일사량(W/m <sup>2</sup> ·s)	외부온도(도)	내부온도(도)	외부풍속(m/s)
1 2021-01-12	635.18	81.0	0.34	110.15	-3.13	14.74	0.2
2 2021-01-13	552.26	84.38	0.0	63.94	1.82	17.46	0.05
3 2021-01-14	532.32	82.04	0.0	173.33	1.74	17.91	0.35
4 2021-01-15	455.24	87.31	0.0	82.31	8.69	17.3	1.84
5 2021-01-16	530.14	83.2	0.0	129.99	-0.93	17.63	0.93

- 데이터를 포괄적으로 정제할 수 있는 함수를 만들고 데이터 정제
- 데이터 누락값이 많은 컬럼 제거, 데이터 정규화 등과 같은 데이터 전처리
- 주 단위로 작성된 데이터셋과 일 단위로 작성된 데이터셋을 결합하기 위해, 주 단위 데이터셋 데이터 증강

# AI 모델링



- LSTM모델을 이용하여 방울토마토 생산량 예측, 모델 성능 개선
  - 시계열 데이터 처리의 가장 기본인 LSTM을 사용했지만 데이터양이 너무 적어서 과적합.
  - 과적합이 나타난 큰 이유중 하나는, 방울토마토의 작물 시즌에 따라 생산량 패턴이 다르게 나왔기 때문.
  - 그래서 데이터를 시즌별로 나누어 주었고, 훈련데이터와 검증데이터는 같은 시즌의 데이터를 사용하여
  - 최근주 생산량 데이터 또한 데이터 셋에 포함시켰더니 정확도가 많이 올라갔다.

# FOSITION의 서비스(점수 예측기 모델)

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z				
1	Species	Owner	Country of Fam	ham	Loc	humb	MI	ICO	Num	Company	Altitude	Region	Producer	Number of Bag	Weight	In Country	Harvest	1	Grading	D	Onset	1	Variety	Processing	Acroma	Flavor	Alterate	Acidity	Body	Balance
2	0	Arabica	metad p/c	ethiopia	metad p/c	metad p/c	2014/2015	metad ag	1950-2200	guy	ham	METAD PL	300 60 kg	METAD AL	2014 April 4th	metad p/c	Other	Washed	/	0.67	0.83	0.67	0.75	0.85	0.85	0.85	0.85	0.85	0.85	0.85
3	1	Arabica	metad p/c	ethiopia	metad p/c	metad p/c	2014/2015	metad ag	1950-2200	guy	ham	METAD PL	300 60 kg	METAD AL	2014 April 4th	metad p/c	Other	Washed	/	0.75	0.87	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85
4	2	Arabica	grounds	Colombia	marco baron	san cristobal	cuch	1800 m																						
5	3	Arabica	yidrekech	ethiopia	yidrekech	dabaw	wideru																							
6	4	Arabica	metad p/c	ethiopia	metad p/c	metad p/c	2014/2015	metad ag	1950-2200	guy	ham	METAD PL	300 60 kg	METAD AL	2014 April 4th	metad p/c	Other	Washed	/	0.17	0.58	0.82	0.82	0.85	0.85	0.85	0.85	0.85	0.85	0.85
7	5	Arabica	j-ae	ahr	brazil																									
8	6	Arabica	hugo	addisu																										
9	7	Arabica	ethiopia	c/ethiopia	aslime																									
10	8	Arabica	ethiopia	c/ethiopia	aslime																									
11	9	Arabica	diamond	ethiopia	tulla	coffee	fam																							
12	10	Arabica	muhammad	ethiopia	fahem	coffee	plantation																							
13	11	Arabica	col	q	col	united	steel	fil																						
14	12	Arabica	col	q	col	united	steel	rod																						
15	13	Arabica	grounds	United	Stannaria	fams																								
16	14	Arabica	ethiopia	c/ethiopia	aslime																									
17	15	Arabica	col	q	col	united	steel	rod																						
18	16	Arabica	grounds	Indonesia	taraca	jaya																								
19	17	Arabica	ethiopia	c/ethiopia																										
20	18	Arabica	yunnan	cc	china	echo	coffee	mill																						
21	19	Arabica	esmeralda	ethiopia	dima	note																								
22	20	Arabica	col	q	col	united	steel	rod																						
23	21	Arabica	the	coffee	Costa	Rica	ceval																							
24	22	Arabica	roberto	ic	Mexico	la	heradura																							
25	23	Arabica	col	q	col	united	steel	rod																						
26	24	Arabica	j-ae	ahr	ethiopia																									
27	25	Arabica	nucoffee	brazil	fazenda	laquand																								
28	26	Arabica	ethiopia	c/ethiopia																										
29	27	Arabica	kabum	ha	guinea	choreon	23	none	kabum	ha																				
30	28	Arabica	bonmarc	chardak	los	hospita																								
31	29	Arabica	lin	che	ha	nawan	toosun	toosun	toosun	toosun																				
32	30	Arabica	nora	zees	Nicaragua																									
33	31	Arabica	specialty	C	tanzania	United	Republic	Of																						

```

x_data = xy[:, 1: -1]
y_data = xy[:, [-1]]

[ ] tf.compat.v1.disable_eager_execution()
X = tf.compat.v1.placeholder(shape=[None, 4], dtype=tf.float32)
Y = tf.compat.v1.placeholder(shape=[None, 1], dtype=tf.float32)
W = tf.Variable(tf.random.normal([4,1]), name='weight')
b = tf.Variable(tf.random.normal([1]), name='bias')

[ ] hypothesis = tf.matmul(X, W) + b

[ ] cost = tf.reduce_mean(tf.square(hypothesis - Y))
optimizer = tf.train.GradientDescentOptimizer(learning_rate=0.000005)
train = optimizer.minimize(cost)

[ ] sess = tf.Session()

[ ] sess.run(tf.global_variables_initializer())

[ ] for step in range(100001):
    cost, hypo_0 = sess.run([cost, hypothesis, train], feed_dict = {X : x_data, Y : y_data})
    if step % 500 == 0:
        print("#", step, "손실 비용", cost)
        print("- 커피점수", hypo_0)

```

저희는 스마트 농장의 고객 유치를 위해서 커피의 퀄리티에 대한 점수 예측기를 만들기 위해서 CQI(Coffee Quality Institution)에서 얻은 커피 데이터를 기반으로 커피의 중요한 특성의 Correlationship을 찾아 주요 특성을 추출하여 다항회귀 사용해 커피의 점수를 예측하는 모델을 만들어 보았습니다.

# 커피 점수 예측기



현재 저희 학원 커피 머신앞에 붙여져 있는 커피 점수 판입니다. 이 판을 가지고 점수를 예측하게 되면

# 저희 학원의 커피의 점수는??



```
Aroma = float(input("향미 점수 "))  
Flavor = float(input("맛 점수 "))  
Acidity = float(input("산도 점수 "))  
body = float(input("바디감 점수 "))
```



```
향미 점수 7  
맛 점수 9  
산도 점수 4  
바디감 점수 9
```

```
[18] with tf.Session() as sess:  
    sess.run(model)  
  
    save_path = "/content/drive/MyDrive/데이터 프로젝트/커피의 맛/데이터/학습데이터/saved.cpkt"  
    saver.restore(sess, save_path)  
  
    data=((Aroma, Flavor, Acidity, body), )  
    arr = np.array(data, dtype=np.float32)  
  
    x_data =arr[0:4]  
    dict = sess.run(hypothesis, feed_dict = {X: x_data})  
    print(dict[0])
```

```
[73.87462]
```

보시는 바와 같이 저희 학원 커피의 점수는 73점입니다.

# 결과물(App Service) 시연



저희는 결과물 시연을 아두이노에서 데이터가 쌓여지는 모습, 아두이노 기기의 실동작을 보여 드리겠습니다.

# 극복하기 어려웠던 점

## - 장비의 문제

- 현재 가지고 있는 장비의 수가 극히 적어 AI로 분석을 하는데 많은 어려움이 있었습니다. 저희가 분석한 데이터는 co2데이터가 가장 중요한데 저희는 CO2 장비가 없음
- 대부분의 스마트 팜 시스템 업체에서는 CCTV 구현을 실시하고 있으나 저희는 장비의 문제로 CCTV구현의 어려움이 있다.

## - 데이터의 문제

- 누구나 같은 문제를 가지고 있다고 생각합니다. 하지만 저희는 정말로 제한된 데이터 속에서 최대한의 능력을 뽑아내기 위해서 많은 논문과 실험을 하였으며 확실한 데이터 몇개만있었다라면 더욱 정확한 AI분석을 보여드릴수 있었다고 생각합니다.
- 시계열 데이터의 주기의 문제 또한 저희를 괴롭혔습니다. 어떠한 농가는 주차로 저장되어 있는 반면 다른 농가는 일차로 저장되어 있어 이 두가지의 농가의 데이터를 합치는데에 상당히 많은 어려움이 있었습니다.

# 더욱이 추가하고 싶었던점

- 웹서비스 구현
  - Flask를 통하여 웹서비스 연동을 하여 사용자가 편리하게 확인이 가능한 웹서비스를 구현
- 커피농장의 전자상거래화
  - 커피를 만드는것도 중요하지만 파는것이 더 중요하다고 생각하였습니다. 웹 서비스에 자사의 서비스를 이용하는 농장주들에게 각자의 커피콩을 소비자에게 직접 팔수있는 기회를 제공
- 데이터를 통하여 사람들에게 커피 추천시스템 제공하여 나에게 맞는 커피가 무엇인지를 인식
  - Surpise 라이브러리를 사용하여 CQI데이터를 기준으로 맛의 요소별로 추천시스템을 예측



# 질문



감사합니다

