**[프로젝트 보고서]**

|  |  |
| --- | --- |
| **프로젝트 명** | 실시간 날씨 데이터를 통한 토마토 가격 예측 및 스마트팜의 수확시기 추천 |
| **팀 명** | Intelligent Farm (인텔리전트팜) |
| **기간** | 2021년 02월 25일 ~ 2021년 03월 04일 |
| **최종결과물** | <https://enrbs22.aty.kr/1614489833346/main#> |
| **프로젝트 매니저** | 박경태 (인) |
| 팀원 | | | |
| NO. | 성명 | 연락처 | E-mail |
| 1 | 박경태 | 010-8685-3426 | pgt0409@gmail.com |
| 2 | 이주희 | 010-6783-0342 | dailylog.a@gmail.com |
| 3 | 정현모 | 010-4180-6116 | jhm6116@gmail.com |
| 4 | 윤두균 | 010-7799-7819 | enrbs22@gmail.com |
| 5 | 김인기 | 010-9307-9486 | elvins1107@gmail.com |



**1. 팀 소개**

**- 팀명 : Intelligent Farm (스마트팜을 넘어 인텔리전트팜으로!)**

ICT를 접목한 스마트팜이 보편적으로 확산되면서 노동에너지 등 투입 요소의 최적 사용을 통해 우리 농업의 경쟁력을 한층 높였다.

이제는 스마트팜을 넘어 인텔리전트팜으로 성장할 시대가 왔다. 우리나라 농업에 새로운 바람과 같은 존재를 필요로 하는 시대가 왔다. 그게 우리다.

**- 역할 :**

박경태 : PM

이주희 : DB 구축(MongoDB, MySQL) / DB 설계(MySQL)

정현모 : 분석 모델 개발(Pandas)

윤두균 : Front-End 개발(Web, PowerBI, Kibana, Grafana)

김인기 : 데이터 전처리(Logstash) / 클러스터 구축(ELK)

**2. Project 주제 :** “실시간 날씨 데이터를 통한 토마토 가격 예측 및 수확시기 추천”

- 많은 사람들에게 사랑받고 있는 토마토는 과거엔 토마토를 주로 생과로 소비했었습니다. 그러나 현재에는 가공 식품에서 뿐만 아니라 여러 요리의 주 재료로써 사랑받고 있습니다. 토마토의 소비가 늘어나면서 자연스럽게 토마토의 생산과 가격에 관심이 생겼습니다.

- 이런 토마토는 외부 환경에 비교적 영향을 약하게 받는 시설에서 재배를 하고 있습니다. 그러나 토마토의 생산에는 외부 환경인 날씨여 영향을 받는 것으로 보였고, 이 부분을 알아내기 위해 토마토의 가격과 날씨 데이터를 토대로 가격을 예측하고, 토마토 농장에 정보를 주고자 합니다.

**3. Project 개요**

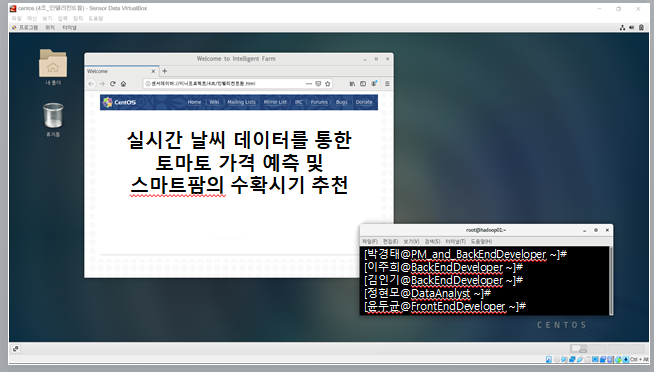
**- 소개 :**

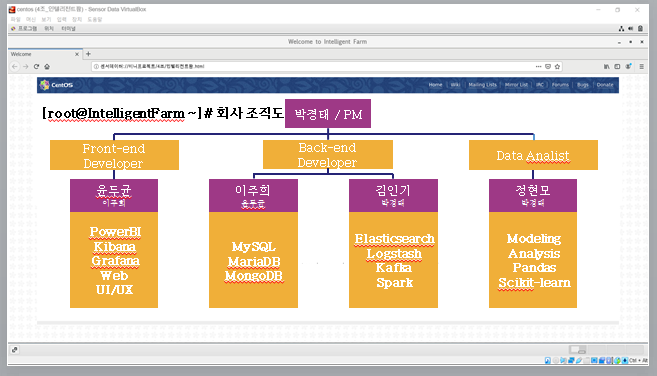
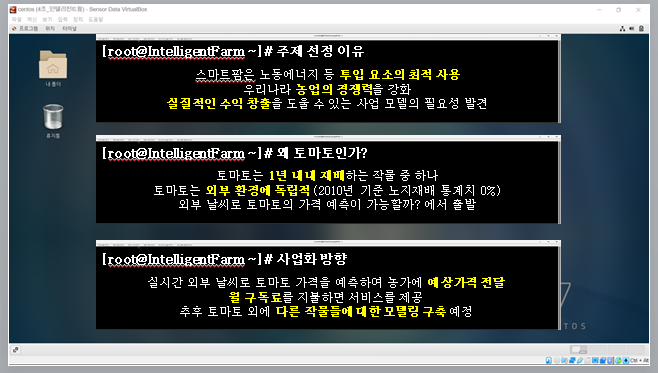
"Intelligent Farm”은 노지재배가 사라진 토마토 농사에 외부 요인이 영향을 미치는지, 미친다면 얼마만큼 미치는지를 가격과 날씨요인을 통해 분석해보고 이를 예측해 보고자 한다.

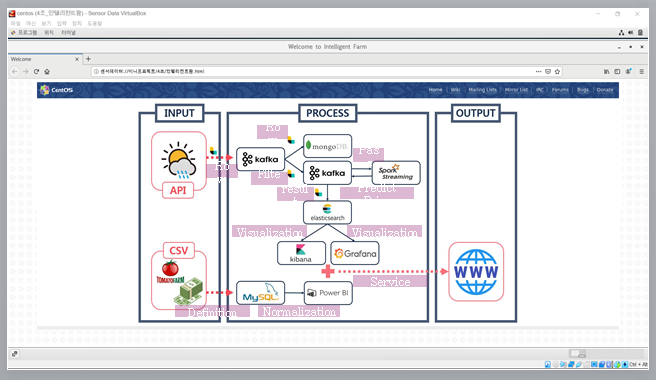
**- 주제 선정 이유 :**

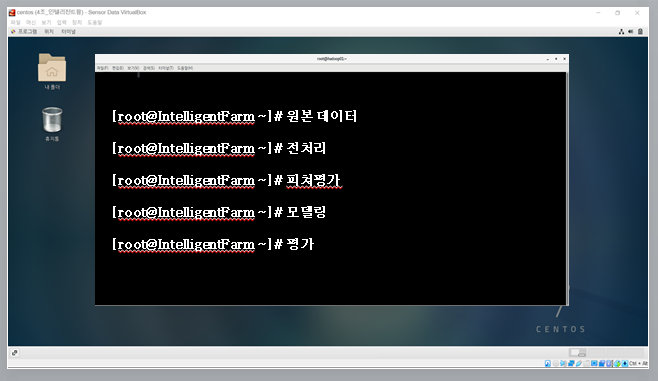
토마토는 2009년 이후로 국내에서 노지재배가 사라진 작물이다. 그렇기에 모든 생산량이 통제된 환경에서 생산이 되고 있다. 그러나 토마토의 가격은 날씨와 계절에 따라 달라지는 경향성을 보여주었다.

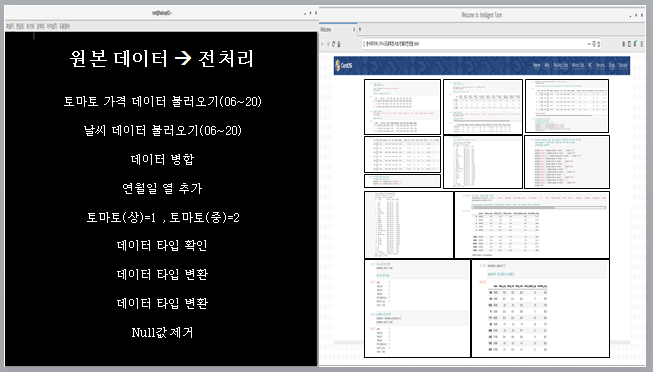
그렇기에 실시간 데이터를 활용할 수 있는 날씨데이터와 토마토 도매가격을 통해서 외부 요소인 날씨가 얼마나 토마토 작물에 영향을 미치는지 알아보고, 이를 토대로 토마토의 가격을 예측하여 토마토를 키우는 농장에 도움이 되고자 한다.

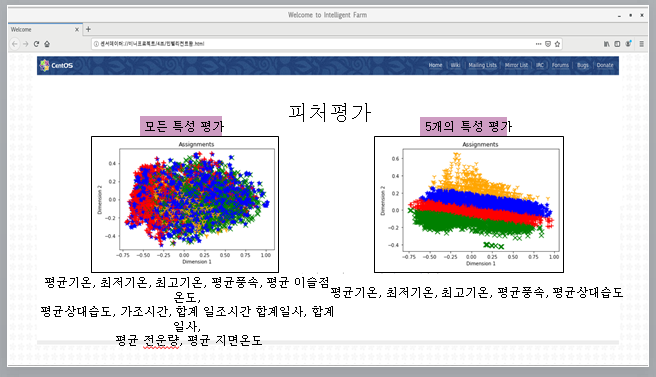
****

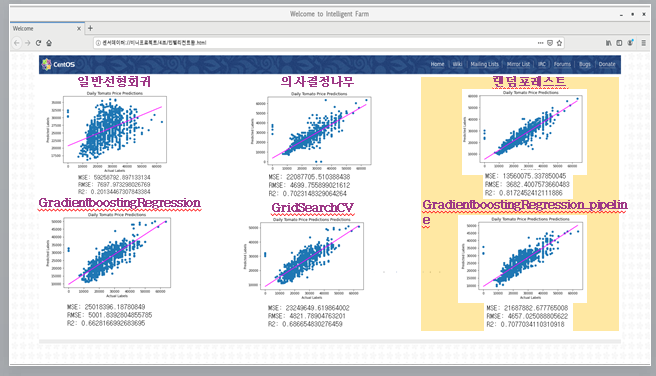
****

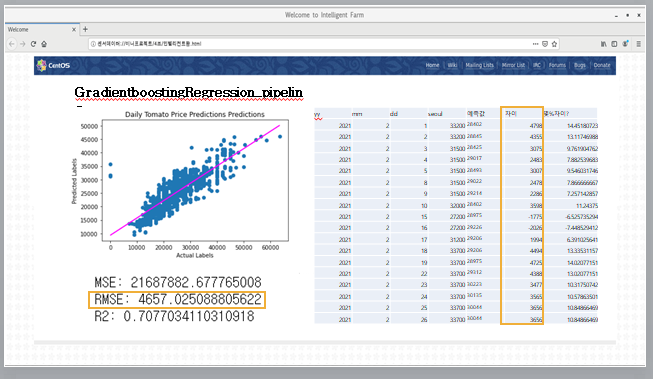
****

****

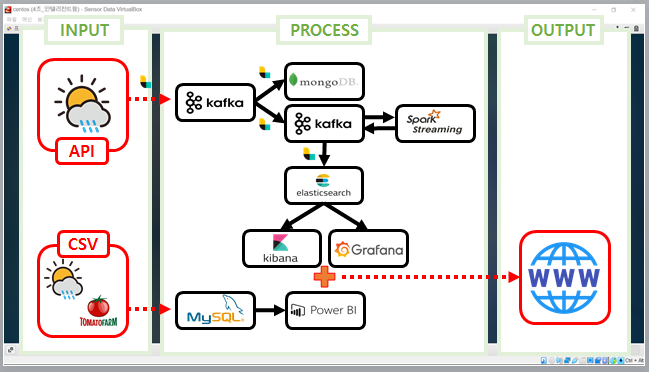
****

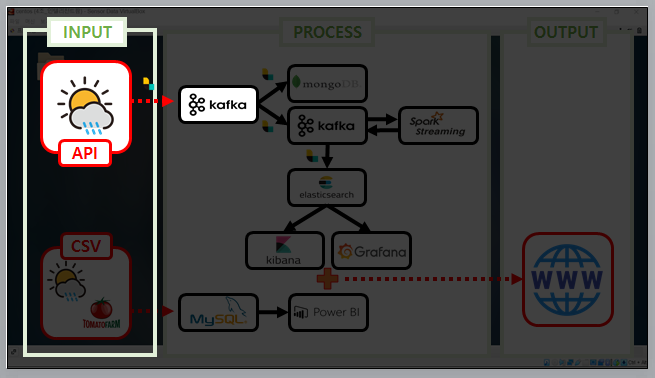
****

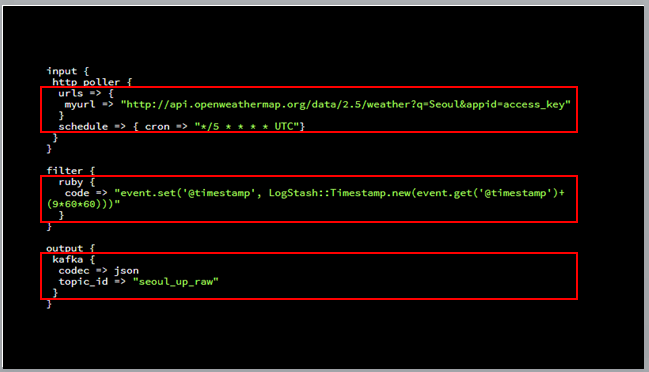
****

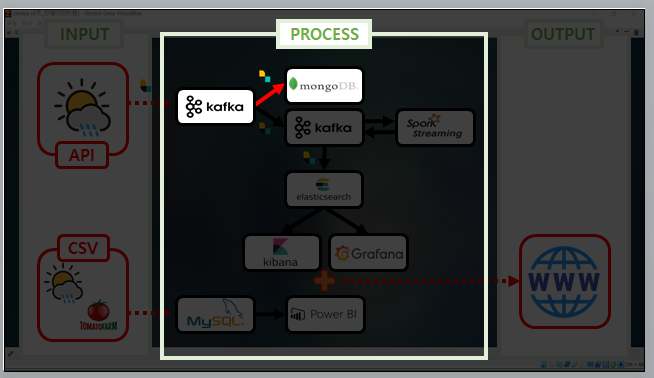
****

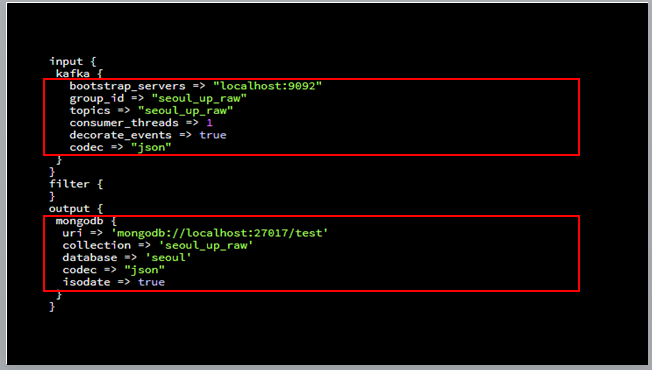
****

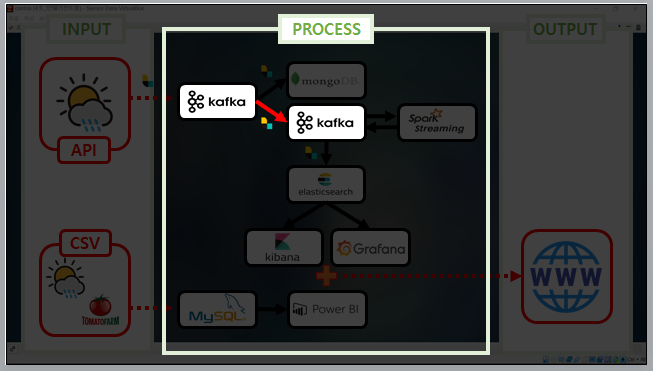
****

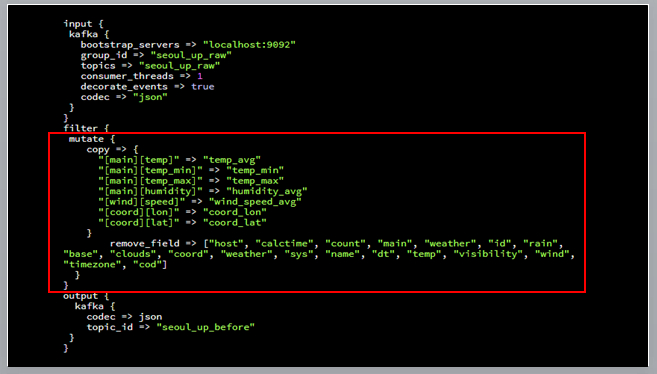
****

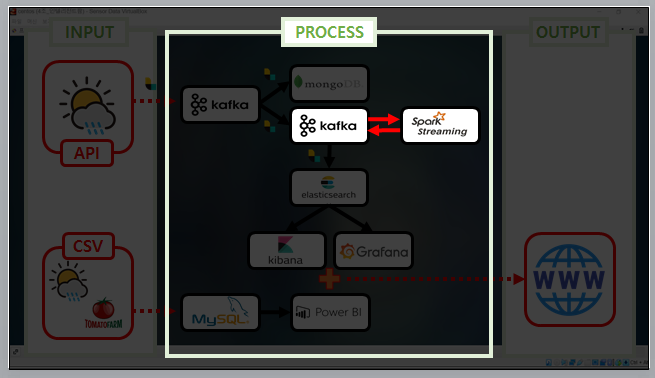
****

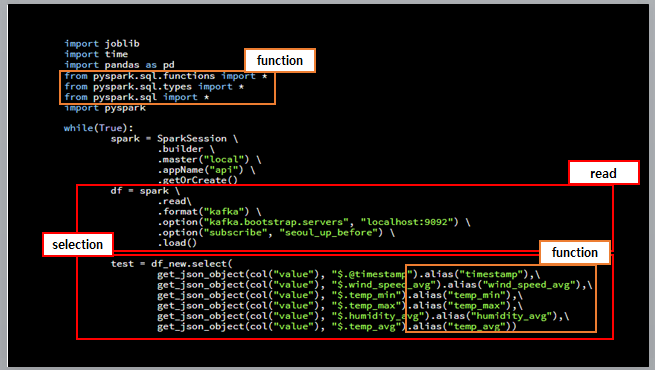
****

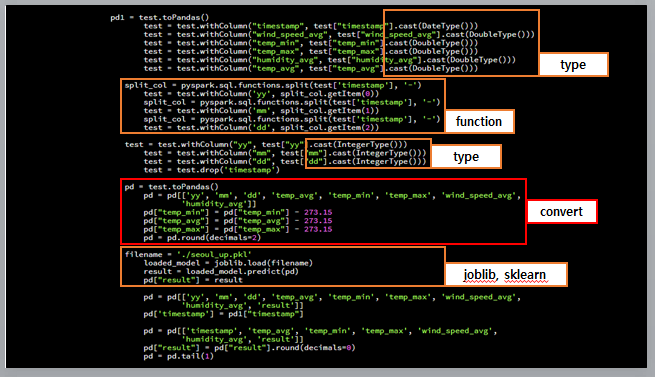
****

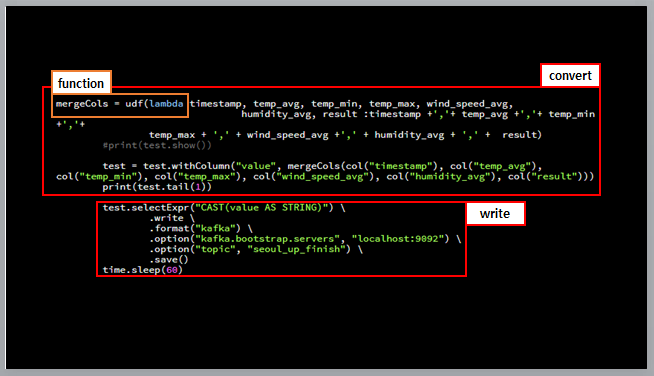
****

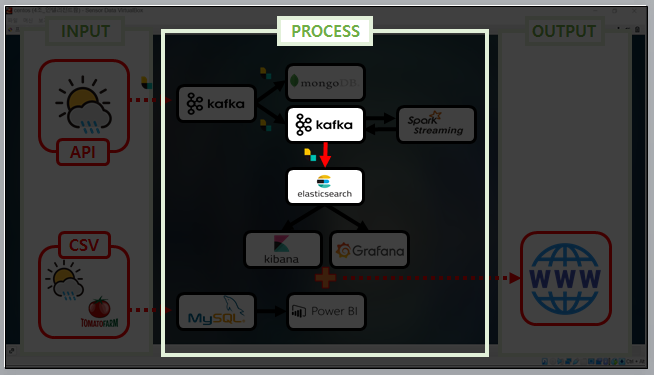
****

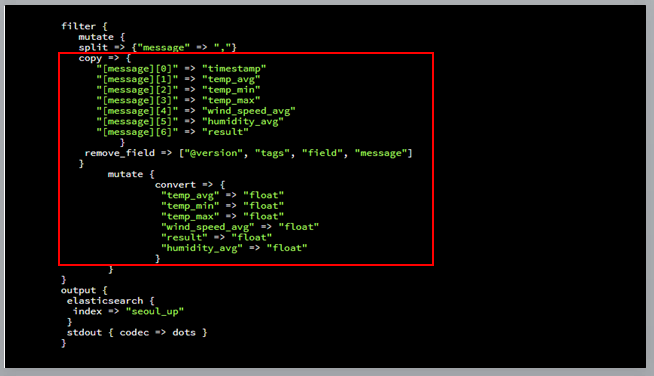
****

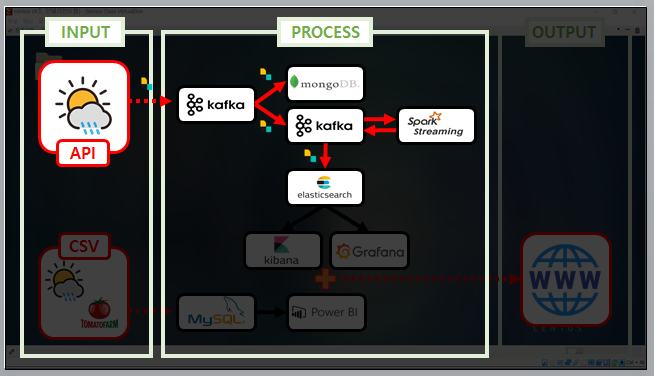
****

****

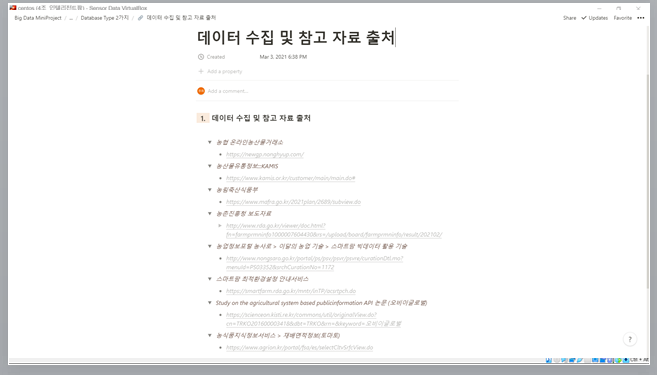
****

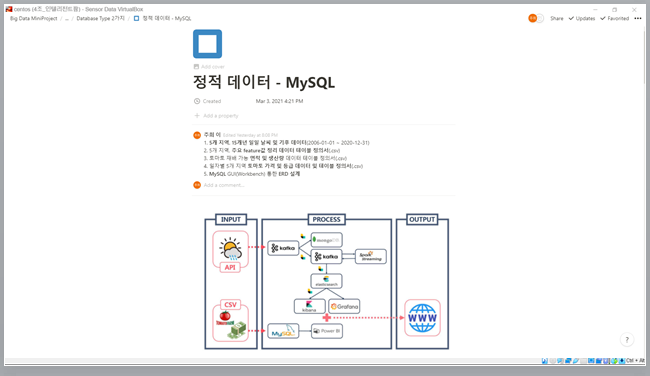
****

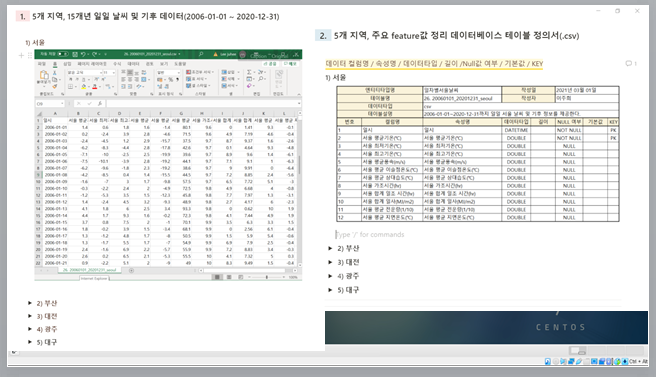
****

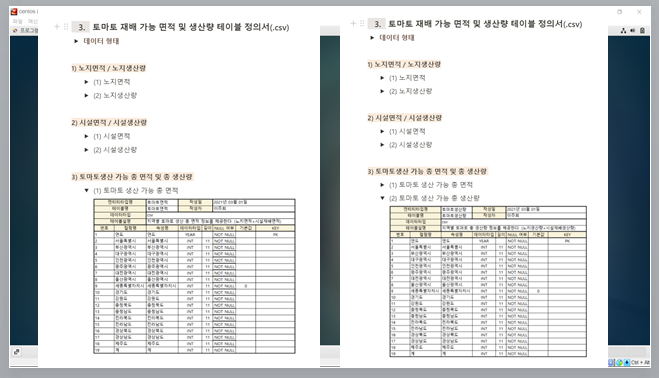
****

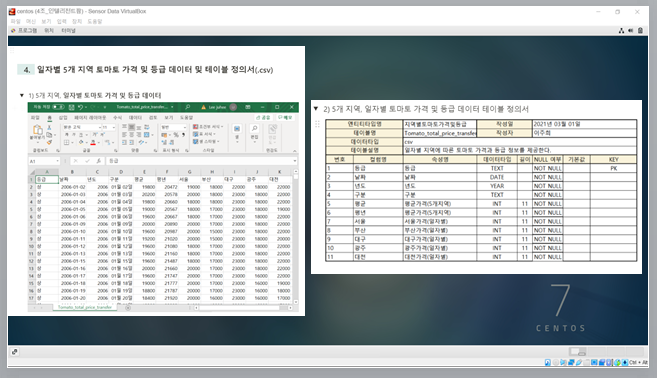
****

****

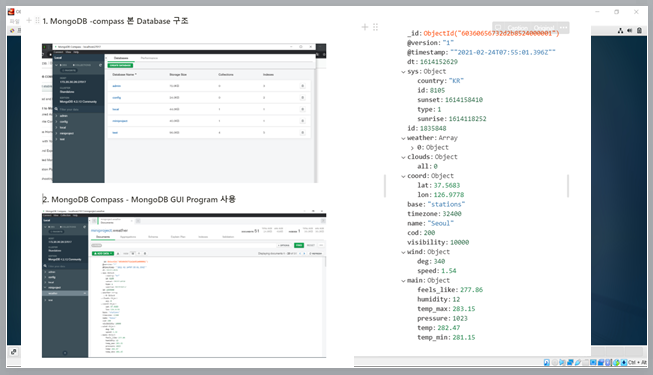
****

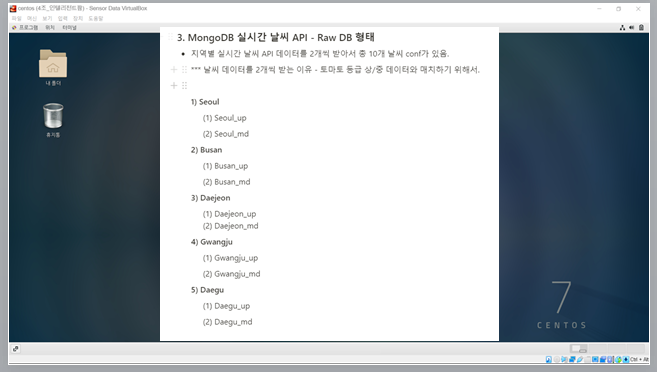
****

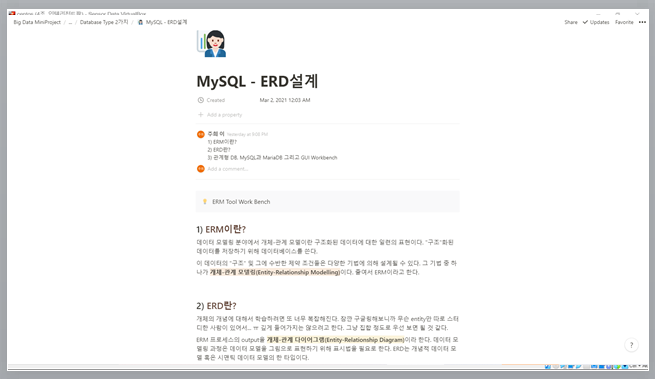
****

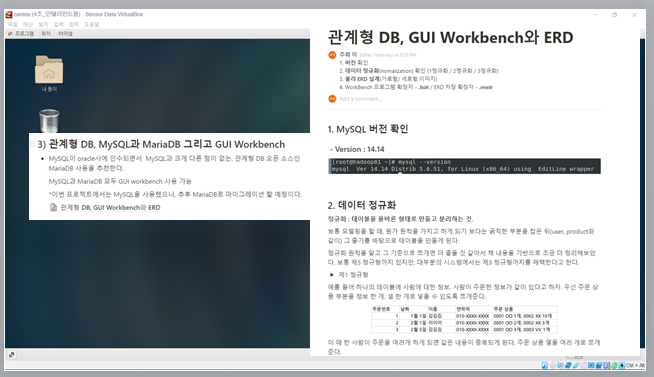
****

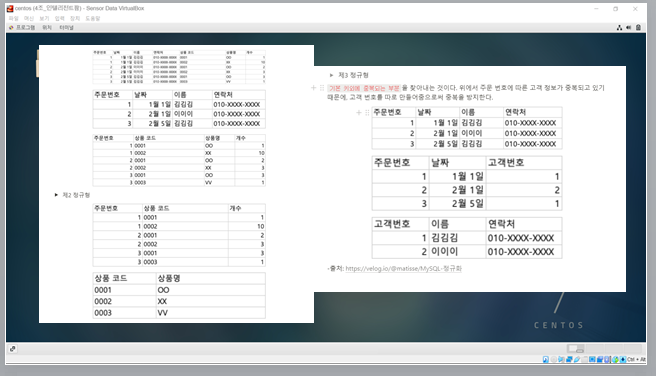
****

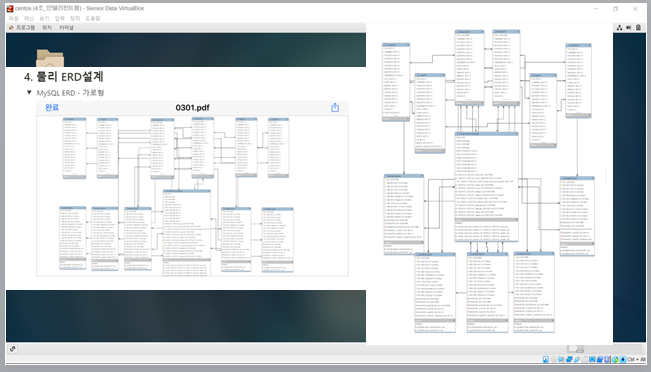
****

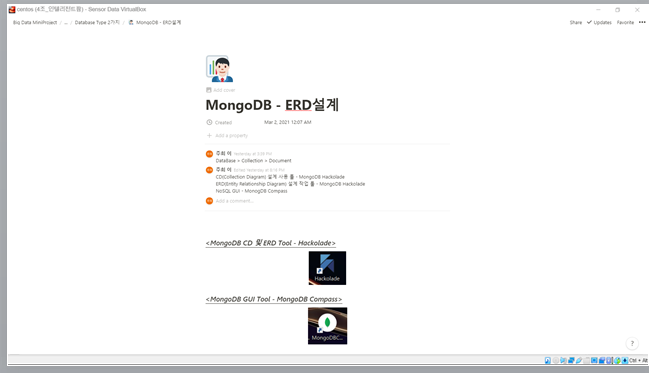
****

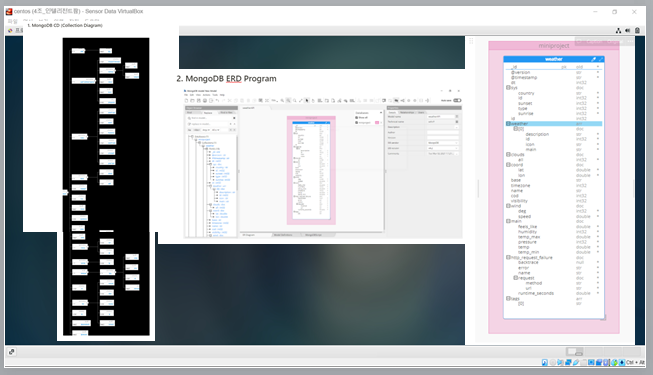
****

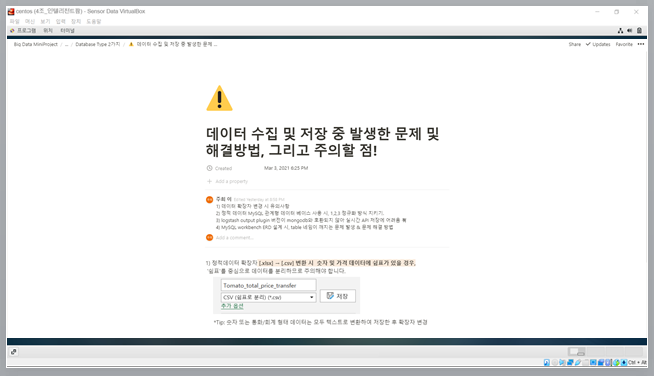
****

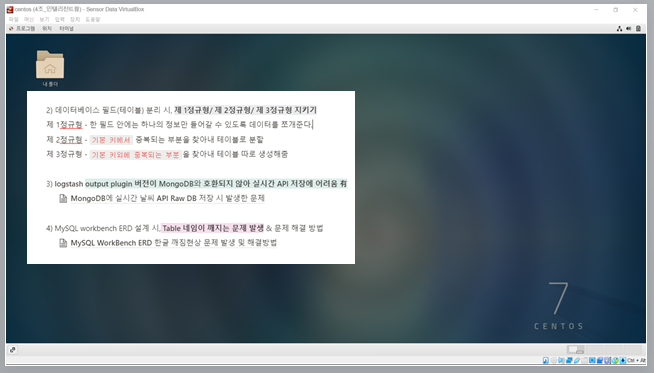
****

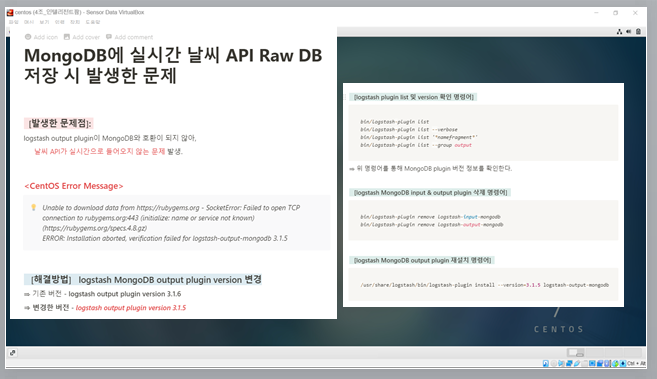
****

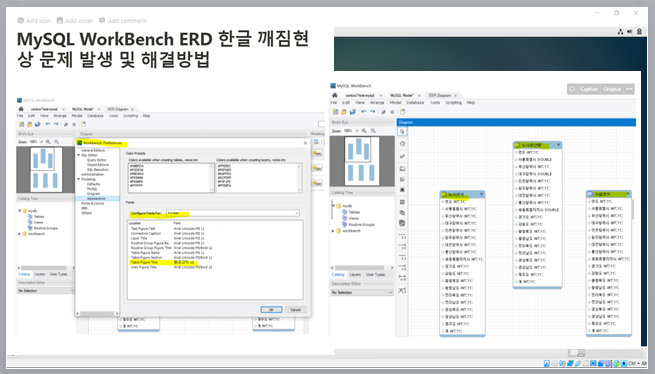
****

****

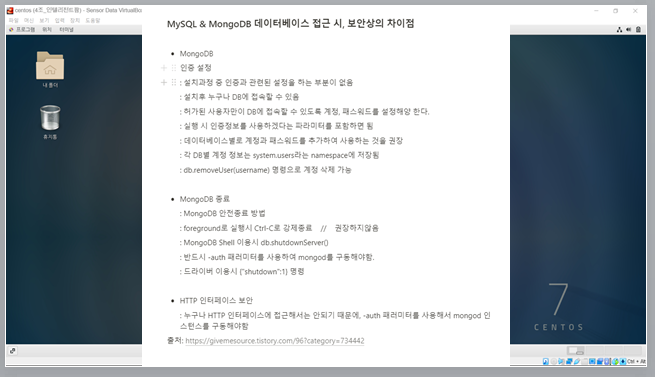
****

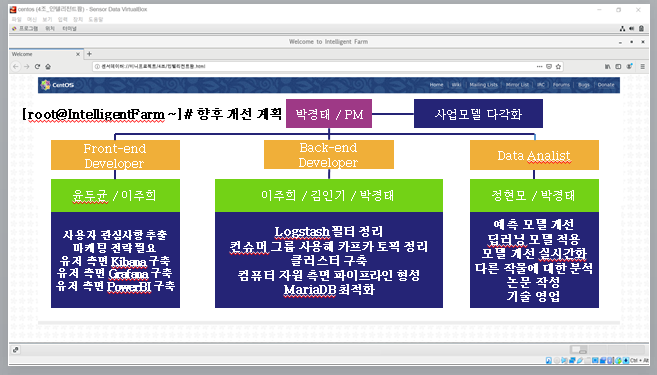
****

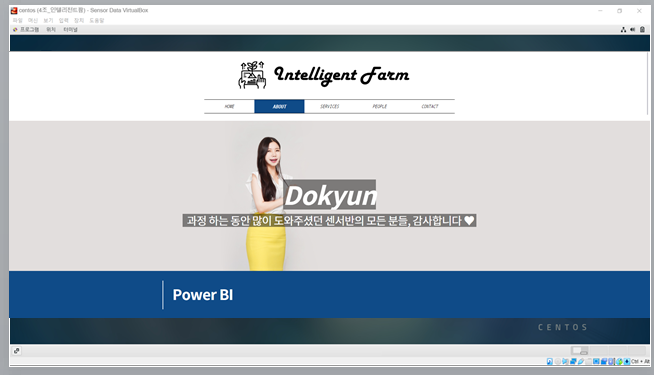
****

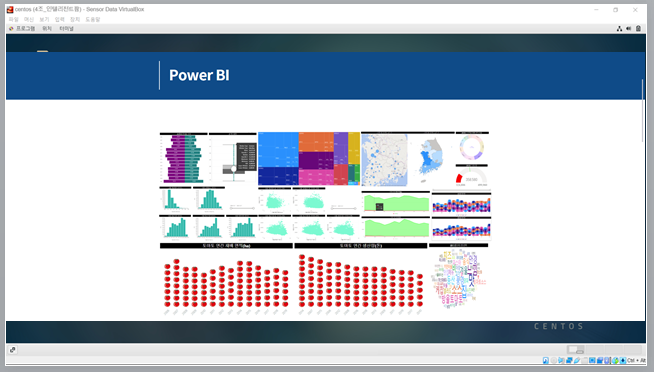
****

****

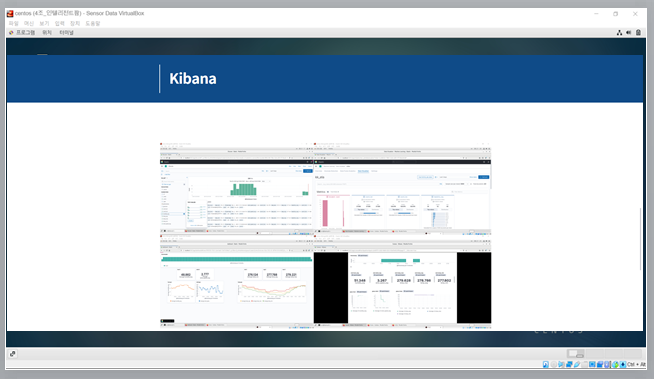
****

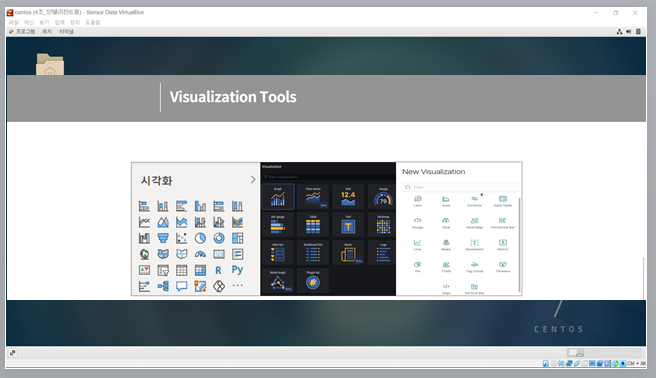
****

****

****

****

****

****

**4. 활용 범위**

[한층 더 지혜로워진 농업]

- 우리나라의 모든 농업 종사자에게 예측 가능한 수익모델을 선물한다. 스마트팜을 이용해 농작물 수확의 편리함을 가져다주었다면, 실질적으로 피부에 와닿는 경제성이 필요하다.

- 스마트팜은 고도화된 센서데이터 인프라와 작물 데이터를 활용하여 기존의 농업 패러다임을 바꿨습니다. 작물 재배에 있어 날씨를 비롯한 외부 환경 요인에 영향을 덜 받게 되었으며 원하는 시기에 원하는 작물을 재배할 수 있게 되었습니다. 하지만, 재배환경이 고도화되었음에도 도소매 과정은 여전히 도소매상들과 대형유통사가 주도권을 쥐고 가격을 결정하고 있습니다. 이에 가격 예측 모델을 개발하여 개별 농가들이 작물의 가격을 직접 예측 가능하도록 만들고자 합니다.

- 가격 예측 모델은 중간 유통 가격 협상 과정에서 스마트팜 농가에 가격 예측 데이터를 제공하여 협상에 유리하게 만듭니다. 더 나아가, 농가에 수익이 많이 발생하는 생육, 재배 시기를 추천하여 개별 농가들의 수익성을 높일 수 있습니다.

**5. DB 구축 및 데이터 분석 내용**

● **논리 구성**



- 정적데이터 : 서울, 대전, 대구, 부산, 광주 날씨 데이터 (2006년~2020년)

-> 머신러닝 모델링

* 정적데이터 : 서울, 대전, 대구, 부산, 광주 토마토 가격 데이터 (2006년~2020년)

-> 머신러닝 모델링

동적데이터 : 서울, 대전, 대구, 부산, 광주 날씨 데이터 (실시간 API)

● **사용 기술**

|  |  |  |  |
| --- | --- | --- | --- |
| **구분** | **이름** | **버전** | **목적** |
| 개발환경 | Oracle VM | 6.1.16 | 가상머신 |
| CentOS7 | CentOS-7\_x86\_64 | OS |
| RAM | - | 8G |
| 사용 Tool | OpenJDK | 1.8.0\_282 | - |
| Logstash | 7.10.2 | 수집 및 저장 |
| MongoDB | 4.2.11 | Rawdata 관리 |
| Kafka | 2.4.1 | 메시지 관리 |
| Zookeeper | 3.5.6 | 카프카 관리 |
| Spark | 2.4.7 | 모델 적용 |
| Elasticsearch | 7.10.2 | 동적데이터 저장 |
| Kibana | 7.10.2 | 동적데이터 시각화 |
| Grafana | 7.3.7 | 동적데이터 시각화 |
| MySQL | 14.14 | 정적데이터 저장 |
| PowerBI | 2.90.782.0 | 동적데이터 시각화 |