Time Series

윤재경 정재원 전혜민
윤빈나 권지혜 최성웅

KU-BIG

INDEX

1. Project
- Data
- Visualization

2. Model Introduction
- ARIMA
- RNN based models

3. Future Schedule
- Model Study
- Data Preprocessing

KU-BIG

Project - Great Energy Predictor III

https://www.kaggle.com/c/ashrae-energy-prediction/overview

KU-BIG

Project - Data

building_meta.csv

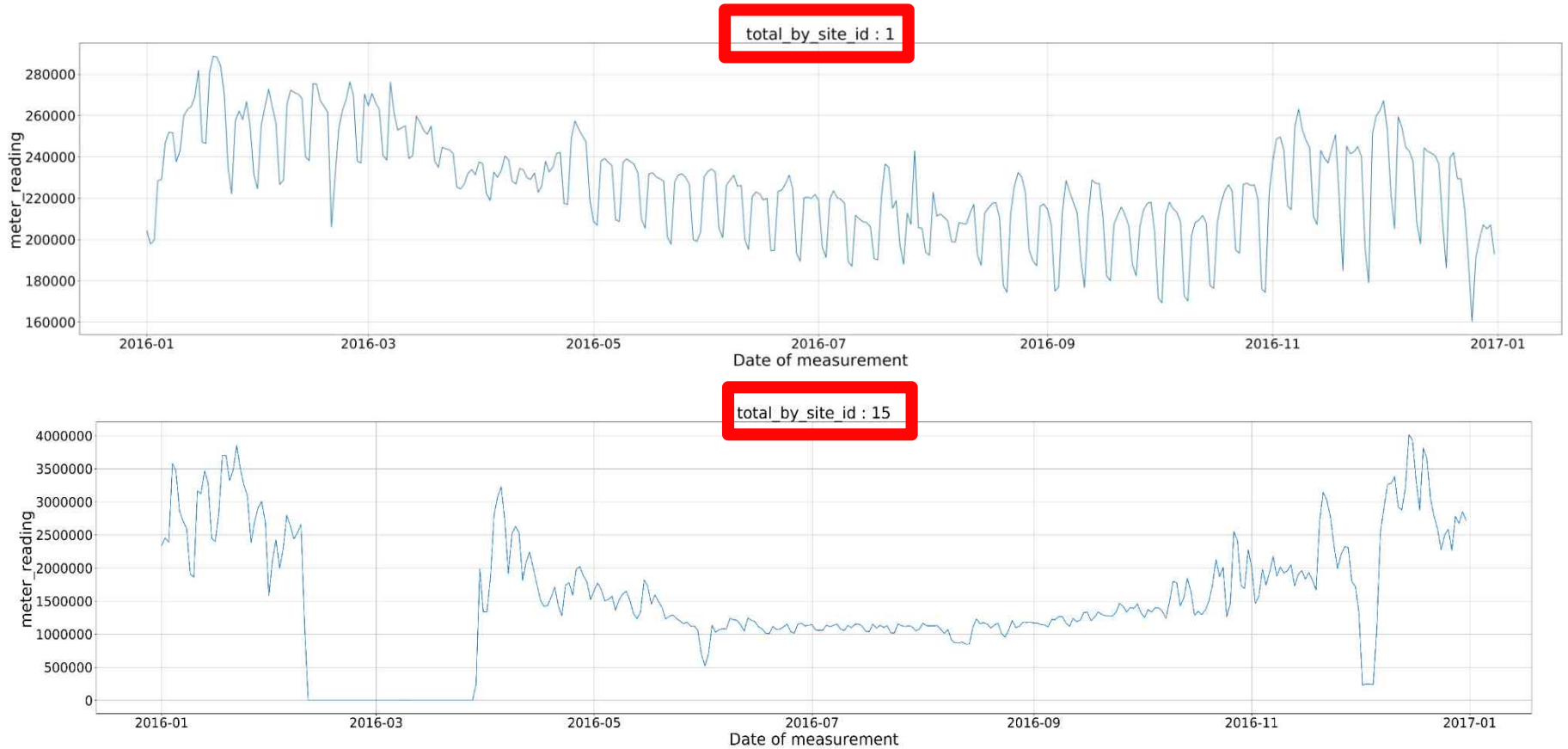
- `site_id` - Foreign key for the weather files.
- `building_id` - Foreign key for `training.csv`
- `primary_use` - Indicator of the primary category of activities for the building based on EnergyStar property type definitions
- `square_feet` - Gross floor area of the building
- `year_built` - Year building was opened
- `floor_count` - Number of floors of the building

weather_[train/test].csv

Weather data from a meteorological station as close as possible to the site.

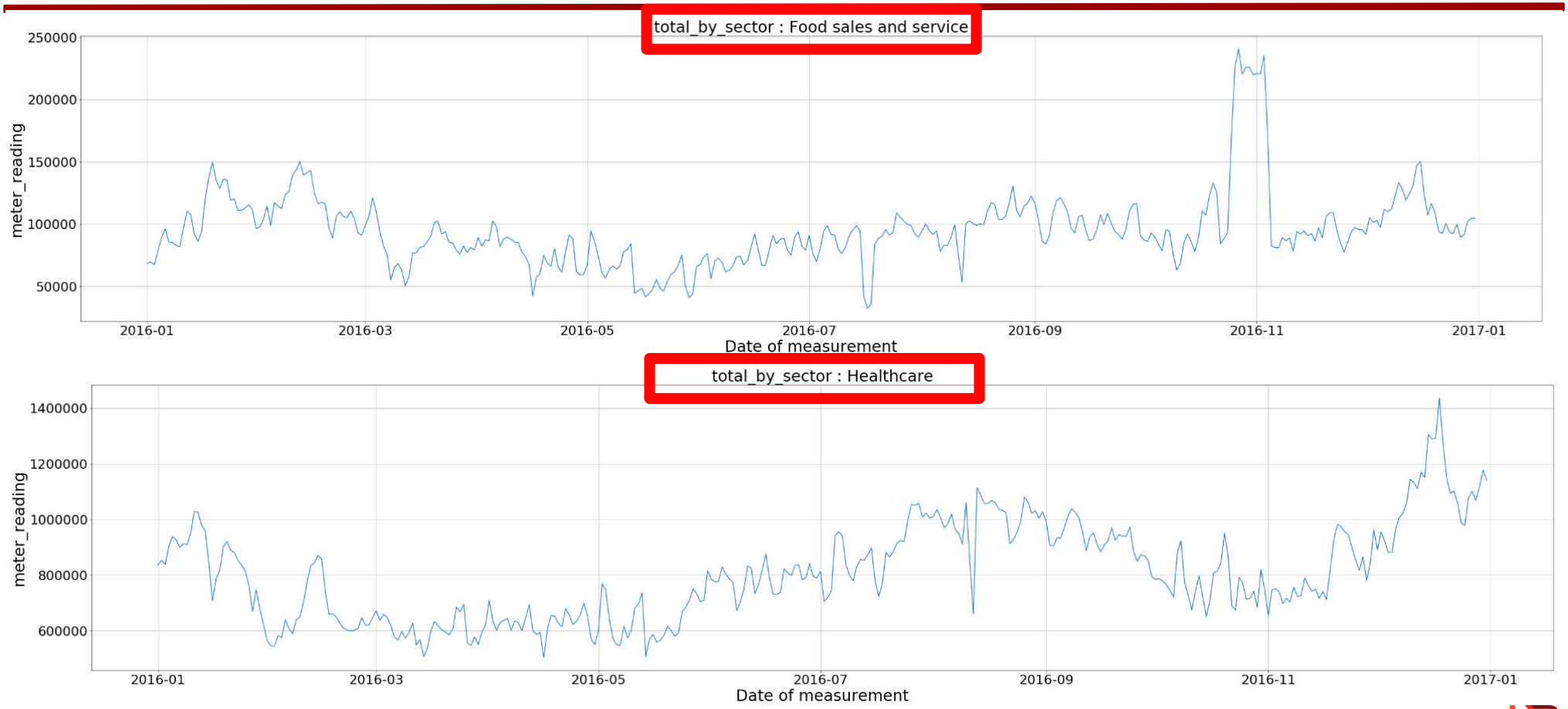
- `site_id`
- `air_temperature` - Degrees Celsius
- `cloud_coverage` - Portion of the sky covered in clouds, in oktas
- `dew_temperature` - Degrees Celsius
- `precip_depth_1_hr` - Millimeters
- `sea_level_pressure` - Millibar/hectopascals
- `wind_direction` - Compass direction (0-360)
- `wind_speed` - Meters per second

Project - Visualization (Per site)



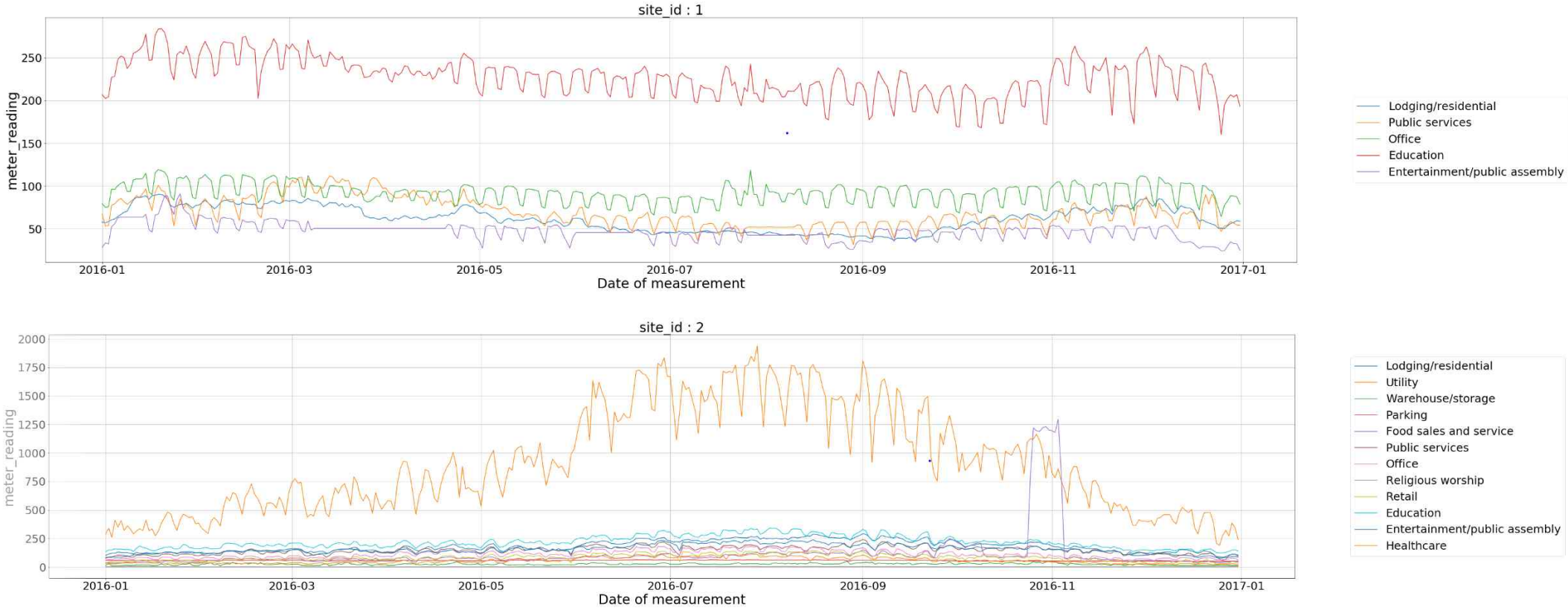
KU-BIG

Project - Visualization (Per sector)



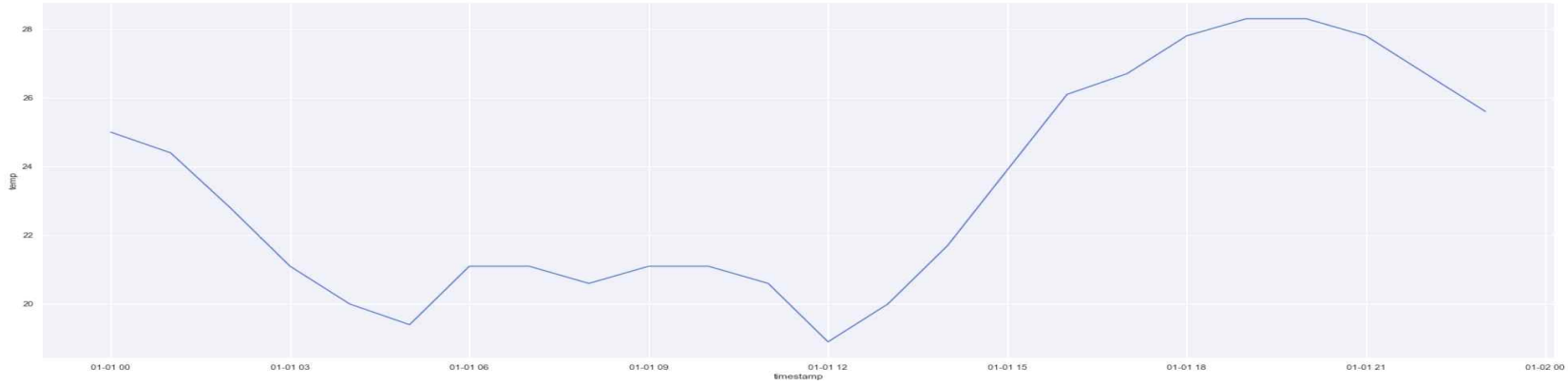
KU-BIG

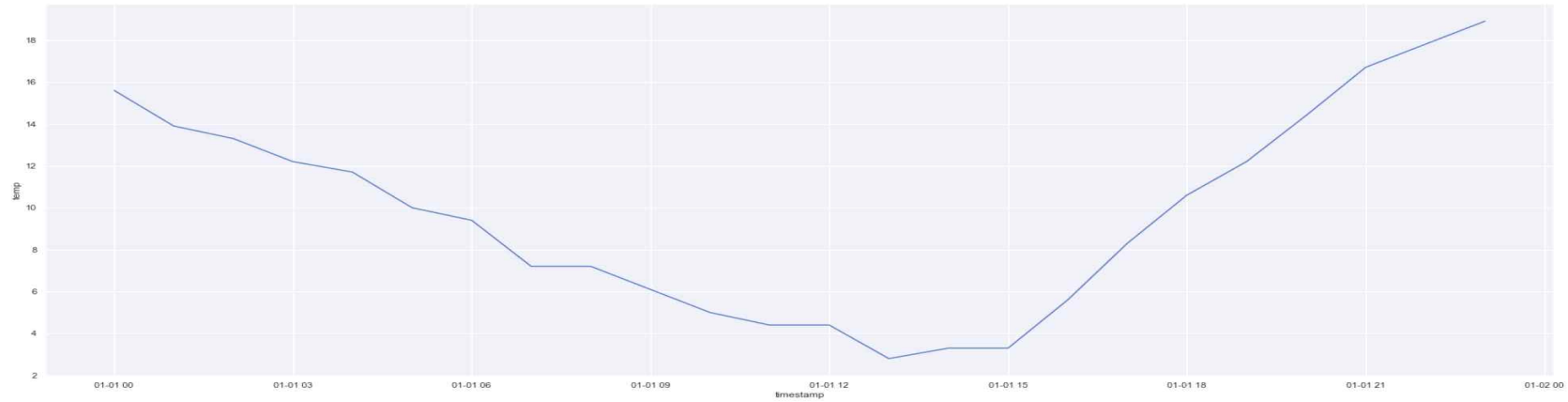
Project - Visualization



KU-BIG

Project - Visualization (Temperature in Jan 1st, 2016)





KU-BIG

Project - Visualization

기온 데이터의 PCA 결과를
기반으로 한 시각화 결과

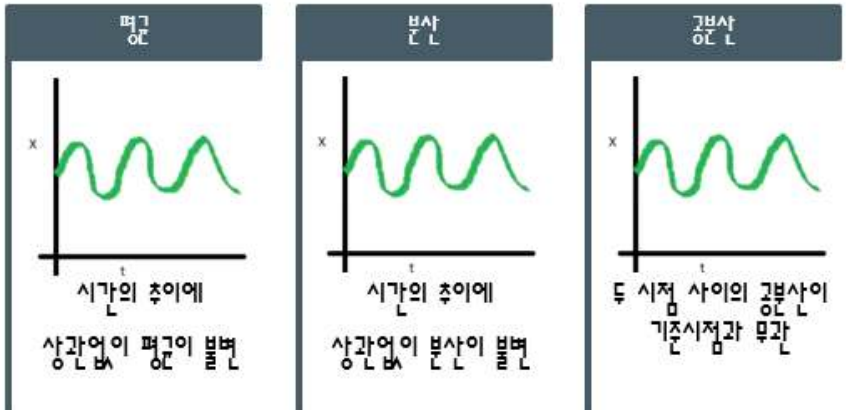
=> 미국과 유사한 형태를 띄고 있다는
내용 발견

이외에도 다양한 시각화 및 EDA를 통해
각 변수의 특성 및 관계를 분석해나가는
것이 관건

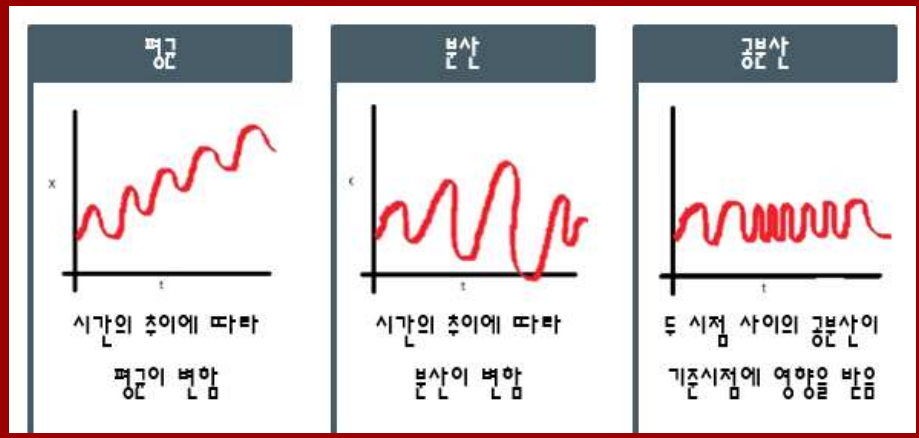
KU-BIG

Model 1 – ARIMA
(Auto-Regressive Integrated Moving-Average)

Stationary Series



Non-Stationary Series



KU-BIG

Model 1 – ARIMA (Auto-Regressive Integrated Moving-Average)

AR (Auto-Regression) :
이전 관측값이 이후 관측값에 영향을 주는 모형

$$AR(1) : X_t = \phi X_{t-1} + \epsilon_t$$

I (Integrated) : 누적을 의미

MA (Moving-Average) :
관측값이 이전의 연속적인 오차항의 영향을 받는 모형

$$MA(1) : X_t = \epsilon_t - \beta_1 \epsilon_{t-1}$$

ARIMA(p, d, q)

$$\hat{y}_t = \mu + \phi_1 y_{t-1} + \ldots + \phi_p y_{t-p} - \beta_1 \epsilon_{t-1} - \ldots - \beta_q \epsilon_{t-q}$$

- $\mu = constant$
- $\phi_1 y_{t-1} + \ldots + \phi_p y_{t-p} : ARterms(laggedvaluesofy)$
- $-\beta_1 \epsilon_{t-1} - \ldots - \beta_q \epsilon_{t-q} : MAterms(laggedvaluesofy)$
- $\hat{y}_t = Y_t, \ if \ d = 0$
- $\hat{y}_t = Y_t - Y_{t-1}, \ if \ d = 1$
- $\hat{y}_t = (Y_t - Y_{t-1}) - (Y_{t-1} - Y_{t-2}), \ if \ d = 2$

KU-BIG

Model2 – RNN based models

We can process a sequence of vectors **x** by applying a **recurrence formula** at every time step:

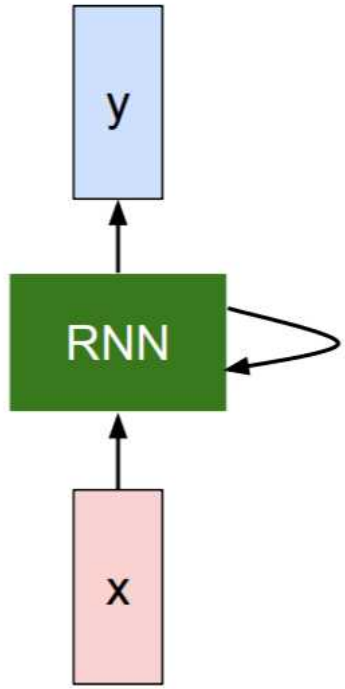
$$h_t = f_W(h_{t-1}, x_t)$$

new state

some function with parameters W

old state

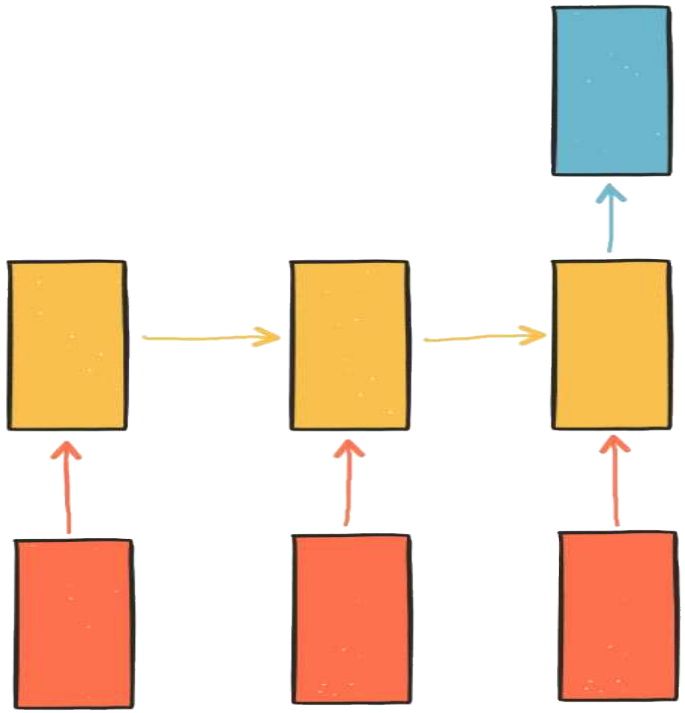
input vector at some time step



출처 : http://cs231n.stanford.edu/slides/2019/cs231n_2019_lecture10.pdf

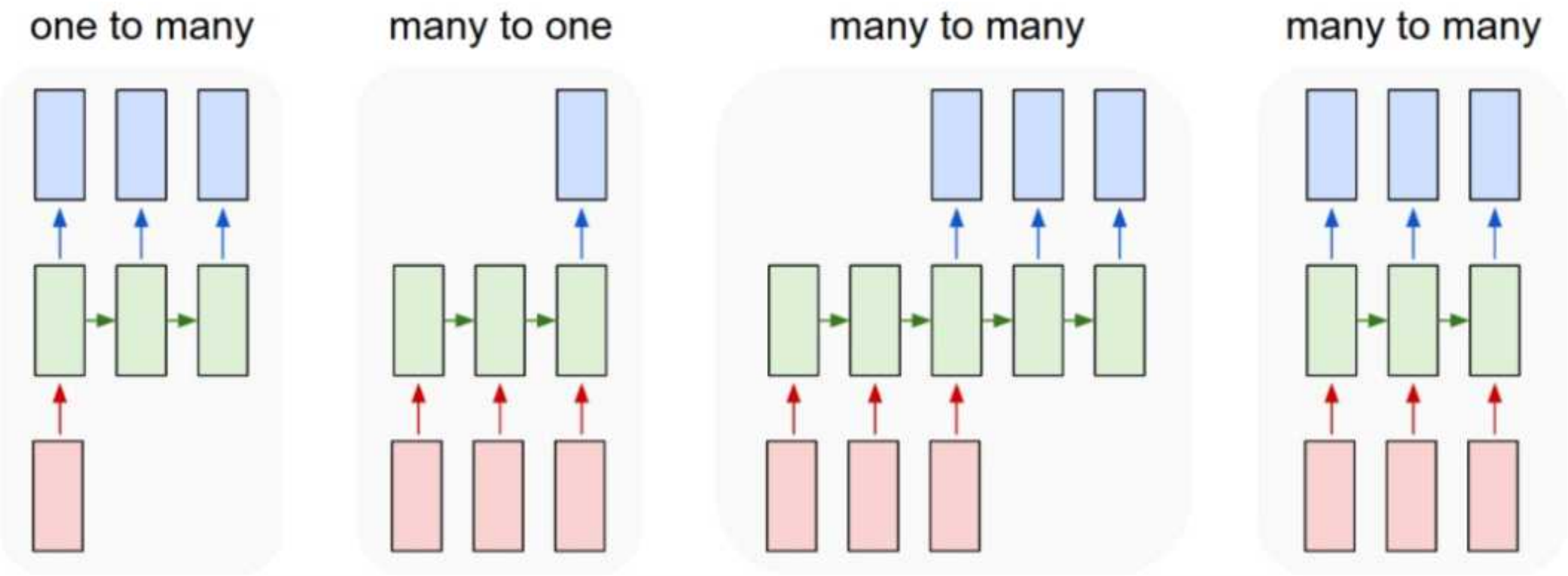
KU-BIG

Model2 – RNN based models

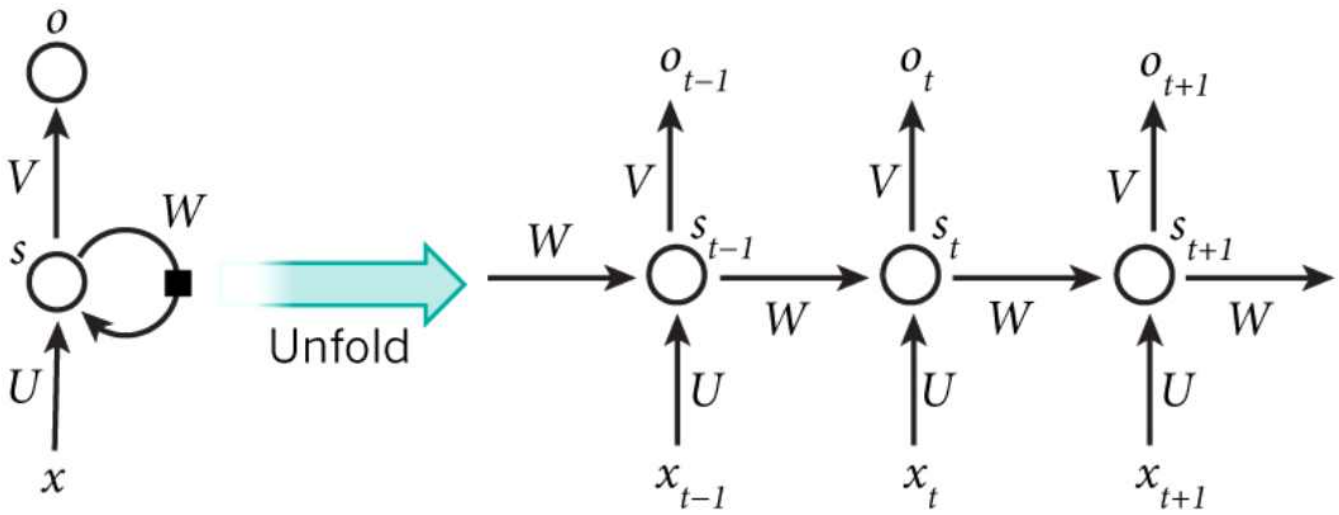


RNN Diagram

Model2 – RNN based models



Model2 – RNN based models



출처 : https://aikorea.org/blog/rnn-tutorial-1/

KU-BIG

1. Model study

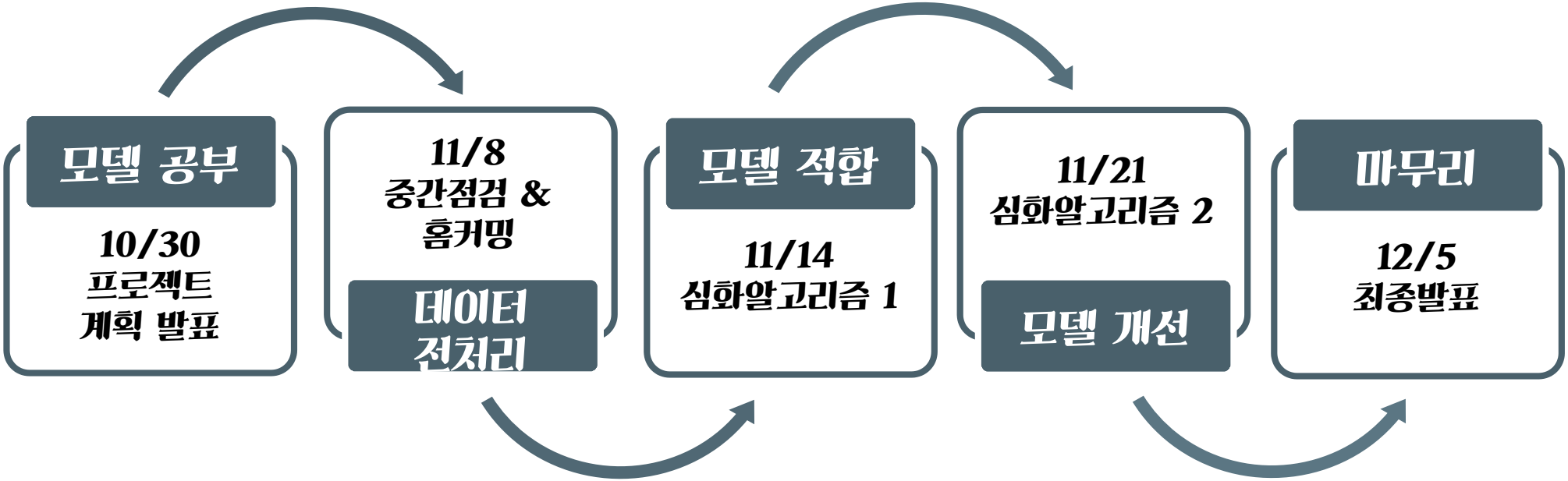
2. Processing & EDA

DATA CAMP

COUSERA

KU-BIG

Project Plan



KU-BIG

Do you
have any
question?
Thank you
for your attention.