

# Status report

JUNHO LEE

2018.07

# Outline

- ▶ 1. Additional web to be scrapped. → hourly Data
- ▶ 2. 饿了么, 京东。
- ▶ 3. Study of Air in SEOUL

# 1. Additional Web page :: Hourly Data of SEOUL Air pollutants

- ▶ Macro Location : Project-pre/func/w0\_SEOUL\_pollute\_hourly.py
- ▶ Adjustable params : Starting Date(yyyy/mm/dd), duration, output text filename(stacking, no overriding).
- ▶ Only after 2008/01/01, no data before .

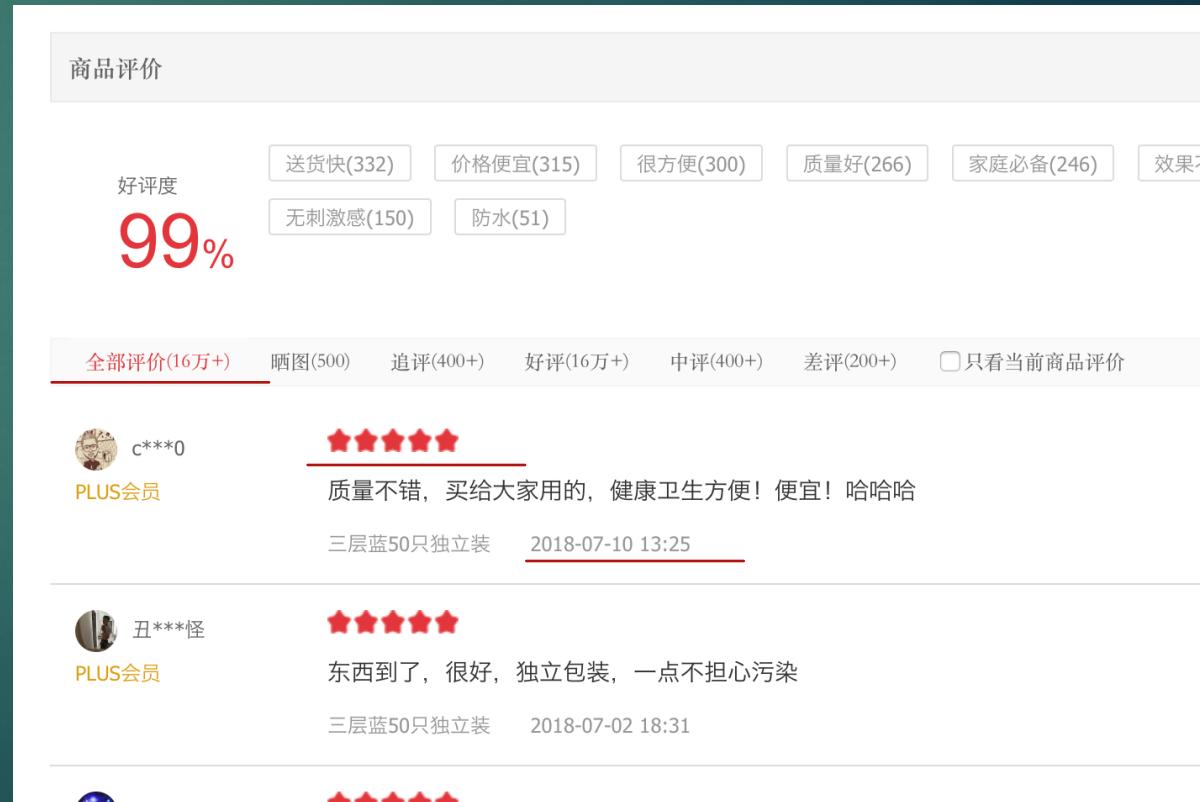
※ 0 ~ 10분 사이에는 대기오염측정소에서 자료수집 중이므로 이전 시간의 자료가 조회됩니다.

■ 측정소명 : 측정소별 전체 , 측정단위 : 시간평균 , 측정일자 : 2008.07.15

날짜	측정소명	미세먼지	초미세먼지	오존	이산화질소	일산화탄소	아황산가스	통합대기환경지수		
		PM-10( $\mu\text{g}/\text{m}^3$ )	PM-2.5( $\mu\text{g}/\text{m}^3$ )	O <sub>3</sub> (ppm)	NO <sub>2</sub> (ppm)	CO(ppm)	SO <sub>2</sub> (ppm)	등급	지수	결정물질
2008-07-15 05:00	평균	71	43	0.008	0.026	0.4	0.006	-	-	-
2008-07-15 05:00	종로구	74	40	0.006	0.025	0.3	0.011	-	-	-
2008-07-15 05:00	중구	55	44	0.008	0.025	0.6	0.005	-	-	-

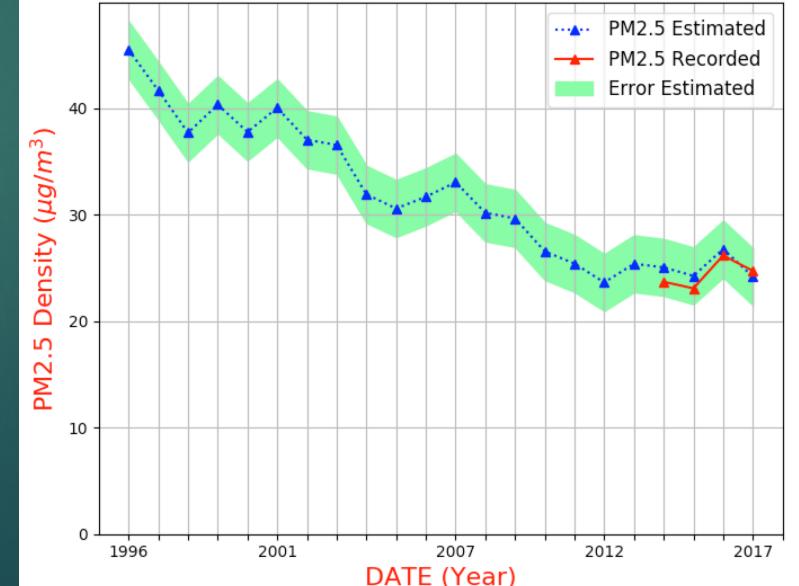
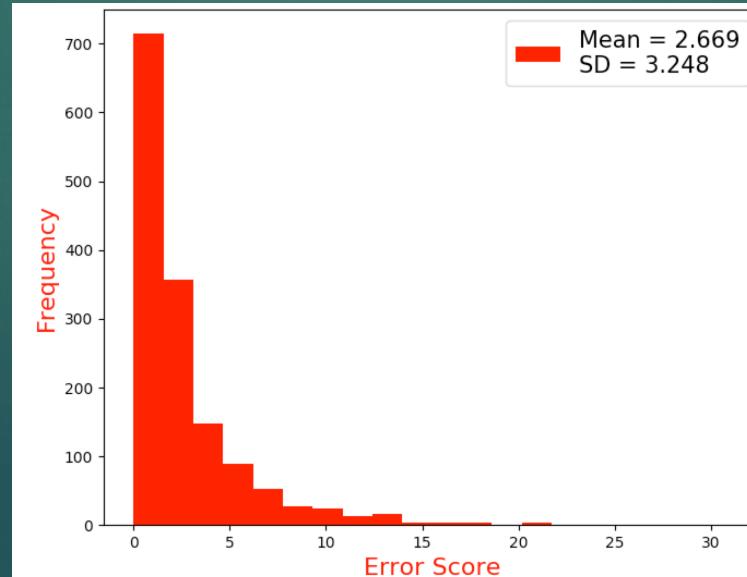
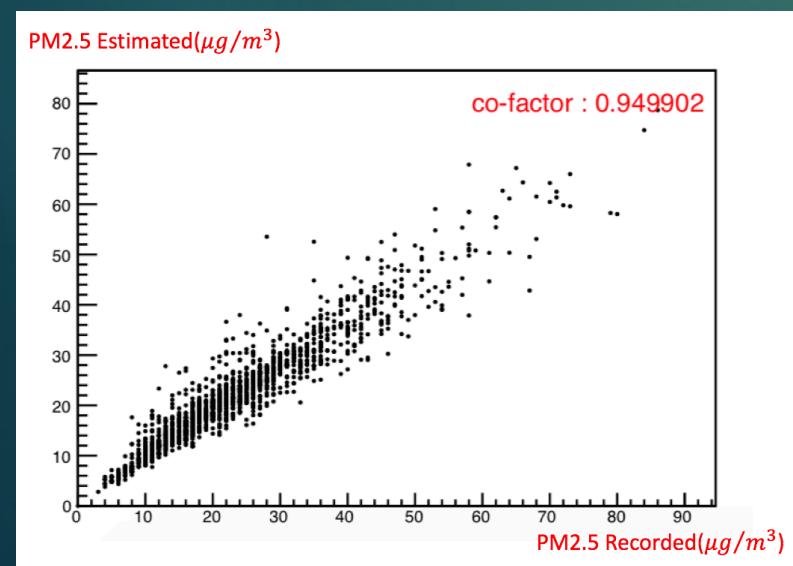
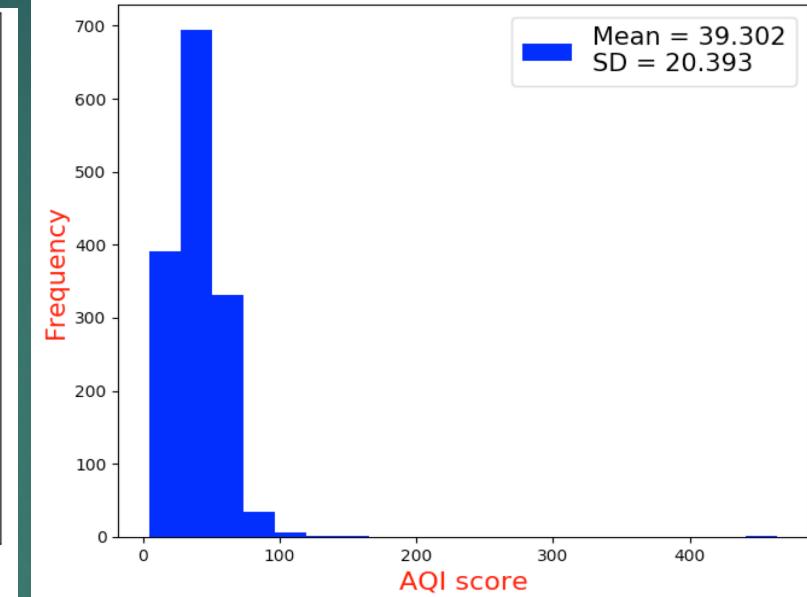
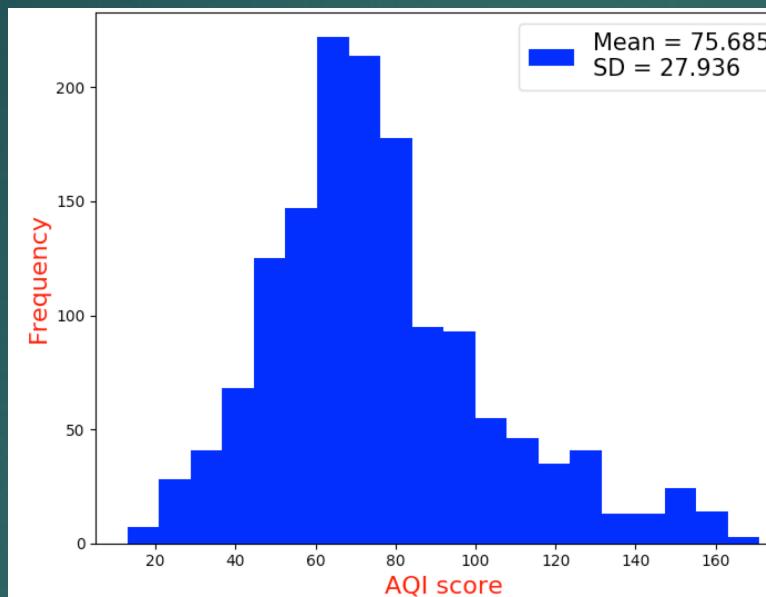
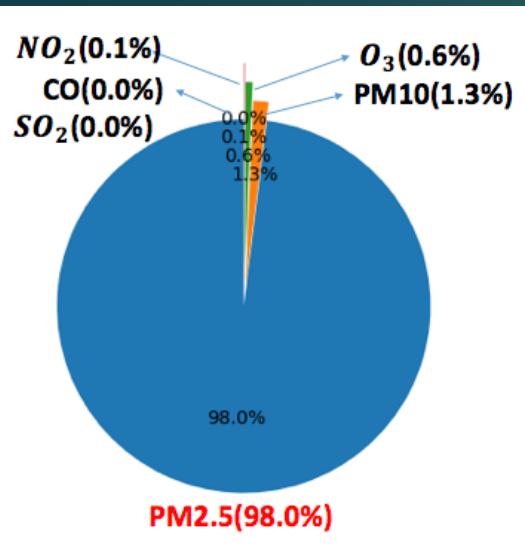
## 2. 饿了么, 京东 and other info

- ▶ 1. I don't think DATA is published and achievable, in terms of bill, for both of them.
- ▶ But for 京东, maybe we can extract what we want from reported “评价”, based on daily, since the number may have relation between number of product sold. → But do we need this info?
- ▶ Similar limitation on “美团”, “百度外卖”...
- ▶ Useful DATA website : <http://www.idataapi.cn/>



# 3. Study on SEOUL Air

Based on averaged Daily DATA



# 4. TO DO

- ▶ 1. Apply the constructed model on BEIJING Air ??
  - ▶ Does it make sense?
  - ▶ Data type correspondancy
- ▶ 2. Gather the hourly based Data → 30 times larger statistic, at least!!!
  - ▶ With higher statistic, much more accurate DNN model??
- ▶ 3. More statistics tool to try on Data analysis?