해커톤1. 레스토랑 방문자 수 예측하기

데이터셋: [Recruit Restaurant Visitor Forecasting | Kaggle](https://www.kaggle.com/c/recruit-restaurant-visitor-forecasting/data)

문제 이해하기

* a time-series forecasting problem
* 2개 사이트에서 얻은 데이터셋이다.
* hpg (Hot Pepper Gourmet): 레스토랑을 찾거나 예약할 수 있는 사이트
* air (AirReGI/Restaurant Borad): 예약 관리, 사전계산 가능한 사이트
* 훈련데이터는 2016년부터 2017년4월까지, 테스트데이터는 2016년 4월마지막주부터 5월동안의 데이터이다. (테스트데이터는 의도적으로 휴가철 데이터를 설정하였다.)
* 테스트셋에서 레스토랑이 닫거나 방문자가 없는 날은 scoring되지 않는다. 훈련데이터셋에서는 레스토랑이 닫은 날은 생략했다.

데이터셋 이해하기 (8개)

* 각 레스토랑은 air\_store\_id와 hpd\_store\_id를 가진다.

1. air\_reserve: air사이트에서 이뤄진 예약들

칼럼: air\_store\_id, visit\_datetime, reserve\_datetime, reserve\_visitors

1. hpg\_reserve: hpg사이트에서 이뤄진 예약들

칼럼: hpg\_store\_id, visit\_datetime, reserve\_datetime, reserve\_visitors

1. air\_store\_info: air사이트의 레스토랑 정보

칼럼: air\_store\_id, air\_genre\_name, air\_area\_name, latitude(위도), longitude(경도)

1. hpg\_store\_info: hpg사이트의 레스토랑 정보

칼럼: hpg\_store\_id, hpg\_genre\_name, hpg\_area\_name, latitude, longitude

1. store\_id\_relation

칼럼: hpg\_store\_id, air\_store\_id

1. air\_visit\_data

칼럼: air\_store\_id, visit\_date, visitors

1. sample\_submission: 우리가 예측해야하는 정답의 포맷을 보여줌.

칼럼: id, visitors

1. date\_info

칼럼: calendar\_date, day\_of\_week, holiday\_fig(일본의 휴일)