DigitalHouse >
Coding School

¿Comemos afuera?

Predicción de visitas a restaurants

Gastón Domínguez

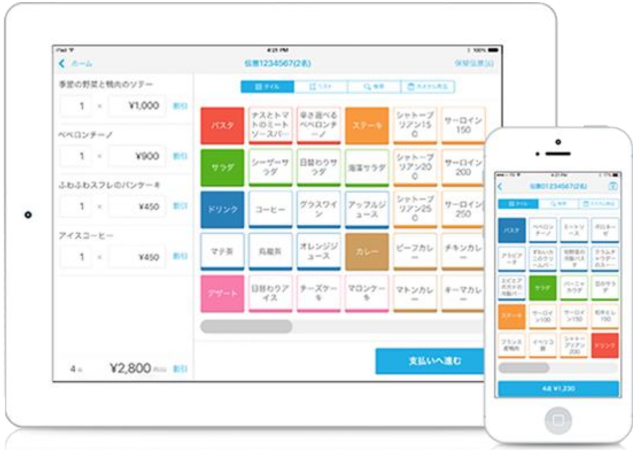
Data Science

Objetivo

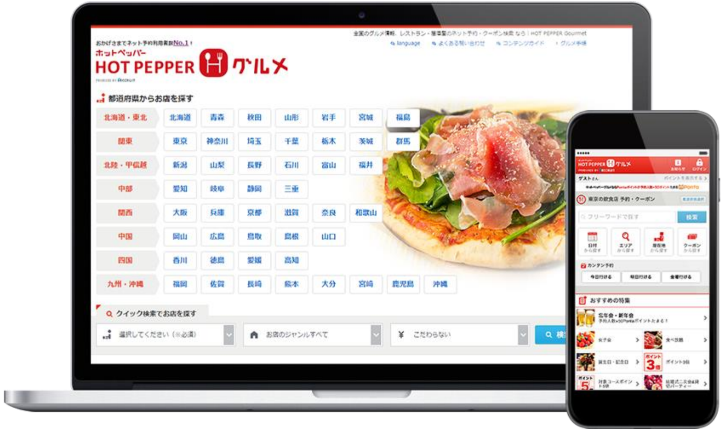
- Predicción de visitas a restaurant
- Caracterización como Serie de Tiempo
- Aplicación de modelos idóneos
- Incorporación de información complementaria

GENERALIDADES

Información disponible



RESERVAS Y VISITANTES



RESERVAS

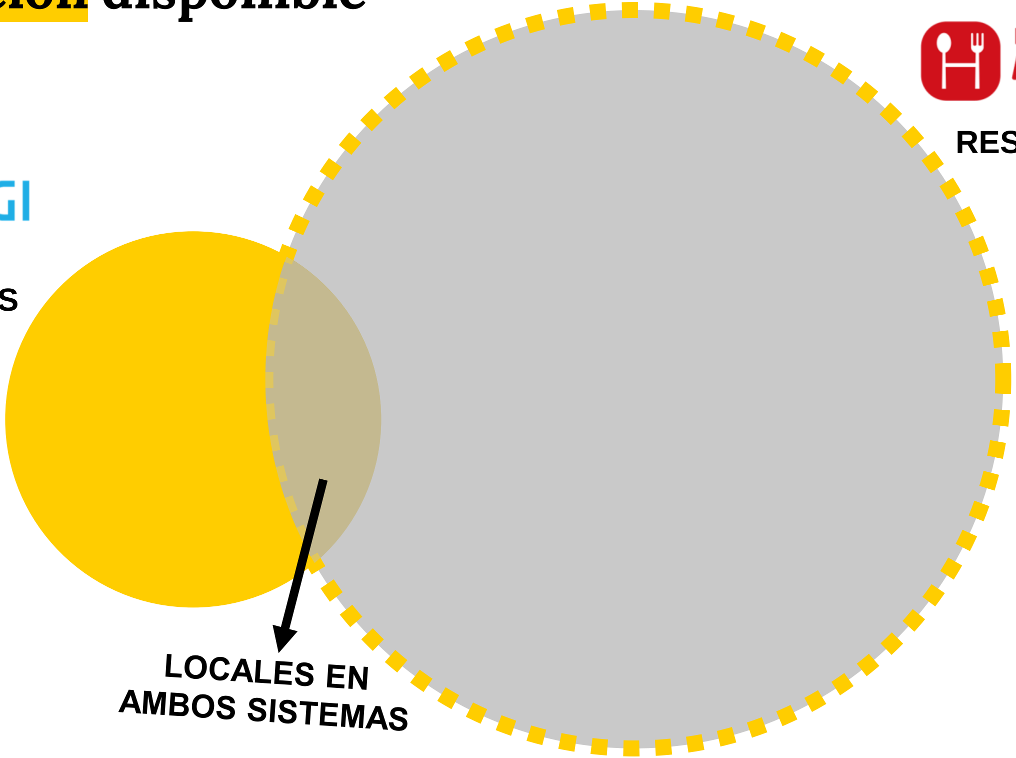
Información disponible

HOT PEPPER グルメ

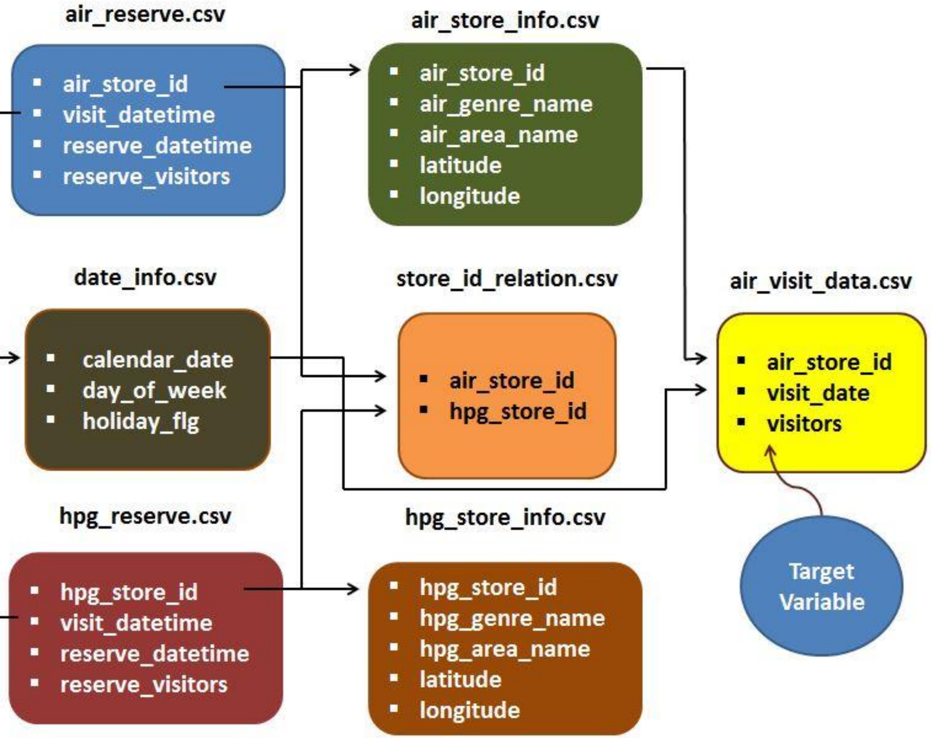
RESERVAS

AirREGI

RESERVAS
Y VISITANTES

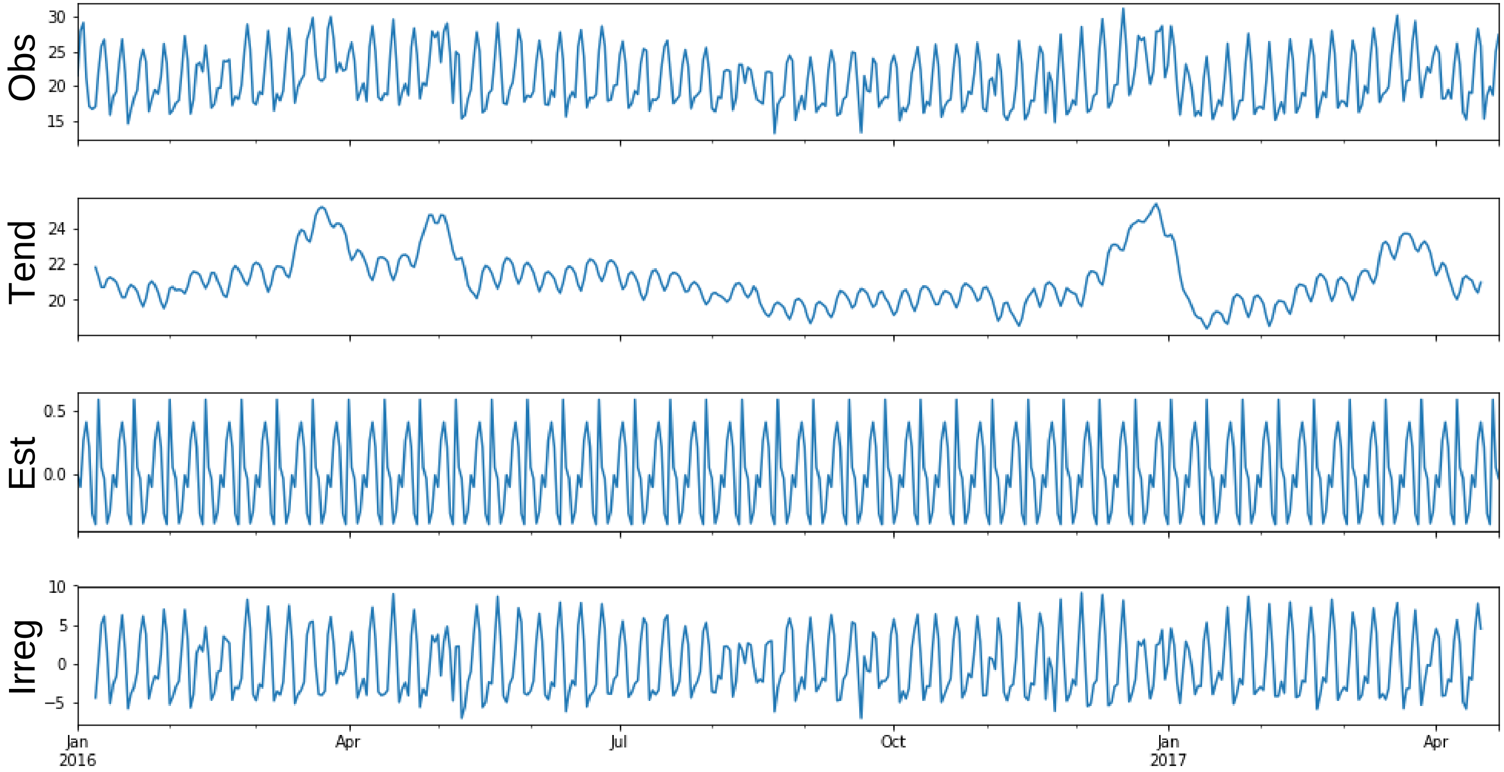


Información disponible



Series de Tiempo

seasonal_decompose(air.visit_mean, freq=12)

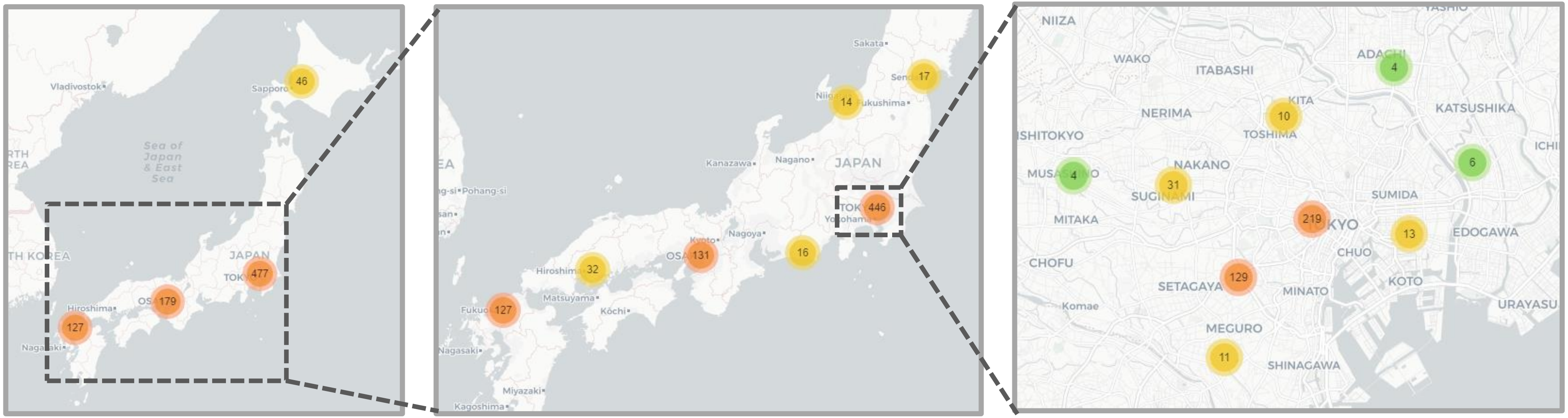


Series de Tiempo

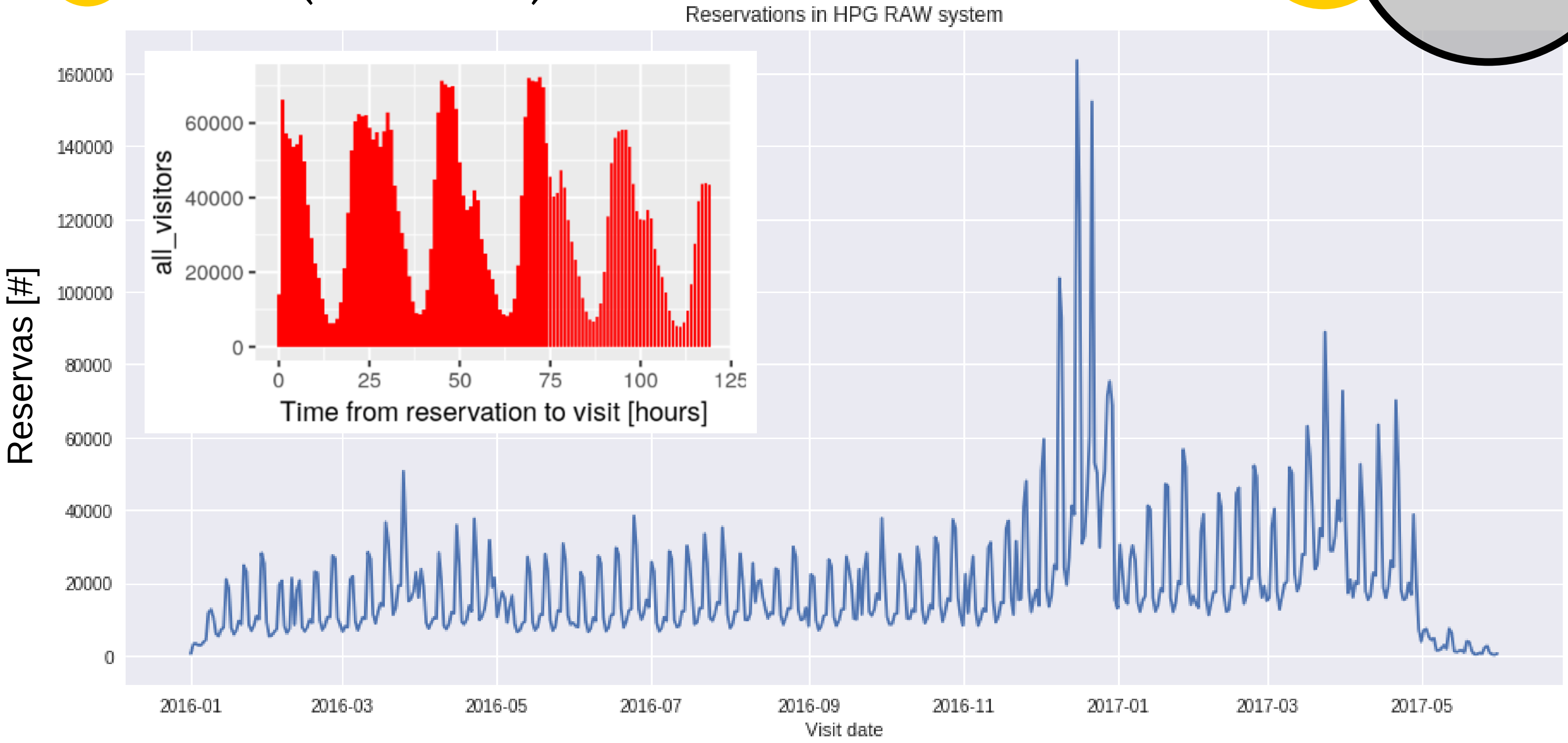
Tendencia	• Incremento/disminución a largo plazo
Estacional	• Serie influenciada por patrones estacionales
Cíclico	• Aumento/Caída no propios al período fijo (>2años)
Irregular	• Variaciones aleatorias o residuales

ANÁLISIS EXPLORATORIO

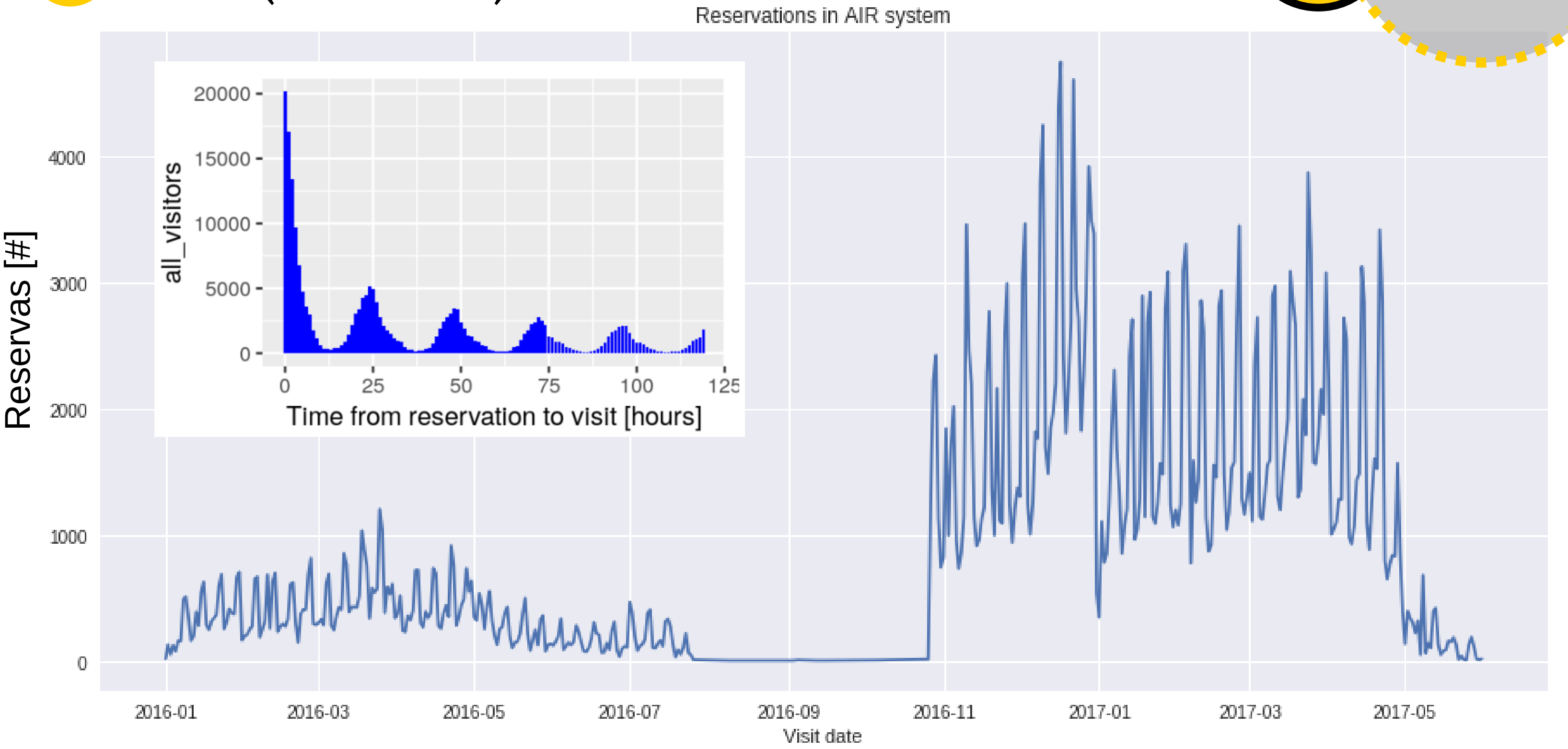
Ubicación de locales



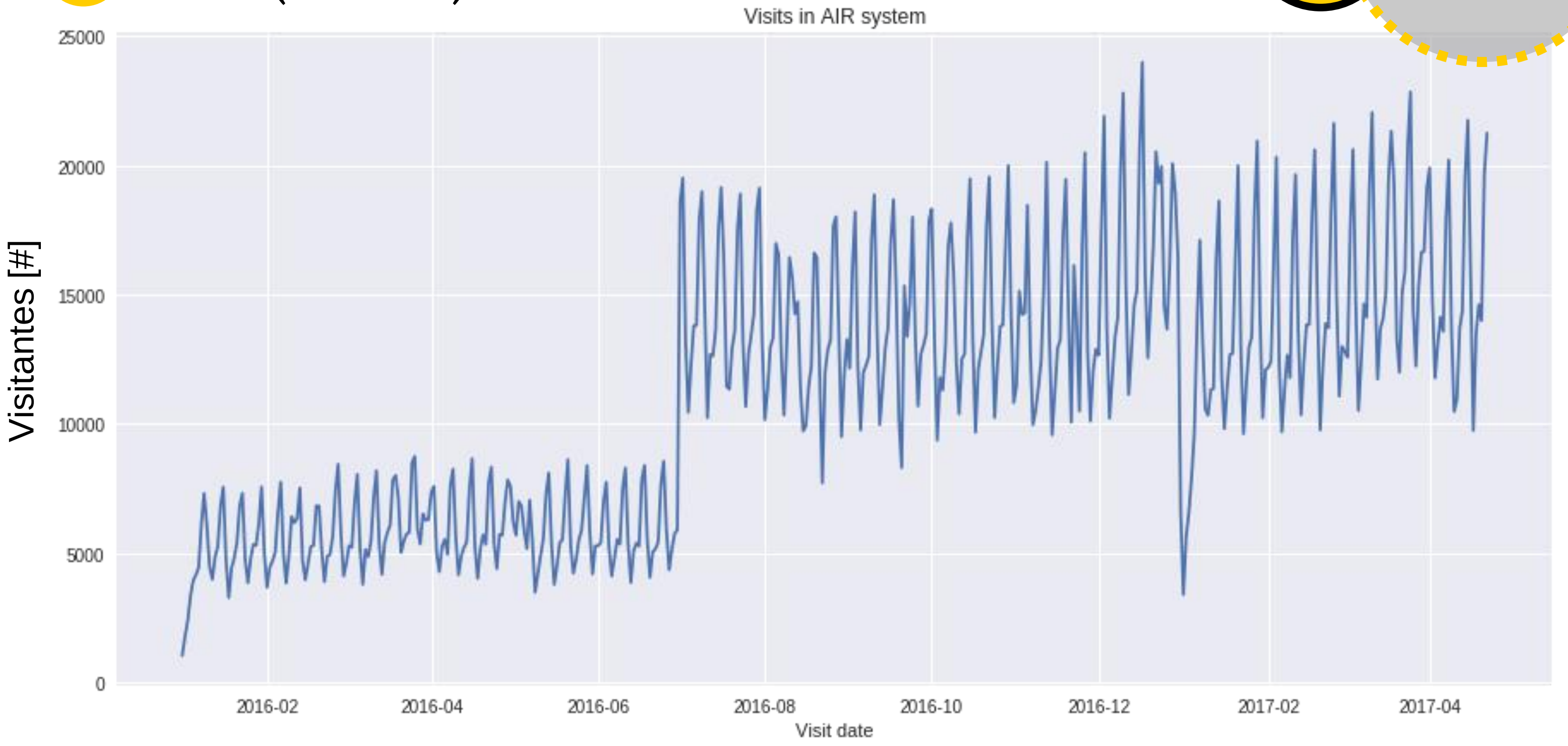
HPG (Reservas)



AIR (Reservas)

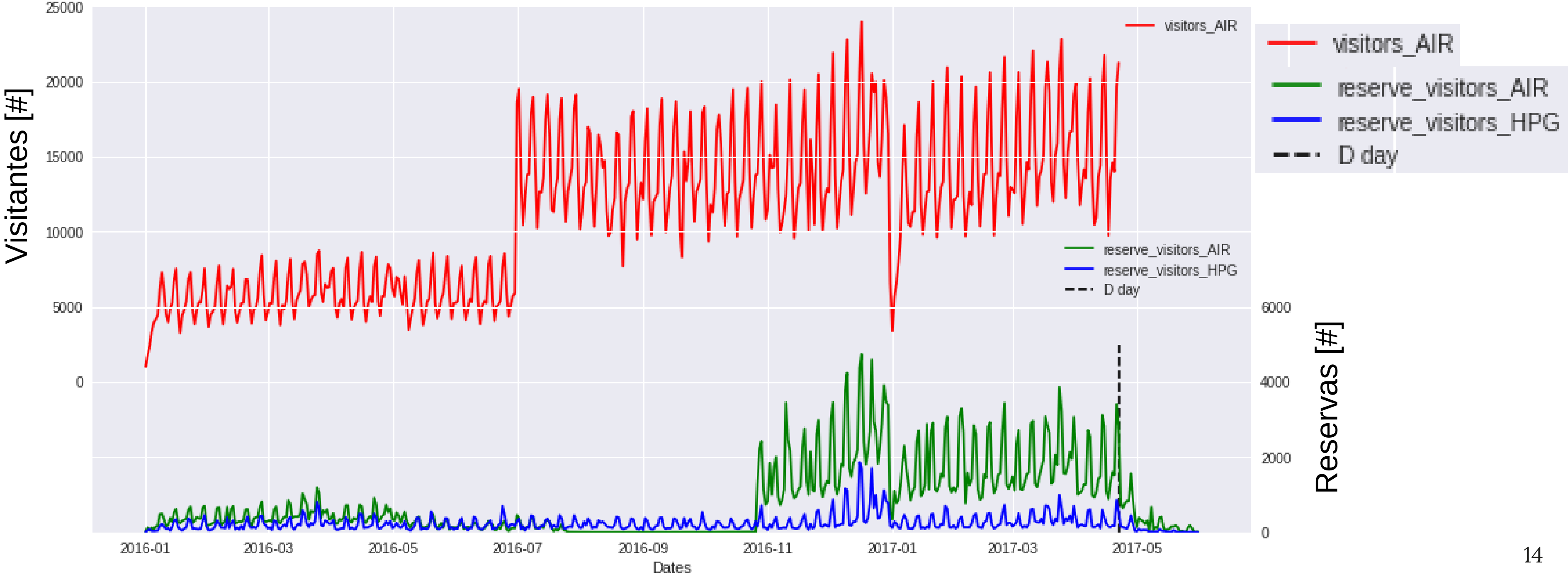


AIR (Visitas)

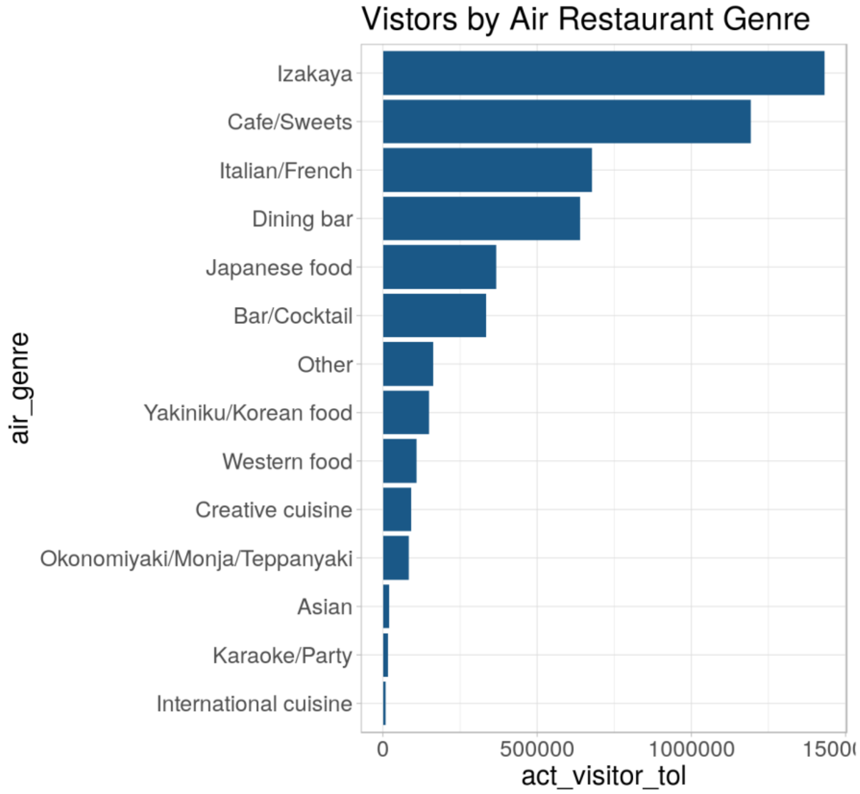

Visits in AIR system
Visitantes [#]
25000
20000
15000
10000
5000
0
2016-02
2016-04
2016-06
2016-08
2016-10
2016-12
2017-02
2017-04
Visit date

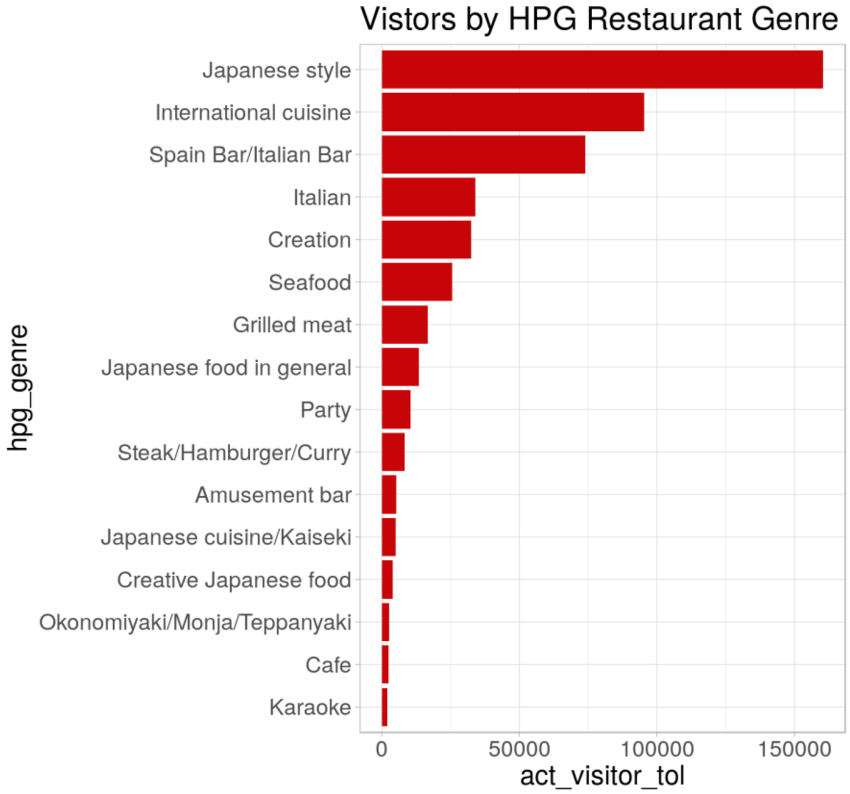
AIR & HPG (Reservas y Visitas)

Visitors against reservations in AIR & HPG systems

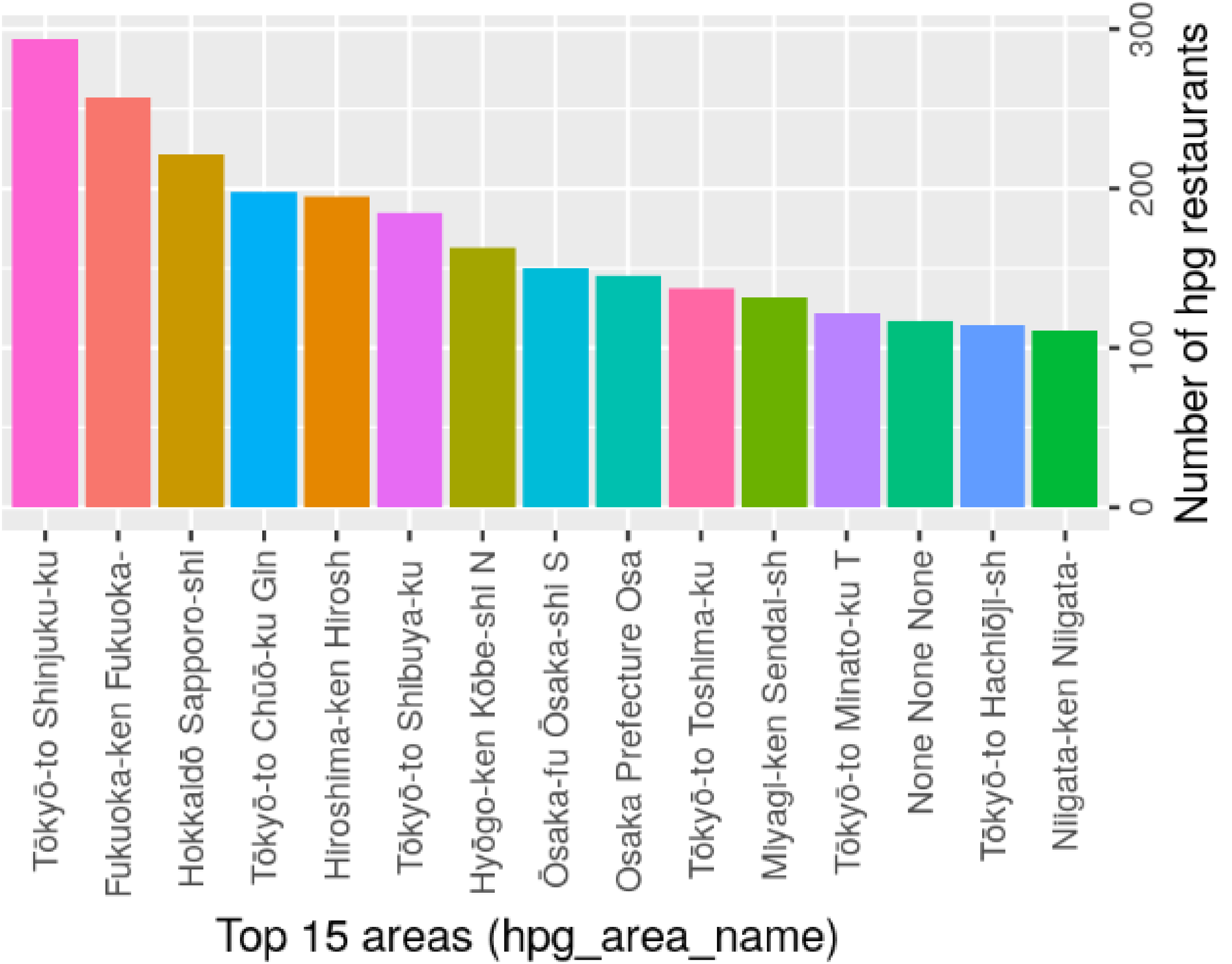


Tipo de cocina



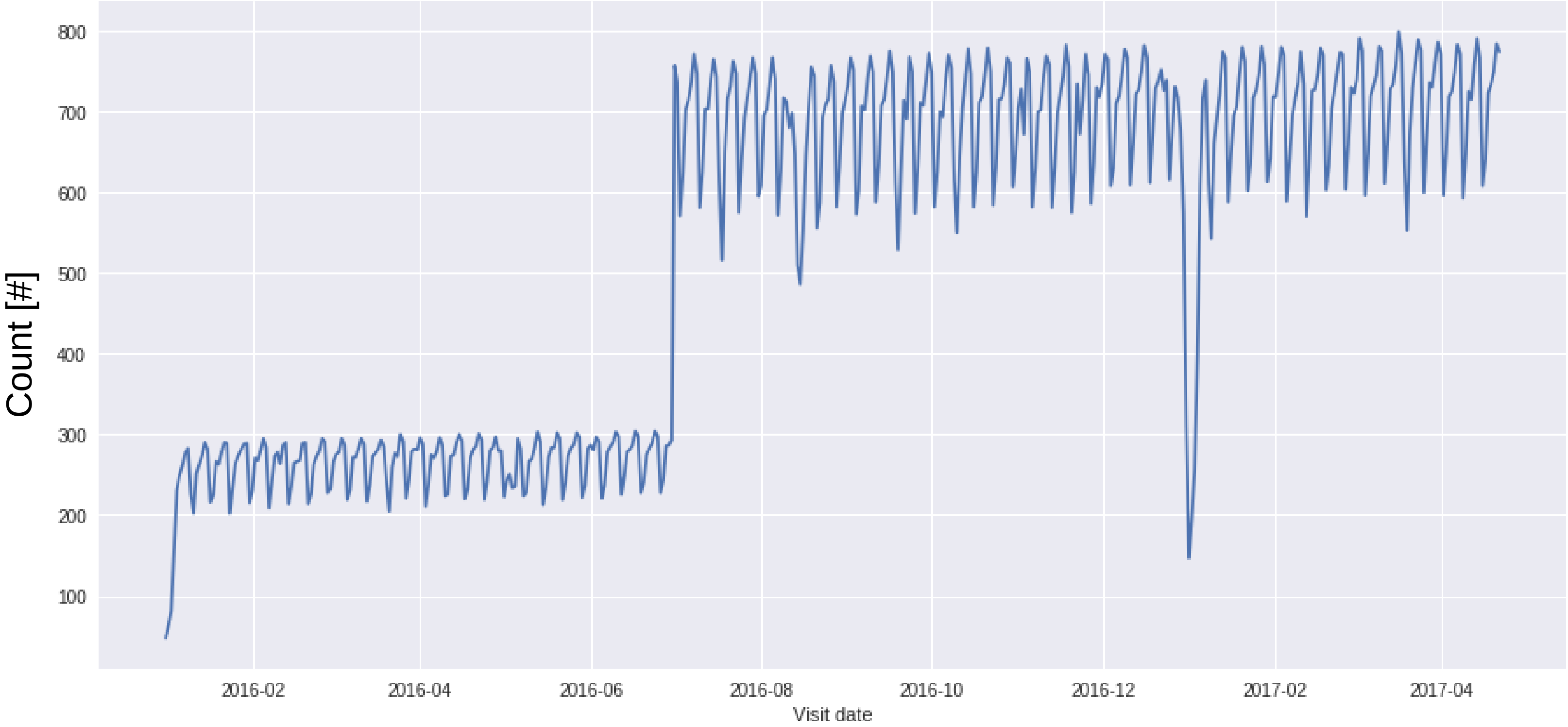


Top Areas (HPG)

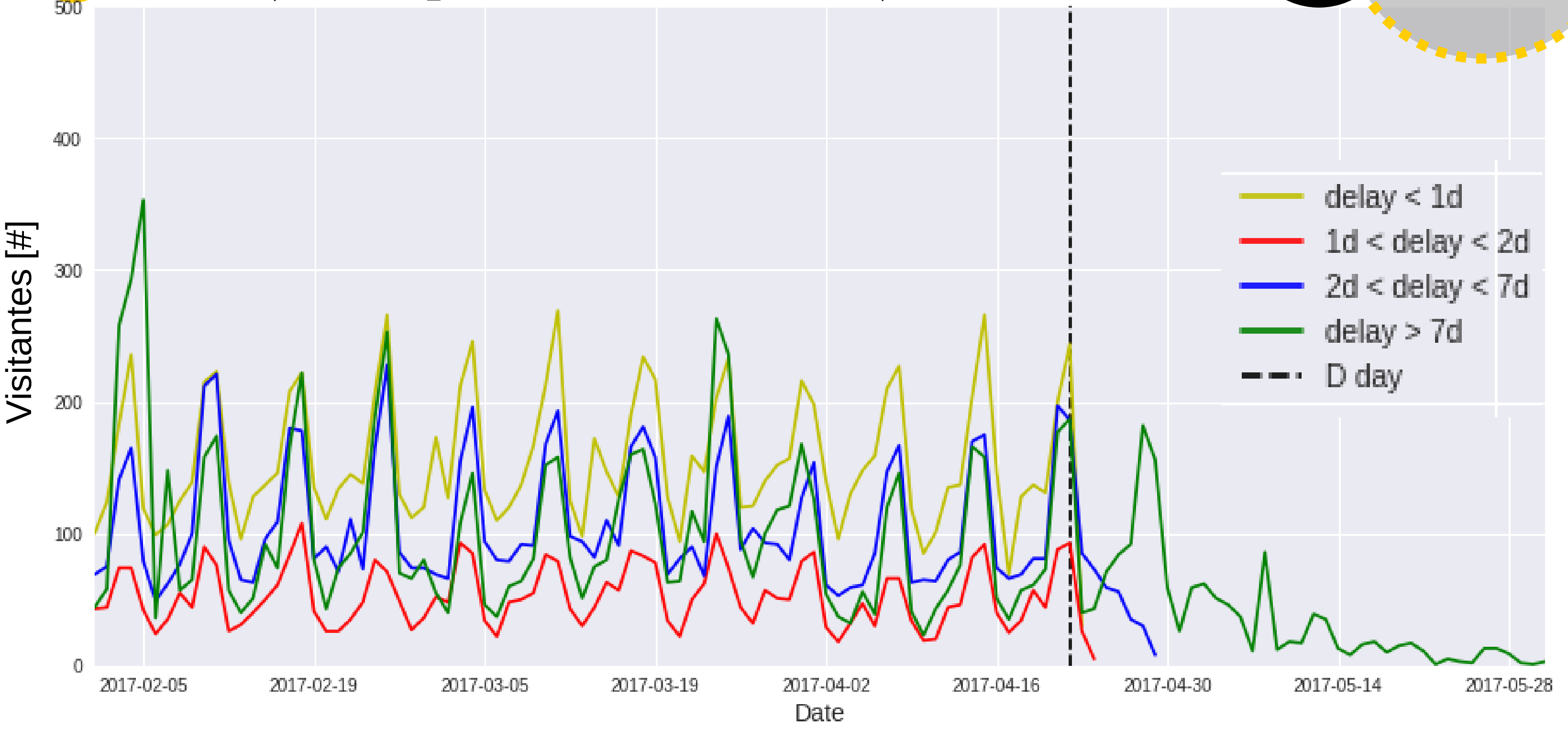


INFORMACIÓN COMPLEMENTARIA

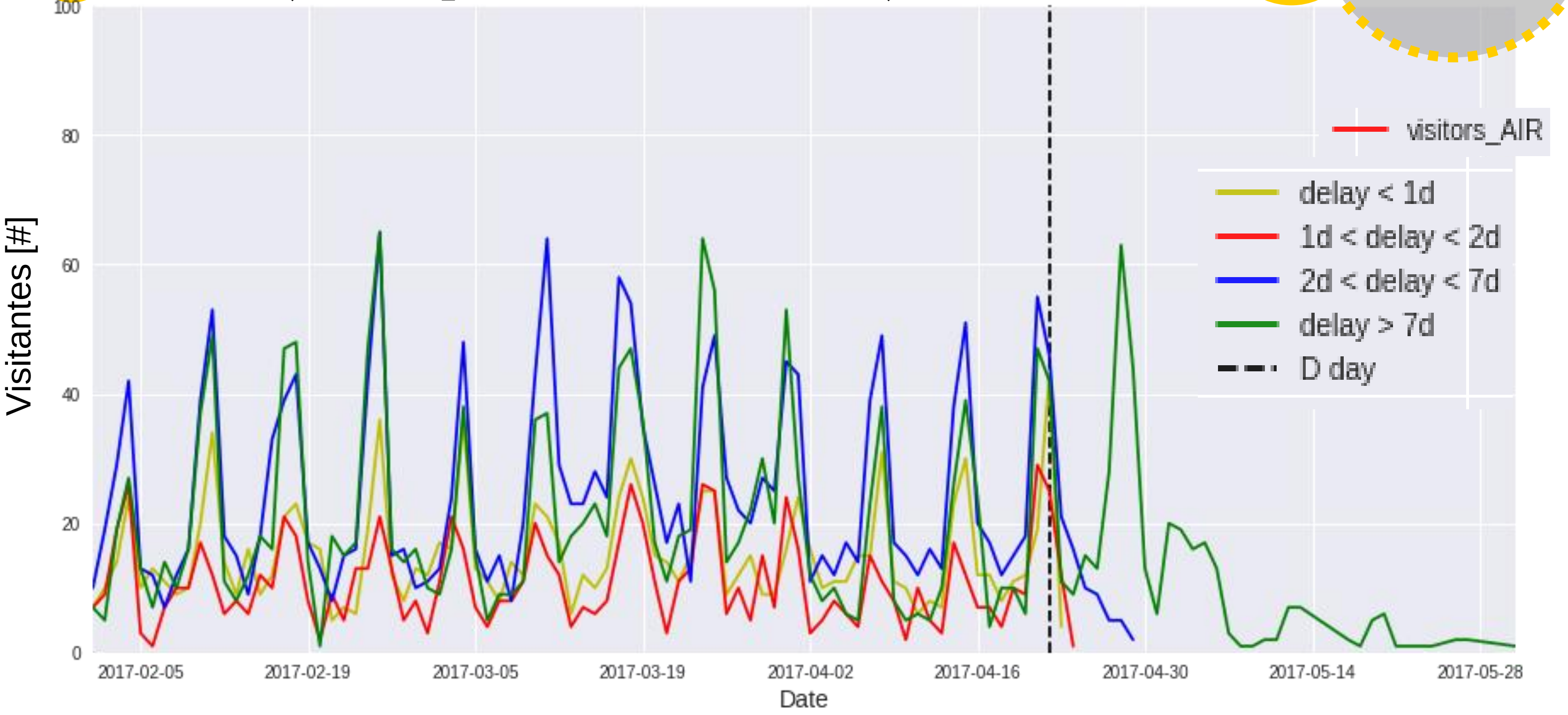
Total de Locales



AIR (Anticipación de reservas)

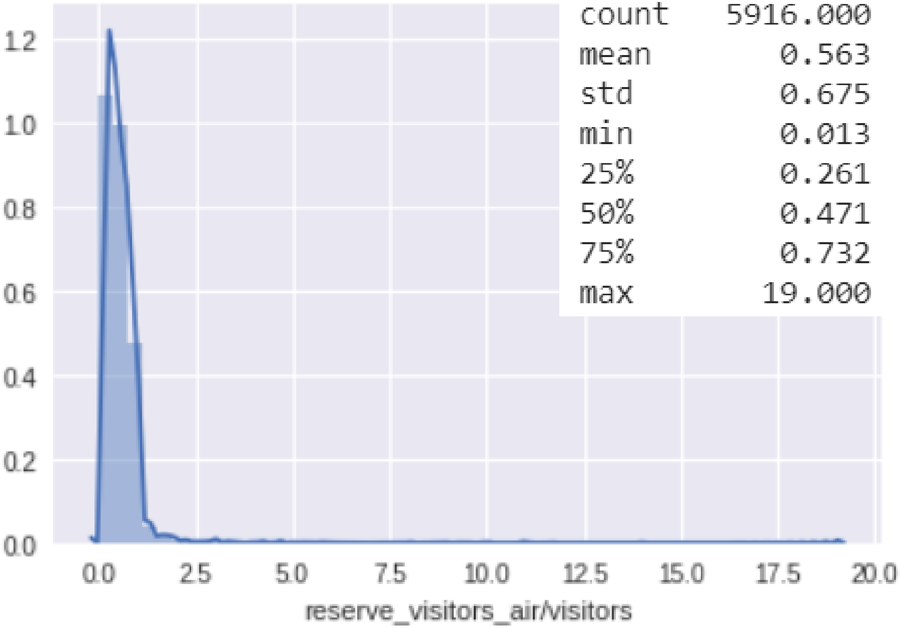


HPG (Anticipación de reservas)

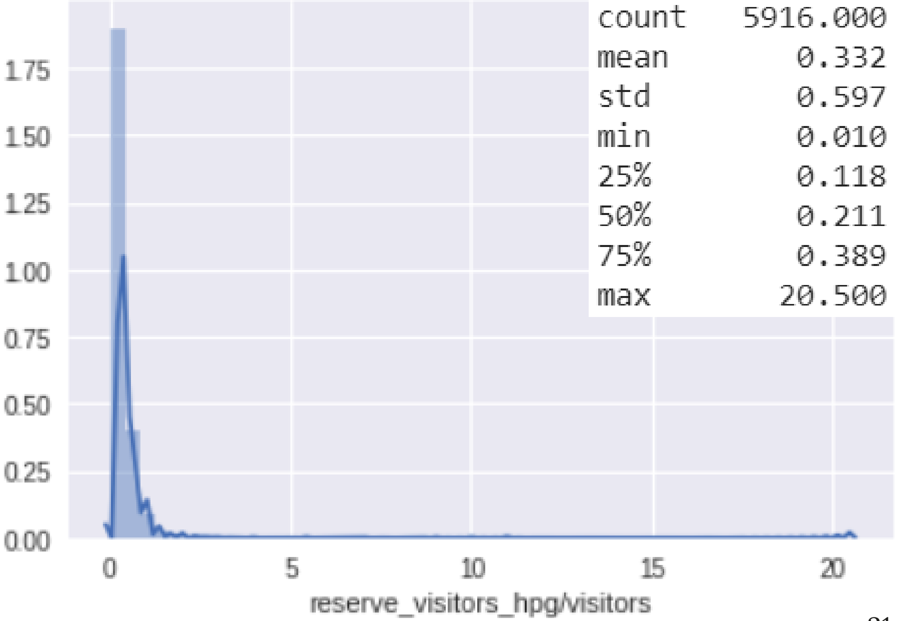


Overbooking

AIR

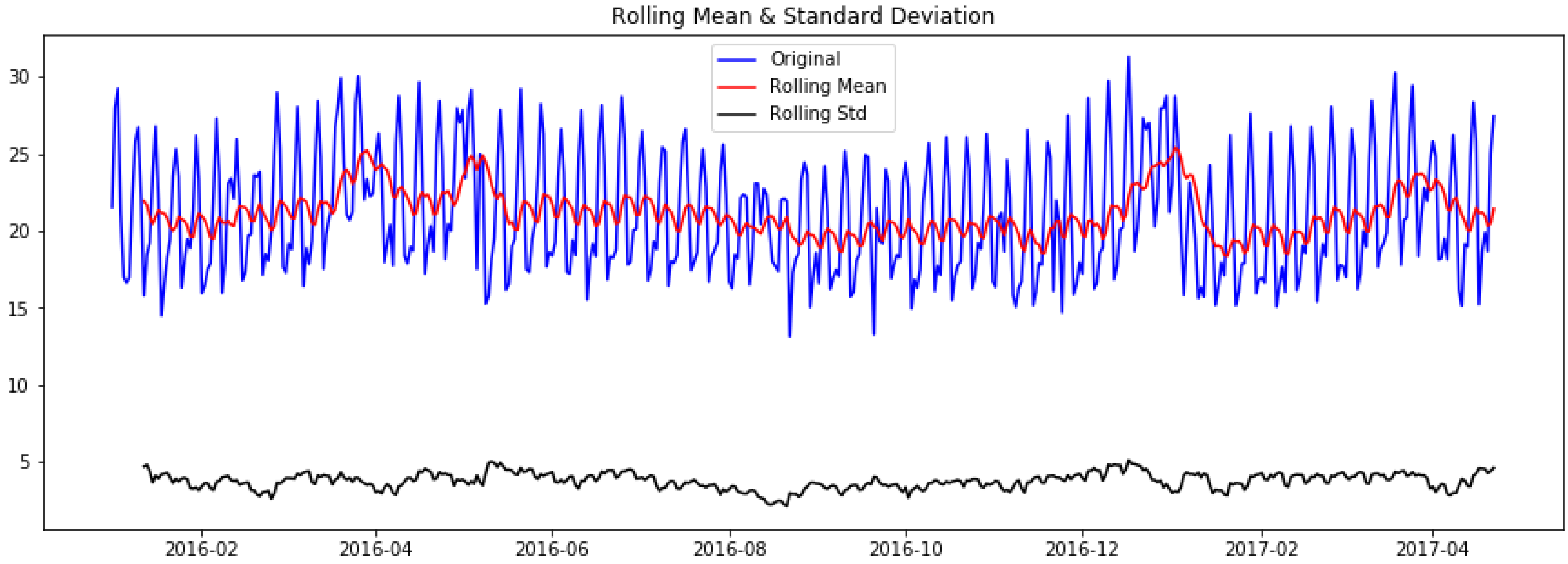


HPG



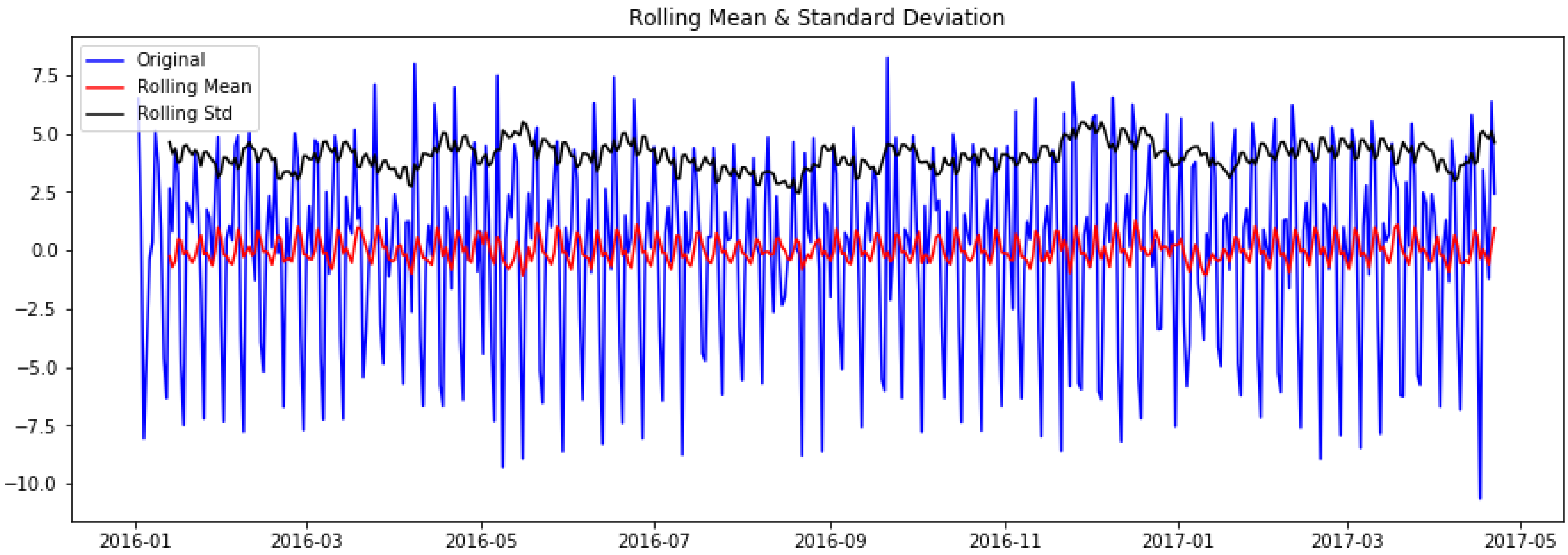
test_stationarity(air.**visit_mean**)

Estacionariedad



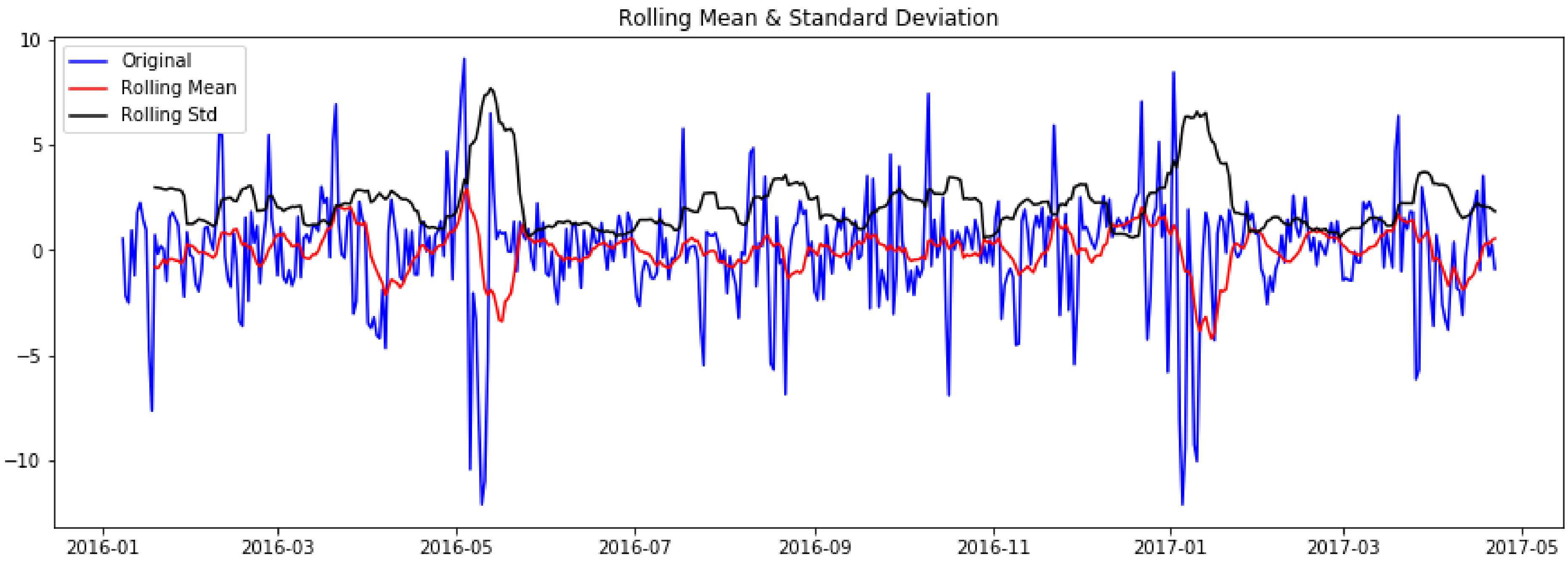
Estacionariedad

```
air['visit_mean_diff'] = air.visit_mean - air.visit_mean.shift(1)
test_stationarity(air.visit_mean_diff.dropna(inplace=False))
```



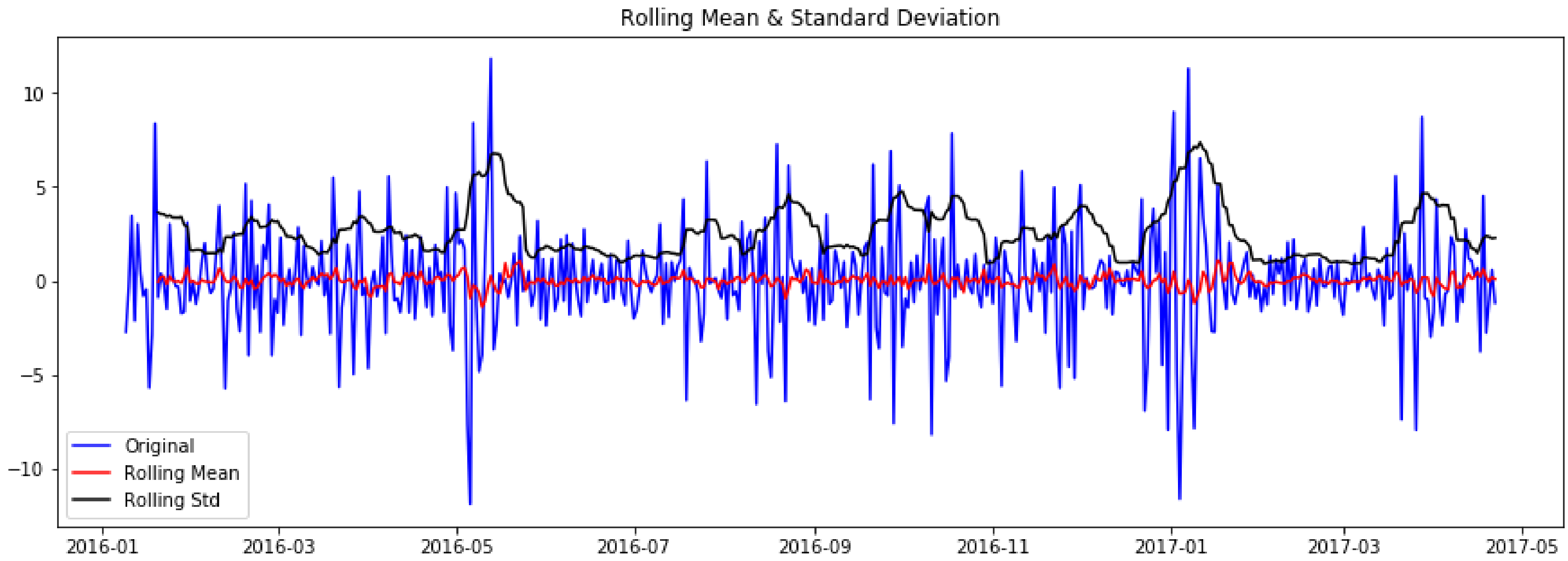
air['**visit_mean_seasonal**'] = air.visit_mean - air.visit_mean.shift(**7**)
test_stationarity(air.visit_mean_seasonal.dropna(inplace=False))

Estacionariedad



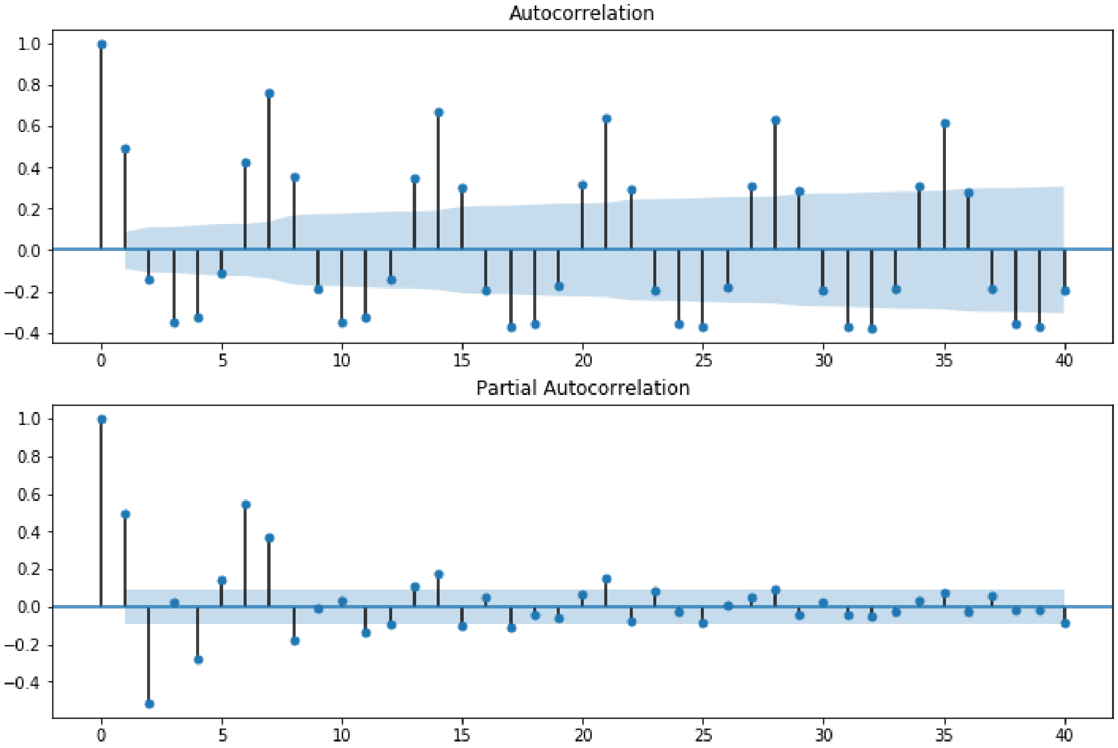
Estacionariedad

```
air['visit_mean_seasonal_diff'] = air.visit_mean_diff - air.visit_mean_diff.shift(7)
test_stationarity(air.visit_mean_seasonal_diff.dropna(inplace=False))
```



```
fig1 = sm.graphics.tsa.plot_acf(air.visit_mean, lags=40, alpha=.05, ax=ax1)
fig2 = sm.graphics.tsa.plot_pacf(air.visit_mean, lags=40, alpha=.05, ax=ax2)
```

Estacionariedad



Features adicionales

GENERALES

- Día / Mes / Año
- Feriados
- Cant. Restaurants (x zona)
- Total de locales
- Total reservas (HPG / AIR)
- Clima

PARTICULARES

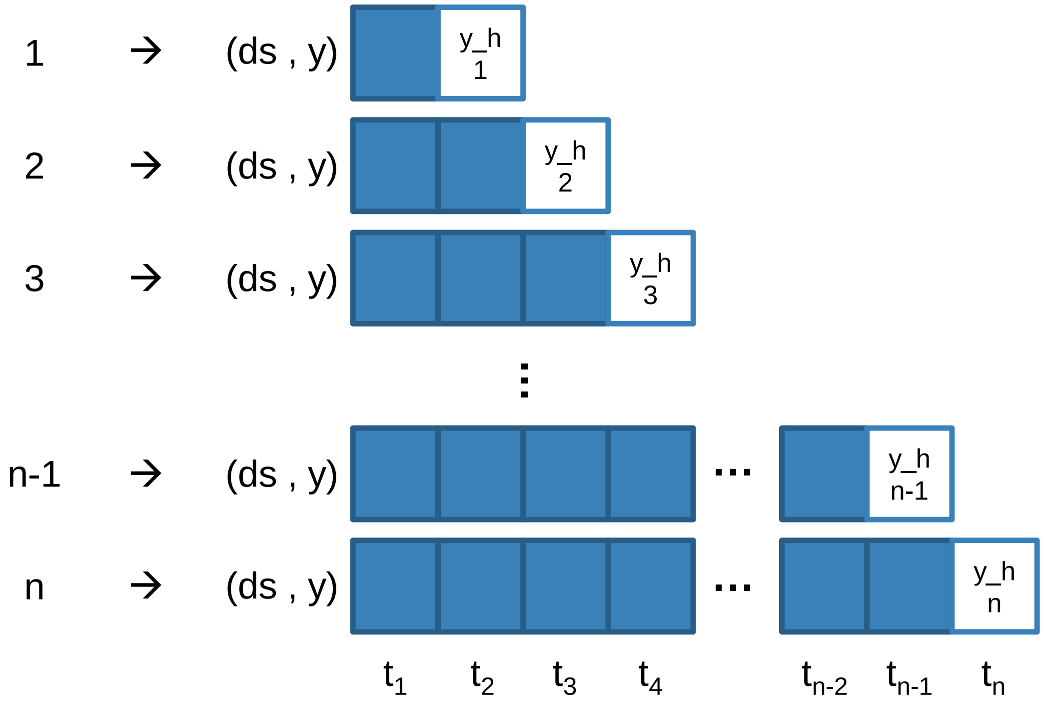
- Tipo de comida
- Ciudad / Área
- Latitud / Longitud
- Anticipación reservas
  1, 2, 7 o más días

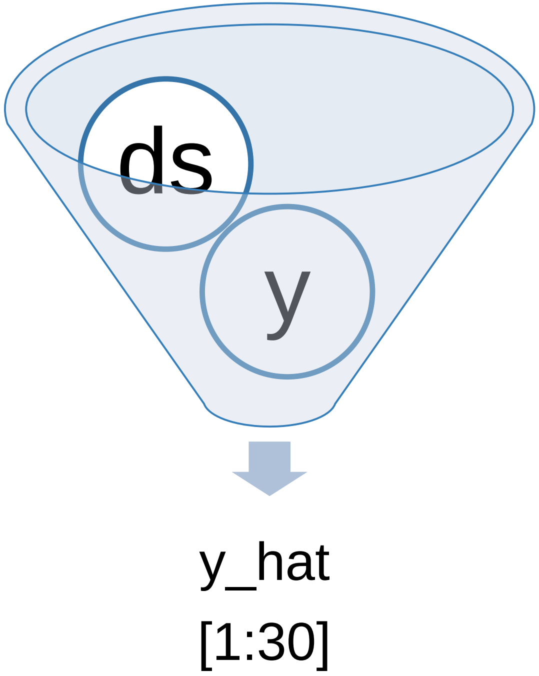
AUTO Regresor

- Predicción de visitas en el corto plazo (1~30 días)

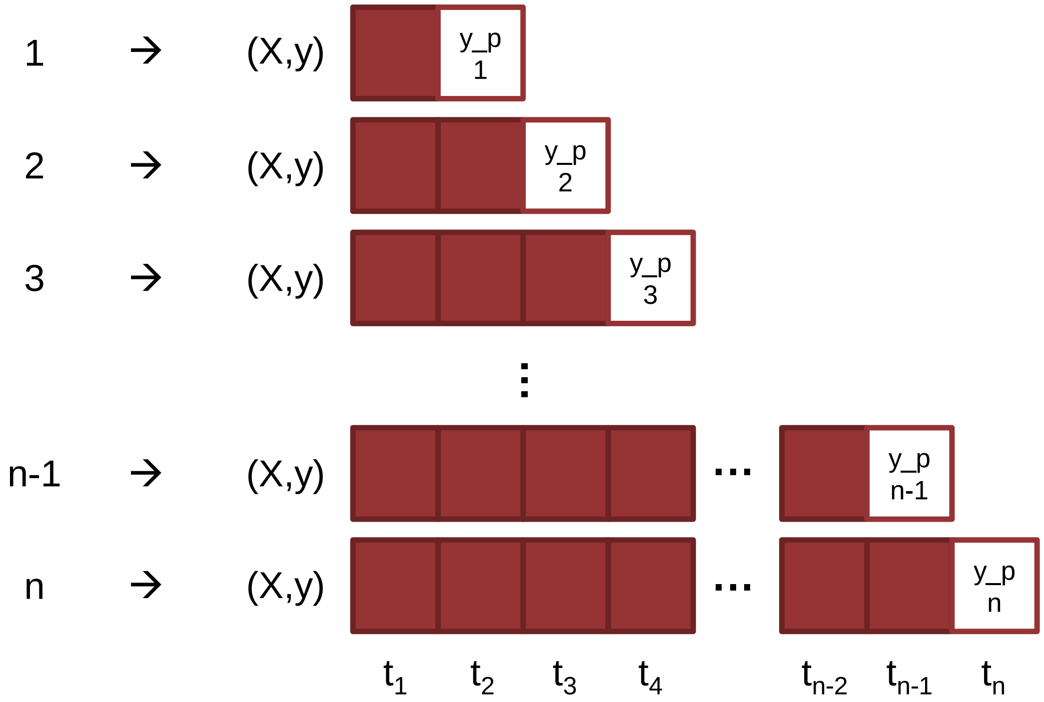
METODOLOGÍA

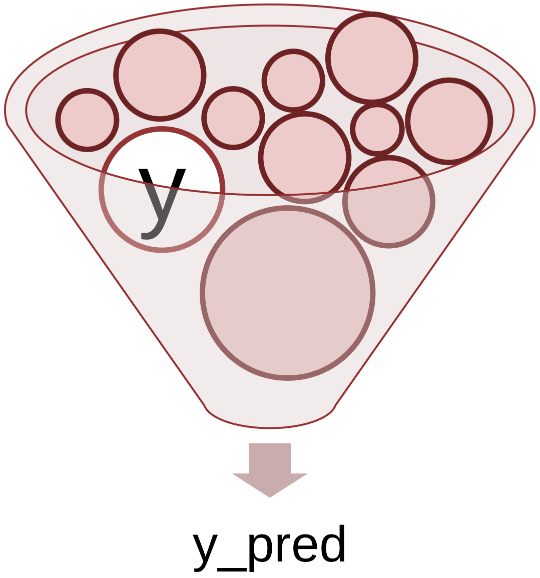
Prophet (facebook)



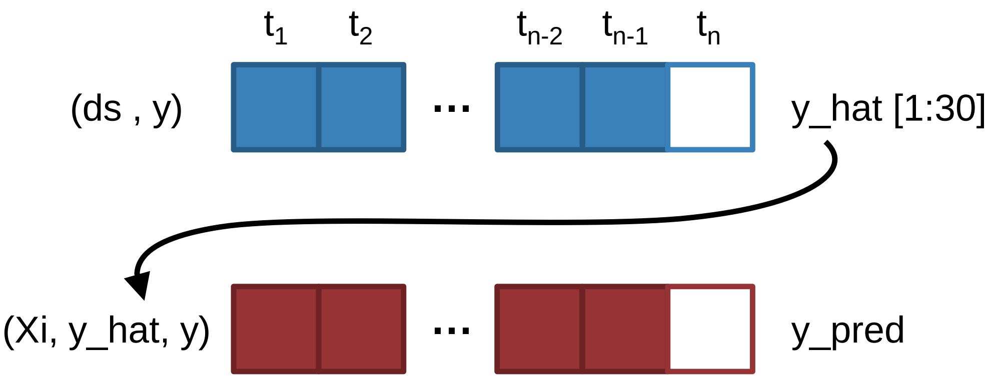


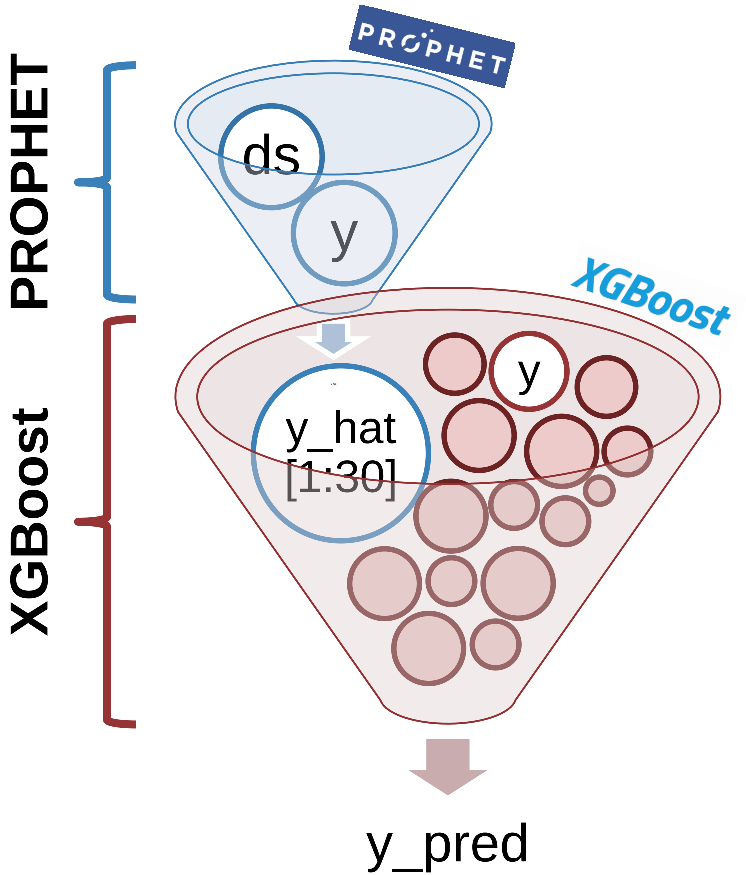
XGBoost
(eXtreme Gradient Boosting)



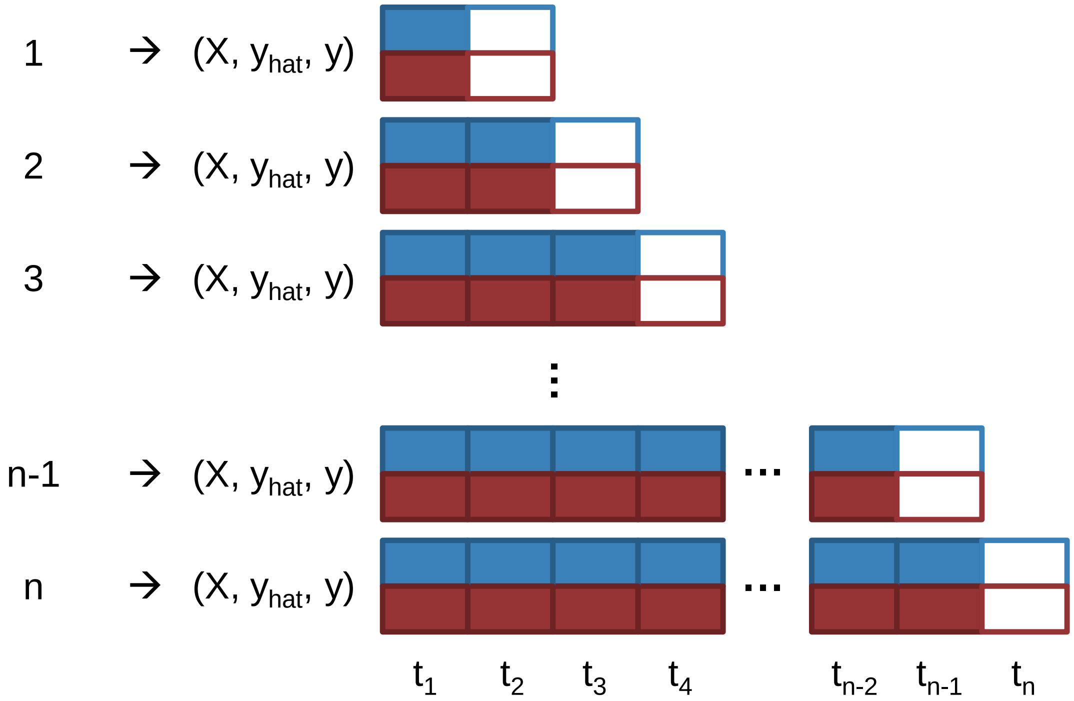


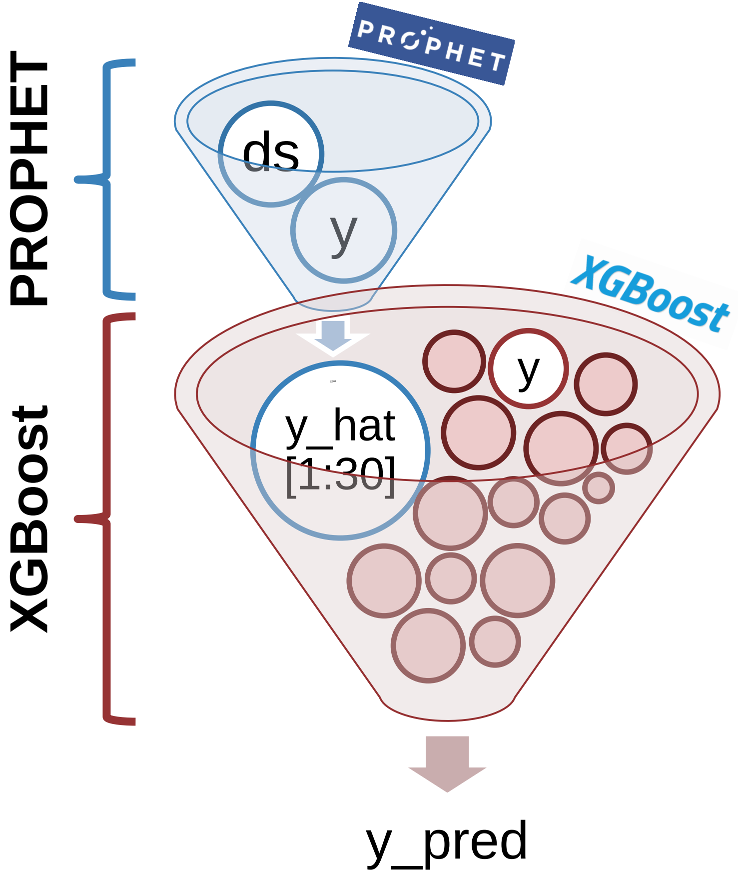
Prophet + XGBoost



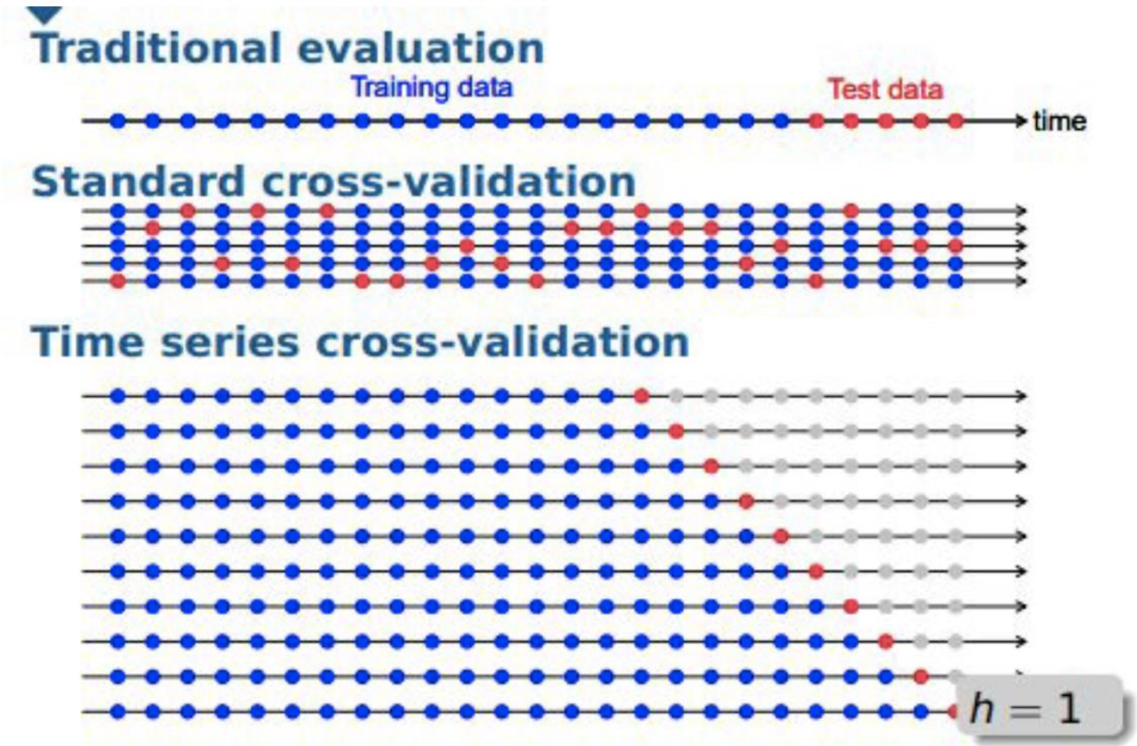


Prophet + XGBoost

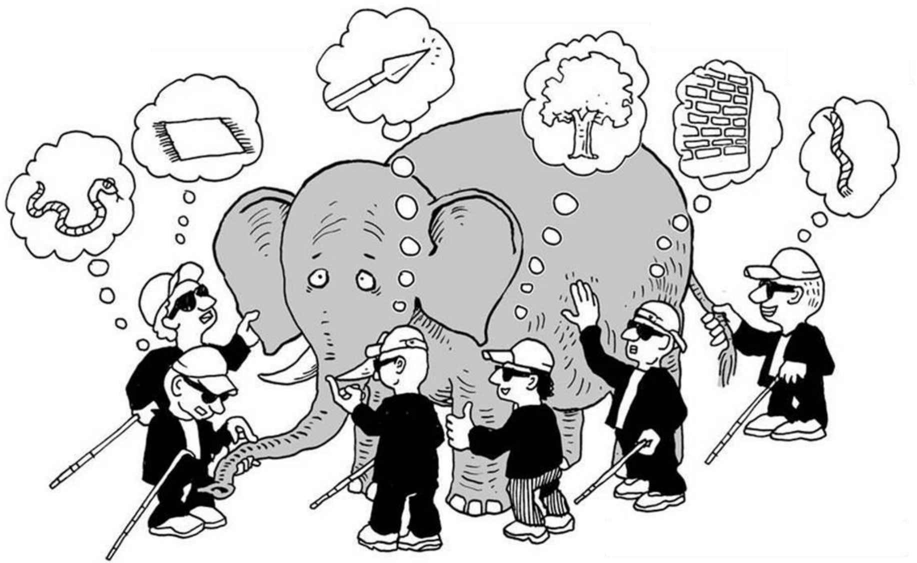


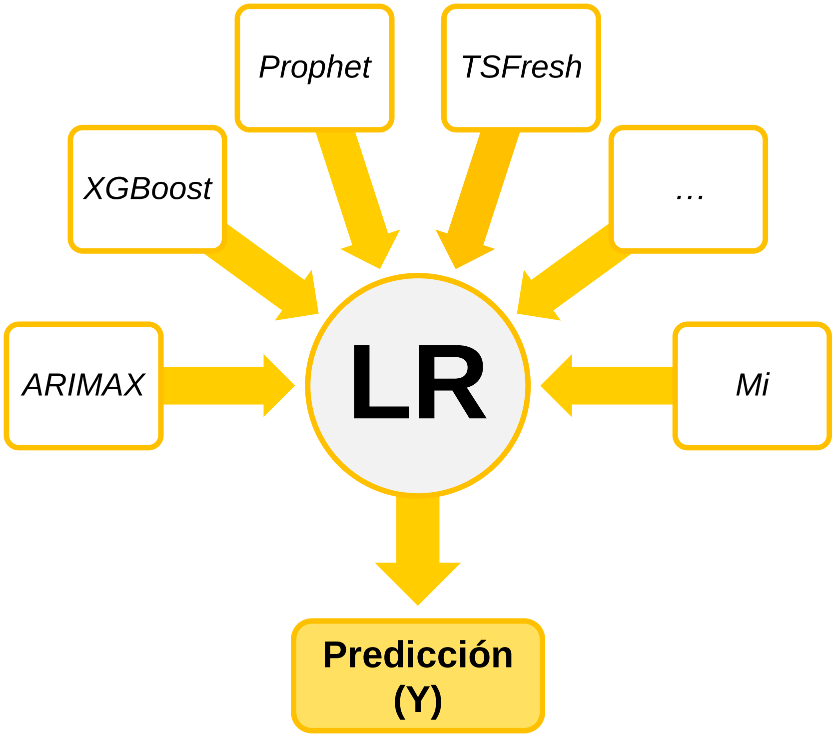


Validación Cruzada



Otro enfoque

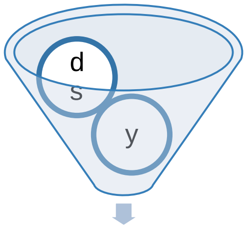


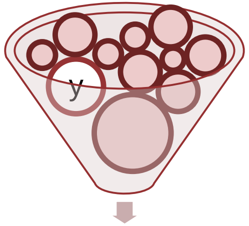


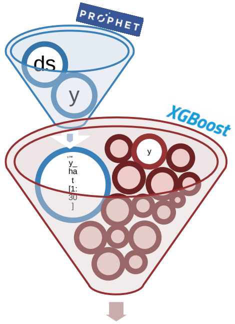
RESULTADOS

(Score = RMSE)

Resultados







LOCAL	vis./test	vis./día	Prophet	XGBoost	P+XGB
air_e55abd740f93ecc4	*1195*	*40*	18.48	9.47	**9.25**
air_681f96e6a6595f82	*1184*	*39*	16.14	14.47	**13.82**
air_399904bdb7685ca0	*1129*	*38*	**15.38**	15.93	17.29
air_03963426c9312048	*1038*	*35*	26.21	**10.30**	11.56
air_4cca5666eaf5c709	*815*	*27*	6.47	7.35	**6.40**
air_506fe758114df773	*573*	*19*	10.18	9.71	**9.51**
air_c6ffd6a93e6b68d6	*377*	*12*	9.86	5.80	**4.68**
air_800c02226e2e0288	*231*	*8*	4.16	4.04	**4.02**

CONCLUSIONES

Conclusiones

- Múltiples enfoques posibles (modelos)
- Importancia del análisis temporal de los datos (TS)
- Features adicionales propios de la Serie de Tiempo
- Información útil a incluir:
  - "capacidad máxima del local"
  - "reserva concretada o no" (Flag_reserva_caida)
- A partir de eso, se podría estimar:
  - Qué reservas se podrían caer
  - Determinar hasta cuánto hacer overbooking

Aplicaciones múltiples

- Reservas turísticas
  - Hotelería
  - airbnb
  - Excursiones
- Pasajes aéreos
- Turnos médicos
- Espectáculos
- Otros

Muchas Gracias!

Preguntas ?