

# TOP Y NOT DE LAS AERONAVES DEL MUNDO: UN MODELO DE PREDICCIÓN CON BIBLIOTECA "RANDOM FOREST"

Por: Juan Carlos Cárdenas S.



WEIGHT

ORIG. TRIP

SEOUL

SOUTH KOREA

FLIGHT NO.

DATE

90-46-28

SEOUL

# OBJETIVO DEL MODELO

- Predecir qué modelo de aeronave es más probable que se utilice en una ruta aleatoria.
- **BONUS:**
- Conocer ubicaciones de aeropuertos en México, Colombia y Francia.

# NUMERALIA

- 6,072 aerolíneas,
- 67,664 rutas entre
- 3,321 aeropuertos operadas por
- 548 aerolíneas
- 3 bases de datos
- 1 fuente: Open Flights



# CÓDIGO IATA, ¿QUÉ ES?

- Por sus siglas en inglés: *International Air Transport Association*.
- Se utiliza para identificar aeropuertos en el mundo.
- Consta de 3 caracteres alfabéticos.
- Ejemplos: NLU (Aeropuerto Santa Lucía),
- TIJ (Aeropuerto Tijuana).



# PROGRAMAS USADOS



Python



Tableau

# MÉTODO ELEGIDO: XGBOOST

- Permite manejar bases de datos con muestreos del orden de los miles de millones.
- Funciona para tratar valores numéricos y categóricos.
- Usando "X" y "y" numéricos obtuvimos:
  - R2 entrenado obtenida: .8274
  - R2 prueba obtenida: .5835
- Tamaño de la prueba: 30% de la data.





# CONCLUSIÓN

EL MODELO PRESENTA APROXIMACIÓN QUE NOS PERMITE  
ACERTAR EN UN 52% DE LAS VECES.

TRAIN SCORE: 0.5494

TEST SCORE: 0.5835



! GRACIAS POR SU  
ATENCIÓN!

