

TRAVEL DATASET

OBJETIVO:

- Con este modelo se busca predecir la satisfacción de pasajeros de líneas aéreas



ABSTRACT:

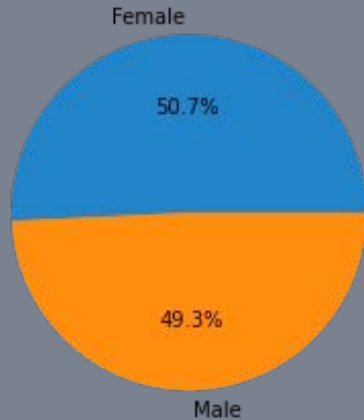
- El dataset se trata de encuestas a pasajeros de aerolíneas, sus respuestas van de 1 a 5 según satisfacción. Cada fila del conjunto son los datos de respuesta de cada pasajero y según sus respuestas a servicios o características vamos a saber si está satisfecho o no

TRAVEL DATASET

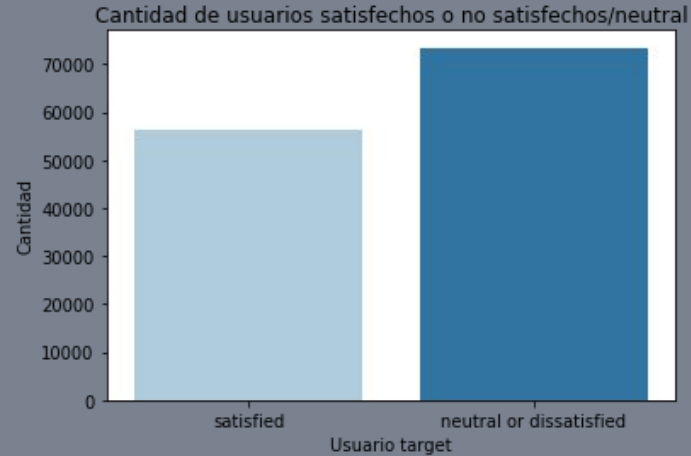


EXPLORACIÓN DE DATOS:

El conjunto contiene 129.880 filas y 25 columnas

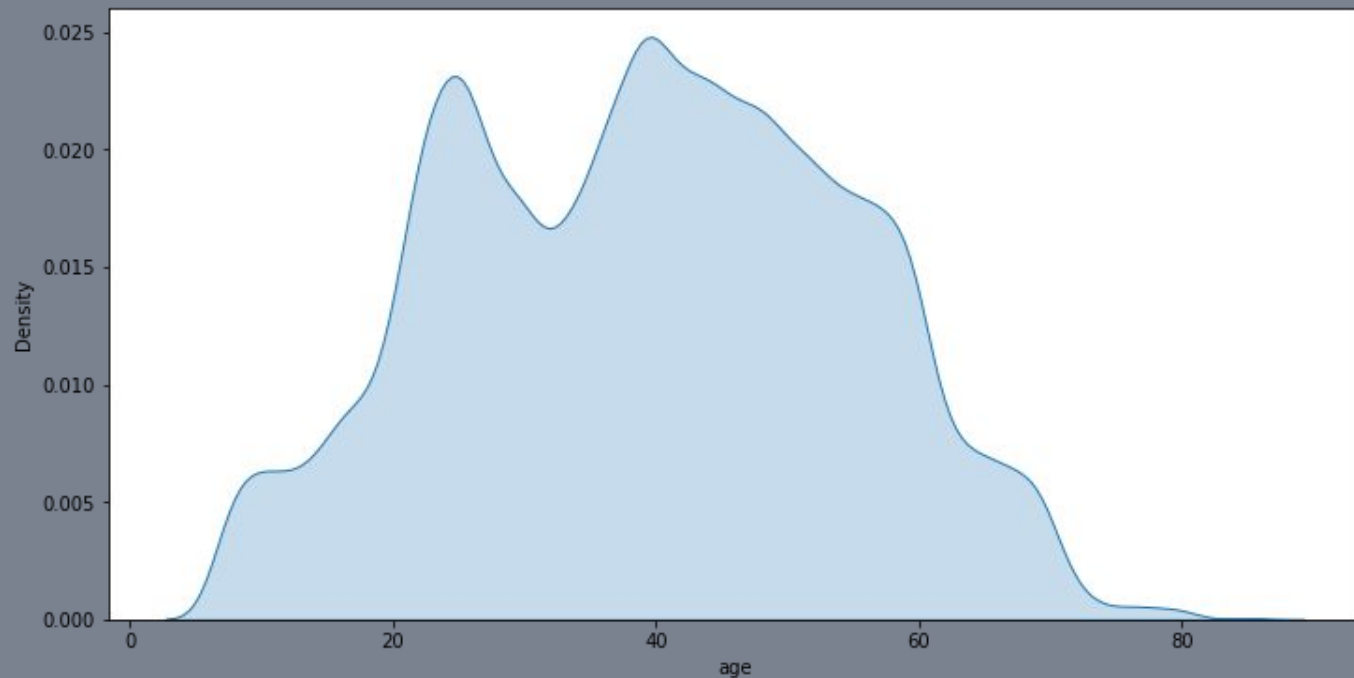


dentro del dataset vimos que el nivel de satisfacción es levemente superior en el género femenino



Notamos que predominan los usuarios insatisfechos/neutral, además se observa una diferencia superior a 10.000 pasajeros a simple vista

TRAVEL DATASET

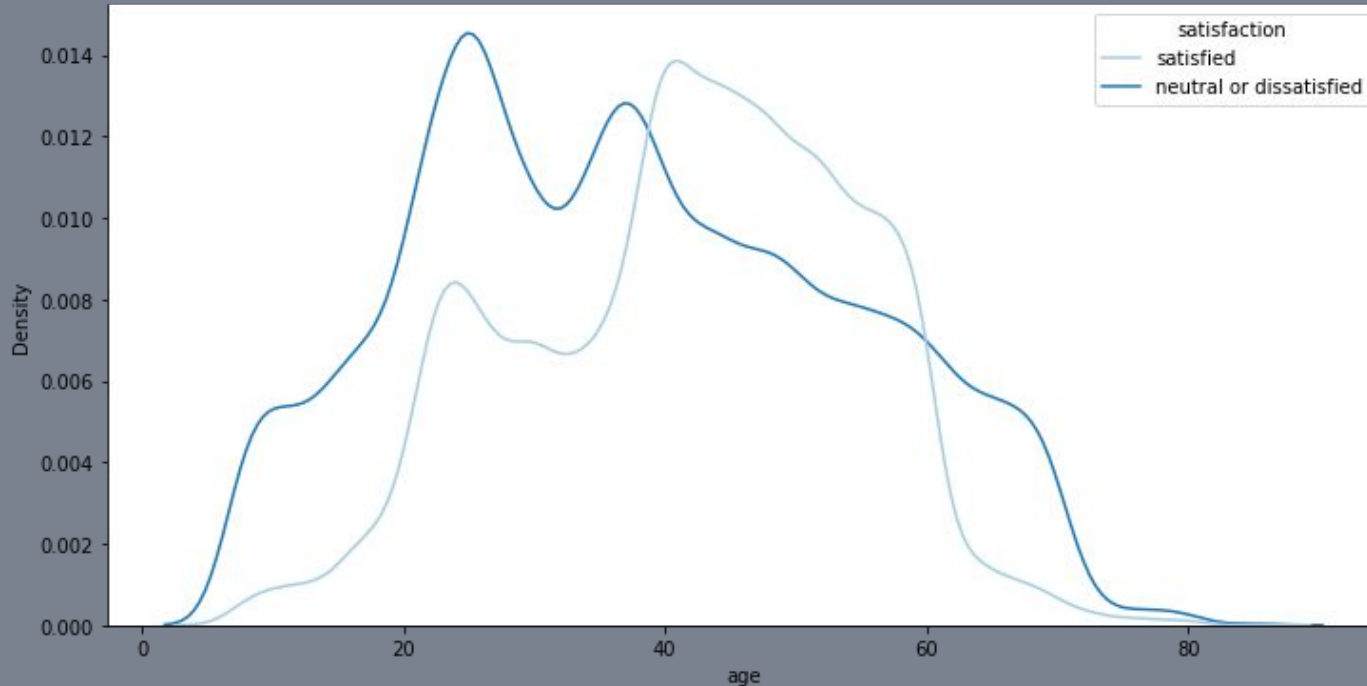


Se pueden observar más respuestas de pasajeros en el rango de 20, 40 y 50 años, luego disminuye la cantidad



TRAVEL DATASET

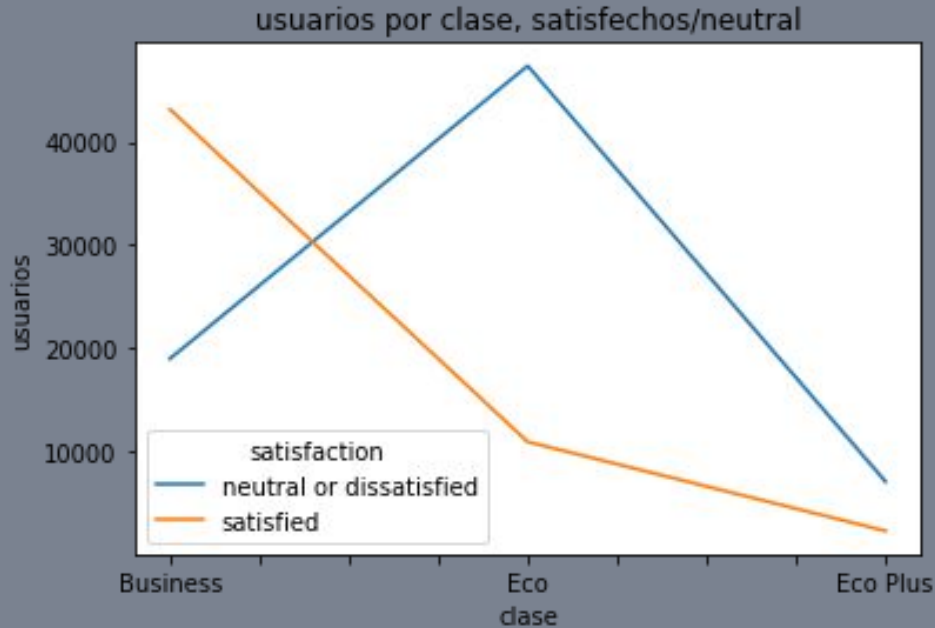
Satisfaction segun edad



Ahora queríamos dividir los datos un poco más, y notamos que hay más pasajeros satisfechos de 40 a 60 años aprox, y entre los 20 y 40 hay más cantidad de pasajeros insatisfechos

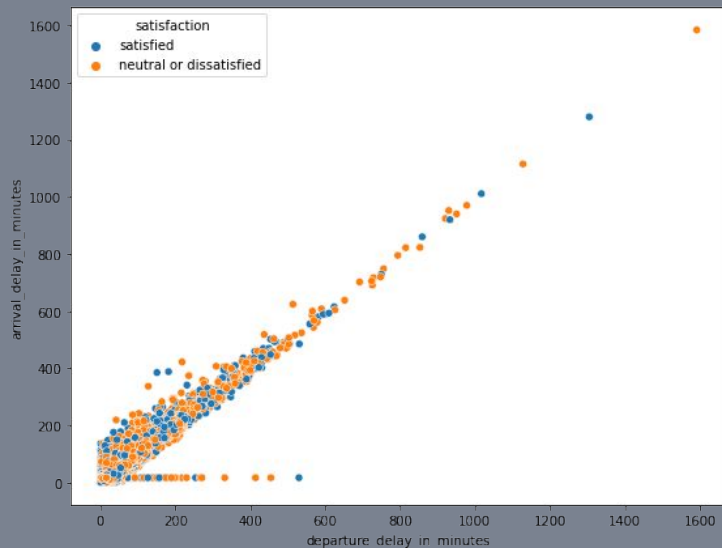


TRAVEL DATASET



Notamos que la clase business es donde se encuentra la mayor cantidad de pasajeros satisfechos por encima de 4000 y la eco clase es donde mas pasajeros estan insatisfechos , tambien notamos que la clase eco plus está bastante similar entre satisfecho y no satisfechos/neutral

TRAVEL DATASET



vemos la fuerte correlación entre el retraso de salida y el retraso de llegada; también podemos ver si el pasajero quedó satisfecho o no



se observa una correlación fuerte entre departure_delay, arrival_delay y delay_t (esta última variable fue creada por nosotros)