

En aquesta **TASCA** has d'entregar un document PDF d'un màxim de 4 pàgines en el qual aportis la següent informació sobre el teu projecte final:

- **Presentació del conjunt de dades escollit:** explicació i observacions generals del conjunt de dades a utilitzar per a fer el projecte final.
- **Característiques generals:** sense necessitat d'entrar molt en detall, explicar les principals característiques que defineixen aquest conjunt de dades. Tipologia, sector, tipus de dades, font, context, etc.
- **Definició de les variables:** explicació teòrica de les principals variables que conté el conjunt de dades.
- **Presentació dels objectius:** detallar els objectius inicials marcats de cara a extreure informació rellevant del conjunt de dades.

#### Base de dades:



Per al projecte faré servir una base de dades d'una línia aèria espanyola amb dades del 2022 fins al maig, i 230.000 registres (salts=vols).

#### Definició de les variables:

Les dades són reals i l'estructura del fitxer és aquesta:

DATE	Data
FLT	Vol
REG	Matrícula
AC	Tipus d'avió
DEP	Codi Aeroport Sortida
ARR	Codi Aeroport Arribada
STD	Hora estimada de sortida
STA	Hora estimada d'arribada
TKof	Hora de enlairament a capçalera pista
TDwn	H. de aterratge en la pista
ATD	Hora real de sortida

ATA	Hora real de arribada
BLOCK	T. estimat de vol
FLThr	Temps de vol real
ACT PAX	Quantitats de passatgers en l'avió
Taxi-out	Temps de taxi de sortida
Taxi-In	Temps de taxi a la arribada
SLOT	H. teòric donat per el controlador
C1	Codi Rebuig 1
DLY1	Minuts amd aquest codic
Sub1	criticitat
C2	Codi rebuig 2
DLY2	Minuts amb aquest codi
Sub2	crtiticitat
C3	Codi rebuig 3
DLY3	Minuts amb aquest codi
Sub3	crtiticitat
C4	Codi rebuig 4
DLY4	Minuts amb aquest codi
Sub4	crtiticitat
Close Pax Door	H. porta tancament passatgers
Close Cargo Door	H. tancament bodega de l'avió
Open Cargo/Pax Door	H. obertura carga avió
close pax door2	H. obertura bodega avió

Aquesta base de dades es complementa amb altres bases de dades del 2019 i del 2020 de la mateixa companyia. El principal problema és que l'operativa de la companyia en aquest període, avants del COVID, ha canviat, per la qual cosa la informació és difícil poder contrastar-la perquè el període és molt diferent. Per minimitzar l'impacte dels retards el que han fet és posar un buffer de 1 hora després de 3 salts que reduïu l'impacte del retard per eixample. Però, ha sigut efectiu?

Una altra base de dades serà d'una altra companyia que opera a Barcelona però que principalment opera a la terminal 2 per poder fer un benchmarking.

Les columnes són molt semblants entre cada base de dades al contingut de la columna, però no iguals. Una té la columna data i en una altra té el camp hora. Hi ha altres bases que a la columna té el camp data i hora junts que ajuda molt en els càlculs.

L'estructura de la base de dades per cada línia aquesta informació principal de temps d'un salt d'avió. Per la referència podem saber en aquell dia que vols (salts) ha fet. Si filtrem per data, matrícula d'avió, tenim tots els salts que ha fet un avió aquell dia.

Com està ordenat per STD o per l'ATD, podem saber que el que té les STD més baixa (hora més primerenca) és el primer vol del dia de aquell avió, el segon és el vol següent, i així successivament.

És important identificar que un avió pot arribar al dia següent i això cal tenir-ho en compte perquè la columna data només dona el dia d'enlairament. Poden sortir temps de vol absurds.

La base de dades es té que netejar perquè hi ha registres amb un valor de 0, o no hi són, o no tenen raó de ser perquè són temps negatius per exemple.



El retard és motivat perquè l'avió surt tard, i s'endarrereix més perquè perd el SLOT donat pel controlador aeri (ATC), cosa que obliga l'avió a esperar fins que l'ATC doni un altre SLOT per minimitzar l'impacte de sobrecàrrega per avions volant si surtis en aquell moment.

Les fases de vol són embarcament passatgers i maletes, tancament portes, taxi a capçalera pista, enlairament, vol, aterratge, taxi fins a terminal i obertura de portes.

Al vol d'un avió hi ha uns temps pre-definits que són l'STD i l'STA, (l'hora d'enlairament i hora d'aterratge), que es pot veure a les pantalles dels aeroports o pàgines de compra de bitllets. Aquests horaris són fixos i són les referències que hi ha de temps.

Hem d'analitzar si un avió surt abans de la seva STD, que ho pot fer si teniu permís de l'ATC i tots els passatgers/maletes estan dins de l'avió. Aquesta situació sempre passa al primer salt que l'avió fa al dia. Als salts posteriors hi ha l'impacte en el retard que l'avió acumula dels assalts anteriors.

Cal analitzar si el temps de vol és possible reduir el retard anant l'avió més ràpid o escurçant en l'aproximació en no haver de fer voltes (amb autorització de l'ATC), o perquè el temps calculat és més llarg que la mitjana del vol normal .

Els temps de boarding, en teoria, són dependents de l'avió i l'aeroport que estigui l'avió, però sabem que per exemple a l'aeroport de Barcelona aquests temps són més llargs (aeroports gran) que aeroports més petits perquè l'impacte en el control dels SLOT per congestions del trànsit aeri és menor i el personal de terra pot controlar millor el temps de boarding i landed

Una cosa que ens trobarem és que als retards hi ha molts outliers pel que fa a la manera d'operar normal d'un avió. Es pot donar el cas que l'avió pateixi un retard important perquè el motor s'ha fet malbé, i decideixo que aquestes situacions s'escapen a l'estudi de normal de treball i eliminaré tots els vols que el seu retard i avenç del vol sigui superior a 30 minuts. Aquests outliers només donen soroll i molt poca informació. Analitzaré quin percentatge de vols estan afectats per aquest motiu, i si és inferior al 5% consideraré la meua assumptió es ok.

Sobre la qualitat de la informació (MSA, measure sistema analysis) he de dir que els pilots no tots procedeixen de la mateixa manera, perquè com que estan molt controlats en els seus objectius personals, el temps de retard, controlen el moment 0 traient el fre, i immediatament canvia d'estatus l'avió i imputa aquest retard a un altre motiu, i cal tenir-ho en compte a l'hora de crear el model.

Un tema important és saber quina és la porta d'embarcament perquè el temps d'embarcament desembarcament no és igual si es fa amb finger o a la pista a què s'ha d'agafar un autobús.

Un avió, com a element individual, té poc marge de maniobra i es veu supeditat a les ordres de l'ATC i els slots. Per exemple, amb un cotxe podem sortir tard, i podem escollir anar per autopista o per carreteres, o escollir els carrers més amples o buscar dreceres, o parar si volem o necessitem.

A l'aviació no es pot. Si hi ha molts avions en circulació per seguretat cal separar-los a l'espai 3D i això es fa per temps. Els avions sempre estan volant el que implica desplaçament i consum de combustible. En conclusió, que el nombre d'avions volant en un sector aeri mai no pot ser infinit, sinó que finit.

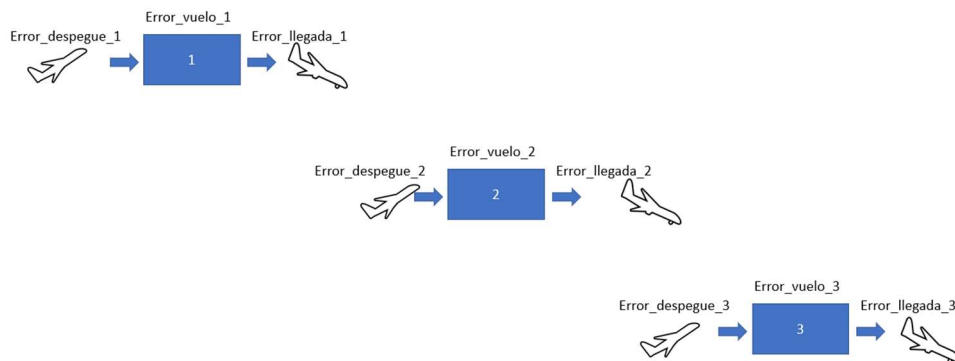
Un altre element que distorsiona molt, és el temps atmosfèric, si hi ha boira, pluja dia nit, visibilitat etc. En situació extrema s'obliga els avions a augmentar aquesta distància de separació, cosa que força haver-hi més retards. Aquests endarreriments impacte en altres aeroports que poden tenir unes condicions de sol i de visibilitat superior a 30 km.

És clar que els mesos d'hivern les condicions ambientals són pitjors en general a Europa, però no als països d'altres latituds. I en períodes d'estiu i en dates molt assenyalades com pot ser principi de setmana, cap de setmana, Nadal, etc els avions se saturen més. Hi ha molts més SLOTS per hora, cosa que provoca interferència en les puntualitats entre companyies. És important fer un mapa de calor de retards per veure si hi ha períodes (mes/setmana/dia específic) amb més impacte.

El que és clar que al primer salt, el pilot té la capacitat de poder sortir 10 minuts abans que és el temps que obliguen el passatge a estar abans de l'embarcament i no es veu influenciat per exemple amb passatgers que van tard per retards en altres vols en salts posteriors. i que l'avió ha d'esperar que tot el passatge/equipatge estigui a bord.

Objectius:

Analitzar les causes generals del retard dels avions, i focalitzant-nos al 3r salt, (tercer vol). Si el retard és superior a 10 minuts, que és un temps acordat per les companyies com **"No retard"**, saber quin percentatge d'avions arriben a l'hora .



Em vull focalitzar especialment a l'aeroport de Barcelona i veure l'impacte del retard que hi ha.