1. **Abstract**: Resum del vostre projecte. Què conté, objectiu i resultat. Molt resumit

Tinc que esbrinar quin es la puntualitat d’una companyia aèria en lo que ells denomina’n el tercer salt, fent servri una base de dades proporcionada per ells

1. **Introducció**: Aquí fem una introducció al nostre problema dins del marc de la Data Science. Així com objectius que tenim i els requisits que hem de complir.

La base de dades te varies columnes amb les dates de enlairament i aterrament de cada avió. Un avió en un dia fa varis “salts”. Primer tinc que adequar els camps temps a un format date per poder fer la resat de temps. Tinc que identificar el vols de cada avio per dia, ordenar-los per hora de sortida per poder identificar cada salt . Això obligarà a treballar en un format de línia perquè el temps calculats de cada salt estiguin en una sola línia per després poder filtrar el fitxer per tenir tota la informació de un avió per dia, el tres salts. Amb això reduirem el temps de càlcul al simplificar el tamany del arxiu en un 80% accelerant el procés. Pesem de 500.000 Líneas a 80.000.

Una vegada el arxiu està muntat podrem fer els principals càlculs i poder fer estudis de capacitat, calculant el Cp y Cpk. Obtenint el valor de % de vols que la seva puntualitat es més gran de 0 o 10 minuts.

Tinc que esbrinar si les dades de enlairament aterrament, temps de boarding i temps de vols estàn relacionades entre salts. Tinc que crear una regressió logística en que una sortida es 0 que l’avió arriba a l’hora o 1 que arriba 10 minuts tard del temps planificat.

La base de dades està molt des-balancejada perquè el 9% del avions no arriben a l’hora per lo que si volem tenir un resultat creïble esté que balancejar. Farem servir el mètode ....... per tenir un resultat mes o menys del 50% per el 0 i 1.

1. **State of Art**: Aquí comentem com s'ha resolt en el nostre problema amb anterioritat, els mètodes ja utilitzats i els resultats ja extrets. D'aquesta manera donem una perspectiva de quins objectius volem arribar a batre així com intentar no tornar a fer servir elements ja utilitzats.

No em consta haver vist una base de dades semblant a aquesta perquè son dades bastant confidencials del 2022 i 2019. Segur que les companyies fan anàlisis interns per esbrinar el seu comportament amb la puntualitat. Però intentaré arribar una mica més juny comparant la puntualitat amb un altre companyia per exemple.

Principalment la idea es saber si la puntualitat que tenen es deguda a una raó i esbrinar quina és aquesta raó.

1. **Metodologia**: Aquí explicarem el preprocessat, els models, les mètriques així com llibreries si s'escau.
2. **Resultats**: Aquí és hora de mostrar resultats, taules i gràfiques. Així com remarcar els millors resultats obtinguts.
3. **Conclusió**: Aquí tornarem a portar els objectius i determinar si els hem complert o no. Comparar l'state of art així com presentar totes les conclusions extretes.
4. **Referències**: És important documentar totes les vostres referències!!!