

REPUBLIQUE D'HAITI



AKADEMI



RAPPORT D'ANALYSE DES AERONERFS LES MOINS RISQUÉS

PRÉSENTÉ PAR : DEBREUS Monitès

PROFESSEURS : JEROME Wedter, LAGUERRE Geovany

Date : 09 Juin 2025

Information personnelle

Email : monites.debreus@student.ueh.edu.ht

Profil linkedin : www.linkedin.com/in/monites-debreus-7b4b7a1ab

Table des matières

1. Introduction	5
1.2. Objectifs.....	5
1.2.1. Objectif principal :	5
1.2.2. Objectifs spécifiques :.....	5
2. Public cible :	6
3. Compréhension des données disponibles :	6
4. Méthodologie	7
4.1. Matériels utilisés	7
4.2. Méthode :	7
4.2.1. Lecture du fichier	7
4.2.2. Nettoyage des données	8
4.2.3. Résultat et visualisation.....	10
5. Résultat et Analyse.....	10
6. Conclusions et recommandations.....	19

LISTE DES FIGURES

Figure 1: les éléments de risques investigués.....	11
Figure 2:Evolution des accidents.....	12
Figure 3: Nombre de personnes impliquées dans les accidents	12
Figure 4: Accidents en fonction du but de vol	13
Figure 5: Fatalité par phase de vol	14
Figure 6:Fatalité selon les conditions météorologiques	15
Figure 7:Degre de fatalité et nombre de moteurs	16
Figure 8: Les aéronefs les plus impliqués dans les accidents.....	16
Figure 9: Aéronefs avec plus de personnes indemnes.....	17
Figure 10:top 5 des plus sur par accidents.....	18

Liste des tableaux

Tableau 1:Aéronefs les plus surs	18
--	----

1. Introduction

La "bibliographie" de l'industrie aéronautique (composée des rapports des grands acteurs et institutions) dépeint un secteur riche économiquement , en croissance constante sur le long terme, et doté d'une résilience avérée. Ce secteur, bien que porteur d'opportunités, est aussi connu pour sa complexité, ses exigences réglementaires strictes et surtout, les risques associés à la sécurité aérienne. Avant de se lancer dans ce secteur, il est recommandé de s'appuyer sur des données pour faire des analyses pour anticiper les dangers et prendre de très bonnes décisions sur les types d'aéronefs à acquérir.

Dans cette optique, pour répondre aux exigences de la compagnie pour l'achat et l'exploitation d'avions pour des entreprises commerciales et privées afin de se développer dans de nouvelles industries et de diversifier son portefeuille, on s'appuie sur un vaste jeu de données que la « National Transportation Safety Board (NTSB) » a mis à la disposition du grand public, couvrant plus de soixante ans d'accidents d'aviation civile (de 1962 à 2023), aussi bien sur le territoire des États-Unis que dans les eaux internationales pour tirer des conclusions et les mettre à la disposition de la compagnie pour mieux orienter ses choix.

1.2. Objectifs

Les objectifs de ce travail sont classés en deux catégories : objectif principal et objectifs spécifiques

1.2.1. Objectif principal :

L'objectif principal de ce projet est de fournir des recommandations pertinentes fondées sur l'analyse des données historiques, afin d'aider l'entreprise à sélectionner les aéronefs les moins risqués dans le cadre de son expansion vers l'aviation commerciale et privée.

1.2.2. Objectifs spécifiques :

En vue d'atteindre l'objectif principal, on se charge :

- D'identifier les types de risques liés aux activités aéronautiques de la base de données
- d'évaluer l'évolution des risques dans le temps et leur degré de sévérité
- de regarder la relation des risques avec d'autres facteurs pouvant influencer les vols
- de proposer un classement des aéronefs les moins à risque basé sur un indice de risque intégré ;

2. Public cible

Ce travail s'adresse avant tout à la direction stratégique de la nouvelle division aviation de l'entreprise, en particulier à son chef de projet et/ou à son directeur exécutif, qui va devoir prendre des décisions d'investissement sur la base des conclusions de ce rapport et également à tous les partenaires institutionnels susceptibles d'être impliqués dans l'achat, l'exploitation ou l'assurance des aéronefs.

3. Compréhension des données disponibles :

Cette base de données constitue une source d'information essentielle pour la prise de décisions concernant les activités aéronautiques. Avec ses 88 889 lignes et 31 colonnes, elle offre une richesse statistique couvrant des aspects techniques, opérationnels, géographiques, environnementaux et humains qui sont impliqués dans les activités aéronautiques. Grâce à des décennies de données, elle permet une exploration des évolutions temporelles des différents accidents. Cependant, elle contient de nombreuses valeurs manquantes ou mal typées nécessitant des travaux de nettoyage, d'imputation, d'agrégation pertinente et de visualisation pour en exploiter tout le potentiel. En procédant ainsi, on peut identifier les modèles d'avion les plus sûrs, ceux présentant les taux de personnes indemnes par accidents ou encore une moindre gravité des risques recensés.

4. Méthodologie

Cette section comprend les matériels, logiciels et extension qu'on a utilisé pour la réalisation de ce travail ainsi que la méthode que l'on a suivi pour parvenir aux résultats escomptés.

4.1. Matériels utilisés

- Ordinateur portatif (hp, core i5, 7^e generation): principal support sur lequel on a installé tous les logiciels et extensions nécessaires pour faire l'analyse
- Excel (Ms 365) : utilisé pour une première visualisation des données et se renseigner sur les infos que contiennent chaque colonne.
- Git (version 2.49.0), Anaconda (version 2024.10-1), Jupyter notebook :respectivement des logiciels et environnement interactif qui permet d'exécuter du code dans des notebooks organisés en cellules permettant de clean les données, de faire des analyses et des graphiques.
- Power BI desktop (version 2.143.1378.0) : pour la réalisation des graphiques interactives.
- Chatbot : pour le dépannage des erreurs et des conseils d'écritures de codes
- Les supports du cours pour la révision de l'utilité de certaines fonctions et méthodes en python.
- Word (Ms 365) : pour la rédaction du rapport final.

4.2. Méthode :

4.2.1. Lecture du fichier

Pour la réalisation de ce travail, on a premièrement ouvert le jeu de données contenant 88 889 lignes d'enregistrements d'accidents/incidents aériens sur Excel, afin d'inspecter les colonnes, leurs contenus et avoir une première idée des variables d'intérêt.

Cela n'étant pas suffisant pour identifier les données pertinentes pour l'analyse, on a donc importé toutes les bibliothèques jugées nécessaires sur 'jupyter notebook' et chargé ensuite le fichier en utilisant 'pd_read_csv('nom du fichier')' et lui assigner un nom :df. Malheureusement, l'encodage n'était pas de type UTF-8, on a dû recourir à des « chatbots » pour résoudre le problème de lecture du fichier.

Après avoir lu convenablement le fichier, on utilise une ligne de code qui tient à ce qu'aucun avertissement ne soit pas affiché sur jupyter notebook toutefois qu'il y aurait des changements de comportement prévu dans une future version. Ensuite, on a créé une copie du fichier et faire les manipulations de bases pour vérifier la forme, le nom de toutes les colonnes présentes et le type de données que contient chaque colonne. Ensuite, on a procédé à une analyse descriptive sommaire des colonnes de type « float » à l'aide de la fonction « describe () ».

4.2.2. Nettoyage des données

Le nettoyage des données commence à la suppression des potentielles espaces vides compris au début et à la fin de chaque nom de colonne et calculer ensuite la somme des données manquantes dans chacune et les triés par ordre décroissant.

Cela étant fait, on a éliminé les colonnes avec plus de 60% données manquantes puisque toute tentative d'imputation risquerait d'introduire des biais majeurs dans les données. On a aussi supprimé d'autres colonnes jugées non-pertinentes pour l'analyse que l'on souhaite faire. Comme : Report.Status, Registration.Number...

Après quoi, on a procédé à la vérification des doublons dans les données à l'aide de certaines colonnes clés comme 'Event.Id'. Après avoir trouvé les doublons, on les a éliminés ainsi que la colonne « Event.Id » puisqu'elle n'est pas jugée pertinentes pour la poursuite du travail et procédé à la conversion des informations de la colonne « Event.Date » en date tout en créant une nouvelle colonne à partir d'elle contenant que les années.

Pour les données restantes, jugées pertinentes pour la suite du travail, on les a classées en deux grandes catégories :

- données discrètes : Total.Serious.Injuries, Total.Minor.Injuries...
- données catégorielles : Broad.phase.of.flight, Engine.Type, Purpose.of.flight, weather.condition...

On fait des fillna('unknown') dans certaines données catégorielles tout en s'assurant de l'homogénéité des chaînes de caractère en minuscule. Cependant dans « weather.Condition » pourvu qu'il y a 3 catégories de données collectées qui sont: UNK,IMC,VMC traduisant respectivement les conditions: « Inconnu, Instrument Meteorological Conditions, Visual Meteorological Condition », les données manquantes de cette colonne seront remplacées par « unk » pour éviter toutes ambiguïtés dans l'analyse des résultats.

Pour les données discrètes, on fait des « fillna(0) » puisque dans le contexte des accidents d'aviation, une valeur manquante pour le nombre de blessés ou de décès signifie probablement qu'il n'y a eu aucune blessure ou aucune donnée de décès enregistré pour cet événement donc imputer par la moyenne ou la médiane fausserait à la hausse les statistiques de sécurité, ce qui serait très problématique pour notre objectif d'évaluation des risques. Cependant pour les données concernant le nombre de moteurs, on a supprimé les lignes vides pour ne pas fausser les analyses puis qu'il y a aussi des appareils qui ont zéro moteur. Étant donné que la compagnie ne va pas investir dans des types d'aéronefs dont les informations sur le nombre de moteurs ne sont pas fournies(ce qui semble logique), dropper les lignes NaN ne paraît pas trop non-judicieux pour la poursuite de l'analyse quoique cela engendrera la perte de quelques lignes d'informations.

On a finalement enregistré le travail qu'on vient de faire dans un format « csv » et lui attribuer un autre nom.

4.2.3. Résultat et visualisation

Après avoir fait ce travail de « data cleaning », on a utilisé « Jupyter Notebook » et Power BI Desktop pour procéder à l'analyse et la visualisation des résultats. Les analyses se portent sur trois grands axes :

1. D'abord avoir une vue globale des accidents. Cela permet de suivre l'évolution des accidents dans le temps, le nombre de personnes impliquées par an tout en identifiant les buts de ces vols, le degré de gravité de ces accidents en fonction des phases de vol ainsi que les conditions météorologiques dans lesquelles ils ont été réalisées.
2. En second lieu, on a regardé la relation « caractéristiques des aeronefs & accident » dans laquelle on prendra en compte des liens existants entre le nombre de moteurs des aéronefs et le nombre d'accident.
3. En troisième lieu, on a analysé le nombre de personnes non-blessées par make modele. On a stocké ces modèles dans une liste et calculer un ratio (indemnité pour trier les moins risqués de ces modèles.

5. Résultat et Analyse

- a) Quantification des risques liés aux activités aéronautiques de la base de données :

L'analyse des données se porte sur deux risques combinés, il s'agit des accidents qui représentent près de 97% des données traitées contre près de 3% pour les incidents.

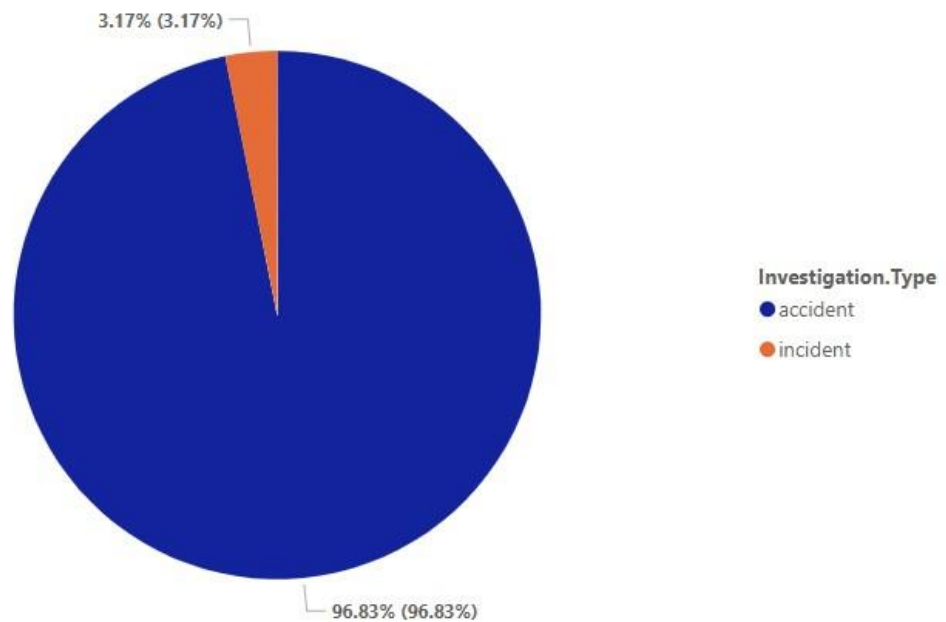


Figure 1: les éléments de risques investigués

NB : pour la continuité du travail, on va faire mention que des accidents pour faire références aux deux risques.

b) Vue globale des accidents dans le temps

Parlant des risques, on voit leur évolution diminuer au fil des ans pour passer de plus de 3500 accidents par an à moins de 1500/an. Cela suggère une amélioration globale de la sécurité aérienne au fil du temps.

Le graphique ci-dessous donne une idée de l'évolution des risques (accidents/incidents) par an.

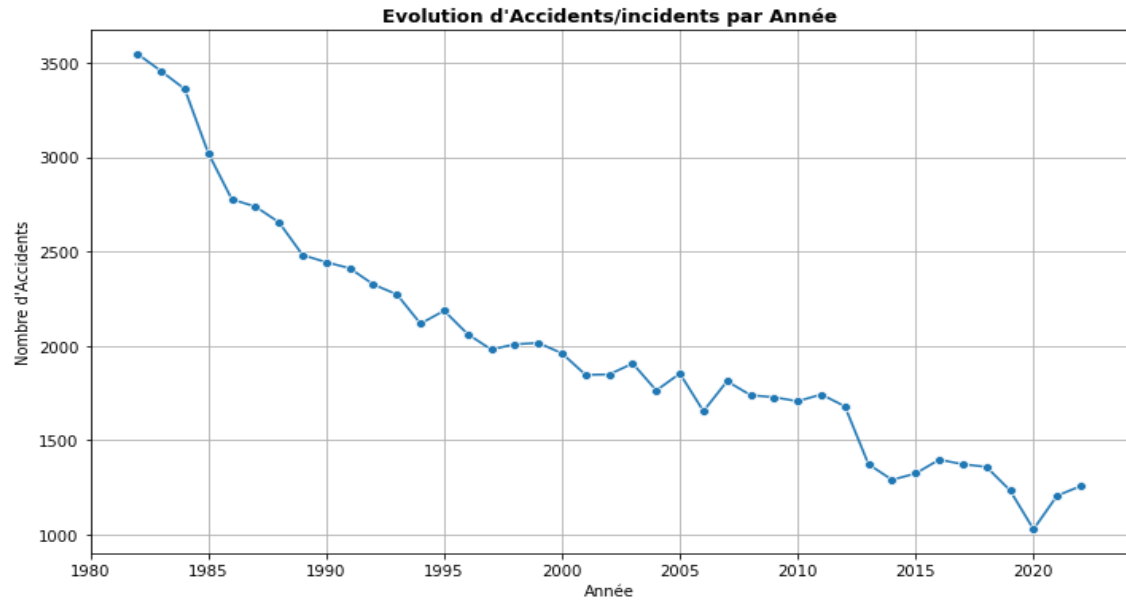


Figure 2: Evolution des accidents

Aussi bien que les risques diminuent, le nombre de personnes impliquées dans les accidents diminuent également, bien qu'avec une pente plus douce. De façon plus approfondie, on peut observer que le nombre de personnes impliquées dans les accidents sans avoir eu la moindre blessure est de loin supérieur au nombre de personnes qui sont mortes ou blessées. Le graphique suivant permet d'apprécier cette diminution.

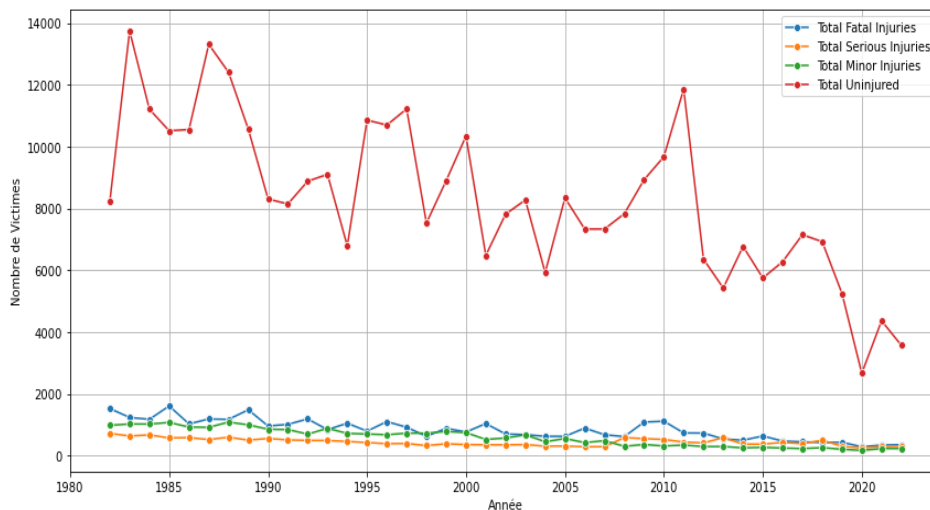


Figure 3: Nombre de personnes impliquées dans les accidents

En se basant sur le graphique ci-dessus, dans une certaine mesure, on voit que les accidents des aéronefs sur cette période n'ont pas été toujours mortels ou du moins ont

été moins mortels qu'on le pense, sans ignorer pourtant le poids des cas de blessures et de fatalités.

Après avoir regardé, l'évolution des accidents dans le secteur aérien, on se pose les questions suivantes :quels ont été les buts de ces vols ? quelles ont été les conditions météorologiques au moment des accidents ? Dans quelles phases les accidents surviennent-ils le plus ? Ont-ils été fatal ou non ?

La réponse à ces questions seraient utiles à l'entreprise pour mieux appréhender les risques du secteur aéronautique, mieux orienter ses choix de vols et prendre en considération les protocoles de sécurité fixés par l'OACI.

Les graphiques suivants permettent d'avoir une idée du nombre d'accidents en fonction des buts de vols.

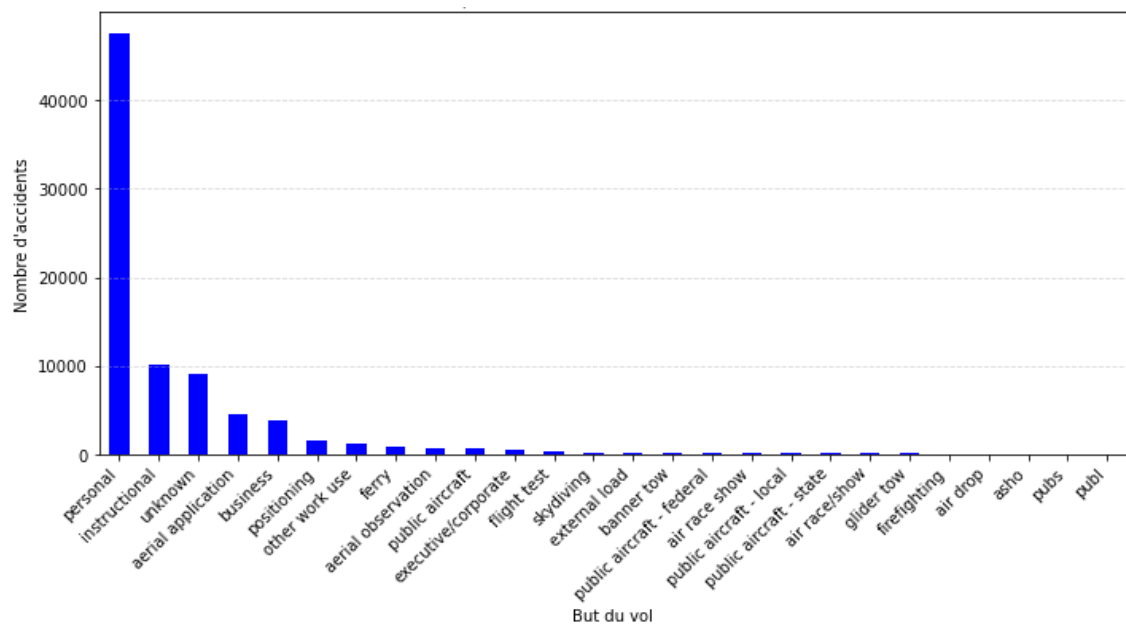


Figure 4: Accidents en fonction du but de vol

Selon les données, de tous les accidents enregistrés, on remarque que les 5 buts de vol ayant le plus grand nombre d'accidents sont : «personnal, instructionnel, aerial application», business et une quantité d'accidents dont les motifs n'ont pas été enregistré...Les motifs de ces vols sont bien similaires à ceux de l'entreprise, soulignant

davantage l'intérêt de l'entreprise de se baser sur les données pour choisir les meilleurs aéronefs pour ses ambitions de vols commerciales et privés.

Les résultats ont également montré que ces accidents ont été beaucoup plus fatales dans les phases « cruise, manoeuvring, takeoff, approche et climb »

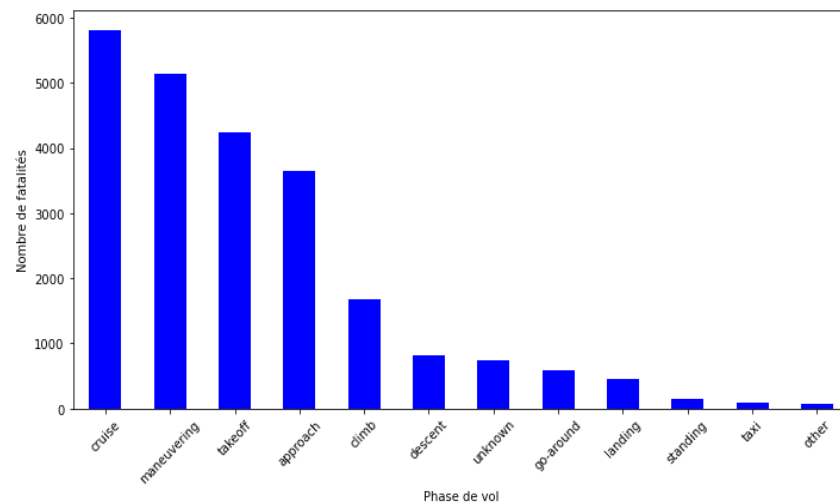


Figure 5: Fatalité par phase de vol

Et ce, environ 68% dans des conditions météorologiques VMC¹ contre environ 27% en IMC² comme c'est décrit sur le graphique suivant.

¹ VMC : Visual Meteorological Conditions

² IMC : Instruments Meteorological Conditions

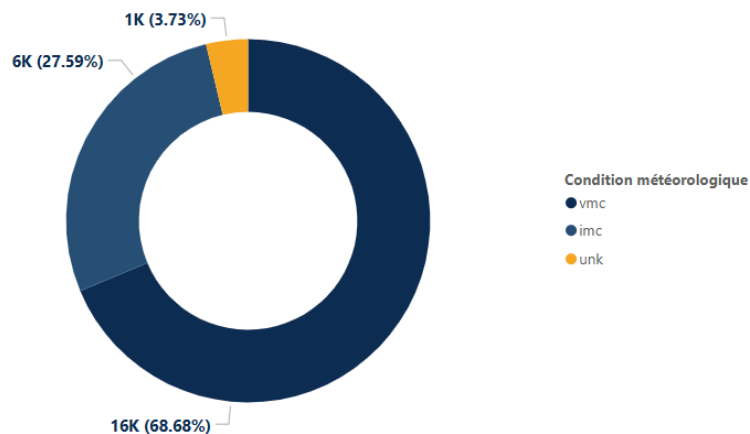


Figure 6: Fatalité selon les conditions météorologiques

Cela paraît contre-intuitif puisque les conditions météorologiques VMC (ciel clair) sont favorables aux vols à vue (VFR³) tandis que dans les conditions IMC, cela nécessite forcément de passer en IFR⁴, mais cela a du sens lorsqu'on tient compte qu'en VFR, le pilotage dépend de la vision du pilote pour naviguer et éviter les obstacles, certains pilotes peuvent potentiellement prendre plus de risque, sous-estimer le danger alors qu'en IFR, ce sont les instruments de bord qui présentent à l'équipage technique toutes les informations utiles au pilotage et au fonctionnement de l'aéronef, la gestion du risque est donc plus stricte.

c) Relation entre caractéristiques des aéronefs & accident

Un peu plus loin, les analyses faites sur le nombre d'accident en fonction du nombre de moteurs des aéronefs, le nombre de personnes victimes parmi ceux qui ont été à bord en fonction du nombre de moteur et le nombre de personnes mortes (fatalité), ont montré que ces aéronefs ont majoritairement un seul moteur.

³ VFR: Règle de vol qui offre la simplicité et la liberté de la navigation à vue, idéale par temps clair.

⁴ IFR : Règle de vol privilégiée pour le vol aux instruments, essentielle par faible visibilité, notamment par temps de brouillard, de pluie ou de nuit.

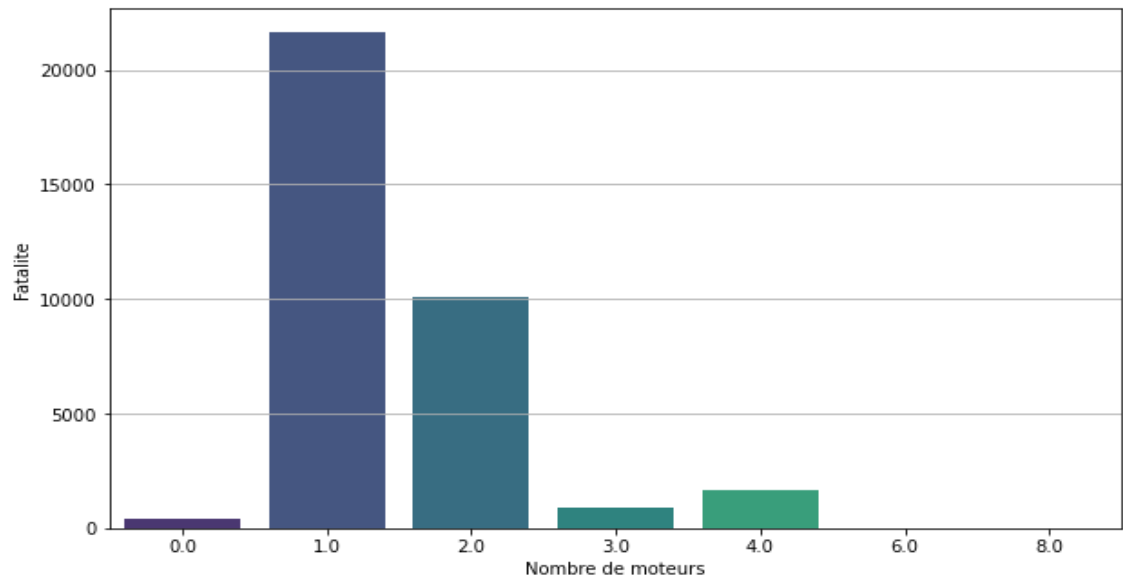


Figure 7: Degré de fatalité et nombre de moteurs

Selon les données, ces aéronefs d'un seul moteur sont les plus nombreux dans le ciel, souvent pilotés par des pilotes moins expérimentés selon la littérature et volent souvent à basse altitude en VFR ce qui peut causer des collisions avec le relief et tant d'autres accidents. En voici une idée de 20 marques et modèles les plus impliqués dans les accidents.

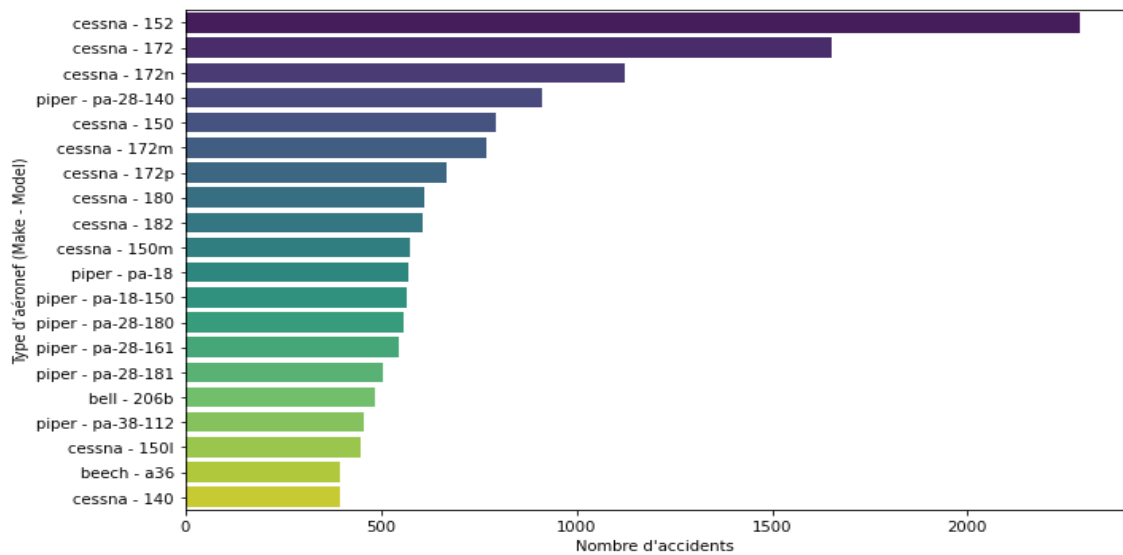


Figure 8: Les aéronefs les plus impliqués dans les accidents

Le graphique ci-dessus présente les résultats de mesure de fréquence des accidents et non de gravité. C'est-à-dire, un type d'aéronefs peut avoir 200 accidents, mais avec très peu de personnes blessées ou dans le cas contraire avec beaucoup de cas de fatalités. De ce fait, présentons ceux qui ont le plus de personnes indemnes.

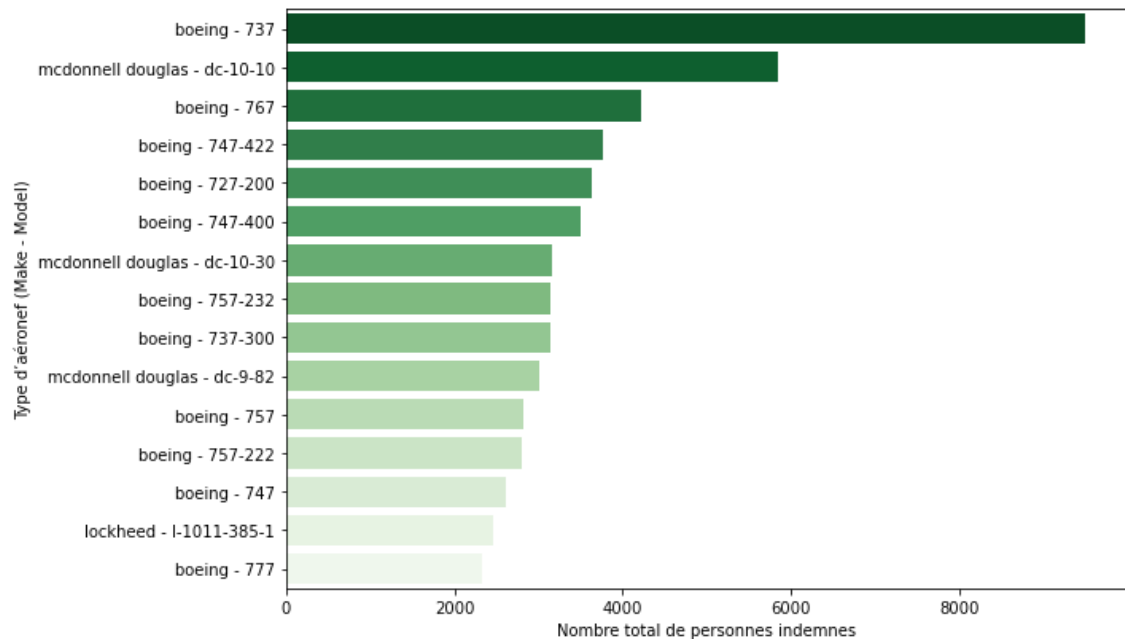


Figure 9: Aéronefs avec plus de personnes indemnes

Maintenant, on sait ceux qui ont les moins cas de blessures en fonction du nombre d'accidents qu'ils ont eu. A l'aide du ratio mettant en relation le nombre de personnes indemnes et le nombre totale de personnes qui ont été à bord, présentons le top 5 de ceux qui ont le ratio d'indemnité les plus élevés par accidents :

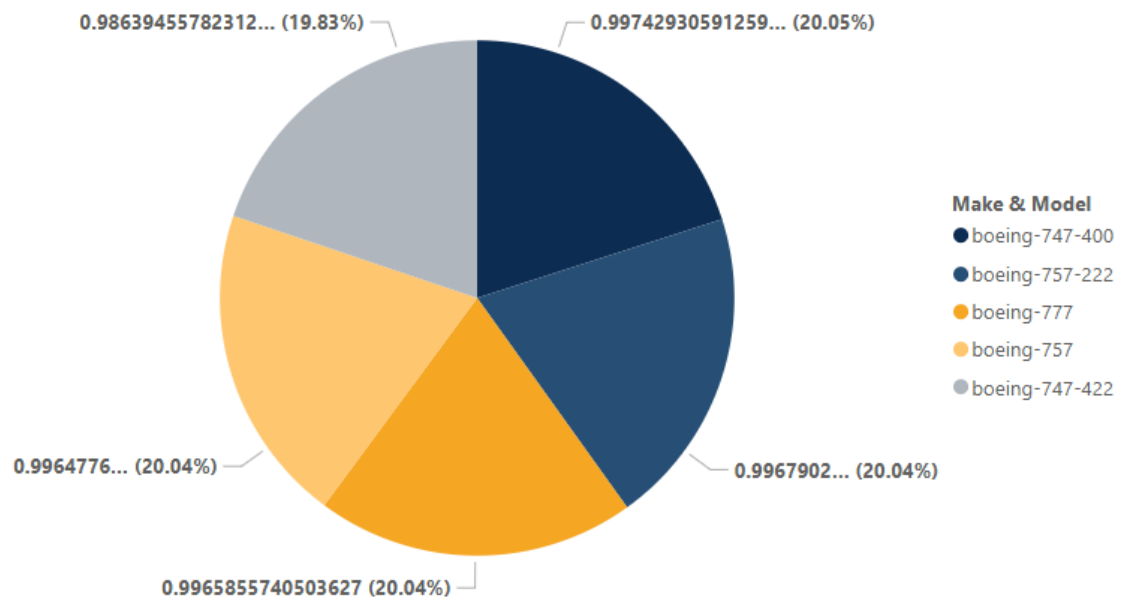


Figure 10: top 5 des plus sur par accidents

Le tableau suivant présente le nombre de moteur de chaque type d'aéronefs :

	Make_Model	ratio	Number.of.Engines
0	boeing-747-400	0.997429	4.0
1	boeing-757-222	0.996790	2.0
2	boeing-777	0.996586	2.0
3	boeing-757	0.996478	2.0
4	boeing-747-422	0.986395	4.0

Tableau 1: Aéronefs les plus surs

6. Conclusions et recommandations

L'analyse des données historiques issues de la NTSB révèle que les accidents constituent la quasi-totalité des événements à risque, avec près de 97% des cas étudiés. Heureusement, la tendance générale est à une nette de la sécurité aérienne au fil des années, avec une diminution progressive du nombre d'accidents annuels, ainsi que du nombre de personnes. Non seulement le nombre total d'accidents diminue, mais la gravité aussi. Cette évolution conforte l'idée que l'industrie aéronautique devient de plus en plus sécurisée, ouvrant ainsi des perspectives favorables aux investissements dans l'aviation commerciale et privée.

Cependant, la vigilance doit rester de mise, car il y a un nombre significatif de blessés et de fatalités subsistantes pour les vols de tout type, notamment avec des avions monomoteurs, qui représentent aussi la majeure partie de la flotte et dans certaines phases critiques du vol comme le décollage, les manœuvres ,cruise etc...

Ces accidents sont produits beaucoup plus en conditions VMC (ciel clair), probablement parce que les pilotes en vol VFR peuvent sous-estimer les risques et adopter des comportements plus risqués, qu'en conditions IMC, car la navigation instrumentale impose un protocole plus strict et permet de mieux gérer les risques. Cela met en évidence, la nécessité d'avoir un équipage très professionnel pour diminuer les risques d'accidents.

Enfin, en se concentrant sur la sécurité relative des modèles, grâce à la méthodologie appliquée, on a démarqué 5 modèles d'avions du point de vue de rapport indemnité et total de personnes à bord après avoir récupéré ceux qui ont eu le plus de personnes indemnes à savoir le Boeing 747-400 ayant 4 moteurs, le Boeing 757-222 de 2 moteurs, le Boeing 777 de 2 moteurs, le Boeing 757 de 2 moteurs et le Boeing 747-422 ayant 4 moteurs.

Ce classement final donne une orientation à la compagnie sur quel type d'avions qu'il doit se pencher pour une plus grande marge de sûreté de la vie de ses futurs clients. Ainsi,

l'entreprise peut envisager son expansion dans l'aviation commerciale et privée pour enfin diversifier son portefeuille. Toutefois, Il lui est fortement recommandé de:

- Prioriser ces 5 premiers aéronefs tels que car ils ont le meilleur ratio parmi ceux qui ont les plus grands nombres de cas indemnes.
- Prioriser les aéronefs comprenant les dernières technologies pour pouvoir voler en IFR en utilisant les instruments de bord pour naviguer car la gestion du risque est plus stricte et cela permettra aux pilotes de voler dans presque toutes les conditions météorologiques ;
- S'assurer d'avoir un équipage expérimenté qui maîtrise les différentes phases de vols pour éviter les accidents ;
- Mettre en place un bon programme de maintenance et un suivi rigoureux des conditions d'exploitation pour contribuer à la minimisation des risques d'accidents et d'incidents ;
- Compléter l'analyse par des critères contextuels en croisant ces résultats avec :
Les profils d'usage visés (vols courts, vols longs, fret, passagers...), la disponibilité des pièces de rechange et du personnel qualifié et la conformité aux réglementations locales et internationales.