

# **MyAkademi**

## **Bootcamp DataScience**

### **Analyse des accidents aériens de 1962 à 2023**

Mackenson JOSEPH

Juin 2025

## Objectif

Ce projet vise à identifier les avions, constructeurs et types de moteur les plus sûrs à partir de données historiques d'accidents aériens.

## Contexte

Dans le cadre de sa stratégie de diversification, notre entreprise envisage d'entrer dans le secteur de l'aviation commerciale et privée. Elle souhaite proposer une flotte d'avions afin de proposer des services de transport aérien. Toutefois, avant de lancer cette nouvelle activité, une étude des risques liés à l'exploitation aérienne est nécessaire.

En particulier, la direction souhaite identifier **les modèles d'avions les plus**, sur la base sûre d'un historique fiable d'incidents et d'accidents. Nous avons été mandatés pour mener cette analyse en utilisant un ensemble de données fournies par le **National Transportation Safety Board (NTSB)**, couvrant les accidents d'aviation civile entre **1962 et 2023**.

# Méthodologie

- **Nettoyage des données :** Certaines colonnes sont supprimées du fait de leur peu de valeurs ajoutées à notre analyse. (*`Calendrier`*, *`Transporteur aérien`*, *`Description FAR`*, *`Catégorie d'aéronef`*, *`Code d'aéroport`*, *`Nom d'aéroport`*, *`Date de publication`*, *`État du rapport`*, *`Gravité de la blessure`* et *`Numéro d'enregistrement`*).

Les valeurs manquantes pour les colonnes catégorielles sont gérées en les remplaçant par *Unknown* pour éviter de supprimer toute la ligne et perdre des données sur d'autres colonnes vu l'enjeu je préfère ne pas faire de supposition, mais seront ignorées lors des analyses. Pour les colonnes numériques elles sont remplacées par la valeur étant donnée que les distributions en question sont asymétriques.

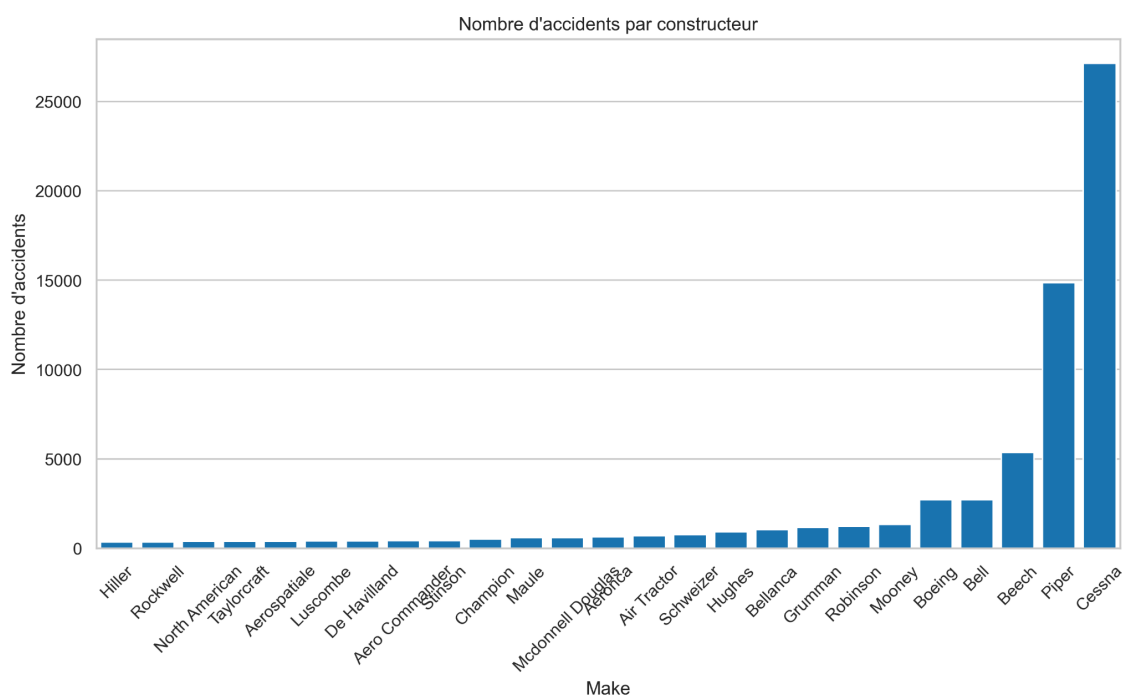
- **Variables retenues :** Les colonnes principalement utilisées dans ce projet sont (*Total.Fatal.Injuries*, *Total.Serious.Injuries*, *Total.Minor.Injuries*, *Make*, *Model*, *Event.Date*, *Engine.Type*).
- **Critère de sélection :** Pour chaque Constructeur, modèle et Type de moteur analysés, j'ai fait le choix de mettre de côté ceux avec moins de 300 accidents pour avoir un résultat le moins biaisé que possible.

# Analyse des résultats

## 1. Par constructeur (Make)

- Fréquence des accidents par constructeur

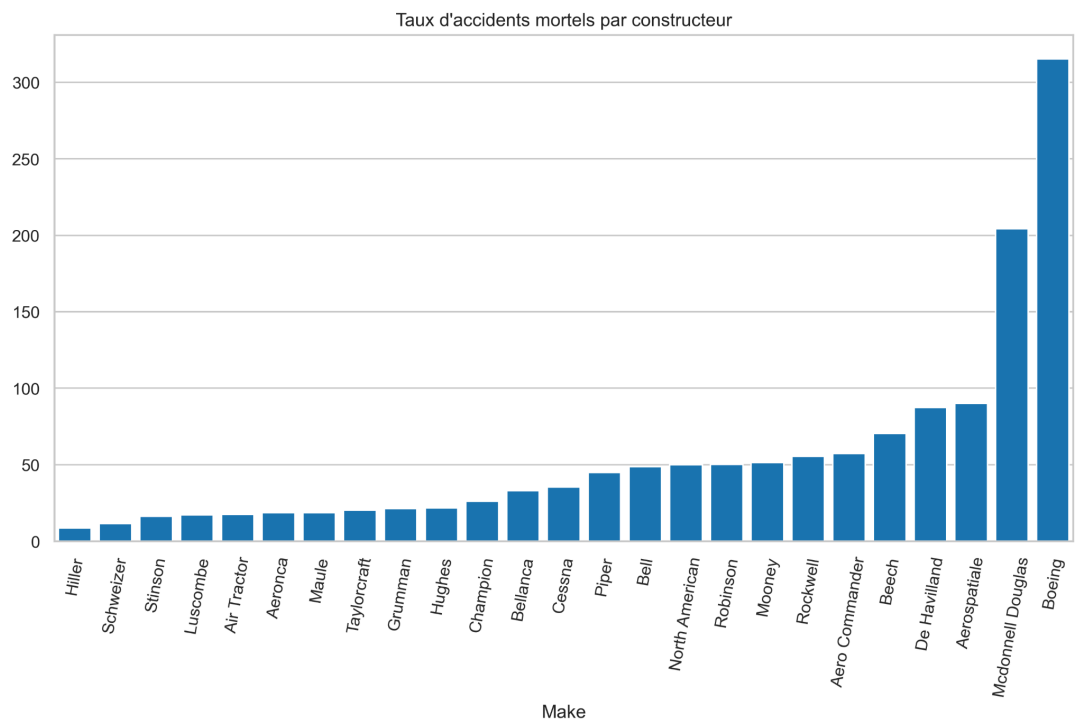
**Objectif :** Identifier quels constructeurs ont été impliqués dans le plus grand nombre d'accidents, sans distinction initiale de gravité.



Le constructeur **Cessna** arrive en tête avec plus de 27 000 accidents recensés, suivi de **Piper** et **Beech**. Plutard dans ce documents on verra que le d'accidents recensé par constructeur ne dit pas tout sur l'objectif de notre analyse.

- Taux d'accidents mortels et score de gravité

Le graphique ci-dessous montre un classement des différents constructeurs par rapport au taux d'accident mortels (ce dernier est multiplié par cent pour plus de facilité lors de la lecture, donc les valeurs affichées représentent le nombre de morts par cent accidents)



Contrairement à ce qu'on pourrait croire bien que le **Cessna** ait le plus grand nombre d'accident il n'est pas celui avec le haut **Taux d'accident mortels par constructeur**, la place est détenue par **Boeing** avec une moyenne de 315 morts par centaines d'accidents, soit 10 fois plus que **Cessna** avec une moyenne de 35 morts par centaines d'accidents.

Une métrique utilisée pour déterminer la dangerosité des accidents chez un constructeur est le **Score de gravité**. Calculé comme suit, on donne un poids de 4 au Total.Fatal.Injuries, 2 au Total.Serious.Injuries, 1 au Total.Minor.Injuries.

$$\text{Score de gravité} = \frac{4 * \text{Total.Fatal.Inju} + 2 * \text{Total.Serious.Inju} + 1 * \text{Total.Minor.Inju}}{\text{Total.Fatal.Inju} + \text{Total.Serious.Inju} + \text{Total.Minor.Inju}}$$

Si le score de gravité se rapproche de 4 ca veut dire que c'est très Dangereux, 2 moyennement dangereux, 1 faiblement dangereux. Cette même métrique sera utilisé dans les analyse par type de moteurs et par model.

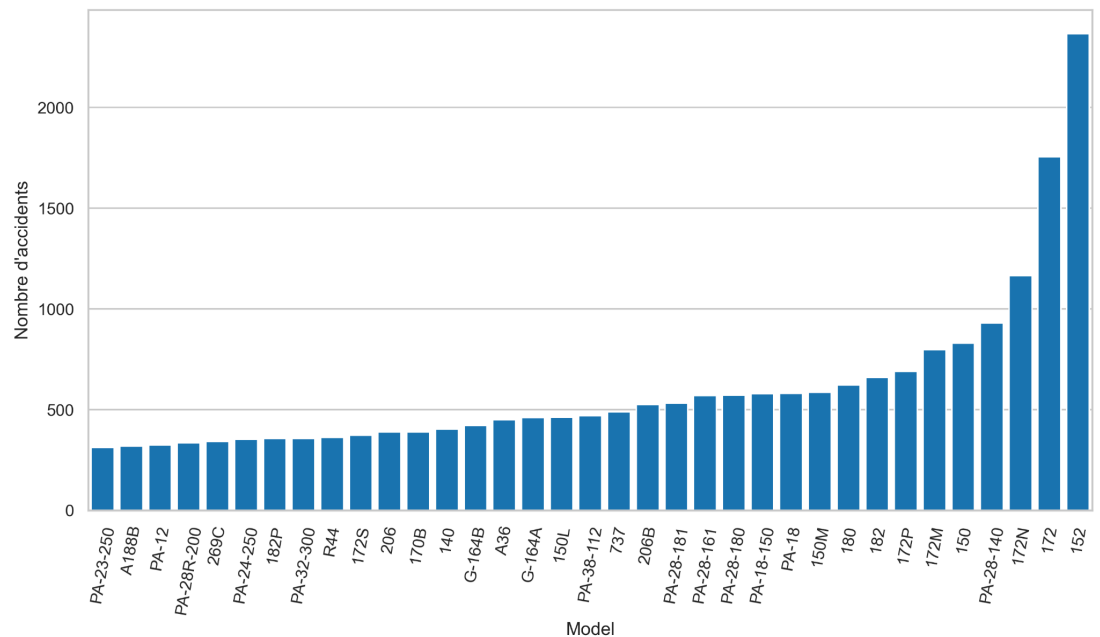
Ce le tableau ci-dessous fera un résumé des visuels précédent, affichera aussi le score de gravité de ces models et classé ces models par ordre croissant de leur score de gravité.

Make	Gravity Score	Nombre d'accidents	Taux mortels
Hiller	1.8857	348	8.62
Stinson	2.0478	439	16.17
Schweizer	2.0686	772	11.40
Hughes	2.0799	932	21.78
Taylorcraft	2.1220	383	20.37
Aeronca	2.1538	636	18.55
Luscombe	2.1883	414	17.15
Mcdonnell Douglas	2.2897	602	204.15
Grumman	2.3275	1172	21.16
Air Tractor	2.4267	691	17.51
Maule	2.4502	589	18.68
Bell	2.4615	2721	48.84
Champion	2.4715	519	26.01
Bellanca	2.5782	1045	33.01
Cessna	2.5784	27127	35.45
Aerospatiale	2.6864	388	89.95
Piper	2.7107	14860	44.87
De Havilland	2.7151	421	87.41
Mooney	2.7394	1334	51.35
Robinson	2.7913	1228	50.24
Aero Commander	2.9368	429	57.34
North American	2.9936	383	49.87
Beech	3.0006	5367	70.41
Rockwell	3.0455	354	55.37
Boeing	3.0633	2706	315.23

## 2. Par modèle (Model)

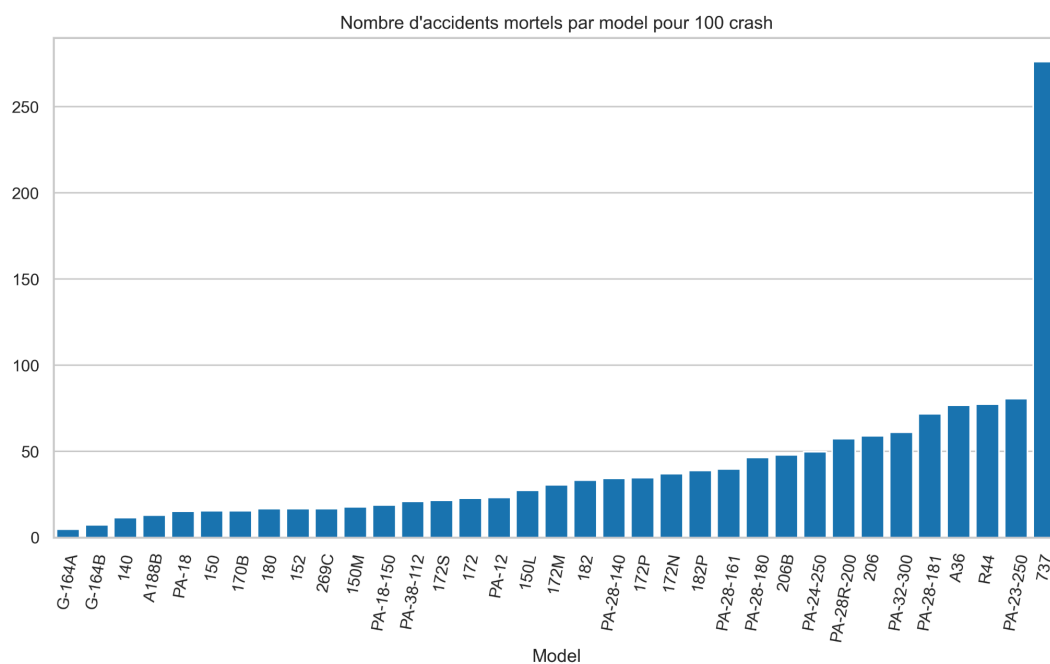
- Fréquence des accidents par modèles

**Objectif :** Identifier quels modèles ont été impliqués dans le plus grand nombre d'accidents, sans distinction initiale de gravité.



Les modèles comme **152, 172 et 172N** dominent ce classement toutes les trois des modèles du constructeur **Cessna**. Elles représentent à eux 3 environ 25 % des accidents.

- Taux d'accidents mortels et score de gravité



Ce graphique vient encore de confirmer qu' avoir le plus grand nombre de crash recensé ne veut pas dire que c'est le plus mortel.



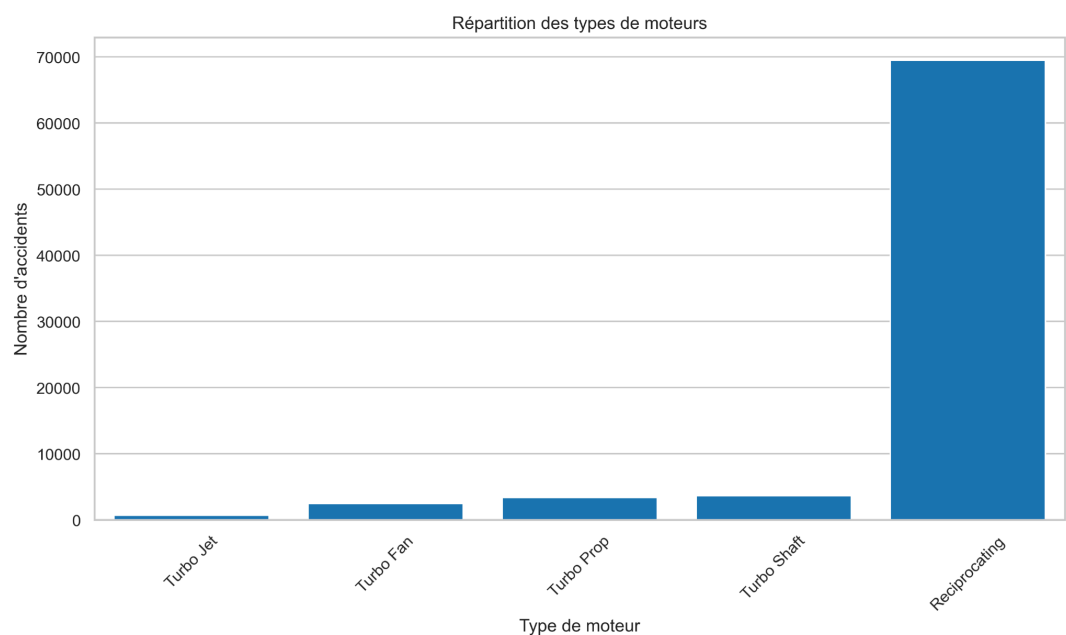
Tableau récapitulatif des modèles classés par ordre croissant du score de gravité

Model	Gravity Score	Nombre d'accidents	Taux mortels
G-164A	1.80	460	4.78
G-164B	1.97	420	7.14
140	2.10	402	11.44
150	2.11	829	15.32
269C	2.11	341	16.72
150M	2.17	585	17.61
PA-28-140	2.23	930	34.19
PA-18	2.24	581	15.15
A188B	2.29	318	12.89
170B	2.34	389	15.42
152	2.36	2367	16.56
172M	2.36	798	30.45
172	2.37	1755	22.79
150L	2.38	461	27.33
182P	2.39	356	38.76
206B	2.45	524	47.90
172P	2.47	689	34.54
PA-12	2.47	324	23.15
182	2.48	659	33.08
PA-38-112	2.49	469	20.90
172N	2.53	1164	36.94
180	2.53	622	16.56
PA-18-150	2.54	578	18.69
PA-28-180	2.58	572	46.33
PA-28-161	2.61	569	39.72
PA-32-300	2.64	356	60.96
PA-24-250	2.70	352	49.72
172S	2.74	373	21.45
206	2.79	389	58.87
PA-28-181	2.85	532	71.62
PA-28R-200	2.90	335	57.31
R44	2.94	361	77.29
A36	2.97	450	76.67
PA-23-250	3.05	311	80.39
737	3.43	488	276.23

### 3. Par Type de moteurs (Engine.Type)

- Fréquence des accidents par type de moteur

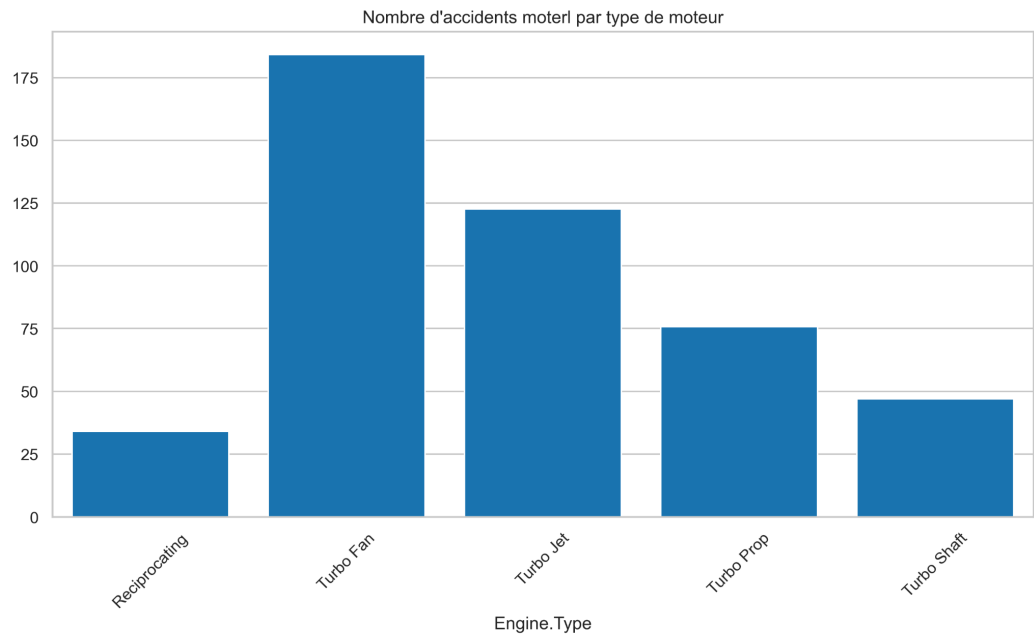
L'une des choses qui caractérise un moteur est bien sûr son moteur. Il est donc important de prendre en compte les types de moteur présentant le moins de risque que possible. Dans ce qui va suivre nous allons voir les types de moteurs présentant le moins de risque indépendamment du constructeur ou du modèle.



Lors de cette analyse trois types de moteurs sont ignorés à cause du peu d'information qu'on dispose, Moteur Hybrid Rocket avec 1 cas recensé, Electrique 10 et Geared Turbo Fan 12.

Les Reciprocating Engine dominent en terme de fréquence de d'incident avec plus de 87% à son actif. Au même lieu c'est un type de moteur très utilisé. Peut-être que ce sont aussi les plus sûrs, s'ils sont aussi utilisés. Poursuivons donc notre analyse pour voir ce qu'il en est.

- Taux d'accidents mortels et score de gravité



Bien que les moteurs Reciprocating ont eu le plus grand cas d'incident recensé, ce n'est pas le plus mortel contrairement à ce qu'on pourrait pensé. Il laisse la place au Turbo Fan avec moyenne de 185 morts pour 100 incidents contre 34 pour Reciprocating.

Voyons maintenant où se place chaque type de moteur selon leur score de gravité toujours avec la même métrique.

Engine.Type	Gravity Score	Nombre d'accidents	Taux mortels
Turbo Shaft	2.41	3607	46.96
Reciprocating	2.55	69484	33.99
Turbo Fan	2.58	2472	184.14
Turbo Jet	2.93	703	122.62
Turbo Prop	3.02	3390	75.69

Encore une fois le moteur Reciprocating reste très prometteur

## Croisement des résultats.

Les étapes vu plus haut montre que montrent que le fait q'un constructeurs, un modele ou un type de moteur a le plus grand cas d'incident recensé ne signifait pas pour autant que c'est le plus dangereux. Par exemple le modèle **152** de **Cessna** a le plus grand nombre d'incidents recensé mais fait partie des modèles avec le moins de morts en moyenne. Le plus judicieux à faire après ces analyses c'est de croiser les résultats pour voir quelle combinaison de constructeur, modèle et type de moteur présente le moins de risque.

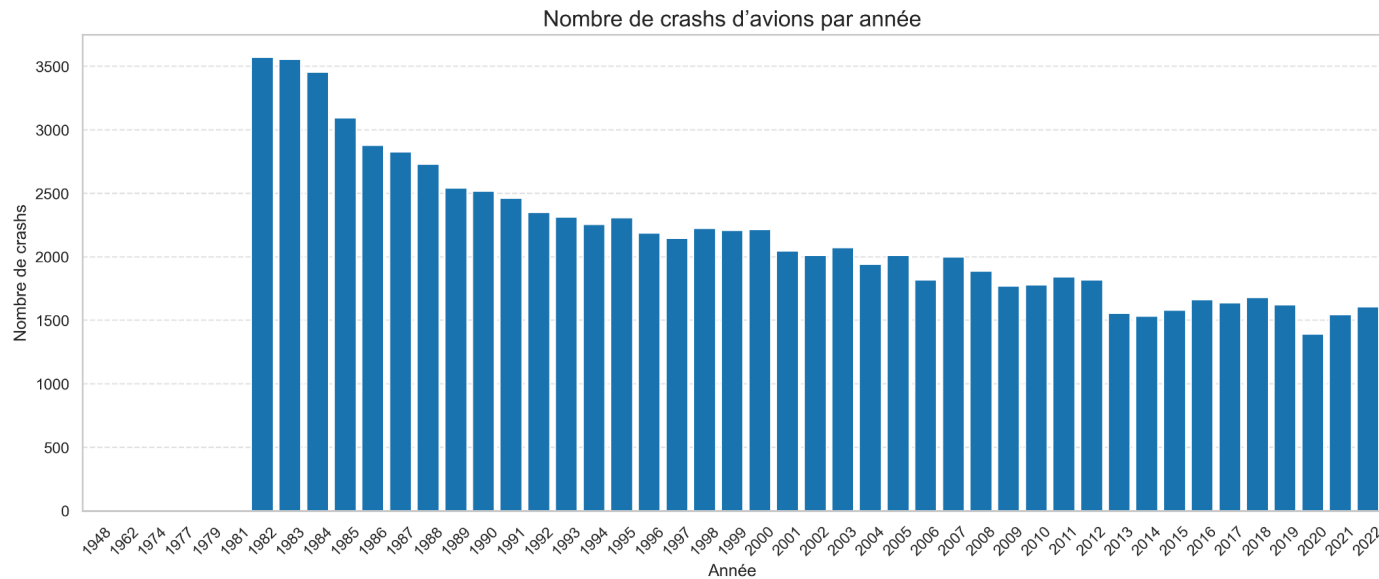
Le tableau suivant va afficher un classement 20 meilleurs candidats trier par constructeur, modèle puis type de moteur, tout en les triant par ordre croissant de leur score de gravité et taux mortel et par ordre décroissant de leur nombre d'accidents recensé. On va donc cherché parmi ceux qui ont le plus d'incidents recensé lesquels ont le bas score de gravité et de taux mortel.

Make	Model	Engine.Type	Nombre.d.accidents	Total.Fatal.Injuries	Total.Serious.Injuries	Total.Minor.Injuries	Taux.Mortel (%)	Score_gravite
Grumman	G-164A	Reciprocating	354	17	29	52	4.80	1.82
Cessna	150	Reciprocating	787	95	114	197	12.07	1.98
Cessna	140	Reciprocating	388	40	55	71	10.31	2.05
Cessna	150M	Reciprocating	580	101	66	150	17.41	2.16
Cessna	172	Reciprocating	1613	275	281	359	17.05	2.21
Piper	PA-28-140	Reciprocating	892	300	253	397	33.63	2.21
Cessna	A188B	Reciprocating	308	36	40	45	11.69	2.22
Piper	PA-18	Reciprocating	565	85	63	110	15.04	2.23
Cessna	182	Reciprocating	590	141	118	163	23.90	2.28
Cessna	152	Reciprocating	2322	363	191	420	15.63	2.31
Cessna	182P	Reciprocating	342	120	64	136	35.09	2.33
Cessna	172M	Reciprocating	772	228	167	240	29.53	2.34
Cessna	150L	Reciprocating	450	122	83	123	27.11	2.37
Bell	206B	Turbo Shaft	473	209	149	201	44.19	2.39
Cessna	170B	Reciprocating	381	59	45	55	15.49	2.40
Cessna	172P	Reciprocating	674	219	113	208	32.49	2.43
Piper	PA-12	Reciprocating	315	69	33	66	21.90	2.43
Piper	PA-38-112	Reciprocating	459	90	57	79	19.61	2.45
Cessna	172N	Reciprocating	1136	403	194	336	35.48	2.50
Piper	PA-18-150	Reciprocating	560	107	80	72	19.11	2.55

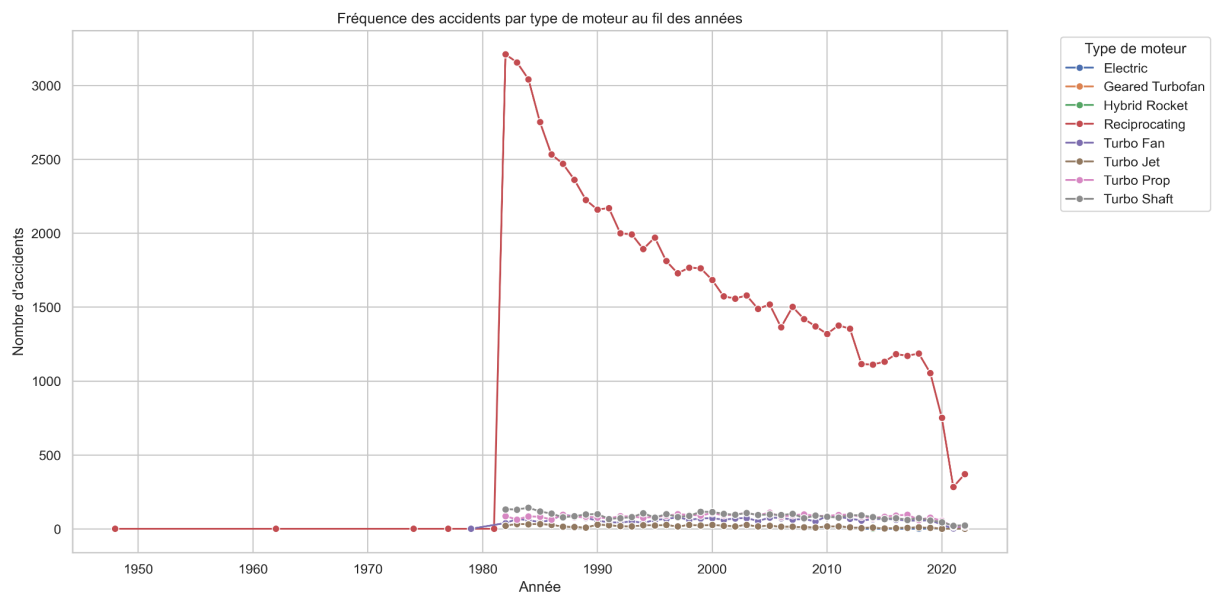
Ce tableau montre que parmi les accidents d'avions recensé, les constructeurs **Grumman, Cessna, Bell et Piper** dominent ce classement. De plus les moteurs **Reciprocating** sont ceux qui presente les meilleurs résultats representant 95% des vingts types d'avions présentant le moins de risques de lors accidents.

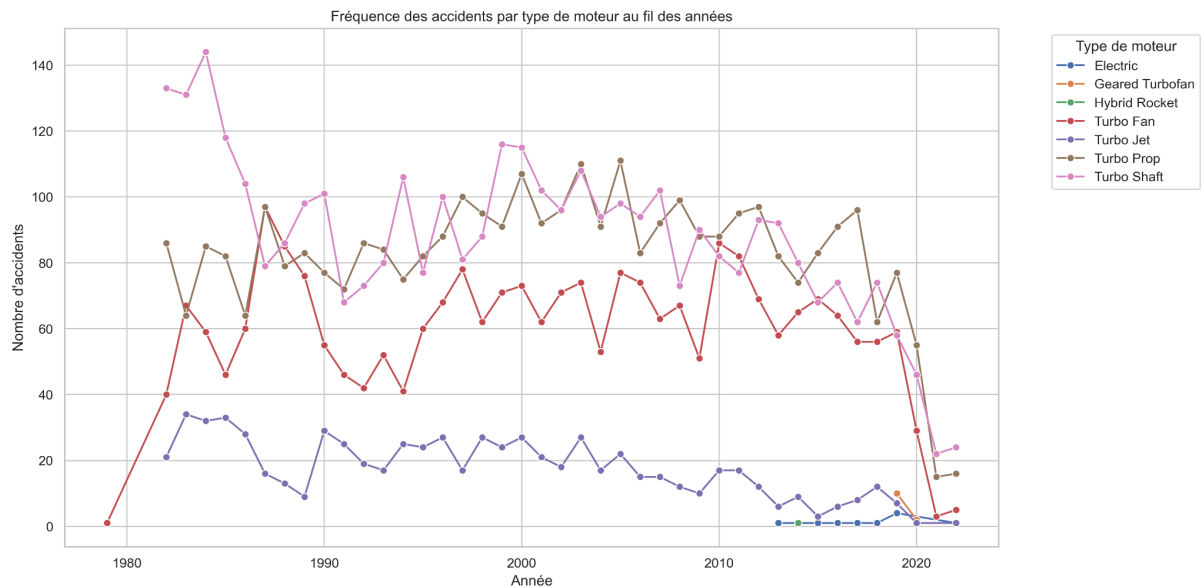
## Autres analyses

Dans les sections plus haut nous avons vu les types de moteurs présentant les moins de risque en se basant sur les monbres de morts, de blessés grave et mineurs. L'étape qui va suivre nous permettrons de voir l'évolution des constructeurs, des types de moteurs dans le temps et la fréquences des accidents avec le temps.

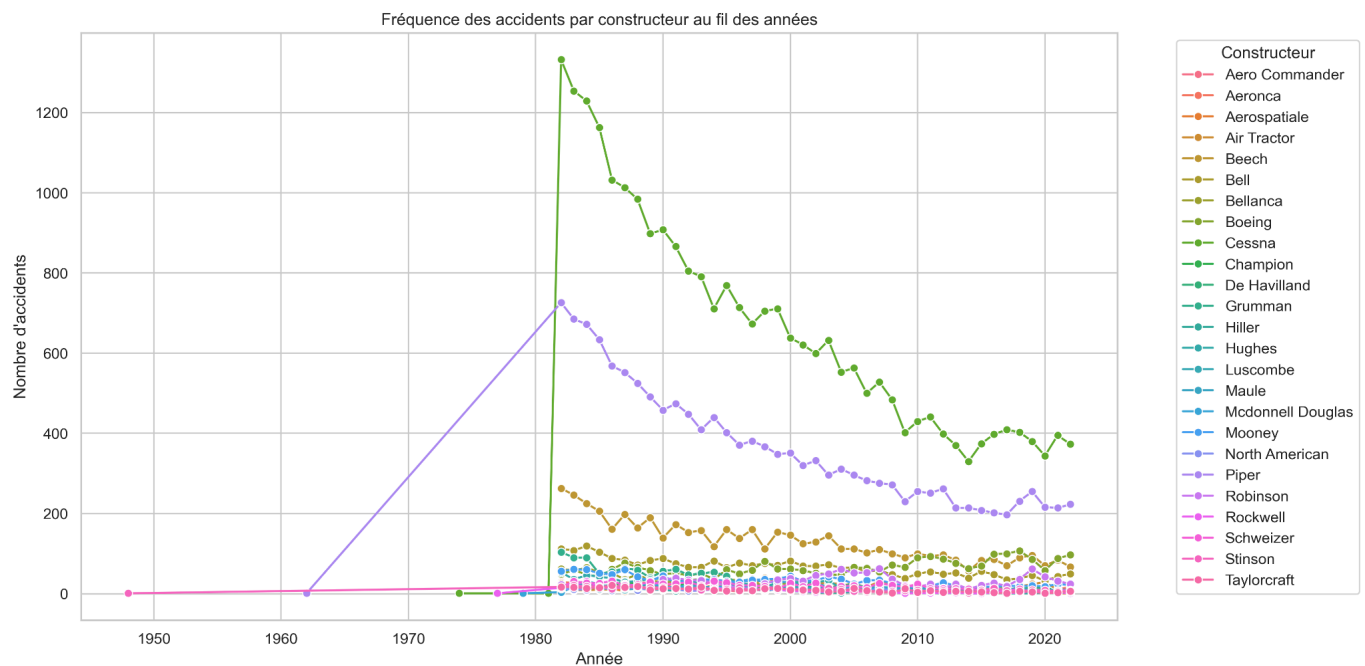


Globalement la fréquence des accidents d'avions diminue avec le temps. Ce qui peut traduire par des meilleurs protocoles de sécurité au fil du temps rendant les voyages plus sûr.





Ces deux graphiques montrent l'évolution de la fréquence des accidents par type de moteurs au fil des années à la seule différence près dans le deuxième le moteur Recoprocating est filtré juste par soucis de visibilité. Mais elles montrent que la fréquence des accidents diminue avec le temps, potentiellement grâce à une certaine maturité des types de moteurs.



La même tendance se confirme pour les constructeur. Une tendance baissiere de la fréquence des accidents avec le temps.

# Mes recommandation

## Résumé des principales observations

- Les moteurs Reciprocating et Turbo Shaft affichent les taux de mortalité les plus faibles.
- Certains modèles d'avions comme le 152 et le 172 du constructeur Cessna ont une gravité d'accidents très basse, même avec un nombres d'incidents relativement élevé – Preuve de robustesse.
- Les constructeur Cesnna, Piper montrent un historique d'accidents maîtrisé avec une tendance à la baisse ces 10 dernières années.

## Recommandations stratégiques

1. Privilégier les modèles à moteur Reciprocating
  - Ce sont les moteurs qui présentent les meilleurs scores de sécurité sur le long terme.
  - Idéal pour les petits trajets régionaux.
2. Éviter (ou surveiller) les modèles avec moteurs *Turbo Fan* dans une première phase
  - Ce type de moteur présente une gravité d'accident très élevée. Il est plus courant sur les gros porteurs.
  - À réserver pour une phase d'expansion, avec un budget sécurité + formation conséquent.
3. Choisir des modèles ayant un historique d'usage dans des contextes similaires
  - Rechercher des modèles utilisés dans des régions ou segments similaires pour minimiser les incertitudes.

# Limite de l'analyse et mise en garde

## 1. Incomplétude des données

- Certains enregistrements manquent d'informations essentielles comme le type de moteur, le modèle exact, ou encore les conditions météo.
- Ces valeurs manquantes peuvent biaiser certaines statistiques ou masquer des tendances réelles.

## 2. Données historiques non équilibrées

- La période couverte par le jeu de données (1962–2023) inclut des changements majeurs dans l'aviation (évolutions technologiques, réglementations, sécurité, etc.).
- Les modèles d'avions plus anciens ont eu plus de temps pour accumuler des incidents, ce qui peut fausser le calcul des taux d'accidents si on ne prend pas en compte la date de mise en service.

## 3. Fréquences d'utilisation inconnue

- Le dataset n'indique pas combien d'avions de chaque modèle sont en circulation, ni combien d'heures de vol ont été effectuées.
- Donc un modèle peut avoir plus d'accidents simplement parce qu'il est plus utilisé, pas forcément parce qu'il est moins sûr.

## 4. Événements extrêmes

- Certains accidents peuvent être dus à des facteurs externes non liés à l'avion (erreur humaine, sabotage, météo extrême, etc.).
- Ces événements peuvent fausser les scores de gravité ou de mortalité pour un modèle donné.



## Mes infos de contact

E-mail : [mackensonsh@gmail.com](mailto:mackensonsh@gmail.com)

Telephone : +509 46-98-7948 / +509 42-64- 8619

Linkedin : [www.linkedin.com/in/mackenson-joseph-186492222](https://www.linkedin.com/in/mackenson-joseph-186492222)

Github : <https://github.com/MackSH>