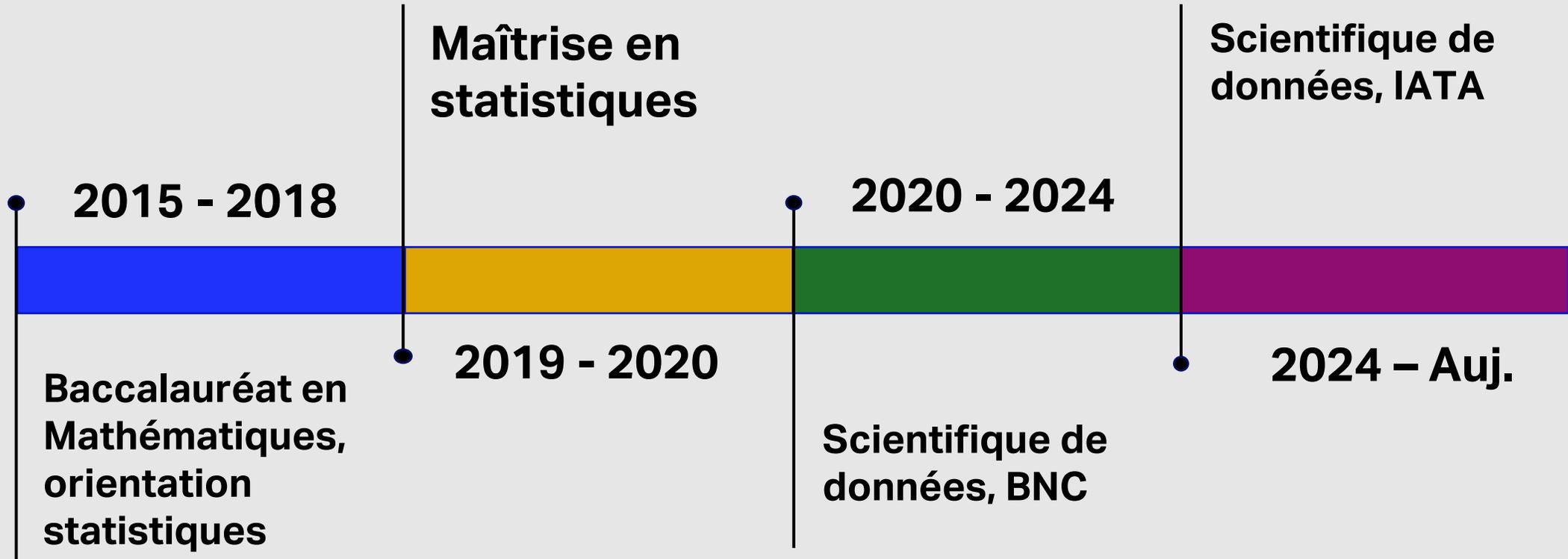


# De la donnée brute aux insights, l'impact de la science des données sur la sécurité aérienne

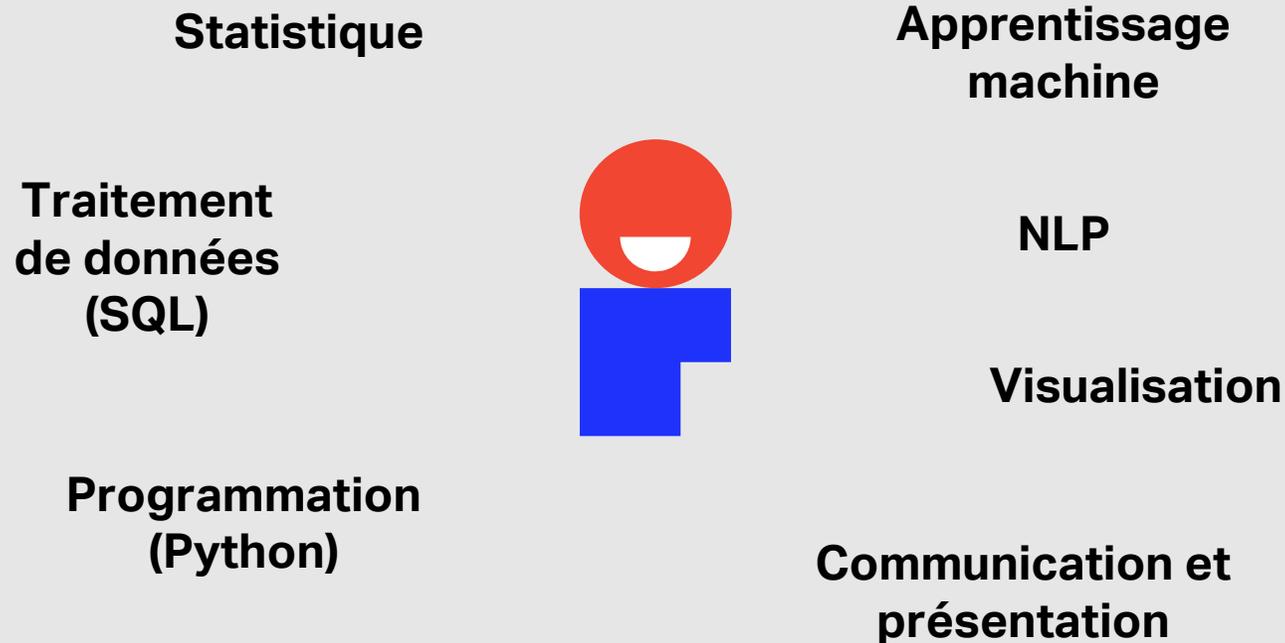
Katia Badaoui, Scientifique de données, IATA

# Mon parcours



# Qu'est-ce qu'un(e) scientifique de données

La **science des données** est l'étude des données afin d'en **extraire des informations significatives** pour les entreprises. Il s'agit d'une approche pluridisciplinaire qui combine des principes et des pratiques issus des domaines des mathématiques, des statistiques, de l'intelligence artificielle et du génie informatique, en vue d'analyser de grands volumes de données



To represent,  
lead and serve  
the airline  
industry

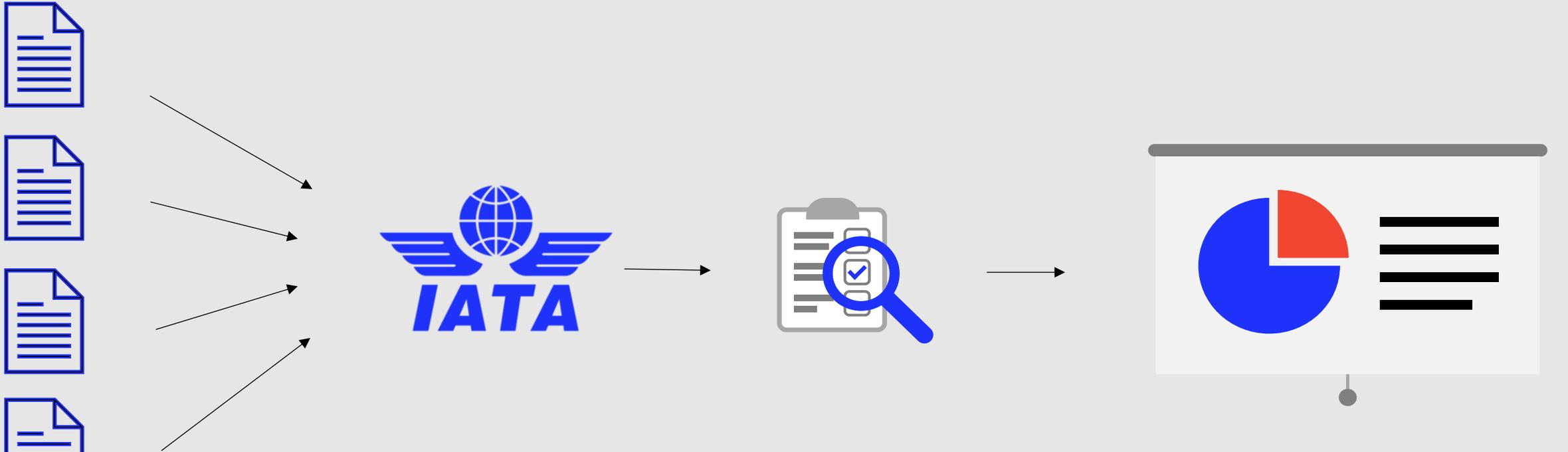
# Mission



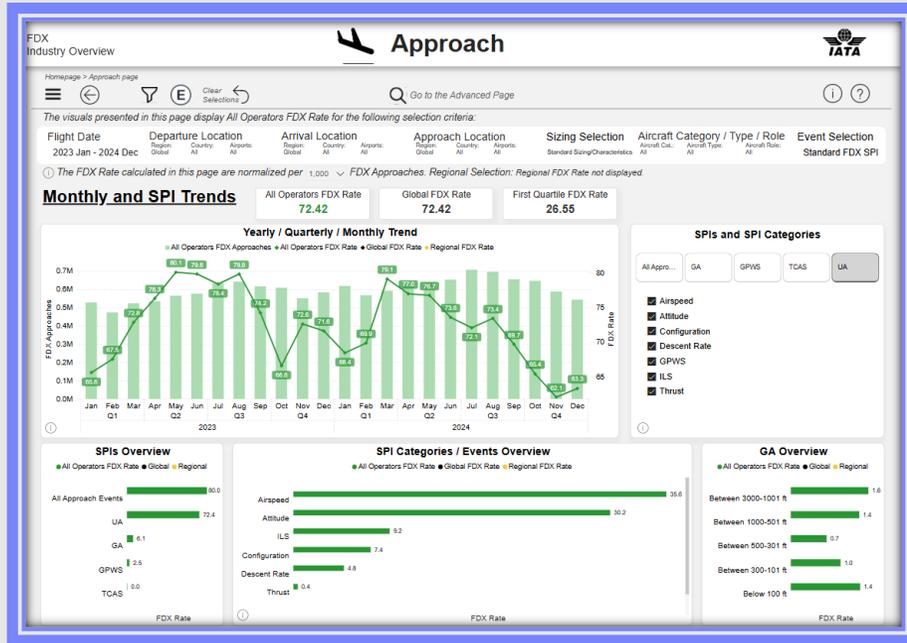
# La sécurité aérienne et les données

## La sécurité est la priorité numéro 1 de IATA

Programmes d'échanges de données

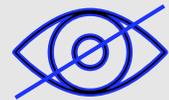


# Exemple 1 de projets: "actionable insights"



FLIGHT_PHASE	Operator	SPI Level Selection	time interval	My Trend %	Industry Trend %	top aircraft
Takeoff	XYZ	SPI: TCAS	12	1,898.76 ↑	-19.11 ↓	Aircraft X
Takeoff	XYZ	global	3yoy	464.47 ↓	19.40 ↑	Aircraft Y
Takeoff	XYZ	SPI: All Takeoff Events	3yoy	464.47 ↓	19.40 ↑	Aircraft X
Takeoff	XYZ	SPI: GPWS	3	339.29 ↓	4.26 →	Aircraft Z
In-Flight	XYZ	SPI: GPWS	3yoy	280.72 ↓	8.93 →	Aircraft Y
Takeoff	XYZ	SPI: GPWS	12	251.77 ↓	27.38 ↑	Aircraft X
In-Flight	XYZ	SPI: GPWS	12	240.22 ↓	11.70 →	Aircraft X
Takeoff	XYZ	global	3	136.24 ↓	4.35 →	Aircraft Z
Takeoff	XYZ	SPI: All Takeoff Events	3	136.24 ↓	4.35 →	Aircraft W

**Unstable approach** events increase of **12%** in the last **12 months**, compared to an **8.3%** decrease at the industry level. The main driver of this increase is **Rate of descent high 500-50ft.**



- Les tableaux de bord offrent une multitude d'informations ; il peut être difficile de savoir où prêter attention.
- Certaines tendances sont cachées.
- Il peut être difficile de comprendre les facteurs associés à l'augmentation d'une tendance.



Les "Actionable insights" peuvent aider à:

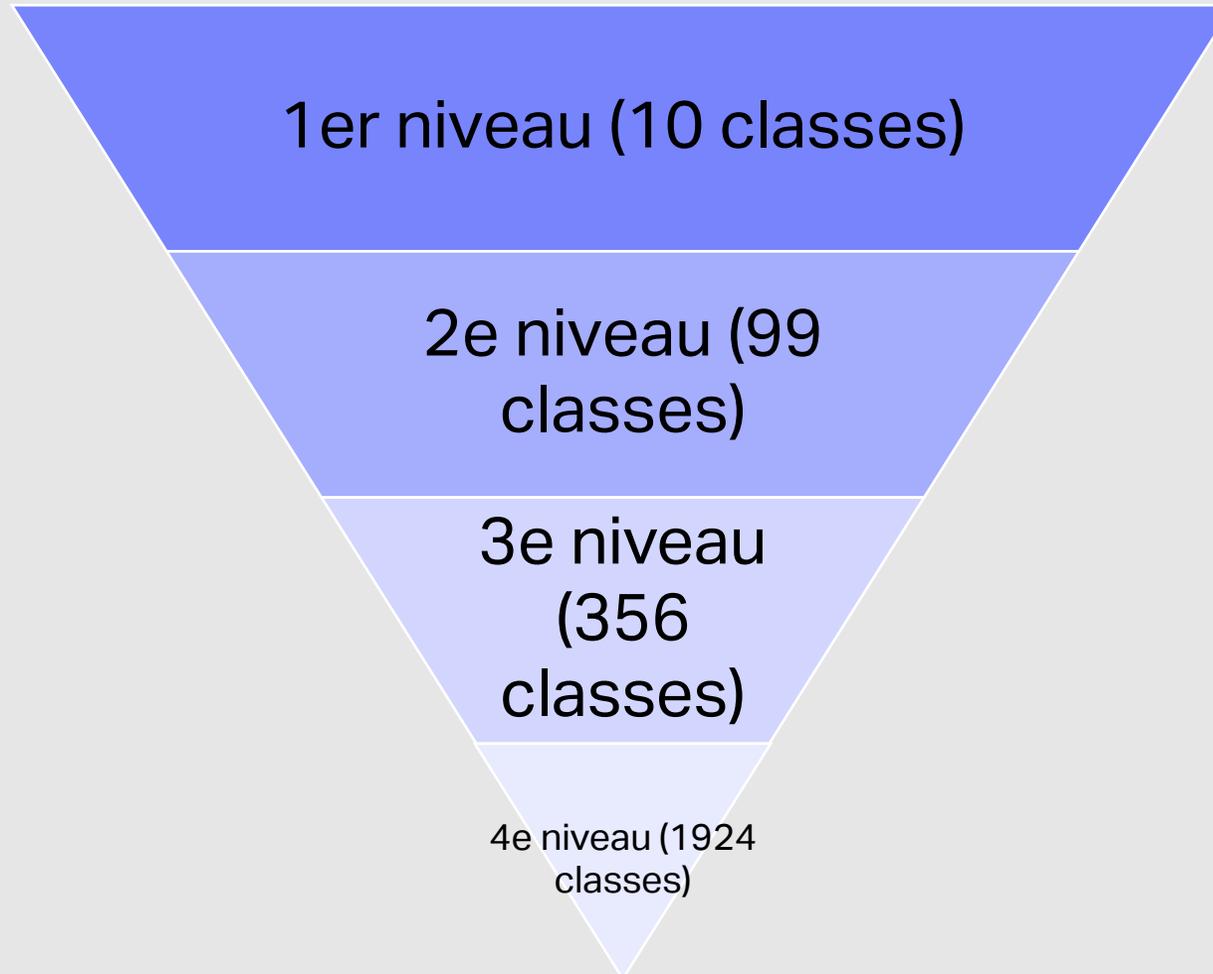
- Démystifier les tendances
- Révéler les patterns cachés.
- Mettre en évidence les tendances positives et négatives, leur impact et leur importance



# Exemple 2 de projets: Modèle de classification de texte

## Une taxonomie

Les données d'incidents signalées par les compagnies aériennes présentent une multitude de structures. Pour partager les informations issues de ces données, il est nécessaire de les uniformiser en une seule structure (taxonomie).



**La taxonomie utilisées**  
4 niveaux,  
Un total de **~2800**  
combinaisons possibles

## Exemple:

1. Sécurité en cabine ;
2. Passager indiscipliné ;
3. Non-conformité ;
4. Ceinture de sécurité

# Exemple 2 de projets: Modèle de classification de texte

## Défis du projet

Il existe près de **2800** combinaisons possibles pour classer un incident :

**Confusion** – Avec des milliers de choix, le modèle commet **davantage d’erreurs**. Même pour les humains, cette tâche peut s’avérer complexe.

**Catégories similaires** – Certaines classes peuvent être très similaires, ce qui les rend difficiles à distinguer.

**Manque d’exemples** – Si certaines catégories apparaissent rarement dans les données d’entraînement, le modèle ne les apprend pas correctement.

# Exemple 2 de projets: Modèle de classification de texte

## Solution: Roberta, un modèle de classification de texte

La solution est construite de manière itérative à l'aide de 3 modèles

### **Modèle 1 : 2e niveau de la taxonomie**

Intrants pour le finetuning: descripteurs envoyés par la compagnie aérienne concaténés

### **Modèle 2 : 3e niveau de la taxonomie**

Intrants pour le finetuning : **Output du 1er modèle** + descripteurs envoyés par la compagnie aérienne concaténés

### **Modèle 3 : 4e niveau de la taxonomie**

Intrants pour le finetuning : **Output du 1er modèle** + **Output du 2e modèle** + descripteurs envoyés par la compagnie aérienne concaténés

# Conclusion : ce que vos études vous apportent... et au-delà

## Compétences académiques directement applicables

- Séries chronologiques
- Régressions et modèles prédictifs
- Tests statistiques
- Comprendre les biais statistiques
- etc

Au-delà de ces méthodes et algorithmes, vos études vous préparent aussi à : **savoir chercher, apprendre et résoudre des problèmes.**

### **Comprendre des recherches et s'adapter à de nouveaux sujets**

Lire et interpréter des articles scientifiques, comprendre des approches nouvelles

Apprendre à implémenter des méthodes avancées que vous n'avez pas vues en cours

### **S'attaquer à des problèmes sans solution toute faite**

Identifier les bonnes méthodes et adapter celles qui existent

Expérimenter, tester différentes approches et apprendre de ses erreurs

### **Développer une pensée critique et une rigueur scientifique**

Vérifier ses hypothèses, remettre en question ses résultats

Prendre du recul pour éviter des conclusions biaisées

# Questions?

