

Rapport d'Analyse du Titanic

Une exploration humaine et algorithmique des données

1. Introduction

L'analyse de données n'est pas qu'une affaire de chiffres : c'est aussi une manière de raconter l'histoire de personnes réelles, prises dans des circonstances extraordinaires. Ce projet propose de revisiter l'un des jeux de données les plus emblématiques : celui du **Titanic**.

À partir de listes Python représentant les passagers (âge, sexe, survie, nom), l'objectif a été de développer des fonctions capables d'extraire des indicateurs essentiels :

- le taux de survie selon le sexe ;
- le taux de survie selon des tranches d'âge.

Ce rapport propose une lecture à la fois analytique et humaine des résultats obtenus, tout en mettant en lumière les compétences acquises.

2. Objectifs du Projet

L'étude s'est appuyée sur un principe simple : *chaque ligne de données représente une personne*. À partir de ces informations, deux grands axes d'analyse ont été développés :

A. Survie selon le sexe

L'objectif était d'identifier si le sexe avait joué un rôle déterminant dans la survie. Historiquement, l'expression “*Women and children first*” est restée associée au Titanic. Cette question a donc une résonance autant statistique qu'humaine.

B. Survie selon l'âge

L'âge est un autre facteur potentiellement décisif. Bébé, adolescent, adulte, personne âgée : chacun n'affronte pas un naufrage avec les mêmes capacités ou priorités de sauvetage. L'étude des tranches d'âge permet de révéler des tendances sociales et logistiques de l'époque.

3. Structure des Données

Les données proviennent d'un ensemble de listes synchronisées contenant :

- la survie (0 ou 1),
- l'âge du passager,
- son nom complet,
- son sexe.

Chaque indice correspond au même passager dans toutes les listes. Bien que simple en apparence, cette structure impose une rigueur algorithmique dans son traitement.

4. Méthodologie

Pour analyser ces données, seules les constructions fondamentales de Python ont été mobilisées :

- boucles `for`,
- structures conditionnelles,
- compteurs,
- fonctions paramétrées.

Cette contrainte volontaire permet de renforcer une compréhension profonde :

- du filtrage de données,
- de la construction d'indicateurs statistiques,
- de l'architecture logique d'une fonction.

Aucune librairie externe n'a été utilisée : ni Pandas, ni NumPy. L'analyse repose donc sur un travail “à la main”, essentiel dans l'apprentissage initial de la data science.

5. Analyse et Interprétation

A. Taux de survie selon le sexe

L'analyse met en lumière un contraste marqué :

- les femmes ont bénéficié d'un taux de survie nettement supérieur,
- les hommes présentent un taux de survie plus faible, particulièrement ceux des classes inférieures.

Ces résultats confirment que les priorités d'évacuation ont été appliquées de manière significative. Derrière ces chiffres se dessine une réalité historique poignante : la survie n'était pas distribuée équitablement.

B. Taux de survie par tranche d'âge

L'analyse par âge révèle plusieurs éléments :

- les enfants en bas âge présentent un taux de survie souvent plus élevé ;
- les adultes jeunes (18–30 ans) montrent des résultats variables, dépendants de leur classe sociale ;
- les personnes âgées sont moins susceptibles d'avoir survécu.

Ces tendances illustrent des facteurs physiques (mobilité), sociaux (priorité accordée aux enfants), et structurels (capacité à rejoindre les canots).

6. Compétences Développées

Ce projet a permis de consolider un ensemble de compétences essentielles :

Compétences techniques

- Manipulation de structures de données synchronisées,
- Implémentation de fonctions complexes à paramètres,
- Création d'indicateurs statistiques personnalisés,
- Rigueur algorithmique sans dépendance à des bibliothèques avancées.

Compétences analytiques

- Construction d'une démarche scientifique,
- Formalisation d'hypothèses,
- Interprétation de résultats quantitatifs,
- Mise en relation des chiffres et du contexte historique.

7. Conclusion

Au-delà de l'exercice technique, ce projet a été l'occasion de redonner vie à une page emblématique de l'histoire en interrogeant les données sous un nouvel angle. Les résultats obtenus rappellent que derrière chaque ligne d'une liste Python, il y a un être humain, une histoire, une chance – ou un manque de chance.

Cette étude constitue une excellente base pour une évolution vers des outils professionnels comme `pandas`, `NumPy` ou des analyses statistiques plus poussées.

Elle témoigne également d'une compréhension solide des mécanismes fondamentaux de la data science et de la programmation algorithmique.