

Ce projet a été réalisé à l'initiative personnelle dans le cadre de mon développement de compétences en Analyse de données et dans le but d'enrichir mon portfolio professionnel.

Analyse des Données du Titanic

Objectif : Créer un tableau de bord interactif pour explorer les facteurs de survie des passagers

André Mamdouba Bangoura

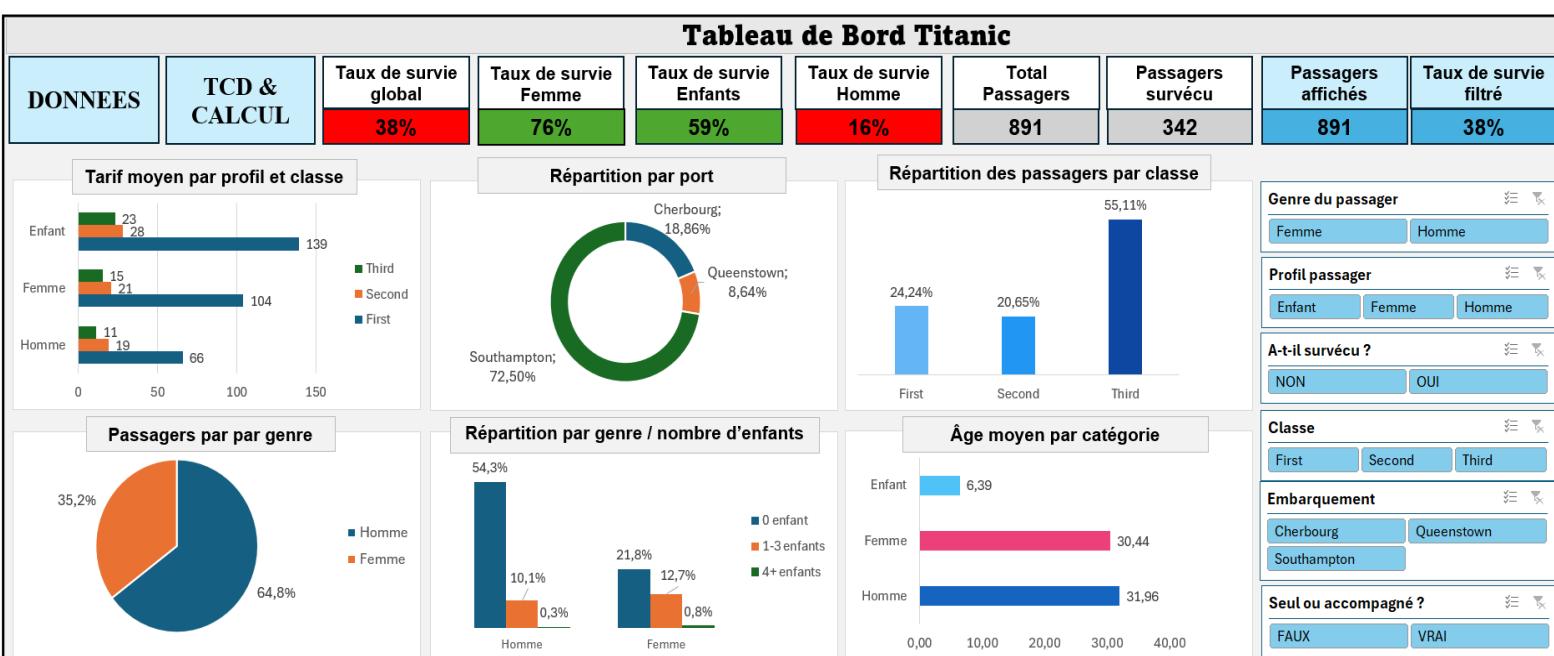
Table des matières

Tableau de Bord Titanic - Analyse des Données	2
A propos :	2
1. Objectif :.....	3
2. Source des données :.....	3
3. Outils utilisés :	3
4. Méthodologie et transformations :	3
4.1 Nettoyage :	3
4.2 Transformations :	4
5. Analyse et indicateurs :.....	4
6. Visualisation :	4
7. Résultat final	4
8. Conclusion de l'analyse Titanic	5
8.1 Le facteur clé : le profil du passager	5
8.2 Le rôle de la classe sociale	5
8.3 L'embarquement et le prix du billet.....	6
8.4 Famille et survie	6
9. RESUME :	6
10. ANNEXES :	7

Tableau de Bord Titanic - Analyse des Données

A propos :

Ce projet a été réalisé à l'initiative personnelle dans le cadre de mon développement de compétences en Analyse de données et dans le but d'enrichir mon portfolio professionnel.



A partir des données ouvertes du Titanic, j'ai conçu ce tableau de bord interactif sous Excel, en appliquant une démarche complète d'analyse de données :

- Nettoyage des données et traitement des valeurs manquantes
- Transformation des variables pour enrichir l'analyse
- Exploration statistique et création d'indicateurs clés
- Visualisation dynamique via tableaux croisés, graphiques et filtres interactifs

Ce projet démontre ma capacité à :

- ✓ Travailler de manière autonome
- ✓ Structurer une démarche d'analyse pertinente

- ✓ Utiliser efficacement Excel, Python et les TCD
- ✓ Produire des visualisations professionnelles adaptées à la prise de décision

👉 Ce travail s'intègre pleinement dans ma démarche de montée en compétences vers un poste de Data Analyst, et reflète ma motivation à progresser dans le domaine de la data.

1. Objectif :

Identifier les facteurs influençant la survie des passagers (sexe, classe, port d'embarquement, situation familiale, etc.), et d'en extraire des insights visuels compréhensibles rapidement.

2. Source des données :

Jeu de données Titanic (version Seanborn), disponible dans les bibliothèques de données comme Seanborn :

```
# Exemple : jeu de données Titanic (via seaborn)
import seaborn as sns
df = sns.load_dataset('titanic')

# Export en CSV pour Excel
df.to_csv('titanic.csv', index=False)
```

3. Outils utilisés :

Python (Seanborn pour récupérer les données)

Excel (TCD, Filtres, graphiques dynamiques)

4. Méthodologie et transformations :

4.1 Nettoyage :

- Etat des données manquantes :
 - ‘age’ : 177 valeurs manquantes
 - ‘embarked’ : 2 valeurs manquantes
 - ‘embark_town’ : 2 valeurs manquantes
 - ‘deck’ : 688 valeurs manquantes (beaucoup)
- Méthodes statistiques pour remplir :
 - Age :

Remplir par la médiane par groupe (ex : par sex et par classe), car l'âge suit souvent une distribution asymétrique et dépend de la classe et du sexe.

- **Embarked / Embark_town :**

Remplacer par la valeur la plus fréquente (mode) car il n'y a que 2 manquants.

- **Deck :**

688 manquants sur ~900 lignes → trop de valeurs manquantes (76%). Colonne supprimer.

4.2 Transformations :

- Traduction des colonnes (ex. : sex → sexe, survived → survécu)
- Création d'une variable qui : Homme / Femme / Enfant
- Catégorisation du nombre d'enfants (0, 1-3, 4+)
- Détection des passagers seuls (seul = True/False)

5. Analyse et indicateurs :

- Taux de survie global et par profil
- Répartition des passagers par classe, port, sexe
- Moyenne d'âge et de tarif par catégorie
- Analyse croisée : port + classe + genre, configuration familiale

6. Visualisation :

- Tableaux croisés dynamiques Excel
- Graphiques dynamiques (Histogrammes, barres, camemberts, Anneau)
- Filtres interactifs (segments) et indicateurs colorés

7. Résultat final

Un tableau de bord clair, interactif et filtrable permettant d'analyser rapidement :

- ✓ Qui a survécu et pourquoi ?

- ✓ Quel rôle ont joué la classe, le sexe ou la famille ?
- ✓ Existe-t-il un lien entre le prix du billet et les chances de survie ?

8. Conclusion de l'analyse Titanic

Cette analyse basée sur les données du naufrage du Titanic met en évidence plusieurs insights clés en matière de survie, de profil des passagers et de disparités structurelles à bord :

⌚ 8.1 Le facteur clé : le profil du passager

- Le **taux de survie global** s'élève à **38 %**.
 - Les **enfants (59%)** et **les femmes (76%)** ont survécu en bien plus grand nombre que les **hommes (16%)**, illustrant clairement la règle "Women and children first".
 - La répartition par profil (Homme, Femme, Enfant) permet une analyse plus fine qu'un simple découpage par sexe.
-

✿ 8.2 Le rôle de la classe sociale

- Les passagers de **1ère classe** représentent **24 % des passagers**, mais concentrent une part élevée des survivants.
- Les taux de survie varient significativement :
 - 1ère classe : **62 %** de survie
 - 2e classe : **47 %**
 - 3e classe : **24 %**
- Les passagers des classes inférieures étaient moins susceptibles de survivre, probablement en raison d'un accès plus difficile aux canots de sauvetage.



8.3 L'embarquement et le prix du billet

- Southampton concentre 72 % des embarquements, avec un taux de survie plus faible.
 - Les passagers embarqués à Cherbourg affichent de meilleurs taux de survie, liés à une plus grande proportion de femmes de 1ère classe.
 - Le tarif moyen du billet varie fortement selon le profil et la classe :
 - Femmes en 1ère classe : plus de 106 unités
 - Hommes en 3e classe : environ 12 unités
-



8.4 Famille et survie

- Les passagers voyageant seuls représentent 60 %, avec un taux de survie inférieur à ceux accompagnés.
 - Les femmes voyageant avec 1 à 3 enfants ont un taux de survie supérieur aux autres groupes familiaux.
-

9. RESUME :

Cette analyse démontre l'intérêt de croiser plusieurs dimensions (sexe, âge, classe, embarquement, configuration familiale) pour comprendre les comportements de survie.

Elle met également en lumière des inégalités structurelles à bord du Titanic, renforcées par la classe sociale et le genre, confirmant ainsi les priorités d'évacuation et les inégalités d'accès aux ressources de secours.

10. ANNEXES :

Feuille de calcul et Tableaux croisés dynamiques Excel

		DONNEES			TABLEAU DE BORD	
Nombre de passagers	891	Étiquettes de colonne ▾			Étiquettes de lignes ▾ Passagers (%)	
	891	NON OUI Total général			Homme 64,8%	
	Taux de survie	61,62% 38,38% 100,00%			Femme 35,2%	
	Taux de survie	62%	38%		Total général 100,0%	
Nombre de Passager Survécu	342	Étiquettes de lignes ▾ Moyenne de age			Étiquettes de lignes ▾ Passagers (%)	
Taux de survie Global	38%	Homme 31,96			First 24,24%	
Nombre de Passager	891	Femme 30,44			Second 20,65%	
Nombre de Passager Filtré	891	Enfant 6,39			Third 55,11%	
Taux de survie Global Homme	16%	Total général 29,11			Total général 100,00%	
Taux de survie Global Femme	76%					
Taux de survie Global Enfants	59%	Moyenne de tarif Étiquettes de lignes ▾			Étiquettes de lignes ▾ First Second Third Total général	
Étiquettes de ligne: ▾ Passagers (%)	Cherbourg 18,86%	Homme 66			Homme 66 19 11 25	
	Queenstown 8,64%	Femme 104			Femme 104 21 15 47	
	Southampton 72,50%	Enfant 139			Enfant 139 28 23 33	
Total général	100,00%	Total général 84			Total général 84 21 14 32	
seul	(Tous)					
Passagers	Étiquettes d' ▾					
Étiquettes de ligne: ▾	0 enfant 1-3 enfants 4+ enfants Total général					
Homme	54,3%	10,1% 0,3% 64,8%				
Femme	21,8%	12,7% 0,8% 35,2%				
Total général	76,1%	22,8% 1,1% 100,0%				