

## Étape 6 : Conclusion et Documentation

### Conclusion :

Après avoir analysé les données du Titanic, nous pouvons tirer les conclusions suivantes :

1. **Nombre de femmes de moins de 18 ans ayant survécu :** Neuf (9) femmes de moins de 18 ans ont survécu au naufrage du Titanic.
2. **Répartition par classe parmi les femmes de moins de 18 ans ayant survécu :**
  - 7 femmes de moins de 18 ans ayant survécu étaient en classe 1.
  - 1 femme de moins de 18 ans ayant survécu était en classe 3.
  - 1 femme de moins de 18 ans ayant survécu était en classe 2.
3. **Influence du port d'embarquement sur la survie :**
  - À partir du port de départ 'C', 48 passagers ont survécu et 17 sont décédés.
  - À partir du port de départ 'Q', 1 passager a survécu et 1 est décédé.
  - À partir du port de départ 'S', 73 passagers ont survécu et 42 sont décédés.
4. **Répartition par sexe et par âge des passagers :** Le graphique montre la répartition par sexe et par âge des passagers. On peut voir qu'il y avait un nombre significatif de passagers masculins et féminins dans différentes tranches d'âge.

### Documentation du Pipeline DataOps :

Nous avons mis en place un pipeline DataOps pour traiter les données du Titanic en suivant les étapes suivantes :

1. **Récupération des données :** Les données brutes ont été récupérées à partir d'un fichier CSV en utilisant la bibliothèque **pandas**.
2. **Création du modèle de données :** Nous avons créé un modèle de données JSON à partir des données brutes en sélectionnant les colonnes pertinentes.
3. **Nettoyage et formatage des données :** Les données ont été nettoyées en supprimant les valeurs manquantes, en convertissant les types de données et en supprimant les valeurs aberrantes.
4. **Enregistrement des données :** Les données nettoyées ont été enregistrées dans un nouveau fichier CSV.
5. **Analyse des données :** Nous avons répondu aux questions posées en utilisant les données nettoyées et avons généré des visualisations pour la répartition par sexe et par âge des passagers.
6. **Conclusion :** Une conclusion a été tirée en fonction des résultats de l'analyse.