

## Aufgabe: Analyse der Überlebenden auf der Titanic

Wir wollen herausfinden, wie sich Alter, Klasse und Familiengröße auf die Überlebenschance ausgewirkt haben.

### Ziel:

1. Daten untersuchen (unique, value\_counts).
2. Neue Spalten an spezifischen Positionen erstellen (insert).
3. Das Alter in Kategorien einteilen (cut).
4. Statistiken pro Gruppe erstellen (groupby, agg).
5. Eine Übersichtstabelle der Überlebensraten erstellen (pivot\_table).

## Erklärungen der verwendeten Methoden

Hier ist eine kurze Zusammenfassung, warum wir welche Methode benutzt haben:

1. **unique()**: Gibt uns ein Array aller einzigartigen Werte in einer Spalte. Perfekt, um zu sehen, welche Kategorien (z.B. Städte, Decks) existieren, ohne Duplikate.
2. **value\_counts()**: Zählt, wie oft jeder Wert vorkommt. Das ist der schnellste Weg, um die Verteilung von Daten zu prüfen (z.B. "Wie viele Männer vs. Frauen?").
3. **insert(loc, column, value)**: Im Gegensatz zur normalen Zuweisung (df['neu'] = ...), die die Spalte immer ans Ende hängt, können wir mit insert bestimmen, wo die Spalte stehen soll (hier an Index 7).
4. **cut()**: Ein mächtiges Tool, um kontinuierliche Zahlen (wie Alter oder Preis) in "Bins" (Schubladen) zu stecken. Es wandelt Zahlen in Kategorien um.
5. **groupby() & agg()**:
  - groupby zerlegt die Daten in Stapel (hier nach Altersgruppe).
  - agg (aggregate) erlaubt es uns, mehrere Berechnungen gleichzeitig durchzuführen (z.B. Mittelwert für Preis UND Summe für Überlebende).
6. **pivot\_table()**: Erstellt eine Excel-ähnliche Kreuztabelle. Sie ist übersichtlicher als ein groupby, wenn man zwei Dimensionen (Zeilen & Spalten) vergleichen will (hier Klasse vs. Geschlecht).

## Challenge

"Wie hoch ist der durchschnittliche Fahrpreis (fare) für Alleinreisende im Vergleich zu Familien?"

### Tipps:

1. Benutze cut auf der Spalte family\_size (die wir erstellt haben), um zwei Gruppen zu bilden: "Allein" (Größe 1) und "Familie" (Größe 2-15).
2. Benutze groupby auf dieser neuen Spalte.
3. Benutze agg oder mean(), um den Preis zu berechnen.