# 타이타닉호 생존자 분석 보고서

본 보고서는 1912년 4월 15일 침몰한 타이타닉호의 승객 데이터를 분석하여 생존에 영향을 미친 요인들을 조사한 결과를 담고 있습니다. 이 분석은 seaborn 라이브러리의 타이타닉 데이터셋을 기반으로 하였으며, 총 891명의 승객 데이터를 포함하고 있습니다.

## 데이터 개요

|  |  |  |
| --- | --- | --- |
| 변수명 | 데이터 타입 | 설명 |
| 생존여부 (survived) | 범주형 (0/1) | 0: 사망, 1: 생존 |

### 주요 변수 설명  
  
| 변수명 | 데이터 타입 | 설명 |  
|--------|------------|------|  
| 생존여부 (survived) | 범주형 (0/1) | 0: 사망, 1: 생존 |  
| 승객등급 (pclass) | 범주형 (1/2/3) | 1등석, 2등석, 3등석 |  
| 성별 (sex) | 범주형 | male/female |  
| 연령 (age) | 수치형 | 승객의 나이 |  
| 형제/배우자수 (sibsp) | 수치형 | 함께 탑승한 형제/배우자 수 |  
| 부모/자녀수 (parch) | 수치형 | 함께 탑승한 부모/자녀 수 |  
| 요금 (fare) | 수치형 | 티켓 요금 |  
| 선실등급 (class) | 범주형 | First/Second/Third |  
| 탑승도시 (embark\_town) | 범주형 | 승선 항구 (Southampton, Cherbourg, Queenstown) |

## 주요 분석 결과

### 1. 성별에 따른 생존율  
- 여성 승객의 생존율(약 74%)이 남성 승객(약 19%)보다 현저히 높았습니다.  
- 이는 '여성과 어린이 먼저'라는 구조 원칙이 적용된 결과로 보입니다.  
  
### 2. 승객 등급별 생존율  
- 1등석 승객의 생존율(약 63%)이 2등석(약 47%), 3등석(약 24%)에 비해 월등히 높았습니다.  
- 고가의 티켓을 구매한 승객일수록 생존 가능성이 높은 경향을 보였습니다.  
  
### 3. 연령대별 생존율  
- 어린이(12세 미만)의 생존율이 성인에 비해 상대적으로 높았습니다.  
- 특히 여성과 어린이를 우선적으로 구조한 정황이 확인되었습니다.  
  
### 4. 동반 가족 수와의 관계  
- 혼자 여행한 승객보다 가족과 함께 여행한 승객의 생존율이 다소 높은 경향을 보였습니다.  
- 특히 1-3명의 가족과 동반한 경우 생존율이 높았습니다.  
  
### 5. 탑승 항구별 특징  
- Cherbourg에서 탑승한 승객의 생존율(약 55%)이 다른 항구에서 탑승한 승객들에 비해 높았습니다.  
- 이는 Cherbourg에서 탑승한 승객 중 1등석 승객의 비율이 높았기 때문으로 분석됩니다.

## 결론

타이타닉호의 생존에 영향을 미친 주요 요인은 성별, 승객 등급, 연령, 동반 가족 수, 탑승 항구 등으로 나타났습니다. 특히 '여성과 어린이 먼저'라는 구조 원칙이 뚜렷하게 적용되었으며, 사회경제적 지위가 높을수록 생존 가능성이 높은 경향을 보였습니다. 이러한 분석 결과는 당시의 사회적 계층 구조와 구조 과정에서의 우선순위가 생존 여부에 큰 영향을 미쳤음을 보여줍니다.

## 참고문헌