**"Дейтасет Титаник"**

Потъването на Титаник е едно от най-печално известните корабокрушения в историята.

На 15 април 1912 г., по време на своето първо пътуване, широко приетият за 'непотопим' RMS Титаник потъва след сблъсък с айсберг. За съжаление, няма достатъчно спасителни лодки за всички на борда, което довежда до смъртта на 1502 от общо 2224 пътници и екипаж.

Една от причините, поради които корабокрушението довежда до толкова много жертви, е липсата на достатъчно спасителни лодки за пътниците и екипажа. Въпреки че има елемент на късмет при оцеляването след потъването, изглежда, че някои групи хора бяха по-склонни да оцелеят от други, като например жени, деца и горния клас.

Данните на дейтасета Титаник описват статуса на оцелели сред пътниците на Титаник. Данните не съдържат информация за екипажа, но съдържат реални възрасти на половината от пътниците. Основният източник на данни за пътниците на Титаник е Encyclopedia Titanica. Един от оригиналните източници на дейтасета е Ейтън и Хаас (1994) "Титаник: Триумф и Трагедия", издателство Патрик Стивънс Лтд, който включва списък на пътниците, създаден от множество изследователи и редактиран от Майкъл А. Финдлей.

Променливите в набора от данни са pclass, survived, name, age, embarked, home.dest, room, ticket, boat и sex. Pclass се отнася до класа на пътниците (1-вa, 2-рa, 3-тa) и е прокси за социално-икономически клас. Възрастта е в години, а някои бебета имат дробни стойности.

Pclass: Прокси за социално-икономически статус (СИС)

1-ва = Upper

2-ра = Middle

3-та = Lower

Възраст: Възрастта е десетична, ако е по-малка от 1. Ако възрастта е преоценена, тя е във формата на xx.5

Sibsp: Наборът от данни дефинира семейните връзки по следния начин...

Sibling = брат, сестра, страшен брат, страшна сестра

Spouse = съпруг, съпруга (любовниците и жените на бъдещите са игнорирани)

Parch: Наборът от данни дефинира семейните връзки по следния начин...

Parent = майка, баща

Child = дъщеря, син, страшна дъщеря, страшен син

Някои деца пътуват само с бавачка, следователно parch=0 за тях.