# Titanic

舍弃部分feature：重复度高或包含很多NAN的

创造部分feature：如把兄弟姐妹和父母，创造一个family；把连续的，划分出几个分段，如年龄；从名字中提取出，头衔，这个feature

对没有空NAN值的数据进行比例分析，如

|  | **Pclass** | **Survived** |
| --- | --- | --- |
| **0** | 1 | 0.629630 |
| **1** | 2 | 0.472826 |
| **2** | 3 | 0.242363 |