

pranshu.R

Pranshu

Sun Jun 26 15:31:34 2016

```
setwd("~/Research/Test2")
titanic3 <- read.csv("~/Research/Test2/titanic3.csv", na.strings = "")
library(dplyr)
```

```
## Warning: package 'dplyr' was built under R version 3.2.5
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##      filter, lag
```

```
## The following objects are masked from 'package:base':
##
##      intersect, setdiff, setequal, union
```

```
titanic3 %>% filter(titanic3$embarked==NA)
```

```
##   [1] pclass   survived  name      sex      age      sibsp   parch
##   [8] ticket   fare      cabin    embarked boat     body     home.dest
## <0 rows> (or 0-length row.names)
```

```
titanic3$age[which(is.na(titanic3$age))]<-mean(titanic3$age, na.rm=TRUE)
View(titanic3)
titanic4<-titanic3 %>% mutate(Has_cab_number=ifelse(!is.na(titanic3$cabin),1,0))
View(titanic4)
str(titanic4)
```

```
## 'data.frame':    1310 obs. of  15 variables:
##  $ pclass       : int   1 1 1 1 1 1 1 1 1 1 ...
##  $ survived     : int   1 1 0 0 0 1 1 0 1 0 ...
##  $ name         : Factor w/ 1307 levels "Abbing, Mr. Anthony",...: 22 24 25 26 27
31 46 47 51 55 ...
##  $ sex          : Factor w/ 2 levels "female","male": 1 2 1 2 1 2 1 2 1 2 ...
##  $ age          : num   29 0.917 2 30 25 ...
##  $ sibsp        : int    0 1 1 1 1 0 1 0 2 0 ...
##  $ parch        : int    0 2 2 2 2 0 0 0 0 0 ...
##  $ ticket       : Factor w/ 929 levels "110152","110413",...: 188 50 50 50 50 125
93 16 77 826 ...
##  $ fare         : num   211 152 152 152 152 ...
##  $ cabin        : Factor w/ 186 levels "A10","A11","A14",...: 44 80 80 80 80 150 1
46 16 62 NA ...
##  $ embarked    : Factor w/ 3 levels "C","Q","S": 3 3 3 3 3 3 3 3 3 1 ...
##  $ boat        : Factor w/ 27 levels "1","10","11",...: 12 3 NA NA NA 13 2 NA 27
NA ...
##  $ body        : int    NA NA NA 135 NA NA NA NA NA 22 ...
##  $ home.dest    : Factor w/ 369 levels "?Havana, Cuba",...: 309 231 231 231 231 23
7 162 24 22 229 ...
##  $ Has_cab_number: num    1 1 1 1 1 1 1 1 1 0 ...
```

```
setwd("~/Research/Test2")
write.csv(titanic4, file="titanic4.csv")
```