Untitled

KanKan Yang

March 9, 2018

## R Markdown

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see <http://rmarkdown.rstudio.com>.

When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

hmda\_df<-read.csv("ny\_hmda\_2015.csv")  
  
hmda\_df$Loan\_Approval<-NA  
unique(hmda\_df[,c('action\_taken','action\_taken\_name')])

## action\_taken action\_taken\_name  
## 1 1 Loan originated  
## 213117 2 Application approved but not accepted  
## 226297 3 Application denied by financial institution  
## 301402 4 Application withdrawn by applicant  
## 337937 5 File closed for incompleteness  
## 353302 6 Loan purchased by the institution  
## 410232 7 Preapproval request denied by financial institution

hmda\_df$Loan\_Approval[hmda\_df$action\_taken==2|hmda\_df$action\_taken==6]<-"Approved"  
hmda\_df$Loan\_Approval[hmda\_df$action\_taken==3|hmda\_df$action\_taken==7]<-"Declined"

第19行读入数据，从21行道25行是回答第6个问题。

首先要根据 action\_taken 来判断哪些是 approved 的。

22 行找出数据中 action\_taken 和相应的名字。这7个actions里 1，4，5 我感觉是无关的

（我assum你们的目的是分析数据帮助贷款公司判断具体case是如何做决定。如果这种情况由于申请人没能提供完整信息，或者还没处理，或者直接取消的话 这些数据也没什么作用）。

我感觉2，6 是approved 3，7 是declined。 你们仔细研究一下

## Including Plots

You can also embed plots, for example:

Note that the echo = FALSE parameter was added to the code chunk to prevent printing of the R code that generated the plot.