

eda variables by event and non-event and variables by event

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:

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

```
npf_original <- read.csv("npf_train.csv")
npf <- npf_original
```

Clean data making date as index and removing columns (id and oartlybad)

```
rownames(npf) <- npf[, "date"]
npf <- npf[, -(1:2)]
npf <- npf[, -2]
```

```
npf$class2 <- factor("event", levels=c("nonevent", "event"))
npf$class2[npf$class4=="nonevent"] <- "nonevent"
```

```
library(ggplot2)
library(grid)
library(gridExtra)
for (i in 3:ncol(npf)-1) {
  p1 <- ggplot(data=npf, aes(x=npf[, i], group=class2, fill=class2)) +
    geom_density(adjust=1.5, alpha=.4) +
    labs(title=names(npf)[i], x=names(npf)[i], y="Density")

  p2 <- ggplot(data=npf, aes(x=npf[, i], group=class4, fill=class4)) +
    geom_density(adjust=1.5, alpha=.4) +
    labs(title=names(npf)[i], x=names(npf)[i], y="Density")

  print(p1)
  print(p2)
}
```















































































































































































































































































































































































































