RQ2: Does the interest differ based on the type of TD?

Yasutaka Kamei

Feb 11th, 2016

Read Data

```
data <- read.csv("/Users/kamei/Research/techdebt/msr16_td_interest/datasets/CSV/technical_debt_summary.

# choose one of duplicated method and version name
method_and_version_name <- paste(data$Method_Signature, data$v1, sep="")
data <- data[!duplicated(method_and_version_name), ]

# only use technical debt including metrics
a <- data[(data[, "CountInput_v1"] != -1 & data[, "CountInput_v2"] != -1), ]
# only use technical debt including non 0 for division
b <- a[(a[,"CountInput_v1"] != 0), ]
b <- cbind(b, interest = (b$CountInput_v2 / b$CountInput_v1))
```

The number of technical debt in each type

```
tmp <- b[b[,"Project"] == "apache-ant", ]
fc <- factor(tmp$Type)
tapply(tmp$interest, fc, length)
```

```
##      DEFECT      DESIGN REQUIREMENT        TEST
##           8          46           5           5
```

```
tapply(tmp$interest, fc, function(x){length(x)/length(tmp$interest) * 100})
```

```
##      DEFECT      DESIGN REQUIREMENT        TEST
##     12.5000     71.8750      7.8125      7.8125
```

```
tmp <- b[b[,"Project"] == "apache-jmeter", ]
fc <- factor(tmp$Type)
tapply(tmp$interest, fc, length)
```

```
##        DEFECT        DESIGN DOCUMENTATION   REQUIREMENT          TEST
##             8           134             2            12             3
```

```
tapply(tmp$interest, fc, function(x){length(x)/length(tmp$interest) * 100})
```

```
##        DEFECT        DESIGN DOCUMENTATION   REQUIREMENT          TEST
##      5.031447     84.276730      1.257862      7.547170      1.886792
```

```r
tmp <- b[b[,"Project"] == "jruby", ]
fc <- factor(tmp$Type)
tapply(tmp$interest, fc, length)
```

```
##        DEFECT        DESIGN DOCUMENTATION   REQUIREMENT          TEST
##            87           132             1            41             3
```

```r
tapply(tmp$interest, fc, function(x){length(x)/length(tmp$interest) * 100})
```

```
##        DEFECT        DESIGN DOCUMENTATION   REQUIREMENT          TEST
##    32.9545455    50.0000000     0.3787879    15.5303030     1.1363636
```

Observation

- In jruby, three types of technical dept are included than 10%.
- We use DEFECT, DESIGN and REQ in jruby.

```r
tmp <- b[b[,"Project"] == "jruby", ]
tmp <- tmp[(tmp[,"Type"] == "DEFECT" |  tmp[,"Type"] == "DESIGN" |  tmp[,"Type"] == "REQUIREMENT"),]
fc <- factor(tmp$Type)

tapply(tmp$interest, fc, summary)
```

```
## $DEFECT
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.8889  1.0000  1.0000  1.1220  1.0000  5.0000
##
## $DESIGN
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.3333  1.0000  1.0000  1.2010  1.2000  9.0000
##
## $REQUIREMENT
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.3125  0.8000  1.0000  1.1290  1.0000  5.5000
```

```r
defect <- subset(tmp$interest, tmp[,"Type"] == "DEFECT")
design <- subset(tmp$interest, tmp[,"Type"] == "DESIGN")
req <- subset(tmp$interest, tmp[,"Type"] == "REQUIREMENT")

library(vioplot)
```

```
## Loading required package: sm
```

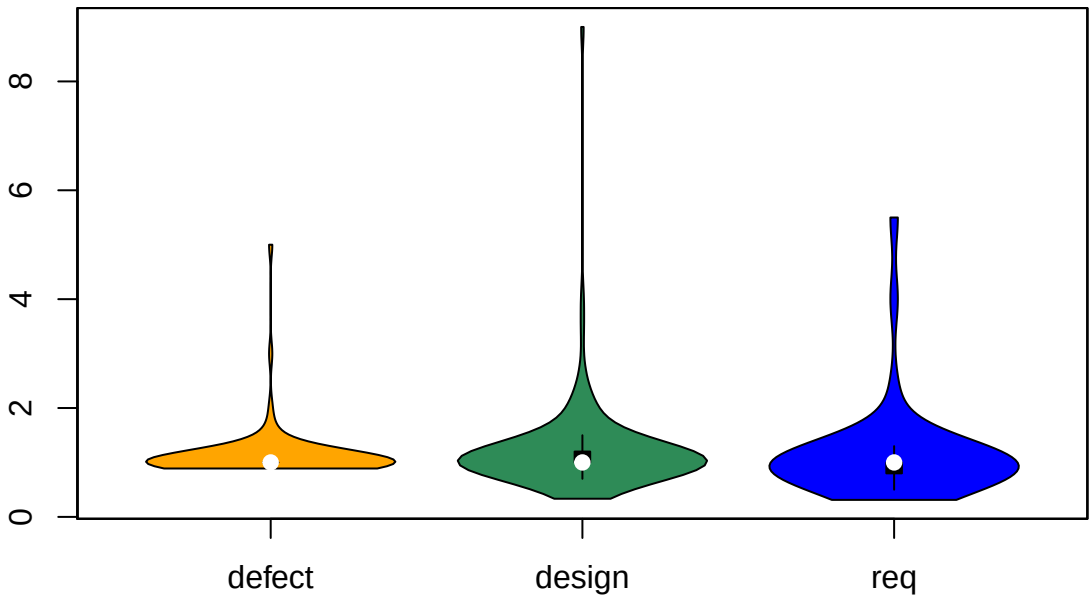
```
## Package 'sm', version 2.2-5.4: type help(sm) for summary information
```

```r
plot(0, 0, type = "n", xlab = "", ylab = "", axes = FALSE,
     xlim = c(0.5, 3.5), ylim = range(c(defect, design, req)))

axis(side = 1, at = 1:3, labels = c("defect", "design", "req"))
```

```
axis(side = 2)

vioplot(defect, at = 1, col = "orange", add = TRUE)
vioplot(design, at = 2, col = "seagreen", add = TRUE)
vioplot(req, at = 3, col = "blue", add = TRUE)
```



Observation

- There is no difference in each category.