Exploration

Nick Lauerman

Contents

last Run 23 jan 18 @ 10:19

Purpose

To explore metrics options based on data collected by Software Quality Assurance in a spreadsheet for document review and approval.

Libraries used

No libriries used yet

```
library(lubridate)
```

```
##
## Attaching package: 'lubridate'
```

```
## The following object is masked from 'package:base':
##
##     date
```

Data

The data is presented n a CVS file (saved from an Excel workbook) with the following columns and data chartistics

SQA SQAE Preforming the review
Project ID Uniquie ID assigned tot he project may come from NPV or ePAS
Application Application Name (limited selection)
Doc Version Version number of the document, also know as the revision number
Doc ID, if needed unique identifier to distingous between documnets fo the same deliver type within a project
Approver or Disapprov? Approval status from the following list
* **A** Approved
* **A-FP** Approved on fist pass
* **D** Disapproved
Date of Approval or Disapproval Date action was taken

Reason for Disapproval: Selected from the following list

- Inaccurate information
- Incorrect scope
- Insufficient testing
- Insufficient traceability
- N/A
- Not compliant with procedure
- Not following configuration management
- Not following document version control
- Not using template
- Other
- Requirement deficiency

Inport the data

The data is read in and variables are assigned more programicaly useful names.

```
FPA.Raw <- read.csv(file = "./data/First Pass Acceptance.csv",
                    stringsAsFactors = FALSE,
                    col.names = c("sqa",
                                  "project",
                                  "application",
                                  "deliverable",
                                  "version",
                                  "docID",
                                  "status",
                                  "date",
                                  "reason",
                                  "comments"))
FPA <- FPA.Raw
```

format the data

factors

Set the following data as factors:

*SQA

*Application

*Deliverable

*status

```
FPA$sqa <- as.factor(FPA$sqa)
FPA$application <- as.factor(FPA$application)
FPA$deliverable <- as.factor(FPA$deliverable)
FPA$status <- as.factor(FPA$status)
```

Clean up reason

convert blank fields in status to NA when status is converted to a factor

```
FPA$reason <- as.factor(FPA$reason)
levels(FPA$reason)
```

```
## [1] ""
## [2] "Inaccurate information"
## [3] "Incorrect scope"
## [4] "Not compliant with procedure"
## [5] "Not following document version control"
## [6] "Requirement deficiency"
```

```
levels(FPA$reason)[1] <- NA
levels(FPA$reason)
```

```
## [1] "Inaccurate information"
## [2] "Incorrect scope"
## [3] "Not compliant with procedure"
## [4] "Not following document version control"
## [5] "Requirement deficiency"
```

Date

```
FPA$date <- as.Date(FPA$date, format = "%d-%B-%y")
```

Add new fields

Month and year

From date create month and year fields

```
FPA$month <- month(FPA$date, label = TRUE)
FPA$year <- year(FPA$date)
```

deliverable UID

create an unique identification (UID) for each processed item so that each deliverable is uniquely identified accross the entire data set. to do this the UID will consist of the Project, deliverable, docID and version seperated with a dash

```
FPA$UID <- paste(FPA$project,
                 FPA$deliverable,
                 FPA$docID,
                 FPA$version,
                 sep = "-"
                 )
```

Simple count metrics

Projects worked

```
length(unique(FPA$project))
```

```
## [1] 15
```

```
as.character(unique(FPA$application))
```

```
##  [1] "Groninger"                   "Assay File Database "
##  [3] "PCN/SCN"                     "WWLIMS"
##  [5] "Pulse"                       "DFCS"
##  [7] "DPW"                         "QIMS"
##  [9] "Abbott Transfusion Medicine" "Apollo/PHM"
## [11] "AFMS"                        "SAS"
## [13] "Metrics Library"             "DaVinci"
```

Number of reviews

```
nrow(FPA)
```

```
## [1] 66
```

Number of Documents Reviewed

```
length(unique(FPA$UID))
```

```
## [1] 57
```

Number of Projects Started

```
length(unique(
     subset(
          subset(FPA,
```

```
                subset = ((status == "A" | status == "A-FP") & version == 1)),
          subset = deliverable == "Software Change Request")$project))
```

```
## [1] 2
```

Number of Projects Completed

```
length(unique(
     subset(
          subset(FPA,
                 subset = ((status == "A" | status == "A-FP"))),
          subset = deliverable == "System Certification Summary")$project))
```

```
## [1] 0
```

Approval rate

probability of a document being approved when reviewed

```
nrow(unique(
     subset(FPA, subset = (status == "A" | status == "A-FP"))
))/nrow(FPA)
```

```
## [1] 0.7575758
```

First Pass Acceptance

probability of a document being approved when reviewed the first time

```
nrow(unique(
     subset(FPA, subset = status == "A-FP")
))/nrow(FPA)
```

```
## [1] 0.5454545
```