

Functions

Nick Lauerman

January 22, 2018 @ 4:00 PM

Contents

Data Input	1
Test	1
count	1
cleaning Data	2
Factors and Date	2
Test	2
Processing	4
Test	4

This is to document the development of functions derived from various other activities to support the FPA project. It is important to note that functions developed or documented here will have to be included in an R file so they may be easily brought into the R environment in the future.

Data Input

this function will read the csv file in and will make required conversions to factors or date

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

Test

count

First the function is run creating the FPA data frame. The the number of rows in the data frame is checked again the source data (66 at this time) as well as the number of columns (10)

```
FPA <- get.FPA()
nrow(FPA)
```

```
## [1] 66
ncol(FPA)

## [1] 10
dim(FPA)

## [1] 66 10
str(FPA)

## 'data.frame': 66 obs. of 10 variables:
## $ sqa : chr "Wayne" "Wayne" "Wayne" "Wayne" ...
## $ project : chr "LC-S 01073.000" "LC-S 01104.000" "LC-S 01104.000" "LC-S 00986.005" ...
## $ application: chr "Groninger" "Assay File Database " "Assay File Database " "PCN/SCN" ...
## $ deliverable: chr "Traceability" "Software Change Request" "Software Change Request" "Software Change Request" ...
## $ version : int 1 1 1 NA 1 1 1 1 1 1 ...
## $ docID : chr "Protrace-01" "SCR-01" "SCR-01" "SCA-01" ...
## $ status : chr "A-FP" "D" "A" "D" ...
## $ date : chr "16-Jan-18" "17-Jan-18" "17-Jan-18" "18-Jan-18" ...
## $ reason : chr "" "Not compliant with procedure" "" "Incorrect scope" ...
## $ comments : chr "" "" "" "Project name includes SMF, but Assessment links only PCN SCN" ...
```

cleaning Data

Factors and Date

```
clean.FPA <- function(FPA){
  require(lubridate)
  FPA$sqa <- as.factor(FPA$sqa)
  FPA$application <- as.factor(FPA$application)
  FPA$deliverable <- as.factor(FPA$deliverable)
  FPA$status <- as.factor(FPA$status)
  FPA$reason <- as.factor(FPA$reason)
  levels(FPA$reason)[1] <- NA
  FPA$date <- as.Date(FPA$date, format = "%d-%B-%y")
  return(FPA)
}
```

Test

Now a count of each factor will be tabulated to verify against the know data and look at data frame structure

```
FPA <- clean.FPA(FPA = FPA)
```

```
## Loading required package: lubridate
##
## Attaching package: 'lubridate'
## The following object is masked from 'package:base':
##
##     date
```

```
str(FPA)
```

```
## 'data.frame':   66 obs. of  10 variables:
## $ sqa          : Factor w/ 3 levels "Beilah","Nick",...: 3 3 3 3 1 1 1 1 1 1 ...
## $ project      : chr  "LC-S 01073.000" "LC-S 01104.000" "LC-S 01104.000" "LC-S 00986.005" ...
## $ application: Factor w/ 14 levels "Abbott Transfusion Medicine",...: 8 4 4 10 14 14 14 14 11 11 ...
## $ deliverable: Factor w/ 14 levels "CII","Design Verification",...: 11 6 6 7 9 9 9 9 4 9 ...
## $ version      : int   1 1 1 NA 1 1 1 1 1 1 ...
## $ docID        : chr   "Protrac-01" "SCR-01" "SCR-01" "SCA-01" ...
## $ status       : Factor w/ 3 levels "A","A-FP","D": 2 3 1 3 1 1 3 1 2 2 ...
## $ date         : Date, format: "2018-01-16" "2018-01-17" ...
## $ reason       : Factor w/ 5 levels "Inaccurate information",...: NA 3 NA 2 NA NA 1 NA NA NA ...
## $ comments     : chr   "" "" "" "Project name includes SMF, but Assessment links only PCN SCN" ...
```

```
table(FPA$sqa)
```

```
##
## Beilah   Nick   Wayne
##      37     24     5
```

```
table(FPA$application)
```

```
##
## Abbott Transfusion Medicine      AFMS
##              1                    9
##              Apollo/PHM          Assay File Database
##              9                    2
##              DaVinci              DFCS
##              4                    6
##              DPW                  Groninger
##              3                    1
##              Metrics Library      PCN/SCN
##              1                    2
##              Pulse                QIMS
##              10                   4
##              SAS                  WWLIMS
##              1                    13
```

```
table(FPA$deliverable)
```

```
##
##              CII              Design Verification
##              5                2
##              FRS              IIVP results
##              4                7
##              Project Plan      Software Change Request
##              4                5
## Software Compliance Assessment  System Certification Summary
##              3                1
##              Test Protocol      Test Protocol Results
##              21                 6
##              Traceability       URS
##              3                  2
##              User Acceptance Protocol  Validation Plan
##              2                  1
```

```
table(FPA$status)
```

```
##
##      A A-FP      D
##     14   37   15
```

```
table(FPA$reason)
```

```
##
##              Inaccurate information
##                      7
##              Incorrect scope
##                      1
##      Not compliant with procedure
##                      5
## Not following document version control
##                      1
##              Requirement deficiency
##                      1
```

```
table(FPA$date)
```

```
##
## 2018-01-03 2018-01-09 2018-01-11 2018-01-15 2018-01-16 2018-01-17
##           2           1           1           14           17           11
## 2018-01-18 2018-01-19 2018-01-22
##           16           3           1
```

Processing

Create a separate value for Month and Year for the date value.

Create a UID for each deliverable

```
process <- function(FPA){
  require(lubridate)
  FPA$month <- month(FPA$date, label = TRUE)
  FPA$year <- year(FPA$date)
  FPA$UID <- paste(FPA$project,
                  FPA$deliverable,
                  FPA$docID,
                  FPA$version,
                  sep = "-")
}
```

Test

inspect the structure of the data frame

```
str(FPA)
```

```
## 'data.frame':   66 obs. of  10 variables:
## $ sqa          : Factor w/ 3 levels "Beilah","Nick",...: 3 3 3 3 1 1 1 1 1 1 ...
## $ project      : chr  "LC-S 01073.000" "LC-S 01104.000" "LC-S 01104.000" "LC-S 00986.005" ...
## $ application: Factor w/ 14 levels "Abbott Transfusion Medicine",...: 8 4 4 10 14 14 14 14 11 11 ...
```

```
## $ deliverable: Factor w/ 14 levels "CII","Design Verification",...: 11 6 6 7 9 9 9 9 4 9 ...
## $ version      : int   1 1 1 NA 1 1 1 1 1 1 ...
## $ docID        : chr   "Protrace-01" "SCR-01" "SCR-01" "SCA-01" ...
## $ status       : Factor w/ 3 levels "A","A-FP","D": 2 3 1 3 1 1 3 1 2 2 ...
## $ date         : Date, format: "2018-01-16" "2018-01-17" ...
## $ reason       : Factor w/ 5 levels "Inaccurate information",...: NA 3 NA 2 NA NA 1 NA NA NA ...
## $ comments     : chr   "" "" "" "Project name includes SMF, but Assessment links only PCN SCN" ...
```

Now look at at the new values

```
table(FPA$month)
```

```
## < table of extent 0 >
```

```
table(FPA$year)
```

```
## < table of extent 0 >
```

```
table((FPA$UID))
```

```
## < table of extent 0 >
```