

---

# DataFax Reporting Tools Using R

**Zekai Otles<sup>1</sup>, Pascal Groenen<sup>2</sup>,  
David Gaston<sup>1</sup> and Joris Bannenberg<sup>2</sup>**

*<sup>1</sup>Frontier Science & Technology Research Foundation*

*<sup>2</sup>Factory CRO, Clinical Research for Medical Devices*



## Objectives

---

- Display or tabulate DataFax reports in a comprehensible fashion.
- Eliminate the need for additional resources.
- Generate new tools and enhance existing tools to display data in R.

## Background

---

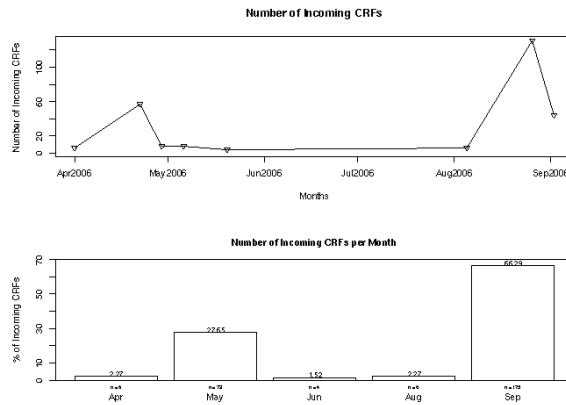
- Gulati V. and Green K. (DFUG-2001); Using DataFAX Reports to Monitor and Improve Site Performance.
- Otles Z. et al. (DFUG-2006); Graphical Representation of DataFAX Reports Using R-Project Statistical Software.
- R is an open source statistical computation and graphics software (<http://www.r-project.org/>).
- Factory has a need for efficient and enhanced reporting of study data.
- Standard DataFAX reports are limited
  - Layout of DataFAX reports
  - Out-of-environment applications are required, i.e. SAS®, Excel
  - Data Management receives many requests from
    - Sponsors – Study metrics
    - CRAs – Study metrics, detailed visit scheduling

## Approach

---

- Compilation CRF set Mock study.
- Setup Mock study in DataFAX.
- Populate database with 14 fictitious subjects.
- Populate database with ~50 QC Notes.
- Factory Mock study exported to Frontier Science to develop R tools.
- Monthly web conferences to exchange programs, ideas, specifications and provide feedback.
- Initial specifications for the functions determined based on the needs of Factory and Frontier Science.
- The tools presented here are developed in R only.

## Weekly incoming CRFs (DF\_WFcrfsperwk)



- Snap shots of incoming CRFs
- CRF transmission by the month

## Visit Schedule

### Visit schedule output from SAS program

visit(Visit Number)	v_done		Total
Frequency	No	Yes	
Percent			
Row Pct			
Col Pct			
<b>Screening</b>			
	1	12	13
	1.62	23.08	25.00
	7.69	92.31	16.67
	16.67	26.09	
<b>2 Weeks Follow-Up</b>			
	1	12	13
	1.92	23.08	25.00
	7.69	92.31	16.67
	16.67	26.09	
<b>6 weeks Follow-Up</b>			
	2	11	13
	3.85	21.15	25.00
	15.38	84.62	16.67
	33.33	23.91	
<b>Study Termination</b>			
	2	11	13
	3.85	21.15	25.00
	15.38	84.62	16.67
	33.33	23.91	
<b>Total</b>			
	6	46	52
	11.54	88.46	100.00

Subject ID	Visit Number	Start of Visit Window	Target Visit Date	Actual Date	End of Visit Window	Indicator
1001	Screening	24/09/2005	01/07/2005	01/07/2005	01/07/2005	
1001	2 Weeks Follow-Up	08/07/2005	10/07/2005	14/07/2005	11/07/2005	Out of Window
1001	6 weeks Follow-Up	13/08/2005	14/08/2005	13/08/2005	15/08/2005	
1001	Study Termination	13/08/2005	14/08/2005	14/08/2005	15/08/2005	
1002	Screening	24/09/2005	01/07/2005	01/07/2005	01/07/2005	
1002	2 Weeks Follow-Up	08/07/2005	08/07/2005	16/07/2005	10/07/2005	Out of Window
1002	6 weeks Follow-Up	12/08/2005	13/08/2005	12/08/2005	14/08/2005	
1002	Study Termination	12/08/2005	13/08/2005	03/09/2005	14/09/2005	
1003	Screening	24/09/2005	01/07/2005	01/07/2005	01/07/2005	
1003	2 Weeks Follow-Up	08/07/2005	10/07/2005	17/07/2005	11/07/2005	Out of Window
1003	6 weeks Follow-Up	13/08/2005	14/08/2005	13/08/2005	15/08/2005	
1003	Study Termination	13/08/2005	14/08/2005	15/08/2005	15/08/2005	
1004	Screening	24/09/2005	01/07/2005	01/07/2005	01/07/2005	
1004	2 Weeks Follow-Up	07/07/2005	08/07/2005	15/07/2005	09/07/2005	Out of Window
1004	6 weeks Follow-Up	11/08/2005	12/08/2005	11/08/2005	13/08/2005	
1004	Study Termination	11/08/2005	13/08/2005	11/08/2005	13/08/2005	

Subject ID	Visit Number	Start of Visit Window	Target Visit Date	Actual Date	End of Visit Window	Indicator
2001	Screening	10/04/2006	17/04/2006	10/04/2006	17/04/2006	
2001	2 Weeks Follow-Up	23/04/2006	24/04/2006	03/05/2006	25/04/2006	Out of Window
2001	6 weeks Follow-Up	28/05/2006	28/05/2006	26/05/2006	30/05/2006	
2001	Study Termination	28/05/2006	28/05/2006	30/05/2006	30/05/2006	
2002	Screening	05/02/2006	15/02/2006	10/02/2006	12/02/2006	
2002	2 Weeks Follow-Up	18/02/2006	18/02/2006	28/02/2006	20/02/2006	Out of Window
2002	6 weeks Follow-Up	24/02/2006	24/02/2006	27/02/2006	27/02/2006	Late
2002	Study Termination	24/02/2006	24/02/2006	27/02/2006	27/02/2006	Late

- Export data to SAS from DataFax
- SAS program is not in the same OS
- RTF produced by SAS code and displayed different OS
- No graphical output

# Visit Schedule

## Visit schedule output from DfRTools

```

rParWc: Display Number of Incoming CRFs
cSParWc: Display Number of CRFs per plate
cSParC: Display Number of CRFs per subject
VisitSchedule: Display Visit and Appointment Scheduling
Visit: Center subject's tracking for a specified visit and center
VCR: Summary report of records received for each center
VCR: Summary report of external QC entries for each center
V: Visit: Display some statistics on single visits
Options: --i: /d:/a2/ --t: /T/
Run Report Explain Reset
Start: Thu Jan 25 15:13:50 2007
Output From: VisitSchedule: Display tables and pie Charts visit nap schedu
studyNumber=220
current working dir=/u/home/oties
current studies dir=/u/ops/studies
T/m/y

```

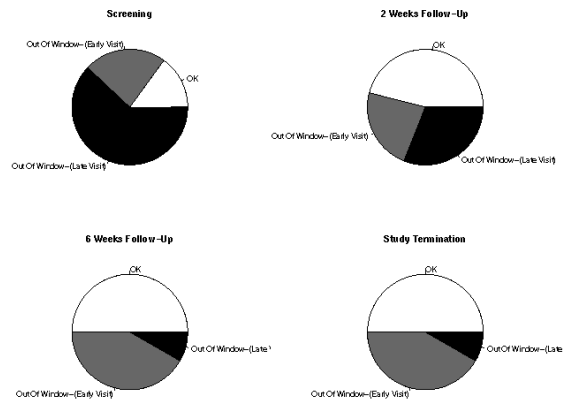
Table 2: Visit Scheduling

StudyID	VisitID	VisitLabel	TargetDate	TargetDate	VisitDate	SubjectID	Status
1	1	03001	Screening	26/06/06	26/06/06	02/07/2006 26/06/06	Out Of Window (Late)
2	1	03001	2 Weeks Follow-Up	30/07/06	17/07/06	14/07/2006 30/07/06	Out Of Window (Earl)
3	1	03001	6 Weeks Follow-Up	15/08/06	14/08/06	13/08/2006 15/08/06	OIE
4	1	03001	Study Termination	10/09/06	14/09/06	14/09/2006 10/09/06	OIE
5	1	03002	Screening	26/06/06	23/06/06	01/07/2006 26/06/06	Out Of Window (Late)
6	1	03002	2 Weeks Follow-Up	30/07/06	19/07/06	06/07/2006 30/07/06	OIE
7	1	03002	6 Weeks Follow-Up	15/08/06	13/08/06	12/08/2006 15/08/06	OIE
8	1	03002	Study Termination	10/09/06	12/09/06	12/09/2006 10/09/06	OIE
9	1	03003	Screening	26/06/06	26/06/06	05/07/2006 26/06/06	Out Of Window (Late)
10	1	03003	2 Weeks Follow-Up	30/07/06	17/07/06	17/07/2006 30/07/06	OIE
11	1	03003	6 Weeks Follow-Up	15/08/06	14/08/06	13/08/2006 15/08/06	OIE
12	1	03003	Study Termination	10/09/06	14/09/06	13/09/2006 10/09/06	OIE
13	1	03004	Screening	24/06/06	24/06/06	01/07/2006 24/06/06	Out Of Window (Late)
14	1	03004	2 Weeks Follow-Up	34/07/06	13/07/06	13/07/2006 34/07/06	OIE
15	1	03004	6 Weeks Follow-Up	11/08/06	12/08/06	11/08/2006 11/08/06	OIE
16	1	03004	Study Termination	11/09/06	12/09/06	11/09/2006 11/09/06	OIE
17	2	02011	Screening	20/04/06	10/04/06	05/04/2006 20/04/06	OIE
18	2	02011	2 Weeks Follow-Up	20/04/06	01/04/06	05/03/2006 20/04/06	Out Of Window (Late)
19	2	02011	6 Weeks Follow-Up	20/04/06	29/03/06	29/03/2006 20/04/06	OIE
20	2	02011	Study Termination	20/04/06	29/03/06	29/03/2006 20/04/06	OIE
21	2	02012	Screening	05/02/06	03/02/06	30/02/2006 05/02/06	Out Of Window (Late)
22	2	02012	2 Weeks Follow-Up	25/02/06	25/02/06	25/02/2006 25/02/06	Out Of Window (Late)
23	2	NA	NA	25/03/06	25/03/06	NA	Missing Place/Late Vis
24	2	NA	NA	25/03/06	25/03/06	NA	Missing Place/Late Vis
25	2	02011	Screening	05/04/06	03/04/06	14/03/2006 05/04/06	Out Of Window (Late)
26	2	02011	2 Weeks Follow-Up	25/04/06	24/04/06	24/04/2006 25/04/06	OIE
27	2	02011	6 Weeks Follow-Up	21/05/06	22/05/06	20/05/2006 21/05/06	Out Of Window (Earl)
28	2	02011	Study Termination	21/05/06	22/05/06	20/05/2006 21/05/06	Out Of Window (Earl)
29	4	04011	Screening	20/03/06	20/03/06	NA	20/03/06
30	4	04011	2 Weeks Follow-Up	17/04/06	10/04/06	NA	10/04/06
31	4	04011	6 Weeks Follow-Up	10/05/06	10/05/06	NA	10/05/06
32	4	04011	Study Termination	10/06/06	10/06/06	NA	10/06/06
33	5	05011	Screening	09/07/06	09/07/06	01/07/2006 09/07/06	Out Of Window (Earl)
34	5	05011	2 Weeks Follow-Up	29/07/06	29/07/06	21/07/2006 29/07/06	Out Of Window (Late)
35	5	05011	6 Weeks Follow-Up	24/08/06	25/08/06	25/08/2006 24/08/06	OIE
36	5	05011	Study Termination	24/09/06	29/09/06	29/09/2006 24/09/06	OIE
37	5	05012	Screening	09/07/06	09/07/06	12/07/2006 09/07/06	Out Of Window (Late)
38	5	05012	2 Weeks Follow-Up	29/07/06	29/07/06	21/07/2006 29/07/06	Out Of Window (Late)
39	5	05012	6 Weeks Follow-Up	24/08/06	24/08/06	24/08/2006 24/08/06	Out Of Window (Earl)
40	5	05012	Study Termination	24/09/06	29/09/06	24/09/2006 24/09/06	Out Of Window (Earl)
41	5	05013	Screening	24/07/06	24/07/06	22/07/2006 24/07/06	Out Of Window (Earl)
42	5	05013	2 Weeks Follow-Up	19/08/06	17/08/06	17/08/2006 19/08/06	OIE
43	5	05013	6 Weeks Follow-Up	10/09/06	11/09/06	09/09/2006 10/09/06	Out Of Window (Earl)
44	5	05013	Study Termination	10/10/06	11/10/06	09/10/2006 10/10/06	Out Of Window (Earl)
45	9	09011	Screening	01/01/06	01/01/06	01/01/2006 01/01/06	OIE
46	9	09011	2 Weeks Follow-Up	21/01/06	20/01/06	18/01/2006 21/01/06	Out Of Window (Earl)
47	9	09011	6 Weeks Follow-Up	10/02/06	13/02/06	24/01/2006 10/02/06	Out Of Window (Earl)
48	9	09011	Study Termination	10/02/06	13/02/06	29/01/2006 10/02/06	Out Of Window (Earl)
49	9	09012	Screening	10/02/06	10/02/06	18/02/2006 10/02/06	Out Of Window (Late)
50	9	09012	2 Weeks Follow-Up	30/04/06	10/04/06	02/04/2006 30/04/06	OIE
51	9	09012	6 Weeks Follow-Up	01/04/06	01/04/06	05/03/2006 01/04/06	Out Of Window (Late)
52	9	09012	Study Termination	30/04/06	01/05/06	05/03/2006 01/05/06	Out Of Window (Late)

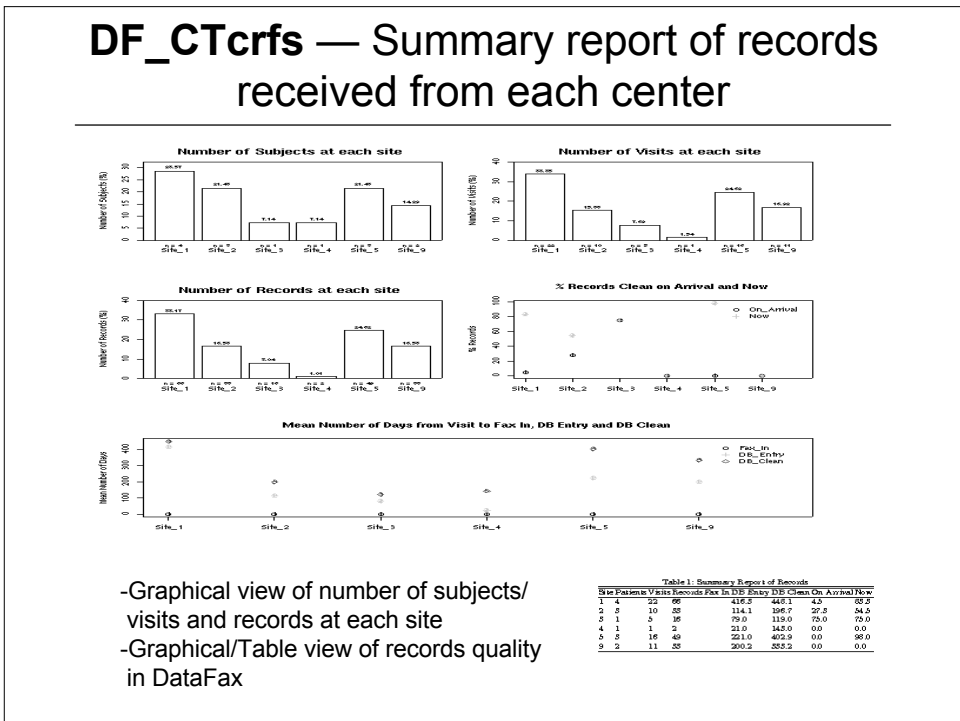
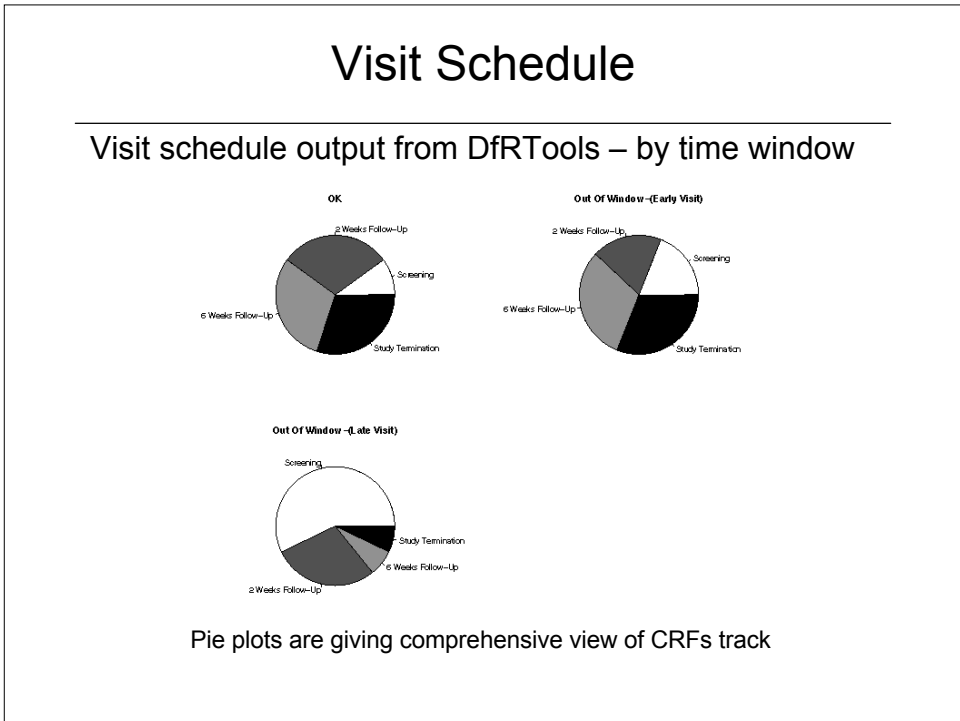
- DfRTools function uses DfExport.rpc to avoid data transfer
- DfRTools embedded into the DataFax
- The report outputs produced as an image, and directly displayed on the OS
- Graphical outputs are part of the visit schedule report
- No need to SAS, Excel, data transfer to other OS, etc.

# Visit Schedule

## Visit schedule output from DfRTools – by visit



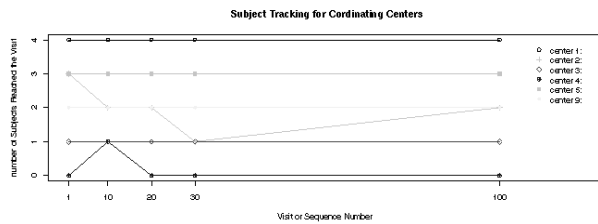
- Data are still superficial, and for demonstrations
- Late/early visits are later/earlier than targeted date,



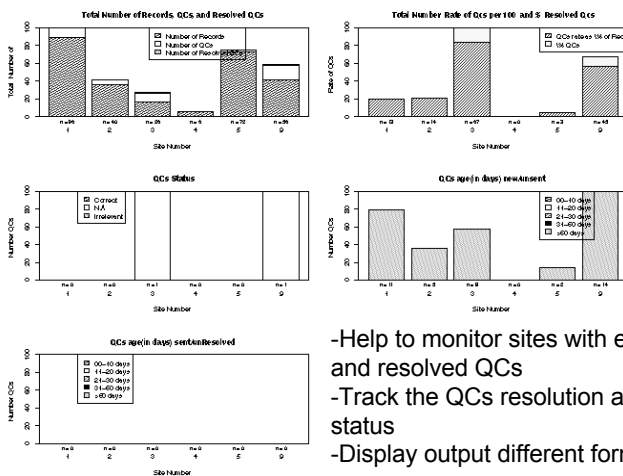
## DF\_CTvisits — Number of patients who have reached each visit at each center

```

ctrlPerk: Display Number of Incoming CTs
qcPerkStat: Display Number of QCs per site
qcPerSub: Display Number of QCs per subject
visitSchedule: Display tables and pie charts visit map scheduling
QCPerk: Display Number of QCs per visit
tOfCs: Summary report of records received for each center
eQCs: Summary report of external QC notes for each center
DF_stats: Display some statistics on single site
Options: -8 1,10,20,30,100
Run Report Explain Reset
Start: Thu Jan 25 15:30:34 2007
Output From: "ctVisits: Center subject Tracking for a specified visits and cen
studyNumber=220
siteNumbers=1,10,20,30,100
centerNumbers:
    
```



## DF\_CTqcs — Summary report of external QC notes for each center



- Help to monitor sites with existing and resolved QCs
- Track the QCs resolution and the status
- Display output different formats

## DF\_CTqcs — Summary report of external QC notes for each center

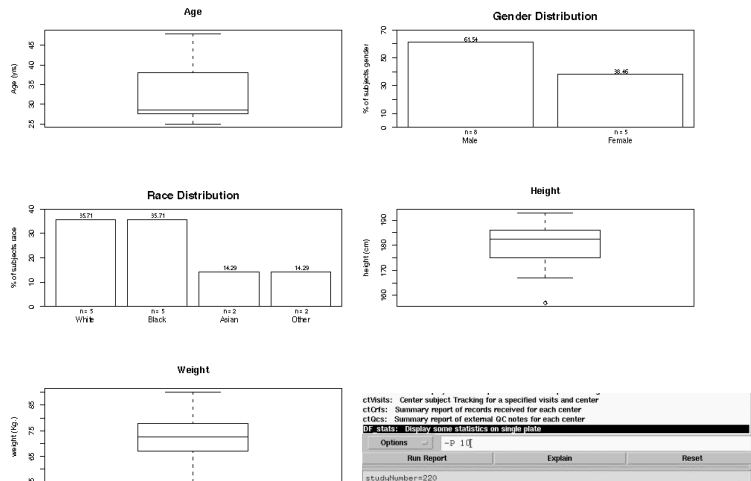
Table 1: Summary Report of Records

Site	Number of Records	Number of QC's	QC's rate as % of Records	Number of Retrieved QC's	% QC's Correct	N/A	Unknown number
1	85	11	13	0	0	0	0
2	56	5	9	14	0	0	0
3	16	9	56	1	11	0	1
4	6	0	0	0	0	0	0
5	70	2	3	0	0	0	0
9	40	15	38	1	7	0	1

Table 2: Summary Report of Records

Site	0-10 days	11-20 days	21-30 days	31-60 days	60 days-10 days	11-20 days	21-30 days	31-60 days	60 days
1	0	0	0	0	11	0	0	0	0
2	0	0	0	0	8	0	0	0	0
3	0	0	0	0	8	0	0	0	0
4	0	0	0	0	2	0	0	0	0
5	0	0	0	0	14	0	0	0	0
9	0	0	0	0	0	0	0	0	0

## DF\_stats - Display simple variable statistics for a single plate



- Comprehensible view of graphs will improve the quality of data
- Simple observations of outliers, error in the units of measurement will be caught

## Remarks

---

The integrated DfRTools with DataFax helps

- Study PIs
  - Monitor accrual rates
  - Identify or correct where QCs are (site or CRF based) and what causes them
  - Creates snap shots and quality of their clinical data
  - Provide statistics in a graphical or table format
- Data Manager
  - Helps determine the work load
  - Detailed information for visit scheduling
  - Quickly identifies sites needing further training or complicated plates requiring extra detailed attention
- Clinical Research Associates
  - Helps to decrease QC rates by paying attention to the specific plate/s or field/s
  - Detailed information for visit scheduling

## Summary

---

- Collaborative effort between Frontier Science and Factory.
- Factory provided the Mock Study and helped to develop specifications for the functions.
- Frontier science develop R functions with specifications from Factory and Frontier science needs.
- Goal is to produce DataFax reports in a comprehensible way.
- Eliminate data transfer between systems and applications.
- DfRTools are embedded into DataFax
- DfRTools scripts can be easily extended/modified to customize report.
- DfRTools package has been implemented at Factory and used in "live" studies.

## Future plans

- DfRTools will be extended for generic reports.
- DfRTools are designed to be community tools.
- It is planned to be distributed as an open source package.
- If it's agreed, the DfRTools can be extended using CVS repository, any new functionality needs to be tested and validated to be included in DfRTools.

## R-information

```

"dfStats" <- function(matrix=plateM)
{
  layout(matrix(c(1,2,3,4,5,6), 3, 2, byrow = TRUE))
  age<-as.numeric(na.omit(plateM[,20]))
  minAge<-min(age,na.rm=)
  maxAge<-max(age,na.rm=)
  minAge<-10*floor(minAge/10)
  maxAge<-10*ceiling(maxAge/10)
  boxplot(age,ylab="Age (yrs)",main="Age",yLim=c(minAge,maxAge))
  #
  gender<-as.factor(na.omit(plateM[,12]))
  levels(gender)<-c("Male","Female")
  tableFillName<-"age"
  genderMat<-matrix(0,nrow=1,ncol=2)
  #colnames(genderMat)<-levels(gender)
  genderMat[1,1]<-length(na.omit(gender[gender=="Male"]))
  genderMat[1,2]<-length(na.omit(gender[gender=="Female"]))
  crfBarPlot(genderMat,titleOfPlot="Gender Distribution",yLabel="% of subjects gender")
  #Race
  race<-as.factor(na.omit(plateM[,13]))
  levels(race)<-c("White","Black","Asian","Other")
  raceMat<-matrix(0,nrow=1,ncol=4)
  colnames(raceMat)<-levels(race)
  #filling up the bins
  raceMat[1,1]<-length(na.omit(race[race=="White"]))
  raceMat[1,2]<-length(na.omit(race[race=="Black"]))
  raceMat[1,3]<-length(na.omit(race[race=="Asian"]))
  raceMat[1,4]<-length(na.omit(race[race=="Other"]))
  crfBarPlot(raceMat,titleOfPlot="Race Distribution",yLabel="% of subjects race")
  #Height
  height<-as.numeric(na.omit(plateM[,15]))
  minHt<-min(height,na.rm=)
  maxHt<-max(height,na.rm=)
  minHt<-10*floor(minHt/10)
  maxHt<-10*ceiling(maxHt/10)
  boxplot(height,ylab="height (cm)",main="Height",yLim=c(minHt,maxHt))
  #Weight
  weight<-as.numeric(na.omit(plateM[,16]))
  minWt<-min(weight,na.rm=)
  maxWt<-max(weight,na.rm=)
  minWt<-10*floor(minWt/10)
  maxWt<-10*ceiling(maxWt/10)
  boxplot(weight,ylab="weight (Kg.)",main="Weight",yLim=c(minWt,maxWt))
}

```

# R-information

```

* checking for file 'DIRTOOLS/DESCRIPTION' ... OK
* Installing *source* package 'dfrTools' ...
** R
** data
chmod: WARNING: can't access /tmp/Rbuild.5201/dfrTools/data/*
** help
>>> Building/Updating help pages for package 'dfrTools'
      Formats: text html latex example
barProbs          text      html      latex    example
colPercentsPie   text      html      latex    example
crfBarPlot        text      html      latex    example
crfPerWk          text      html      latex    example
ctCrfFs           text      html      latex    example
ctQcs             text      html      latex    example
ctVisits          text      html      latex    example
dfBarPlot         text      html      latex    example
dfStats           text      html      latex    example
freqTable         text      html      latex    example
legScales         text      html      latex    example
legendsFs         text      html      latex    example
matPlotDf         text      html      latex    example
matPoints         text      html      latex    example
mergeDataModule   text      html      latex    example
qcsByField        text      html      latex    example
qcsPerSub         text      html      latex    example
rowPercentsPie    text      html      latex    example
visitSchedule     text      html      latex    example
wkToDate          text      html      latex    example
* DONE (dfrTools)

* building 'dfrTools_1.0.3_R_sparc-sun-solaris2.9.tar.gz'

* Installing *binary* package 'dfrTools' ...
* DONE (dfrTools)
gabriel1{otles}46: R CMD INSTALL dfrTools_1.0.3_R_sparc-sun-solaris2.9.tar.gz

```

# R-information

