

project1.R

hahsler

Fri Aug 25 10:08:39 2017

Federally funded Data file from <https://archive.org/download/opm-federal-employment-data/data/2014-09-to-2016-09/non-dod/status/>

```
#dat <- read.table("Non_DoD_201512.txt", sep = "|")
```

```
dat <- read.csv("Non_DoD_201512.txt", sep = "|")
```

```
head(dat)
```

```
##          UPADHYAYA          SUNIT X201512
## 1 NAME WITHHELD BY OPM NAME WITHHELD BY OPM 201512
## 2          MAJUMDAR          RAJTILAK 201512
## 3   CIOBAN CHERESTESIU          CARMEN 201512
## 4 NAME WITHHELD BY OPM NAME WITHHELD BY OPM 201512
## 5          SHAHI          RAJU 201512
## 6 NAME WITHHELD BY OPM NAME WITHHELD BY OPM 201512
##      TD.DEPARTMENT.OF.TRANSPORTATION
## 1 HS-DEPARTMENT OF HOMELAND SECURITY
## 2      AG-DEPARTMENT OF AGRICULTURE
## 3   VA-DEPARTMENT OF VETERANS AFFAIRS
## 4      TR-DEPARTMENT OF THE TREASURY
## 5   VA-DEPARTMENT OF VETERANS AFFAIRS
## 6      TR-DEPARTMENT OF THE TREASURY
##  TD01.OFFICE.OF.THE.SECRETARY.OF.TRANSPORTATION X25 X25.29 X0...4.years
## 1   HSBC-TRANSPORTATION SECURITY ADMINISTRATION ## 30-34 0 - 4 years
## 2          AG03-AGRICULTURAL RESEARCH SERVICE 22 35-39 0 - 4 years
## 3          VATA-VETERANS HEALTH ADMINISTRATION 50 40-44 20 - 24 years
## 4          TR93-INTERNAL REVENUE SERVICE 25 20-24 Unspecified
## 5          VATA-VETERANS HEALTH ADMINISTRATION 25 40-44 15 - 19 years
## 6          TR93-INTERNAL REVENUE SERVICE 25 50-54 Unspecified
##  X13 GS X12      X3...4.years X2210.INFORMATION.TECHNOLOGY.MANAGEMENT A
## 1  13 SV  E      1 - 2 years 1802-COMPLIANCE INSPECTION AND SUPPORT T
## 2  21 GS 11      1 - 2 years          0434-PLANT PATHOLOGY P
## 3  06 VN 02      1 - 2 years          0610-NURSE P
## 4  04 GS 05 Less than 1 year          0962-CONTACT REPRESENTATIVE T
## 5  13 GS 06      3 - 4 years          0681-DENTAL ASSISTANT T
## 6  04 GS 06      1 - 2 years          0592-TAX EXAMINING T
##  X76735.00 X8 X10 F X1
## 1 40002.00 8 38 P 2
## 2 60514.00 4 42 F 2
## 3 71269.00 8 38 F 1
## 4 34922.00 8 15 G 2
## 5 45417.00 8 38 F 1
## 6 38928.00 8 15 F 1
```

Manually edited from <https://archive.org/download/opm-federal-employment-data/data/2014-09-to-2016-09/non-dod/documentation/>

```
dat_header <- read.csv("Data_Record_Format.csv")
```

```
colnames(dat) <- dat_header[,1]
```

```
head(dat)
```

```
##          Family Name          Given Name File Date
## 1 NAME WITHHELD BY OPM NAME WITHHELD BY OPM 201512
## 2          MAJUMDAR          RAJTILAK 201512
## 3    CIOBAN CHERESTESIU          CARMEN 201512
## 4 NAME WITHHELD BY OPM NAME WITHHELD BY OPM 201512
## 5          SHAHI          RAJU 201512
## 6 NAME WITHHELD BY OPM NAME WITHHELD BY OPM 201512
##
##          Agency
## 1 HS-DEPARTMENT OF HOMELAND SECURITY
## 2          AG-DEPARTMENT OF AGRICULTURE
## 3 VA-DEPARTMENT OF VETERANS AFFAIRS
## 4          TR-DEPARTMENT OF THE TREASURY
## 5 VA-DEPARTMENT OF VETERANS AFFAIRS
## 6          TR-DEPARTMENT OF THE TREASURY
##
##          SubAgency State AgeRange          YSD
## 1 HSBC-TRANSPORTATION SECURITY ADMINISTRATION ## 30-34 0 - 4 years
## 2          AG03-AGRICULTURAL RESEARCH SERVICE 22 35-39 0 - 4 years
## 3          VATA-VETERANS HEALTH ADMINISTRATION 50 40-44 20 - 24 years
## 4          TR93-INTERNAL REVENUE SERVICE 25 20-24 Unspecified
## 5          VATA-VETERANS HEALTH ADMINISTRATION 25 40-44 15 - 19 years
## 6          TR93-INTERNAL REVENUE SERVICE 25 50-54 Unspecified
##
## Education PayPlan Grade          LOS
## 1          13      SV      E      1 - 2 years
## 2          21      GS      11      1 - 2 years
## 3          06      VN      02      1 - 2 years
## 4          04      GS      05 Less than 1 year
## 5          13      GS      06      3 - 4 years
## 6          04      GS      06      1 - 2 years
##
##          Occupation OccupationalCategory
## 1 1802-COMPLIANCE INSPECTION AND SUPPORT T
## 2          0434-PLANT PATHOLOGY P
## 3          0610-NURSE P
## 4          0962-CONTACT REPRESENTATIVE T
## 5          0681-DENTAL ASSISTANT T
## 6          0592-TAX EXAMINING T
##
## AdjustedBasicPay Supervisor TOA Work Schedule NSFTP
## 1          40002.00          8 38          P 2
## 2          60514.00          4 42          F 2
## 3          71269.00          8 38          F 1
## 4          34922.00          8 15          G 2
## 5          45417.00          8 38          F 1
## 6          38928.00          8 15          F 1
```

```
summary(dat)
```

```
##          Family Name          Given Name
## NAME WITHHELD BY OPM :269602 NAME WITHHELD BY OPM :269602
## NAME WITHHELD BY AGENCY: 69442 NAME WITHHELD BY AGENCY: 69442
## SMITH          : 8580 MICHAEL          : 18515
## JOHNSON          : 7620 JOHN          : 14640
## WILLIAMS          : 6604 DAVID          : 14540
## (Other)          :970747 (Other)          :945856
```

```

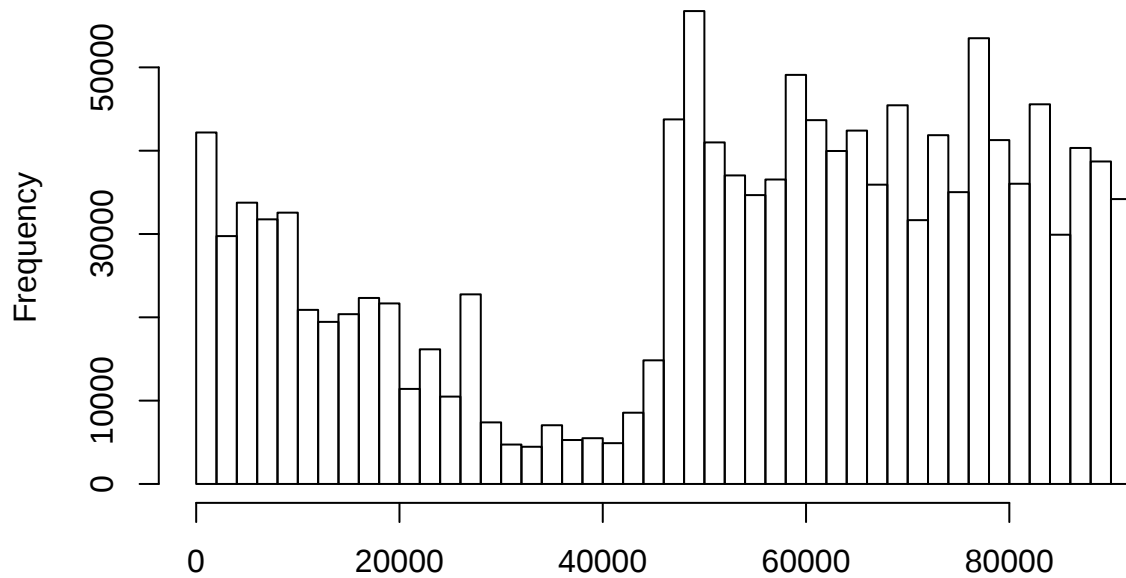
## NA's          :      6  NA's          :      6
##   File Date          Agency
## Min.   :201512  VA-DEPARTMENT OF VETERANS AFFAIRS      :366343
## 1st Qu.:201512  HS-DEPARTMENT OF HOMELAND SECURITY      :188527
## Median :201512  DJ-DEPARTMENT OF JUSTICE                  :115498
## Mean   :201512  TR-DEPARTMENT OF THE TREASURY                  : 95003
## 3rd Qu.:201512  AG-DEPARTMENT OF AGRICULTURE                    : 87305
## Max.   :201512  HE-DEPARTMENT OF HEALTH AND HUMAN SERVICES: 85937
##                                     (Other)              :393988
##                                     SubAgency            State
## VATA-VETERANS HEALTH ADMINISTRATION      :328985  ##      :269386
## TR93-INTERNAL REVENUE SERVICE              : 80267  11      :128945
## SZ00-SOCIAL SECURITY ADMINISTRATION        : 65518  24      : 86118
## HSBD-CUSTOMS AND BORDER PROTECTION         : 59400  06      : 77119
## HSBC-TRANSPORTATION SECURITY ADMINISTRATION: 58753  48      : 63129
## TD03-FEDERAL AVIATION ADMINISTRATION      : 45756  12      : 46603
## (Other)                                   :693922  (Other):661301
##   AgeRange          YSD          Education          PayPlan
## 50-54 :208111  Unspecified :504021  13      :361455  GS      :937346
## 45-49 :189148   5 - 9 years :149992  04      :268538  VN      : 63420
## 55-59 :182750  10 - 14 years:133335  17      :180305  SV      : 58611
## 35-39 :160524  15 - 19 years:120267  10      : 82771  WG      : 38122
## 40-44 :160244   0 - 4 years : 98230  15      : 73747  AD      : 35310
## 30-34 :144972  20 - 24 years: 96278  07      : 64545  (Other):197674
## (Other):286852  (Other)      :230478  (Other):301240  NA's    : 2118
##   Grade          LOS
## 12      :179170   5 - 9 years :322631
## 13      :178456  10 - 14 years:217917
## 11      :117337  15 - 19 years:132402
## 14      : 99829   1 - 2 years :120396
## 07      : 85545  25 - 29 years:120240
## 09      : 84683   3 - 4 years :108303
## (Other):587581  (Other)      :310712
##                                     Occupation
## 0610-NURSE                      : 72876
## 0301-MISCELLANEOUS ADMINISTRATION AND PROGRAM: 61652
## 1802-COMPLIANCE INSPECTION AND SUPPORT      : 50694
## 2210-INFORMATION TECHNOLOGY MANAGEMENT      : 46473
## 0343-MANAGEMENT AND PROGRAM ANALYSIS        : 44304
## 1811-CRIMINAL INVESTIGATION                 : 40998
## (Other)                                   :1015604
## OccupationalCategory  AdjustedBasicPay  Supervisor  TOA
## *: 217                158700.00: 16112  *: 82  10      :720589
## A:511271              79554.00 : 14197  2: 166691  38      :366518
## B: 56523              77213.00 : 12851  4: 15310  15      : 93379
## C: 72308              31944.00 : 10316  5: 10682  48      : 54451
## O: 57612              139523.00: 8811   6: 7672  30      : 42366
## P:379520              118069.00: 7938   7: 6843  20      : 17913
## T:255150              (Other) :1262376  8:1125321 (Other): 37385
## Work Schedule      NSFTP
## F      :1226779  Min.   :1.000
## P      : 44678  1st Qu.:1.000
## I      : 40017  Median :1.000
## G      : 18876  Mean   :1.116

```

```
## J      : 1514 3rd Qu.:1.000
## Q      : 655 Max.   :2.000
## (Other): 82
```

```
dat$AdjustedBasicPay <- as.numeric(dat$AdjustedBasicPay)
hist(dat$AdjustedBasicPay, breaks = 50)
```

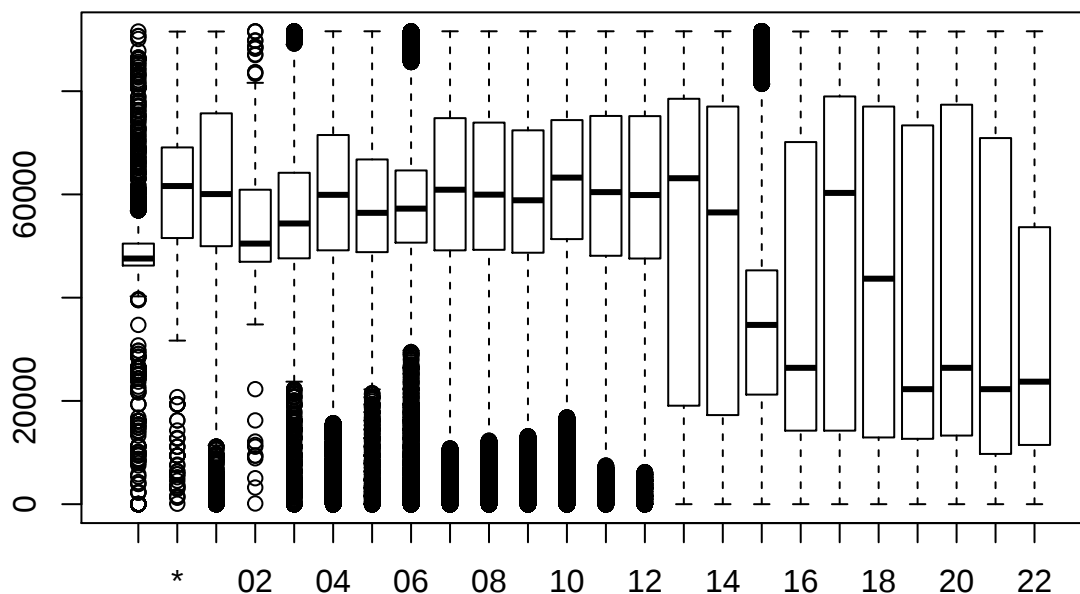
Histogram of dat\$AdjustedBasicPay



dat\$AdjustedBasicPay

Education <https://ia800608.us.archive.org/16/items/opm-federal-employment-data/data/2014-09-to-2016-09/non-dod/translations/Education%20Translation.txt>

```
boxplot(AdjustedBasicPay~Education, data = dat)
```



```
med_pay_by_agency <- aggregate(AdjustedBasicPay~Agency, data = dat, median)
med_pay_by_agency[head(order(med_pay_by_agency$AdjustedBasicPay)),]
```

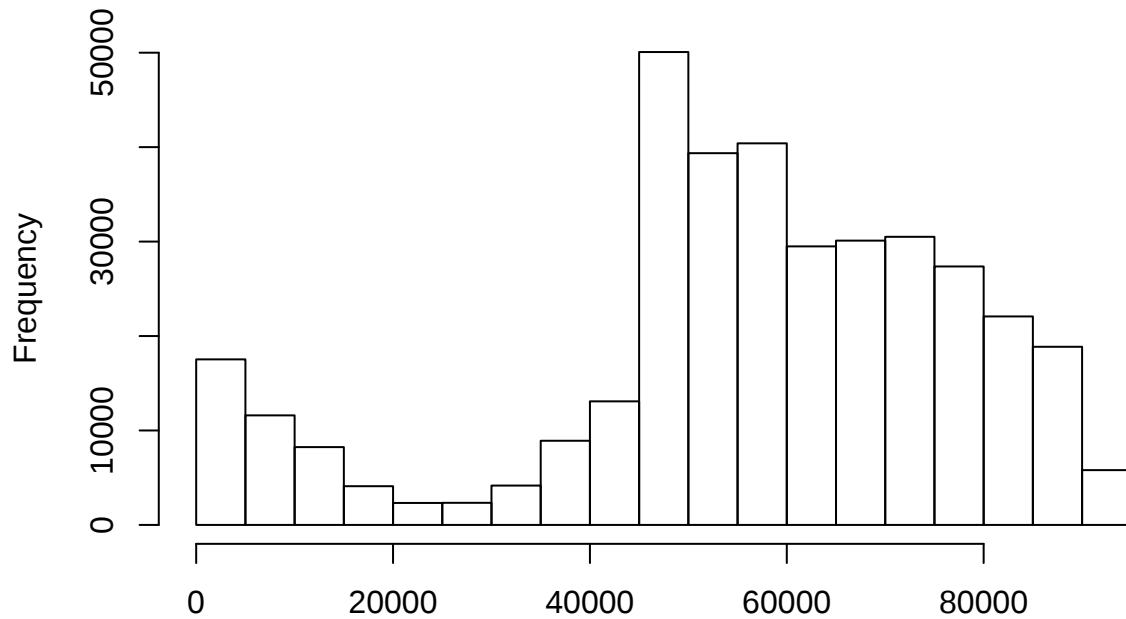
```
##
## 17 BZ-CHRISTOPHER COLUMBUS FELLOWSHIP FOUNDATION
## 51 GE-BARRY GOLDWATER SCHOLARSHIP AND EXCELLENCE IN EDUCATION FOUNDATION
## 47 FR-FEDERAL RESERVE SYSTEM
## 24 CX-NATIONAL COMMISSION ON LIBRARIES AND INFORMATION SCIENCE
## 116 VB-COMMISSION TO ELIMINATE CHILD ABUSE AND NEGLECT FATALITIES
## 25 DB-PUBLIC INTEREST DECLASSIFICATION BOARD
## AdjustedBasicPay
## 17 1.0
## 51 1.0
## 47 2.0
## 24 11227.0
## 116 12090.5
## 25 12803.5
```

```
med_pay_by_agency[head(order(med_pay_by_agency$AdjustedBasicPay,
decreasing = TRUE)),]
```

```
##
## 46 FQ-COURT SERVICES AND OFFENDER SUPERVISION AGENCY FOR THE DISTRICT OF COLUMBIA
## 70 IB-BROADCASTING BOARD OF GOVERNORS
## 59 GX-INTERNATIONAL BOUNDARY COMMISSION: UNITED STATES AND CANADA
## 90 OM-OFFICE OF PERSONNEL MANAGEMENT
## 53 GJ-PRESIDIO TRUST
## 75 KS-CORPORATION FOR NATIONAL AND COMMUNITY SERVICE
## AdjustedBasicPay
## 46 77258
## 70 77258
## 59 72538
## 90 69105
## 53 68622
## 75 67965
```

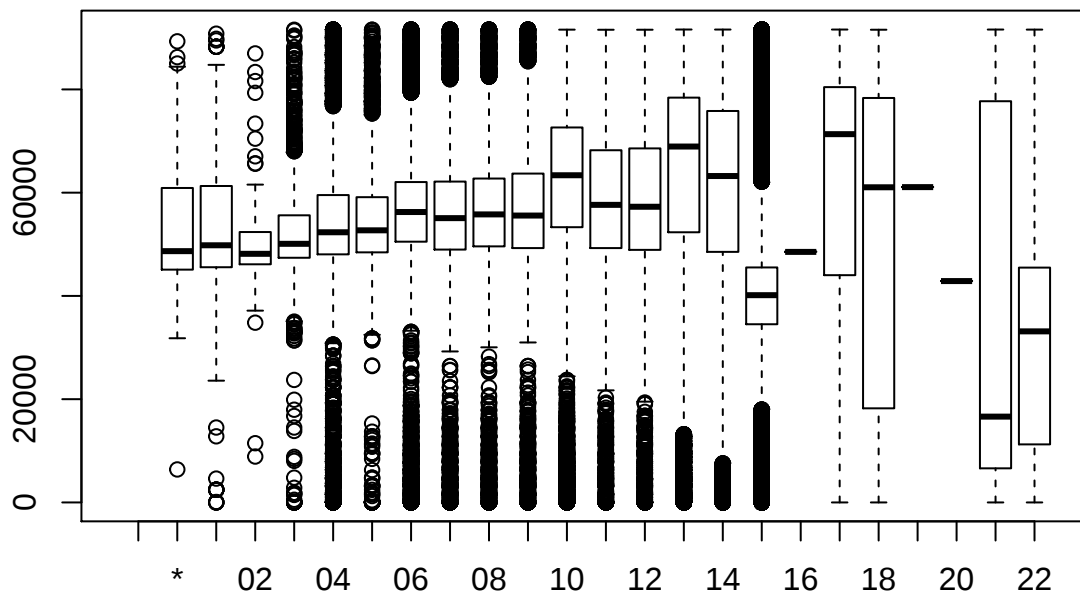
```
va <- dat[dat$Agency=="VA-DEPARTMENT OF VETERANS AFFAIRS", ]
hist(va$AdjustedBasicPay)
```

Histogram of va\$AdjustedBasicPay



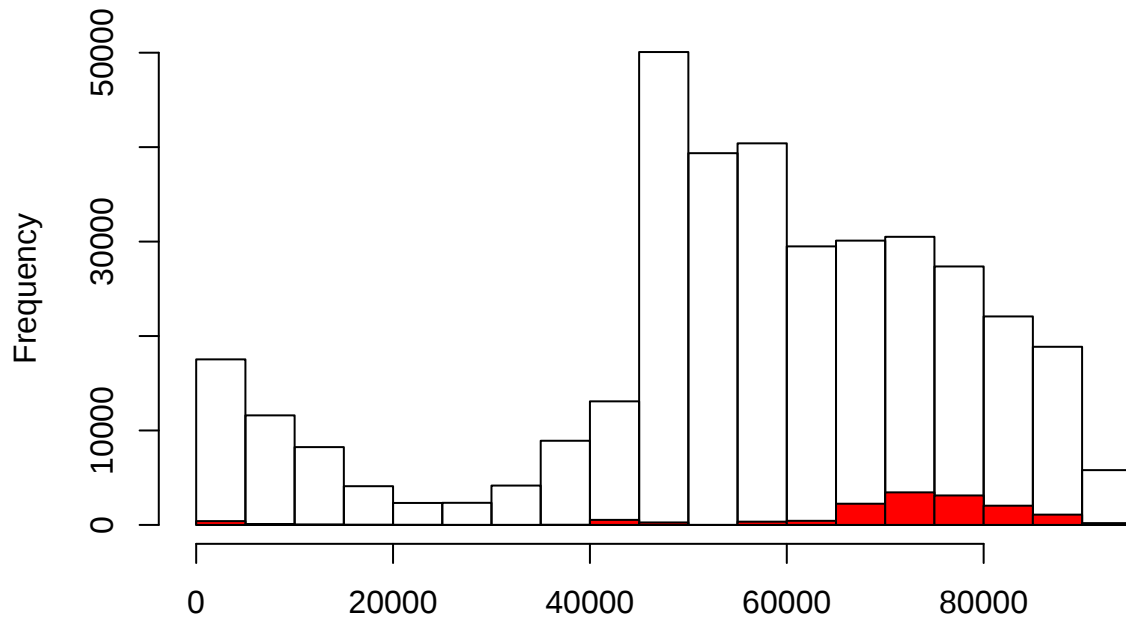
va\$AdjustedBasicPay

```
boxplot(AdjustedBasicPay~Education, data = va)
```



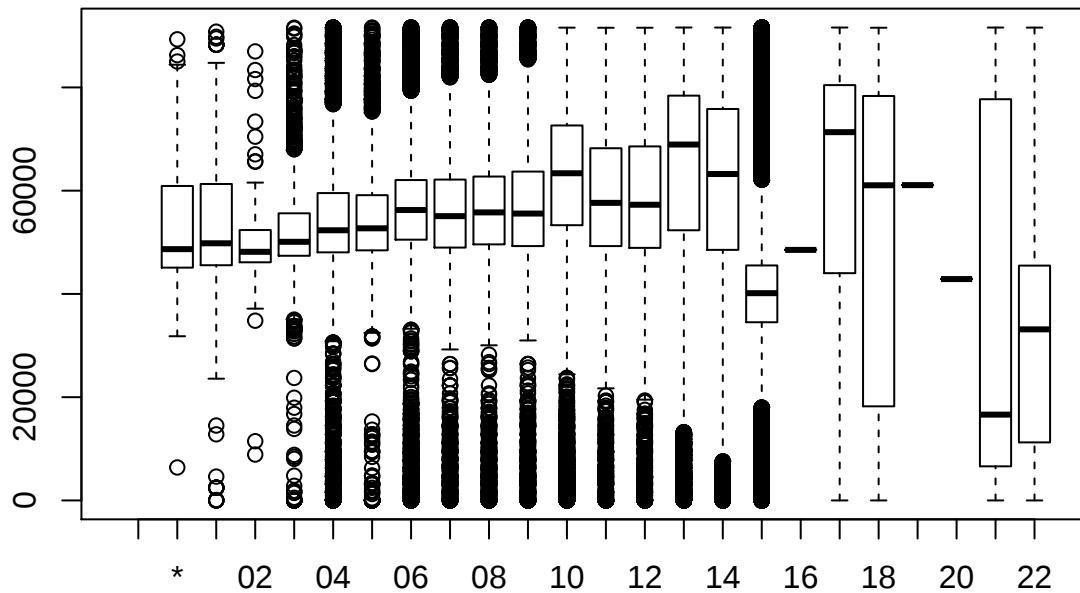
```
va_sw <- dat[dat$Agency=="VA-DEPARTMENT OF VETERANS AFFAIRS" & dat$Occupation=="0185-SOCIAL WORK", ]
hist(va$AdjustedBasicPay)
hist(va_sw$AdjustedBasicPay, add = TRUE, col = "red")
```

Histogram of va\$AdjustedBasicPay



va\$AdjustedBasicPay

```
boxplot(AdjustedBasicPay~Education, data = va)
```

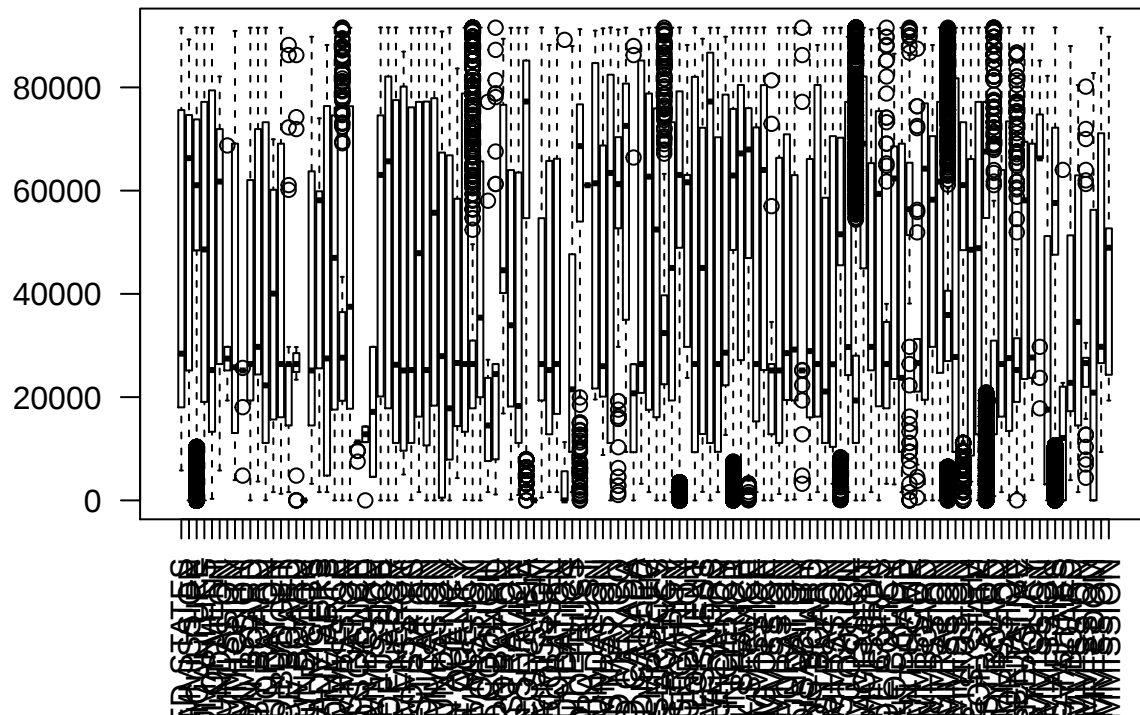


```
employees_by_agency <- aggregate(AdjustedBasicPay~Agency, data = dat, length)
med_pay_by_agency[head(order(employees_by_agency$AdjustedBasicPay,
decreasing = TRUE), n = 10),]
```

##	Agency	AdjustedBasicPay
## 115	VA-DEPARTMENT OF VETERANS AFFAIRS	57608
## 66	HS-DEPARTMENT OF HOMELAND SECURITY	63038
## 27	DJ-DEPARTMENT OF JUSTICE	63047
## 111	TR-DEPARTMENT OF THE TREASURY	58116

```
## 3          AG-DEPARTMENT OF AGRICULTURE          61051
## 63 HE-DEPARTMENT OF HEALTH AND HUMAN SERVICES    52466
## 106        SZ-SOCIAL SECURITY ADMINISTRATION      67571
## 73         IN-DEPARTMENT OF THE INTERIOR          62932
## 109        TD-DEPARTMENT OF TRANSPORTATION       27562
## 21         CM-DEPARTMENT OF COMMERCE             46975
```

```
boxplot(AdjustedBasicPay~Agency, data = dat, las = 2)
```



```
dat_sep <- read.csv("Separations Non_DoD FY2015Q4.txt", sep = "|")
head(dat_sep)
```

```
## TANIGUCHI ALINE X20150828 VA.DEPARTMENT.OF.VETERANS.AFFAIRS
## 1 KEEN WAYNE 20150930 VA-DEPARTMENT OF VETERANS AFFAIRS
## 2 RENNIE DALE 20150930 IN-DEPARTMENT OF THE INTERIOR
## 3 DANIELS LEANNE 20150930 VA-DEPARTMENT OF VETERANS AFFAIRS
## 4 CHEVALIER RAYMOND 20150930 HS-DEPARTMENT OF HOMELAND SECURITY
## 5 CYR JEANNINE 20150823 HE-DEPARTMENT OF HEALTH AND HUMAN SERVICES
## 6 ROBINSON JOSEPH 20150808 VA-DEPARTMENT OF VETERANS AFFAIRS
## VATA.VETERANS.HEALTH.ADMINISTRATION SD X37
## 1 VATA-VETERANS HEALTH ADMINISTRATION SD 06
## 2 IN05-BUREAU OF LAND MANAGEMENT SD 41
## 3 VATA-VETERANS HEALTH ADMINISTRATION SD 33
## 4 HSCB-FEDERAL EMERGENCY MANAGEMENT AGENCY SD 24
## 5 HE10-OFFICE OF THE SECRETARY OF HEALTH AND HUMAN SERVICES SJ 11
## 6 VATA-VETERANS HEALTH ADMINISTRATION SJ 45
## X65.or.more X15...19.years X17 AD X03 X30...34.years
## 1 65 or more 30 - 34 years 17 GS 13 20 - 24 years
## 2 65 or more 35 - 39 years 13 GS 09 10 - 14 years
## 3 55-59 25 - 29 years 06 GS 12 35 or more
## 4 60-64 20 - 24 years 13 GS 14 35 or more
## 5 60-64 Unspecified 07 GS 14 15 - 19 years
```

```
## 6      60-64  25 - 29 years  13 AD  12      5 - 9 years
##                               X0610.NURSE P X127388.00 X48 F X2
## 1                               1102-CONTRACTING P  114665.00  38 F  1
## 2      2210-INFORMATION TECHNOLOGY MANAGEMENT A    67977.00  10 F  1
## 3                               1102-CONTRACTING P    92085.00  10 F  1
## 4                               1712-TRAINING INSTRUCTION A  128790.00  10 F  1
## 5 0301-MISCELLANEOUS ADMINISTRATION AND PROGRAM A  128790.00  20 I  2
## 6                               0603-PHYSICIAN ASSISTANT A   91255.00  48 F  2
```

```
dat_dyn_header <- read.csv("Data_Record_Format_dynamic.csv")
colnames(dat_sep) <- dat_dyn_header[,1]
head(dat_sep)
```

```
##  FamilyName GivenName EffectiveDate
## 1      KEEN      WAYNE      20150930
## 2      RENNIE      DALE      20150930
## 3      DANIELS     LEANNE     20150930
## 4  CHEVALIER     RAYMOND     20150930
## 5      CYR  JEANNINE     20150823
## 6  ROBINSON     JOSEPH     20150808
##
##                               Agency
## 1      VA-DEPARTMENT OF VETERANS AFFAIRS
## 2      IN-DEPARTMENT OF THE INTERIOR
## 3      VA-DEPARTMENT OF VETERANS AFFAIRS
## 4      HS-DEPARTMENT OF HOMELAND SECURITY
## 5  HE-DEPARTMENT OF HEALTH AND HUMAN SERVICES
## 6      VA-DEPARTMENT OF VETERANS AFFAIRS
##
##                               SubAgency AccSep State
## 1      VATA-VETERANS HEALTH ADMINISTRATION      SD    06
## 2      INO5-BUREAU OF LAND MANAGEMENT          SD    41
## 3      VATA-VETERANS HEALTH ADMINISTRATION      SD    33
## 4      HSCB-FEDERAL EMERGENCY MANAGEMENT AGENCY SD    24
## 5  HE10-OFFICE OF THE SECRETARY OF HEALTH AND HUMAN SERVICES SJ    11
## 6      VATA-VETERANS HEALTH ADMINISTRATION      SJ    45
##
##      Age YearsSinceDegree EducationLevel PayPlan Grade      LOS
## 1 65 or more      30 - 34 years          17      GS   13 20 - 24 years
## 2 65 or more      35 - 39 years          13      GS   09 10 - 14 years
## 3   55-59      25 - 29 years            06      GS   12   35 or more
## 4   60-64      20 - 24 years            13      GS   14   35 or more
## 5   60-64      Unspecified              07      GS   14 15 - 19 years
## 6   60-64      25 - 29 years            13      AD   12   5 - 9 years
##
##                               Occupation OccupationalCategory
## 1                               1102-CONTRACTING      P
## 2      2210-INFORMATION TECHNOLOGY MANAGEMENT      A
## 3                               1102-CONTRACTING      P
## 4                               1712-TRAINING INSTRUCTION      A
## 5 0301-MISCELLANEOUS ADMINISTRATION AND PROGRAM      A
## 6                               0603-PHYSICIAN ASSISTANT      A
##
##  AdjustedBasicPay TOA WorkSchedule NSFTP
## 1      114665.00  38      F      1
## 2      67977.00  10      F      1
## 3      92085.00  10      F      1
## 4     128790.00  10      F      1
## 5     128790.00  20      I      2
## 6     91255.00  48      F      2
```