

# A Pandoc Markdown Article Starter and Template \*

**James Harper, PE, ENV SP**     *University of Colorado Boulder*

---

This document provides an introduction to R Markdown, argues for its benefits, and presents a sample manuscript template intended for an academic audience. I include basic syntax to R Markdown and a minimal working example of how the analysis itself can be conducted within R with the knitr package.

*Keywords:* pandoc, r markdown, knitr

---

## Results

```
## Loading required package: rio

## Warning: package 'rio' was built under R version 3.4.4

## Loading required package: gplots

##
## Attaching package: 'gplots'

## The following object is masked from 'package:stats':
##
##      lowess

## Loading required package: Amelia

## Loading required package: Rcpp

## Warning: package 'Rcpp' was built under R version 3.4.4

## ##
## ## Amelia II: Multiple Imputation
## ## (Version 1.7.4, built: 2015-12-05)
## ## Copyright (C) 2005-2018 James Honaker, Gary King and Matthew Blackwell
## ## Refer to http://gking.harvard.edu/amelia/ for more information
## ##

## Loading required package: car

## Warning: package 'car' was built under R version 3.4.4

## Loading required package: carData
```

---

\*Replication files are available on the author's Github account (<http://github.com/svmiller>). **Current version:** August 21, 2018; **Corresponding author:** [svmille@clemsun.edu](mailto:svmille@clemsun.edu).

```
## Warning: package 'carData' was built under R version 3.4.4
## Loading required package: fastDummies
## Warning: package 'fastDummies' was built under R version 3.4.4
## Loading required package: corrplot
## corrplot 0.84 loaded
##
## Attaching package: 'corrplot'
##
## The following object is masked _by_ '.GlobalEnv':
##
##      cor.mtest
##
## Loading required package: FactoMineR
## Warning: package 'FactoMineR' was built under R version 3.4.4
## Loading required package: factoextra
## Loading required package: ggplot2
## Welcome! Related Books: 'Practical Guide To Cluster Analysis in R' at https://goo.gl/13EFCZ
## Loading required package: lavaan
## Warning: package 'lavaan' was built under R version 3.4.4
## This is lavaan 0.6-2
## lavaan is BETA software! Please report any bugs.
```

```
##      Date                      LBO                      Prov
## Min.      :2014-10-26  Mang Srey(Kandal)1: 188  Banteay Meanchey: 638
## 1st Qu.   :2016-10-11  Ream Ny (Kandal) : 169  Kampong Thom    : 709
## Median    :2017-01-25  Mang Srey(Kandal)2: 142  Kandal          :1278
## Mean      :2017-01-09  Ngim Houn        : 119  Oddar Meanchey : 387
## 3rd Qu.   :2017-06-15  Sorn Kuon        : 112  Prey Veng       : 977
## Max.      :2017-10-27  Hong Davy        : 104  Siem Reap      : 788
##              (Other)          :4635  Svay Rieng     : 692
##
##      Dist      Comm      Vill      IDPoor
## Kaoh Thum      : 239  Prasat   : 93  Thmei      : 83  No :4205
## Preah Sdach    : 195  Ampil    : 74  Prasat     : 40  Yes:1264
## Ponhea Lueu     : 191  Leuk Daek: 66  Chheu Teal: 35
## Svay Chrum      : 189  Beng     : 65  Kandal     : 34
## Banteay Ampil   : 184  Mkak     : 61  Svay       : 28
## Preah Netr Preah: 168  Chhveang : 60  Sambuor    : 26
## (Other)         :4303  (Other)  :5050  (Other)    :5223
```

```

## IDPoorTyp      M01      M1824      M217
## IDP1 : 378  Min.   :0.00000  Min.   : 0.000  Min.   :0.0000
## IDP2 : 758  1st Qu.:0.00000  1st Qu.: 0.000  1st Qu.:0.0000
## UnkIDP: 128  Median :0.00000  Median : 0.000  Median :1.0000
## No    :4205  Mean   :0.09088  Mean   : 0.434  Mean   :0.8073
##          3rd Qu.:0.00000  3rd Qu.: 1.000  3rd Qu.:1.0000
##          Max.   :3.00000  Max.   :18.000  Max.   :9.0000
##
##          NA's    :1
##      M2559      M60      F01      F1824
## Min.   :0.000  Min.   :0.0000  Min.   :0.00000  Min.   :0.0000
## 1st Qu.:1.000  1st Qu.:0.0000  1st Qu.:0.00000  1st Qu.:0.0000
## Median :1.000  Median :0.0000  Median :0.00000  Median :0.0000
## Mean   :1.027  Mean   :0.1514  Mean   :0.08246  Mean   :0.3959
## 3rd Qu.:1.000  3rd Qu.:0.0000  3rd Qu.:0.00000  3rd Qu.:1.0000
## Max.   :6.000  Max.   :8.0000  Max.   :8.00000  Max.   :8.0000
##
##      F217      F2559      F60      LivRP
## Min.   :0.0000  Min.   : 0.000  Min.   :0.0000  No :4410
## 1st Qu.:0.0000  1st Qu.: 1.000  1st Qu.:0.0000  Pond: 414
## Median :1.0000  Median : 1.000  Median :0.0000  Rivr: 645
## Mean   :0.8104  Mean   : 1.123  Mean   :0.2293
## 3rd Qu.:1.0000  3rd Qu.: 1.000  3rd Qu.:0.0000
## Max.   :9.0000  Max.   :38.000  Max.   :2.0000
## NA's    :1
## VillOD      DateSlabPur      LatInst      BelowGrndInst
## Most: 453  Min.   :2011-01-01  Installed :2531  In-Process : 11
## None: 268  1st Qu.:2016-01-31  Installing : 91  Installed :1308
## Some:4748  Median :2016-05-24  Uninstalled:1390  Uninstalled: 138
##          Mean   :2016-04-17  NA's      :1457  NA's      :4012
##          3rd Qu.:2016-10-18
##          Max.   :2017-12-31
##
## DateBelowGrndInst      ShltrInst      DateInstComp
## Min.   :2016-10-18  In-process : 24  Min.   :2014-01-01
## 1st Qu.:2017-02-14  Installed  : 959  1st Qu.:2016-03-09
## Median :2017-03-10  Uninstalled: 325  Median :2016-06-20
## Mean   :2017-03-13  NA's      :4161  Mean   :2016-05-28
## 3rd Qu.:2017-04-10          3rd Qu.:2016-11-24
## Max.   :2017-12-28          Max.   :2019-04-01
## NA's    :5320          NA's    :1981
##
##      RDefBefor
## Bsh/Fld      :2707
## NeiToi       : 397
## Othr         : 316
## NeiToi;Bsh/Fld: 152
## NA/AlwysToi  : 91
## (Other)      : 80
## NA's        :1726

```



```

##                               RoofMat
## Galvanized steel              :2667
## No roof                      : 349
## Bamboo / Palm Leaves / Thatch: 278
## Plastic Sheet                : 129
## Other                        : 31
## (Other)                     : 34
## NA's                        :1981
##
## <U+1780><U+17D2><U+179A><U+178E><U+17B6><U+178F><U+17CB>
## <U+178F><U+1784><U+17CB>
## <U+1792><U+17D2><U+179C><U+17BE><U+1780><U+17D2><U+179A><U+17C4><U+1798><U+1795><U+17D2><U+1792><U+179C><U+17BE><U+179B><U+17BE><U+1795><U+17D2><U+1791><U+17C7>
## <U+1780><U+17D2><U+1793><U+17BB><U+1784><U+1795><U+17D2><U+1791><U+17C7>
## (Other)
## NA's
## MatsPurTgthr      Cost      CostInclInst  CostPitSlab
## Same time :1368    Min.      :      0    No : 402    Min.      :      0
## Separately:2091    1st Qu.: 900000    Yes : 965    1st Qu.: 230000
## NA's       :2010    Median : 9999999    NA's:4102    Median : 300000
##                               Mean      : 6325986    Mean      :3557193
##                               3rd Qu.: 9999999    3rd Qu.:9999999
##                               Max.      :10800000    Max.      :9999999
##                               NA's      :2196      NA's      :1837
## CostPitSlabKnwn CostShltrMats
## Don't Know: 69    Min.      :      0
## Riel         :2204    1st Qu.: 120000
## Yes          : 21    Median :1200000
## NA's         :3175    Mean      :4089093
##                               3rd Qu.:9999999
##                               Max.      :9999999
##                               NA's      :1943
##
##                               CostShltrMatsKnwn
## <U+1785><U+17C6><U+1793><U+17BD><U+1793><U+200B> (<U+179A><U+17C0><U+179B>):1066
## Don't Know                                           : 229
## Riel                                               : 977
## Yes                                              : 30
## NA's                                             :3167
##
##
## LabPitShltrPurTgthr      CostLabLat      CostLabPitSlab
## Did not pay any installation: 150    Min.      :      0    Min.      :      0
## Did not pay for installation:1362    1st Qu.:9999999    1st Qu.:9999999
## Same time                : 55    Median :9999999    Median :9999999
## Separately                : 923    Mean      :9774832    Mean      :9676899
## NA's                      :2979    3rd Qu.:9999999    3rd Qu.:9999999
##                               Max.      :9999999    Max.      :9999999
##                               NA's      :3019      NA's      :3005

```

```

##                               CostLabPitSlabKnwn  CostLabShltr
## 99                               :    1      Min.    :    0
## Already included in latrine price: 808      1st Qu.: 457500
## Don't Know                               :   20      Median :9999999
## Riel                                     :   52      Mean   :6992244
## Yes                                      :   14      3rd Qu.:9999999
## NA's                                    :4574      Max.    :9999999
##                                         NA's     :2641
##      CostLabShltrKnwn
## Don't Know: 62
## Riel       : 818
## Yes        :   4
## NA's       :4585
##
##
##
## <U+1794><U+17D2><U+179A><U+17BE><U+179A><U+17BD><U+1798><U+1782><U+17D2><U+1793><U+17B6><U+1
## Dry Latrine
## No Latrine
## Shared Latrine with Others
## Wet Latrine
## NA's
##
## IntndChngDich          IntndChng
## No :1616      Shltr          : 640
## Yes :1872      Shltr;Shwr;WtrRes: 292
## NA's:1981      Shltr;WtrRes    : 216
##                               Shltr;Shwr      : 125
##                               Shltr;Pit       : 101
##                               (Other)        : 498
##                               NA's           :3597
##
##                                         IntndChngOthr
## Painting the walls and make new floor          :    9
## <U+179A><U+17C0><U+1794><U+1780><U+17B6><U+179A><U+17C9><U+17BC>          :    5
## <U+1780><U+17D2><U+179A><U+17B6><U+179B><U+1780><U+17B6><U+179A><U+17C9><U+17BC>:    3
## <U+179B><U+17B6><U+1794><U+1790><U+17D2><U+1793><U+17B6><U+17C6>          :    3
## Painting the walls                             :    3
## (Other)                                         :   71
## NA's                                           :5375
## AdltUseLat  ChldUseLat  InfLatDump  IntndPitFull
## Freq:3398    Freq :2821    Freq : 144    DK      : 613
## DK :    1    NoChld:  20    NoInf:  89    EmpSlf: 782
## Rare:  22    DK/NA : 343    DK/NA:2314    Pit     : 883
## Some:  68    Rare  :  89    Rare  : 586    Othr   :  65
## NA's:1980    Some   : 163    Some  : 115    Pay    :1301
##                               NA's   :2033    NA's  :2221    Stop   :  77
##                               NA's   :2033    NA's  :2221    NA's   :1748

```



```

## IntndPitFullDes IntndPitFullPit IntndPitFullEmpSlf IntndPitFullDK
## 0 :1537 0 :2838 0 :2939 0 :3108
## 1 :2184 1 : 883 1 : 782 1 : 613
## NA's:1748 NA's:1748 NA's:1748 NA's:1748
##
##
##
## IntndPitFullOthr IntndPitFullPay IntndPitFullStop ChlngsNoFlsh
## 0 :3656 0 :2420 0 :3644 0 :3251
## 1 : 65 1 :1301 1 : 77 1 : 234
## NA's:1748 NA's:1748 NA's:1748 NA's:1984
##
##
##
## ChlngsFlood ChlngsOthr ChlngsFulOvrFlw ChlngsNoWtr ChlngsSmels
## 0 :3452 0 :3404 0 :3463 0 :3458 0 :3401
## 1 : 33 1 : 81 1 : 22 1 : 27 1 : 84
## NA's:1984 NA's:1984 NA's:1984 NA's:1984 NA's:1984
##
##
##
## ChlngsOK SatisColaps SatisSupColaps SatisColapsMore SatisSupColapsMore
## 0 : 426 1 : 189 1 : 234 1 : 877 1 :1213
## 1 :3059 2 : 688 2 : 979 2 :4252 2 :4100
## NA's:1984 3 :4252 3 :4100 DK : 86 DK : 154
## DK : 86 DK : 154 NA's: 254 NA's: 2
## NA's: 254 NA's: 2
##
##
## [1] "Date" "LB0" "Prov"
## [4] "Dist" "Comm" "Vill"
## [7] "IDPoor" "IDPoorTyp" "M01"
## [10] "M1824" "M217" "M2559"
## [13] "M60" "F01" "F1824"
## [16] "F217" "F2559" "F60"
## [19] "LivRP" "VillOD" "DateSlabPur"
## [22] "LatInst" "BelowGrndInst" "DateBelowGrndInst"
## [25] "ShltrInst" "DateInstComp" "RDefBefor"
## [28] "RDefBeforOthr" "FreqNeiToi" "WhoInstLat"
## [31] "KnwSubsdy" "RecSubsdy" "BorwLat"
## [34] "CanBuyLat" "UseFincOthr" "SlabTil"
## [37] "Npits" "PitConfig" "NringsDir"
## [40] "NringsOff" "NringsOthr" "WallMat"
## [43] "WallMatOthr" "RoofMat" "RoofMatOthr"

```

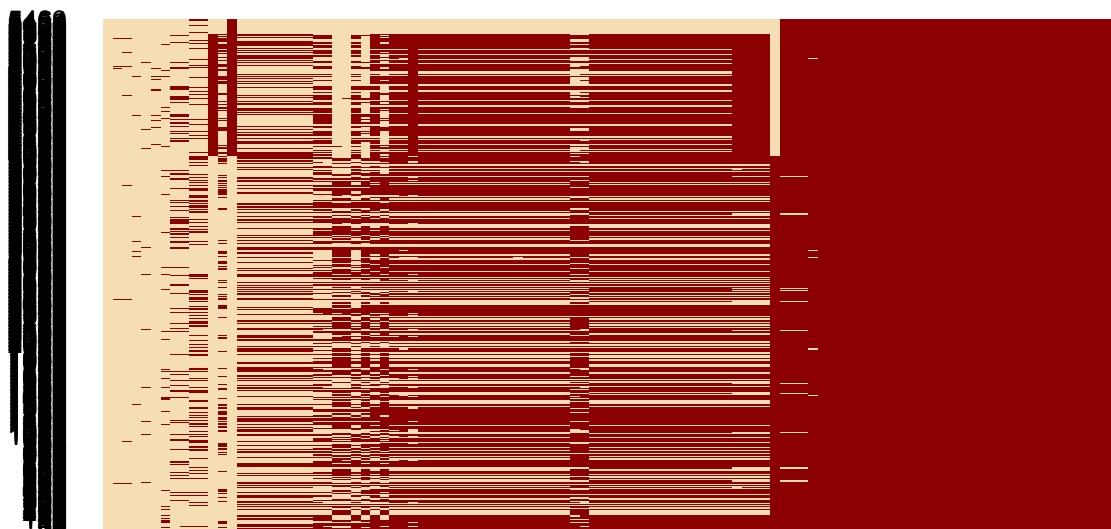


##	[46]	"MatsPurTgthr"	"Cost"	"CostInclInst"
##	[49]	"CostPitSlab"	"CostPitSlabKnwn"	"CostShltrMats"
##	[52]	"CostShltrMatsKnwn"	"LabPitShltrPurTgthr"	"CostLabLat"
##	[55]	"CostLabPitSlab"	"CostLabPitSlabKnwn"	"CostLabShltr"
##	[58]	"CostLabShltrKnwn"	"LatTypOwndBefor"	"IntndChngDich"
##	[61]	"IntndChng"	"IntndChngOthr"	"AdltUseLat"
##	[64]	"ChldUseLat"	"InfLatDump"	"IntndPitFull"
##	[67]	"IntndPitFullOthrAns"	"Chlngs"	"Satis"
##	[70]	"Rec"	"SatisSup"	"RecSup"
##	[73]	"DateSurvCreated"	"Yr"	"Mnth"
##	[76]	"YrMnth"	"RDefBefor_BshFld"	"RDefBefor_RivPnd"
##	[79]	"RDefBefor_NeiToi"	"RDefBefor_Othr"	"RDefBefor_NAAlwysToi"
##	[82]	"IntndChng_Shltr"	"IntndChng_Shwr"	"IntndChng_Sink"
##	[85]	"IntndChng_WtrRes"	"IntndChng_Pit"	"IntndChng_Othr"
##	[88]	"IntndChng_NAAlwysToi"	"IntndPitFullDes"	"IntndPitFullPit"
##	[91]	"IntndPitFullEmpSlf"	"IntndPitFullDK"	"IntndPitFullOthr"
##	[94]	"IntndPitFullPay"	"IntndPitFullStop"	"ChlngsNoFlsh"
##	[97]	"ChlngsFlood"	"ChlngsOthr"	"ChlngsFulOvrFlw"
##	[100]	"ChlngsNoWtr"	"ChlngsSmels"	"ChlngsOK"
##	[103]	"SatisColaps"	"SatisSupColaps"	"SatisColapsMore"
##	[106]	"SatisSupColapsMore"		

##	Date	LBO	Prov
##	0	0	0
##	Dist	Comm	Vill
##	0	0	0
##	IDPoor	IDPoorTyp	M01
##	0	0	0
##	M1824	M217	M2559
##	1	0	0
##	M60	F01	F1824
##	0	0	0
##	F217	F2559	F60
##	1	0	0
##	LivRP	VillOD	DateSlabPur
##	0	0	0
##	LatInst	BelowGrndInst	DateBelowGrndInst
##	1457	4012	5320
##	ShltrInst	DateInstComp	RDefBefor
##	4161	1981	1726
##	RDefBeforOthr	FreqNeiToi	WhoInstLat
##	5138	1726	1979
##	KnwSubsdy	RecSubsdy	BorwLat
##	5	0	36
##	CanBuyLat	UseFincOthr	SlabTil
##	4161	4164	1632
##	Npits	PitConfig	NringsDir
##	1631	1630	2002

##	NringsOff	NringsOthr	WallMat
##	1642	5469	1980
##	WallMatOthr	RoofMat	RoofMatOthr
##	5406	1981	5438
##	MatsPurTgthr	Cost	CostInclInst
##	2010	2196	4102
##	CostPitSlab	CostPitSlabKnwn	CostShltrMats
##	1837	3175	1943
##	CostShltrMatsKnwn	LabPitShltrPurTgthr	CostLabLat
##	3167	2979	3019
##	CostLabPitSlab	CostLabPitSlabKnwn	CostLabShltr
##	3005	4574	2641
##	CostLabShltrKnwn	LatTypOwndBefor	IntndChngDich
##	4585	1980	1981
##	IntndChng	IntndChngOthr	AdltUseLat
##	3597	5375	1980
##	ChldUseLat	InfLatDump	IntndPitFull
##	2033	2221	1748
##	IntndPitFullOthrAns	Chlngs	Satis
##	5405	1984	254
##	Rec	SatisSup	RecSup
##	1985	2	10
##	DateSurvCreated	Yr	Mnth
##	7	0	0
##	YrMnth	RDefBefor_BshFld	RDefBefor_RivPnd
##	0	1726	1726
##	RDefBefor_NeiToi	RDefBefor_Othr	RDefBefor_NAAlwysToi
##	1726	1726	1726
##	IntndChng_Shltr	IntndChng_Shwr	IntndChng_Sink
##	3597	3597	3597
##	IntndChng_WtrRes	IntndChng_Pit	IntndChng_Othr
##	3597	3597	3597
##	IntndChng_NAAlwysToi	IntndPitFullDes	IntndPitFullPit
##	3597	1748	1748
##	IntndPitFullEmpSlf	IntndPitFullDK	IntndPitFullOthr
##	1748	1748	1748
##	IntndPitFullPay	IntndPitFullStop	ChlngsNoFlsh
##	1748	1748	1984
##	ChlngsFlood	ChlngsOthr	ChlngsFulOvrFlw
##	1984	1984	1984
##	ChlngsNoWtr	ChlngsSmels	ChlngsOK
##	1984	1984	1984
##	SatisColaps	SatisSupColaps	SatisColapsMore
##	254	2	254
##	SatisSupColapsMore		
##	2		

## Missing Values in Variables

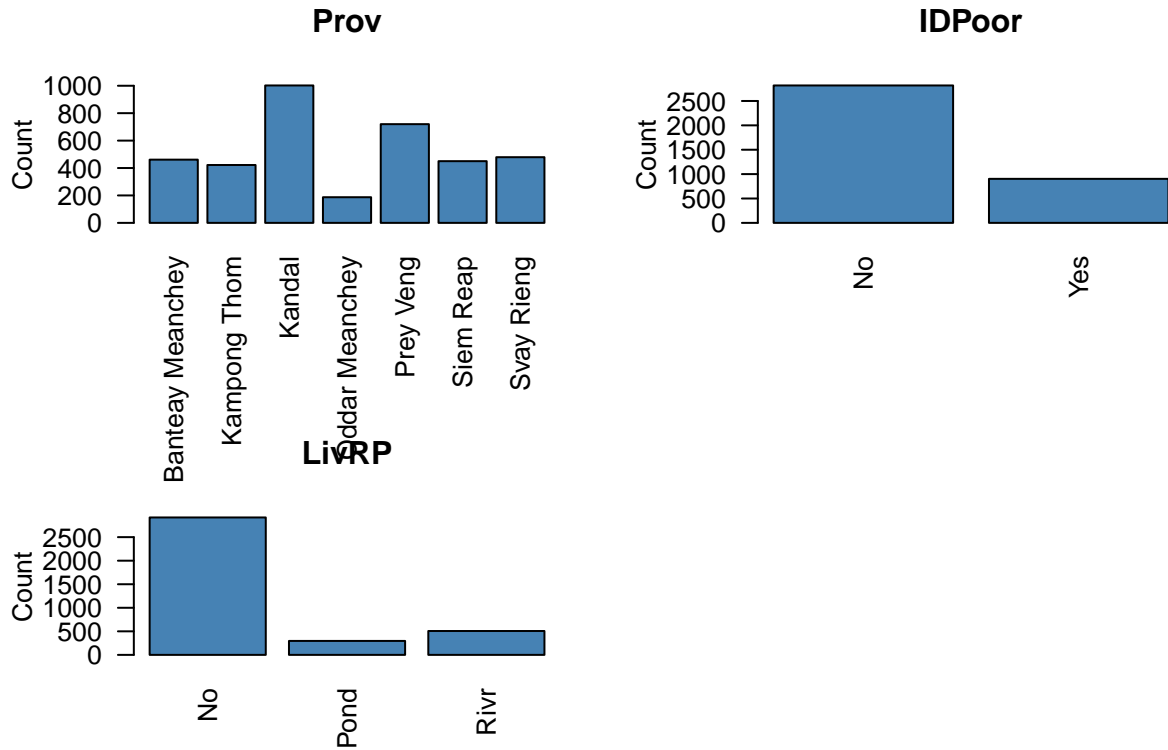


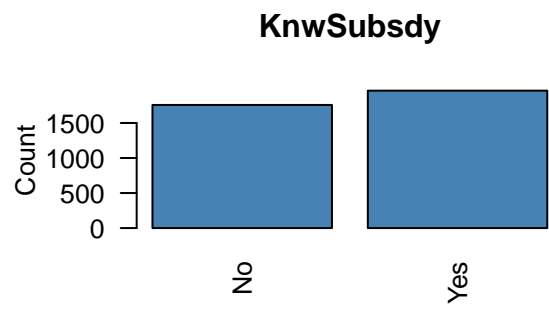
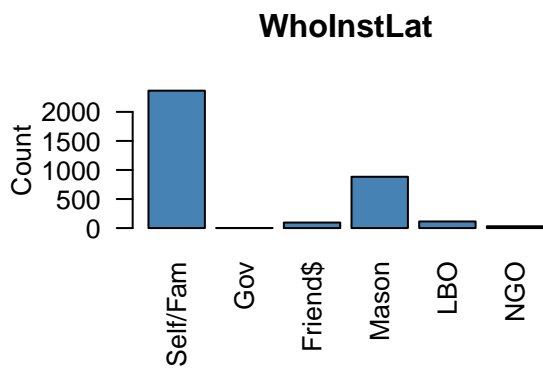
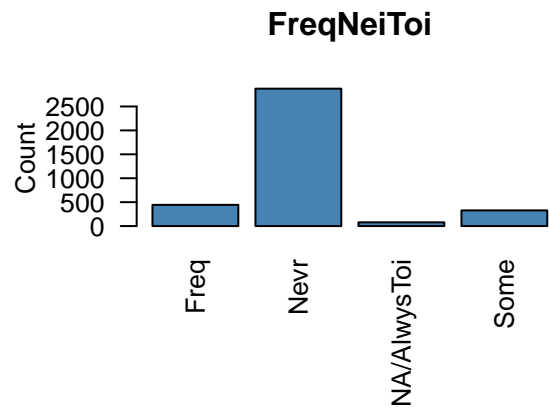
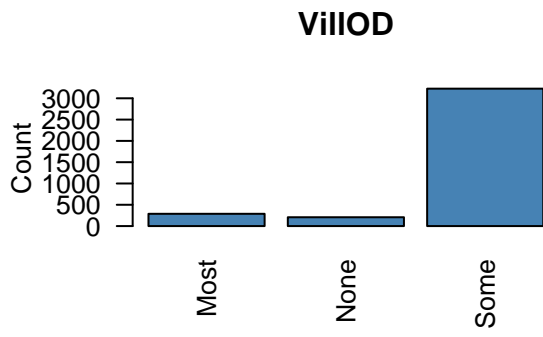
Variable names (rotated) corresponding to the columns of the heatmap:

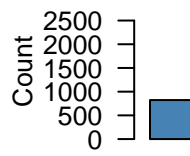
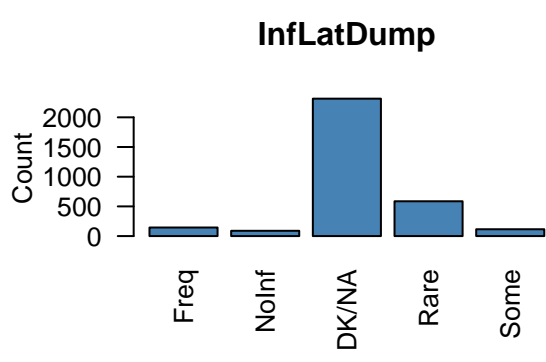
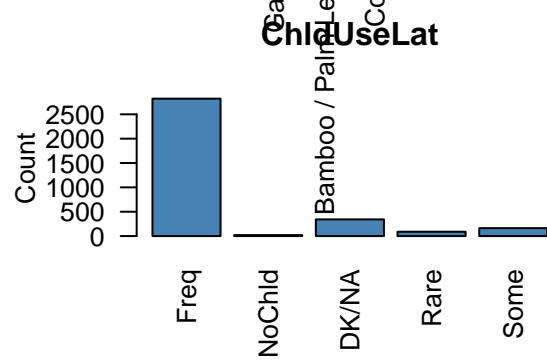
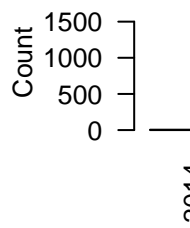
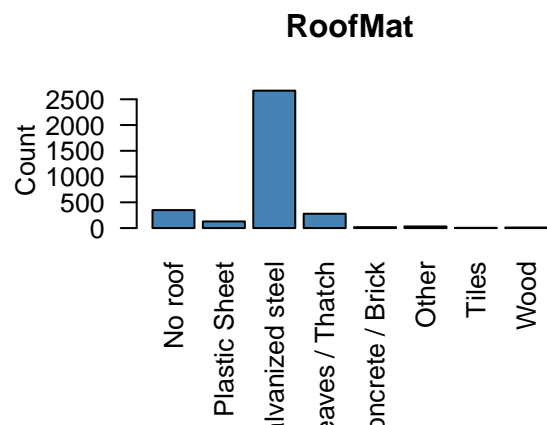
Date, Dist, IDPoor, M1824, M60, F217, LivRP, LatInst, ShltrInst, RDefBeforOthr, KnwSubsdy, CanBuyLat, Npits, NringsOff, WallMatOthr, MatsPurTgthr, CostPitSlab, CostShltrMatsKnwn, CostLabPitSlab, CostLabShltrKnwn, IntndChng, ChldUseLat, IntndPitFullOthrAns, LB0, Comm, IDPoorTyp, M217, F01, F2559, VillOD, BelowGrndInst, DateInstComp, FreqNeiToi, RecSubsdy, UseFincOthr, PitConfig, NringsOthr, RoofMat, Cost, CostPitSlabKnwn, LabPitShltrPurTgthr, CostLabPitSlabKnwn, LatTypOwndBefor, IntndChngOthr, InfLatDump, Chlngs, Prov, Vill, M01, M2559, F1824, F60, DateSlabPur, DateBelowGrndInst, RDefBefor, WhoInstLat, BorwLat, SlabTil, NringsDir, WallMat, RoofMatOthr, CostInclInst, CostShltrMats, CostLabLat, CostLabShltr, IntndChngDich, AdltUseLat, IntndPitFull, Satis

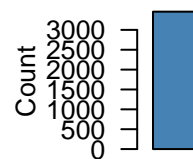
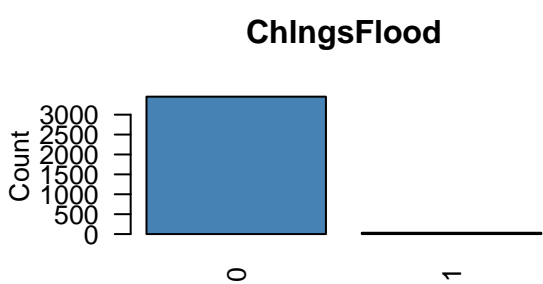
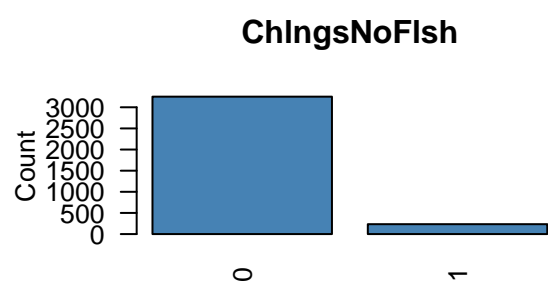
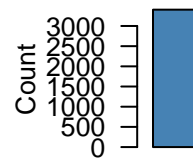
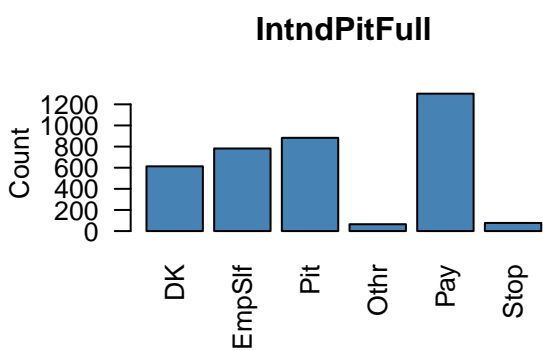
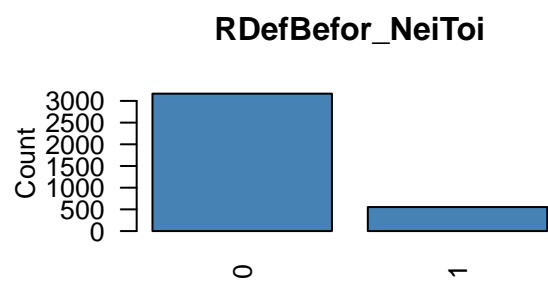
```
## [1] "Date"          "LB0"           "Prov"
## [4] "Dist"          "Comm"          "Vill"
## [7] "IDPoor"        "IDPoorTyp"     "M01"
## [10] "M1824"         "M217"          "M2559"
## [13] "M60"           "F01"           "F1824"
## [16] "F217"          "F2559"         "F60"
## [19] "LivRP"         "VillOD"        "DateSlabPur"
## [22] "LatInst"       "BelowGrndInst" "DateBelowGrndInst"
## [25] "ShltrInst"     "DateInstComp"  "RDefBefor"
## [28] "RDefBeforOthr" "FreqNeiToi"    "WhoInstLat"
## [31] "KnwSubsdy"     "RecSubsdy"     "BorwLat"
## [34] "CanBuyLat"     "UseFincOthr"   "SlabTil"
## [37] "Npits"         "PitConfig"     "NringsDir"
## [40] "NringsOff"     "NringsOthr"    "WallMat"
## [43] "WallMatOthr"   "RoofMat"       "RoofMatOthr"
## [46] "MatsPurTgthr"  "Cost"          "CostInclInst"
## [49] "CostPitSlab"   "CostPitSlabKnwn" "CostShltrMats"
## [52] "CostShltrMatsKnwn" "LabPitShltrPurTgthr" "CostLabLat"
## [55] "CostLabPitSlab" "CostLabPitSlabKnwn" "CostLabShltr"
## [58] "CostLabShltrKnwn" "LatTypOwndBefor" "IntndChngDich"
## [61] "IntndChng"     "IntndChngOthr" "AdltUseLat"
## [64] "ChldUseLat"    "InfLatDump"    "IntndPitFull"
## [67] "IntndPitFullOthrAns" "Chlngs"        "Satis"
```

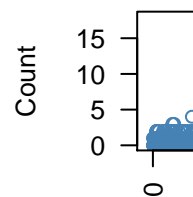
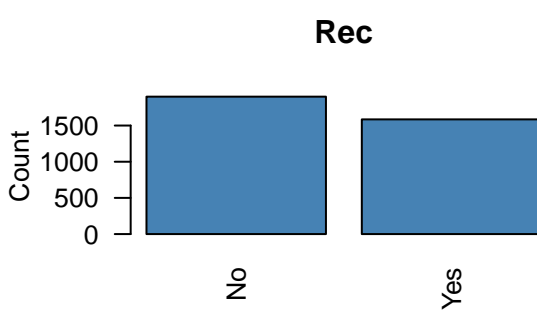
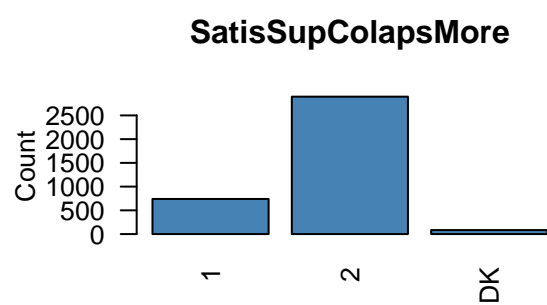
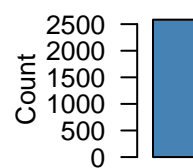
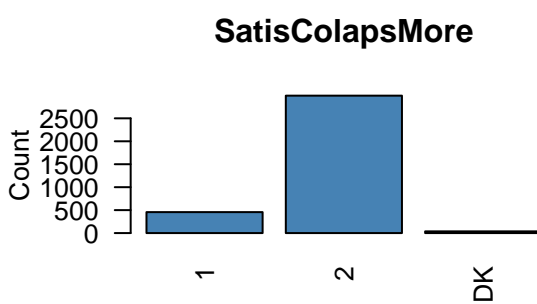
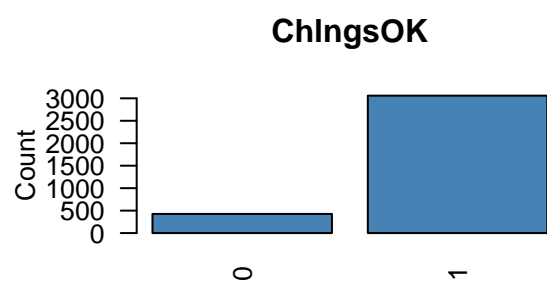
##	[70]	"Rec"	"SatisSup"	"RecSup"
##	[73]	"DateSurvCreated"	"Yr"	"Mnth"
##	[76]	"YrMnth"	"RDefBefor_BshFld"	"RDefBefor_RivPnd"
##	[79]	"RDefBefor_NeiToi"	"RDefBefor_Othr"	"RDefBefor_NAAlwysToi"
##	[82]	"IntndChng_Shlttr"	"IntndChng_Shwr"	"IntndChng_Sink"
##	[85]	"IntndChng_WtrRes"	"IntndChng_Pit"	"IntndChng_Othr"
##	[88]	"IntndChng_NAAlwysToi"	"IntndPitFullDes"	"IntndPitFullPit"
##	[91]	"IntndPitFullEmpSlf"	"IntndPitFullDK"	"IntndPitFullOthr"
##	[94]	"IntndPitFullPay"	"IntndPitFullStop"	"ChlngsNoFlsh"
##	[97]	"ChlngsFlood"	"ChlngsOthr"	"ChlngsFulOvrFlw"
##	[100]	"ChlngsNoWtr"	"ChlngsSmels"	"ChlngsOK"
##	[103]	"SatisColaps"	"SatisSupColaps"	"SatisColapsMore"
##	[106]	"SatisSupColapsMore"		



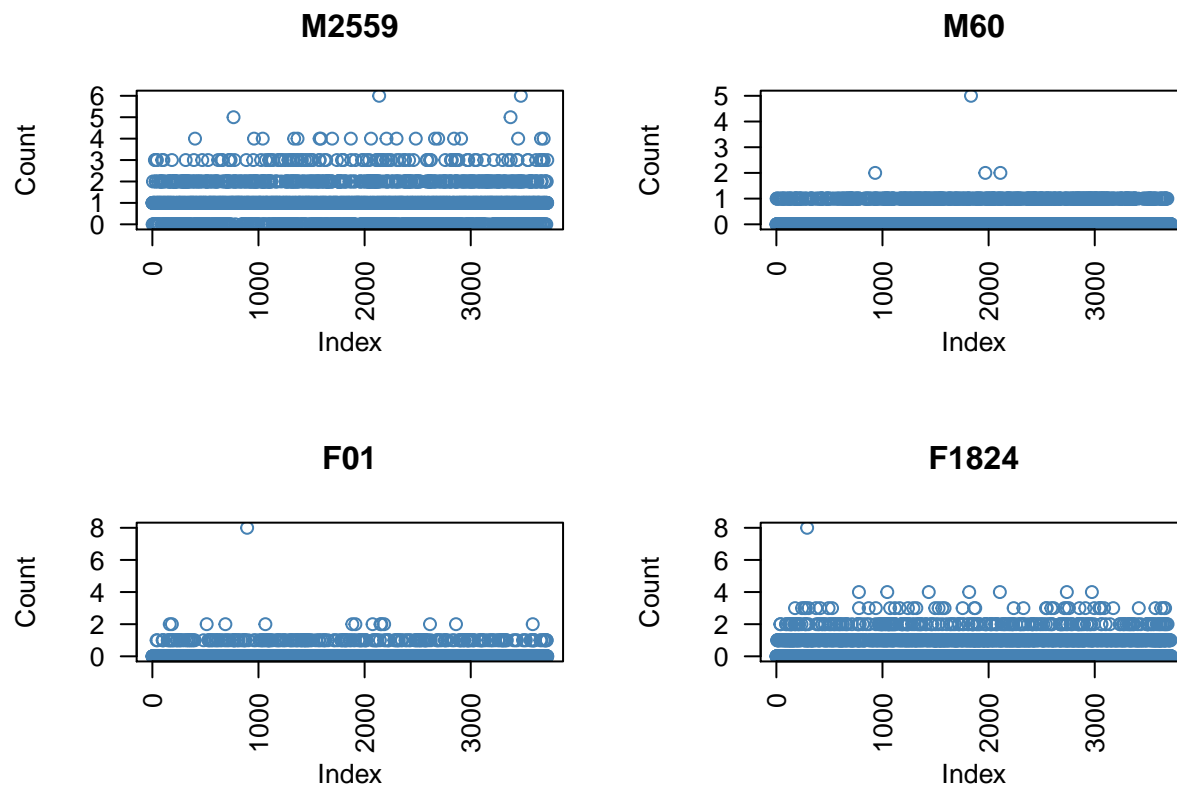












```
##          Prov      IDPoor      LivRP      VillOD
## Banteay Meanchey: 461    No :2818    No  :2918    Most: 288
## Kampong Thom      : 422    Yes: 903    Pond: 296    None: 207
## Kandal            :1002                Rivr: 507    Some:3226
## Oddar Meanchey   : 187
## Prey Veng         : 720
## Siem Reap         : 450
## Svay Rieng        : 479
##
##
##
##
##      FreqNeiT0i      WhoInstLat      KnwSubsdy      RecSubsdy      BorwLat
## Freq      : 443    Self/Fam:2365    No  :1757    DK   :   0    Yes  : 943
## Nev       :2872    Gov      :   2    Yes :1961    No   :3273    DK   :   0
## NA/AlwysT0i: 78    Friend$ :  95    NA's:   3    Yfull:  33    No   :2744
## Some      : 328    Mason   : 883                Ypart: 415    NA's:  34
##
##          LBO      : 114
##          NGO      :  31
##          NA's     : 231
##
##
##
```

```

##
##
##
## SlabTil                               WallMat
## No : 84   Bamboo / Palm Leaves / Thatch: 374
## Yes :3404 Concrete / Brick                :1488
## NA's: 233 Galvanized steel                :1066
##           No walls                        : 15
##           Other                          : 68
##           Plastic Sheet                  : 338
##           Wood                          : 140
##           NA's                          : 232
##
##
##
##
##           RoofMat      AdltUseLat      ChldUseLat
## No roof                : 349   Freq:3397   Freq :2821
## Plastic Sheet          : 129   DK : 1     NoChld: 20
## Galvanized steel       :2667   Rare: 22   DK/NA : 343
## Bamboo / Palm Leaves / Thatch: 278   Some: 68   Rare : 89
## Concrete / Brick       : 18   NA's: 233   Some : 163
## Other                  : 31           NA's : 285
## Tiles                  : 5
## Wood                   : 11
## NA's                   : 233
##
##
##
## InfLatDump      Yr      Mnth      RDefBefor_BshFld RDefBefor_RivPnd
## Freq : 144      2014: 5    01:196   0: 823           0:3678
## NoInf: 89       2015: 620  02:198   1:2898           1: 43
## DK/NA:2314      2016:1195  03:203
## Rare : 586      2017:1901  04: 77
## Some : 115              05:267
## NA's : 473              06:242
##                          07:176
##                          08:529
##                          09:361
##                          10:411
##                          11:598
##                          12:463
## RDefBefor_NeiToi IntndPitFull ChlngsNoFlsh ChlngsFlood ChlngsOthr
## 0:3167            DK : 613   0 :3251   0 :3452   0 :3404
## 1: 554            EmpSlf: 782 1 : 234   1 : 33    1 : 81
##                  Pit : 883   NA's: 236   NA's: 236   NA's: 236
##                  Othr : 65
##                  Pay :1301

```



```

## 1st Qu.:0.0000 1st Qu.:0.0000 1st Qu.:0.0000 1st Qu.:1.000
## Median :0.0000 Median :0.0000 Median :1.0000 Median :1.000
## Mean :0.0774 Mean :0.4058 Mean :0.8282 Mean :1.121
## 3rd Qu.:0.0000 3rd Qu.:1.0000 3rd Qu.:1.0000 3rd Qu.:1.000
## Max. :8.0000 Max. :8.0000 Max. :9.0000 Max. :6.000
## NA's :1
##
##
##
##
##
## F60
## Min. :0.0000
## 1st Qu.:0.0000
## Median :0.0000
## Mean :0.2435
## 3rd Qu.:0.0000
## Max. :2.0000
##
##
##
##
##
##
##
##
##
##
## Prov IDPoor LivWtr VillOD
## BM:440 NotIDPoor:2652 LivWtr : 757 MostVillOD: 220
## KT:385 IDPoor : 785 NoLivWtr:2680 NoVillOD : 197
## K :966 SomeVillOD:3020
## OM:152
## PV:654
## SR:400
## SG:440
##
##
##
##
## FreqNeiTol WhoInstLat KnwSubsdy RecSubsdy
## FreqNeiTol: 413 Slf/FamInstLat:2349 KnwSubsdy:1637 NoSubsdy :3121
## NoNeiTol :2648 FrndInstLat : 95 DKSubsdy :1800 PrtSubsdy: 316
## AlwysTol : 76 MasonInstLat : 870
## SomeNeiTol: 300 LBOInstLat : 110
## NGOInstLat : 13
##
##
##
##
##

```

```

##
##
##
##      BorwLat      SlabTil      WallMat      RoofMat
##      Borw  : 858    NoSlabTil: 81    WoodWal  : 508    NoRoof   : 346
##      NoBorw:2579    SlabTil   :3356    MasnWal  :1473    PlstcRoof: 129
##                                     StelWal  :1038    StelRoof  :2626
##                                     NoWal    : 15    WoodRoof  : 284
##                                     OthrWal  : 68    MasnRoof  : 16
##                                     PlstcWal: 335    OthrRoof  : 36
##
##
##
##
##
##
##      AdltUseLat      ChldUseLat      InfLatDump      Yr
##      AdltFreqLat:3347    ChldFreqLat:2777    InfFreqLat: 141    2015: 367
##      AdltRarLat : 22    ChldDKLat  : 413    InfDKLat   :2608    2016:1183
##      AdltSomLat : 68    ChldRarLat : 87    InfRarLat  : 574    2017:1887
##                                     ChldSomLat : 160    InfSomLat  : 114
##
##
##
##
##
##
##
##
##
##      Mnth      RDefBefor_NeiToi    IntndPitFull    ChlngsNoFlsh
##      01:192    NoDefBeforNeiToi:2926    DK      : 571    FlshOK:3204
##      02:198    DefBeforNeiToi  : 511    EmpSlf: 702    NoFlsh: 233
##      03:203                                     Pit      : 844
##      04: 77                                     Othr     : 64
##      05:264                                     Pay      :1190
##      06:240                                     Stop     : 66
##      07:174
##      08:526
##      09:357
##      10:373
##      11:482
##      12:351
##      ChlngsFlood      ChlngsOthr      ChlngsFulOvrFlw    ChlngsNoWtr
##      NoFlood:3406    NoOthrChlngs:3358    NoFulOvrflw:3416    WtrOK:3411
##      Flood  : 31    OthrChlngs  : 79    FulOvrflw  : 21    NoWtr: 26
##
##
##

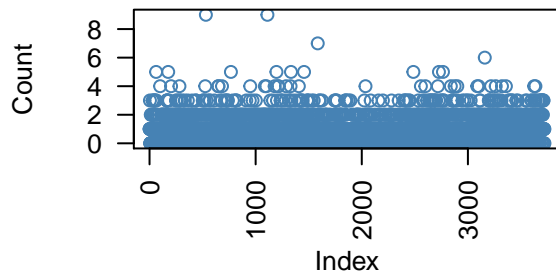
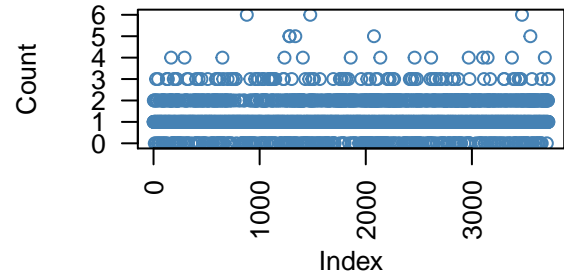
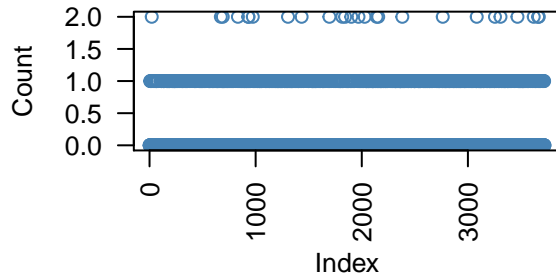
```



##	1st Qu.:	0.0000	1st Qu.:	0.0000	1st Qu.:	1.000	1st Qu.:	0.0000
##	Median	:0.0000	Median	:1.0000	Median	:1.000	Median	:0.0000
##	Mean	:0.4038	Mean	:0.8303	Mean	:1.128	Mean	:0.2459
##	3rd Qu.:	1.0000	3rd Qu.:	1.0000	3rd Qu.:	1.000	3rd Qu.:	0.0000
##	Max.	:8.0000	Max.	:9.0000	Max.	:6.000	Max.	:2.0000

```
##          ODBefor
## NoODBefor: 754
## ODBefor   :2683
```

##	Prov	IDPoor	LivWtr	Vill10D
##	0	0	0	0
##	FreqNeiToi	WhoInstLat	KnwSubsdy	RecSubsdy
##	0	0	0	0
##	BorwLat	SlabTil	WallMat	RoofMat
##	0	0	0	0
##	AdltUseLat	ChldUseLat	InfLatDump	Yr
##	0	0	0	0
##	Mnth RDefBefor_NeiToi	IntndPitFull	ChlngsNoFlsh	
##	0	0	0	0
##	ChlngsFlood	ChlngsOthr	ChlngsFulOvrFlw	ChlngsNoWtr
##	0	0	0	0
##	ChlngsSmels	ChlngsOK	Satis	SatisSup
##	0	0	0	0
##	Rec	RecSup	M01	M1824
##	0	0	0	0
##	M217	M2559	M60	F01
##	0	0	0	0
##	F1824	F217	F2559	F60
##	0	0	0	0
##	ODBefor			
##	0			

**F217****F2559****F60**

##		name	idx	nobs	type	exo	user	mean	var	nlev
## 1		Prov	1	3437	factor	0	0	NA	NA	7
## 2		IDPoor	2	3437	factor	0	0	NA	NA	2
## 3		LivWtr	3	3437	factor	0	0	NA	NA	2
## 4		VillOD	4	3437	factor	0	0	NA	NA	3
## 5		FreqNeiToi	5	3437	factor	0	0	NA	NA	4
## 6		WhoInstLat	6	3437	factor	0	0	NA	NA	5
## 7		KnwSubsdy	7	3437	factor	0	0	NA	NA	2
## 8		RecSubsdy	8	3437	factor	0	0	NA	NA	2
## 9		BorwLat	9	3437	factor	0	0	NA	NA	2
## 10		SlabTil	10	3437	factor	0	0	NA	NA	2
## 11		WallMat	11	3437	factor	0	0	NA	NA	6
## 12		RoofMat	12	3437	factor	0	0	NA	NA	6
## 13		AdltUseLat	13	3437	factor	0	0	NA	NA	3
## 14		ChldUseLat	14	3437	factor	0	0	NA	NA	4
## 15		InfLatDump	15	3437	factor	0	0	NA	NA	4
## 16		Yr	16	3437	factor	0	0	NA	NA	3
## 17		Mnth	17	3437	factor	0	0	NA	NA	12
## 18	RDefBefor_NeiToi		18	3437	factor	0	0	NA	NA	2
## 19		IntndPitFull	19	3437	factor	0	0	NA	NA	6
## 20		ChlngsNoFlsh	20	3437	factor	0	0	NA	NA	2
## 21		ChlngsFlood	21	3437	factor	0	0	NA	NA	2
## 22		ChlngsOthr	22	3437	factor	0	0	NA	NA	2



## 23	ChlngsFulOvrFlw	23	3437	factor	0	0	NA	NA	2
## 24	ChlngsNoWtr	24	3437	factor	0	0	NA	NA	2
## 25	ChlngsSmels	25	3437	factor	0	0	NA	NA	2
## 26	ChlngsOK	26	3437	factor	0	0	NA	NA	2
## 27	Satis	27	3437	factor	0	0	NA	NA	3
## 28	SatisSup	28	3437	factor	0	0	NA	NA	3
## 29	Rec	29	3437	factor	0	0	NA	NA	2
## 30	RecSup	30	3437	factor	0	0	NA	NA	2
## 31	M01	31	3437	numeric	0	0	0.091	0.089	0
## 32	M1824	32	3437	numeric	0	0	0.445	0.653	0
## 33	M217	33	3437	numeric	0	0	0.812	0.875	0
## 34	M2559	34	3437	numeric	0	0	1.033	0.458	0
## 35	M60	35	3437	numeric	0	0	0.166	0.146	0
## 36	F01	36	3437	numeric	0	0	0.077	0.095	0
## 37	F1824	37	3437	numeric	0	0	0.404	0.475	0
## 38	F217	38	3437	numeric	0	0	0.830	0.900	0
## 39	F2559	39	3437	numeric	0	0	1.128	0.386	0
## 40	F60	40	3437	numeric	0	0	0.246	0.199	0
## 41	ODBefor	41	3437	factor	0	0	NA	NA	2
##									lnam
## 1							BM KT K OM PV SR SG		
## 2							NotIDPoor IDPoor		
## 3							LivWtr NoLivWtr		
## 4							MostVillOD NoVillOD SomeVillOD		
## 5							FreqNeiToi NoNeiToi AlwysToi SomeNeiToi		
## 6	Slf/FamInstLat FrndInstLat MasonInstLat LB0InstLat NG0InstLat								
## 7							KnwSubsdy DKSubsdy		
## 8							NoSubsdy PrtSubsdy		
## 9							Borw NoBorw		
## 10							NoSlabTil SlabTil		
## 11							WoodWal MasnWal StelWal NoWal OthrWal PlstcWal		
## 12	NoRoof PlstcRoof StelRoof WoodRoof MasnRoof OthrRoof								
## 13							AdltFreqLat AdltRarLat AdltSomLat		
## 14							ChldFreqLat ChldDKLat ChldRarLat ChldSomLat		
## 15							InfFreqLat InfDKLat InfRarLat InfSomLat		
## 16							2015 2016 2017		
## 17							01 02 03 04 05 06 07 08 09 10 11 12		
## 18							NoDefBeforNeiToi DefBeforNeiToi		
## 19							DK EmpSlf Pit Othr Pay Stop		
## 20							FlshOK NoFlsh		
## 21							NoFlood Flood		
## 22							NoOthrChlngs OthrChlngs		
## 23							NoFulOvrflw FulOvrflw		
## 24							WtrOK NoWtr		
## 25							NoSmel Smel		
## 26							Chlngs NoChlngs		
## 27							UnsatLat SatLat DKSatLat		
## 28							UnsatSup SatSup DKSatSup		

```
## 29                                     NoRecLat|RecLat
## 30                                     NoRecSup|RecSup
## 31
## 32
## 33
## 34
## 35
## 36
## 37
## 38
## 39
## 40
## 41                                     NoODBefor|ODBefor
```

```
par(mfrow = c(1, 1))
print(res.mca)
```

```
## **Results of the Multiple Correspondence Analysis (MCA)**
## The analysis was performed on 3437 individuals, described by 41 variables
## *The results are available in the following objects:
##
##      name
## 1  "$eig"
## 2  "$var"
## 3  "$var$coord"
## 4  "$var$cos2"
## 5  "$var$contrib"
## 6  "$var$v.test"
## 7  "$ind"
## 8  "$ind$coord"
## 9  "$ind$cos2"
## 10 "$ind$contrib"
## 11 "$quanti.sup"
## 12 "$quanti.sup$coord"
## 13 "$quali.sup"
## 14 "$quali.sup$coord"
## 15 "$quali.sup$cos2"
## 16 "$quali.sup$v.test"
## 17 "$call"
## 18 "$call$marge.col"
## 19 "$call$marge.li"
##      description
## 1  "eigenvalues"
## 2  "results for the variables"
## 3  "coord. of the categories"
## 4  "cos2 for the categories"
## 5  "contributions of the categories"
## 6  "v-test for the categories"
```

```

## 7 "results for the individuals"
## 8 "coord. for the individuals"
## 9 "cos2 for the individuals"
## 10 "contributions of the individuals"
## 11 "results for the supplementary quantitative variables"
## 12 "coord. of the supplementary quantitative variables"
## 13 "results for the supplementary categorical variables"
## 14 "coord. for the supplementary categories"
## 15 "cos2 for the supplementary categories"
## 16 "v-test for the supplementary categories"
## 17 "intermediate results"
## 18 "weights of columns"
## 19 "weights of rows"

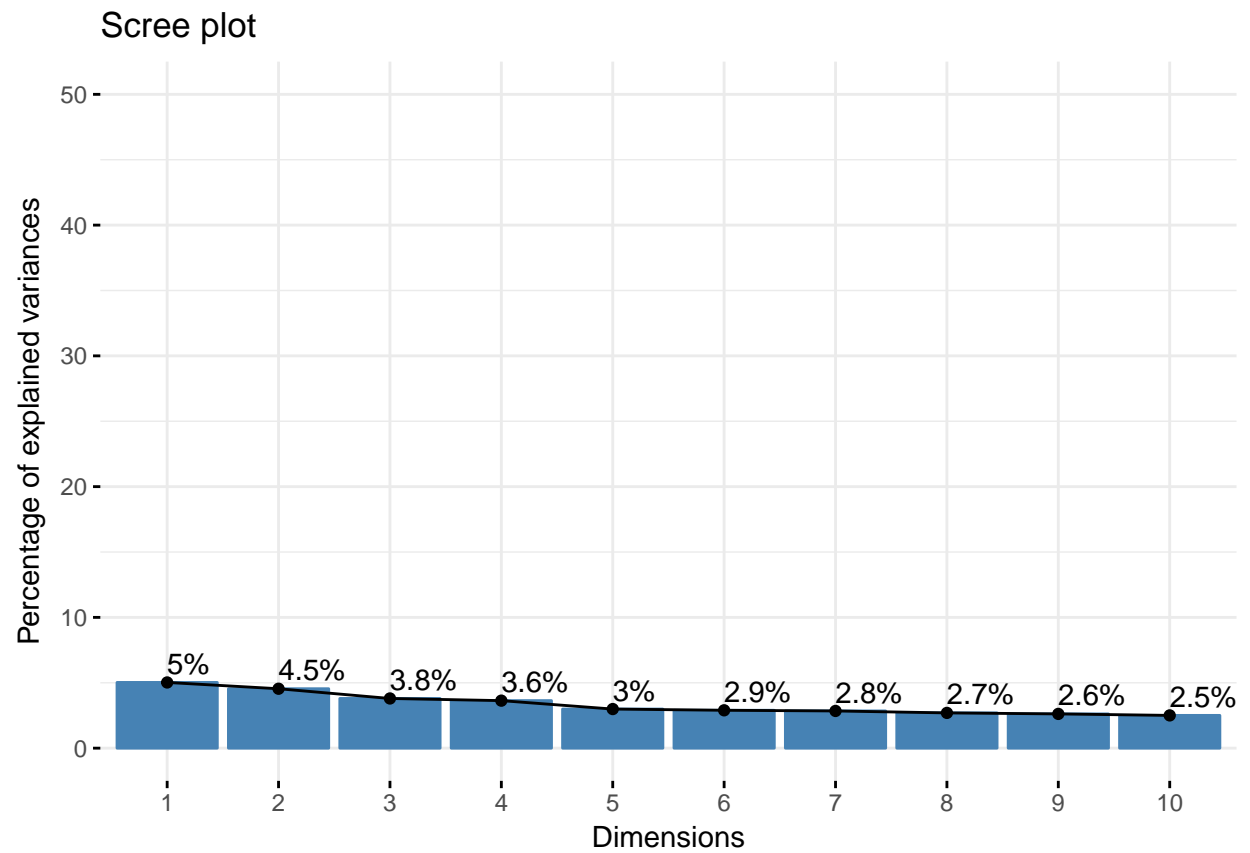
```

```
get_eigenvalue(res.mca)
```

##	eigenvalue	variance.percent	cumulative.variance.percent
## Dim.1	0.095099282	5.02411301	5.024113
## Dim.2	0.086015833	4.54423269	9.568346
## Dim.3	0.071923454	3.79972966	13.368075
## Dim.4	0.068784294	3.63388723	17.001963
## Dim.5	0.056595599	2.98995619	19.991919
## Dim.6	0.054810324	2.89563974	22.887559
## Dim.7	0.053854318	2.84513376	25.732692
## Dim.8	0.051034490	2.69616176	28.428854
## Dim.9	0.049579420	2.61929010	31.048144
## Dim.10	0.047411947	2.50478213	33.552926
## Dim.11	0.046394441	2.45102707	36.003953
## Dim.12	0.044354061	2.34323343	38.347187
## Dim.13	0.043292758	2.28716459	40.634351
## Dim.14	0.042900221	2.26642675	42.900778
## Dim.15	0.040519824	2.14066995	45.041448
## Dim.16	0.040134419	2.12030893	47.161757
## Dim.17	0.039361587	2.07948004	49.241237
## Dim.18	0.038428707	2.03019582	51.271433
## Dim.19	0.038214325	2.01887000	53.290303
## Dim.20	0.037544405	1.98347802	55.273781
## Dim.21	0.037097664	1.95987660	57.233657
## Dim.22	0.036602357	1.93370942	59.167367
## Dim.23	0.036454241	1.92588443	61.093251
## Dim.24	0.035490894	1.87499064	62.968242
## Dim.25	0.035380699	1.86916901	64.837411
## Dim.26	0.034627493	1.82937701	66.666788
## Dim.27	0.033470631	1.76825975	68.435048
## Dim.28	0.033313788	1.75997370	70.195021
## Dim.29	0.032910341	1.73865951	71.933681
## Dim.30	0.032513003	1.71766807	73.651349
## Dim.31	0.032262282	1.70442247	75.355771

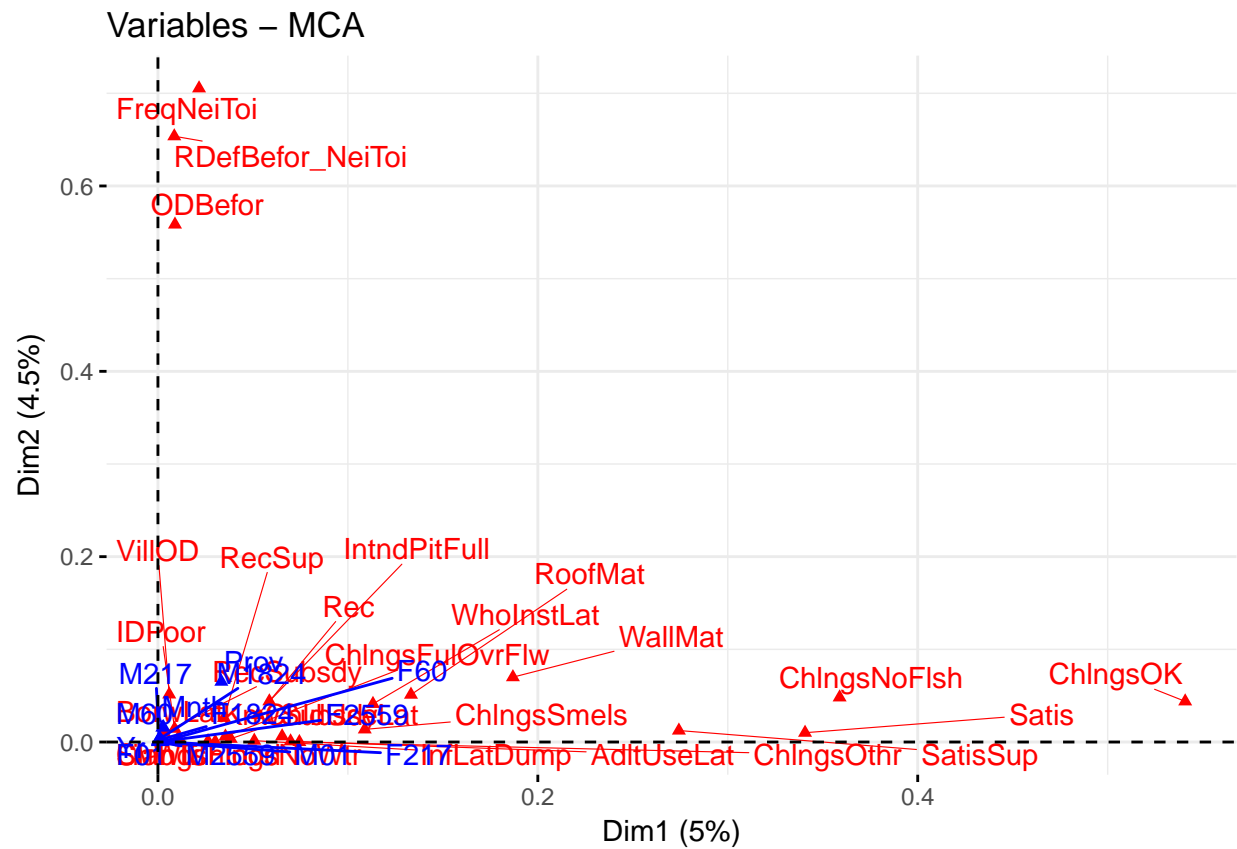
## Dim.32	0.031466206	1.66236558	77.018137
## Dim.33	0.031181360	1.64731715	78.665454
## Dim.34	0.030784046	1.62632696	80.291781
## Dim.35	0.030587267	1.61593108	81.907712
## Dim.36	0.029235096	1.54449563	83.452208
## Dim.37	0.028623443	1.51218189	84.964390
## Dim.38	0.027770623	1.46712726	86.431517
## Dim.39	0.026321098	1.39054858	87.822066
## Dim.40	0.024757767	1.30795752	89.130023
## Dim.41	0.023983643	1.26706040	90.397084
## Dim.42	0.023087391	1.21971124	91.616795
## Dim.43	0.021674107	1.14504717	92.761842
## Dim.44	0.020782383	1.09793724	93.859779
## Dim.45	0.018985773	1.00302197	94.862801
## Dim.46	0.017933536	0.94743211	95.810233
## Dim.47	0.016481836	0.87073851	96.680972
## Dim.48	0.015427169	0.81502025	97.495992
## Dim.49	0.014014066	0.74036574	98.236358
## Dim.50	0.012840028	0.67834110	98.914699
## Dim.51	0.012093731	0.63891408	99.553613
## Dim.52	0.006771343	0.35773131	99.911344
## Dim.53	0.001678127	0.08865577	100.000000

```
fviz_screplot(res.mca, addlabels = TRUE, ylim = c(0, 50))
```



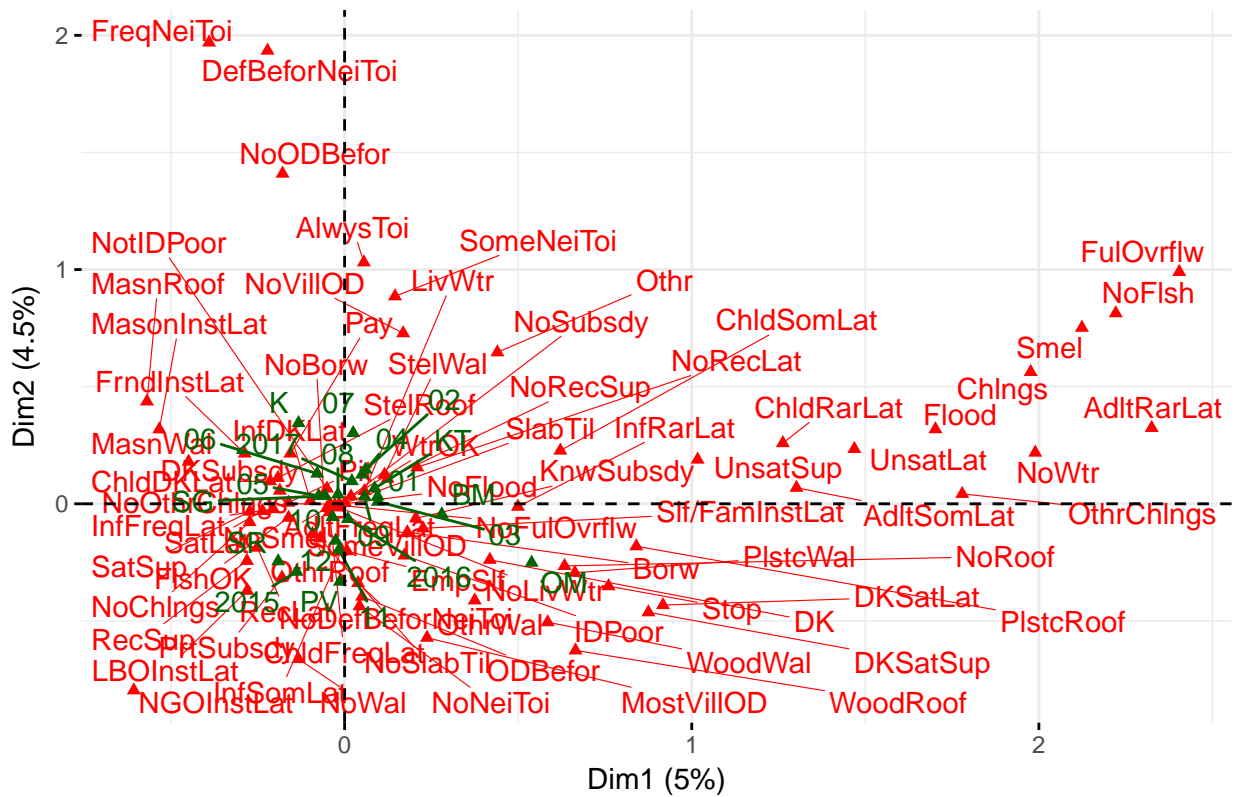
```
fviz_mca_biplot(res.mca, repel = FALSE, ggtheme = theme_minimal())
```





```
fviz_mca_var(res.mca, repel = TRUE, ggtheme = theme_minimal())
```

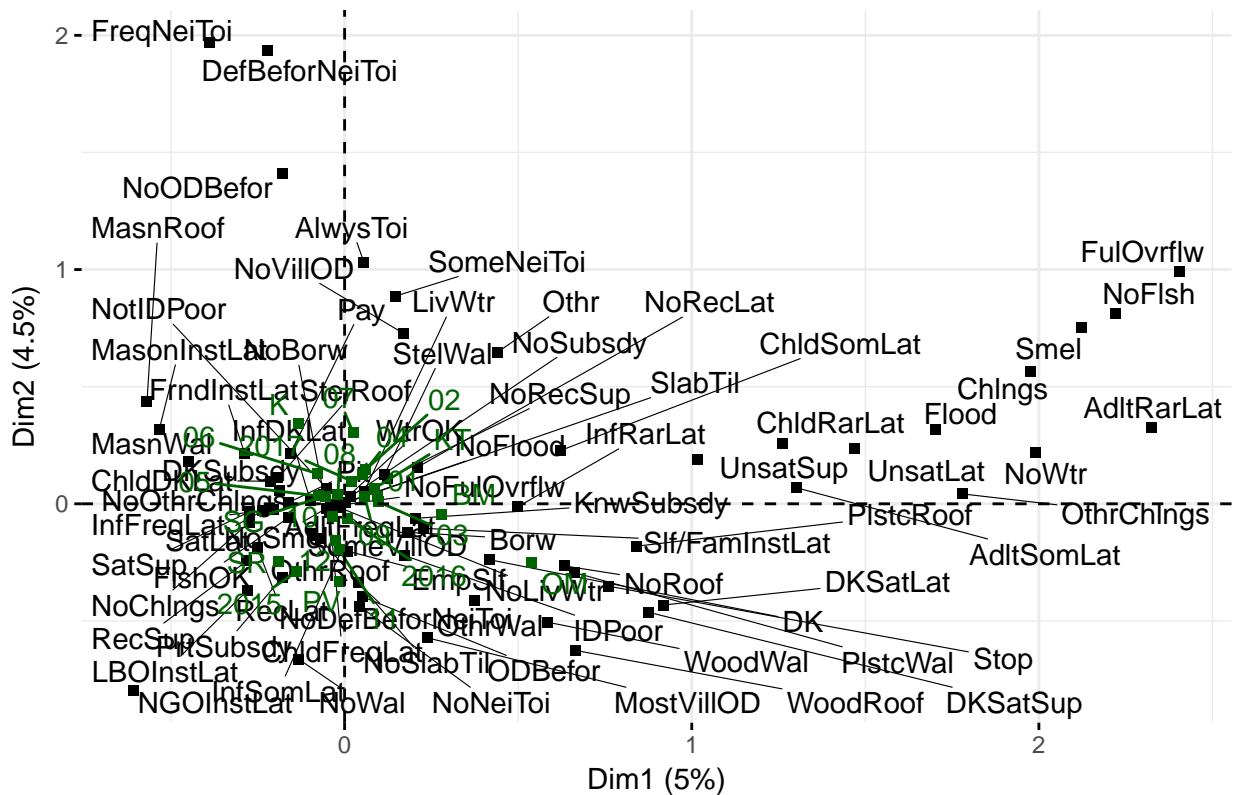
## Variable categories – MCA



```
fviz_mca_var(res.mca, col.var = "black", shape.var = 15, repel = TRUE)
```



## Variable categories – MCA



```
head(round(res.mca.vars$coord, 2))
```

```
##          Dim 1 Dim 2 Dim 3 Dim 4 Dim 5
## NotIDPoor  -0.05  0.07 -0.02 -0.15  0.06
## IDPoor     0.17 -0.22  0.05  0.52 -0.22
## LivWtr     0.06  0.05 -0.15  0.15 -0.04
## NoLivWtr   -0.02 -0.02  0.04 -0.04  0.01
## MostVillOD 0.24 -0.57  0.28 -0.29  0.14
## NoVillOD   0.17  0.73  0.27  0.66 -0.08
```

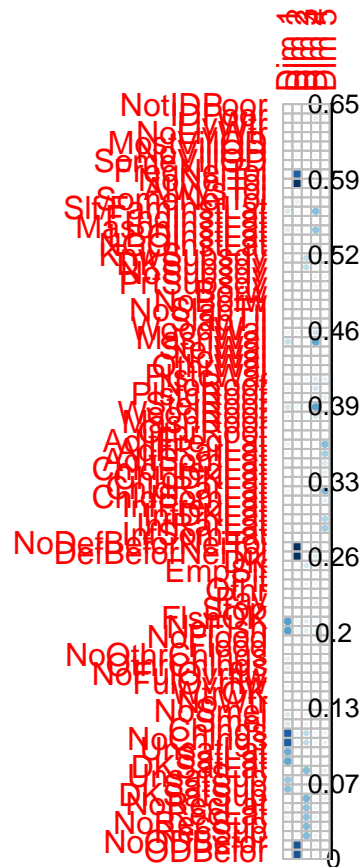
```
head(round(res.mca.vars$coord, 2))
```

```
##          Dim 1 Dim 2 Dim 3 Dim 4 Dim 5
## NotIDPoor  -0.05  0.07 -0.02 -0.15  0.06
## IDPoor     0.17 -0.22  0.05  0.52 -0.22
## LivWtr     0.06  0.05 -0.15  0.15 -0.04
## NoLivWtr   -0.02 -0.02  0.04 -0.04  0.01
## MostVillOD 0.24 -0.57  0.28 -0.29  0.14
## NoVillOD   0.17  0.73  0.27  0.66 -0.08
```

```
head(round(res.mca.vars$cos2, 4))
```

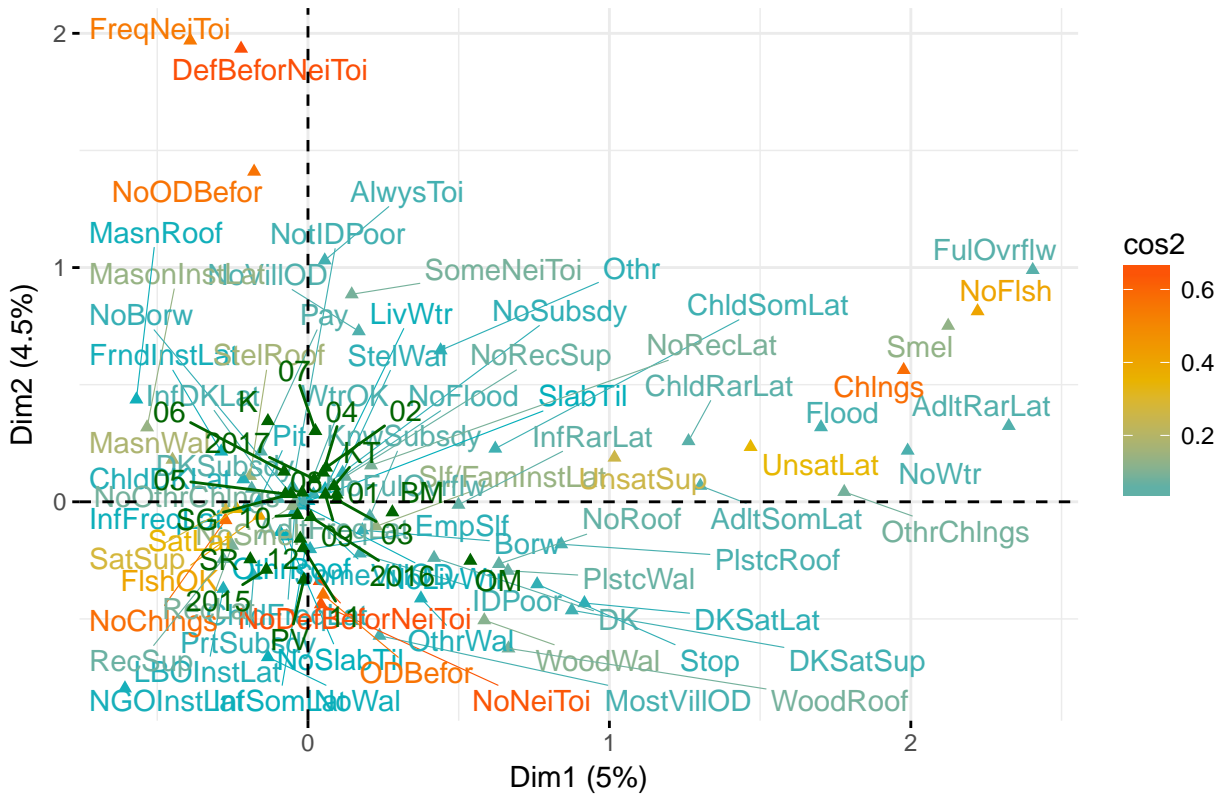
```
##          Dim 1  Dim 2  Dim 3  Dim 4  Dim 5
## NotIDPoor  0.0088 0.0144 0.0009 0.0812 0.0138
## IDPoor     0.0088 0.0144 0.0009 0.0812 0.0138
## LivWtr     0.0009 0.0008 0.0063 0.0061 0.0005
## NoLivWtr   0.0009 0.0008 0.0063 0.0061 0.0005
## MostVillOD 0.0039 0.0224 0.0055 0.0056 0.0013
## NoVillOD   0.0017 0.0323 0.0045 0.0261 0.0004
```

```
corrplot(res.mca.vars$cos2, is.corr = FALSE)
```



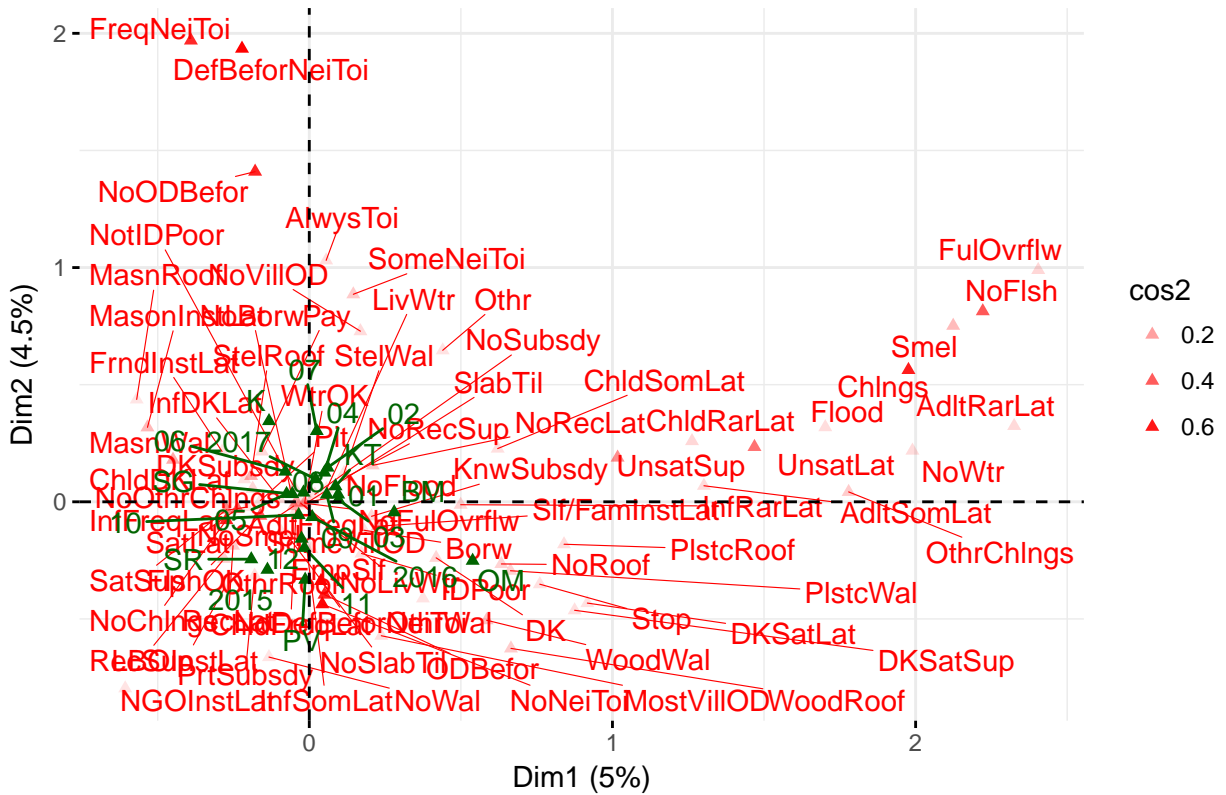
```
fviz_mca_var(res.mca, col.var = "cos2",
  gradient.cols = c("#00AFBB", "#E7B800", "#FC4E07"),
  repel = TRUE,
  ggtheme = theme_minimal())
```

## Variable categories – MCA



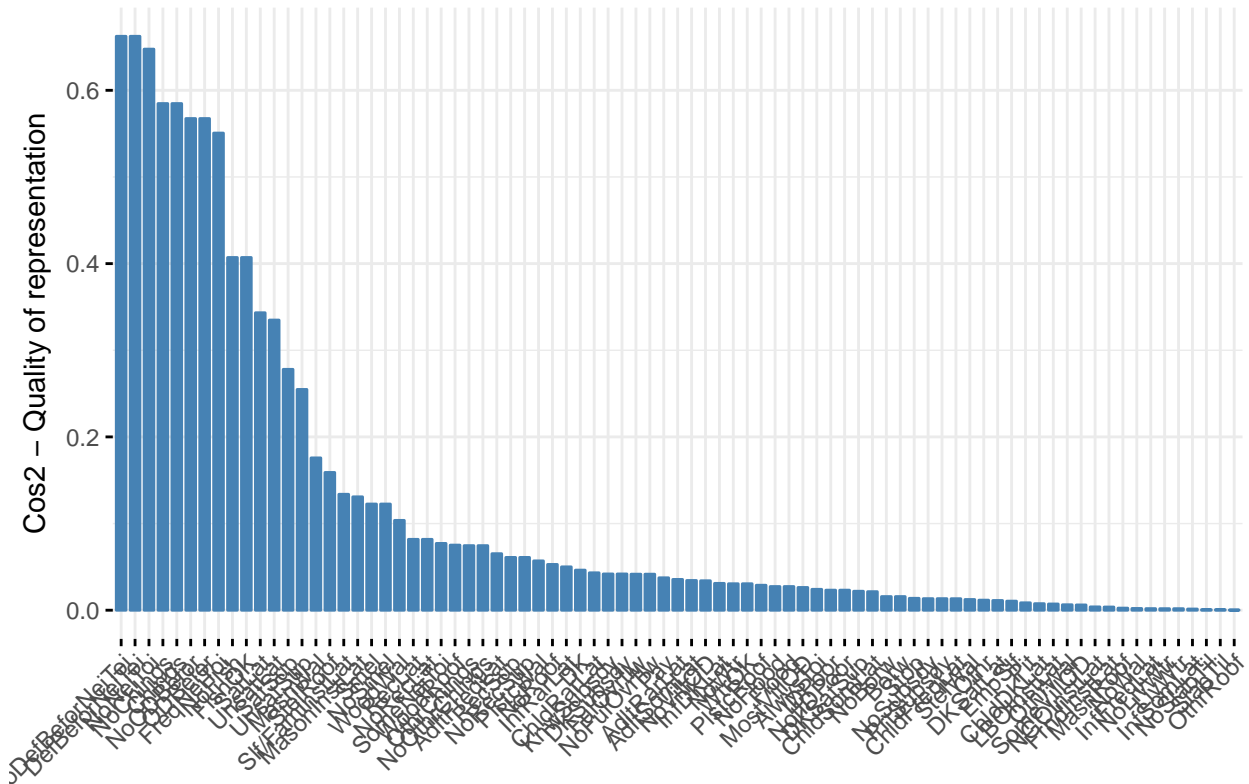
```
fviz_mca_var(res.mca, alpha.var = "cos2",
              repel = TRUE,
              ggtheme = theme_minimal())
```

## Variable categories – MCA



```
fviz_cos2(res.mca, choice = "var", axes = 1:2)
```

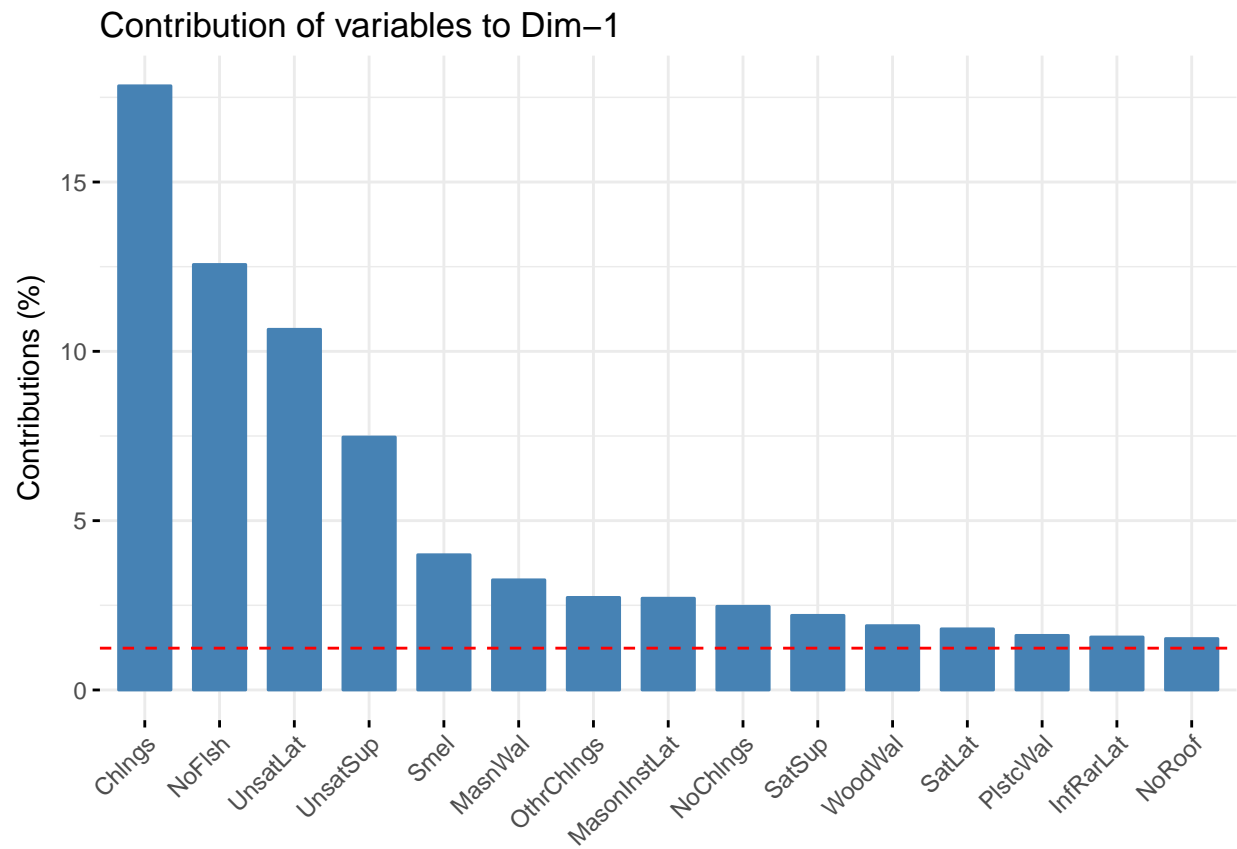
## Cos2 of variables to Dim-1-2



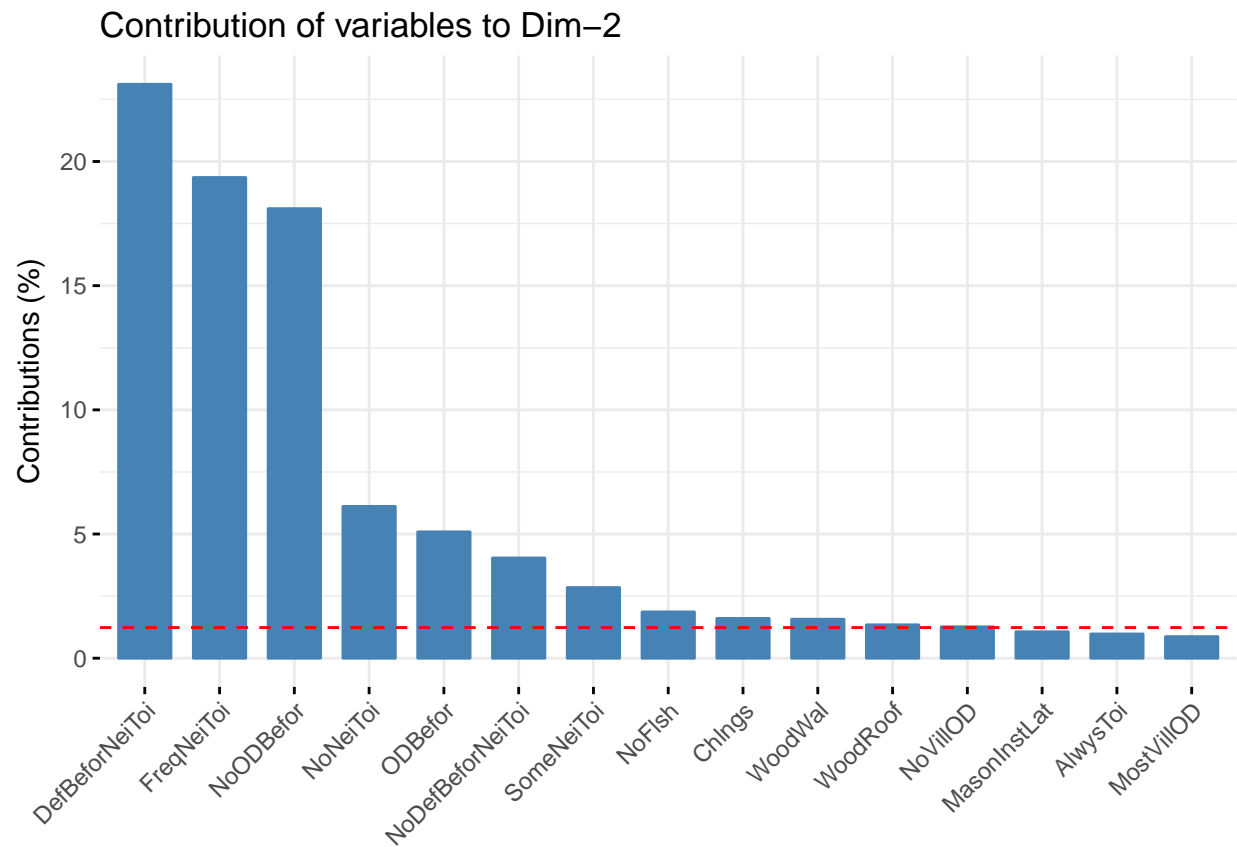
```
head(round(res.mca.vars$contrib, 3))
```

##	Dim 1	Dim 2	Dim 3	Dim 4	Dim 5
## NotIDPoor	0.075	0.137	0.010	0.962	0.198
## IDPoor	0.254	0.461	0.033	3.251	0.671
## LivWtr	0.027	0.027	0.244	0.248	0.023
## NoLivWtr	0.008	0.008	0.069	0.070	0.007
## MostVillOD	0.135	0.869	0.254	0.272	0.074
## NoVillOD	0.061	1.263	0.208	1.279	0.024

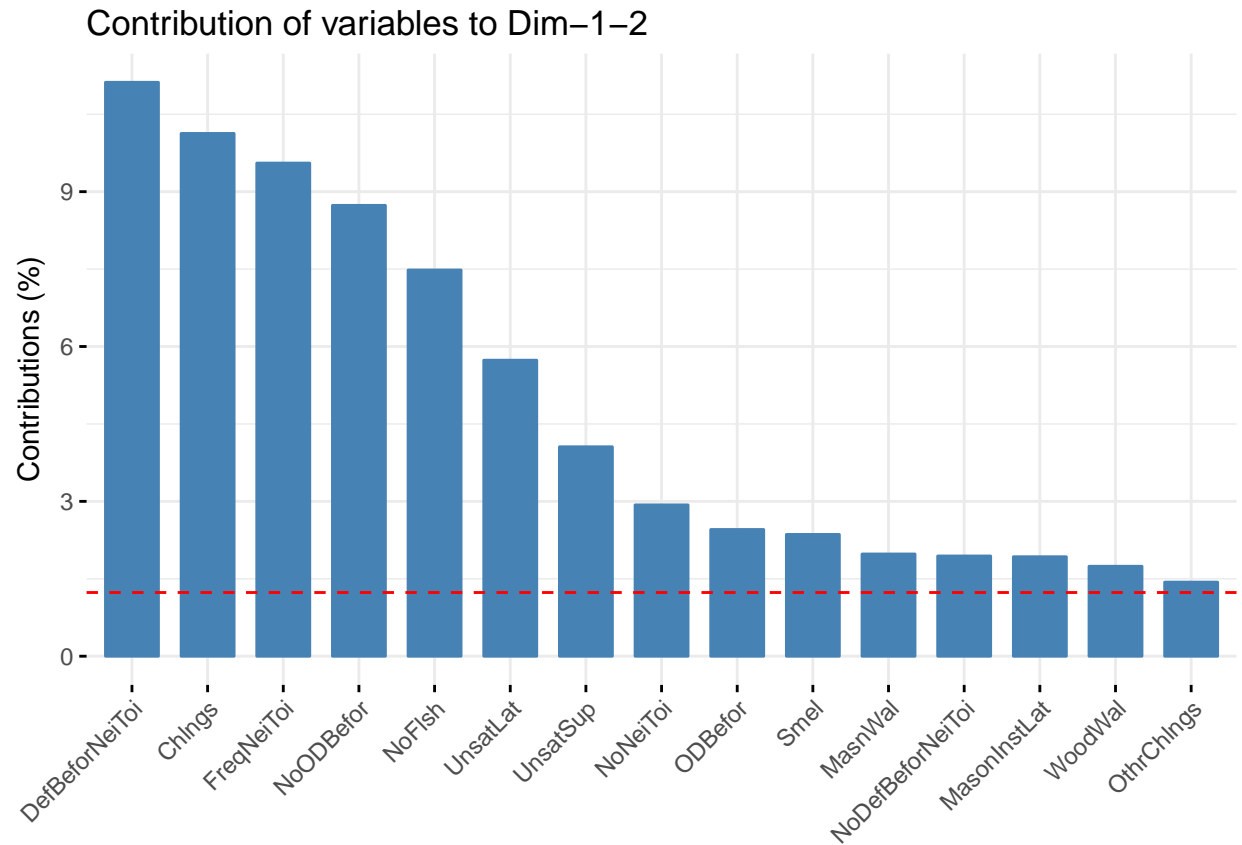
```
fviz_contrib(res.mca, choice = "var", axes = 1, top = 15)
```



```
fviz_contrib(res.mca, choice = "var", axes = 2, top = 15)
```



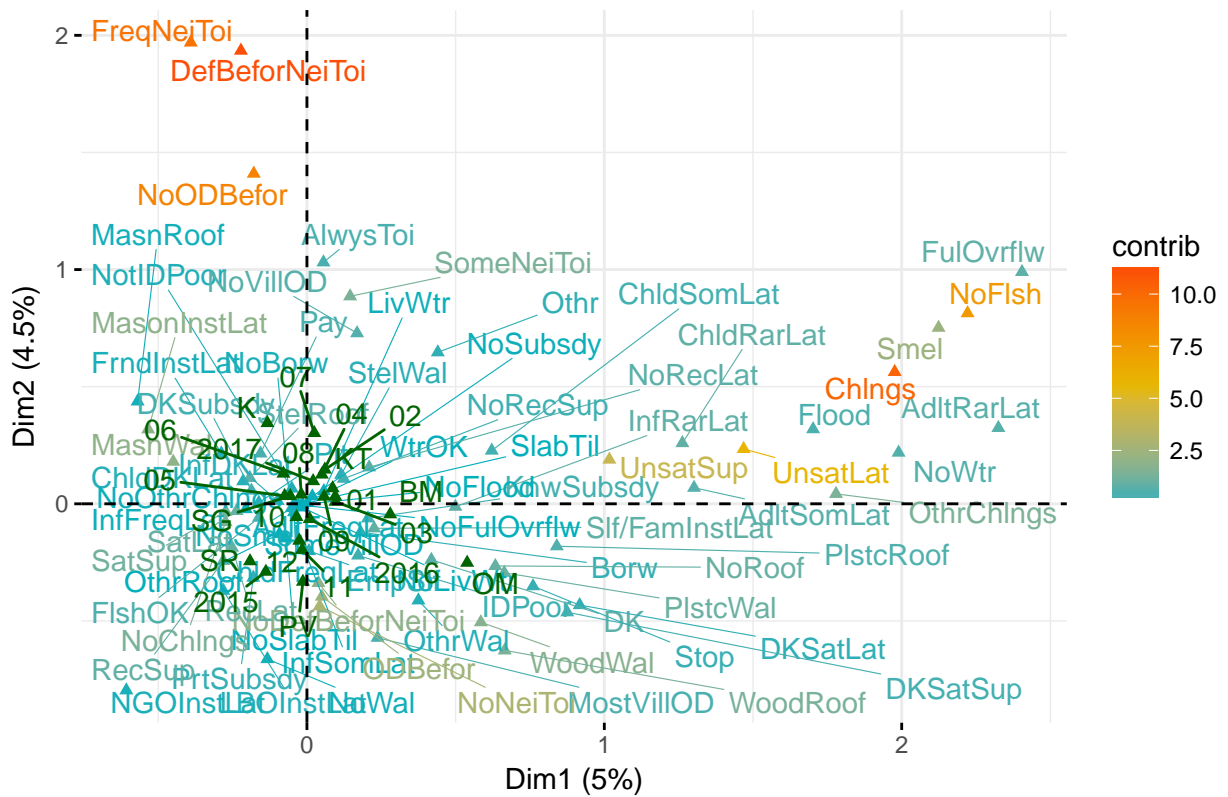
```
fviz_contrib(res.mca, choice = "var", axes = 1:2, top = 15)
```



```
fviz_mca_var(res.mca, col.var = "contrib",
  gradient.cols = c("#00AFBB", "#E7B800", "#FC4E07"),
  repel = TRUE,
  ggtheme = theme_minimal())
```

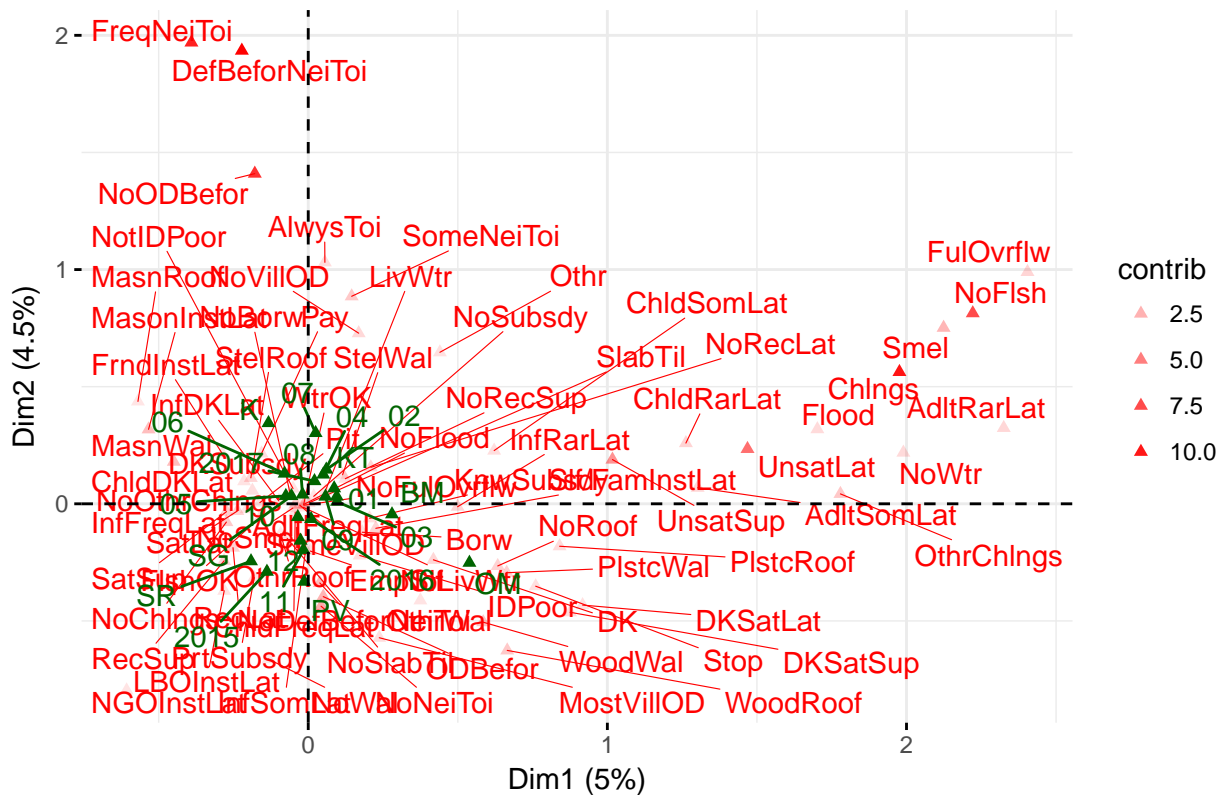


## Variable categories – MCA



```
fviz_mca_var(res.mca, alpha.var = "contrib",
  repel = TRUE,
  ggtheme = theme_minimal())
```

## Variable categories – MCA



```
res.mca.ind = get_mca_ind(res.mca)
print(res.mca.ind)
```

```
## Multiple Correspondence Analysis Results for individuals
## =====
##   Name      Description
## 1 "$coord"   "Coordinates for the individuals"
## 2 "$cos2"    "Cos2 for the individuals"
## 3 "$contrib" "contributions of the individuals"
```

```
head(res.mca.ind$coord)
```

```
##           Dim 1      Dim 2      Dim 3      Dim 4      Dim 5
## 154 -0.240932934  0.27914189 -0.09847273  0.13755096 -0.004505266
## 155 -0.140426748 -0.25611141 -0.19948985  0.25789534 -0.151063757
## 156  0.501385740 -0.21275166 -0.41714123  0.12991465 -0.096105515
## 157  0.028059233 -0.17517766  0.18410213  0.19425809 -0.217321753
## 158 -0.131075218 -0.07847411  0.04550284 -0.10031911 -0.037179985
## 159  0.006733437 -0.14859122 -0.33604633  0.02250878 -0.096037171
```

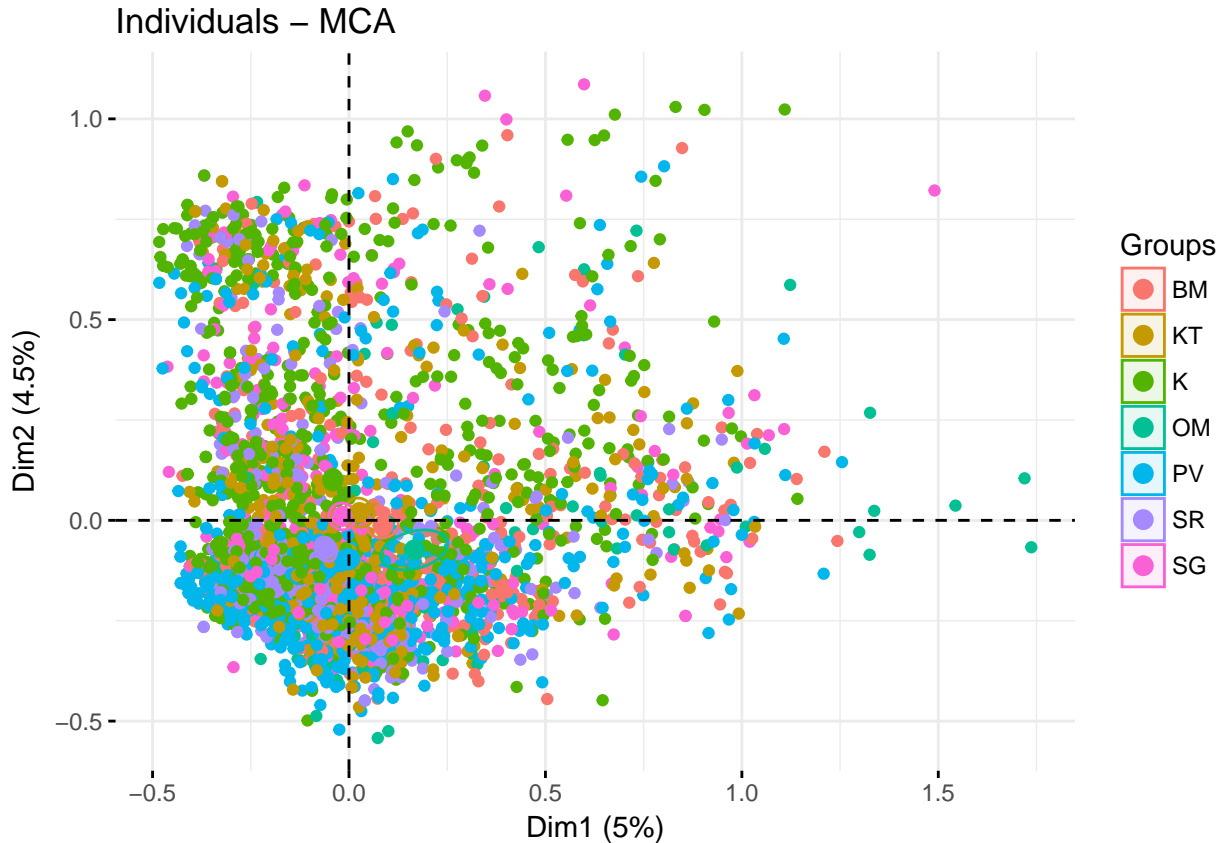
```
head(res.mca.ind$contrib)
```

```
##           Dim 1      Dim 2      Dim 3      Dim 4      Dim 5
## 154 1.775970e-02 0.026356764 0.0039226707 0.0080031039 1.043467e-05
## 155 6.033133e-03 0.022187068 0.0160987269 0.0281331379 1.173163e-02
## 156 7.691077e-02 0.015310451 0.0703908438 0.0071391658 4.748259e-03
## 157 2.408767e-04 0.010380052 0.0137109509 0.0159620780 2.427975e-02
## 158 5.256352e-03 0.002083027 0.0008375814 0.0042569496 7.106508e-04
## 159 1.387129e-05 0.007468414 0.0456823288 0.0002143064 4.741508e-03
```

```
head(res.mca.ind$cos2)
```

```
##           Dim 1      Dim 2      Dim 3      Dim 4      Dim 5
## 154 5.539237e-02 0.07435456 0.009253149 0.0180544760 1.936862e-05
## 155 2.226336e-02 0.07405406 0.044929610 0.0750892973 2.576390e-02
## 156 8.971537e-02 0.01615358 0.062099650 0.0060233652 3.296245e-03
## 157 8.492946e-04 0.03310276 0.036561535 0.0407066180 5.094636e-02
## 158 4.632887e-02 0.01660593 0.005583257 0.0271380022 3.727597e-03
## 159 3.172318e-05 0.01544861 0.079013528 0.0003544925 6.453300e-03
```

```
# fviz_mca_ind(res.mca, col.ind = "cos2",
#             gradient.cols = c("#00AFBB", "#E7B800", "#FC4E07"),
#             repel = TRUE,
#             ggtheme = theme_minimal())
fviz_mca_ind(res.mca,
  label = "none", # hide individual labels
  habillage = data.mca$Prov, # color by groups
  # palette = c("#00AFBB", "#E7B800"),
  addEllipses = TRUE, ellipse.type = "confidence",
  ggtheme = theme_minimal())
```



## Introduction

Academic workflow, certainly in political science, is at a crossroads. The *American Journal of Political Science* (*AJPS*) announced a (my words) “[show your work](#)” initiative in which authors who are tentatively accepted for publication at the journal must hand over the raw code and data that produced the results shown in the manuscript. The editorial team at *AJPS* then reproduces the code from the manuscript. Pending successful replication, the manuscript moves toward publication. The *AJPS* might be at the fore of this movement, and it could be the most aggressive among political science journals, but other journals in our field have signed the joint [Data Access & Research Transparency](#) (DART) initiative. This, at a bare minimum, requires uploading code from quantitatively-oriented published articles to in-house directories hosted by the journal or to services like [Dataverse](#).

There are workflow implications to the Lacour controversy as well. Political science, for the foreseeable future, will struggle with the extent of [the data fraud perpetrated by Michael Lacour](#) in an article co-authored with Donald P. Green in *Science*, the general scientific journal of record in the United States. A failure to reproduce LaCour’s results with different samples uncovered a comprehensive effort by LaCour to “fake” data that provided results to what we felt or believed to be true (i.e. “truthiness”). However, [fake data can have real consequences](#) for both the researcher and those who want to learn from it and use it for various purposes. Even research done honestly may suffer the same fate if researchers are not diligent in their workflow.

These recent events underscore the DART push and cast a shadow over our workflow. However, good workflow has always been an issue in our discipline. Cloud storage services like [Drop-](#)

box are still relatively new among political scientists. Without cloud storage, previous workflow left open the possibility that work between a home computer and an office computer was lost as a function of a corrupted thumb drive, an overheated power supply, or, among other things, the wave of viruses that [would particularly affect Microsoft users every summer](#). Social sciences, [unlike engineering](#), have traditionally relied on software like Microsoft Word for manuscript preparation though any word processor reduces workflow to a series of clicks and strokes on a keyboard. This is [a terrible way to track changes](#) or maintain version control. The addition of collaborators only compounds all the aforementioned issues. The proverbial left hand may not know what the right hand is doing.

I think there is reason for optimism. We only struggle with it now because we have tools like [R Markdown](#) and [Pandoc](#), more generally, that make significant strides in workflow. LaTeX resolved earlier issues of corrupted binary files by reducing documents to raw markup that was little more than raw text and revisions that could be easily kept as [“commented” text](#). However, for all its benefits (including pretty PDFs), [LaTeX is ugly code](#) and does not provide means of seamlessly working with the actual data analysis itself. R Markdown both eliminates markup and allows the author and her collaborators to write and reproduce the manuscript in one fell swoop.

## Getting Started with YAML

The lion’s share of a R Markdown document will be raw text, though the front matter may be the most important part of the document. R Markdown uses [YAML](#) for its metadata and the fields differ from [what an author would use for a Beamer presentation](#). I provide a sample YAML metadata largely taken from this exact document and explain it below.

```
---
output:
  pdf_document:
    citation_package: natbib
    keep_tex: true
    fig_caption: true
    latex_engine: pdflatex
    template: ~/Dropbox/miscelanea/svm-r-markdown-templates/svm-latex-ms.tex
title: "A Pandoc Markdown Article Starter and Template"
thanks: "Replication files are available on the author's Github account..."
author:
- name: Steven V. Miller
  affiliation: Clemson University
- name: Mary Margaret Albright
  affiliation: Pendelton State University
- name: Rembrandt Q. Einstein
  affiliation: Springfield University
abstract: "This document provides an introduction to R Markdown, argues for its..."
keywords: "pandoc, r markdown, knitr"
date: "'r format(Sys.time(), '%B %d, %Y')'"
geometry: margin=1in
fontfamily: mathpazo
fontsize: 11pt
# spacing: double
```

```
bibliography: ~/Dropbox/master.bib
biblio-style: apsr
---
```

output: will tell R Markdown we want a PDF document rendered with LaTeX. Since we are adding a fair bit of custom options to this call, we specify pdf\_document: on the next line (with, importantly, a two-space indent). We specify additional output-level options underneath it, each are indented with four spaces. citation\_package: natbib tells R Markdown to use natbib to handle bibliographic citations.<sup>1</sup> Thereafter, the next line (keep\_tex: true) tells R Markdown to render a raw .tex file along with the PDF document. This is useful for both debugging and the publication stage, when the editorial team will ask for the raw .tex so that they could render it and later provide page proofs. The next line fig\_caption: true tells R Markdown to make sure that whatever images are included in the document are treated as figures in which our caption in brackets in a Markdown call is treated as the caption in the figure. The next line (latex\_engine: pdflatex) tells R Markdown to use pdflatex and not some other option like lualatex. For my template, I'm pretty sure this is mandatory.<sup>2</sup>

The next line (template: ...) tells R Markdown to use my custom LaTeX template.<sup>3</sup> While I will own any errors in the code, I confess to “Frankensteining” this template from [the default LaTeX template](#) from Pandoc, [Kieran Healy's LaTeX template](#), and liberally using raw TeX from the [Association for Computing Machinery's \(ACM\) LaTeX template](#). I rather like that template since it resembles standard manuscripts when they are published in some of our more prominent journals. I will continue with a description of the YAML metadata in the next paragraph, though invite the curious reader to scroll to the end of the accompanying post to see the PDF this template produces.

The next fields get to the heart of the document itself. title: is, intuitively, the title of the manuscript. Do note that fields like title: do not have to be in quotation marks, but must be in quotation marks if the title of the document includes a colon. That said, the only reason to use a colon in an article title is if it is followed by a subtitle, hence the optional field (subtitle:). Notice I “comment out” the subtitle in the above example with a pound sign since this particular document does not have a subtitle. If thanks: is included and has an accompanying entry, the ensuing title of the document gets an asterisk and a footnote. This field is typically used to advise readers that the document is a working paper or is forthcoming in a journal.

The next field (author:) is a divergence from standard YAML, but I think it is useful. I will also confess to pilfering this idea from Kieran Healy's template. Typically, multiple authors for a given document are separated by an \and in this field. However, standard LaTeX then creates a tabular field separating multiple authors that is somewhat restrictive and not easy to override. As a result, I use this setup (again, taken from Kieran Healy) to sidestep the restrictive rendering of authors in the standard \maketitle tag. After author:, enter - name: (no space before the dash) and fill in

---

<sup>1</sup>R Markdown can use Pandoc's native bibliography management system or even biblatex, but I've found that it chokes with some of the more advanced stuff I've done with my .bib file over the years. For example, I've been diligent about special characters (e.g. umlauts and acute accents) in author names in my .bib file, but Pandoc's native citation system will choke on these characters in a .bib file. I effectively need natbib for my own projects.

<sup>2</sup>The main reason I still use pdflatex (and most readers probably do as well) is because of LaTeX fonts. [Unlike others](#), I find standard LaTeX fonts to be appealing.

<sup>3</sup>Notice that the path is relative. The user can, if she wishes, install this in the default Pandoc directory. I don't think this is necessary. Just be mindful of wherever the template is placed. Importantly, ~ is used in R to find the home directory (not necessarily the working directory). It is equivalent to saying /home/steve in Linux, or /Users/steve on a Mac, in my case.

the field with the first author. On the next line, enter two spaces, followed by affiliation: and the institute or university affiliation of the first author.

Do notice this can be repeated for however many co-authors there are to a manuscript. The rendered PDF will enter each co-author in a new line in a manner similar to journals like *American Journal of Political Science*, *American Political Science Review*, or *Journal of Politics*.

The next two fields pertain to the frontmatter of a manuscript. They should also be intuitive for the reader. abstract should contain the abstract and keywords should contain some keywords that describe the research project. Both fields are optional, though are practically mandatory. Every manuscript requires an abstract and some journals—especially those published by Sage—request them with submitted manuscripts. My template also includes these keywords in the PDF’s metadata.

date comes standard with R Markdown and you can use it to enter the date of the most recent compile. I typically include the date of the last compile for a working paper in the thanks: field, so this field currently does not do anything in my Markdown-LaTeX manuscript template. I include it in my YAML as a legacy, basically.

The next items are optional and cosmetic. geometry: is a standard option in LaTeX. I set the margins at one inch, and you probably should too. fontfamily: is optional, but I use it to specify the Palatino font. The default option is Computer Modern Roman. fontsize: sets, intuitively, the font size. The default is 10-point, but I prefer 11-point. spacing: is an optional field. If it is set as “double”, the ensuing document is double-spaced. “single” is the only other valid entry for this field, though not including the entry in the YAML metadata amounts to singling the document by default. Notice I have this “commented out” in the example code.

The final two options pertain to the bibliography. bibliography: specifies the location of the .bib file, so the author could make citations in the manuscript. biblio-style specifies the type of bibliography to use. You’ll typically set this as APSR. You could also specify the relative path of [my Journal of Peace Research .bst file](#) if you are submitting to that journal.

## Getting Started with Markdown Syntax

There are a lot of cheatsheets and reference guides for Markdown (e.g. [Adam Prichard](#), [Assemble](#), [Rstudio](#), [Rstudio again](#), [Scott Boms](#), [Daring Fireball](#), among, I’m sure, several others). I encourage the reader to look at those, though I will retread these references here with a minimal working example below.

```
# Introduction

**Lorem ipsum** dolor *sit amet*.

- Single asterisks italicize text *like this*.
- Double asterisks embolden text **like this**.

Start a new paragraph with a blank line separating paragraphs.

- This will start an unordered list environment, and this will be the first item.
- This will be a second item.
- A third item.
  - Four spaces and a dash create a sublist and this item in it.
```

```

- The fourth item.

1. This starts a numerical list.
2. This is no. 2 in the numerical list.

# This Starts A New Section
## This is a Subsection
### This is a Subsubsection
#### This starts a Paragraph Block.

> This will create a block quote, if you want one.

Want a table? This will create one.

Table Header | Second Header
----- | -----
Table Cell | Cell 2
Cell 3 | Cell 4

Note that the separators do not have to be aligned.

Want an image? This will do it.

![caption for my image](path/to/image.jpg)

'fig_caption: yes' will provide a caption. Put that in the YAML metadata.

Almost forgot about creating a footnote.[^1] This will do it again.[^2]

[^1]: The first footnote
[^2]: The second footnote

Want to cite something?

- Find your biblatexkey in your bib file.
- Put an @ before it, like @smith1984, or whatever it is.
- @smith1984 creates an in-text citation (e.g. Smith (1984) says...)
- [@smith1984] creates a parenthetical citation (Smith, 1984)

That'll also automatically create a reference list at the end of the document.

[In-text link to Google](http://google.com) as well.

```

That's honestly it. Markdown takes the chore of markup from your manuscript (hence: "Markdown").

On that note, you could easily pass most LaTeX code through Markdown if you're writing a LaTeX document. However, you don't need to do this (unless you're using the math environment) and probably shouldn't anyway if you intend to share your document in HTML as well.



## Using R Markdown with Knitr

Perhaps the greatest intrigue of R Markdown comes with the [knitr package](#) provided by ?. In other words, the author can, if she chooses, do the analysis in the Markdown document itself and compile/execute it in R.

Take, for example, this simple exercise using the `voteincome` data from the `Zelig` package. Suppose I want to explain the decision to vote using data from this package. I load in the data, clean the data, run the analyses, and present the results as a coefficient plot.

Here's what this code looks like. All I did was create a code display, which starts with three *backticks* (i.e. those ticks next to the number 1 key on your keyboard) and ends with three backticks on another line. On the first line of backticks (i.e. to start the code display) enter `{r, eval=FALSE, tidy=TRUE}`. The `eval=FALSE` option just displays the R code (and does not run it), `tidy=TRUE` wraps long code so it does not run off the page.

Within that code display, I enter my R code like this.

```
suppressMessages(library(Zelig))
suppressMessages(library(arm))
suppressMessages(library(coefplot))

data(voteincome)

voteincome$z.age <- arm::rescale(voteincome$age)
voteincome$z.education <- arm::rescale(voteincome$education)
voteincome$z.income <- arm::rescale(voteincome$income)

M1 <- glm(vote ~ z.age + female + z.education + z.income, data = voteincome,
         family = binomial)

coefplot(M1)
```

The implications for workflow are fairly substantial. Authors can rather quickly display the code they used to run the analyses in the document itself (likely in the appendix). As such, there's little guesswork for reviewers and editors in understanding what the author did in the analyses reported in the manuscript.

It doesn't end there. In fact, here's what happens when `eval=FALSE` is omitted or changed to `eval=TRUE`. Now, the code runs within R. Observe.

```
suppressMessages(library(Zelig))

## Warning: package 'Zelig' was built under R version 3.4.4

suppressMessages(library(arm))

## Warning: package 'arm' was built under R version 3.4.4

## Warning: package 'lme4' was built under R version 3.4.4
```

```
data(voteincome)

voteincome$z.age <- arm::rescale(voteincome$age)
voteincome$z.education <- arm::rescale(voteincome$education)
voteincome$z.income <- arm::rescale(voteincome$income)

M1 <- glm(vote ~ z.age + female + z.education + z.income, data = voteincome,
  family = binomial)

arm::coefplot(M1)
```

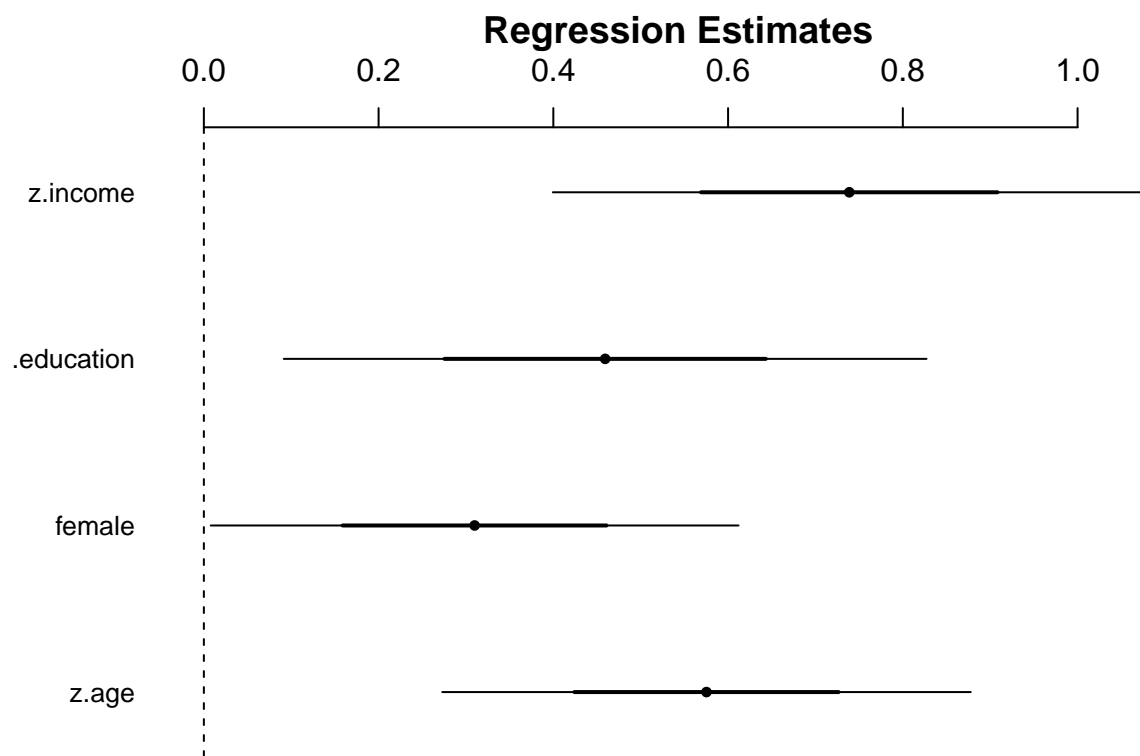


Figure 1: A Coefficient Plot

To get knitr to present the results of a table, add `results="asis"` to the brackets to start the R code chunk. The ensuing output will look like this (though the table may come on the next page).

```
suppressMessages(library(Zelig))
suppressMessages(library(stargazer))
suppressMessages(library(arm))

data(voteincome)
```

```

voteincome$z.age <- arm::rescale(voteincome$age)
voteincome$z.education <- arm::rescale(voteincome$education)
voteincome$z.income <- arm::rescale(voteincome$income)

M1 <- glm(vote ~ z.age + female + z.education + z.income, data = voteincome,
  family = binomial)

stargazer(M1, title = "A Handsome Table", header = FALSE)

```

Table 1: A Handsome Table	
	<i>Dependent variable:</i>
	vote
z.age	0.575*** (0.151)
female	0.310** (0.151)
z.education	0.459** (0.184)
z.income	0.739*** (0.170)
Constant	1.706*** (0.110)
Observations	1,500
Log Likelihood	−592.801
Akaike Inf. Crit.	1,195.602
<i>Note:</i> *p<0.1; **p<0.05; ***p<0.01	

Adding `echo="FALSE"` inside the brackets to start the R chunk will omit the presentation of the R commands. It will just present the table. This provides substantial opportunity for authors in doing their analyses. Now, the analysis and presentation in the form of a polished manuscript can be effectively simultaneous.<sup>4</sup>

<sup>4</sup>I'm not sure if I'm ready to commit to this myself since my workflow is still largely derived from [Rob J. Hyndman's example](#). However, *knitr* has endless potential, especially when analyses can be stored in cache, saved as chunks, or loaded in the preamble of a document to reference later in the manuscript.