FunModelling

Rob Eva

2022-04-14

HAVEN Data Profiling

# Read the test file into a variable  
my\_test <- read.csv('example\_data\_set.csv', header=TRUE)  
  
# Transform columns to dates  
my\_test$DissolutionDate <- as.Date(my\_test$DissolutionDate, format = "%d/%m/%y")  
my\_test$IncorporationDate <- as.Date(my\_test$IncorporationDate, format = "%d/%m/%y")  
my\_test$Accounts.NextDueDate <- as.Date(my\_test$Accounts.NextDueDate, format = "%d/%m/%y")  
my\_test$Accounts.LastMadeUpDate <- as.Date(my\_test$Accounts.LastMadeUpDate, format = "%d/%m/%y")  
my\_test$Returns.NextDueDate <- as.Date(my\_test$Returns.NextDueDate, format = "%d/%m/%y")  
my\_test$Returns.LastMadeUpDate <- as.Date(my\_test$Returns.LastMadeUpDate, format = "%d/%m/%y")  
my\_test$ConfStmtNextDueDate <- as.Date(my\_test$ConfStmtNextDueDate, format = "%d/%m/%y")  
my\_test$ConfStmtLastMadeUpDate <- as.Date(my\_test$ConfStmtLastMadeUpDate, format = "%d/%m/%y")  
  
# Glimpse comes from the dply package  
glimpse(my\_test)

## Rows: 10,001  
## Columns: 55  
## $ CompanyName <chr> "E & N LOFTS LTD", "E & N METAL AND~  
## $ CompanyNumber <chr> "13484610", "11093347", "SO306680",~  
## $ RegAddress.CareOf <chr> "", "MR E POWELL", "", "", "", "", ~  
## $ RegAddress.POBox <chr> "", "", "", "", "", "", "", "", "",~  
## $ RegAddress.AddressLine1 <chr> "JUBILEE HOUSE", "1 CAE CADNO", "1 ~  
## $ RegAddress.AddressLine2 <chr> "EAST BEACH", "CHURCH VILLAGE", "",~  
## $ RegAddress.PostTown <chr> "LYTHAM ST.ANNES", "PONTYPRIDD", "P~  
## $ RegAddress.County <chr> "LANCASHIRE", "MID GLAMORGAN", "AYR~  
## $ RegAddress.Country <chr> "ENGLAND", "WALES", "UNITED KINGDOM~  
## $ RegAddress.PostCode <chr> "FY8 5FT", "CF38 1UL", "KA9 2TA", "~  
## $ CompanyCategory <chr> "Private Limited Company", "Private~  
## $ CompanyStatus <chr> "Active", "Active - Proposal to Str~  
## $ CountryOfOrigin <chr> "United Kingdom", "United Kingdom",~  
## $ DissolutionDate <date> NA, NA, NA, NA, NA, NA, NA, NA, NA~  
## $ IncorporationDate <date> 2020-06-30, 2020-12-04, 2020-03-29~  
## $ Accounts.AccountRefDay <int> 31, 31, 31, 31, 30, 31, 31, 31, 31,~  
## $ Accounts.AccountRefMonth <int> 7, 12, 3, 5, 9, 10, 10, 10, 12, 6, ~  
## $ Accounts.NextDueDate <date> 2020-03-31, 2020-12-31, 2020-12-31~  
## $ Accounts.LastMadeUpDate <date> NA, 2020-12-31, 2020-03-31, 2020-0~  
## $ Accounts.AccountCategory <chr> "NO ACCOUNTS FILED", "MICRO ENTITY"~  
## $ Returns.NextDueDate <date> 2020-07-28, 2020-01-01, 2020-04-26~  
## $ Returns.LastMadeUpDate <date> NA, NA, NA, 2020-05-18, NA, 2020-1~  
## $ Mortgages.NumMortCharges <int> 0, 0, 6, 1, 0, 0, 0, 0, 0, 0, 1, 0,~  
## $ Mortgages.NumMortOutstanding <int> 0, 0, 6, 1, 0, 0, 0, 0, 0, 0, 0, 0,~  
## $ Mortgages.NumMortPartSatisfied <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,~  
## $ Mortgages.NumMortSatisfied <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0,~  
## $ SICCode.SicText\_1 <chr> "43320 - Joinery installation", "46~  
## $ SICCode.SicText\_2 <chr> "", "", "", "", "78200 - Temporary ~  
## $ SICCode.SicText\_3 <chr> "", "", "", "", "", "", "", "", "",~  
## $ SICCode.SicText\_4 <chr> "", "", "", "", "", "", "", "", "",~  
## $ LimitedPartnerships.NumGenPartners <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,~  
## $ LimitedPartnerships.NumLimPartners <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,~  
## $ URI <chr> "http://business.data.gov.uk/id/com~  
## $ PreviousName\_1.CONDATE <chr> "", "12/12/2017", "", "", "06/10/20~  
## $ PreviousName\_1.CompanyName <chr> "", "E & M METAL AND WASTE RECYCLIN~  
## $ PreviousName\_2.CONDATE <chr> "", "", "", "", "", "", "", "", "31~  
## $ PreviousName\_2.CompanyName <chr> "", "", "", "", "", "", "", "", "50~  
## $ PreviousName\_3.CONDATE <chr> "", "", "", "", "", "", "", "", "",~  
## $ PreviousName\_3.CompanyName <chr> "", "", "", "", "", "", "", "", "",~  
## $ PreviousName\_4.CONDATE <chr> "", "", "", "", "", "", "", "", "",~  
## $ PreviousName\_4.CompanyName <chr> "", "", "", "", "", "", "", "", "",~  
## $ PreviousName\_5.CONDATE <chr> "", "", "", "", "", "", "", "", "",~  
## $ PreviousName\_5.CompanyName <chr> "", "", "", "", "", "", "", "", "",~  
## $ PreviousName\_6.CONDATE <lgl> NA, NA, NA, NA, NA, NA, NA, NA, NA,~  
## $ PreviousName\_6.CompanyName <lgl> NA, NA, NA, NA, NA, NA, NA, NA, NA,~  
## $ PreviousName\_7.CONDATE <lgl> NA, NA, NA, NA, NA, NA, NA, NA, NA,~  
## $ PreviousName\_7.CompanyName <lgl> NA, NA, NA, NA, NA, NA, NA, NA, NA,~  
## $ PreviousName\_8.CONDATE <lgl> NA, NA, NA, NA, NA, NA, NA, NA, NA,~  
## $ PreviousName\_8.CompanyName <lgl> NA, NA, NA, NA, NA, NA, NA, NA, NA,~  
## $ PreviousName\_9.CONDATE <lgl> NA, NA, NA, NA, NA, NA, NA, NA, NA,~  
## $ PreviousName\_9.CompanyName <lgl> NA, NA, NA, NA, NA, NA, NA, NA, NA,~  
## $ PreviousName\_10.CONDATE <lgl> NA, NA, NA, NA, NA, NA, NA, NA, NA,~  
## $ PreviousName\_10.CompanyName <lgl> NA, NA, NA, NA, NA, NA, NA, NA, NA,~  
## $ ConfStmtNextDueDate <date> 2020-07-23, 2020-12-03, 2020-04-11~  
## $ ConfStmtLastMadeUpDate <date> 2020-07-09, 2020-10-22, 2020-03-28~

# Remove some extraneous previous name fields  
my\_test <- my\_test %>%  
 select(-PreviousName\_2.CONDATE, -PreviousName\_2.CompanyName, -PreviousName\_3.CONDATE, -PreviousName\_3.CompanyName,  
 -PreviousName\_4.CONDATE, -PreviousName\_4.CompanyName, -PreviousName\_5.CONDATE, -PreviousName\_5.CompanyName,   
 -PreviousName\_6.CONDATE, -PreviousName\_6.CompanyName, -PreviousName\_7.CONDATE, -PreviousName\_7.CompanyName,  
 -PreviousName\_8.CONDATE, -PreviousName\_8.CompanyName, -PreviousName\_9.CONDATE, -PreviousName\_9.CompanyName,  
 -PreviousName\_10.CONDATE, -PreviousName\_10.CompanyName)  
  
  
# Changes blank fields to NA  
my\_test[my\_test==""] <- NA  
  
# Data Integrity  
data\_integrity(my\_test)

## $vars\_num\_with\_NA  
## variable q\_na p\_na  
## Accounts.AccountRefDay Accounts.AccountRefDay 82 0.00819918  
## Accounts.AccountRefMonth Accounts.AccountRefMonth 82 0.00819918  
##   
## $vars\_cat\_with\_NA  
## variable q\_na p\_na  
## RegAddress.POBox RegAddress.POBox 9975 0.99740026  
## RegAddress.CareOf RegAddress.CareOf 9906 0.99050095  
## SICCode.SicText\_4 SICCode.SicText\_4 9703 0.97020298  
## SICCode.SicText\_3 SICCode.SicText\_3 9361 0.93600640  
## PreviousName\_1.CONDATE PreviousName\_1.CONDATE 9060 0.90590941  
## PreviousName\_1.CompanyName PreviousName\_1.CompanyName 9060 0.90590941  
## SICCode.SicText\_2 SICCode.SicText\_2 8485 0.84841516  
## RegAddress.County RegAddress.County 6281 0.62803720  
## RegAddress.AddressLine2 RegAddress.AddressLine2 3542 0.35416458  
## RegAddress.Country RegAddress.Country 3121 0.31206879  
## RegAddress.PostTown RegAddress.PostTown 137 0.01369863  
## RegAddress.PostCode RegAddress.PostCode 64 0.00639936  
## RegAddress.AddressLine1 RegAddress.AddressLine1 18 0.00179982  
##   
## $vars\_cat\_high\_card  
## variable unique  
## CompanyName CompanyName 10001  
## CompanyNumber CompanyNumber 10001  
## URI URI 10001  
## RegAddress.AddressLine1 RegAddress.AddressLine1 8276  
## RegAddress.PostCode RegAddress.PostCode 7951  
## RegAddress.AddressLine2 RegAddress.AddressLine2 4688  
## RegAddress.PostTown RegAddress.PostTown 1579  
## PreviousName\_1.CompanyName PreviousName\_1.CompanyName 938  
## PreviousName\_1.CONDATE PreviousName\_1.CONDATE 795  
## SICCode.SicText\_1 SICCode.SicText\_1 566  
## RegAddress.County RegAddress.County 319  
## SICCode.SicText\_2 SICCode.SicText\_2 319  
## SICCode.SicText\_3 SICCode.SicText\_3 208  
## SICCode.SicText\_4 SICCode.SicText\_4 129  
## RegAddress.CareOf RegAddress.CareOf 89  
##   
## $MAX\_UNIQUE  
## [1] 35  
##   
## $vars\_one\_value  
## character(0)  
##   
## $vars\_cat  
## [1] "CompanyName" "CompanyNumber"   
## [3] "RegAddress.CareOf" "RegAddress.POBox"   
## [5] "RegAddress.AddressLine1" "RegAddress.AddressLine2"   
## [7] "RegAddress.PostTown" "RegAddress.County"   
## [9] "RegAddress.Country" "RegAddress.PostCode"   
## [11] "CompanyCategory" "CompanyStatus"   
## [13] "CountryOfOrigin" "Accounts.AccountCategory"   
## [15] "SICCode.SicText\_1" "SICCode.SicText\_2"   
## [17] "SICCode.SicText\_3" "SICCode.SicText\_4"   
## [19] "URI" "PreviousName\_1.CONDATE"   
## [21] "PreviousName\_1.CompanyName"  
##   
## $vars\_num  
## [1] "Accounts.AccountRefDay" "Accounts.AccountRefMonth"   
## [3] "Mortgages.NumMortCharges" "Mortgages.NumMortOutstanding"   
## [5] "Mortgages.NumMortPartSatisfied" "Mortgages.NumMortSatisfied"   
## [7] "LimitedPartnerships.NumGenPartners" "LimitedPartnerships.NumLimPartners"  
##   
## $vars\_char  
## [1] "CompanyName" "CompanyNumber"   
## [3] "RegAddress.CareOf" "RegAddress.POBox"   
## [5] "RegAddress.AddressLine1" "RegAddress.AddressLine2"   
## [7] "RegAddress.PostTown" "RegAddress.County"   
## [9] "RegAddress.Country" "RegAddress.PostCode"   
## [11] "CompanyCategory" "CompanyStatus"   
## [13] "CountryOfOrigin" "Accounts.AccountCategory"   
## [15] "SICCode.SicText\_1" "SICCode.SicText\_2"   
## [17] "SICCode.SicText\_3" "SICCode.SicText\_4"   
## [19] "URI" "PreviousName\_1.CONDATE"   
## [21] "PreviousName\_1.CompanyName"  
##   
## $vars\_factor  
## character(0)  
##   
## $vars\_other  
## [1] "DissolutionDate" "IncorporationDate"   
## [3] "Accounts.NextDueDate" "Accounts.LastMadeUpDate"  
## [5] "Returns.NextDueDate" "Returns.LastMadeUpDate"   
## [7] "ConfStmtNextDueDate" "ConfStmtLastMadeUpDate"

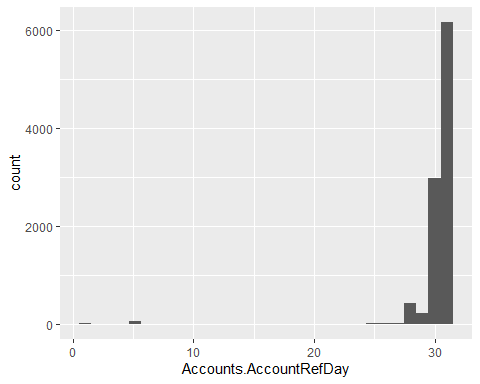
# Descriptive Statistics  
summary(my\_test)

## CompanyName CompanyNumber RegAddress.CareOf RegAddress.POBox   
## Length:10001 Length:10001 Length:10001 Length:10001   
## Class :character Class :character Class :character Class :character   
## Mode :character Mode :character Mode :character Mode :character   
##   
##   
##   
##   
## RegAddress.AddressLine1 RegAddress.AddressLine2 RegAddress.PostTown  
## Length:10001 Length:10001 Length:10001   
## Class :character Class :character Class :character   
## Mode :character Mode :character Mode :character   
##   
##   
##   
##   
## RegAddress.County RegAddress.Country RegAddress.PostCode CompanyCategory   
## Length:10001 Length:10001 Length:10001 Length:10001   
## Class :character Class :character Class :character Class :character   
## Mode :character Mode :character Mode :character Mode :character   
##   
##   
##   
##   
## CompanyStatus CountryOfOrigin DissolutionDate IncorporationDate   
## Length:10001 Length:10001 Min. :NA Min. :2018-01-05   
## Class :character Class :character 1st Qu.:NA 1st Qu.:2020-02-13   
## Mode :character Mode :character Median :NA Median :2020-05-25   
## Mean :NA Mean :2020-05-05   
## 3rd Qu.:NA 3rd Qu.:2020-09-15   
## Max. :NA Max. :2020-12-31   
## NA's :10001 NA's :2   
## Accounts.AccountRefDay Accounts.AccountRefMonth Accounts.NextDueDate  
## Min. : 1.00 Min. : 1.000 Min. :2019-01-24   
## 1st Qu.:30.00 1st Qu.: 3.000 1st Qu.:2020-04-26   
## Median :31.00 Median : 6.000 Median :2020-07-31   
## Mean :30.25 Mean : 6.438 Mean :2020-07-25   
## 3rd Qu.:31.00 3rd Qu.:10.000 3rd Qu.:2020-10-31   
## Max. :31.00 Max. :12.000 Max. :2020-12-31   
## NA's :82 NA's :82 NA's :110   
## Accounts.LastMadeUpDate Accounts.AccountCategory Returns.NextDueDate   
## Min. :2019-01-31 Length:10001 Min. :2019-01-12   
## 1st Qu.:2020-03-31 Class :character 1st Qu.:2020-04-02   
## Median :2020-06-30 Mode :character Median :2020-06-30   
## Mean :2020-07-09 Mean :2020-07-01   
## 3rd Qu.:2020-10-31 3rd Qu.:2020-10-06   
## Max. :2020-12-31 Max. :2020-12-31   
## NA's :2400 NA's :145   
## Returns.LastMadeUpDate Mortgages.NumMortCharges Mortgages.NumMortOutstanding  
## Min. :2019-01-03 Min. : 0.0000 Min. : 0.0000   
## 1st Qu.:2020-03-27 1st Qu.: 0.0000 1st Qu.: 0.0000   
## Median :2020-06-20 Median : 0.0000 Median : 0.0000   
## Mean :2020-06-25 Mean : 0.8085 Mean : 0.4137   
## 3rd Qu.:2020-09-30 3rd Qu.: 0.0000 3rd Qu.: 0.0000   
## Max. :2020-12-31 Max. :401.0000 Max. :241.0000   
## NA's :5315   
## Mortgages.NumMortPartSatisfied Mortgages.NumMortSatisfied SICCode.SicText\_1   
## Min. :0e+00 Min. : 0.0000 Length:10001   
## 1st Qu.:0e+00 1st Qu.: 0.0000 Class :character   
## Median :0e+00 Median : 0.0000 Mode :character   
## Mean :4e-04 Mean : 0.3941   
## 3rd Qu.:0e+00 3rd Qu.: 0.0000   
## Max. :1e+00 Max. :258.0000   
##   
## SICCode.SicText\_2 SICCode.SicText\_3 SICCode.SicText\_4   
## Length:10001 Length:10001 Length:10001   
## Class :character Class :character Class :character   
## Mode :character Mode :character Mode :character   
##   
##   
##   
##   
## LimitedPartnerships.NumGenPartners LimitedPartnerships.NumLimPartners  
## Min. :0.000000 Min. : 0.0000   
## 1st Qu.:0.000000 1st Qu.: 0.0000   
## Median :0.000000 Median : 0.0000   
## Mean :0.006199 Mean : 0.0223   
## 3rd Qu.:0.000000 3rd Qu.: 0.0000   
## Max. :2.000000 Max. :65.0000   
##   
## URI PreviousName\_1.CONDATE PreviousName\_1.CompanyName  
## Length:10001 Length:10001 Length:10001   
## Class :character Class :character Class :character   
## Mode :character Mode :character Mode :character   
##   
##   
##   
##   
## ConfStmtNextDueDate ConfStmtLastMadeUpDate  
## Min. :2020-01-01 Min. :2020-01-01   
## 1st Qu.:2020-04-01 1st Qu.:2020-03-30   
## Median :2020-07-02 Median :2020-06-29   
## Mean :2020-07-03 Mean :2020-06-29   
## 3rd Qu.:2020-10-06 3rd Qu.:2020-09-30   
## Max. :2020-12-31 Max. :2020-12-31   
## NA's :52 NA's :1796

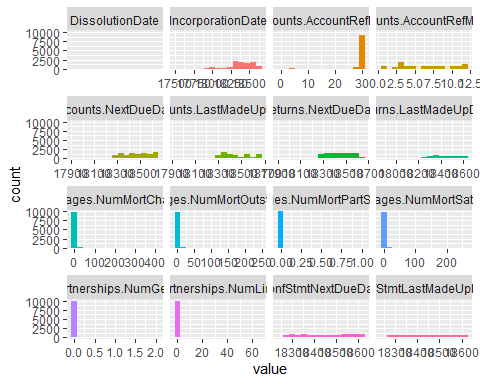
# Profile\_num  
# Get a metric table with many indicators for all numerical variables,   
# automatically skipping the non-numerical variables.   
# Current metrics are: mean, std\_dev: standard deviation,   
# all the p\_XX: percentile at XX number, skewness, kurtosis,   
# iqr: inter quartile range,   
# variation\_coef: the ratio of sd/mean, range\_98 is the limit for which the 98  
profiling\_num(my\_test)

## variable mean std\_dev variation\_coef  
## 1 Accounts.AccountRefDay 30.24548846 2.63814212 0.08722432  
## 2 Accounts.AccountRefMonth 6.43804819 3.59179007 0.55790046  
## 3 Mortgages.NumMortCharges 0.80851915 6.14664391 7.60234798  
## 4 Mortgages.NumMortOutstanding 0.41365863 3.30014688 7.97794753  
## 5 Mortgages.NumMortPartSatisfied 0.00039996 0.01999600 49.99499900  
## 6 Mortgages.NumMortSatisfied 0.39406059 3.60130542 9.13896359  
## 7 LimitedPartnerships.NumGenPartners 0.00619938 0.08100348 13.06638400  
## 8 LimitedPartnerships.NumLimPartners 0.02229777 0.75100117 33.68055030  
## p\_01 p\_05 p\_25 p\_50 p\_75 p\_95 p\_99 skewness kurtosis iqr range\_98  
## 1 18.18 28 30 31 31 31 31 -8.6907518 84.795545 1 [18.18, 31]  
## 2 1.00 1 3 6 10 12 12 0.1702749 1.665484 7 [1, 12]  
## 3 0.00 0 0 0 0 4 13 45.5095331 2669.867702 0 [0, 13]  
## 4 0.00 0 0 0 0 2 6 49.9718973 3236.478600 0 [0, 6]  
## 5 0.00 0 0 0 0 0 0 49.9724964 2498.250400 0 [0, 0]  
## 6 0.00 0 0 0 0 2 9 43.1470028 2748.771216 0 [0, 9]  
## 7 0.00 0 0 0 0 0 0 13.6932503 204.832839 0 [0, 0]  
## 8 0.00 0 0 0 0 0 0 69.8351345 5720.938479 0 [0, 0]  
## range\_80  
## 1 [30, 31]  
## 2 [2, 12]  
## 3 [0, 2]  
## 4 [0, 1]  
## 5 [0, 0]  
## 6 [0, 0]  
## 7 [0, 0]  
## 8 [0, 0]

# ggplot2 histogram of Accounts.AccountRefDay  
ggplot(my\_test, aes(x=Accounts.AccountRefDay)) + geom\_histogram()

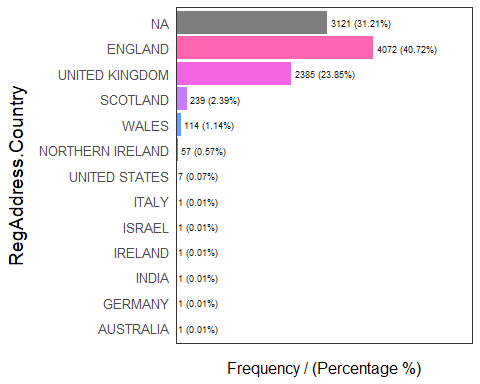


# Use plot\_num to profile continuous variables  
plot\_num(my\_test, bins = 15)

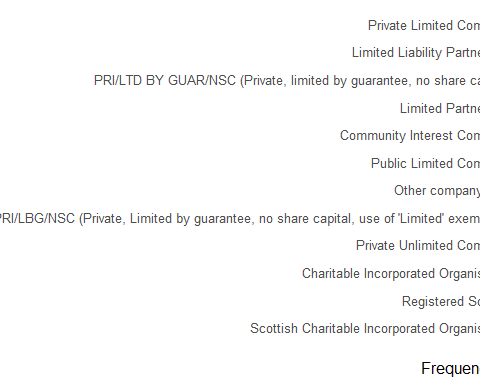


Change the options here to NOT include the code

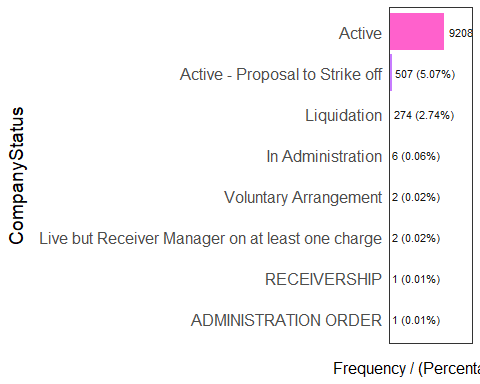
# Frequency Distributions



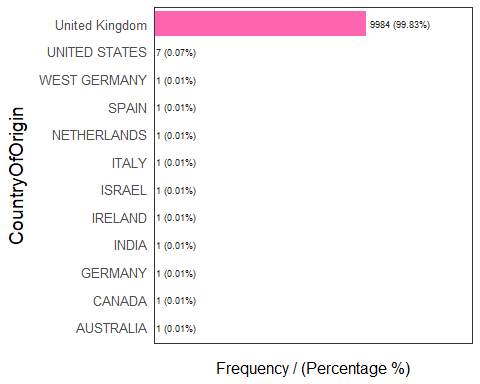
## RegAddress.Country frequency percentage cumulative\_perc  
## 1 ENGLAND 4072 40.72 40.72  
## 2 <NA> 3121 31.21 71.93  
## 3 UNITED KINGDOM 2385 23.85 95.78  
## 4 SCOTLAND 239 2.39 98.17  
## 5 WALES 114 1.14 99.31  
## 6 NORTHERN IRELAND 57 0.57 99.88  
## 7 UNITED STATES 7 0.07 99.95  
## 8 AUSTRALIA 1 0.01 99.96  
## 9 GERMANY 1 0.01 99.97  
## 10 INDIA 1 0.01 99.98  
## 11 IRELAND 1 0.01 99.99  
## 12 ISRAEL 1 0.01 100.00  
## 13 ITALY 1 0.01 100.00



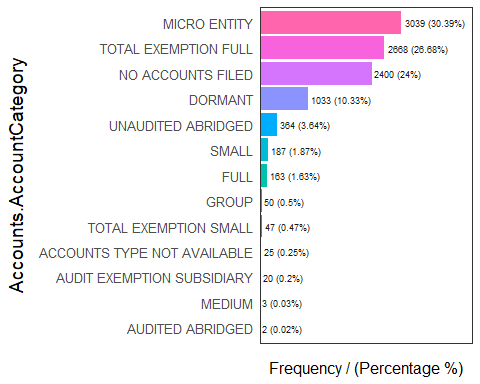
## CompanyCategory  
## 1 Private Limited Company  
## 2 Limited Liability Partnership  
## 3 PRI/LTD BY GUAR/NSC (Private, limited by guarantee, no share capital)  
## 4 Limited Partnership  
## 5 Community Interest Company  
## 6 Public Limited Company  
## 7 Other company type  
## 8 PRI/LBG/NSC (Private, Limited by guarantee, no share capital, use of 'Limited' exemption)  
## 9 Private Unlimited Company  
## 10 Charitable Incorporated Organisation  
## 11 Registered Society  
## 12 Scottish Charitable Incorporated Organisation  
## frequency percentage cumulative\_perc  
## 1 9690 96.89 96.89  
## 2 82 0.82 97.71  
## 3 67 0.67 98.38  
## 4 65 0.65 99.03  
## 5 21 0.21 99.24  
## 6 18 0.18 99.42  
## 7 17 0.17 99.59  
## 8 15 0.15 99.74  
## 9 13 0.13 99.87  
## 10 10 0.10 99.97  
## 11 2 0.02 99.99  
## 12 1 0.01 100.00



## CompanyStatus frequency percentage  
## 1 Active 9208 92.07  
## 2 Active - Proposal to Strike off 507 5.07  
## 3 Liquidation 274 2.74  
## 4 In Administration 6 0.06  
## 5 Live but Receiver Manager on at least one charge 2 0.02  
## 6 Voluntary Arrangement 2 0.02  
## 7 ADMINISTRATION ORDER 1 0.01  
## 8 RECEIVERSHIP 1 0.01  
## cumulative\_perc  
## 1 92.07  
## 2 97.14  
## 3 99.88  
## 4 99.94  
## 5 99.96  
## 6 99.98  
## 7 99.99  
## 8 100.00



## CountryOfOrigin frequency percentage cumulative\_perc  
## 1 United Kingdom 9984 99.83 99.83  
## 2 UNITED STATES 7 0.07 99.90  
## 3 AUSTRALIA 1 0.01 99.91  
## 4 CANADA 1 0.01 99.92  
## 5 GERMANY 1 0.01 99.93  
## 6 INDIA 1 0.01 99.94  
## 7 IRELAND 1 0.01 99.95  
## 8 ISRAEL 1 0.01 99.96  
## 9 ITALY 1 0.01 99.97  
## 10 NETHERLANDS 1 0.01 99.98  
## 11 SPAIN 1 0.01 99.99  
## 12 WEST GERMANY 1 0.01 100.00



## Accounts.AccountCategory frequency percentage cumulative\_perc  
## 1 MICRO ENTITY 3039 30.39 30.39  
## 2 TOTAL EXEMPTION FULL 2668 26.68 57.07  
## 3 NO ACCOUNTS FILED 2400 24.00 81.07  
## 4 DORMANT 1033 10.33 91.40  
## 5 UNAUDITED ABRIDGED 364 3.64 95.04  
## 6 SMALL 187 1.87 96.91  
## 7 FULL 163 1.63 98.54  
## 8 GROUP 50 0.50 99.04  
## 9 TOTAL EXEMPTION SMALL 47 0.47 99.51  
## 10 ACCOUNTS TYPE NOT AVAILABLE 25 0.25 99.76  
## 11 AUDIT EXEMPTION SUBSIDIARY 20 0.20 99.96  
## 12 MEDIUM 3 0.03 99.99  
## 13 AUDITED ABRIDGED 2 0.02 100.00

## [1] "Variables processed: RegAddress.Country, CompanyCategory, CompanyStatus, CountryOfOrigin, Accounts.AccountCategory"

# Probability Density  
plotar(my\_test, target = 'CompanyStatus', plot\_type = "histdens")

