ML1

2024-01-10

## Loading required package: car

## Warning: package 'car' was built under R version 4.3.3

## Loading required package: carData

## Warning: package 'carData' was built under R version 4.3.3

## Loading required package: ggplot2

## Warning: package 'ggplot2' was built under R version 4.3.3

## Loading required package: ROI

## Warning: package 'ROI' was built under R version 4.3.3

## ROI: R Optimization Infrastructure

## Registered solver plugins: nlminb, quadprog, symphony.

## Default solver: auto.

## Loading required package: tidyverse

## Warning: package 'tidyverse' was built under R version 4.3.3

## Warning: package 'tibble' was built under R version 4.3.3

## Warning: package 'tidyr' was built under R version 4.3.3

## Warning: package 'readr' was built under R version 4.3.3

## Warning: package 'purrr' was built under R version 4.3.3

## Warning: package 'dplyr' was built under R version 4.3.3

## Warning: package 'stringr' was built under R version 4.3.3

## Warning: package 'forcats' was built under R version 4.3.3

## Warning: package 'lubridate' was built under R version 4.3.3

## ── Attaching core tidyverse packages ──────────────────────── tidyverse 2.0.0 ──  
## ✔ dplyr 1.1.4 ✔ readr 2.1.5  
## ✔ forcats 1.0.0 ✔ stringr 1.5.1  
## ✔ lubridate 1.9.3 ✔ tibble 3.2.1  
## ✔ purrr 1.0.2 ✔ tidyr 1.3.1

## ── Conflicts ────────────────────────────────────────── tidyverse\_conflicts() ──  
## ✖ dplyr::filter() masks stats::filter()  
## ✖ dplyr::lag() masks stats::lag()  
## ✖ dplyr::recode() masks car::recode()  
## ✖ purrr::some() masks car::some()  
## ℹ Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors  
## Loading required package: corrplot

## Warning: package 'corrplot' was built under R version 4.3.3

## corrplot 0.94 loaded  
## Loading required package: GGally

## Warning: package 'GGally' was built under R version 4.3.3

## Registered S3 method overwritten by 'GGally':  
## method from   
## +.gg ggplot2  
## Loading required package: mgcv  
## Loading required package: nlme  
##   
## Attaching package: 'nlme'  
##   
## The following object is masked from 'package:dplyr':  
##   
## collapse  
##   
## This is mgcv 1.9-0. For overview type 'help("mgcv-package")'.  
## Loading required package: nnet  
##   
## Attaching package: 'nnet'  
##   
## The following object is masked from 'package:mgcv':  
##   
## multinom  
##   
## Loading required package: e1071

## Warning: package 'e1071' was built under R version 4.3.3

## Registered S3 methods overwritten by 'proxy':  
## method from   
## print.registry\_field registry  
## print.registry\_entry registry  
## Loading required package: MASS  
##   
## Attaching package: 'MASS'  
##   
## The following object is masked from 'package:dplyr':  
##   
## select  
##   
## Loading required package: arm

## Warning: package 'arm' was built under R version 4.3.3

## Loading required package: Matrix  
##   
## Attaching package: 'Matrix'  
##   
## The following objects are masked from 'package:tidyr':  
##   
## expand, pack, unpack  
##   
## Loading required package: lme4

## Warning: package 'lme4' was built under R version 4.3.3

## Warning in check\_dep\_version(): ABI version mismatch:   
## lme4 was built with Matrix ABI version 1  
## Current Matrix ABI version is 0  
## Please re-install lme4 from source or restore original 'Matrix' package

##   
## Attaching package: 'lme4'  
##   
## The following object is masked from 'package:nlme':  
##   
## lmList  
##   
##   
## arm (Version 1.14-4, built: 2024-4-1)  
##   
## Working directory is C:/Users/lucar/switchdrive/SyncVM/Sem 2/ML1/assignment/machine-learning-1  
##   
##   
## Attaching package: 'arm'  
##   
## The following object is masked from 'package:corrplot':  
##   
## corrplot  
##   
## The following object is masked from 'package:car':  
##   
## logit

## Warning: package 'caret' was built under R version 4.3.3

## Loading required package: lattice  
##   
## Attaching package: 'caret'  
##   
## The following object is masked from 'package:purrr':  
##   
## lift

Notizen Rafi SVM Teil: Cross validation or other methods for model comparing must be used on one single methods (e.g. you use Cross Validation to compare 2-3 SVM models). • Students are free to choose a measure of fit that they find more appropriate. • In case students cannot find an appropriate measure of fit, they can use the Root Mean Squared Error (RMSE).

Schliesslich ist zu beachten, dass die Zusammenstellung von Dokumenten einige Zeit in Anspruch nehmen kann… insbesondere wenn komplexe Modelle angepasst werden - In diesen Fällen können Sie die Argumentationsoption cache = TRUE verwenden, so dass ein Chunk nur dann neu ausgewertet wird nur dann neu ausgewertet wird, wenn er seit der letzten Kompilierung geändert wurde. Wenn der Chunk unverändert blieb unverändert, dann werden die alten Ergebnisse verwendet

# Data Preprocessing

Aus Quelle: Some predictors such as carrying capacity and seat number are removed from the dataset prior to data analysis and modeling since they are not correctly coded.

CHATGPT CCM TON Es macht keinen Sinn, dass Fahrzeuge einen Wert von 0 für die Variable CCM\_TON haben, wenn diese Variable den Hubraum oder das Gewicht des Motors in Kubikzentimetern (ccm) oder Tonnen angibt. Warum? Der Hubraum (ccm) gibt das Volumen der Zylinder eines Verbrennungsmotors an. Ein Wert von 0 wäre unplausibel, da ein Fahrzeug ohne Hubraum keinen funktionsfähigen Motor hätte. Wenn CCM\_TON das Gewicht des Motors in Tonnen angibt, wäre ebenfalls ein Wert von 0 unplausibel, da ein Fahrzeug ohne Motorgewicht nicht funktionsfähig wäre.

## [1] 508499 16

## spc\_tbl\_ [508,499 × 16] (S3: spec\_tbl\_df/tbl\_df/tbl/data.frame)  
## $ SEX : num [1:508499] 0 0 0 0 0 0 0 0 1 1 ...  
## $ INSR\_BEGIN : chr [1:508499] "08-AUG-17" "08-AUG-16" "08-AUG-15" "08-AUG-14" ...  
## $ INSR\_END : chr [1:508499] "07-AUG-18" "07-AUG-17" "07-AUG-16" "07-AUG-15" ...  
## $ EFFECTIVE\_YR : chr [1:508499] "08" "08" "08" "08" ...  
## $ INSR\_TYPE : num [1:508499] 1202 1202 1202 1202 1202 ...  
## $ INSURED\_VALUE : num [1:508499] 519755 519755 519755 519755 1400000 ...  
## $ PREMIUM : num [1:508499] 5098 6557 6557 5103 13305 ...  
## $ OBJECT\_ID : num [1:508499] 5e+09 5e+09 5e+09 5e+09 5e+09 ...  
## $ PROD\_YEAR : num [1:508499] 2007 2007 2007 2007 2010 ...  
## $ SEATS\_NUM : num [1:508499] 4 4 4 4 4 4 4 4 0 0 ...  
## $ CARRYING\_CAPACITY: chr [1:508499] "6" "6" "6" "6" ...  
## $ TYPE\_VEHICLE : chr [1:508499] "Pick-up" "Pick-up" "Pick-up" "Pick-up" ...  
## $ CCM\_TON : num [1:508499] 3153 3153 3153 3153 2494 ...  
## $ MAKE : chr [1:508499] "NISSAN" "NISSAN" "NISSAN" "NISSAN" ...  
## $ USAGE : chr [1:508499] "Own Goods" "Own Goods" "Own Goods" "Own Goods" ...  
## $ CLAIM\_PAID : num [1:508499] NA NA NA NA NA ...  
## - attr(\*, "spec")=  
## .. cols(  
## .. SEX = col\_double(),  
## .. INSR\_BEGIN = col\_character(),  
## .. INSR\_END = col\_character(),  
## .. EFFECTIVE\_YR = col\_character(),  
## .. INSR\_TYPE = col\_double(),  
## .. INSURED\_VALUE = col\_double(),  
## .. PREMIUM = col\_double(),  
## .. OBJECT\_ID = col\_double(),  
## .. PROD\_YEAR = col\_double(),  
## .. SEATS\_NUM = col\_double(),  
## .. CARRYING\_CAPACITY = col\_character(),  
## .. TYPE\_VEHICLE = col\_character(),  
## .. CCM\_TON = col\_double(),  
## .. MAKE = col\_character(),  
## .. USAGE = col\_character(),  
## .. CLAIM\_PAID = col\_double()  
## .. )  
## - attr(\*, "problems")=<externalptr>

## Anzahl der entfernten Duplikate: 113

##   
## Legal entity Male Female   
## 247026 217734 43626

## Fehlende Werte in INSURED\_VALUE: 0

## Zusammenfassung der statistischen Kennzahlen von INSURED\_VALUE:

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 0 0 150000 563768 850000 67824388

## Anzahl der Einträge mit dem Wert 0 in INSURED\_VALUE: 231296

## Anzahl der Datensätze mit INSURED\_VALUE = 0: 231296

## Verteilung der Versicherungstypen (INSR\_TYPE) bei INSURED\_VALUE = 0:

##   
## Private Commercial Motor trade road risk   
## 123271 384945 170

##   
## Private Commercial Motor trade road risk   
## 45475 185776 45

##   
## Verteilung der Fahrzeugtypen (TYPE\_VEHICLE) bei INSURED\_VALUE = 0:

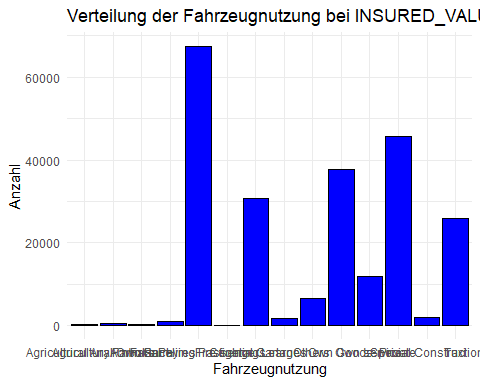
##   
## Automobile Bus Motor-cycle   
## 32740 31066 90720   
## Pick-up Special construction Station Wagones   
## 25818 3090 6757   
## Tanker Tractor Trailers and semitrailers   
## 1872 657 2655   
## Truck   
## 35921

##   
## Verteilung der Fahrzeugnutzung (USAGE) bei INSURED\_VALUE = 0:

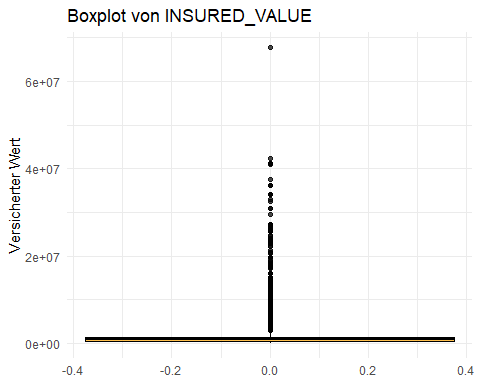
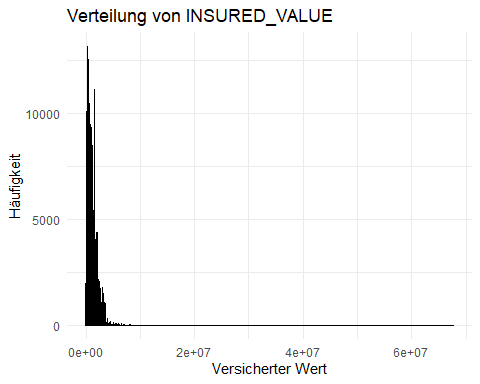
##   
## Agricultural Any Farm Agricultural Own Farm Ambulance   
## 247 475 234   
## Car Hires Fare Paying Passengers Fire fighting   
## 989 67427 7   
## General Cartage Learnes Others   
## 30696 1740 6492   
## Own Goods Own service Private   
## 37619 11836 45675   
## Special Construction Taxi   
## 1907 25952

##   
## Zusammenfassung der Prämien (PREMIUM) bei INSURED\_VALUE = 0:

## Min. 1st Qu. Median Mean 3rd Qu. Max. NA's   
## 0.0 347.7 659.3 1397.2 1830.5 59969.9 5

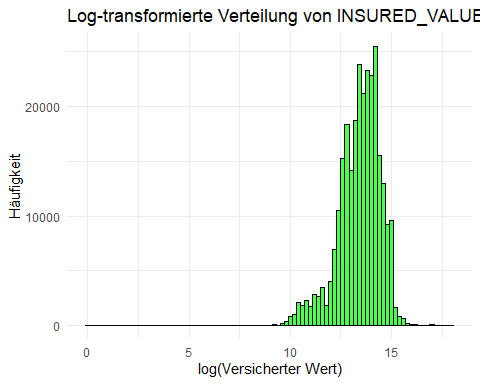


## Anzahl der verbleibenden Datensätze: 277090



## Zusammenfassung der statistischen Kennzahlen ohne 0-Werte:

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 1 400000 780000 1034363 1440000 67824388



## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 1 400000 780000 1034363 1440000 67824388

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 2665 400000 780000 1034408 1440000 67824388

## PREMIUM\_0\_Percent PREMIUM\_NA\_Percent PREMIUM\_MORE\_Percent  
## 1 0.0025 0.0029 99.9975

## Anzahl der entfernten Duplikate: 0

## Die OBJECT\_IDs sind NICHT einmalig.  
## Anzahl der Duplikate: 171346   
## Durchschnittliche Häufigkeit der OBJECT\_ID: 2.621   
## Maximale Häufigkeit der OBJECT\_ID: 12   
## Durchschnittliche Häufigkeit der Kombination (OBJECT\_ID, INSR\_BEGIN, INSR\_END, INSURED\_VALUE, PREMIUM): 1   
## Maximale Häufigkeit der Kombination (OBJECT\_ID, INSR\_BEGIN, INSR\_END, INSURED\_VALUE, PREMIUM): 2

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 1960 2004 2010 2008 2014 2018

## SEATS\_NUM\_0 SEATS\_NUM\_NA SEATS\_NUM\_OTHER  
## 1 27313 13 249687

## SEATS\_NUM\_0\_or\_NA\_Percent SEATS\_NUM\_OTHER\_Percent  
## 1 9.864519 90.13548

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 0.000 1.000 4.000 6.245 4.000 198.000

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 0 1298 2779 3838 4334 19798

## CCM\_TON\_0\_Percent CCM\_TON\_MORE\_Percent  
## 1 13.48653 86.51347

##   
## Automobile Bus Motor-cycle   
## 914 3866 969   
## Pick-up Special construction Station Wagones   
## 1830 1643 800   
## Tanker Tractor Trailers and semitrailers   
## 640 4736 17482   
## Truck   
## 4477

##   
## Automobile Bus Motor-cycle   
## 39262 30630 14246   
## Pick-up Special construction Station Wagones   
## 62490 3480 28420   
## Tanker Tractor Trailers and semitrailers   
## 3924 2118 1403   
## Truck   
## 53665

## Anzahl der entfernten Duplikate: 0

##   
## A   
## 1   
## ABAY   
## 825   
## ACHIVER   
## 2   
## ADDIS GEELY   
## 46   
## ADDIS GELLY   
## 4   
## ADGE   
## 817   
## AEOLUS   
## 76   
## AFRO   
## 1242   
## AIRCARGO MOBILE TRUCK   
## 8   
## ALAMI   
## 7   
## ALFAROMEO   
## 4   
## AMBULANCE   
## 34   
## AMI   
## 2   
## APACHE   
## 6   
## APACHERTE   
## 1   
## ARBE EMERET   
## 3   
## AREB EMERATE   
## 2   
## ARTICULATED DUMP TRUCK   
## 3   
## ASNAKE ENGINERING   
## 2   
## ASNAKE ENGNERING   
## 2   
## ASTRA   
## 953   
## ATOZ   
## 205   
## ATZ   
## 20   
## AU   
## 3   
## AUDI   
## 43   
## AUTO   
## 19   
## AUTOMOBIL   
## 1   
## AUTOMOBILE   
## 6   
## AWASH   
## 66   
## AXION   
## 8   
## BAIC   
## 20   
## BAIC AUTOMOBIL   
## 22   
## BAJAJ   
## 449   
## BAJAJI   
## 2542   
## BARTOLETI   
## 140   
## BAYBEN HIGHBAD   
## 3   
## BAYBEN HIGHBAD TRAILER   
## 2   
## BAYBEN TRUCK HIGHBED   
## 4   
## BEBEN   
## 549   
## BEBEN HIGHBAD   
## 3   
## BEBEN SEMI TRAILER   
## 5   
## BEBIEN TANKER   
## 34   
## BEL TRACTOR   
## 4   
## BELARUS   
## 930   
## BELARUS TRACTOR   
## 260   
## BELL   
## 41   
## BELL TRACTOR   
## 11   
## BEYBEN TRUCK   
## 1   
## BISHEFTU   
## 4   
## BISHOFTU   
## 5064   
## BISHOFTU P/UP D/CAB   
## 10   
## BISHOFTU PICK UP   
## 1   
## BISHOFTU/FAW   
## 136   
## BISHOFTUKAMA   
## 274   
## BJC   
## 2   
## BMB   
## 1   
## BMP SONIC   
## 3   
## BMW   
## 533   
## BMW AUTO   
## 3   
## BOX   
## 56   
## BOXER   
## 437   
## BRIDGE   
## 4   
## BUS   
## 30   
## BYD   
## 57   
## CACCIAMALLI   
## 14   
## CADILLAC   
## 33   
## CALABRASE   
## 807   
## CALABRESE   
## 2018   
## CANEHAULAGE   
## 5   
## CARENZI   
## 30   
## CARGO   
## 2   
## CASE   
## 2   
## CAT   
## 41   
## CAT DOZER   
## 74   
## CATERPILLAR   
## 4   
## CATERPILLAR TRACTOR   
## 6   
## CATO   
## 4   
## CHANA   
## 82   
## CHANGHE   
## 4   
## CHARRY   
## 30   
## CHENGLONG MOTOR OF CHINA   
## 4   
## CHERRY   
## 55   
## CHEVROLET   
## 383   
## CHINA   
## 7   
## CHINA - BBN   
## 5   
## CHINA SELONG   
## 1   
## CHINA SPECIAL POWER TRUCK   
## 36   
## CHINA SPECIAL SEMI TRAILER   
## 100   
## CHINA ZENGIZO   
## 1   
## CITROEN   
## 9   
## CLASS   
## 176   
## CLASS COMBINE   
## 5   
## CO   
## 8   
## CO BUS   
## 8   
## COASTER   
## 3   
## COASTER BUS   
## 3   
## COMBI   
## 6   
## COMPACT YARIS   
## 1   
## CORDES   
## 62   
## CORE DRILLING RING   
## 6   
## CRANE   
## 24   
## CRANE ZUMLIN   
## 1   
## CRANE ZUMLIN 70 TON   
## 2   
## DACIA   
## 107   
## DAEWOO   
## 1609   
## DAF   
## 409   
## DAHATSUN   
## 45   
## DAIHATSU   
## 8   
## DAIHATSU TERIOS   
## 96   
## DAMAS   
## 58   
## DANDO GEATECH 7.5 HYDOLIC TOP ROTATYING   
## 4   
## DATSUN   
## 28   
## DAW BUS   
## 8   
## DAWOO   
## 345   
## DAWWO   
## 1   
## DAYUN   
## 16   
## DEAWOO   
## 1   
## DEAWOO USE   
## 1   
## DEUTZ FAHR   
## 129   
## DFAC   
## 1   
## DFM   
## 3   
## DGOIX   
## 3   
## DHATSU   
## 4   
## DIAHATSU   
## 243   
## DIATSU   
## 1   
## DISCOVERY   
## 10   
## DOCC   
## 7   
## DONFING   
## 323   
## DONG FENG   
## 48   
## DONG FENGSHEN   
## 1   
## DONGFANG   
## 5   
## DONGFENG   
## 4   
## DORSEY   
## 4   
## DOZER   
## 71   
## DSIT   
## 18   
## DUBI   
## 5   
## DUNGFINF   
## 71   
## DUNGFING   
## 32   
## EICHER   
## 72   
## EMGRAND   
## 7   
## EMGTAND   
## 4   
## ENGLAND   
## 2   
## ENGLAND TRACTOR   
## 17   
## ETHIOPIA   
## 13   
## EURO TRUCKER   
## 4   
## EXCAVATOR   
## 33   
## FARID   
## 483   
## FAW   
## 799   
## FAWBELLA   
## 4   
## FENGXING   
## 4   
## FIAT   
## 2238   
## FOED   
## 3   
## FORCE   
## 350   
## FORD   
## 3032   
## FORD CARGO   
## 30   
## FORD CORGO   
## 2   
## FORD TRUCK   
## 1   
## FORKLIFT   
## 9   
## FORLAND   
## 132   
## FORSCHE   
## 4   
## FOTON   
## 224   
## FOTTON   
## 39   
## FPRD   
## 5   
## FRANKON   
## 2   
## FRANKUN   
## 4   
## FRANKUN ET   
## 3   
## FRANKUN IVECO   
## 4   
## G9   
## 4   
## GEELY   
## 1334   
## GEEP   
## 3   
## GELION   
## 727   
## GELYION   
## 4   
## GENLION   
## 2   
## GENLYON   
## 42   
## GENLYONIVECO   
## 4   
## GETZ   
## 1   
## GLEEY   
## 58   
## GMC   
## 102   
## GMS   
## 3   
## GOLZ-PLUS   
## 2   
## GORICA   
## 101   
## GRADER   
## 48   
## GREAT WALL   
## 44   
## H.H   
## 2   
## HAFEI RULY   
## 6   
## HAFREI   
## 4   
## HANVE   
## 5   
## HERO   
## 183   
## HH   
## 10   
## HIGBAN HIGHBAD   
## 3   
## HIGER   
## 77   
## HIGER BUS   
## 23   
## HIGH BED   
## 4   
## HIGH BED TRAILER   
## 31   
## HIGHBED   
## 7   
## HIGHBENCARGOTRAUCK   
## 2   
## HIGHBIN HIGHBAD   
## 3   
## HIGHER   
## 59   
## HILUX   
## 3   
## HINO   
## 414   
## HOLAND CAR   
## 4   
## HONDA   
## 448   
## HONGYAN   
## 6   
## HOVER   
## 20   
## HOWO   
## 121   
## HOYGYAN   
## 8   
## HUANGHA   
## 4   
## HUMMER   
## 4   
## HUYBED   
## 3   
## HYDROLIC   
## 1   
## HYUNDAI   
## 3164   
## HYUNDI GETZ   
## 3   
## ILSBO   
## 240   
## INDOFARMO   
## 2   
## INFINITY   
## 3   
## INTERNATIONAL   
## 46   
## INTERNATIONAL USE   
## 2227   
## ISUSU   
## 8   
## ISUZU   
## 23154   
## ISUZU FVR   
## 4   
## ITALY   
## 2   
## IVECO   
## 10388   
## IVECO/CHINA   
## 21   
## JAC   
## 8   
## JAK   
## 3   
## JCB WORK MAX   
## 4   
## JEEP   
## 22   
## JERMEN   
## 1   
## JIEFANG   
## 195   
## JILI SABA   
## 4   
## JIN BEI   
## 231   
## JMC   
## 24   
## JOHN DEER   
## 644   
## JOHNDEER   
## 22   
## KAINUO   
## 10   
## KAMA   
## 1   
## KAMA MINI TRUCK   
## 13   
## KAMA NINI TRUCK   
## 2   
## KAMAMI   
## 1   
## KAMAZ   
## 9   
## KAMZ   
## 9   
## KAT   
## 24   
## KAT TRACTOR   
## 2   
## KAT TRAILER   
## 2   
## KG   
## 2   
## KIA   
## 360   
## KING LONG   
## 19   
## KM.UAG   
## 3   
## KOMATSU   
## 2   
## KOREA   
## 24   
## KORIA   
## 4   
## KORYA   
## 2   
## KUBOTA   
## 1   
## KYRON   
## 3   
## LADA   
## 19   
## LAND CRUISER   
## 4   
## LAND ROVER   
## 223   
## LAND SCRAPER   
## 1   
## LANDINI   
## 249   
## LANDINI DT125   
## 1   
## LANDROVER   
## 390   
## LANJIAN   
## 28   
## LEXUS   
## 11   
## LIBERR MOBILE CRANE   
## 1   
## LIBERR MOBILECRANE   
## 1   
## LIEBERR MOBILE CRANE   
## 1   
## LIFAN   
## 3261   
## LIFAN 520   
## 4   
## LIFAN AUTOMOBILE   
## 3   
## LISBO   
## 23   
## LITON   
## 15   
## LIZE   
## 1   
## LOADER   
## 38   
## LOADER POERR PLUS   
## 1   
## LOADER POWER PLUS   
## 1   
## LOBADE TRUCK   
## 2   
## LOBED   
## 24   
## LOBED TRAILER   
## 1   
## LOGAN   
## 4   
## LONG BASE TRAILER   
## 1   
## LONG JIANG   
## 7   
## LONGJIANG   
## 29   
## LOW BED   
## 21   
## LOW LOADER SEMI-TRAILERS(AXEL WHEEL 12   
## 1   
## LOWBED   
## 289   
## MACK   
## 402   
## MAHANDRA   
## 209   
## MAHINDRA   
## 156   
## MAMMUT   
## 18   
## MAN   
## 169   
## MARU   
## 712   
## MASIL FERGUSAN   
## 109   
## MASSY FUREGUSON   
## 520   
## MATIZ   
## 6   
## MAZ   
## 33   
## MAZDA   
## 1931   
## ME   
## 4   
## MERCEDES   
## 1308   
## MERCEDICE   
## 3   
## MERCEEDES   
## 1186   
## MERCEEDICE   
## 366   
## MERCHEDES   
## 5   
## MESFIN   
## 8397   
## MESIFIN   
## 4   
## MF5340   
## 4   
## MIE   
## 69   
## MINI BUS   
## 3   
## MISTIBUSH   
## 1   
## MITSUBISHI   
## 8722   
## MITSUBISHI\*   
## 4   
## MIXER   
## 2   
## MOBILE GUARAGE   
## 1   
## MOTOR CYCLE   
## 17   
## MOTOR CYCLE (TWOCYCLE)   
## 14   
## MOTORCYCLE   
## 14   
## MTE   
## 80   
## MUSSO   
## 7   
## NAM   
## 1   
## NAMI   
## 347   
## NATFA   
## 22   
## NEW HOLAND   
## 2   
## NEW HOLLAND   
## 345   
## NIO   
## 6   
## NISAN   
## 7   
## NISSAN   
## 15603   
## NISSAN SUNNY   
## 2   
## NISSAN UD   
## 270   
## NISSAN X-TRIAL   
## 1   
## NISSAN\*   
## 56   
## NIVA   
## 18   
## NKG ENG   
## 2   
## OD   
## 4   
## OHNDEERE   
## 6   
## OPEL   
## 59   
## ORAL   
## 186   
## OTOYOL   
## 52   
## OZSAN TREYLER   
## 1   
## P/UP   
## 14   
## PAGOT   
## 2   
## PEJOT   
## 1   
## PEUGEOT   
## 452   
## PEUGEOT AUTOMOBILE   
## 4   
## PEUGEOUT   
## 2   
## PLATENA   
## 139   
## PORCHE   
## 3   
## PORSCHE   
## 2   
## POWER PLUS   
## 5   
## POWER PLUS DAM   
## 2   
## POWER PLUS DAM TRUCK   
## 1   
## POWER PLUS DAMP   
## 3   
## POWER PLUS DOSER   
## 9   
## POWER PLUS TRUCK   
## 7   
## POWRPLUS TRUCK   
## 2   
## PREGIO   
## 9   
## R425DOHC   
## 1   
## RANDON   
## 77   
## RANGE ROVER   
## 4   
## RANGEROVER   
## 24   
## RAV4   
## 1   
## RAVA   
## 4   
## RED FOX   
## 125   
## RENALT   
## 27   
## RENAULT   
## 1498   
## RENAULT\*   
## 4   
## RENGE ROVER   
## 10   
## RENUALT   
## 3   
## REXTON   
## 9   
## RIG   
## 25   
## RIO LS   
## 10   
## RIO JAMES   
## 5   
## RIO JAMES TRUCK PALLET   
## 2   
## ROLD   
## 3   
## ROLF   
## 2   
## ROLFO   
## 767   
## ROLLER   
## 56   
## ROZA   
## 42   
## S/W   
## 28   
## SAMI   
## 93   
## SANIA   
## 2   
## SANY   
## 5   
## SCANIA   
## 907   
## SCHACMAN   
## 26   
## SCHMITZ   
## 142   
## SCRAPER   
## 1   
## SEDEN   
## 1   
## SEECOME   
## 3   
## SHACMAN   
## 106   
## SHNAY   
## 72   
## SINALIKE   
## 39   
## SINO   
## 1368   
## SINO HOWO   
## 10083   
## SINO TRUCK   
## 9   
## SINOTRUK   
## 72   
## SINOTRUK HOWO   
## 2   
## SKODA   
## 56   
## SKY BUS   
## 20   
## SMART   
## 2   
## SOCOOL   
## 2   
## SONALIKA   
## 128   
## SPAIN   
## 11   
## SPORTAGE   
## 5   
## STAYER   
## 9   
## STEYER   
## 134   
## SUGERCANE TRAILER   
## 93   
## SUNLONG   
## 291   
## SUNLONGBUS   
## 28   
## SUV   
## 17   
## SUZIKE   
## 4   
## SUZUKI   
## 3997   
## SUZUKI GRAND VITARA   
## 4   
## T0Y0TA   
## 109   
## TAIWAN   
## 3   
## TALER   
## 4   
## TALIAN   
## 2   
## TATA   
## 1421   
## TEKEZE   
## 5   
## TERIOS   
## 10   
## TICO   
## 8   
## TOMSON   
## 4   
## TOYATA   
## 2   
## TOYOTA   
## 102796   
## TOYOTA AUTOMOBILE   
## 4   
## TOYOTA YARIS   
## 1   
## TOYOTA 4 RUNNER   
## 1   
## TOYOTA AUTOMOBILE   
## 3   
## TOYOTA COROLLA   
## 5   
## TOYOTA HIACE   
## 2   
## TOYOTA HILUX   
## 4   
## TOYOTA L/C PRADO   
## 3   
## TOYOTA L/CRUISER   
## 4   
## TOYOTA MERCHEDIS   
## 1   
## TOYOTA MINIBUS   
## 1   
## TOYOTA P/UP   
## 3   
## TOYOTA PICK-UP   
## 3   
## TOYOTA PLATZ   
## 2   
## TOYOTA RAV4   
## 12   
## TOYOTA RAVA4   
## 1   
## TOYOTA VANZE   
## 2   
## TOYOTA VITZ   
## 14   
## TOYOTA YARIS   
## 10   
## TOYOTA\*   
## 42   
## TOYOTAA   
## 4   
## TOYTA   
## 3   
## TRACTOR   
## 405   
## TRACTOR 4WD   
## 32   
## TRACTOR BELARUS   
## 12   
## TRACTOR TRAILER   
## 43   
## TRACTOR4WD   
## 4   
## TRAILED TANKER WITH FIRE EXTINGUISHER   
## 4   
## TRAILED TANKKER WITH FIRE FIRE EXTINGUIS   
## 2   
## TRAILER   
## 1562   
## TRAKER   
## 73   
## TRAKKER   
## 468   
## TRUCK   
## 18   
## TURBO   
## 5   
## TURBO BUS   
## 147   
## TVS   
## 1098   
## TVS125   
## 4   
## UAE   
## 6   
## UAI   
## 4   
## URA   
## 2   
## URAL   
## 167   
## URSUS   
## 209   
## URSUS TRACTOR   
## 15   
## URSUS TRACTOR URSUS TRACTOR   
## 2   
## USA   
## 6   
## VALTRA TRACTOR   
## 2   
## VAN TRUCK   
## 8   
## VERCYA   
## 77   
## VERSATILE   
## 157   
## VERSATILE TRACTOR   
## 6   
## VERYCA   
## 2   
## VIBERTI   
## 283   
## VITZ   
## 1457   
## VITZ AUTOMOBILE   
## 3   
## VOLKS WAGON   
## 72   
## VOLKSWAGEN   
## 20   
## VOLKSWAGON   
## 146   
## VOLVO   
## 1241   
## WAFA   
## 2   
## WAZ   
## 7   
## WETER TRUCK STAYER   
## 2   
## WHEEL LOADER   
## 176   
## WINEGEL   
## 16   
## WUCING   
## 4   
## X60   
## 3   
## XERION TARCTOR   
## 1   
## XERION TRACTOR   
## 1   
## YAMAHA   
## 4773   
## YAMHA   
## 4   
## YARIS   
## 9   
## YOUNGMAN   
## 4   
## YOUTOGNMIDBUS   
## 1   
## YOUTONG   
## 6   
## YOUTONG BUS   
## 2   
## YTO   
## 39   
## YTO TRACTOR   
## 6   
## YUTONG   
## 8   
## YVS   
## 14   
## ZALANGE   
## 1   
## ZAMAJ   
## 40   
## ZEPPLIN   
## 67   
## ZILE SHOPAN   
## 5   
## ZNA   
## 15   
## ZOBLE   
## 19   
## ZONGSHEN   
## 91   
## ZONGUSHEN   
## 100   
## ZOOM LION CRANE   
## 22   
## ZORZI   
## 80   
## ZOTYE   
## 40   
## ZOTYE, NOMAD II   
## 4   
## ZOYTE   
## 25   
## ZOYTE, NOMAD II   
## 4   
## ZTLTRUCK   
## 1   
## ZUMLIN CRANE   
## 1   
## ZUNGSHUN   
## 26   
## ZX-TOP   
## 2   
## ZX TOP   
## 1   
## ZZ   
## 39

## `summarise()` has grouped output by 'MAKE'. You can override using the  
## `.groups` argument.

## CLAIM\_PAID\_0 CLAIM\_PAID\_MORE\_THAN\_0  
## 1 179443 28213

## CLAIM\_PAID\_0\_Percent CLAIM\_PAID\_MORE\_THAN\_0\_Percent  
## 1 86.41359 13.58641

## CLAIM\_PAID\_0 CLAIM\_PAID\_MORE\_THAN\_0  
## 1 229930 31417

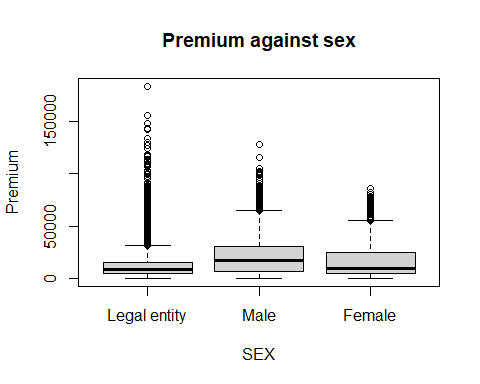
## CLAIM\_PAID\_0\_Percent CLAIM\_PAID\_MORE\_THAN\_0\_Percent  
## 1 87.97882 12.02118

## SEX INSR\_BEGIN INSR\_END INSR\_TYPE INSURED\_VALUE   
## 0 0 0 0 0   
## PREMIUM OBJECT\_ID PROD\_YEAR SEATS\_NUM TYPE\_VEHICLE   
## 0 0 0 0 0   
## CCM\_TON MAKE USAGE CLAIM\_PAID CLAIM\_PAID\_USD   
## 0 0 0 0 0

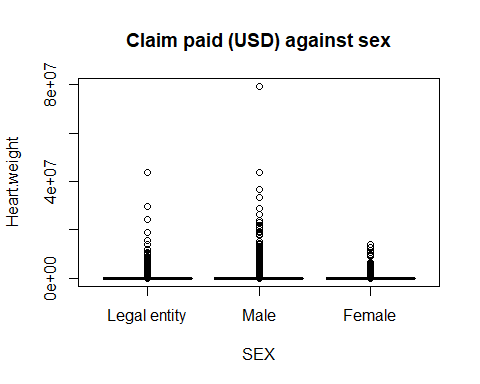
## SEX INSR\_BEGIN INSR\_END INSR\_TYPE INSURED\_VALUE   
## 0 0 0 0 0   
## PREMIUM OBJECT\_ID SEATS\_NUM TYPE\_VEHICLE CCM\_TON   
## 0 0 0 0 0   
## MAKE USAGE CLAIM\_PAID CLAIM\_PAID\_USD AGE\_VEHICLE   
## 0 0 0 0 0

# Graphical Data Analysis

boxplot(PREMIUM ~ SEX, data = clean\_dat\_motor,  
main = "Premium against sex",  
ylab = "Premium")



boxplot(CLAIM\_PAID\_USD ~ SEX, data = clean\_dat\_motor,  
main = "Claim paid (USD) against sex",  
ylab = "Heart.weight") #TODO should be USD?!



# Models

## Linear Model

#Modell erstellen  
lm\_model <- lm(PREMIUM ~ SEX + INSR\_TYPE + MAKE + USAGE + TYPE\_VEHICLE + INSURED\_VALUE + CLAIM\_PAID\_USD + AGE\_VEHICLE + SEATS\_NUM + CCM\_TON, data = clean\_dat\_motor)  
  
#Modellzusammenfassung anzeigen  
summary(lm\_model)

##   
## Call:  
## lm(formula = PREMIUM ~ SEX + INSR\_TYPE + MAKE + USAGE + TYPE\_VEHICLE +   
## INSURED\_VALUE + CLAIM\_PAID\_USD + AGE\_VEHICLE + SEATS\_NUM +   
## CCM\_TON, data = clean\_dat\_motor)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -261717 -2830 -18 2353 120698   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)  
## (Intercept) 9.196e+03 2.791e+02 32.941 < 2e-16  
## SEXMale 3.196e+02 5.652e+01 5.654 1.57e-08  
## SEXFemale 3.820e+02 8.010e+01 4.769 1.86e-06  
## INSR\_TYPEPrivate -1.415e+03 2.146e+02 -6.591 4.37e-11  
## MAKEDAEWOO 3.862e+03 2.115e+02 18.260 < 2e-16  
## MAKEFIAT -1.341e+02 2.167e+02 -0.619 0.53592  
## MAKEFORD 4.908e+02 1.858e+02 2.641 0.00826  
## MAKEGEELY -3.713e+02 2.498e+02 -1.487 0.13708  
## MAKEGENLYON -5.461e+03 3.093e+02 -17.656 < 2e-16  
## MAKEHYUNDAI -1.248e+03 1.850e+02 -6.742 1.56e-11  
## MAKEISUZU 2.554e+03 1.407e+02 18.144 < 2e-16  
## MAKEIVECO -1.653e+03 1.609e+02 -10.274 < 2e-16  
## MAKELIFAN -2.841e+02 1.897e+02 -1.498 0.13426  
## MAKEMAZDA -5.164e+02 2.142e+02 -2.411 0.01591  
## MAKEMERCEDES 1.672e+03 1.882e+02 8.885 < 2e-16  
## MAKEMITSUBISHI 4.401e+02 1.467e+02 3.000 0.00270  
## MAKENISSAN 5.592e+02 1.371e+02 4.079 4.53e-05  
## MAKERENAULT -1.695e+03 2.403e+02 -7.053 1.75e-12  
## MAKESINO -1.606e+03 2.526e+02 -6.360 2.02e-10  
## MAKESINO HOWO -8.153e+01 1.577e+02 -0.517 0.60519  
## MAKESUZUKI 4.997e+00 2.260e+02 0.022 0.98236  
## MAKETATA 3.366e+03 2.342e+02 14.373 < 2e-16  
## MAKETOYOTA 5.786e+02 1.257e+02 4.602 4.18e-06  
## MAKEVOLVO 5.284e+03 2.614e+02 20.214 < 2e-16  
## USAGECar Hires 5.497e+02 4.174e+02 1.317 0.18785  
## USAGEFare Paying Passengers 5.160e+03 3.194e+02 16.158 < 2e-16  
## USAGEGeneral Cartage 4.445e+03 5.289e+02 8.404 < 2e-16  
## USAGEOwn Goods -7.520e+03 5.240e+02 -14.349 < 2e-16  
## USAGEOwn service -2.344e+03 3.043e+02 -7.702 1.35e-14  
## USAGEPrivate -5.909e+03 3.083e+02 -19.170 < 2e-16  
## TYPE\_VEHICLEBus 6.039e+02 2.413e+02 2.503 0.01231  
## TYPE\_VEHICLEMotor-cycle -1.933e+03 2.296e+02 -8.417 < 2e-16  
## TYPE\_VEHICLEPick-up -1.110e+03 4.799e+02 -2.313 0.02070  
## TYPE\_VEHICLESpecial construction -6.023e+02 4.572e+02 -1.317 0.18772  
## TYPE\_VEHICLEStation Wagones 2.554e+03 7.519e+01 33.968 < 2e-16  
## TYPE\_VEHICLETanker 4.072e+03 5.084e+02 8.010 1.15e-15  
## TYPE\_VEHICLETrailers and semitrailers -8.297e+03 6.125e+02 -13.545 < 2e-16  
## TYPE\_VEHICLETruck 2.464e+03 4.906e+02 5.023 5.10e-07  
## INSURED\_VALUE 8.566e-03 3.095e-05 276.755 < 2e-16  
## CLAIM\_PAID\_USD 4.847e-04 3.836e-05 12.636 < 2e-16  
## AGE\_VEHICLE -2.557e+01 3.119e+00 -8.200 2.41e-16  
## SEATS\_NUM -5.269e+01 2.603e+00 -20.240 < 2e-16  
## CCM\_TON -9.857e-02 1.143e-02 -8.624 < 2e-16  
##   
## (Intercept) \*\*\*  
## SEXMale \*\*\*  
## SEXFemale \*\*\*  
## INSR\_TYPEPrivate \*\*\*  
## MAKEDAEWOO \*\*\*  
## MAKEFIAT   
## MAKEFORD \*\*   
## MAKEGEELY   
## MAKEGENLYON \*\*\*  
## MAKEHYUNDAI \*\*\*  
## MAKEISUZU \*\*\*  
## MAKEIVECO \*\*\*  
## MAKELIFAN   
## MAKEMAZDA \*   
## MAKEMERCEDES \*\*\*  
## MAKEMITSUBISHI \*\*   
## MAKENISSAN \*\*\*  
## MAKERENAULT \*\*\*  
## MAKESINO \*\*\*  
## MAKESINO HOWO   
## MAKESUZUKI   
## MAKETATA \*\*\*  
## MAKETOYOTA \*\*\*  
## MAKEVOLVO \*\*\*  
## USAGECar Hires   
## USAGEFare Paying Passengers \*\*\*  
## USAGEGeneral Cartage \*\*\*  
## USAGEOwn Goods \*\*\*  
## USAGEOwn service \*\*\*  
## USAGEPrivate \*\*\*  
## TYPE\_VEHICLEBus \*   
## TYPE\_VEHICLEMotor-cycle \*\*\*  
## TYPE\_VEHICLEPick-up \*   
## TYPE\_VEHICLESpecial construction   
## TYPE\_VEHICLEStation Wagones \*\*\*  
## TYPE\_VEHICLETanker \*\*\*  
## TYPE\_VEHICLETrailers and semitrailers \*\*\*  
## TYPE\_VEHICLETruck \*\*\*  
## INSURED\_VALUE \*\*\*  
## CLAIM\_PAID\_USD \*\*\*  
## AGE\_VEHICLE \*\*\*  
## SEATS\_NUM \*\*\*  
## CCM\_TON \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 7709 on 207613 degrees of freedom  
## Multiple R-squared: 0.6368, Adjusted R-squared: 0.6367   
## F-statistic: 8666 on 42 and 207613 DF, p-value: < 2.2e-16

coef(lm\_model)

## (Intercept) SEXMale   
## 9.195570e+03 3.195761e+02   
## SEXFemale INSR\_TYPEPrivate   
## 3.819654e+02 -1.414642e+03   
## MAKEDAEWOO MAKEFIAT   
## 3.862149e+03 -1.341312e+02   
## MAKEFORD MAKEGEELY   
## 4.907941e+02 -3.713335e+02   
## MAKEGENLYON MAKEHYUNDAI   
## -5.460538e+03 -1.247516e+03   
## MAKEISUZU MAKEIVECO   
## 2.553660e+03 -1.653457e+03   
## MAKELIFAN MAKEMAZDA   
## -2.840871e+02 -5.164460e+02   
## MAKEMERCEDES MAKEMITSUBISHI   
## 1.671880e+03 4.400717e+02   
## MAKENISSAN MAKERENAULT   
## 5.591584e+02 -1.694731e+03   
## MAKESINO MAKESINO HOWO   
## -1.606261e+03 -8.152798e+01   
## MAKESUZUKI MAKETATA   
## 4.996835e+00 3.365612e+03   
## MAKETOYOTA MAKEVOLVO   
## 5.785557e+02 5.284152e+03   
## USAGECar Hires USAGEFare Paying Passengers   
## 5.496642e+02 5.160394e+03   
## USAGEGeneral Cartage USAGEOwn Goods   
## 4.444877e+03 -7.519562e+03   
## USAGEOwn service USAGEPrivate   
## -2.343966e+03 -5.909297e+03   
## TYPE\_VEHICLEBus TYPE\_VEHICLEMotor-cycle   
## 6.039085e+02 -1.932814e+03   
## TYPE\_VEHICLEPick-up TYPE\_VEHICLESpecial construction   
## -1.110197e+03 -6.022632e+02   
## TYPE\_VEHICLEStation Wagones TYPE\_VEHICLETanker   
## 2.554197e+03 4.072001e+03   
## TYPE\_VEHICLETrailers and semitrailers TYPE\_VEHICLETruck   
## -8.297097e+03 2.464055e+03   
## INSURED\_VALUE CLAIM\_PAID\_USD   
## 8.566353e-03 4.847543e-04   
## AGE\_VEHICLE SEATS\_NUM   
## -2.557399e+01 -5.269237e+01   
## CCM\_TON   
## -9.857030e-02

Anova(lm\_model, type="II") # Type II oder III je nach Modellstrukt

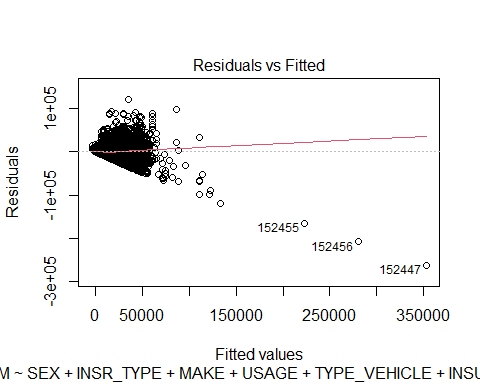
## Warning in printHypothesis(L, rhs, names(b)): one or more coefficients in the hypothesis include  
## arithmetic operators in their names;  
## the printed representation of the hypothesis will be omitted

## Anova Table (Type II tests)  
##   
## Response: PREMIUM  
## Sum Sq Df F value Pr(>F)   
## SEX 2.2923e+09 2 19.287 4.212e-09 \*\*\*  
## INSR\_TYPE 2.5819e+09 1 43.447 4.366e-11 \*\*\*  
## MAKE 2.2574e+11 20 189.935 < 2.2e-16 \*\*\*  
## USAGE 1.2672e+12 6 3553.824 < 2.2e-16 \*\*\*  
## TYPE\_VEHICLE 1.6624e+11 8 349.678 < 2.2e-16 \*\*\*  
## INSURED\_VALUE 4.5517e+12 1 76593.314 < 2.2e-16 \*\*\*  
## CLAIM\_PAID\_USD 9.4883e+09 1 159.663 < 2.2e-16 \*\*\*  
## AGE\_VEHICLE 3.9963e+09 1 67.248 2.408e-16 \*\*\*  
## SEATS\_NUM 2.4344e+10 1 409.647 < 2.2e-16 \*\*\*  
## CCM\_TON 4.4200e+09 1 74.377 < 2.2e-16 \*\*\*  
## Residuals 1.2338e+13 207613   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

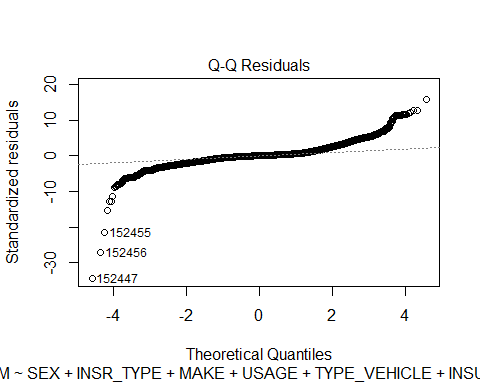
vif(lm\_model)

## GVIF Df GVIF^(1/(2\*Df))  
## SEX 2.327511 2 1.235159  
## INSR\_TYPE 34.168295 1 5.845365  
## MAKE 38.381508 20 1.095476  
## USAGE 29406.998329 6 2.357061  
## TYPE\_VEHICLE 81977.657969 8 2.028178  
## INSURED\_VALUE 2.422496 1 1.556437  
## CLAIM\_PAID\_USD 1.008157 1 1.004070  
## AGE\_VEHICLE 1.965687 1 1.402030  
## SEATS\_NUM 2.870067 1 1.694127  
## CCM\_TON 5.391321 1 2.321922

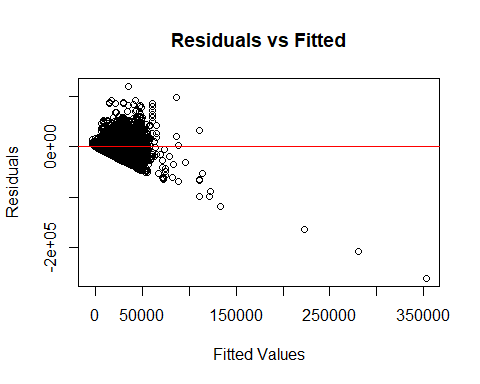
# Residuen plotten  
plot(lm\_model, which=1) # Residuals vs Fitted



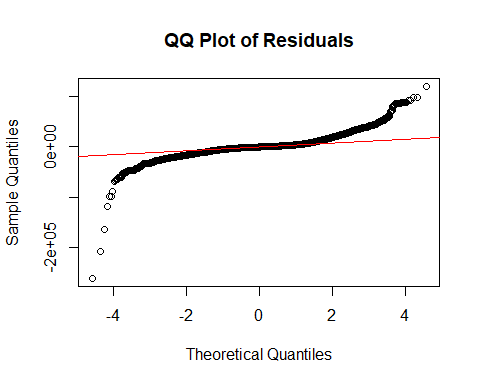
plot(lm\_model, which=2) # Normal Q-Q



#Residuen-Plots für Modell-Diagnose  
#Plot der Residuen vs. Fit-Werte  
plot(lm\_model$fitted.values, lm\_model$residuals,   
 xlab = "Fitted Values",   
 ylab = "Residuals",   
 main = "Residuals vs Fitted")  
abline(h = 0, col = "red")



#QQ-Plot der Residuen (Überprüfung der Normalverteilung)  
qqnorm(lm\_model$residuals, main = "QQ Plot of Residuals")  
qqline(lm\_model$residuals, col = "red")



#Modell-Performance Metriken (z.B. R² und MSE)  
#Berechnung des Mean Squared Error (MSE)  
mse <- mean(lm\_model$residuals^2)  
cat("Mean Squared Error (MSE):", mse, "\n")

## Mean Squared Error (MSE): 59414379

#Berechnung des R² (wird auch in summary(lm\_model) angezeigt)  
r\_squared <- summary(lm\_model)$r.squared  
cat("R-squared:", r\_squared, "\n")

## R-squared: 0.6367643

* R-squared= 0,6368, dh. 63,68% der Varianz in der Zielvariablen PREMIUM wird durch die erklärenden Variablen erklärt. Gut, hat aber Verbesserungspotential
* Alle Prädiktoren sind signifikant Pr(>F) <0,05
* Die Estimate-Werte zeigen die Richtung und Stärke des Effekts der jeweiligen Variable auf die PREMIUM
* INSR\_TYPEPrivate hat tiefere Premium wie commercial
* Make: Die Marke Toyota hat die teuerste Praemie. Die guenstigste Premium hat die Marke HOWO
* Guenstigsten USAGEOwn Goods teuerste Premium Car Hires
* teuerste Premium Type: Bus und guenstigsten Trailers and semitrailers
* Insured value und CLAIM\_PAID\_USD positiver effekt auf premium war zu erwarten

VIF: Die VIF-Werte (Variance Inflation Factor) geben an, wie stark die Multikollinearität in deinem Modell ist. Hohe VIF-Werte (normalerweise über 5 oder 10) deuten darauf hin, dass einige Prädiktoren stark miteinander korrelieren, was die Stabilität und Interpretierbarkeit des Modells beeinträchtigen kann. - INSR\_TYPE (5.85): Ein VIF-Wert von über 5 deutet auf eine mögliche Multikollinearität hin. Es könnte sinnvoll sein, diese Variable genauer zu untersuchen.

Residuen Der Residuals vs Fitted-Plot sollte keine Muster aufweisen (d.h. die Residuen sollten zufällig verteilt sein), und der Normal Q-Q-Plot sollte annähernd eine Gerade bilden, um die Normalverteilung der Residuen zu bestätigen.

Der Residuals vs Fitted Plot zeigt eine leicht gebogene Linie sowie eine Ansammlung von Punkten nahe dem Wert 0, begleitet von einigen extremen Ausreissern (mit ID-Nummern markiert). Dieses Muster deutet darauf hin, dass die Annahme der Homoskedastizität, also der gleichmässigen Varianz der Residuen, in diesem Modell nicht erfüllt ist. In einem idealen Modell sollten die Residuen zufällig und gleichmässig um die horizontale Linie bei 0 verteilt sein. Darüber hinaus sind einige Residuen-Punkte deutlich vom Hauptcluster entfernt. Diese Ausreisser könnten extreme Werte darstellen, die das Modell möglicherweise verzerren. Eine genauere Untersuchung dieser Punkte ist empfehlenswert, um zu entscheiden, ob sie aus dem Modell entfernt oder separat behandelt werden sollten. Bei höheren Prämienwerten zeigen die Residuen zudem eine ungleichmässige Verteilung, was darauf hindeutet, dass das Modell für diese höheren Werte keine ausreichende Passgenauigkeit aufweist.

Im Normal Q-Q Plot lässt sich erkennen, dass die Residuen nicht vollständig auf der Linie liegen, die eine perfekte Normalverteilung darstellt. Dies weist darauf hin, dass die Annahme der Normalverteilung der Residuen verletzt ist. Besonders auffällig sind die deutlichen Abweichungen sowohl am unteren als auch am oberen Ende des Plots. Dies deutet auf das Vorhandensein von Ausreissern hin und darauf, dass die Verteilung der Residuen „schwerere Enden“ aufweist, d. h. mehr extreme Werte enthält als in einer Normalverteilung zu erwarten wäre. Die Punkte am Rand des Plots repräsentieren extreme Werte, die signifikant von der angenommenen Normalverteilung abweichen.

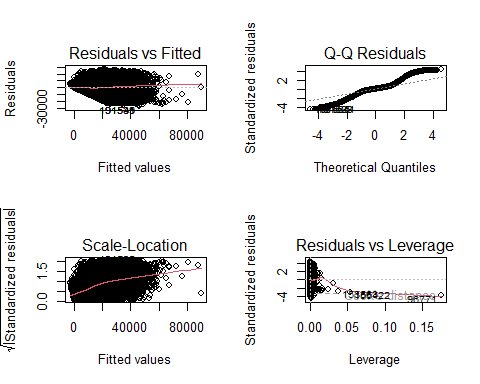
Massnahmen:Transformation der Zielvariablen, Ausreisserbehandlung

### Massnahme 1): Ausreisserbehandlung

# Standardisierte Residuen berechnen  
standardized\_resid <- rstandard(lm\_model)  
  
# Schwellenwert für Ausreisser setzen (z.B. Residuen grösser als 3 oder kleiner als -3)  
threshold <- 3  
  
# Indizes der Ausreisser identifizieren  
outliers <- which(abs(standardized\_resid) > threshold)  
  
# Daten ohne Ausreisser  
clean\_data\_no\_outliers <- clean\_dat\_motor[-outliers, ]  
  
# Neues Modell ohne Ausreisser anpassen  
lm\_model\_no\_outliers <- lm(PREMIUM ~ SEX + INSR\_TYPE + MAKE + USAGE + TYPE\_VEHICLE + INSURED\_VALUE + CLAIM\_PAID\_USD + AGE\_VEHICLE + SEATS\_NUM + CCM\_TON, data = clean\_data\_no\_outliers)  
  
# Zusammenfassung des neuen Modells anzeigen  
summary(lm\_model\_no\_outliers)

##   
## Call:  
## lm(formula = PREMIUM ~ SEX + INSR\_TYPE + MAKE + USAGE + TYPE\_VEHICLE +   
## INSURED\_VALUE + CLAIM\_PAID\_USD + AGE\_VEHICLE + SEATS\_NUM +   
## CCM\_TON, data = clean\_data\_no\_outliers)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -27187.2 -2359.7 172.4 2324.6 26329.5   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)  
## (Intercept) 7.788e+03 2.257e+02 34.501 < 2e-16  
## SEXMale 3.762e+01 4.570e+01 0.823 0.41029  
## SEXFemale 1.748e+02 6.489e+01 2.694 0.00707  
## INSR\_TYPEPrivate -1.186e+03 1.724e+02 -6.879 6.06e-12  
## MAKEDAEWOO 1.639e+03 1.736e+02 9.440 < 2e-16  
## MAKEFIAT -1.151e+03 1.749e+02 -6.580 4.73e-11  
## MAKEFORD 3.916e+02 1.491e+02 2.627 0.00860  
## MAKEGEELY 1.230e+02 2.002e+02 0.614 0.53900  
## MAKEGENLYON -5.062e+03 2.509e+02 -20.178 < 2e-16  
## MAKEHYUNDAI -1.084e+03 1.486e+02 -7.300 2.89e-13  
## MAKEISUZU 1.429e+03 1.143e+02 12.505 < 2e-16  
## MAKEIVECO -2.333e+03 1.306e+02 -17.867 < 2e-16  
## MAKELIFAN 1.848e+02 1.522e+02 1.214 0.22462  
## MAKEMAZDA -7.724e+02 1.717e+02 -4.498 6.85e-06  
## MAKEMERCEDES -9.775e+01 1.532e+02 -0.638 0.52352  
## MAKEMITSUBISHI 3.167e+02 1.179e+02 2.685 0.00725  
## MAKENISSAN 2.925e+02 1.103e+02 2.652 0.00801  
## MAKERENAULT -2.050e+03 1.930e+02 -10.620 < 2e-16  
## MAKESINO -1.168e+03 2.035e+02 -5.737 9.65e-09  
## MAKESINO HOWO 5.130e+02 1.277e+02 4.016 5.92e-05  
## MAKESUZUKI -8.431e+01 1.811e+02 -0.465 0.64160  
## MAKETATA 1.121e+03 1.951e+02 5.745 9.18e-09  
## MAKETOYOTA 2.946e+02 1.011e+02 2.913 0.00358  
## MAKEVOLVO 4.130e+03 2.103e+02 19.635 < 2e-16  
## USAGECar Hires -7.268e+02 3.445e+02 -2.110 0.03487  
## USAGEFare Paying Passengers 5.507e+03 2.582e+02 21.328 < 2e-16  
## USAGEGeneral Cartage 3.200e+03 4.295e+02 7.452 9.25e-14  
## USAGEOwn Goods -7.739e+03 4.254e+02 -18.193 < 2e-16  
## USAGEOwn service -2.589e+03 2.458e+02 -10.533 < 2e-16  
## USAGEPrivate -6.066e+03 2.488e+02 -24.381 < 2e-16  
## TYPE\_VEHICLEBus -8.161e+02 1.953e+02 -4.179 2.92e-05  
## TYPE\_VEHICLEMotor-cycle -1.245e+03 1.839e+02 -6.771 1.29e-11  
## TYPE\_VEHICLEPick-up -9.936e+02 3.897e+02 -2.549 0.01079  
## TYPE\_VEHICLESpecial construction -1.456e+03 3.683e+02 -3.953 7.71e-05  
## TYPE\_VEHICLEStation Wagones 1.728e+03 6.068e+01 28.468 < 2e-16  
## TYPE\_VEHICLETanker 1.827e+03 4.142e+02 4.410 1.03e-05  
## TYPE\_VEHICLETrailers and semitrailers -6.816e+03 4.950e+02 -13.770 < 2e-16  
## TYPE\_VEHICLETruck 1.984e+03 3.988e+02 4.974 6.57e-07  
## INSURED\_VALUE 9.241e-03 2.649e-05 348.835 < 2e-16  
## CLAIM\_PAID\_USD 5.073e-04 3.264e-05 15.540 < 2e-16  
## AGE\_VEHICLE 3.369e+01 2.552e+00 13.199 < 2e-16  
## SEATS\_NUM 2.966e+00 2.405e+00 1.233 0.21750  
## CCM\_TON 1.641e-02 9.411e-03 1.743 0.08127  
##   
## (Intercept) \*\*\*  
## SEXMale   
## SEXFemale \*\*   
## INSR\_TYPEPrivate \*\*\*  
## MAKEDAEWOO \*\*\*  
## MAKEFIAT \*\*\*  
## MAKEFORD \*\*   
## MAKEGEELY   
## MAKEGENLYON \*\*\*  
## MAKEHYUNDAI \*\*\*  
## MAKEISUZU \*\*\*  
## MAKEIVECO \*\*\*  
## MAKELIFAN   
## MAKEMAZDA \*\*\*  
## MAKEMERCEDES   
## MAKEMITSUBISHI \*\*   
## MAKENISSAN \*\*   
## MAKERENAULT \*\*\*  
## MAKESINO \*\*\*  
## MAKESINO HOWO \*\*\*  
## MAKESUZUKI   
## MAKETATA \*\*\*  
## MAKETOYOTA \*\*   
## MAKEVOLVO \*\*\*  
## USAGECar Hires \*   
## USAGEFare Paying Passengers \*\*\*  
## USAGEGeneral Cartage \*\*\*  
## USAGEOwn Goods \*\*\*  
## USAGEOwn service \*\*\*  
## USAGEPrivate \*\*\*  
## TYPE\_VEHICLEBus \*\*\*  
## TYPE\_VEHICLEMotor-cycle \*\*\*  
## TYPE\_VEHICLEPick-up \*   
## TYPE\_VEHICLESpecial construction \*\*\*  
## TYPE\_VEHICLEStation Wagones \*\*\*  
## TYPE\_VEHICLETanker \*\*\*  
## TYPE\_VEHICLETrailers and semitrailers \*\*\*  
## TYPE\_VEHICLETruck \*\*\*  
## INSURED\_VALUE \*\*\*  
## CLAIM\_PAID\_USD \*\*\*  
## AGE\_VEHICLE \*\*\*  
## SEATS\_NUM   
## CCM\_TON .   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 6168 on 203479 degrees of freedom  
## Multiple R-squared: 0.7175, Adjusted R-squared: 0.7175   
## F-statistic: 1.231e+04 on 42 and 203479 DF, p-value: < 2.2e-16

# Residuenplots für das neue Modell erstellen  
par(mfrow = c(2, 2)) # Mehrere Plots auf einer Seite  
plot(lm\_model\_no\_outliers)

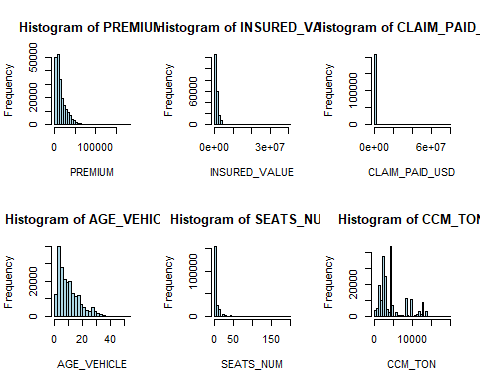


Nach dem Entfernen der Ausreisser zeigen die Residuenplots zwar eine Verbesserung, jedoch bleiben einige Auffälligkeiten bestehen. Der Residuals vs Fitted Plot weist weiterhin eine Trichterform auf, was auf eine Verletzung der Homoskedastizitätsannahme hindeutet. Dies deutet darauf hin, dass die Varianz der Residuen nicht konstant ist. Im Normal Q-Q Plot sind an den Enden deutliche Abweichungen von der theoretischen Normalverteilung zu erkennen, was auf verbleibende Ausreisser oder eine unzureichende Anpassung an die Normalverteilungsannahme schliessen lässt. Der Scale-Location Plot bestätigt ebenfalls eine ungleichmässige Streuung der Residuen, insbesondere bei höheren fitted values. Insgesamt legen diese Ergebnisse nahe, dass eine Transformation der Zielvariablen, wie eine Log-Transformation, sinnvoll sein könnte, um die Modellgüte weiter zu verbessern und die Annahmen der linearen Regression besser zu erfüllen.

Die Modellgüte hat sich verbessert: Der Residual Standard Error (RSE) verringerte sich von über 7700 auf 6168, und das R-squared stieg von 0,6368 auf 0,7175. Dies bedeutet, dass nun 71,75% der Varianz in der Prämie durch die erklärenden Variablen erklärt werden. Die meisten Variablen, wie INSURED\_VALUE und CLAIM\_PAID\_USD, bleiben weiterhin signifikant.

### Massnahme 2): Transformation

#Analyse Variablen  
numeric\_vars <- c("PREMIUM", "INSURED\_VALUE", "CLAIM\_PAID\_USD", "AGE\_VEHICLE", "SEATS\_NUM", "CCM\_TON")  
  
#Histogramme erstellen  
par(mfrow = c(2, 3)) # 2 Zeilen, 3 Spalten für die Plots  
for (var in numeric\_vars) {  
 hist(clean\_dat\_motor[[var]], main = paste("Histogram of", var), xlab = var, col = "lightblue", breaks = 30)  
}



par(mfrow = c(1, 1)) # Zurück zu einem einzelnen Plot pro Seite  
  
#Massnahme: Transformation der Variablen  
# Log-Transformation für stark rechtsschiefe Variablen  
clean\_dat\_motor$PREMIUM\_LOG <- log(clean\_dat\_motor$PREMIUM + 1) # Hinzufügen von +1 um log(0) zu vermeiden  
clean\_dat\_motor$INSURED\_VALUE\_LOG <- log(clean\_dat\_motor$INSURED\_VALUE + 1)  
clean\_dat\_motor$CLAIM\_PAID\_USD\_LOG <- log(clean\_dat\_motor$CLAIM\_PAID\_USD + 1)  
clean\_dat\_motor$CCM\_TON\_LOG <- log(clean\_dat\_motor$CCM\_TON + 1)

Die Histogramme der Variablen zeigen, dass einige von ihnen, insbesondere PREMIUM, INSURED\_VALUE, CLAIM\_PAID\_USD und CCM\_TON, stark rechtsschief verteilt sind. Diese Verteilungen deuten darauf hin, dass eine log-Transformation sinnvoll sein könnte, um die Daten zu normalisieren und die Varianz zu stabilisieren. Durch die Transformation würde die Schiefe verringert, was zu einer besseren Modellanpassung führen könnte.

Variablen wie AGE\_VEHICLE und SEATS\_NUM zeigen hingegen eine weniger ausgeprägte Schiefe, sodass eine Transformation hier weniger notwendig erscheint. Eine log-Transformation der stark schiefen Variablen wird daher empfohlen, um die Modellgüte weiter zu verbessern.

# Zunächst prüfen, ob es 0-Werte gibt, die das Logarithmieren verhindern  
# Kleine Konstante (1) hinzufügen, um mit möglichen 0-Werten in den Variablen umzugehen  
clean\_data\_no\_outliers$LOG\_PREMIUM <- log(clean\_data\_no\_outliers$PREMIUM + 1)  
clean\_data\_no\_outliers$LOG\_INSURED\_VALUE <- log(clean\_data\_no\_outliers$INSURED\_VALUE + 1)  
clean\_data\_no\_outliers$LOG\_CLAIM\_PAID\_USD <- log(clean\_data\_no\_outliers$CLAIM\_PAID\_USD + 1)  
clean\_data\_no\_outliers$LOG\_CCM\_TON <- log(clean\_data\_no\_outliers$CCM\_TON + 1)  
  
# Neues Modell mit den logarithmierten Variablen  
lm\_model\_log <- lm(LOG\_PREMIUM ~ SEX + INSR\_TYPE + MAKE + USAGE + TYPE\_VEHICLE +   
 LOG\_INSURED\_VALUE + LOG\_CLAIM\_PAID\_USD + AGE\_VEHICLE + SEATS\_NUM + LOG\_CCM\_TON,   
 data = clean\_data\_no\_outliers)  
  
# Zusammenfassung des neuen Modells anzeigen  
summary(lm\_model\_log)

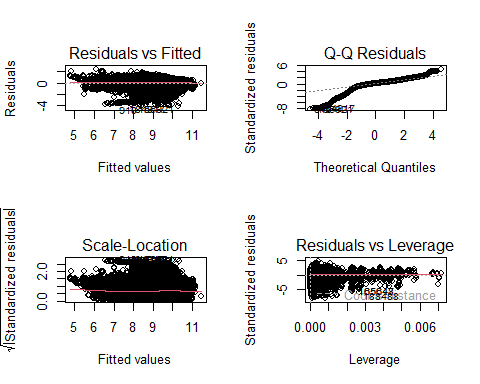
##   
## Call:  
## lm(formula = LOG\_PREMIUM ~ SEX + INSR\_TYPE + MAKE + USAGE + TYPE\_VEHICLE +   
## LOG\_INSURED\_VALUE + LOG\_CLAIM\_PAID\_USD + AGE\_VEHICLE + SEATS\_NUM +   
## LOG\_CCM\_TON, data = clean\_data\_no\_outliers)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -4.1595 -0.1433 0.0826 0.2783 2.3608   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)  
## (Intercept) -1.5507286 0.0404780 -38.310 < 2e-16  
## SEXMale -0.0777382 0.0038301 -20.297 < 2e-16  
## SEXFemale -0.0754770 0.0054269 -13.908 < 2e-16  
## INSR\_TYPEPrivate 0.0544011 0.0143413 3.793 0.000149  
## MAKEDAEWOO 0.1291477 0.0144484 8.939 < 2e-16  
## MAKEFIAT 0.0227382 0.0145176 1.566 0.117291  
## MAKEFORD 0.0460143 0.0124035 3.710 0.000207  
## MAKEGEELY 0.0639265 0.0166982 3.828 0.000129  
## MAKEGENLYON -0.2157021 0.0208925 -10.324 < 2e-16  
## MAKEHYUNDAI 0.0195631 0.0123573 1.583 0.113395  
## MAKEISUZU 0.1681810 0.0094500 17.797 < 2e-16  
## MAKEIVECO -0.1103374 0.0107017 -10.310 < 2e-16  
## MAKELIFAN 0.1147476 0.0127153 9.024 < 2e-16  
## MAKEMAZDA 0.0300363 0.0143014 2.100 0.035709  
## MAKEMERCEDES 0.0322226 0.0127848 2.520 0.011723  
## MAKEMITSUBISHI 0.0414783 0.0098127 4.227 2.37e-05  
## MAKENISSAN 0.0931459 0.0092010 10.123 < 2e-16  
## MAKERENAULT -0.0991103 0.0159972 -6.195 5.82e-10  
## MAKESINO 0.0052687 0.0169332 0.311 0.755692  
## MAKESINO HOWO -0.0237596 0.0106582 -2.229 0.025800  
## MAKESUZUKI 0.0264112 0.0150741 1.752 0.079758  
## MAKETATA 0.0033828 0.0162316 0.208 0.834908  
## MAKETOYOTA 0.0600979 0.0084271 7.131 9.96e-13  
## MAKEVOLVO 0.2559883 0.0174842 14.641 < 2e-16  
## USAGECar Hires -0.1108774 0.0286677 -3.868 0.000110  
## USAGEFare Paying Passengers 0.4824817 0.0214852 22.457 < 2e-16  
## USAGEGeneral Cartage 0.0909504 0.0357339 2.545 0.010922  
## USAGEOwn Goods -0.5445243 0.0354040 -15.380 < 2e-16  
## USAGEOwn service -0.0886068 0.0204484 -4.333 1.47e-05  
## USAGEPrivate -0.4766147 0.0207043 -23.020 < 2e-16  
## TYPE\_VEHICLEBus 0.0848840 0.0162907 5.211 1.88e-07  
## TYPE\_VEHICLEMotor-cycle 0.0389572 0.0168846 2.307 0.021041  
## TYPE\_VEHICLEPick-up 0.0404827 0.0324690 1.247 0.212468  
## TYPE\_VEHICLESpecial construction -0.0064395 0.0307428 -0.209 0.834086  
## TYPE\_VEHICLEStation Wagones 0.1840035 0.0054475 33.778 < 2e-16  
## TYPE\_VEHICLETanker 0.3502200 0.0343543 10.194 < 2e-16  
## TYPE\_VEHICLETrailers and semitrailers -0.1260780 0.0468171 -2.693 0.007082  
## TYPE\_VEHICLETruck 0.3112175 0.0331343 9.393 < 2e-16  
## LOG\_INSURED\_VALUE 0.8004438 0.0024869 321.865 < 2e-16  
## LOG\_CLAIM\_PAID\_USD 0.0127854 0.0003052 41.892 < 2e-16  
## AGE\_VEHICLE 0.0103368 0.0002413 42.842 < 2e-16  
## SEATS\_NUM -0.0008042 0.0001971 -4.079 4.52e-05  
## LOG\_CCM\_TON -0.0128429 0.0031585 -4.066 4.78e-05  
##   
## (Intercept) \*\*\*  
## SEXMale \*\*\*  
## SEXFemale \*\*\*  
## INSR\_TYPEPrivate \*\*\*  
## MAKEDAEWOO \*\*\*  
## MAKEFIAT   
## MAKEFORD \*\*\*  
## MAKEGEELY \*\*\*  
## MAKEGENLYON \*\*\*  
## MAKEHYUNDAI   
## MAKEISUZU \*\*\*  
## MAKEIVECO \*\*\*  
## MAKELIFAN \*\*\*  
## MAKEMAZDA \*   
## MAKEMERCEDES \*   
## MAKEMITSUBISHI \*\*\*  
## MAKENISSAN \*\*\*  
## MAKERENAULT \*\*\*  
## MAKESINO   
## MAKESINO HOWO \*   
## MAKESUZUKI .   
## MAKETATA   
## MAKETOYOTA \*\*\*  
## MAKEVOLVO \*\*\*  
## USAGECar Hires \*\*\*  
## USAGEFare Paying Passengers \*\*\*  
## USAGEGeneral Cartage \*   
## USAGEOwn Goods \*\*\*  
## USAGEOwn service \*\*\*  
## USAGEPrivate \*\*\*  
## TYPE\_VEHICLEBus \*\*\*  
## TYPE\_VEHICLEMotor-cycle \*   
## TYPE\_VEHICLEPick-up   
## TYPE\_VEHICLESpecial construction   
## TYPE\_VEHICLEStation Wagones \*\*\*  
## TYPE\_VEHICLETanker \*\*\*  
## TYPE\_VEHICLETrailers and semitrailers \*\*   
## TYPE\_VEHICLETruck \*\*\*  
## LOG\_INSURED\_VALUE \*\*\*  
## LOG\_CLAIM\_PAID\_USD \*\*\*  
## AGE\_VEHICLE \*\*\*  
## SEATS\_NUM \*\*\*  
## LOG\_CCM\_TON \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.5132 on 203479 degrees of freedom  
## Multiple R-squared: 0.7108, Adjusted R-squared: 0.7107   
## F-statistic: 1.191e+04 on 42 and 203479 DF, p-value: < 2.2e-16

Anova(lm\_model\_log, type="II") # Type II oder III je nach Modellstrukt

## Warning in printHypothesis(L, rhs, names(b)): one or more coefficients in the hypothesis include  
## arithmetic operators in their names;  
## the printed representation of the hypothesis will be omitted

## Anova Table (Type II tests)  
##   
## Response: LOG\_PREMIUM  
## Sum Sq Df F value Pr(>F)   
## SEX 117 2 221.970 < 2.2e-16 \*\*\*  
## INSR\_TYPE 4 1 14.389 0.0001487 \*\*\*  
## MAKE 732 20 138.986 < 2.2e-16 \*\*\*  
## USAGE 4083 6 2583.851 < 2.2e-16 \*\*\*  
## TYPE\_VEHICLE 742 8 351.968 < 2.2e-16 \*\*\*  
## LOG\_INSURED\_VALUE 27283 1 103596.916 < 2.2e-16 \*\*\*  
## LOG\_CLAIM\_PAID\_USD 462 1 1754.963 < 2.2e-16 \*\*\*  
## AGE\_VEHICLE 483 1 1835.403 < 2.2e-16 \*\*\*  
## SEATS\_NUM 4 1 16.641 4.518e-05 \*\*\*  
## LOG\_CCM\_TON 4 1 16.534 4.781e-05 \*\*\*  
## Residuals 53587 203479   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

# Residuenplots für das neue Modell erstellen  
par(mfrow = c(2, 2)) # Mehrere Plots auf einer Seite  
plot(lm\_model\_log)



#Modell-Performance Metriken (z.B. R² und MSE)  
#Berechnung des Mean Squared Error (MSE)  
mse <- mean(lm\_model\_log$residuals^2)  
cat("Mean Squared Error (MSE):", mse, "\n")

## Mean Squared Error (MSE): 0.2633004

#Berechnung des R² (wird auch in summary(lm\_model\_log) angezeigt)  
r\_squared <- summary(lm\_model\_log)$r.squared  
cat("R-squared:", r\_squared, "\n")

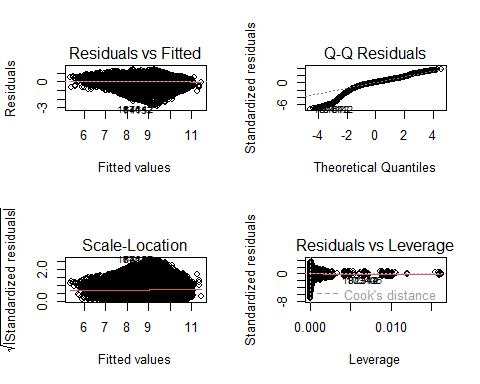
## R-squared: 0.710784

### Massnahme 3): Ausreisser entfernen mit cooks distance

# Berechnung von Cook's Distance für das Modell  
cooksd <- cooks.distance(lm\_model\_log)  
  
# Setze einen Schwellenwert, um Ausreisser zu identifizieren (z.B. 4/n, wobei n die Anzahl der Datenpunkte ist)  
threshold <- 4 / nrow(clean\_data\_no\_outliers)  
  
# Identifiziere potenziell einflussreiche Punkte (die Cook's Distance über dem Schwellenwert haben)  
influential <- as.numeric(names(cooksd)[(cooksd > threshold)])  
  
# Daten ohne diese einflussreichen Punkte  
clean\_data\_no\_outliers\_influential\_removed <- clean\_data\_no\_outliers[-influential, ]  
  
# Erstelle das Modell erneut ohne die einflussreichen Ausreisser  
lm\_model\_no\_influential <- lm(LOG\_PREMIUM ~ SEX + INSR\_TYPE + MAKE + USAGE + TYPE\_VEHICLE +   
 LOG\_INSURED\_VALUE + LOG\_CLAIM\_PAID\_USD + AGE\_VEHICLE + SEATS\_NUM + LOG\_CCM\_TON,   
 data = clean\_data\_no\_outliers\_influential\_removed)  
  
# Zusammenfassung des neuen Modells  
summary(lm\_model\_no\_influential)

##   
## Call:  
## lm(formula = LOG\_PREMIUM ~ SEX + INSR\_TYPE + MAKE + USAGE + TYPE\_VEHICLE +   
## LOG\_INSURED\_VALUE + LOG\_CLAIM\_PAID\_USD + AGE\_VEHICLE + SEATS\_NUM +   
## LOG\_CCM\_TON, data = clean\_data\_no\_outliers\_influential\_removed)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -2.8053 -0.1657 0.0442 0.2345 1.3917   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)  
## (Intercept) -1.5441527 0.0329472 -46.867 < 2e-16  
## SEXMale -0.0449710 0.0030289 -14.848 < 2e-16  
## SEXFemale -0.0088517 0.0043110 -2.053 0.040047  
## INSR\_TYPEPrivate 0.0175625 0.0135802 1.293 0.195929  
## MAKEDAEWOO 0.0559122 0.0116572 4.796 1.62e-06  
## MAKEFIAT -0.0429623 0.0118031 -3.640 0.000273  
## MAKEFORD 0.0414327 0.0098865 4.191 2.78e-05  
## MAKEGEELY 0.0518550 0.0133238 3.892 9.95e-05  
## MAKEGENLYON -0.2216161 0.0185070 -11.975 < 2e-16  
## MAKEHYUNDAI -0.0447319 0.0097398 -4.593 4.38e-06  
## MAKEISUZU 0.0602176 0.0075108 8.017 1.09e-15  
## MAKEIVECO -0.1713086 0.0084934 -20.170 < 2e-16  
## MAKELIFAN 0.0617431 0.0100582 6.139 8.34e-10  
## MAKEMAZDA -0.0154240 0.0112691 -1.369 0.171097  
## MAKEMERCEDES -0.0199245 0.0103209 -1.931 0.053546  
## MAKEMITSUBISHI -0.0050344 0.0077507 -0.650 0.515994  
## MAKENISSAN 0.0362261 0.0072856 4.972 6.62e-07  
## MAKERENAULT -0.1699752 0.0129577 -13.118 < 2e-16  
## MAKESINO -0.0374221 0.0141801 -2.639 0.008315  
## MAKESINO HOWO -0.0268106 0.0085299 -3.143 0.001672  
## MAKESUZUKI -0.0260147 0.0123060 -2.114 0.034518  
## MAKETATA 0.0311377 0.0138329 2.251 0.024388  
## MAKETOYOTA 0.0062055 0.0066934 0.927 0.353868  
## MAKEVOLVO 0.1932949 0.0140567 13.751 < 2e-16  
## USAGECar Hires -0.1583097 0.0391685 -4.042 5.31e-05  
## USAGEFare Paying Passengers 0.5786295 0.0185353 31.218 < 2e-16  
## USAGEGeneral Cartage 0.0129865 0.0406088 0.320 0.749123  
## USAGEOwn Goods -0.5605961 0.0404173 -13.870 < 2e-16  
## USAGEOwn service -0.0928364 0.0177593 -5.227 1.72e-07  
## USAGEPrivate -0.4808035 0.0183065 -26.264 < 2e-16  
## TYPE\_VEHICLEBus 0.0480948 0.0144820 3.321 0.000897  
## TYPE\_VEHICLEMotor-cycle 0.0097493 0.0140213 0.695 0.486856  
## TYPE\_VEHICLEPick-up -0.0172244 0.0386931 -0.445 0.656209  
## TYPE\_VEHICLESpecial construction -0.1302352 0.0327519 -3.976 7.00e-05  
## TYPE\_VEHICLEStation Wagones 0.1624936 0.0043265 37.558 < 2e-16  
## TYPE\_VEHICLETanker 0.3304403 0.0397450 8.314 < 2e-16  
## TYPE\_VEHICLETrailers and semitrailers -0.2947348 0.0508822 -5.792 6.95e-09  
## TYPE\_VEHICLETruck 0.3168220 0.0390910 8.105 5.32e-16  
## LOG\_INSURED\_VALUE 0.8183431 0.0019979 409.593 < 2e-16  
## LOG\_CLAIM\_PAID\_USD 0.0094368 0.0002368 39.851 < 2e-16  
## AGE\_VEHICLE 0.0123493 0.0001937 63.770 < 2e-16  
## SEATS\_NUM -0.0015769 0.0001740 -9.064 < 2e-16  
## LOG\_CCM\_TON -0.0273860 0.0026878 -10.189 < 2e-16  
##   
## (Intercept) \*\*\*  
## SEXMale \*\*\*  
## SEXFemale \*   
## INSR\_TYPEPrivate   
## MAKEDAEWOO \*\*\*  
## MAKEFIAT \*\*\*  
## MAKEFORD \*\*\*  
## MAKEGEELY \*\*\*  
## MAKEGENLYON \*\*\*  
## MAKEHYUNDAI \*\*\*  
## MAKEISUZU \*\*\*  
## MAKEIVECO \*\*\*  
## MAKELIFAN \*\*\*  
## MAKEMAZDA   
## MAKEMERCEDES .   
## MAKEMITSUBISHI   
## MAKENISSAN \*\*\*  
## MAKERENAULT \*\*\*  
## MAKESINO \*\*   
## MAKESINO HOWO \*\*   
## MAKESUZUKI \*   
## MAKETATA \*   
## MAKETOYOTA   
## MAKEVOLVO \*\*\*  
## USAGECar Hires \*\*\*  
## USAGEFare Paying Passengers \*\*\*  
## USAGEGeneral Cartage   
## USAGEOwn Goods \*\*\*  
## USAGEOwn service \*\*\*  
## USAGEPrivate \*\*\*  
## TYPE\_VEHICLEBus \*\*\*  
## TYPE\_VEHICLEMotor-cycle   
## TYPE\_VEHICLEPick-up   
## TYPE\_VEHICLESpecial construction \*\*\*  
## TYPE\_VEHICLEStation Wagones \*\*\*  
## TYPE\_VEHICLETanker \*\*\*  
## TYPE\_VEHICLETrailers and semitrailers \*\*\*  
## TYPE\_VEHICLETruck \*\*\*  
## LOG\_INSURED\_VALUE \*\*\*  
## LOG\_CLAIM\_PAID\_USD \*\*\*  
## AGE\_VEHICLE \*\*\*  
## SEATS\_NUM \*\*\*  
## LOG\_CCM\_TON \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.3933 on 194348 degrees of freedom  
## Multiple R-squared: 0.8028, Adjusted R-squared: 0.8027   
## F-statistic: 1.883e+04 on 42 and 194348 DF, p-value: < 2.2e-16

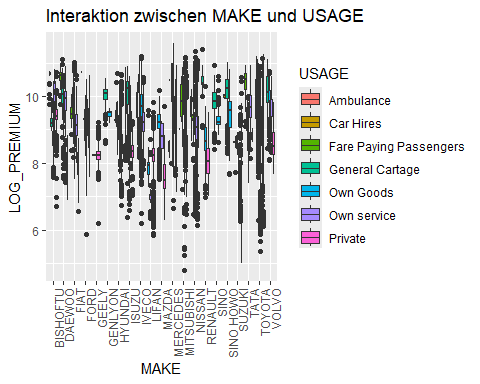
# Residuenplots für das neue Modell  
par(mfrow = c(2, 2))  
plot(lm\_model\_no\_influential)



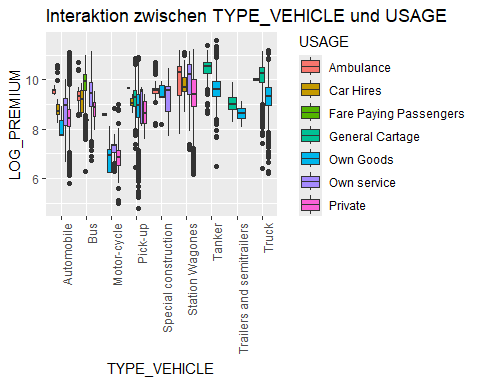
Nach dem Entfernen der einflussreichen Ausreisser zeigt das Modell eine deutliche Verbesserung. Die Residuen im Residuals vs Fitted Plot sind gleichmässiger um die Nulllinie verteilt, was auf eine verbesserte Homoskedastizität hinweist. Der Normal Q-Q Plot zeigt eine gute Annäherung an die Normalverteilung, mit nur geringen Abweichungen an den Enden. Der Scale-Location Plot bestätigt eine weitgehend konstante Varianz der Residuen, während im Residuals vs Leverage Plot nur wenige einflussreiche Datenpunkte verbleiben. Mit einem R-squared von 0,8028 erklärt das Modell nun 80,28% der Varianz, was eine signifikante Verbesserung darstellt.

### Massnahme 4): Interaktionen? noch nicht fix

# Boxplot von LOG\_PREMIUM nach MAKE und USAGE  
ggplot(clean\_data\_no\_outliers\_influential\_removed, aes(x=MAKE, y=LOG\_PREMIUM, fill=USAGE)) +   
 geom\_boxplot() +  
 theme(axis.text.x = element\_text(angle = 90, hjust = 1)) +  
 labs(title = "Interaktion zwischen MAKE und USAGE", x = "MAKE", y = "LOG\_PREMIUM")

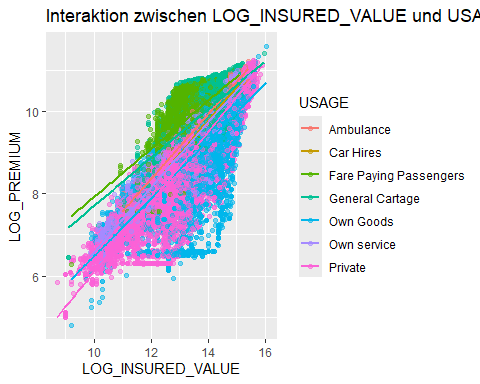


# Boxplot von LOG\_PREMIUM nach TYPE\_VEHICLE und USAGE  
ggplot(clean\_data\_no\_outliers\_influential\_removed, aes(x=TYPE\_VEHICLE, y=LOG\_PREMIUM, fill=USAGE)) +   
 geom\_boxplot() +  
 theme(axis.text.x = element\_text(angle = 90, hjust = 1)) +  
 labs(title = "Interaktion zwischen TYPE\_VEHICLE und USAGE", x = "TYPE\_VEHICLE", y = "LOG\_PREMIUM")



# Scatterplot von LOG\_INSURED\_VALUE und LOG\_PREMIUM nach USAGE  
ggplot(clean\_data\_no\_outliers\_influential\_removed, aes(x=LOG\_INSURED\_VALUE, y=LOG\_PREMIUM, color=USAGE)) +   
 geom\_point(alpha=0.5) +   
 geom\_smooth(method="lm", se=FALSE) +  
 labs(title = "Interaktion zwischen LOG\_INSURED\_VALUE und USAGE", x = "LOG\_INSURED\_VALUE", y = "LOG\_PREMIUM")

## `geom\_smooth()` using formula = 'y ~ x'



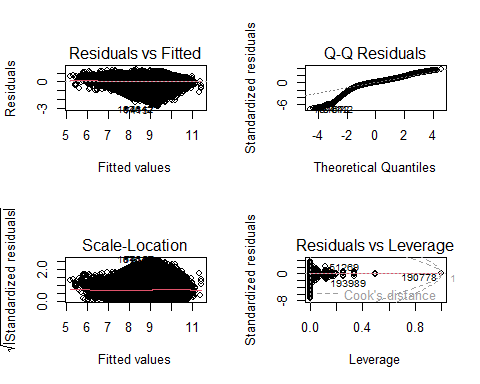
# Anpassung eines linearen Modells mit den identifizierten Interaktionen  
lm\_model\_interactions <- lm(LOG\_PREMIUM ~ SEX + INSR\_TYPE + MAKE \* USAGE + TYPE\_VEHICLE \* USAGE +   
 LOG\_INSURED\_VALUE \* USAGE + LOG\_CLAIM\_PAID\_USD + AGE\_VEHICLE + SEATS\_NUM +   
 LOG\_CCM\_TON, data = clean\_data\_no\_outliers\_influential\_removed)  
  
# Zusammenfassung des neuen Modells mit Interaktionen  
summary(lm\_model\_interactions)

##   
## Call:  
## lm(formula = LOG\_PREMIUM ~ SEX + INSR\_TYPE + MAKE \* USAGE + TYPE\_VEHICLE \*   
## USAGE + LOG\_INSURED\_VALUE \* USAGE + LOG\_CLAIM\_PAID\_USD +   
## AGE\_VEHICLE + SEATS\_NUM + LOG\_CCM\_TON, data = clean\_data\_no\_outliers\_influential\_removed)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -2.76942 -0.15003 0.04162 0.22142 1.35493   
##   
## Coefficients: (81 not defined because of singularities)  
## Estimate  
## (Intercept) -1.8246018  
## SEXMale -0.0291072  
## SEXFemale 0.0054850  
## INSR\_TYPEPrivate 0.0158815  
## MAKEDAEWOO 0.0598910  
## MAKEFIAT 0.0388500  
## MAKEFORD -0.3414971  
## MAKEGEELY 0.0078622  
## MAKEGENLYON -0.0120883  
## MAKEHYUNDAI -0.0084153  
## MAKEISUZU -0.1745789  
## MAKEIVECO 0.0240013  
## MAKELIFAN 0.0162048  
## MAKEMAZDA -0.0053859  
## MAKEMERCEDES -0.0273660  
## MAKEMITSUBISHI -0.0606001  
## MAKENISSAN -0.0663785  
## MAKERENAULT -0.0954276  
## MAKESINO 0.0691488  
## MAKESINO HOWO 0.0133370  
## MAKESUZUKI -0.0676224  
## MAKETATA 0.0654462  
## MAKETOYOTA -0.0775076  
## MAKEVOLVO -0.0797668  
## USAGECar Hires -0.0238792  
## USAGEFare Paying Passengers 0.2851017  
## USAGEGeneral Cartage 1.7066821  
## USAGEOwn Goods -0.1074615  
## USAGEOwn service -0.3075923  
## USAGEPrivate -1.4565617  
## TYPE\_VEHICLEBus 0.2479227  
## TYPE\_VEHICLEMotor-cycle 0.2361819  
## TYPE\_VEHICLEPick-up 0.2949417  
## TYPE\_VEHICLESpecial construction 0.0109813  
## TYPE\_VEHICLEStation Wagones 0.3579476  
## TYPE\_VEHICLETanker 0.4689762  
## TYPE\_VEHICLETrailers and semitrailers 0.0785203  
## TYPE\_VEHICLETruck 0.4454129  
## LOG\_INSURED\_VALUE 0.8147650  
## LOG\_CLAIM\_PAID\_USD 0.0091670  
## AGE\_VEHICLE 0.0127213  
## SEATS\_NUM 0.0004664  
## LOG\_CCM\_TON -0.0019201  
## MAKEDAEWOO:USAGECar Hires NA  
## MAKEFIAT:USAGECar Hires NA  
## MAKEFORD:USAGECar Hires NA  
## MAKEGEELY:USAGECar Hires -0.2139837  
## MAKEGENLYON:USAGECar Hires NA  
## MAKEHYUNDAI:USAGECar Hires -0.1327399  
## MAKEISUZU:USAGECar Hires NA  
## MAKEIVECO:USAGECar Hires NA  
## MAKELIFAN:USAGECar Hires NA  
## MAKEMAZDA:USAGECar Hires NA  
## MAKEMERCEDES:USAGECar Hires 0.0811028  
## MAKEMITSUBISHI:USAGECar Hires 0.2036921  
## MAKENISSAN:USAGECar Hires 0.0427499  
## MAKERENAULT:USAGECar Hires NA  
## MAKESINO:USAGECar Hires NA  
## MAKESINO HOWO:USAGECar Hires -0.2443172  
## MAKESUZUKI:USAGECar Hires 0.1380395  
## MAKETATA:USAGECar Hires NA  
## MAKETOYOTA:USAGECar Hires NA  
## MAKEVOLVO:USAGECar Hires NA  
## MAKEDAEWOO:USAGEFare Paying Passengers 0.3322447  
## MAKEFIAT:USAGEFare Paying Passengers 0.2453906  
## MAKEFORD:USAGEFare Paying Passengers 1.3201334  
## MAKEGEELY:USAGEFare Paying Passengers NA  
## MAKEGENLYON:USAGEFare Paying Passengers NA  
## MAKEHYUNDAI:USAGEFare Paying Passengers 0.5566476  
## MAKEISUZU:USAGEFare Paying Passengers 0.5582834  
## MAKEIVECO:USAGEFare Paying Passengers 0.2521531  
## MAKELIFAN:USAGEFare Paying Passengers NA  
## MAKEMAZDA:USAGEFare Paying Passengers NA  
## MAKEMERCEDES:USAGEFare Paying Passengers 0.3450260  
## MAKEMITSUBISHI:USAGEFare Paying Passengers 0.6247912  
## MAKENISSAN:USAGEFare Paying Passengers 0.4233597  
## MAKERENAULT:USAGEFare Paying Passengers NA  
## MAKESINO:USAGEFare Paying Passengers NA  
## MAKESINO HOWO:USAGEFare Paying Passengers NA  
## MAKESUZUKI:USAGEFare Paying Passengers 0.2360530  
## MAKETATA:USAGEFare Paying Passengers 0.3177009  
## MAKETOYOTA:USAGEFare Paying Passengers 0.7148300  
## MAKEVOLVO:USAGEFare Paying Passengers NA  
## MAKEDAEWOO:USAGEGeneral Cartage -0.0776452  
## MAKEFIAT:USAGEGeneral Cartage 0.0075832  
## MAKEFORD:USAGEGeneral Cartage 0.2349238  
## MAKEGEELY:USAGEGeneral Cartage NA  
## MAKEGENLYON:USAGEGeneral Cartage 0.0159210  
## MAKEHYUNDAI:USAGEGeneral Cartage -0.2060334  
## MAKEISUZU:USAGEGeneral Cartage 0.4560880  
## MAKEIVECO:USAGEGeneral Cartage 0.0235288  
## MAKELIFAN:USAGEGeneral Cartage NA  
## MAKEMAZDA:USAGEGeneral Cartage NA  
## MAKEMERCEDES:USAGEGeneral Cartage -0.0195780  
## MAKEMITSUBISHI:USAGEGeneral Cartage 0.0200236  
## MAKENISSAN:USAGEGeneral Cartage 0.0533277  
## MAKERENAULT:USAGEGeneral Cartage 0.1658223  
## MAKESINO:USAGEGeneral Cartage 0.0604534  
## MAKESINO HOWO:USAGEGeneral Cartage 0.2100986  
## MAKESUZUKI:USAGEGeneral Cartage 0.0091013  
## MAKETATA:USAGEGeneral Cartage -0.0534586  
## MAKETOYOTA:USAGEGeneral Cartage 0.0171487  
## MAKEVOLVO:USAGEGeneral Cartage 0.2710946  
## MAKEDAEWOO:USAGEOwn Goods 0.0572769  
## MAKEFIAT:USAGEOwn Goods -0.0305374  
## MAKEFORD:USAGEOwn Goods 0.4370446  
## MAKEGEELY:USAGEOwn Goods NA  
## MAKEGENLYON:USAGEOwn Goods NA  
## MAKEHYUNDAI:USAGEOwn Goods -0.0694645  
## MAKEISUZU:USAGEOwn Goods 0.3849563  
## MAKEIVECO:USAGEOwn Goods 0.0105627  
## MAKELIFAN:USAGEOwn Goods -0.1004702  
## MAKEMAZDA:USAGEOwn Goods 0.0731157  
## MAKEMERCEDES:USAGEOwn Goods 0.1187593  
## MAKEMITSUBISHI:USAGEOwn Goods 0.1096107  
## MAKENISSAN:USAGEOwn Goods 0.2254252  
## MAKERENAULT:USAGEOwn Goods 0.1793182  
## MAKESINO:USAGEOwn Goods NA  
## MAKESINO HOWO:USAGEOwn Goods NA  
## MAKESUZUKI:USAGEOwn Goods 0.2001975  
## MAKETATA:USAGEOwn Goods 0.0280819  
## MAKETOYOTA:USAGEOwn Goods 0.1348441  
## MAKEVOLVO:USAGEOwn Goods 0.4365154  
## MAKEDAEWOO:USAGEOwn service 0.0539118  
## MAKEFIAT:USAGEOwn service 0.0446421  
## MAKEFORD:USAGEOwn service 0.4654007  
## MAKEGEELY:USAGEOwn service NA  
## MAKEGENLYON:USAGEOwn service NA  
## MAKEHYUNDAI:USAGEOwn service 0.0680012  
## MAKEISUZU:USAGEOwn service 0.3630418  
## MAKEIVECO:USAGEOwn service NA  
## MAKELIFAN:USAGEOwn service 0.2997647  
## MAKEMAZDA:USAGEOwn service 0.2407098  
## MAKEMERCEDES:USAGEOwn service 0.0817083  
## MAKEMITSUBISHI:USAGEOwn service 0.1028383  
## MAKENISSAN:USAGEOwn service 0.0731062  
## MAKERENAULT:USAGEOwn service NA  
## MAKESINO:USAGEOwn service NA  
## MAKESINO HOWO:USAGEOwn service NA  
## MAKESUZUKI:USAGEOwn service 0.1348490  
## MAKETATA:USAGEOwn service 0.0194954  
## MAKETOYOTA:USAGEOwn service 0.0833600  
## MAKEVOLVO:USAGEOwn service -0.0191972  
## MAKEDAEWOO:USAGEPrivate NA  
## MAKEFIAT:USAGEPrivate NA  
## MAKEFORD:USAGEPrivate 0.2427756  
## MAKEGEELY:USAGEPrivate NA  
## MAKEGENLYON:USAGEPrivate NA  
## MAKEHYUNDAI:USAGEPrivate NA  
## MAKEISUZU:USAGEPrivate NA  
## MAKEIVECO:USAGEPrivate NA  
## MAKELIFAN:USAGEPrivate NA  
## MAKEMAZDA:USAGEPrivate NA  
## MAKEMERCEDES:USAGEPrivate -0.1026148  
## MAKEMITSUBISHI:USAGEPrivate NA  
## MAKENISSAN:USAGEPrivate NA  
## MAKERENAULT:USAGEPrivate NA  
## MAKESINO:USAGEPrivate NA  
## MAKESINO HOWO:USAGEPrivate NA  
## MAKESUZUKI:USAGEPrivate NA  
## MAKETATA:USAGEPrivate NA  
## MAKETOYOTA:USAGEPrivate NA  
## MAKEVOLVO:USAGEPrivate NA  
## USAGECar Hires:TYPE\_VEHICLEBus -0.2306092  
## USAGEFare Paying Passengers:TYPE\_VEHICLEBus NA  
## USAGEGeneral Cartage:TYPE\_VEHICLEBus NA  
## USAGEOwn Goods:TYPE\_VEHICLEBus 0.2081515  
## USAGEOwn service:TYPE\_VEHICLEBus -0.2473694  
## USAGEPrivate:TYPE\_VEHICLEBus -0.2246195  
## USAGECar Hires:TYPE\_VEHICLEMotor-cycle NA  
## USAGEFare Paying Passengers:TYPE\_VEHICLEMotor-cycle NA  
## USAGEGeneral Cartage:TYPE\_VEHICLEMotor-cycle -0.0157121  
## USAGEOwn Goods:TYPE\_VEHICLEMotor-cycle -0.2122101  
## USAGEOwn service:TYPE\_VEHICLEMotor-cycle -0.2784052  
## USAGEPrivate:TYPE\_VEHICLEMotor-cycle NA  
## USAGECar Hires:TYPE\_VEHICLEPick-up -0.3465775  
## USAGEFare Paying Passengers:TYPE\_VEHICLEPick-up NA  
## USAGEGeneral Cartage:TYPE\_VEHICLEPick-up -0.0882346  
## USAGEOwn Goods:TYPE\_VEHICLEPick-up -0.1064181  
## USAGEOwn service:TYPE\_VEHICLEPick-up -0.3147451  
## USAGEPrivate:TYPE\_VEHICLEPick-up -0.3656713  
## USAGECar Hires:TYPE\_VEHICLESpecial construction NA  
## USAGEFare Paying Passengers:TYPE\_VEHICLESpecial construction NA  
## USAGEGeneral Cartage:TYPE\_VEHICLESpecial construction NA  
## USAGEOwn Goods:TYPE\_VEHICLESpecial construction 0.1235966  
## USAGEOwn service:TYPE\_VEHICLESpecial construction -0.1050213  
## USAGEPrivate:TYPE\_VEHICLESpecial construction NA  
## USAGECar Hires:TYPE\_VEHICLEStation Wagones -0.1104604  
## USAGEFare Paying Passengers:TYPE\_VEHICLEStation Wagones NA  
## USAGEGeneral Cartage:TYPE\_VEHICLEStation Wagones NA  
## USAGEOwn Goods:TYPE\_VEHICLEStation Wagones NA  
## USAGEOwn service:TYPE\_VEHICLEStation Wagones -0.1318241  
## USAGEPrivate:TYPE\_VEHICLEStation Wagones -0.2983074  
## USAGECar Hires:TYPE\_VEHICLETanker NA  
## USAGEFare Paying Passengers:TYPE\_VEHICLETanker NA  
## USAGEGeneral Cartage:TYPE\_VEHICLETanker 0.1229410  
## USAGEOwn Goods:TYPE\_VEHICLETanker NA  
## USAGEOwn service:TYPE\_VEHICLETanker NA  
## USAGEPrivate:TYPE\_VEHICLETanker NA  
## USAGECar Hires:TYPE\_VEHICLETrailers and semitrailers NA  
## USAGEFare Paying Passengers:TYPE\_VEHICLETrailers and semitrailers NA  
## USAGEGeneral Cartage:TYPE\_VEHICLETrailers and semitrailers -0.1318514  
## USAGEOwn Goods:TYPE\_VEHICLETrailers and semitrailers NA  
## USAGEOwn service:TYPE\_VEHICLETrailers and semitrailers NA  
## USAGEPrivate:TYPE\_VEHICLETrailers and semitrailers NA  
## USAGECar Hires:TYPE\_VEHICLETruck NA  
## USAGEFare Paying Passengers:TYPE\_VEHICLETruck NA  
## USAGEGeneral Cartage:TYPE\_VEHICLETruck NA  
## USAGEOwn Goods:TYPE\_VEHICLETruck NA  
## USAGEOwn service:TYPE\_VEHICLETruck NA  
## USAGEPrivate:TYPE\_VEHICLETruck NA  
## USAGECar Hires:LOG\_INSURED\_VALUE 0.0005112  
## USAGEFare Paying Passengers:LOG\_INSURED\_VALUE -0.0257800  
## USAGEGeneral Cartage:LOG\_INSURED\_VALUE -0.1376237  
## USAGEOwn Goods:LOG\_INSURED\_VALUE -0.0442619  
## USAGEOwn service:LOG\_INSURED\_VALUE 0.0225932  
## USAGEPrivate:LOG\_INSURED\_VALUE 0.0911856  
## Std. Error  
## (Intercept) 0.2578191  
## SEXMale 0.0030482  
## SEXFemale 0.0042840  
## INSR\_TYPEPrivate 0.0136181  
## MAKEDAEWOO 0.0456870  
## MAKEFIAT 0.0522567  
## MAKEFORD 0.2807398  
## MAKEGEELY 0.0190728  
## MAKEGENLYON 0.1038026  
## MAKEHYUNDAI 0.0184806  
## MAKEISUZU 0.0338905  
## MAKEIVECO 0.0231347  
## MAKELIFAN 0.0170802  
## MAKEMAZDA 0.0531470  
## MAKEMERCEDES 0.1647750  
## MAKEMITSUBISHI 0.0175337  
## MAKENISSAN 0.0161576  
## MAKERENAULT 0.0284842  
## MAKESINO 0.0358918  
## MAKESINO HOWO 0.0138093  
## MAKESUZUKI 0.0186868  
## MAKETATA 0.2234848  
## MAKETOYOTA 0.0154212  
## MAKEVOLVO 0.0837920  
## USAGECar Hires 0.7928446  
## USAGEFare Paying Passengers 0.2430771  
## USAGEGeneral Cartage 0.3732034  
## USAGEOwn Goods 0.3418689  
## USAGEOwn service 0.2755810  
## USAGEPrivate 0.2588213  
## TYPE\_VEHICLEBus 0.1674575  
## TYPE\_VEHICLEMotor-cycle 0.0148082  
## TYPE\_VEHICLEPick-up 0.4171004  
## TYPE\_VEHICLESpecial construction 0.1635090  
## TYPE\_VEHICLEStation Wagones 0.1582414  
## TYPE\_VEHICLETanker 0.2233484  
## TYPE\_VEHICLETrailers and semitrailers 0.2298759  
## TYPE\_VEHICLETruck 0.2231721  
## LOG\_INSURED\_VALUE 0.0146503  
## LOG\_CLAIM\_PAID\_USD 0.0002328  
## AGE\_VEHICLE 0.0001932  
## SEATS\_NUM 0.0002664  
## LOG\_CCM\_TON 0.0027466  
## MAKEDAEWOO:USAGECar Hires NA  
## MAKEFIAT:USAGECar Hires NA  
## MAKEFORD:USAGECar Hires NA  
## MAKEGEELY:USAGECar Hires 0.3954146  
## MAKEGENLYON:USAGECar Hires NA  
## MAKEHYUNDAI:USAGECar Hires 0.3976951  
## MAKEISUZU:USAGECar Hires NA  
## MAKEIVECO:USAGECar Hires NA  
## MAKELIFAN:USAGECar Hires NA  
## MAKEMAZDA:USAGECar Hires NA  
## MAKEMERCEDES:USAGECar Hires 0.4398448  
## MAKEMITSUBISHI:USAGECar Hires 0.4076309  
## MAKENISSAN:USAGECar Hires 0.1456663  
## MAKERENAULT:USAGECar Hires NA  
## MAKESINO:USAGECar Hires NA  
## MAKESINO HOWO:USAGECar Hires 0.3328652  
## MAKESUZUKI:USAGECar Hires 0.3962417  
## MAKETATA:USAGECar Hires NA  
## MAKETOYOTA:USAGECar Hires NA  
## MAKEVOLVO:USAGECar Hires NA  
## MAKEDAEWOO:USAGEFare Paying Passengers 0.0773942  
## MAKEFIAT:USAGEFare Paying Passengers 0.1118091  
## MAKEFORD:USAGEFare Paying Passengers 0.4804325  
## MAKEGEELY:USAGEFare Paying Passengers NA  
## MAKEGENLYON:USAGEFare Paying Passengers NA  
## MAKEHYUNDAI:USAGEFare Paying Passengers 0.1467690  
## MAKEISUZU:USAGEFare Paying Passengers 0.0597164  
## MAKEIVECO:USAGEFare Paying Passengers 0.0957577  
## MAKELIFAN:USAGEFare Paying Passengers NA  
## MAKEMAZDA:USAGEFare Paying Passengers NA  
## MAKEMERCEDES:USAGEFare Paying Passengers 0.1736616  
## MAKEMITSUBISHI:USAGEFare Paying Passengers 0.0791868  
## MAKENISSAN:USAGEFare Paying Passengers 0.1802409  
## MAKERENAULT:USAGEFare Paying Passengers NA  
## MAKESINO:USAGEFare Paying Passengers NA  
## MAKESINO HOWO:USAGEFare Paying Passengers NA  
## MAKESUZUKI:USAGEFare Paying Passengers 0.2014958  
## MAKETATA:USAGEFare Paying Passengers 0.2290779  
## MAKETOYOTA:USAGEFare Paying Passengers 0.0536034  
## MAKEVOLVO:USAGEFare Paying Passengers NA  
## MAKEDAEWOO:USAGEGeneral Cartage 0.1483648  
## MAKEFIAT:USAGEGeneral Cartage 0.1431605  
## MAKEFORD:USAGEGeneral Cartage 0.3347734  
## MAKEGEELY:USAGEGeneral Cartage NA  
## MAKEGENLYON:USAGEGeneral Cartage 0.1690337  
## MAKEHYUNDAI:USAGEGeneral Cartage 0.1459593  
## MAKEISUZU:USAGEGeneral Cartage 0.1366964  
## MAKEIVECO:USAGEGeneral Cartage 0.1344671  
## MAKELIFAN:USAGEGeneral Cartage NA  
## MAKEMAZDA:USAGEGeneral Cartage NA  
## MAKEMERCEDES:USAGEGeneral Cartage 0.2122340  
## MAKEMITSUBISHI:USAGEGeneral Cartage 0.1336558  
## MAKENISSAN:USAGEGeneral Cartage 0.1338408  
## MAKERENAULT:USAGEGeneral Cartage 0.1359893  
## MAKESINO:USAGEGeneral Cartage 0.1376461  
## MAKESINO HOWO:USAGEGeneral Cartage 0.1331393  
## MAKESUZUKI:USAGEGeneral Cartage 0.2331302  
## MAKETATA:USAGEGeneral Cartage 0.2634048  
## MAKETOYOTA:USAGEGeneral Cartage 0.1307733  
## MAKEVOLVO:USAGEGeneral Cartage 0.1617415  
## MAKEDAEWOO:USAGEOwn Goods 0.0490389  
## MAKEFIAT:USAGEOwn Goods 0.0552425  
## MAKEFORD:USAGEOwn Goods 0.2809833  
## MAKEGEELY:USAGEOwn Goods NA  
## MAKEGENLYON:USAGEOwn Goods NA  
## MAKEHYUNDAI:USAGEOwn Goods 0.0242890  
## MAKEISUZU:USAGEOwn Goods 0.0359417  
## MAKEIVECO:USAGEOwn Goods 0.0267997  
## MAKELIFAN:USAGEOwn Goods 0.0626531  
## MAKEMAZDA:USAGEOwn Goods 0.0546483  
## MAKEMERCEDES:USAGEOwn Goods 0.1660355  
## MAKEMITSUBISHI:USAGEOwn Goods 0.0204063  
## MAKENISSAN:USAGEOwn Goods 0.0190347  
## MAKERENAULT:USAGEOwn Goods 0.0427685  
## MAKESINO:USAGEOwn Goods NA  
## MAKESINO HOWO:USAGEOwn Goods NA  
## MAKESUZUKI:USAGEOwn Goods 0.0685711  
## MAKETATA:USAGEOwn Goods 0.2247815  
## MAKETOYOTA:USAGEOwn Goods 0.0178467  
## MAKEVOLVO:USAGEOwn Goods 0.0854624  
## MAKEDAEWOO:USAGEOwn service 0.0490875  
## MAKEFIAT:USAGEOwn service 0.0584397  
## MAKEFORD:USAGEOwn service 0.2911213  
## MAKEGEELY:USAGEOwn service NA  
## MAKEGENLYON:USAGEOwn service NA  
## MAKEHYUNDAI:USAGEOwn service 0.0299575  
## MAKEISUZU:USAGEOwn service 0.0522561  
## MAKEIVECO:USAGEOwn service NA  
## MAKELIFAN:USAGEOwn service 0.0824675  
## MAKEMAZDA:USAGEOwn service 0.1560903  
## MAKEMERCEDES:USAGEOwn service 0.1657941  
## MAKEMITSUBISHI:USAGEOwn service 0.0255584  
## MAKENISSAN:USAGEOwn service 0.0274205  
## MAKERENAULT:USAGEOwn service NA  
## MAKESINO:USAGEOwn service NA  
## MAKESINO HOWO:USAGEOwn service NA  
## MAKESUZUKI:USAGEOwn service 0.0652462  
## MAKETATA:USAGEOwn service 0.2246192  
## MAKETOYOTA:USAGEOwn service 0.0211783  
## MAKEVOLVO:USAGEOwn service 0.0953310  
## MAKEDAEWOO:USAGEPrivate NA  
## MAKEFIAT:USAGEPrivate NA  
## MAKEFORD:USAGEPrivate 0.2812050  
## MAKEGEELY:USAGEPrivate NA  
## MAKEGENLYON:USAGEPrivate NA  
## MAKEHYUNDAI:USAGEPrivate NA  
## MAKEISUZU:USAGEPrivate NA  
## MAKEIVECO:USAGEPrivate NA  
## MAKELIFAN:USAGEPrivate NA  
## MAKEMAZDA:USAGEPrivate NA  
## MAKEMERCEDES:USAGEPrivate 0.1647200  
## MAKEMITSUBISHI:USAGEPrivate NA  
## MAKENISSAN:USAGEPrivate NA  
## MAKERENAULT:USAGEPrivate NA  
## MAKESINO:USAGEPrivate NA  
## MAKESINO HOWO:USAGEPrivate NA  
## MAKESUZUKI:USAGEPrivate NA  
## MAKETATA:USAGEPrivate NA  
## MAKETOYOTA:USAGEPrivate NA  
## MAKEVOLVO:USAGEPrivate NA  
## USAGECar Hires:TYPE\_VEHICLEBus 0.2072255  
## USAGEFare Paying Passengers:TYPE\_VEHICLEBus NA  
## USAGEGeneral Cartage:TYPE\_VEHICLEBus NA  
## USAGEOwn Goods:TYPE\_VEHICLEBus 0.4801219  
## USAGEOwn service:TYPE\_VEHICLEBus 0.1803263  
## USAGEPrivate:TYPE\_VEHICLEBus 0.1812796  
## USAGECar Hires:TYPE\_VEHICLEMotor-cycle NA  
## USAGEFare Paying Passengers:TYPE\_VEHICLEMotor-cycle NA  
## USAGEGeneral Cartage:TYPE\_VEHICLEMotor-cycle 0.4041229  
## USAGEOwn Goods:TYPE\_VEHICLEMotor-cycle 0.2409277  
## USAGEOwn service:TYPE\_VEHICLEMotor-cycle 0.0923414  
## USAGEPrivate:TYPE\_VEHICLEMotor-cycle NA  
## USAGECar Hires:TYPE\_VEHICLEPick-up 0.4440176  
## USAGEFare Paying Passengers:TYPE\_VEHICLEPick-up NA  
## USAGEGeneral Cartage:TYPE\_VEHICLEPick-up 0.4746029  
## USAGEOwn Goods:TYPE\_VEHICLEPick-up 0.4730489  
## USAGEOwn service:TYPE\_VEHICLEPick-up 0.4415476  
## USAGEPrivate:TYPE\_VEHICLEPick-up 0.4221997  
## USAGECar Hires:TYPE\_VEHICLESpecial construction NA  
## USAGEFare Paying Passengers:TYPE\_VEHICLESpecial construction NA  
## USAGEGeneral Cartage:TYPE\_VEHICLESpecial construction NA  
## USAGEOwn Goods:TYPE\_VEHICLESpecial construction 0.3184424  
## USAGEOwn service:TYPE\_VEHICLESpecial construction 0.1822014  
## USAGEPrivate:TYPE\_VEHICLESpecial construction NA  
## USAGECar Hires:TYPE\_VEHICLEStation Wagones 0.1899512  
## USAGEFare Paying Passengers:TYPE\_VEHICLEStation Wagones NA  
## USAGEGeneral Cartage:TYPE\_VEHICLEStation Wagones NA  
## USAGEOwn Goods:TYPE\_VEHICLEStation Wagones NA  
## USAGEOwn service:TYPE\_VEHICLEStation Wagones 0.1728147  
## USAGEPrivate:TYPE\_VEHICLEStation Wagones 0.1583169  
## USAGECar Hires:TYPE\_VEHICLETanker NA  
## USAGEFare Paying Passengers:TYPE\_VEHICLETanker NA  
## USAGEGeneral Cartage:TYPE\_VEHICLETanker 0.0197384  
## USAGEOwn Goods:TYPE\_VEHICLETanker NA  
## USAGEOwn service:TYPE\_VEHICLETanker NA  
## USAGEPrivate:TYPE\_VEHICLETanker NA  
## USAGECar Hires:TYPE\_VEHICLETrailers and semitrailers NA  
## USAGEFare Paying Passengers:TYPE\_VEHICLETrailers and semitrailers NA  
## USAGEGeneral Cartage:TYPE\_VEHICLETrailers and semitrailers 0.0588095  
## USAGEOwn Goods:TYPE\_VEHICLETrailers and semitrailers NA  
## USAGEOwn service:TYPE\_VEHICLETrailers and semitrailers NA  
## USAGEPrivate:TYPE\_VEHICLETrailers and semitrailers NA  
## USAGECar Hires:TYPE\_VEHICLETruck NA  
## USAGEFare Paying Passengers:TYPE\_VEHICLETruck NA  
## USAGEGeneral Cartage:TYPE\_VEHICLETruck NA  
## USAGEOwn Goods:TYPE\_VEHICLETruck NA  
## USAGEOwn service:TYPE\_VEHICLETruck NA  
## USAGEPrivate:TYPE\_VEHICLETruck NA  
## USAGECar Hires:LOG\_INSURED\_VALUE 0.0588271  
## USAGEFare Paying Passengers:LOG\_INSURED\_VALUE 0.0169208  
## USAGEGeneral Cartage:LOG\_INSURED\_VALUE 0.0155819  
## USAGEOwn Goods:LOG\_INSURED\_VALUE 0.0148171  
## USAGEOwn service:LOG\_INSURED\_VALUE 0.0156060  
## USAGEPrivate:LOG\_INSURED\_VALUE 0.0148405  
## t value  
## (Intercept) -7.077  
## SEXMale -9.549  
## SEXFemale 1.280  
## INSR\_TYPEPrivate 1.166  
## MAKEDAEWOO 1.311  
## MAKEFIAT 0.743  
## MAKEFORD -1.216  
## MAKEGEELY 0.412  
## MAKEGENLYON -0.116  
## MAKEHYUNDAI -0.455  
## MAKEISUZU -5.151  
## MAKEIVECO 1.037  
## MAKELIFAN 0.949  
## MAKEMAZDA -0.101  
## MAKEMERCEDES -0.166  
## MAKEMITSUBISHI -3.456  
## MAKENISSAN -4.108  
## MAKERENAULT -3.350  
## MAKESINO 1.927  
## MAKESINO HOWO 0.966  
## MAKESUZUKI -3.619  
## MAKETATA 0.293  
## MAKETOYOTA -5.026  
## MAKEVOLVO -0.952  
## USAGECar Hires -0.030  
## USAGEFare Paying Passengers 1.173  
## USAGEGeneral Cartage 4.573  
## USAGEOwn Goods -0.314  
## USAGEOwn service -1.116  
## USAGEPrivate -5.628  
## TYPE\_VEHICLEBus 1.481  
## TYPE\_VEHICLEMotor-cycle 15.949  
## TYPE\_VEHICLEPick-up 0.707  
## TYPE\_VEHICLESpecial construction 0.067  
## TYPE\_VEHICLEStation Wagones 2.262  
## TYPE\_VEHICLETanker 2.100  
## TYPE\_VEHICLETrailers and semitrailers 0.342  
## TYPE\_VEHICLETruck 1.996  
## LOG\_INSURED\_VALUE 55.614  
## LOG\_CLAIM\_PAID\_USD 39.383  
## AGE\_VEHICLE 65.859  
## SEATS\_NUM 1.750  
## LOG\_CCM\_TON -0.699  
## MAKEDAEWOO:USAGECar Hires NA  
## MAKEFIAT:USAGECar Hires NA  
## MAKEFORD:USAGECar Hires NA  
## MAKEGEELY:USAGECar Hires -0.541  
## MAKEGENLYON:USAGECar Hires NA  
## MAKEHYUNDAI:USAGECar Hires -0.334  
## MAKEISUZU:USAGECar Hires NA  
## MAKEIVECO:USAGECar Hires NA  
## MAKELIFAN:USAGECar Hires NA  
## MAKEMAZDA:USAGECar Hires NA  
## MAKEMERCEDES:USAGECar Hires 0.184  
## MAKEMITSUBISHI:USAGECar Hires 0.500  
## MAKENISSAN:USAGECar Hires 0.293  
## MAKERENAULT:USAGECar Hires NA  
## MAKESINO:USAGECar Hires NA  
## MAKESINO HOWO:USAGECar Hires -0.734  
## MAKESUZUKI:USAGECar Hires 0.348  
## MAKETATA:USAGECar Hires NA  
## MAKETOYOTA:USAGECar Hires NA  
## MAKEVOLVO:USAGECar Hires NA  
## MAKEDAEWOO:USAGEFare Paying Passengers 4.293  
## MAKEFIAT:USAGEFare Paying Passengers 2.195  
## MAKEFORD:USAGEFare Paying Passengers 2.748  
## MAKEGEELY:USAGEFare Paying Passengers NA  
## MAKEGENLYON:USAGEFare Paying Passengers NA  
## MAKEHYUNDAI:USAGEFare Paying Passengers 3.793  
## MAKEISUZU:USAGEFare Paying Passengers 9.349  
## MAKEIVECO:USAGEFare Paying Passengers 2.633  
## MAKELIFAN:USAGEFare Paying Passengers NA  
## MAKEMAZDA:USAGEFare Paying Passengers NA  
## MAKEMERCEDES:USAGEFare Paying Passengers 1.987  
## MAKEMITSUBISHI:USAGEFare Paying Passengers 7.890  
## MAKENISSAN:USAGEFare Paying Passengers 2.349  
## MAKERENAULT:USAGEFare Paying Passengers NA  
## MAKESINO:USAGEFare Paying Passengers NA  
## MAKESINO HOWO:USAGEFare Paying Passengers NA  
## MAKESUZUKI:USAGEFare Paying Passengers 1.172  
## MAKETATA:USAGEFare Paying Passengers 1.387  
## MAKETOYOTA:USAGEFare Paying Passengers 13.336  
## MAKEVOLVO:USAGEFare Paying Passengers NA  
## MAKEDAEWOO:USAGEGeneral Cartage -0.523  
## MAKEFIAT:USAGEGeneral Cartage 0.053  
## MAKEFORD:USAGEGeneral Cartage 0.702  
## MAKEGEELY:USAGEGeneral Cartage NA  
## MAKEGENLYON:USAGEGeneral Cartage 0.094  
## MAKEHYUNDAI:USAGEGeneral Cartage -1.412  
## MAKEISUZU:USAGEGeneral Cartage 3.337  
## MAKEIVECO:USAGEGeneral Cartage 0.175  
## MAKELIFAN:USAGEGeneral Cartage NA  
## MAKEMAZDA:USAGEGeneral Cartage NA  
## MAKEMERCEDES:USAGEGeneral Cartage -0.092  
## MAKEMITSUBISHI:USAGEGeneral Cartage 0.150  
## MAKENISSAN:USAGEGeneral Cartage 0.398  
## MAKERENAULT:USAGEGeneral Cartage 1.219  
## MAKESINO:USAGEGeneral Cartage 0.439  
## MAKESINO HOWO:USAGEGeneral Cartage 1.578  
## MAKESUZUKI:USAGEGeneral Cartage 0.039  
## MAKETATA:USAGEGeneral Cartage -0.203  
## MAKETOYOTA:USAGEGeneral Cartage 0.131  
## MAKEVOLVO:USAGEGeneral Cartage 1.676  
## MAKEDAEWOO:USAGEOwn Goods 1.168  
## MAKEFIAT:USAGEOwn Goods -0.553  
## MAKEFORD:USAGEOwn Goods 1.555  
## MAKEGEELY:USAGEOwn Goods NA  
## MAKEGENLYON:USAGEOwn Goods NA  
## MAKEHYUNDAI:USAGEOwn Goods -2.860  
## MAKEISUZU:USAGEOwn Goods 10.711  
## MAKEIVECO:USAGEOwn Goods 0.394  
## MAKELIFAN:USAGEOwn Goods -1.604  
## MAKEMAZDA:USAGEOwn Goods 1.338  
## MAKEMERCEDES:USAGEOwn Goods 0.715  
## MAKEMITSUBISHI:USAGEOwn Goods 5.371  
## MAKENISSAN:USAGEOwn Goods 11.843  
## MAKERENAULT:USAGEOwn Goods 4.193  
## MAKESINO:USAGEOwn Goods NA  
## MAKESINO HOWO:USAGEOwn Goods NA  
## MAKESUZUKI:USAGEOwn Goods 2.920  
## MAKETATA:USAGEOwn Goods 0.125  
## MAKETOYOTA:USAGEOwn Goods 7.556  
## MAKEVOLVO:USAGEOwn Goods 5.108  
## MAKEDAEWOO:USAGEOwn service 1.098  
## MAKEFIAT:USAGEOwn service 0.764  
## MAKEFORD:USAGEOwn service 1.599  
## MAKEGEELY:USAGEOwn service NA  
## MAKEGENLYON:USAGEOwn service NA  
## MAKEHYUNDAI:USAGEOwn service 2.270  
## MAKEISUZU:USAGEOwn service 6.947  
## MAKEIVECO:USAGEOwn service NA  
## MAKELIFAN:USAGEOwn service 3.635  
## MAKEMAZDA:USAGEOwn service 1.542  
## MAKEMERCEDES:USAGEOwn service 0.493  
## MAKEMITSUBISHI:USAGEOwn service 4.024  
## MAKENISSAN:USAGEOwn service 2.666  
## MAKERENAULT:USAGEOwn service NA  
## MAKESINO:USAGEOwn service NA  
## MAKESINO HOWO:USAGEOwn service NA  
## MAKESUZUKI:USAGEOwn service 2.067  
## MAKETATA:USAGEOwn service 0.087  
## MAKETOYOTA:USAGEOwn service 3.936  
## MAKEVOLVO:USAGEOwn service -0.201  
## MAKEDAEWOO:USAGEPrivate NA  
## MAKEFIAT:USAGEPrivate NA  
## MAKEFORD:USAGEPrivate 0.863  
## MAKEGEELY:USAGEPrivate NA  
## MAKEGENLYON:USAGEPrivate NA  
## MAKEHYUNDAI:USAGEPrivate NA  
## MAKEISUZU:USAGEPrivate NA  
## MAKEIVECO:USAGEPrivate NA  
## MAKELIFAN:USAGEPrivate NA  
## MAKEMAZDA:USAGEPrivate NA  
## MAKEMERCEDES:USAGEPrivate -0.623  
## MAKEMITSUBISHI:USAGEPrivate NA  
## MAKENISSAN:USAGEPrivate NA  
## MAKERENAULT:USAGEPrivate NA  
## MAKESINO:USAGEPrivate NA  
## MAKESINO HOWO:USAGEPrivate NA  
## MAKESUZUKI:USAGEPrivate NA  
## MAKETATA:USAGEPrivate NA  
## MAKETOYOTA:USAGEPrivate NA  
## MAKEVOLVO:USAGEPrivate NA  
## USAGECar Hires:TYPE\_VEHICLEBus -1.113  
## USAGEFare Paying Passengers:TYPE\_VEHICLEBus NA  
## USAGEGeneral Cartage:TYPE\_VEHICLEBus NA  
## USAGEOwn Goods:TYPE\_VEHICLEBus 0.434  
## USAGEOwn service:TYPE\_VEHICLEBus -1.372  
## USAGEPrivate:TYPE\_VEHICLEBus -1.239  
## USAGECar Hires:TYPE\_VEHICLEMotor-cycle NA  
## USAGEFare Paying Passengers:TYPE\_VEHICLEMotor-cycle NA  
## USAGEGeneral Cartage:TYPE\_VEHICLEMotor-cycle -0.039  
## USAGEOwn Goods:TYPE\_VEHICLEMotor-cycle -0.881  
## USAGEOwn service:TYPE\_VEHICLEMotor-cycle -3.015  
## USAGEPrivate:TYPE\_VEHICLEMotor-cycle NA  
## USAGECar Hires:TYPE\_VEHICLEPick-up -0.781  
## USAGEFare Paying Passengers:TYPE\_VEHICLEPick-up NA  
## USAGEGeneral Cartage:TYPE\_VEHICLEPick-up -0.186  
## USAGEOwn Goods:TYPE\_VEHICLEPick-up -0.225  
## USAGEOwn service:TYPE\_VEHICLEPick-up -0.713  
## USAGEPrivate:TYPE\_VEHICLEPick-up -0.866  
## USAGECar Hires:TYPE\_VEHICLESpecial construction NA  
## USAGEFare Paying Passengers:TYPE\_VEHICLESpecial construction NA  
## USAGEGeneral Cartage:TYPE\_VEHICLESpecial construction NA  
## USAGEOwn Goods:TYPE\_VEHICLESpecial construction 0.388  
## USAGEOwn service:TYPE\_VEHICLESpecial construction -0.576  
## USAGEPrivate:TYPE\_VEHICLESpecial construction NA  
## USAGECar Hires:TYPE\_VEHICLEStation Wagones -0.582  
## USAGEFare Paying Passengers:TYPE\_VEHICLEStation Wagones NA  
## USAGEGeneral Cartage:TYPE\_VEHICLEStation Wagones NA  
## USAGEOwn Goods:TYPE\_VEHICLEStation Wagones NA  
## USAGEOwn service:TYPE\_VEHICLEStation Wagones -0.763  
## USAGEPrivate:TYPE\_VEHICLEStation Wagones -1.884  
## USAGECar Hires:TYPE\_VEHICLETanker NA  
## USAGEFare Paying Passengers:TYPE\_VEHICLETanker NA  
## USAGEGeneral Cartage:TYPE\_VEHICLETanker 6.229  
## USAGEOwn Goods:TYPE\_VEHICLETanker NA  
## USAGEOwn service:TYPE\_VEHICLETanker NA  
## USAGEPrivate:TYPE\_VEHICLETanker NA  
## USAGECar Hires:TYPE\_VEHICLETrailers and semitrailers NA  
## USAGEFare Paying Passengers:TYPE\_VEHICLETrailers and semitrailers NA  
## USAGEGeneral Cartage:TYPE\_VEHICLETrailers and semitrailers -2.242  
## USAGEOwn Goods:TYPE\_VEHICLETrailers and semitrailers NA  
## USAGEOwn service:TYPE\_VEHICLETrailers and semitrailers NA  
## USAGEPrivate:TYPE\_VEHICLETrailers and semitrailers NA  
## USAGECar Hires:TYPE\_VEHICLETruck NA  
## USAGEFare Paying Passengers:TYPE\_VEHICLETruck NA  
## USAGEGeneral Cartage:TYPE\_VEHICLETruck NA  
## USAGEOwn Goods:TYPE\_VEHICLETruck NA  
## USAGEOwn service:TYPE\_VEHICLETruck NA  
## USAGEPrivate:TYPE\_VEHICLETruck NA  
## USAGECar Hires:LOG\_INSURED\_VALUE 0.009  
## USAGEFare Paying Passengers:LOG\_INSURED\_VALUE -1.524  
## USAGEGeneral Cartage:LOG\_INSURED\_VALUE -8.832  
## USAGEOwn Goods:LOG\_INSURED\_VALUE -2.987  
## USAGEOwn service:LOG\_INSURED\_VALUE 1.448  
## USAGEPrivate:LOG\_INSURED\_VALUE 6.144  
## Pr(>|t|)   
## (Intercept) 1.48e-12 \*\*\*  
## SEXMale < 2e-16 \*\*\*  
## SEXFemale 0.200425   
## INSR\_TYPEPrivate 0.243532   
## MAKEDAEWOO 0.189894   
## MAKEFIAT 0.457213   
## MAKEFORD 0.223827   
## MAKEGEELY 0.680180   
## MAKEGENLYON 0.907292   
## MAKEHYUNDAI 0.648853   
## MAKEISUZU 2.59e-07 \*\*\*  
## MAKEIVECO 0.299524   
## MAKELIFAN 0.342751   
## MAKEMAZDA 0.919280   
## MAKEMERCEDES 0.868093   
## MAKEMITSUBISHI 0.000548 \*\*\*  
## MAKENISSAN 3.99e-05 \*\*\*  
## MAKERENAULT 0.000808 \*\*\*  
## MAKESINO 0.054032 .   
## MAKESINO HOWO 0.334146   
## MAKESUZUKI 0.000296 \*\*\*  
## MAKETATA 0.769641   
## MAKETOYOTA 5.01e-07 \*\*\*  
## MAKEVOLVO 0.341117   
## USAGECar Hires 0.975973   
## USAGEFare Paying Passengers 0.240843   
## USAGEGeneral Cartage 4.81e-06 \*\*\*  
## USAGEOwn Goods 0.753267   
## USAGEOwn service 0.264355   
## USAGEPrivate 1.83e-08 \*\*\*  
## TYPE\_VEHICLEBus 0.138739   
## TYPE\_VEHICLEMotor-cycle < 2e-16 \*\*\*  
## TYPE\_VEHICLEPick-up 0.479490   
## TYPE\_VEHICLESpecial construction 0.946454   
## TYPE\_VEHICLEStation Wagones 0.023696 \*   
## TYPE\_VEHICLETanker 0.035752 \*   
## TYPE\_VEHICLETrailers and semitrailers 0.732670   
## TYPE\_VEHICLETruck 0.045954 \*   
## LOG\_INSURED\_VALUE < 2e-16 \*\*\*  
## LOG\_CLAIM\_PAID\_USD < 2e-16 \*\*\*  
## AGE\_VEHICLE < 2e-16 \*\*\*  
## SEATS\_NUM 0.080037 .   
## LOG\_CCM\_TON 0.484508   
## MAKEDAEWOO:USAGECar Hires NA   
## MAKEFIAT:USAGECar Hires NA   
## MAKEFORD:USAGECar Hires NA   
## MAKEGEELY:USAGECar Hires 0.588396   
## MAKEGENLYON:USAGECar Hires NA   
## MAKEHYUNDAI:USAGECar Hires 0.738551   
## MAKEISUZU:USAGECar Hires NA   
## MAKEIVECO:USAGECar Hires NA   
## MAKELIFAN:USAGECar Hires NA   
## MAKEMAZDA:USAGECar Hires NA   
## MAKEMERCEDES:USAGECar Hires 0.853708   
## MAKEMITSUBISHI:USAGECar Hires 0.617289   
## MAKENISSAN:USAGECar Hires 0.769157   
## MAKERENAULT:USAGECar Hires NA   
## MAKESINO:USAGECar Hires NA   
## MAKESINO HOWO:USAGECar Hires 0.462960   
## MAKESUZUKI:USAGECar Hires 0.727561   
## MAKETATA:USAGECar Hires NA   
## MAKETOYOTA:USAGECar Hires NA   
## MAKEVOLVO:USAGECar Hires NA   
## MAKEDAEWOO:USAGEFare Paying Passengers 1.76e-05 \*\*\*  
## MAKEFIAT:USAGEFare Paying Passengers 0.028184 \*   
## MAKEFORD:USAGEFare Paying Passengers 0.006000 \*\*   
## MAKEGEELY:USAGEFare Paying Passengers NA   
## MAKEGENLYON:USAGEFare Paying Passengers NA   
## MAKEHYUNDAI:USAGEFare Paying Passengers 0.000149 \*\*\*  
## MAKEISUZU:USAGEFare Paying Passengers < 2e-16 \*\*\*  
## MAKEIVECO:USAGEFare Paying Passengers 0.008458 \*\*   
## MAKELIFAN:USAGEFare Paying Passengers NA   
## MAKEMAZDA:USAGEFare Paying Passengers NA   
## MAKEMERCEDES:USAGEFare Paying Passengers 0.046949 \*   
## MAKEMITSUBISHI:USAGEFare Paying Passengers 3.04e-15 \*\*\*  
## MAKENISSAN:USAGEFare Paying Passengers 0.018832 \*   
## MAKERENAULT:USAGEFare Paying Passengers NA   
## MAKESINO:USAGEFare Paying Passengers NA   
## MAKESINO HOWO:USAGEFare Paying Passengers NA   
## MAKESUZUKI:USAGEFare Paying Passengers 0.241398   
## MAKETATA:USAGEFare Paying Passengers 0.165484   
## MAKETOYOTA:USAGEFare Paying Passengers < 2e-16 \*\*\*  
## MAKEVOLVO:USAGEFare Paying Passengers NA   
## MAKEDAEWOO:USAGEGeneral Cartage 0.600738   
## MAKEFIAT:USAGEGeneral Cartage 0.957756   
## MAKEFORD:USAGEGeneral Cartage 0.482842   
## MAKEGEELY:USAGEGeneral Cartage NA   
## MAKEGENLYON:USAGEGeneral Cartage 0.924960   
## MAKEHYUNDAI:USAGEGeneral Cartage 0.158075   
## MAKEISUZU:USAGEGeneral Cartage 0.000849 \*\*\*  
## MAKEIVECO:USAGEGeneral Cartage 0.861097   
## MAKELIFAN:USAGEGeneral Cartage NA   
## MAKEMAZDA:USAGEGeneral Cartage NA   
## MAKEMERCEDES:USAGEGeneral Cartage 0.926502   
## MAKEMITSUBISHI:USAGEGeneral Cartage 0.880911   
## MAKENISSAN:USAGEGeneral Cartage 0.690306   
## MAKERENAULT:USAGEGeneral Cartage 0.222702   
## MAKESINO:USAGEGeneral Cartage 0.660521   
## MAKESINO HOWO:USAGEGeneral Cartage 0.114559   
## MAKESUZUKI:USAGEGeneral Cartage 0.968859   
## MAKETATA:USAGEGeneral Cartage 0.839173   
## MAKETOYOTA:USAGEGeneral Cartage 0.895670   
## MAKEVOLVO:USAGEGeneral Cartage 0.093721 .   
## MAKEDAEWOO:USAGEOwn Goods 0.242812   
## MAKEFIAT:USAGEOwn Goods 0.580409   
## MAKEFORD:USAGEOwn Goods 0.119850   
## MAKEGEELY:USAGEOwn Goods NA   
## MAKEGENLYON:USAGEOwn Goods NA   
## MAKEHYUNDAI:USAGEOwn Goods 0.004238 \*\*   
## MAKEISUZU:USAGEOwn Goods < 2e-16 \*\*\*  
## MAKEIVECO:USAGEOwn Goods 0.693480   
## MAKELIFAN:USAGEOwn Goods 0.108805   
## MAKEMAZDA:USAGEOwn Goods 0.180920   
## MAKEMERCEDES:USAGEOwn Goods 0.474446   
## MAKEMITSUBISHI:USAGEOwn Goods 7.82e-08 \*\*\*  
## MAKENISSAN:USAGEOwn Goods < 2e-16 \*\*\*  
## MAKERENAULT:USAGEOwn Goods 2.76e-05 \*\*\*  
## MAKESINO:USAGEOwn Goods NA   
## MAKESINO HOWO:USAGEOwn Goods NA   
## MAKESUZUKI:USAGEOwn Goods 0.003506 \*\*   
## MAKETATA:USAGEOwn Goods 0.900579   
## MAKETOYOTA:USAGEOwn Goods 4.18e-14 \*\*\*  
## MAKEVOLVO:USAGEOwn Goods 3.26e-07 \*\*\*  
## MAKEDAEWOO:USAGEOwn service 0.272084   
## MAKEFIAT:USAGEOwn service 0.444927   
## MAKEFORD:USAGEOwn service 0.109900   
## MAKEGEELY:USAGEOwn service NA   
## MAKEGENLYON:USAGEOwn service NA   
## MAKEHYUNDAI:USAGEOwn service 0.023213 \*   
## MAKEISUZU:USAGEOwn service 3.73e-12 \*\*\*  
## MAKEIVECO:USAGEOwn service NA   
## MAKELIFAN:USAGEOwn service 0.000278 \*\*\*  
## MAKEMAZDA:USAGEOwn service 0.123046   
## MAKEMERCEDES:USAGEOwn service 0.622133   
## MAKEMITSUBISHI:USAGEOwn service 5.73e-05 \*\*\*  
## MAKENISSAN:USAGEOwn service 0.007674 \*\*   
## MAKERENAULT:USAGEOwn service NA   
## MAKESINO:USAGEOwn service NA   
## MAKESINO HOWO:USAGEOwn service NA   
## MAKESUZUKI:USAGEOwn service 0.038757 \*   
## MAKETATA:USAGEOwn service 0.930836   
## MAKETOYOTA:USAGEOwn service 8.28e-05 \*\*\*  
## MAKEVOLVO:USAGEOwn service 0.840406   
## MAKEDAEWOO:USAGEPrivate NA   
## MAKEFIAT:USAGEPrivate NA   
## MAKEFORD:USAGEPrivate 0.387951   
## MAKEGEELY:USAGEPrivate NA   
## MAKEGENLYON:USAGEPrivate NA   
## MAKEHYUNDAI:USAGEPrivate NA   
## MAKEISUZU:USAGEPrivate NA   
## MAKEIVECO:USAGEPrivate NA   
## MAKELIFAN:USAGEPrivate NA   
## MAKEMAZDA:USAGEPrivate NA   
## MAKEMERCEDES:USAGEPrivate 0.533308   
## MAKEMITSUBISHI:USAGEPrivate NA   
## MAKENISSAN:USAGEPrivate NA   
## MAKERENAULT:USAGEPrivate NA   
## MAKESINO:USAGEPrivate NA   
## MAKESINO HOWO:USAGEPrivate NA   
## MAKESUZUKI:USAGEPrivate NA   
## MAKETATA:USAGEPrivate NA   
## MAKETOYOTA:USAGEPrivate NA   
## MAKEVOLVO:USAGEPrivate NA   
## USAGECar Hires:TYPE\_VEHICLEBus 0.265778   
## USAGEFare Paying Passengers:TYPE\_VEHICLEBus NA   
## USAGEGeneral Cartage:TYPE\_VEHICLEBus NA   
## USAGEOwn Goods:TYPE\_VEHICLEBus 0.664624   
## USAGEOwn service:TYPE\_VEHICLEBus 0.170131   
## USAGEPrivate:TYPE\_VEHICLEBus 0.215318   
## USAGECar Hires:TYPE\_VEHICLEMotor-cycle NA   
## USAGEFare Paying Passengers:TYPE\_VEHICLEMotor-cycle NA   
## USAGEGeneral Cartage:TYPE\_VEHICLEMotor-cycle 0.968987   
## USAGEOwn Goods:TYPE\_VEHICLEMotor-cycle 0.378425   
## USAGEOwn service:TYPE\_VEHICLEMotor-cycle 0.002571 \*\*   
## USAGEPrivate:TYPE\_VEHICLEMotor-cycle NA   
## USAGECar Hires:TYPE\_VEHICLEPick-up 0.435069   
## USAGEFare Paying Passengers:TYPE\_VEHICLEPick-up NA   
## USAGEGeneral Cartage:TYPE\_VEHICLEPick-up 0.852514   
## USAGEOwn Goods:TYPE\_VEHICLEPick-up 0.822009   
## USAGEOwn service:TYPE\_VEHICLEPick-up 0.475956   
## USAGEPrivate:TYPE\_VEHICLEPick-up 0.386431   
## USAGECar Hires:TYPE\_VEHICLESpecial construction NA   
## USAGEFare Paying Passengers:TYPE\_VEHICLESpecial construction NA   
## USAGEGeneral Cartage:TYPE\_VEHICLESpecial construction NA   
## USAGEOwn Goods:TYPE\_VEHICLESpecial construction 0.697921   
## USAGEOwn service:TYPE\_VEHICLESpecial construction 0.564344   
## USAGEPrivate:TYPE\_VEHICLESpecial construction NA   
## USAGECar Hires:TYPE\_VEHICLEStation Wagones 0.560891   
## USAGEFare Paying Passengers:TYPE\_VEHICLEStation Wagones NA   
## USAGEGeneral Cartage:TYPE\_VEHICLEStation Wagones NA   
## USAGEOwn Goods:TYPE\_VEHICLEStation Wagones NA   
## USAGEOwn service:TYPE\_VEHICLEStation Wagones 0.445580   
## USAGEPrivate:TYPE\_VEHICLEStation Wagones 0.059534 .   
## USAGECar Hires:TYPE\_VEHICLETanker NA   
## USAGEFare Paying Passengers:TYPE\_VEHICLETanker NA   
## USAGEGeneral Cartage:TYPE\_VEHICLETanker 4.72e-10 \*\*\*  
## USAGEOwn Goods:TYPE\_VEHICLETanker NA   
## USAGEOwn service:TYPE\_VEHICLETanker NA   
## USAGEPrivate:TYPE\_VEHICLETanker NA   
## USAGECar Hires:TYPE\_VEHICLETrailers and semitrailers NA   
## USAGEFare Paying Passengers:TYPE\_VEHICLETrailers and semitrailers NA   
## USAGEGeneral Cartage:TYPE\_VEHICLETrailers and semitrailers 0.024962 \*   
## USAGEOwn Goods:TYPE\_VEHICLETrailers and semitrailers NA   
## USAGEOwn service:TYPE\_VEHICLETrailers and semitrailers NA   
## USAGEPrivate:TYPE\_VEHICLETrailers and semitrailers NA   
## USAGECar Hires:TYPE\_VEHICLETruck NA   
## USAGEFare Paying Passengers:TYPE\_VEHICLETruck NA   
## USAGEGeneral Cartage:TYPE\_VEHICLETruck NA   
## USAGEOwn Goods:TYPE\_VEHICLETruck NA   
## USAGEOwn service:TYPE\_VEHICLETruck NA   
## USAGEPrivate:TYPE\_VEHICLETruck NA   
## USAGECar Hires:LOG\_INSURED\_VALUE 0.993067   
## USAGEFare Paying Passengers:LOG\_INSURED\_VALUE 0.127618   
## USAGEGeneral Cartage:LOG\_INSURED\_VALUE < 2e-16 \*\*\*  
## USAGEOwn Goods:LOG\_INSURED\_VALUE 0.002816 \*\*   
## USAGEOwn service:LOG\_INSURED\_VALUE 0.147697   
## USAGEPrivate:LOG\_INSURED\_VALUE 8.04e-10 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.3862 on 194255 degrees of freedom  
## Multiple R-squared: 0.8099, Adjusted R-squared: 0.8098   
## F-statistic: 6132 on 135 and 194255 DF, p-value: < 2.2e-16

# Residuenplots für das neue Modell  
par(mfrow = c(2, 2))  
plot(lm\_model\_interactions)

## Warning: not plotting observations with leverage one:  
## 52401, 57247, 142233, 146420, 169037, 184693

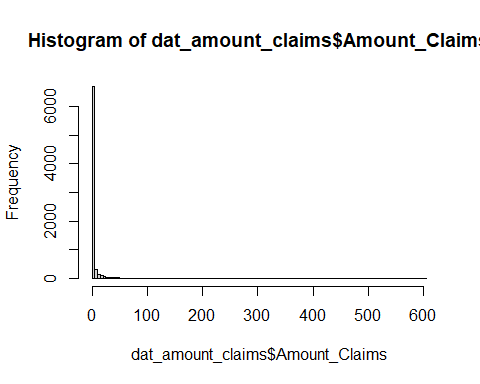
## Warning in sqrt(crit \* p \* (1 - hh)/hh): NaNs produced  
## Warning in sqrt(crit \* p \* (1 - hh)/hh): NaNs produced



Obwohl die Hinzufügung von Interaktionen das Modell in einigen Aspekten verbessert hat, zeigen die Residuenanalysen, dass Heteroskedastizität und Ausreisser weiterhin Probleme darstellen. Mögliche nächste Schritte könnten die Überprüfung und Behandlung dieser einflussreichen Punkte sowie die Anwendung robusterer Regressionstechniken oder zusätzlicher Transformationen umfassen. Darüber hinaus könnte die Untersuchung weiterer Interaktionen oder die Modellierung nichtlinearer Effekte in Betracht gezogen werden, um die Anpassung weiter zu optimieren.

## Poisson

#Data preparation: Aggregate the number of claims per combination of vehicle type, insurance type, and production year  
dat\_amount\_claims <- clean\_dat\_motor %>%  
 group\_by(SEX, TYPE\_VEHICLE, INSR\_TYPE, MAKE, AGE\_VEHICLE, SEATS\_NUM) %>%  
 summarise(Amount\_Claims = sum(CLAIM\_PAID == "YES"), .groups = 'drop')  
  
hist(dat\_amount\_claims$Amount\_Claims, breaks=100)



mean(dat\_amount\_claims$Amount\_Claims) # calculate mean

## [1] 3.739298

var(dat\_amount\_claims$Amount\_Claims)

## [1] 339.3752

#The variance is much greater than the mean, which suggests that we will have over-dispersion in the model.  
  
#Poisson Regression:  
poisson\_model<- glm(Amount\_Claims ~ SEX + TYPE\_VEHICLE + INSR\_TYPE + MAKE + AGE\_VEHICLE + SEATS\_NUM,   
 family = poisson(link = "log"), data = dat\_amount\_claims)  
  
#Summary of the model  
summary(poisson\_model)

##   
## Call:  
## glm(formula = Amount\_Claims ~ SEX + TYPE\_VEHICLE + INSR\_TYPE +   
## MAKE + AGE\_VEHICLE + SEATS\_NUM, family = poisson(link = "log"),   
## data = dat\_amount\_claims)  
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)  
## (Intercept) 2.2281318 0.0432024 51.574 < 2e-16  
## SEXMale -0.2432338 0.0138801 -17.524 < 2e-16  
## SEXFemale -1.1411975 0.0257931 -44.244 < 2e-16  
## TYPE\_VEHICLEBus -0.7778781 0.0304509 -25.545 < 2e-16  
## TYPE\_VEHICLEMotor-cycle -3.5981176 0.3188737 -11.284 < 2e-16  
## TYPE\_VEHICLEPick-up 0.6061555 0.0252585 23.998 < 2e-16  
## TYPE\_VEHICLESpecial construction -1.6118328 0.1284428 -12.549 < 2e-16  
## TYPE\_VEHICLEStation Wagones -0.4477489 0.0225085 -19.892 < 2e-16  
## TYPE\_VEHICLETanker -0.9508582 0.0693549 -13.710 < 2e-16  
## TYPE\_VEHICLETrailers and semitrailers -2.1708040 0.2457193 -8.834 < 2e-16  
## TYPE\_VEHICLETruck 0.7332989 0.0321265 22.825 < 2e-16  
## INSR\_TYPEPrivate 0.0662338 0.0205805 3.218 0.001290  
## MAKEDAEWOO -1.1488748 0.0713240 -16.108 < 2e-16  
## MAKEFIAT -1.4389104 0.0851633 -16.896 < 2e-16  
## MAKEFORD -0.9348718 0.0601163 -15.551 < 2e-16  
## MAKEGEELY 0.3357268 0.0710123 4.728 2.27e-06  
## MAKEGENLYON -0.6643351 0.0977696 -6.795 1.08e-11  
## MAKEHYUNDAI -1.1222073 0.0615915 -18.220 < 2e-16  
## MAKEISUZU 0.4856787 0.0400083 12.139 < 2e-16  
## MAKEIVECO -0.0163640 0.0474467 -0.345 0.730175  
## MAKELIFAN 0.1897465 0.0559560 3.391 0.000696  
## MAKEMAZDA 0.1004829 0.0632107 1.590 0.111914  
## MAKEMERCEDES -1.5241481 0.0737332 -20.671 < 2e-16  
## MAKEMITSUBISHI -0.5359270 0.0462472 -11.588 < 2e-16  
## MAKENISSAN -0.2482406 0.0414576 -5.988 2.13e-09  
## MAKERENAULT -1.1880808 0.0964800 -12.314 < 2e-16  
## MAKESINO -0.7086252 0.0899356 -7.879 3.29e-15  
## MAKESINO HOWO 0.1379077 0.0503280 2.740 0.006141  
## MAKESUZUKI -1.0851625 0.0933207 -11.628 < 2e-16  
## MAKETATA -1.6824072 0.0822529 -20.454 < 2e-16  
## MAKETOYOTA 0.7280048 0.0355472 20.480 < 2e-16  
## MAKEVOLVO -1.4216245 0.1246655 -11.404 < 2e-16  
## AGE\_VEHICLE -0.0807835 0.0008416 -95.986 < 2e-16  
## SEATS\_NUM 0.0069654 0.0006733 10.345 < 2e-16  
##   
## (Intercept) \*\*\*  
## SEXMale \*\*\*  
## SEXFemale \*\*\*  
## TYPE\_VEHICLEBus \*\*\*  
## TYPE\_VEHICLEMotor-cycle \*\*\*  
## TYPE\_VEHICLEPick-up \*\*\*  
## TYPE\_VEHICLESpecial construction \*\*\*  
## TYPE\_VEHICLEStation Wagones \*\*\*  
## TYPE\_VEHICLETanker \*\*\*  
## TYPE\_VEHICLETrailers and semitrailers \*\*\*  
## TYPE\_VEHICLETruck \*\*\*  
## INSR\_TYPEPrivate \*\*   
## MAKEDAEWOO \*\*\*  
## MAKEFIAT \*\*\*  
## MAKEFORD \*\*\*  
## MAKEGEELY \*\*\*  
## MAKEGENLYON \*\*\*  
## MAKEHYUNDAI \*\*\*  
## MAKEISUZU \*\*\*  
## MAKEIVECO   
## MAKELIFAN \*\*\*  
## MAKEMAZDA   
## MAKEMERCEDES \*\*\*  
## MAKEMITSUBISHI \*\*\*  
## MAKENISSAN \*\*\*  
## MAKERENAULT \*\*\*  
## MAKESINO \*\*\*  
## MAKESINO HOWO \*\*   
## MAKESUZUKI \*\*\*  
## MAKETATA \*\*\*  
## MAKETOYOTA \*\*\*  
## MAKEVOLVO \*\*\*  
## AGE\_VEHICLE \*\*\*  
## SEATS\_NUM \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for poisson family taken to be 1)  
##   
## Null deviance: 120841 on 7544 degrees of freedom  
## Residual deviance: 89096 on 7511 degrees of freedom  
## AIC: 99127  
##   
## Number of Fisher Scoring iterations: 7

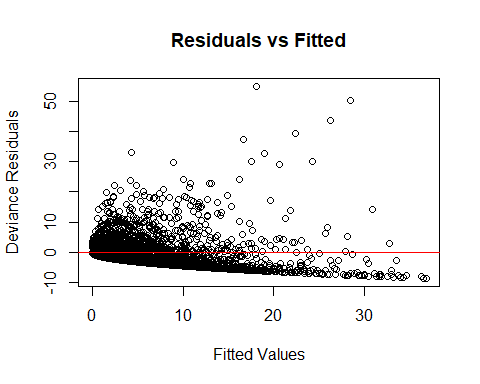
#Coeff  
coef(poisson\_model)

## (Intercept) SEXMale   
## 2.228131804 -0.243233833   
## SEXFemale TYPE\_VEHICLEBus   
## -1.141197538 -0.777878063   
## TYPE\_VEHICLEMotor-cycle TYPE\_VEHICLEPick-up   
## -3.598117557 0.606155519   
## TYPE\_VEHICLESpecial construction TYPE\_VEHICLEStation Wagones   
## -1.611832819 -0.447748898   
## TYPE\_VEHICLETanker TYPE\_VEHICLETrailers and semitrailers   
## -0.950858164 -2.170803990   
## TYPE\_VEHICLETruck INSR\_TYPEPrivate   
## 0.733298948 0.066233846   
## MAKEDAEWOO MAKEFIAT   
## -1.148874804 -1.438910440   
## MAKEFORD MAKEGEELY   
## -0.934871798 0.335726780   
## MAKEGENLYON MAKEHYUNDAI   
## -0.664335100 -1.122207252   
## MAKEISUZU MAKEIVECO   
## 0.485678714 -0.016363985   
## MAKELIFAN MAKEMAZDA   
## 0.189746546 0.100482861   
## MAKEMERCEDES MAKEMITSUBISHI   
## -1.524148075 -0.535926951   
## MAKENISSAN MAKERENAULT   
## -0.248240601 -1.188080790   
## MAKESINO MAKESINO HOWO   
## -0.708625190 0.137907673   
## MAKESUZUKI MAKETATA   
## -1.085162469 -1.682407233   
## MAKETOYOTA MAKEVOLVO   
## 0.728004752 -1.421624510   
## AGE\_VEHICLE SEATS\_NUM   
## -0.080783471 0.006965433

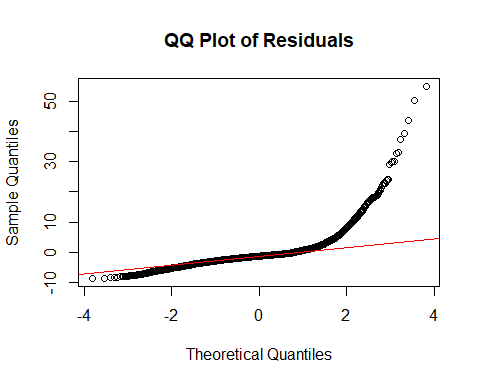
exp(coef(poisson\_model))

## (Intercept) SEXMale   
## 9.28250832 0.78408815   
## SEXFemale TYPE\_VEHICLEBus   
## 0.31943626 0.45937975   
## TYPE\_VEHICLEMotor-cycle TYPE\_VEHICLEPick-up   
## 0.02737521 1.83336948   
## TYPE\_VEHICLESpecial construction TYPE\_VEHICLEStation Wagones   
## 0.19952159 0.63906513   
## TYPE\_VEHICLETanker TYPE\_VEHICLETrailers and semitrailers   
## 0.38640928 0.11408586   
## TYPE\_VEHICLETruck INSR\_TYPEPrivate   
## 2.08193749 1.06847655   
## MAKEDAEWOO MAKEFIAT   
## 0.31699325 0.23718605   
## MAKEFORD MAKEGEELY   
## 0.39263620 1.39895675   
## MAKEGENLYON MAKEHYUNDAI   
## 0.51461558 0.32556041   
## MAKEISUZU MAKEIVECO   
## 1.62527773 0.98376918   
## MAKELIFAN MAKEMAZDA   
## 1.20894315 1.10570469   
## MAKEMERCEDES MAKEMITSUBISHI   
## 0.21780653 0.58512665   
## MAKENISSAN MAKERENAULT   
## 0.78017221 0.30480569   
## MAKESINO MAKESINO HOWO   
## 0.49232058 1.14786957   
## MAKESUZUKI MAKETATA   
## 0.33784689 0.18592587   
## MAKETOYOTA MAKEVOLVO   
## 2.07094444 0.24132167   
## AGE\_VEHICLE SEATS\_NUM   
## 0.92239340 1.00698975

# Model diagnostics and evaluation  
# Residuals vs. Fitted Plot  
plot(poisson\_model$fitted.values, residuals(poisson\_model, type = "deviance"),  
 xlab = "Fitted Values",   
 ylab = "Deviance Residuals",   
 main = "Residuals vs Fitted")  
abline(h = 0, col = "red")



# QQ Plot of residuals (Check for normal distribution)  
qqnorm(residuals(poisson\_model, type = "deviance"), main = "QQ Plot of Residuals")  
qqline(residuals(poisson\_model, type = "deviance"), col = "red")



#Diagnose overdispersion  
deviance(poisson\_model) / df.residual(poisson\_model)

## [1] 11.86207

The residuals vs. fitted diagram shows an increasing dispersion of the residuals with increasing estimated values, which indicates heteroscedasticity. This indicates that the variance of the residuals is not constant and supports the assumption of overdispersion, as the variance is significantly greater than the mean. The QQ plot of the residuals shows a strong deviation from the theoretical normal distribution, especially at the ends, which indicates a lack of normal distribution of the residuals. This deviation is typical for Poisson models, but the significant differences indicate overdispersion or possibly missing predictors. The calculation of the ratio of deviance to degrees of freedom (11.86) confirms the overdispersion, as this value is significantly greater than 1 and thus indicates a higher variance in the data than assumed in the Poisson model.

The significance of the predictors in the model becomes clear from the p-values of the coefficient estimates. Almost all predictors show extremely low p-values (p < 0.001), which indicates that they have a significant influence on the number of claims. This high significance strengthens the significance of the correlations found. Nevertheless, the overdispersion remains problematic, as it indicates that the Poisson model may underestimate the variance in the data. Despite the statistical significance of the predictors, a model change, e.g. to a quassi Poisson model or a negative binomial model, could be necessary to improve the model quality and to address the overdispersion appropriately.

### Massnahme 1): Quasi-Poisson-Regression

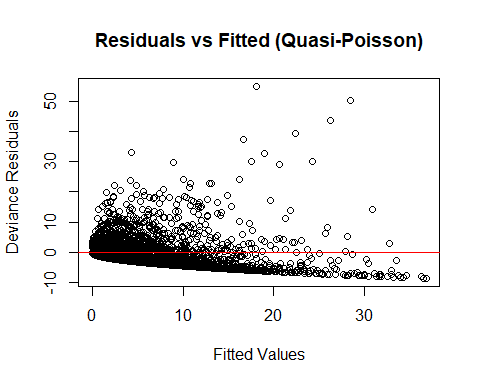
#Quasi-Poisson  
quasi\_poisson\_model <- glm(Amount\_Claims ~ SEX + TYPE\_VEHICLE + INSR\_TYPE + MAKE + AGE\_VEHICLE + SEATS\_NUM,   
 family = quasipoisson, data = dat\_amount\_claims)  
  
# Summary of the model  
summary(quasi\_poisson\_model)

##   
## Call:  
## glm(formula = Amount\_Claims ~ SEX + TYPE\_VEHICLE + INSR\_TYPE +   
## MAKE + AGE\_VEHICLE + SEATS\_NUM, family = quasipoisson, data = dat\_amount\_claims)  
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 2.228132 0.235701 9.453 < 2e-16 \*\*\*  
## SEXMale -0.243234 0.075726 -3.212 0.001324 \*\*   
## SEXFemale -1.141198 0.140720 -8.110 5.88e-16 \*\*\*  
## TYPE\_VEHICLEBus -0.777878 0.166132 -4.682 2.89e-06 \*\*\*  
## TYPE\_VEHICLEMotor-cycle -3.598118 1.739689 -2.068 0.038650 \*   
## TYPE\_VEHICLEPick-up 0.606156 0.137803 4.399 1.10e-05 \*\*\*  
## TYPE\_VEHICLESpecial construction -1.611833 0.700749 -2.300 0.021467 \*   
## TYPE\_VEHICLEStation Wagones -0.447749 0.122801 -3.646 0.000268 \*\*\*  
## TYPE\_VEHICLETanker -0.950858 0.378382 -2.513 0.011993 \*   
## TYPE\_VEHICLETrailers and semitrailers -2.170804 1.340578 -1.619 0.105424   
## TYPE\_VEHICLETruck 0.733299 0.175273 4.184 2.90e-05 \*\*\*  
## INSR\_TYPEPrivate 0.066234 0.112282 0.590 0.555283   
## MAKEDAEWOO -1.148875 0.389125 -2.952 0.003162 \*\*   
## MAKEFIAT -1.438910 0.464628 -3.097 0.001963 \*\*   
## MAKEFORD -0.934872 0.327978 -2.850 0.004378 \*\*   
## MAKEGEELY 0.335727 0.387424 0.867 0.386210   
## MAKEGENLYON -0.664335 0.533404 -1.245 0.213001   
## MAKEHYUNDAI -1.122207 0.336026 -3.340 0.000843 \*\*\*  
## MAKEISUZU 0.485679 0.218274 2.225 0.026105 \*   
## MAKEIVECO -0.016364 0.258856 -0.063 0.949596   
## MAKELIFAN 0.189747 0.305281 0.622 0.534258   
## MAKEMAZDA 0.100483 0.344860 0.291 0.770774   
## MAKEMERCEDES -1.524148 0.402269 -3.789 0.000153 \*\*\*  
## MAKEMITSUBISHI -0.535927 0.252312 -2.124 0.033697 \*   
## MAKENISSAN -0.248241 0.226181 -1.098 0.272445   
## MAKERENAULT -1.188081 0.526369 -2.257 0.024029 \*   
## MAKESINO -0.708625 0.490664 -1.444 0.148720   
## MAKESINO HOWO 0.137908 0.274576 0.502 0.615501   
## MAKESUZUKI -1.085162 0.509133 -2.131 0.033089 \*   
## MAKETATA -1.682407 0.448750 -3.749 0.000179 \*\*\*  
## MAKETOYOTA 0.728005 0.193936 3.754 0.000175 \*\*\*  
## MAKEVOLVO -1.421625 0.680141 -2.090 0.036634 \*   
## AGE\_VEHICLE -0.080783 0.004592 -17.594 < 2e-16 \*\*\*  
## SEATS\_NUM 0.006965 0.003674 1.896 0.057985 .   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for quasipoisson family taken to be 29.765)  
##   
## Null deviance: 120841 on 7544 degrees of freedom  
## Residual deviance: 89096 on 7511 degrees of freedom  
## AIC: NA  
##   
## Number of Fisher Scoring iterations: 7

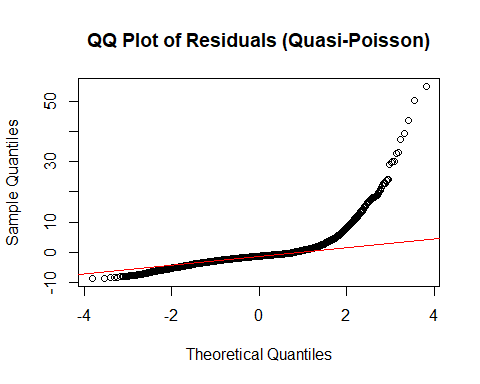
#F-Test, Vergleiche von geschachtelten Quasi-Likelihood-Modellen (Overdispersion)  
drop1(quasi\_poisson\_model, test= "F")

## Single term deletions  
##   
## Model:  
## Amount\_Claims ~ SEX + TYPE\_VEHICLE + INSR\_TYPE + MAKE + AGE\_VEHICLE +   
## SEATS\_NUM  
## Df Deviance F value Pr(>F)   
## <none> 89096   
## SEX 2 91759 112.235 < 2.2e-16 \*\*\*  
## TYPE\_VEHICLE 8 96223 75.102 < 2.2e-16 \*\*\*  
## INSR\_TYPE 1 89106 0.871 0.350707   
## MAKE 20 99860 45.371 < 2.2e-16 \*\*\*  
## AGE\_VEHICLE 1 100388 951.953 < 2.2e-16 \*\*\*  
## SEATS\_NUM 1 89191 8.013 0.004657 \*\*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

# Model diagnostics: Residuals vs. Fitted Plot  
plot(quasi\_poisson\_model$fitted.values, residuals(quasi\_poisson\_model, type = "deviance"),  
 xlab = "Fitted Values",   
 ylab = "Deviance Residuals",   
 main = "Residuals vs Fitted (Quasi-Poisson)")  
abline(h = 0, col = "red")



# QQ plot of residuals (checking normality)  
qqnorm(residuals(quasi\_poisson\_model, type = "deviance"), main = "QQ Plot of Residuals (Quasi-Poisson)")  
qqline(residuals(quasi\_poisson\_model, type = "deviance"), col = "red")

 Switching to a quasi-Poisson model to account for overdispersion led to improvements compared to the original Poisson model. The residuals vs. fitted plot shows a reduced dispersion of the residuals at higher estimated values, which indicates a better fit of the variance, although heteroscedasticity still exists. The QQ plot of the residuals shows an improved fit to the theoretical normal distribution, especially in the middle range, while deviations at the edges remain, indicating extreme values or modelling errors. By adjusting the dispersion parameter (29.765) in the quasi-Poisson model, the increased variance compared to the Poisson model is adequately taken into account. The F-test confirms the significance of the variables ‘SEX’, ‘TYPE\_VEHICLE’, ‘MAKE’, ‘AGE\_VEHICLE’ and ‘SEATS\_NUM’. Despite these improvements, there are still slight anomalies in the residuals.

### Massnahme 2): Negative-Binomial-Modell

# Negative-Binomial-Regression anpassen  
neg\_bin\_model <- glm.nb(Amount\_Claims ~ SEX + TYPE\_VEHICLE + INSR\_TYPE + MAKE + AGE\_VEHICLE + SEATS\_NUM,  
 data = dat\_amount\_claims)  
  
# Zusammenfassung des Modells anzeigen  
summary(neg\_bin\_model)

##   
## Call:  
## glm.nb(formula = Amount\_Claims ~ SEX + TYPE\_VEHICLE + INSR\_TYPE +   
## MAKE + AGE\_VEHICLE + SEATS\_NUM, data = dat\_amount\_claims,   
## init.theta = 0.2601478985, link = log)  
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) 1.593643 0.204559 7.791 6.67e-15 \*\*\*  
## SEXMale -0.294409 0.059278 -4.967 6.81e-07 \*\*\*  
## SEXFemale -1.044152 0.082296 -12.688 < 2e-16 \*\*\*  
## TYPE\_VEHICLEBus -0.199114 0.126434 -1.575 0.115291   
## TYPE\_VEHICLEMotor-cycle -3.218031 0.381609 -8.433 < 2e-16 \*\*\*  
## TYPE\_VEHICLEPick-up 0.905085 0.121015 7.479 7.48e-14 \*\*\*  
## TYPE\_VEHICLESpecial construction -0.975796 0.346122 -2.819 0.004814 \*\*   
## TYPE\_VEHICLEStation Wagones -0.781258 0.089660 -8.714 < 2e-16 \*\*\*  
## TYPE\_VEHICLETanker -0.286033 0.198260 -1.443 0.149101   
## TYPE\_VEHICLETrailers and semitrailers -1.003898 0.444635 -2.258 0.023958 \*   
## TYPE\_VEHICLETruck 1.409117 0.137570 10.243 < 2e-16 \*\*\*  
## INSR\_TYPEPrivate 0.881702 0.098587 8.943 < 2e-16 \*\*\*  
## MAKEDAEWOO -0.964944 0.232020 -4.159 3.20e-05 \*\*\*  
## MAKEFIAT -0.863216 0.227030 -3.802 0.000143 \*\*\*  
## MAKEFORD -0.672220 0.228257 -2.945 0.003229 \*\*   
## MAKEGEELY 0.026850 0.375649 0.071 0.943018   
## MAKEGENLYON -0.689233 0.424798 -1.622 0.104697   
## MAKEHYUNDAI -1.096772 0.214097 -5.123 3.01e-07 \*\*\*  
## MAKEISUZU 0.355738 0.189938 1.873 0.061081 .   
## MAKEIVECO -0.032197 0.207512 -0.155 0.876696   
## MAKELIFAN -0.152712 0.278393 -0.549 0.583317   
## MAKEMAZDA 0.005307 0.322169 0.016 0.986858   
## MAKEMERCEDES -1.221596 0.205633 -5.941 2.84e-09 \*\*\*  
## MAKEMITSUBISHI -0.273297 0.192029 -1.423 0.154676   
## MAKENISSAN -0.098075 0.185962 -0.527 0.597922   
## MAKERENAULT -1.532173 0.290690 -5.271 1.36e-07 \*\*\*  
## MAKESINO -0.583714 0.377357 -1.547 0.121900   
## MAKESINO HOWO -0.007579 0.254208 -0.030 0.976217   
## MAKESUZUKI -0.726496 0.244407 -2.972 0.002954 \*\*   
## MAKETATA -1.392154 0.234004 -5.949 2.69e-09 \*\*\*  
## MAKETOYOTA 1.018970 0.174856 5.827 5.63e-09 \*\*\*  
## MAKEVOLVO -1.250280 0.308946 -4.047 5.19e-05 \*\*\*  
## AGE\_VEHICLE -0.092176 0.002967 -31.062 < 2e-16 \*\*\*  
## SEATS\_NUM 0.012462 0.002582 4.826 1.39e-06 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for Negative Binomial(0.2601) family taken to be 1)  
##   
## Null deviance: 8491.0 on 7544 degrees of freedom  
## Residual deviance: 6177.4 on 7511 degrees of freedom  
## AIC: 26178  
##   
## Number of Fisher Scoring iterations: 1  
##   
##   
## Theta: 0.26015   
## Std. Err.: 0.00602   
##   
## 2 x log-likelihood: -26108.28300

drop1(neg\_bin\_model, test= "LRT")

## Single term deletions  
##   
## Model:  
## Amount\_Claims ~ SEX + TYPE\_VEHICLE + INSR\_TYPE + MAKE + AGE\_VEHICLE +   
## SEATS\_NUM  
## Df Deviance AIC LRT Pr(>Chi)   
## <none> 6177.4 26176   
## SEX 2 6304.5 26299 127.05 < 2.2e-16 \*\*\*  
## TYPE\_VEHICLE 8 6651.9 26635 474.52 < 2.2e-16 \*\*\*  
## INSR\_TYPE 1 6203.5 26200 26.08 3.275e-07 \*\*\*  
## MAKE 20 6869.3 26828 691.90 < 2.2e-16 \*\*\*  
## AGE\_VEHICLE 1 6899.5 26896 722.09 < 2.2e-16 \*\*\*  
## SEATS\_NUM 1 6192.8 26190 15.39 8.756e-05 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

# Exponentierte Koeffizienten anzeigen (um Rate Ratios zu interpretieren)  
exp(coef(neg\_bin\_model))

## (Intercept) SEXMale   
## 4.9216437 0.7449718   
## SEXFemale TYPE\_VEHICLEBus   
## 0.3519901 0.8194563   
## TYPE\_VEHICLEMotor-cycle TYPE\_VEHICLEPick-up   
## 0.0400338 2.4721418   
## TYPE\_VEHICLESpecial construction TYPE\_VEHICLEStation Wagones   
## 0.3768922 0.4578298   
## TYPE\_VEHICLETanker TYPE\_VEHICLETrailers and semitrailers   
## 0.7512378 0.3664483   
## TYPE\_VEHICLETruck INSR\_TYPEPrivate   
## 4.0923385 2.4150066   
## MAKEDAEWOO MAKEFIAT   
## 0.3810046 0.4218033   
## MAKEFORD MAKEGEELY   
## 0.5105737 1.0272140   
## MAKEGENLYON MAKEHYUNDAI   
## 0.5019608 0.3339475   
## MAKEISUZU MAKEIVECO   
## 1.4272333 0.9683155   
## MAKELIFAN MAKEMAZDA   
## 0.8583773 1.0053206   
## MAKEMERCEDES MAKEMITSUBISHI   
## 0.2947592 0.7608666   
## MAKENISSAN MAKERENAULT   
## 0.9065813 0.2160656   
## MAKESINO MAKESINO HOWO   
## 0.5578228 0.9924501   
## MAKESUZUKI MAKETATA   
## 0.4836007 0.2485393   
## MAKETOYOTA MAKEVOLVO   
## 2.7703401 0.2864245   
## AGE\_VEHICLE SEATS\_NUM   
## 0.9119450 1.0125399

# Konfidenzintervalle für die Koeffizienten  
confint(neg\_bin\_model)

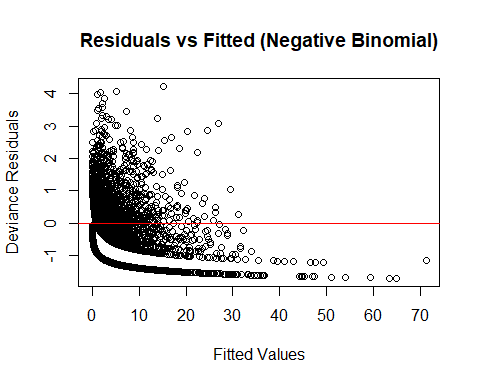
## Waiting for profiling to be done...

## 2.5 % 97.5 %  
## (Intercept) 1.109903018 2.09581017  
## SEXMale -0.427256174 -0.16085729  
## SEXFemale -1.216096881 -0.86980068  
## TYPE\_VEHICLEBus -0.572419236 0.16496590  
## TYPE\_VEHICLEMotor-cycle -4.059047296 -2.44835065  
## TYPE\_VEHICLEPick-up 0.534603259 1.26631305  
## TYPE\_VEHICLESpecial construction -1.702206728 -0.18540763  
## TYPE\_VEHICLEStation Wagones -0.969195254 -0.59410952  
## TYPE\_VEHICLETanker -0.766645736 0.19250704  
## TYPE\_VEHICLETrailers and semitrailers -1.985575796 0.02760646  
## TYPE\_VEHICLETruck 1.013923087 1.79667945  
## INSR\_TYPEPrivate 0.553873879 1.19751197  
## MAKEDAEWOO -1.421685269 -0.50995839  
## MAKEFIAT -1.334550638 -0.40105416  
## MAKEFORD -1.133602262 -0.21623040  
## MAKEGEELY -0.674972955 0.81798326  
## MAKEGENLYON -1.472210390 0.21856631  
## MAKEHYUNDAI -1.530358970 -0.67482034  
## MAKEISUZU -0.036598694 0.73073681  
## MAKEIVECO -0.461399491 0.38454203  
## MAKELIFAN -0.682643504 0.39015911  
## MAKEMAZDA -0.598410117 0.65261867  
## MAKEMERCEDES -1.642465393 -0.81386862  
## MAKEMITSUBISHI -0.666665363 0.10187698  
## MAKENISSAN -0.481264369 0.26463140  
## MAKERENAULT -2.099764795 -0.95249221  
## MAKESINO -1.293733257 0.21400598  
## MAKESINO HOWO -0.510782738 0.50057380  
## MAKESUZUKI -1.238494529 -0.21784589  
## MAKETATA -1.862425424 -0.92331124  
## MAKETOYOTA 0.650527595 1.36342727  
## MAKEVOLVO -1.857773898 -0.62793362  
## AGE\_VEHICLE -0.098771506 -0.08558427  
## SEATS\_NUM 0.006103731 0.01905058

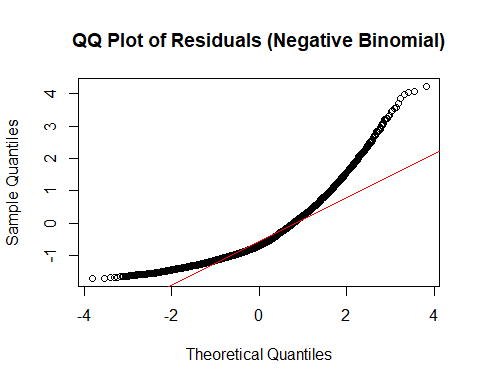
# Devianz und Freiheitsgrade überprüfen (sollte näher an 1 liegen)  
deviance(neg\_bin\_model) / df.residual(neg\_bin\_model)

## [1] 0.8224468

# Residuals vs Fitted Plot  
plot(neg\_bin\_model$fitted.values, residuals(neg\_bin\_model, type = "deviance"),  
 xlab = "Fitted Values",   
 ylab = "Deviance Residuals",   
 main = "Residuals vs Fitted (Negative Binomial)")  
abline(h = 0, col = "red")



# QQ-Plot der Residuen  
qqnorm(residuals(neg\_bin\_model, type = "deviance"), main = "QQ Plot of Residuals (Negative Binomial)")  
qqline(residuals(neg\_bin\_model, type = "deviance"), col = "red")



The fitting of a negative binomial model shows significant improvements compared to the previous models. The residuals vs fitted plot indicates that the residuals are better distributed overall, especially at larger values of the fitted data, with lower heteroscedasticity. However, the QQ plot of the residuals still shows slight deviations from the normal distribution, especially at the extreme values, which indicates remaining modelling anomalies.

The model shows a significant improvement in terms of overdispersion, as suggested by the ratio of deviance to degrees of freedom (0.822), which is closer to 1 and thus significantly reduces overdispersion. Most predictors continue to show high levels of significance, indicating a strong explanatory power for the number of claims.

## Binomial

#Erstellen eines logistischen Regressionsmodells  
fit.binom <- glm(CLAIM\_PAID ~ SEX + INSR\_TYPE + INSURED\_VALUE + PREMIUM + AGE\_VEHICLE +   
 MAKE + SEATS\_NUM + CCM\_TON + USAGE,   
 data = clean\_dat\_motor,   
 family = binomial)  
  
# Zusammenfassung des Modells anzeigen  
summary(fit.binom)

##   
## Call:  
## glm(formula = CLAIM\_PAID ~ SEX + INSR\_TYPE + INSURED\_VALUE +   
## PREMIUM + AGE\_VEHICLE + MAKE + SEATS\_NUM + CCM\_TON + USAGE,   
## family = binomial, data = clean\_dat\_motor)  
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -1.350e+00 8.926e-02 -15.120 < 2e-16 \*\*\*  
## SEXMale -1.190e-01 2.151e-02 -5.532 3.17e-08 \*\*\*  
## SEXFemale -1.697e-01 3.084e-02 -5.502 3.75e-08 \*\*\*  
## INSR\_TYPEPrivate 5.218e-02 8.692e-02 0.600 0.548237   
## INSURED\_VALUE -2.047e-07 1.384e-08 -14.790 < 2e-16 \*\*\*  
## PREMIUM 2.623e-05 7.711e-07 34.019 < 2e-16 \*\*\*  
## AGE\_VEHICLE -3.019e-02 1.320e-03 -22.877 < 2e-16 \*\*\*  
## MAKEDAEWOO -4.336e-01 7.808e-02 -5.553 2.80e-08 \*\*\*  
## MAKEFIAT -5.887e-01 9.317e-02 -6.319 2.63e-10 \*\*\*  
## MAKEFORD -1.023e-01 6.749e-02 -1.515 0.129696   
## MAKEGEELY 1.281e-01 8.134e-02 1.575 0.115252   
## MAKEGENLYON -4.405e-01 1.095e-01 -4.023 5.74e-05 \*\*\*  
## MAKEHYUNDAI -3.234e-01 6.819e-02 -4.742 2.12e-06 \*\*\*  
## MAKEISUZU -6.294e-01 4.868e-02 -12.929 < 2e-16 \*\*\*  
## MAKEIVECO -5.805e-01 5.871e-02 -9.888 < 2e-16 \*\*\*  
## MAKELIFAN -5.149e-02 6.461e-02 -0.797 0.425524   
## MAKEMAZDA 1.990e-01 7.204e-02 2.762 0.005740 \*\*   
## MAKEMERCEDES -7.780e-01 7.934e-02 -9.806 < 2e-16 \*\*\*  
## MAKEMITSUBISHI -2.302e-01 5.281e-02 -4.359 1.31e-05 \*\*\*  
## MAKENISSAN -2.600e-01 4.841e-02 -5.372 7.80e-08 \*\*\*  
## MAKERENAULT -1.180e+00 1.065e-01 -11.086 < 2e-16 \*\*\*  
## MAKESINO -7.940e-01 9.869e-02 -8.046 8.57e-16 \*\*\*  
## MAKESINO HOWO -1.013e+00 5.780e-02 -17.521 < 2e-16 \*\*\*  
## MAKESUZUKI -1.584e+00 9.817e-02 -16.132 < 2e-16 \*\*\*  
## MAKETATA -6.405e-01 8.956e-02 -7.152 8.55e-13 \*\*\*  
## MAKETOYOTA -1.083e-01 4.288e-02 -2.525 0.011575 \*   
## MAKEVOLVO -1.101e+00 1.322e-01 -8.330 < 2e-16 \*\*\*  
## SEATS\_NUM 1.875e-03 8.974e-04 2.090 0.036637 \*   
## CCM\_TON -8.391e-08 3.675e-06 -0.023 0.981784   
## USAGECar Hires -6.619e-01 1.723e-01 -3.843 0.000122 \*\*\*  
## USAGEFare Paying Passengers -1.534e-01 8.239e-02 -1.862 0.062620 .   
## USAGEGeneral Cartage 1.097e-01 8.177e-02 1.342 0.179627   
## USAGEOwn Goods -1.428e-01 7.573e-02 -1.886 0.059263 .   
## USAGEOwn service -1.225e-01 7.946e-02 -1.542 0.123089   
## USAGEPrivate -6.342e-02 1.113e-01 -0.570 0.568776   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for binomial family taken to be 1)  
##   
## Null deviance: 165038 on 207655 degrees of freedom  
## Residual deviance: 160920 on 207621 degrees of freedom  
## AIC: 160990  
##   
## Number of Fisher Scoring iterations: 5

drop1(fit.binom, test= "LRT")

## Single term deletions  
##   
## Model:  
## CLAIM\_PAID ~ SEX + INSR\_TYPE + INSURED\_VALUE + PREMIUM + AGE\_VEHICLE +   
## MAKE + SEATS\_NUM + CCM\_TON + USAGE  
## Df Deviance AIC LRT Pr(>Chi)   
## <none> 160920 160990   
## SEX 2 160963 161029 43.70 3.239e-10 \*\*\*  
## INSR\_TYPE 1 160920 160988 0.36 0.5487   
## INSURED\_VALUE 1 161146 161214 226.28 < 2.2e-16 \*\*\*  
## PREMIUM 1 162075 162143 1155.11 < 2.2e-16 \*\*\*  
## AGE\_VEHICLE 1 161468 161536 548.35 < 2.2e-16 \*\*\*  
## MAKE 20 162166 162196 1246.50 < 2.2e-16 \*\*\*  
## SEATS\_NUM 1 160924 160992 4.31 0.0379 \*   
## CCM\_TON 1 160920 160988 0.00 0.9818   
## USAGE 6 161005 161063 85.11 3.133e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

INSR\_TYPE and CCM\_Ton do not appear to be significant. A new model is adapted without these two variables.

#Erstellen eines logistischen Regressionsmodells  
fit.binom2 <- glm(CLAIM\_PAID ~ SEX + INSURED\_VALUE + PREMIUM + AGE\_VEHICLE +   
 MAKE + SEATS\_NUM + USAGE,   
 data = clean\_dat\_motor,   
 family = binomial)  
  
# Zusammenfassung des Modells anzeigen  
summary(fit.binom2)

##   
## Call:  
## glm(formula = CLAIM\_PAID ~ SEX + INSURED\_VALUE + PREMIUM + AGE\_VEHICLE +   
## MAKE + SEATS\_NUM + USAGE, family = binomial, data = clean\_dat\_motor)  
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) -1.347e+00 8.862e-02 -15.201 < 2e-16 \*\*\*  
## SEXMale -1.186e-01 2.136e-02 -5.551 2.84e-08 \*\*\*  
## SEXFemale -1.692e-01 3.070e-02 -5.512 3.55e-08 \*\*\*  
## INSURED\_VALUE -2.046e-07 1.335e-08 -15.333 < 2e-16 \*\*\*  
## PREMIUM 2.623e-05 7.710e-07 34.016 < 2e-16 \*\*\*  
## AGE\_VEHICLE -3.019e-02 1.290e-03 -23.398 < 2e-16 \*\*\*  
## MAKEDAEWOO -4.344e-01 7.698e-02 -5.644 1.66e-08 \*\*\*  
## MAKEFIAT -5.898e-01 9.165e-02 -6.435 1.24e-10 \*\*\*  
## MAKEFORD -1.026e-01 6.739e-02 -1.522 0.127914   
## MAKEGEELY 1.283e-01 8.130e-02 1.579 0.114410   
## MAKEGENLYON -4.413e-01 1.089e-01 -4.054 5.04e-05 \*\*\*  
## MAKEHYUNDAI -3.237e-01 6.816e-02 -4.749 2.04e-06 \*\*\*  
## MAKEISUZU -6.299e-01 4.855e-02 -12.974 < 2e-16 \*\*\*  
## MAKEIVECO -5.815e-01 5.594e-02 -10.396 < 2e-16 \*\*\*  
## MAKELIFAN -5.142e-02 6.454e-02 -0.797 0.425593   
## MAKEMAZDA 1.986e-01 7.177e-02 2.767 0.005652 \*\*   
## MAKEMERCEDES -7.786e-01 7.914e-02 -9.838 < 2e-16 \*\*\*  
## MAKEMITSUBISHI -2.306e-01 5.275e-02 -4.371 1.24e-05 \*\*\*  
## MAKENISSAN -2.606e-01 4.823e-02 -5.404 6.50e-08 \*\*\*  
## MAKERENAULT -1.181e+00 1.058e-01 -11.162 < 2e-16 \*\*\*  
## MAKESINO -7.946e-01 9.806e-02 -8.103 5.35e-16 \*\*\*  
## MAKESINO HOWO -1.014e+00 5.629e-02 -18.007 < 2e-16 \*\*\*  
## MAKESUZUKI -1.585e+00 9.790e-02 -16.194 < 2e-16 \*\*\*  
## MAKETATA -6.409e-01 8.946e-02 -7.164 7.86e-13 \*\*\*  
## MAKETOYOTA -1.085e-01 4.275e-02 -2.538 0.011151 \*   
## MAKEVOLVO -1.102e+00 1.305e-01 -8.444 < 2e-16 \*\*\*  
## SEATS\_NUM 1.857e-03 8.962e-04 2.072 0.038284 \*   
## USAGECar Hires -6.593e-01 1.722e-01 -3.828 0.000129 \*\*\*  
## USAGEFare Paying Passengers -1.559e-01 8.229e-02 -1.895 0.058123 .   
## USAGEGeneral Cartage 1.067e-01 8.115e-02 1.315 0.188468   
## USAGEOwn Goods -1.453e-01 7.561e-02 -1.922 0.054643 .   
## USAGEOwn service -1.239e-01 7.941e-02 -1.560 0.118798   
## USAGEPrivate -1.427e-02 7.555e-02 -0.189 0.850223   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for binomial family taken to be 1)  
##   
## Null deviance: 165038 on 207655 degrees of freedom  
## Residual deviance: 160920 on 207623 degrees of freedom  
## AIC: 160986  
##   
## Number of Fisher Scoring iterations: 5

drop1(fit.binom2, test= "LRT")

## Single term deletions  
##   
## Model:  
## CLAIM\_PAID ~ SEX + INSURED\_VALUE + PREMIUM + AGE\_VEHICLE + MAKE +   
## SEATS\_NUM + USAGE  
## Df Deviance AIC LRT Pr(>Chi)   
## <none> 160920 160986   
## SEX 2 160964 161026 44.16 2.572e-10 \*\*\*  
## INSURED\_VALUE 1 161163 161227 242.76 < 2.2e-16 \*\*\*  
## PREMIUM 1 162076 162140 1155.55 < 2.2e-16 \*\*\*  
## AGE\_VEHICLE 1 161492 161556 572.39 < 2.2e-16 \*\*\*  
## MAKE 20 162330 162356 1410.25 < 2.2e-16 \*\*\*  
## SEATS\_NUM 1 160924 160988 4.24 0.03957 \*   
## USAGE 6 161064 161118 144.13 < 2.2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

# Wahrscheinlichkeiten vorhersagen  
clean\_dat\_motor$predicted\_probabilities <- predict(fit.binom2, type = "response")  
  
# Klassifikation vorhersagen (0 oder 1)  
clean\_dat\_motor$predicted\_class <- ifelse(clean\_dat\_motor$predicted\_probabilities > 0.5, 1, 0)  
  
# Modellgüte prüfen  
table(clean\_dat\_motor$CLAIM\_PAID, clean\_dat\_motor$predicted\_class)

##   
## 0 1  
## NO 179416 27  
## YES 28204 9

# Optional: Genauigkeit berechnen  
accuracy <- mean(clean\_dat\_motor$CLAIM\_PAID == clean\_dat\_motor$predicted\_class)  
print(paste("Genauigkeit des Modells:", accuracy))

## [1] "Genauigkeit des Modells: 0"

Schlechte Modellgüte

### Massnahme 1): Down-Sampling (Balancing) und weniger Variablen

# Anzahl der YES-Klasse ermitteln  
yes\_class <- clean\_dat\_motor[clean\_dat\_motor$CLAIM\_PAID == "YES", ]  
no\_class <- clean\_dat\_motor[clean\_dat\_motor$CLAIM\_PAID == "NO", ]  
  
# Anzahl der YES-Beobachtungen  
n\_yes <- nrow(yes\_class)  
  
# Zufällige Auswahl aus der NO-Klasse, sodass sie die gleiche Größe wie die YES-Klasse hat  
set.seed(42) # Für Reproduzierbarkeit  
no\_class\_undersampled <- no\_class[sample(1:nrow(no\_class), n\_yes), ]  
  
# Erstellen eines neuen Datensatzes, der YES und die undersampelte NO-Klasse kombiniert  
clean\_dat\_motor\_undersampled <- rbind(yes\_class, no\_class\_undersampled)  
  
# Überprüfen der neuen Klassenverteilung  
table(clean\_dat\_motor\_undersampled$CLAIM\_PAID)

##   
## NO YES   
## 28213 28213

# Logistisches Regressionsmodell mit undersampelten Daten  
fit.binom\_undersampled <- glm(CLAIM\_PAID ~ SEX + INSURED\_VALUE + PREMIUM + AGE\_VEHICLE +   
 MAKE + SEATS\_NUM + USAGE,   
 data = clean\_dat\_motor\_undersampled,   
 family = binomial)  
  
# Zusammenfassung des neuen Modells  
summary(fit.binom\_undersampled)

##   
## Call:  
## glm(formula = CLAIM\_PAID ~ SEX + INSURED\_VALUE + PREMIUM + AGE\_VEHICLE +   
## MAKE + SEATS\_NUM + USAGE, family = binomial, data = clean\_dat\_motor\_undersampled)  
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) 5.918e-01 1.268e-01 4.667 3.05e-06 \*\*\*  
## SEXMale -1.194e-01 2.835e-02 -4.211 2.54e-05 \*\*\*  
## SEXFemale -1.707e-01 4.042e-02 -4.222 2.42e-05 \*\*\*  
## INSURED\_VALUE -2.355e-07 1.896e-08 -12.418 < 2e-16 \*\*\*  
## PREMIUM 2.864e-05 1.166e-06 24.575 < 2e-16 \*\*\*  
## AGE\_VEHICLE -3.139e-02 1.696e-03 -18.514 < 2e-16 \*\*\*  
## MAKEDAEWOO -4.856e-01 1.032e-01 -4.706 2.53e-06 \*\*\*  
## MAKEFIAT -6.536e-01 1.149e-01 -5.689 1.28e-08 \*\*\*  
## MAKEFORD -1.760e-01 9.084e-02 -1.937 0.052729 .   
## MAKEGEELY 3.337e-02 1.163e-01 0.287 0.774225   
## MAKEGENLYON -4.060e-01 1.515e-01 -2.680 0.007357 \*\*   
## MAKEHYUNDAI -3.840e-01 9.139e-02 -4.201 2.65e-05 \*\*\*  
## MAKEISUZU -6.502e-01 6.698e-02 -9.709 < 2e-16 \*\*\*  
## MAKEIVECO -5.788e-01 7.571e-02 -7.645 2.09e-14 \*\*\*  
## MAKELIFAN -1.561e-01 9.006e-02 -1.734 0.082994 .   
## MAKEMAZDA 2.109e-01 1.026e-01 2.056 0.039815 \*   
## MAKEMERCEDES -8.188e-01 1.003e-01 -8.162 3.30e-16 \*\*\*  
## MAKEMITSUBISHI -2.310e-01 7.229e-02 -3.196 0.001395 \*\*   
## MAKENISSAN -3.036e-01 6.676e-02 -4.548 5.41e-06 \*\*\*  
## MAKERENAULT -1.326e+00 1.307e-01 -10.141 < 2e-16 \*\*\*  
## MAKESINO -8.593e-01 1.274e-01 -6.746 1.52e-11 \*\*\*  
## MAKESINO HOWO -1.023e+00 7.585e-02 -13.484 < 2e-16 \*\*\*  
## MAKESUZUKI -1.649e+00 1.140e-01 -14.463 < 2e-16 \*\*\*  
## MAKETATA -7.076e-01 1.180e-01 -5.997 2.00e-09 \*\*\*  
## MAKETOYOTA -1.434e-01 6.032e-02 -2.377 0.017463 \*   
## MAKEVOLVO -1.133e+00 1.541e-01 -7.351 1.97e-13 \*\*\*  
## SEATS\_NUM 9.086e-04 1.229e-03 0.739 0.459688   
## USAGECar Hires -8.124e-01 2.153e-01 -3.773 0.000161 \*\*\*  
## USAGEFare Paying Passengers -2.384e-01 1.180e-01 -2.021 0.043304 \*   
## USAGEGeneral Cartage 3.787e-02 1.163e-01 0.326 0.744662   
## USAGEOwn Goods -1.829e-01 1.092e-01 -1.674 0.094081 .   
## USAGEOwn service -1.806e-01 1.142e-01 -1.582 0.113594   
## USAGEPrivate -3.381e-02 1.094e-01 -0.309 0.757250   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for binomial family taken to be 1)  
##   
## Null deviance: 78223 on 56425 degrees of freedom  
## Residual deviance: 75778 on 56393 degrees of freedom  
## AIC: 75844  
##   
## Number of Fisher Scoring iterations: 4

drop1(fit.binom\_undersampled, test= "LRT") #SEATS\_NUM not sign

## Single term deletions  
##   
## Model:  
## CLAIM\_PAID ~ SEX + INSURED\_VALUE + PREMIUM + AGE\_VEHICLE + MAKE +   
## SEATS\_NUM + USAGE  
## Df Deviance AIC LRT Pr(>Chi)   
## <none> 75778 75844   
## SEX 2 75803 75865 25.34 3.143e-06 \*\*\*  
## INSURED\_VALUE 1 75935 75999 157.64 < 2.2e-16 \*\*\*  
## PREMIUM 1 76420 76484 642.19 < 2.2e-16 \*\*\*  
## AGE\_VEHICLE 1 76127 76191 349.30 < 2.2e-16 \*\*\*  
## MAKE 20 76657 76683 879.56 < 2.2e-16 \*\*\*  
## SEATS\_NUM 1 75778 75842 0.55 0.4595   
## USAGE 6 75877 75931 99.64 < 2.2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

# Logistisches Regressionsmodell mit undersampelten Daten  
fit.binom\_undersampled2 <- glm(CLAIM\_PAID ~ SEX + INSURED\_VALUE + PREMIUM + AGE\_VEHICLE +   
 MAKE + USAGE, data = clean\_dat\_motor\_undersampled,   
 family = binomial)  
  
# Zusammenfassung des neuen Modells  
summary(fit.binom\_undersampled2)

##   
## Call:  
## glm(formula = CLAIM\_PAID ~ SEX + INSURED\_VALUE + PREMIUM + AGE\_VEHICLE +   
## MAKE + USAGE, family = binomial, data = clean\_dat\_motor\_undersampled)  
##   
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)   
## (Intercept) 6.096e-01 1.245e-01 4.896 9.80e-07 \*\*\*  
## SEXMale -1.214e-01 2.822e-02 -4.302 1.69e-05 \*\*\*  
## SEXFemale -1.725e-01 4.035e-02 -4.276 1.90e-05 \*\*\*  
## INSURED\_VALUE -2.316e-07 1.823e-08 -12.706 < 2e-16 \*\*\*  
## PREMIUM 2.858e-05 1.162e-06 24.587 < 2e-16 \*\*\*  
## AGE\_VEHICLE -3.123e-02 1.680e-03 -18.583 < 2e-16 \*\*\*  
## MAKEDAEWOO -4.915e-01 1.028e-01 -4.779 1.76e-06 \*\*\*  
## MAKEFIAT -6.676e-01 1.133e-01 -5.892 3.81e-09 \*\*\*  
## MAKEFORD -1.921e-01 8.820e-02 -2.179 0.029368 \*   
## MAKEGEELY 1.671e-02 1.141e-01 0.146 0.883591   
## MAKEGENLYON -4.227e-01 1.498e-01 -2.821 0.004787 \*\*   
## MAKEHYUNDAI -4.021e-01 8.806e-02 -4.566 4.97e-06 \*\*\*  
## MAKEISUZU -6.638e-01 6.444e-02 -10.300 < 2e-16 \*\*\*  
## MAKEIVECO -5.985e-01 7.093e-02 -8.437 < 2e-16 \*\*\*  
## MAKELIFAN -1.729e-01 8.720e-02 -1.983 0.047394 \*   
## MAKEMAZDA 1.942e-01 1.001e-01 1.940 0.052366 .   
## MAKEMERCEDES -8.294e-01 9.927e-02 -8.355 < 2e-16 \*\*\*  
## MAKEMITSUBISHI -2.468e-01 6.911e-02 -3.571 0.000356 \*\*\*  
## MAKENISSAN -3.214e-01 6.233e-02 -5.156 2.52e-07 \*\*\*  
## MAKERENAULT -1.346e+00 1.278e-01 -10.528 < 2e-16 \*\*\*  
## MAKESINO -8.755e-01 1.255e-01 -6.975 3.06e-12 \*\*\*  
## MAKESINO HOWO -1.039e+00 7.254e-02 -14.325 < 2e-16 \*\*\*  
## MAKESUZUKI -1.669e+00 1.107e-01 -15.078 < 2e-16 \*\*\*  
## MAKETATA -7.051e-01 1.178e-01 -5.983 2.19e-09 \*\*\*  
## MAKETOYOTA -1.621e-01 5.481e-02 -2.957 0.003108 \*\*   
## MAKEVOLVO -1.151e+00 1.521e-01 -7.566 3.84e-14 \*\*\*  
## USAGECar Hires -8.090e-01 2.153e-01 -3.758 0.000171 \*\*\*  
## USAGEFare Paying Passengers -2.201e-01 1.154e-01 -1.908 0.056358 .   
## USAGEGeneral Cartage 3.362e-02 1.161e-01 0.289 0.772232   
## USAGEOwn Goods -1.858e-01 1.092e-01 -1.702 0.088700 .   
## USAGEOwn service -1.673e-01 1.127e-01 -1.484 0.137798   
## USAGEPrivate -3.281e-02 1.094e-01 -0.300 0.764183   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## (Dispersion parameter for binomial family taken to be 1)  
##   
## Null deviance: 78223 on 56425 degrees of freedom  
## Residual deviance: 75778 on 56394 degrees of freedom  
## AIC: 75842  
##   
## Number of Fisher Scoring iterations: 4

drop1(fit.binom\_undersampled2, test= "LRT")

## Single term deletions  
##   
## Model:  
## CLAIM\_PAID ~ SEX + INSURED\_VALUE + PREMIUM + AGE\_VEHICLE + MAKE +   
## USAGE  
## Df Deviance AIC LRT Pr(>Chi)   
## <none> 75778 75842   
## SEX 2 75804 75864 26.28 1.966e-06 \*\*\*  
## INSURED\_VALUE 1 75943 76005 164.99 < 2.2e-16 \*\*\*  
## PREMIUM 1 76421 76483 642.39 < 2.2e-16 \*\*\*  
## AGE\_VEHICLE 1 76130 76192 351.79 < 2.2e-16 \*\*\*  
## MAKE 20 76657 76681 879.07 < 2.2e-16 \*\*\*  
## USAGE 6 75881 75933 102.80 < 2.2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

# Wahrscheinlichkeiten vorhersagen  
clean\_dat\_motor\_undersampled$predicted\_probabilities <- predict(fit.binom\_undersampled2, type = "response")  
  
# Klassifikation vorhersagen (0 oder 1)  
clean\_dat\_motor\_undersampled$predicted\_class <- ifelse(clean\_dat\_motor\_undersampled$predicted\_probabilities > 0.5, 1, 0)  
  
# Modellgüte prüfen  
table(clean\_dat\_motor\_undersampled$CLAIM\_PAID, clean\_dat\_motor\_undersampled$predicted\_class)

##   
## 0 1  
## NO 16260 11953  
## YES 11370 16843

# Genauigkeit berechnen  
accuracy\_undersampled <- mean(clean\_dat\_motor\_undersampled$CLAIM\_PAID == clean\_dat\_motor\_undersampled$predicted\_class)  
print(paste("Genauigkeit des undersampelten Modells:", accuracy\_undersampled))

## [1] "Genauigkeit des undersampelten Modells: 0"

# Umwandlung in Faktoren für die Auswertung  
predicted\_class <- as.factor(clean\_dat\_motor\_undersampled$predicted\_class)  
actual\_class <- as.factor(clean\_dat\_motor\_undersampled$CLAIM\_PAID)  
  
  
# Sicherstellen, dass predicted\_class und actual\_class die gleichen Levels haben  
levels(predicted\_class) <- levels(actual\_class)  
  
# Levels überprüfen  
levels(predicted\_class)

## [1] "NO" "YES"

levels(actual\_class)

## [1] "NO" "YES"

# Berechnung der Metriken  
confusionMatrix(predicted\_class, actual\_class)

## Confusion Matrix and Statistics  
##   
## Reference  
## Prediction NO YES  
## NO 16260 11370  
## YES 11953 16843  
##   
## Accuracy : 0.5867   
## 95% CI : (0.5826, 0.5907)  
## No Information Rate : 0.5   
## P-Value [Acc > NIR] : < 2.2e-16   
##   
## Kappa : 0.1733   
##   
## Mcnemar's Test P-Value : 0.0001384   
##   
## Sensitivity : 0.5763   
## Specificity : 0.5970   
## Pos Pred Value : 0.5885   
## Neg Pred Value : 0.5849   
## Prevalence : 0.5000   
## Detection Rate : 0.2882   
## Detection Prevalence : 0.4897   
## Balanced Accuracy : 0.5867   
##   
## 'Positive' Class : NO   
##

weitere Massnahmen: Schwellenwertanpassubg (hat wenig gebracht) Interaktionen ## Generalised Additive Model (GAM)

# TODO

## Neural Network

# TODO

## Support Vector Machine (SVM)

# TODO

# Conclusion