KU LEUVEN kulak

Informatica
Statistische modellen en data-analyse
Prof. Van Aelst, Stijn Rebry

Opdracht 2

Academiejaar 2018 – 2019

Thomas Bamelis, Michiel Jonckheere

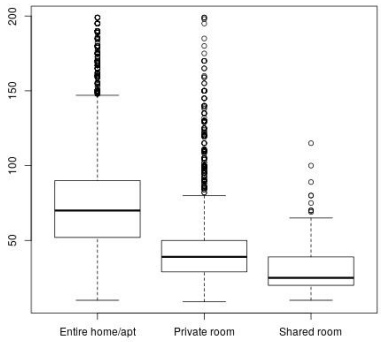
Inleiding

In dit verslag analyseren we de gegevens van airbnb gemeten in 2019 van de steden Antwerpen, Brussel en Gent van 11.000 verblijven. We analyseren de prijs in functie van het type, de stad en de buurt van het bedrijf. Daarna proberen we een regressiemodel op te stellen voor de huurprijs en dit model te evalueren en te interpreteren. Als laatste passen we logistische regressie toe om te voorspellen of een verblijf al dan niet wordt vast verhuurd in plaats van enkel voor korte periodes.
Opmerking: als in dit verslag een p-waarde als nagenoeg 0 wordt beschreven, betekent dit dat de p-waarde voor underflow zorgde in het computersysteem ($<$ 2.2e-16). Indien het significantieniveau niet vermeld staat wordt 0.05 gehanteerd.
De volgende gegevens zijn beschikbaar: id, naam, uitbater id, naam uitbater, buurt, latitude, longitude, type verblijf, huurprijs, minimum aantal nachten, aantal reviews, aantal dagen sinds de laatste review, aantal reviews per maand, aantal verblijven/zoekertjes van de uitbater, hoeveel dagen het verblijf beschikbaar was afgelopen jaar, stad en vol (variabele die aangeeft of een verblijf al dan niet vast verhuurd wordt in plaats van enkel voor korte periodes). Deze laatste werd zelf afgeleid uit het aantal beschikbare dagen.

1 Ligging en type verblijf

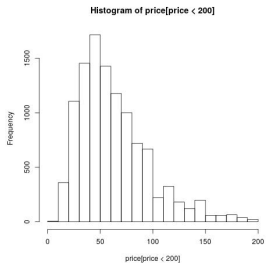
1.1 type verblijf vs huurprijs

Een boxplot (1) zonder de extreme outliers lijkt te suggereren dat er een verschil blijkt te zijn tussen de types, vooral tussen een volledig huis/appartement en de rest. Om overzicht te behouden hebben we de bovenste outliers uit de figuur gelaten, waarvan er nog 705 tussen 150 en 1500 lagen en nog 8 hoger dan 1500 met een maximum van 8944. De prijs lijkt aan de hand van figuur 2a (zonder outliers) een redelijke klokcurve voor te stellen met een zeer lichte linkerstaart en een erg zware rechterstaart. De prijs lijkt niet normaal verdeeld door outliers en de knik in de curve (qqplot 2b). Een schatting van de lambda voor een Box-Cox transformatie raad -0.2549 (afgerond $-\frac{1}{4}$) aan als lambda. Figuur 2d toont een verbetering qua normaliteit. Deze transformatie gaf de beste verbetering van alle geteste transformaties (log10, sqrt, log10(log10)).

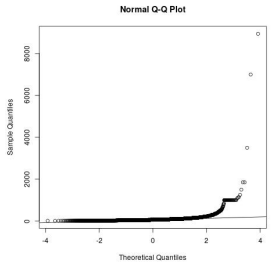


Figuur 1: Boxplot type vs prijs.

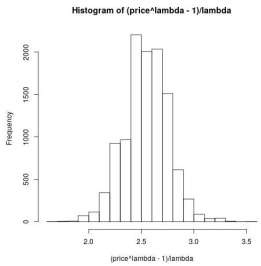
De Box-Cox transformatie komt er als beste uit als we residuals analyseren van een model met de verschillende transformaties (figuur 3a). We nemen voor de rest van deze sectie die transformatie voor de prijs. De Levene test toont aan dat er heteroscedasticiteit is, wat aannemelijk lijkt gegeven figuur 1. Kwantielplot 3a toont dat de residuals niet normaal verdeeld zijn. Het aantal samples is te groot om een Shapiro test op uit te voeren. Een plot tegenover de index 3b toont geen afhankelijkheid “in de tijd”. Niet alle modelveronderstellingen voor de relevante testen zijn voldaan, onze bevindingen moeten daarom met een korrel zout genomen worden.



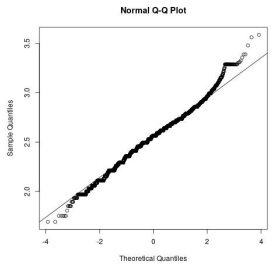
(a) Visualisatie van de prijs.



(b) Kwantielplot prijs



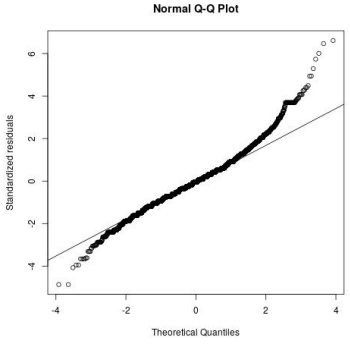
(c) Visualisatie Box-Cox $-\frac{1}{4}$ prijs



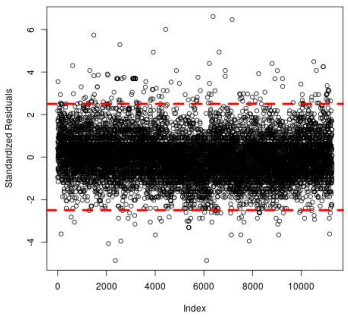
(d) Kwantielplot Box-Cox $-\frac{1}{4}$ prijs

Het model verwierp de f-test en t-test voor alle variabelen. De R-squared en adjusted R-squared waren echter maar 0.268. De weighted least square methode toonde geen verbeteringen op het model na 1 iteratie. De aov methode (partiële F-test met Bonferroni correctie) toont alsook aan dat het type verblijf effectief significant is voor de prijs. De tukey-test toont aan dat de verschillende soorten kamers onderling ook genoeg verschillen van elkaar.

Conclusie: het type verblijf is significant voor de prijs en de types verschillen onderling allemaal in prijs. Volgens de coëfficiënten van het model is de prijs als volgt gerangschikt: volledig huis app. > privé kamer > gedeelde kamer. Deze bevindingen stroken met de vermoedens van de realiteit.



(a) Kwantielplot gestandardiseerde residuals



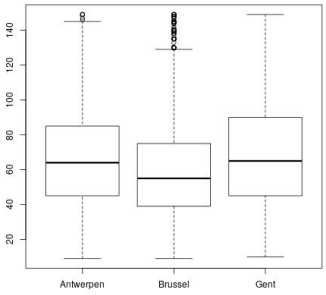
(b) Index vs gest. residuals

1.2 stad vs huurprijs

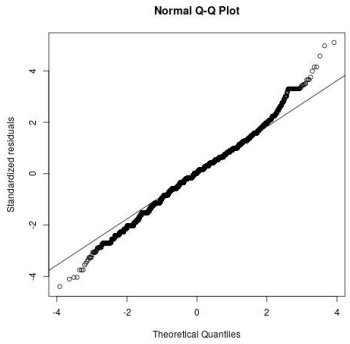
De boxplot zonder outliers (figuur 4a) suggereert dat er niet veel verschil is in de gemiddelde prijs per stad. Antwerpen en Gent zijn nagenoeg het zelfde, enkel de prijzen in Brussel liggen wat lager.

We zien dat de residuals niet normaal verdeeld zijn (figuur 4b), waardoor onze bevindingen opnieuw niet steenhard mogen genomen worden. De Levene test kan echter niet verwerpen dat er homoscedasticiteit.
Het model verwerpt de f-test en t-test voor de variabelen van Antwerpen (= het intercept) en Brussel, maar Gent is niet verworpen met een p-waarde van 0.068. Antwerpen krijgt "voorrang" aangezien het het intercept is, en de gelijkenissen tussen Antwerpen en Gent uit figuur 4a dit kunnen verklaren. De reden dat Gent dus niet verworpen is, is omdat hij heel sterk op Antwerpen lijkt, wat dus betekent dat beiden niet significant verschillen qua prijs. Het kleine verschil in de coëfficiënten bevestigd dat. Aov zegt dat de stad significant is voor de prijs. De residuals zijn niet afhankelijk van de tijd/index (figuur 4c). De R-squared van dit model is 0.0226 en de adjusted is 0.0224, wat dus opnieuw niet goed is. De Tukey-test verwerpt sterk dat er geen verschil zou zijn tussen Brussel en de andere twee steden. Maar Gent en Antwerpen worden niet verworpen met een p-waarde van 0.162, wat opnieuw onze vermoedens

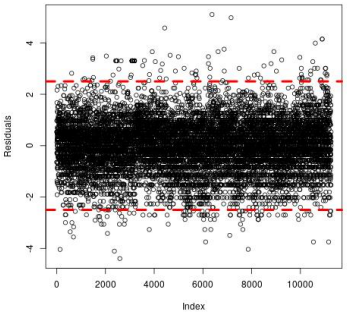
bevestigd. Het effect van Brussel is ook veel sterker dan die van Antwerpen en Gent. Conclusie: de stad is significant voor de prijs. Er is een prijsverschil tussen Brussel en de andere 2 steden, maar er is geen significant prijsverschil tussen Antwerpen en Gent. De coëfficiënten van het model vertellen dat Brussel goedkoper is dan de andere 2 steden, wat strookt met de boxplot op figuur 4a.



(a) Boxplot stad vs prijs.



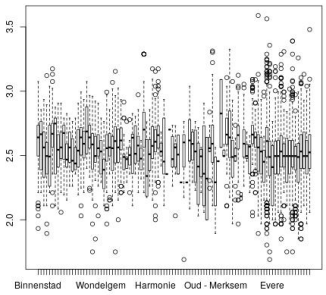
(b) QQplot gestandaardiseerde residuals model



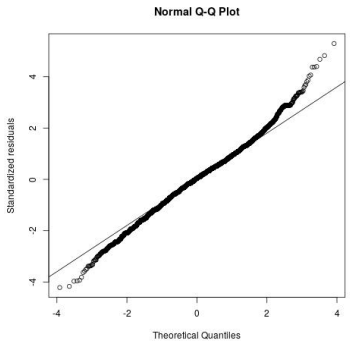
(c) Index vs residuals

1.3 Buurten

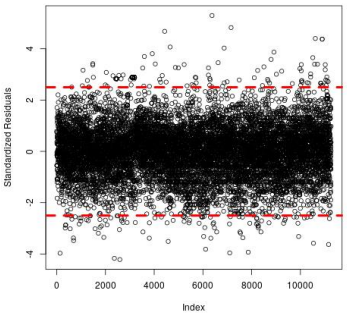
Op het eerste zicht (figuur 5a) lijken de buurten niet zo veel te verschillen. Het model waarbij we enkel neighbourhood meenemen, verwerpt de f-test, maar een groot deel van de neighbourhoods worden niet verworpen door de t-test (43/96). De Levene test en een normale qqplot tonen dat er heteroscedasticiteit is en de residuals niet normaal zijn,dus onze bevindingen moeten daarom met een korrel zout genomen worden. Residual zijn niet afhankelijk van de tijd/index (figuur 5c) Aov toont aan dat de buurt significant. De TukeyTest verteld dat 47,6% van de combinaties van buurten niet significant verschillen. De R-squared is 0.099 en adjusted 0.091, wat opnieuw aan de lage kant is. Alleen lijkt dus dat de buurt significant is in het algemeen, maar zeer veel buurten onderling te weinig verschillen. Dit doet suggereren dat de opdeling per buurt te verfijnt is.



(a) Boxplot buurt vs prijs.



(b) QQplot gestandaardiseerde residuals model

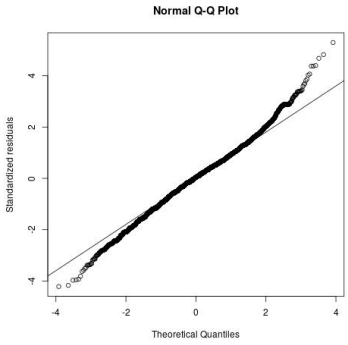


(c) Index vs residuals

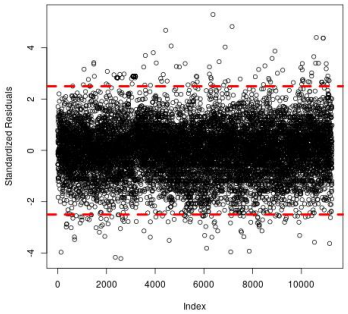
1.4 Buurten en stad

Het model met de buurt en de stad verwerpt de f-test ook met, maar een nog groter deel van wordt niet verworpen door de t-test (54). Meer bepaald is de pwaarde van Gent zeer hoog (0.9) terwijl Antwerpen en Brussel nog steeds verworpen worden. De R-squared is 0.099 en adjusted 0.091 (onveranderd). Dit doet vermoeden dat beide dus sterk verbonden zijn aangezien hun gecombineerd model even slecht is als hun afzonderlijke. Aov verwerpt echter voor beide variabelen. De Tukey test kan nu nog 44,6% van de combinaties verwerpen. Deze bevindingen suggereren dat door de buurt en de stad beiden op te nemen, de significantie van Gent vs. Antwerpen miniem wordt. De Aov verwerpt niet dat een van de

twee overbodig wordt, maar aangezien zeer veel variabelen de t-test falen, er geen verbetering is in de R-squared en het nog steeds hoge percentage van het niet verwerpen van de Tukey test tonen dat het origineel probleem van het te verfijnd zijn van de buurtenverdeling enkel nog maar versterkt wordt door de stad erbij te nemen.

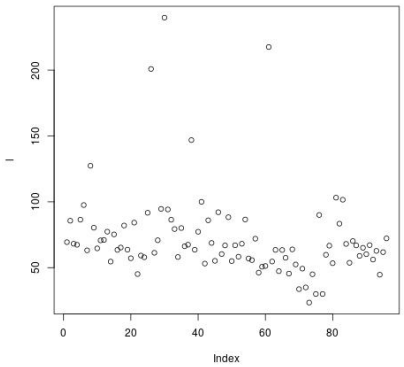


(a) QQplot gestandaardiseerde residuals model



(b) Index vs residuals

Figuur 7 toont dat er een ongeveer 5 buurten zijn die een opvallend hogere prijs hebben dan de rest. De gemiddelde huurprijs is 72,2 euro. Woluwe-Saint-Pierre steekt hier sterk boven met een gemiddelde huurprijs van 240 euro. Daarna volgt Sint Denijs Westrem met 217,5 en Oud-Berchem met 201. Deze drie zijn significant hoger dan de rest in liggen in Brussel, Gent en Antwerpen. Niet in een bepaalde stad dus. Dan zijn er nog 2 minder extreem prijzige gemeenten, namelijk Eilandje met 147 en Polder met 127 (beiden liggen in Antwerpen). De eerste opeenvolgende is nog maar 103. Daarnaast zijn er nog 5 Antwerpse buurten waar de prijs ietwat lager ligt dan de rest met prijzen tussen 23,5 en 35 euro. Deze zijn echter minder een stuk minder ver verwijdert van het gemiddelde in vergelijking met de extreem prijzige outliers.



Figuur 7: Gemiddelde prijs per buurt.

1.5 Buurten en type verblijf

Opnieuw werd de f-test verworpen en 42 t-testen faalden. Aov verwerpt echter voor beide variabelen. De R-squared is sterk gestegen van 0.099 naar 0.3281 en de adjusted van 0.091 naar 0.3222. De R-squared van room-type alleen was ook 0.2678 en de adjusted 0.2677. De combinatie van de 2 variabelen versterken elkaar dus. De buurt en het type verblijf maken elkaar dus niet overbodig.

2 Model voor de huurprijs

Op basis van relevante beschikbare gegevens hebben we geprobeerd een model op te stellen om de prijs van een Airbnb-verblijf zo goed mogelijk te voorspellen. Eerst moesten we daarvoor kijken welke gegevens er relevant zouden zijn voor ons model. Daarna controleerden we of het nodig was om bij bepaalde variabelen een transformatie toe te passen. We bekeken enkele modellen met en zonder transformaties en tot slot bepaalden we het beste model om de prijs te voorspellen.

2.1 Selectie relevante gegevens

Voor we beginnen met een model te zoeken verwijderen we alle rijen die een kolom missen. Het aantal samples gaat dan van 11262 naar 9476.

Om variabelen te selecteren als een relevante variabele voor ons model, keken we naar hoe de prijs zich verhoudt ten opzichte van de variabele. We laten de naam en id van het verblijf en van de host achterwege omdat id willekeurig zijn voor de andere attributen van een zoekertje. De naam (of id) van het verblijf is zinloos omdat dit een 1 op 1 relatie is per zoekertje en dit dus gigantische overfitting zou zijn. De naam van de host (of hun categorische id) zou in principe kunnen meegenomen worden, maar dit doen we niet om 3 redenen:

- kan gezien worden als sterke overfitting (zie puntje 3)
- het model wordt op die manier onoverzichtelijk
- het model heeft op die manier geen enkele betekenis meer als een nieuw verhuurder op de site komt en heeft dus geen voorspellende kracht om te voorspellen.

De variabele "latitude" blijkt wel een in vloed te hebben op de prijs. Als we kijken naar de waarden van de breedtegraden dan zien we dat we deze kunnen indelen in de drie steden. We gaan in plaats van de breedtegraad in ons model te gebruiken, gebruik maken van de variabele "city". "Longitude" wordt om dezelfde reden niet gebruikt in het model, deze variabele deelt de data op in twee delen namelijk als eerste deel Gent en als tweede deel Brussel en Antwerpen. Vervolgens zien we ook dat variabelen "room_type", "minimum_nights", "number_of_reviews", "last_review", "reviews_per_month", "availability_365" en "calculated_host_listings_count" relevant kunnen zijn voor ons model.

Voor de variabelen die iets zeggen over de reviews hebben ook eens gekeken naar de correlatie tussen deze variabelen. We vonden dat er een sterke correlatie bestaat tussen "reviews_per_month" en "number_of_reviews", ook is er een correlatie tussen "reviews_per_month" en "last_review". Om deze reden gaan we zeker al twee verschillende modellen opstellen waarbij enerzijds gewerkt wordt met alle drie de variabelen en anderzijds enkel met de variabele "reviews_per_month".

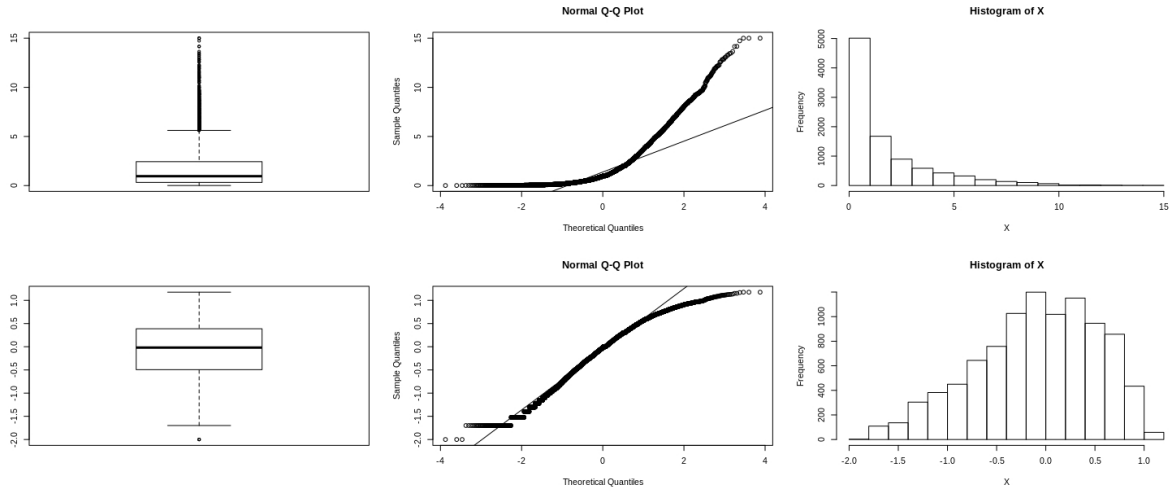
In tabel 1 is een samenvatting te zien van de relevante gegevens die we zullen gebruiken voor ons model op te stellen.

Variabele	**Afkorting**	**Transformatie**
room_type	$R_{private}$, R_{shared},R_{entire}	/
city	C_{Gent}, C_{Antw},$C_{Brussel}$	/
price	P	$\frac{(P^{(-0.25)})-1}{-0.25}$
minimum_nights	MN	$\frac{(MN^{(-0.67)})-1}{-0.67}$
number_of_reviews	$NbRev$	$log_{10}(NbRev)$
last_review	$LRev$	$log_{10}(LRev+1)$
reviews_per_month	$RevMonth$	$log_{10}(RevMonth)$
calculated_host_listings_count	$CHLC$	$\frac{(CHLC^{(-1)})-1}{-1}$
availability_365	Av	/

Tabel 1: Samenvatting van de relevante variabelen en hun transformaties.

2.2 Transformaties

In deze sectie bekijken we of er transformaties zijn die ervoor kunnen zorgen dat we een beter model verkrijgen om de prijs te voorspellen. Het spreekt voor zich dat we hier enkel naar de numerieke variabelen kijken. In tabel 1 is een samenvatting te zien van de gevonden transformaties. Voor elke variabele vergeleken we de identieke met een log10 en Box-Cox transformatie. Hiervoor werd gekeken naar de boxplot, normale kwantielplot en het histogram. In figuur 8 is een voorbeeld te zien van hoe we te werk gingen, de variabele hier is het aantal reviews per maand. De bovenste drie figuren zijn van de identieke, voor de onderste drie is de log10 transformatie toegepast. Het is duidelijk dat de log10 transformatie voor het aantal reviews per maand hier een verbetering geeft.



Figuur 8: Voorbeeld van een verbetering door log10 transformatie op het aantal reviews per maand. Eerste drie grafieken zijn de boxplot, normale kwantielplot en het histogram zonder transformatie, de andere drie zijn de grafieken van de log10 transformatie.

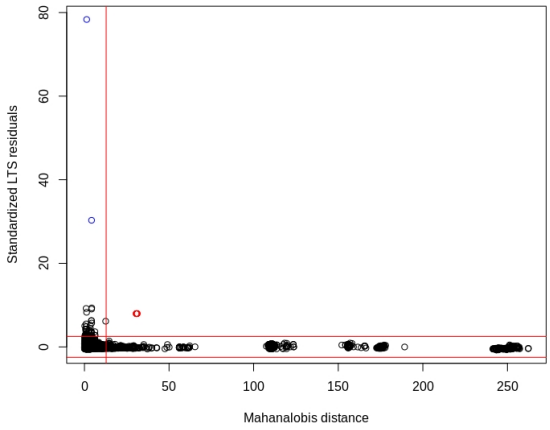
2.3 Modellen

Om een goed model te vinden zijn we als volgt te werk gegaan: ons eerste model ontstond van uit alle relevante niet-getransformeerde variabelen het beste model te nemen. Met een R-squared van 0.032 vonden we dit model zeer slecht en dus gingen we verder naar het model met de getransformeerde variabelen. Hierbij merkten we direct een grote verbetering in de R-squared (= 0.37). In het derde en vierde model dat we bekeken, lieten we de variabelen laatste review en aantal reviews weg. Hier keken we weer naar de modellen met niet-getransformeerde en getransformeerde variabelen. Het derde model bleek gelijk te zijn aan het eerste model, dit is te verklaren omdat bij het opstellen van het eerste model de twee review variabelen uiteindelijk ook weggelaten werden. Tot slot hebben we nog het vierde model, dit model is zo goed als gelijk aan het tweede model met het enige verschil dat de getransformeerde variabele laatste review hier niet in meegenomen is.

Voor de twee beste modellen van hierboven hebben we ook gekeken of er een verbetering was als we met interactietermen werkten. Hierbij zagen we dat die modellen niet beter waren dan de modellen die we al hadden. Bij het kijken naar de multicollineariteit van de modellen zonder en met interactietermen, zagen we dat die zonder interactietermen veel minder multicollineariteit hadden dan die met interactietermen. Hieruit concluderen we dus dat het tweede en vierde model van hierboven beter zijn. Beide modellen hebben geen last van multicollineariteit.

Aangezien het tweede en vierde model enkel verschilde in het al dan niet opnemen van de getransformeerde variabele laatste review, keken we ook eens naar hoe significant de bijdrage was van deze variabele. Hieruit bleek dat deze bijdrage niet significant was en besluiten we dat het vierde model het beste was. De coëfficiënten van beide modellen zijn ook zo goed als gelijk.

We weten dat de prijsvariabele extreme outliers heeft, dit is dan ook de verklaring waarom de twee modellen zonder getransformeerde variabelen zeer slecht waren. Als we keken naar het eerste en vierde model die niet beïnvloed werden door de outliers (met behulp van de LTS regressie), dan bleek dat het eerste model ongeveer een even goede voorspelling geeft voor de prijs als het vierde model (respectievelijk is de R-squared 0.45 en 0.43). Voor het vierde model zonder outliers was er niet zo'n groot verschil te merken in de coëfficiënten met het initiële vierde model, dit komt omdat door de transformatie er al veel outliers onderdrukt zijn.



Figuur 9: Diagnostische plot van het eerste model, geconstrueerd met LTS. de blauwe punten zijn extreme verticale outliers, de rode punten zijn de bad leverage points.

In figuur 9 is de diagnostische plot van het eerste model zonder invloed van outliers te zien. Hierbij merken we eerst op dat er twee extreme verticale outliers zijn. Dit zijn de twee duurste verblijven uit onze dataset. Ook zien we dat er één bad leverage punt is. Als we dit punt onderzochten, vonden we dat er eigenlijk acht bad leverage punten zijn die allemaal samenvallen. Deze acht zijn van dezelfde uitbater, uit dezelfde stad en de prijs van deze verblijven is allemaal €999. Tot slot is het ook opmerkelijk dat er zeer veel good leverage points zijn. De punten met een mahanalobis afstand groter dan 100 zijn allemaal verblijven uit Brussel en dit zijn 361 verblijven. Deze 361 verblijven zijn verdeeld over zes uitbaters. We kunnen besluiten uit deze diagnostische plot dat de meeste reguliere observaties degene zijn van uitbaters die een klein aantal verblijven hebben en een normale prijs vragen.

2.4 Conclusie

We concluderen dat het eerste model het beste model is als de outliers onderdrukt worden. Dit is het gevonden model:

$$price = 62.18 - 26.05 \cdot R_{private} - 44.06 \cdot R_{shared} - 3.04 \cdot C_{Brussel} + 6.04 \cdot C_{Gent} - 0.30 \cdot RevMonth + 1.09 \cdot CHLC + 0.025 \cdot Av \quad (1)$$

Eerst en vooral zien we hier dat de gemiddelde prijs voor het huren van een volledig huis/appartement, in Antwerpen, die altijd verhuurd is, geen reviews per maand heeft en de uitbater maar 1 verblijf heeft, ongeveer €63.27 bedraagd. Daarnaast daalt de voorspelling van de prijs een stuk als er sprake is van een privé of gedeelde kamer. Zoals we in sectie 1.2 al zagen, is hier ook duidelijk dat de prijzen in Gent iets hoger liggen dan in Antwerpen en dat de prijzen in Brussel lager zijn dan in Antwerpen en Gent. De andere drie coëfficiënten zijn moeilijker te interpreteren.

3 Beschikbaarheid van een verblijf

In deze sectie proberen we logistische regressie toe te passen om te voorspellen of een verblijf al dan niet wordt vast verhuurd in plaats van enkel voor korte periodes. We proberen een model te vinden, dit te evalueren en te interpreteren.

3.1 Model selectie

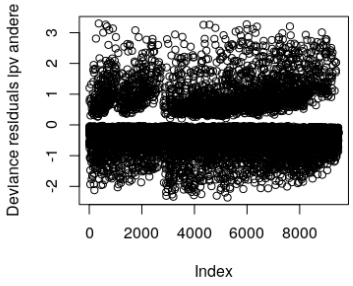
De proporties van categorische veranderlijken worden hierdoor niet significant aangepast. Het aantal beschikbare dagen wordt natuurlijk niet opgenomen als regressor omdat het al dan niet permanent verhuurd zijn direct hieruit is afgeleid. Ook hier laten we de id's en namen achterwege.

We pasten AIC gebaseerde voorwaartse, achterwaartse en stapsgewijze regressie toe. Dit deden we 2 maal, de 2e keer met de variabelen in de omgekeerde volgorde. Hierbij werden de prijs, het type, het minimum aantal nachten, het aantal zoekertjes per host, de datum van de laatste review, het aantal reviews per maand en het aantal reviews van het afgelopen jaar telkens geselecteerd. De buurt werd nooit geselecteerd, de stad 2 keer en de latitude en longitude 4 keer. Stad werd enkel geselecteerd als latitude en longitude niet geselecteerd werden en omgekeerd, wat dus suggereert dat ze dezelfde informatie toevoegen aan het model, wat ook strookt met de realiteit van de soort variabelen (zie sectie 2.1). Stad is de nette classificatie van latitude en longitude, daarom vergelijken we een model met latitude en longitude tegenover een model met de stad in de plaats. De buurt laten we achterwege. De bekomen modellen hebben dezelfde AIC en deviance waarden en resulteren in hetzelfde voor de likelihood ratio test en de goodness of fit test, behalve dat latitude maar zwak verworpen wordt in hun eigen model. Omdat de stad de significantie van latitude en longitude mooi samenvat gaan we verder met de stad en laten we latitude en longitude vallen. In de testen wordt echter niet verworpen dat het type er niet toe doet. Omdat deze echter altijd in de geselecteerde modellen zat en het bijna verworpen was nemen we het toch mee in ons model. Omdat normaliteit van de regressoren de kwaliteit van een model sterk beinvloeden gebruiken we dezelfde transformaties als in sectie 2. Voor de interpretatie hanteren we echter een model zonder transformaties om ons meer zinvol te kunnen uitspreken over het model.

3.2 Model evaluatie

De Wald test verwerpt het niet significant zijn van de enkele variabelen allemaal met voldoende marge. De behaalde deviance is 6305.7 en de AIC 6327.7. Dit is een daling van ongeveer 1000 tegenover het model zonder transformaties.

De likelihood ratio test verwerpt ook voor alle regressoren dat ze niet significant zouden zijn met voldoende marge. De deviance residuals (figuur 10) vertonen niets zorgbarends, behalve dat er een opslitsing te zien is tussen de steden, mede veroorzaakt omdat ze aan elkaar geplakt zijn en dus gegroepeerd in de data. De goodness of fit test verwerpt ook sterk. Qua problemen is er de duidelijke opsplitsing van de deviance residuals, maar de oorzaak is bekend en geeft geen verdere problemen. Doordat de residuals niet normaal verdeeld zijn bij logistische regressie kan er ook niet sterk gecontroleerd worden op problemen met normaliteit of modelveronderstelling om problemen te vinden.



Figuur 10: Deviance residuals.

3.3 Model interpretatie

Het model met dezelfde variabelen maar ongetransformeerd kan als volgt geïnterpreteerd worden:
Als hier gezegd wordt meer kans, wordt veronderstelt dat alle andere variabelen strikt onveranderd blijven en gegeven zijn.
Iemand in Brussel heeft 54 procent meer kans dat het vol zit dan in Antwerpen. Iemand in Gent 73% meer kans dat het vol zit dan in Antwerpen. Iemand die een privé kamer zoekt heeft 22.5 procent meer kans dat het vol zit dan bij een volledig huis of appartement. Iemand die een gedeelde kamer zoekt heeft 16.2 procent minder kans dat het vol zit dan bij een volledig huis of appartement. Je hebt meer kans dat het verblijf permanent verhuurd is als:

- prijs - (1.2)

- laatste review + (0.3)
- aantal reviews - (0.7)
- reviews maand - (22.1)
- minimum nachten - (3)
- zoekertjes uitbater - (4)

waarbij -/+ betekent de kans stijgt als het attribuut daalt/stijgt. Het cijfer tussen de haakjes is het percentage dat de kans stijgt of daalt bij het veranderen van 1 van het attribuut. Als het langer geleden is dat de laatste review werd geschreven heb je een grotere kans dat het verblijf permanent verhuurd is. Als de prijs, het aantal reviews in totaal en per maand, het minimum aantal nachten en het aantal zoekertjes per uitbater lager is, heb je een grotere kans dat het verblijf permanent verhuurd is. Dit strookt met wat je zou denken dat er in de realiteit gebeurt, behalve dan het minimum aantal nachten. Mogelijks is dit omdat je maximum aantal minimum nachten 30 dagen kan zijn en dus niet relevant is voor permanente huurders, waardoor ze het gewoon op 1 laten staan.