Termpaper MSB205

Arbeidskrav (kan finne en annen tittel)

Oppg. 1

Bishop et al. (2020) skriver at den mest direkte illustrasjonen på hvordan private markeder kan avsløre forbrukernes betalingsvillighet er den hedoniske eiendomsverdi modellen. Modellen går ut på at en ser for seg at kjøpere velger eiendommer basert på egenskapene til boligen, som kan for eksempel være antall soverom og bad. Den er også basert på stedsspesifikke fasiliteter for en bolig som for eksempel er luftkvaliteten, nærhet til sentrum/parker og flomrisiko Bishop et al. (2020).

I løpet av de siste årene har forventninger til kvaliteten på dataen og den økonometrisk åpenhet økt. Det har også forbedret forståelse om hvordan en identifisert gjennom kvasi-eksperimentelle forskningsdesign forholder seg til velferdsmål, som vil si mål på betalingsvillighet. Det ble brukt rike data fra boligmarkedet i storbyområdene i en avansert økonomi i de fleste studiene, som etablerte den beste praksisen i den hedoniske modellen. Denne dataen beskriver boligtransaksjoner, egenskaper og fasiliteter som stadig blir mer tilgjengelig rundt om i verden, som gir muligheten for å bruke den hedoniske modellen Bishop et al. (2020).

Den deriverte prisfunksjonen kan tolkes som å indikere tilbudets implisitte priser, denne kan brukes til å beregne husholdningen marginale betalingsvilligheten for tilbudet. Den marginale betalingsvilligheten bidrar med til å informere om politikk. Den moderne hedoniske eiendomsverdimodellen er den fremste tilnærmingene til å verdsette endringer i miljøfasiliteter, da innen områder som akademisk forskning, rettssaker og offentlig politikk Bishop et al. (2020).

Roses (1974) hadde en banebrytende artikkel om det hedoniske rammeverket. Der han delte det inn i to trinn, første trinn handler om definere markedet og andre trinn om å samle inn data. Modellen blir sett på som en likevektsmodell for å kunne forstå hva differensierte produktpriser kunne avsløre om forbrukernes etterspørsel etter produktattributter. Likevekten er forholdet mellom boligpriser og huskarakteristikk, som igjen avslører hver kjøpers marginale betalingsvillighet. Den marginale betalingsvilligheten til kjøpere kan endres over tid, som kan gjenspeiles som en endring i funksjonen for de implisitte prisene for fasilitetene. Grunnen til at den marginale betalingsvilligheten kan endres over tid kan skyldes at en øker arbeidernes produktivitet, induserer migrasjon, gir ny informasjon om fasilitetene eller endrer bekvemmeligheters nivåer, dette kan være for eksempel endringer i regler i luftkvalitet Bishop et al. (2020).

Dette gir et grunnlag for å kunne bruke den hedoniske eiendomsverdimodellen, og en bruker den hedoniske modellen til å estimere marginale betalingsvilligheter for miljøfasiliteter Bishop et al. (2020).

Første trinnet i modellen er å definere et marked. Et marked bør bli definert slik at det tilfredsstiller "loven om én prisfunksjon." Det betyr at identiske hus bør bli solgt for samme pris i et gitt marked. En kan definere marked som et stort område over kort tid, men kan også være et større område over lengre tid. For å følge opp "loven om én prisfunksjon" gjelder det å defineret markedet som et stort område over kort tid. Da unngår en ta stilling til flytting, fordi de fysiske og økonomiske kostnadene ved å flytte ikke viser seg til å endre på ulike destinasjoner innenfor et stort område. I motsetning til dette er det mindre sannsynlig at loven om én prisfunksjon blir tilfredsstilt hvis markedet er definert til å omfatte flere storbyområder og/eller flere år. Om en skulle definert et marked som et større område over lengre tid er at det vil føre til større flyttekostnader, fordi det omfatter flere storbyområder. Dette ville påvirket den marginale betalingsvilligheten til flere, fordi en kan bli påtvunget flytting på grunn av for eksempel jobb, kan også være endringer i lokale skattepolitikk og levekostnader, som gjør at en velger å flytte Bishop et al. (2020).

I trinn to handler det om å samle inn data. Et tilfeldig utvalg er det beste når det kommer til datainnsamling i hedoniske eiendomsverdistudier, da av boligtransaksjonspriser og egenskaper for det aktuelle studieområdet og hovedsakelig for enebolig. Ved innsamling av data kan det oppstå utfordringer som kan oppstå i mindre enn ideelle datainnstillinger, inkludert regulering av priser, sparsomme transaksjoner og mangel på transaksjonspriser Bishop et al. (2020).

Når en skal samle inn data om boligsalg har en forventning på at informasjonen ligge offentlig tilgjengelig, her kan det oppstå problemer som gjør at en ikke fanger opp dette. Ved å identifisere og ta vekk disse dataene kan en redusere muligheten for å få målefeil. Dette kan være for eksempel at kjøper og selger har likt etternavn i et salg, fordi sannsynligheten for at de er slekt er stor. Er også vanlig å fjerne tvangssalg, og kjøp fra eiendomsinvesteringsforetak og uteliggere som tydelig indikerer datainntastingsfeil Bishop et al. (2020).

Det er også viktig å inkludere hvordan kjøpere oppfatter fasilitetsnivåene på hvert boligsted. Der forskerne må inkludere romlig interpolasjon, luftspredningsmodeller eller spådommer fra satellitter for å tildele forurensningsnivåer til hus. En kan også se på nærhet til rekreasjonssteder som strender, innsjøer og parker måles etter geografisk avstand, kjøreavstand, total reisetid eller andelen land som er viet til denne rekreasjonsbruken innenfor et geografisk område rundt et hus, det er for å se på hva som har en betydning for kjøperen Bishop et al. (2020).

Når dataen om salgspriser og kjennetegn ved en enemoligtransaskjon ikke er tilgjengelig, kan en bruke data om anslåtte priser, leiepriser og salg av barmark, samt romlig aggregerte oppsummeringsmål som middel eller medianer. Dette kan gi utfordringer for å kunne tolke prisfunksjonsparametere som mål på den marginale betalingsvilligheten. Det kan også være ideelt å bruke spørreundersøkelser for å kunne estimere en verdi. Det er også mulig å bruke eiendomsvurderinger eller andre selskaper. Transaksjonpriser er foretrukket fremfor en anslått pris, dette fordi predikerte priser er at de inkluderer målefeil, som korrelerer med kjøperens demografi boligkarakteristikker og nabolagsfasiliteter, og fører dermed til skjevheter i prisfunksjonens parameterestimater. Under leiepriser kan det oppstå uklarheter fordi det er opp til den som leier ut hva som er viktig, da om leieren betaler for vedlikehold og utstyr. Ofte leier en i kortere periode, og derfor prioritereren ikke fasiliteter eller naboelaget Bishop et al. (2020).

For å velge en økonometrisk spesifikasjon for den hedoniske prisfunksjonen, bør en først se på at prisfunksjonen antas å være ikke-lineær. Dette er fordi dette gir en mer nøyaktige estimater av gjennomsnittlig marginal betalingsvillighet for boligkarakteristikker, enn om det skulle vært lineære og lig-lineær. Det tillater også at markedslikevekt gjenspeiler komplementaritet mellom fasiliteter. Med tanke på kriminalitet, støy og luftkvalitet. Det er også viktig å se på at ingen informasjon er utelatt, som da er vitkig for huskjøperen. Da med tanke på skoler, nabolag og natur. Det viktigste med denne modellen er å se på hva kjøperens betallingsvilligheten er for miljøfasiliter Bishop et al. (2020).

Oppg. 2

I

Lastet først ned datasettet House Sales in King County, USA fra Kaggle. Deretter sjekket vi definisjonene om at de var riktige.

II

Leser inn hus salgene i King County i USA som vi har lastet ned fra Kaggle.

```
kc_house_data <- read_csv("kc_house_data.csv")
```

```
Rows: 21613 Columns: 21
-- Column specification --------------------------------------------------------
Delimiter: ","
chr   (1): id
dbl  (19): price, bedrooms, bathrooms, sqft_living, sqft_lot, floors, waterf...
dttm  (1): date

i Use `spec()` to retrieve the full column specification for this data.
i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

III

Nå sorterer vi salgene etter dato.

```
kc_house_data <- arrange(kc_house_data, desc(date))
```

IV

I denne omgang bruker vi dplyr:: distinct til å velge siste salg der vi multiple salg av samme eiendom.

```
kc_house_data <- kc_house_data %>%
  distinct(id, .keep_all = TRUE)
```

V og VI

Bruker st_as_sf() til å konvertere house data til et sf objekt vha. long lat og setter til geografisk projekson.

```
kc_house_data_sf <- st_as_sf(kc_house_data,coords = c(x = "long", y= "lat"), crs = 4326) %
  st_transform(2926)
```

VII

Koordinater Seattle som er hentet fra Wikipedia er : 47.3622, -122,1955

```
cbd <- st_sfc(st_point(c(-122.1955, 47.3622)), crs = 4326) %>%
  st_transform(2926)
```

VIII

Her finner vi avstanden mellom punktet EPSG:2926 og samtlige hus i datasettet i luftlinje. Deretter konverterer vi det til km og legger dem inn i variabelen dest_CBD.

```
kc_house_data_sf <- kc_house_data_sf %>% mutate(
  dist_cbd = st_distance(cbd, ., by_element = TRUE),
  dist_cbd_km = set_units(dist_cbd, km)
)
```

Oppg. 3

I og III

Leser inn filen WADOH King County.

```
kc_wadoh_map <- here("WADOH_Environmental_Health_Disparities_Index_Calculated_for_King_Cou
  st_read() %>%
  st_transform(2926)
```

```
Reading layer `WADOH_Environmental_Health_Disparities_Index_Calculated_for_King_County___wad
  using driver `ESRI Shapefile'
Simple feature collection with 398 features and 192 fields
Geometry type: MULTIPOLYGON
Dimension:     XY
Bounding box:  xmin: -122.528 ymin: 47.08446 xmax: -121.0657 ymax: 47.78058
Geodetic CRS:  WGS 84
```

II

Plukker ut variablene som er angitt i oppgaveteksten.

```
kc_wadoh_map <- kc_wadoh_map %>%
     select(
     GEO_ID_TRT,
     EHD_percen,#Environmental Health Index, weighted score many vars
     linguist_2,#Pop. age 5+ speaking English less than "very well"
     poverty_pe,#Percentage people living in poverty
     POC_percen,#People of Color in percentage of pop. in tract
     transporta,#% of income spent on transportation median family in tract
     unemploy_2,#percentage unemployed
     housing_pe,#% of households in group "Unaffordable Housing" (>30% inc.)
     traffic_pe,#% of pop. near heavy traffic roadways
     diesel,# nox consentration
     ozone,# ozone consentration
     PM25, # consentration of Particulate Matter in air
     toxic_rele, # Toxic release from factories
     hazardous_, # Hazardous Waste Treatment Storage and disposal Facilities
     lead_perce, # measure of Lead paint in houses
     superfund, # Proximity to contaminated sites on national list
     facilities, # Proximity to Risk Management Plan Facilities
```

```
        wastewater, # Proximity to wastewater facilities
        sen_pop_pe, # % pop. over 65
        socio_perc # score social economic determants, low best
        )
```

IV

Laster inn acs_b19101_familyincome.

```
acs_b19101_fam_inc <- read.dbf("../Maps/censusSHP/acs_b19101_familyincome.dbf")
```

Lager inntektsvariabelen.

```
acs_b19101_fam_inc <- acs_b19101_fam_inc %>%
  mutate(low = (E19101138 + E19101139 + E19101140 + E19101141 + E19101142 + E19101143)/E19
  mutate(mid = (E19101144 + E19101145 + E19101146 + E19101147 + E19101148 + E19101149)/E19
  mutate(high = (E19101150 + E19101151 + E19101152 + E19101153)/E19101137)
```

```
acs_b19101_fam_inc <- acs_b19101_fam_inc %>%
  select(GEOIDTRT, low, mid, high) %>%
  rename(GEO_ID_TRT = GEOIDTRT)
```

```
kc_wadoh_map_2 <- left_join(acs_b19101_fam_inc, st_drop_geometry(kc_wadoh_map), by = "GEO_
```

```
kc_tracts10 <- here("../Maps/censusSHP/tracts10.shp") %>%
  st_read() %>%
  st_transform(2926)
```

```
Reading layer `tracts10' from data source
  `/Users/kinemakestad/Documents/Master i sivilokonom/3. Semester/Boligmarked og spatial økon
  using driver `ESRI Shapefile'
Simple feature collection with 398 features and 22 fields
Geometry type: POLYGON
Dimension:     XY
Bounding box:  xmin: 1217085 ymin: 31406.52 xmax: 1583210 ymax: 287947.2
Projected CRS: NAD83(HARN) / Washington North (ftUS)
```

```r
kc_tracts10_shore <- here("../Maps/censusSHP/tracts10_shore.shp") %>%
  st_read() %>%
  st_transform(2926)
```

```
Reading layer `tracts10_shore' from data source
  `/Users/kinemakestad/Documents/Master i sivilokonom/3. Semester/Boligmarked og spatial øko
  using driver `ESRI Shapefile'
Simple feature collection with 398 features and 22 fields
Geometry type: MULTIPOLYGON
Dimension:     XY
Bounding box:  xmin: 1220306 ymin: 31406.52 xmax: 1583210 ymax: 287675.5
Projected CRS: NAD83(HARN) / Washington North (ftUS)
```

Nå blir det brukt left join for å legge dataene inn i WADOH King County.

```r
kc_tracts10_env_data <- left_join(kc_tracts10, kc_wadoh_map_2, by = "GEO_ID_TRT")
```

```r
kc_tracts10_shore_env_data <- left_join(kc_tracts10_shore, kc_wadoh_map_2, by= "GEO_ID_TRT
```

```r
summary(kc_tracts10)
```

```
  GEO_ID_TRT          FEATURE_ID      TRACT_LBL          TRACT_STR        
 Length:398         Min.   :10153   Length:398         Length:398        
 Class :character   1st Qu.:25818   Class :character   Class :character  
 Mode  :character   Median :44344   Mode  :character   Mode  :character  
                    Mean   :36731                                        
                    3rd Qu.:45226                                        
                    Max.   :45837                                        
   TRACT_INT         TRACT_FLT         TRACT_DEL          TRTLABEL_F       
 Min.   :   100   Min.   :   1.00   Length:398         Length:398        
 1st Qu.:  9625   1st Qu.:  96.25   Class :character   Class :character  
 Median : 24150   Median : 241.50   Mode  :character   Mode  :character  
 Mean   : 23022   Mean   : 230.22                                        
 3rd Qu.: 30076   3rd Qu.: 300.76                                        
 Max.   :990100   Max.   :9901.00                                        
  TRTLABEL_C          TRTLABEL_T          COUNTY_STR           COUNTY_INT
 Length:398         Length:398         Length:398         Min.   :33  
 Class :character   Class :character   Class :character   1st Qu.:33  
 Mode  :character   Mode  :character   Mode  :character   Median :33  
                                                          Mean   :33  
```

```
                                                         3rd Qu.:33
                                                         Max.   :33
  STATE_STR           STATE_INT     LEVEL_1             LEVEL_2
 Length:398         Min.   :53   Length:398         Length:398
 Class :character   1st Qu.:53   Class :character   Class :character
 Mode  :character   Median :53   Mode  :character   Mode  :character
                    Mean   :53
                    3rd Qu.:53
                    Max.   :53
   LEVEL_3            TRACT_AREA          TRACT_PERI       LOGRECNO
 Length:398         Min.   :2.466e+06   Min.   :  7060   Length:398
 Class :character   1st Qu.:1.933e+07   1st Qu.: 20586   Class :character
 Mode  :character   Median :3.362e+07   Median : 29573   Mode  :character
                    Mean   :1.616e+08   Mean   : 44019
                    3rd Qu.:5.601e+07   3rd Qu.: 43667
                    Max.   :1.526e+10   Max.   :738820
   Shape_area          Shape_len                 geometry
 Min.   :2.466e+06   Min.   :  7060   POLYGON       :398
 1st Qu.:1.933e+07   1st Qu.: 20586   epsg:2926     :  0
 Median :3.362e+07   Median : 29573   +proj=lcc ...:  0
 Mean   :1.616e+08   Mean   : 44019
 3rd Qu.:5.601e+07   3rd Qu.: 43667
 Max.   :1.526e+10   Max.   :738820
```

```r
kc_house_env_var <- st_join(kc_house_data_sf, kc_tracts10_env_data)
kc_tracts10_shore_env_var <- st_join(kc_house_data_sf, kc_tracts10_shore_env_data)
```

```r
st_write(kc_house_data, "../Maps/kc_house_data.gpkg", append = FALSE)
```

```
Deleting layer `kc_house_data' using driver `GPKG'
Writing layer `kc_house_data' to data source
  `../Maps/kc_house_data.gpkg' using driver `GPKG'
Writing 21436 features with 21 fields without geometries.
```

```r
st_write(kc_tracts10, "../Maps/kc_tracts10.gpkg", append = FALSE)
```

```
Deleting layer `kc_tracts10' using driver `GPKG'
Writing layer `kc_tracts10' to data source
  `../Maps/kc_tracts10.gpkg' using driver `GPKG'
Writing 398 features with 22 fields and geometry type Polygon.
```

```r
st_write(kc_tracts10_shore, "../Maps/kc_tracts10_shore.gpkg", append = FALSE)
```

```
Deleting layer `kc_tracts10_shore' using driver `GPKG'
Writing layer `kc_tracts10_shore' to data source
  `../Maps/kc_tracts10_shore.gpkg' using driver `GPKG'
Writing 398 features with 22 fields and geometry type Multi Polygon.
```

```r
st_write(kc_house_env_var, "../Maps/kc_house_env_var.gpkg", append = FALSE)
```

```
Deleting layer `kc_house_env_var' using driver `GPKG'
Writing layer `kc_house_env_var' to data source
  `../Maps/kc_house_env_var.gpkg' using driver `GPKG'
Writing 21436 features with 65 fields and geometry type Point.
```

```r
st_write(kc_tracts10_shore_env_var, "../Maps/kc_tracts10_shore_env_var.gpkg", append = FAL
```

```
Deleting layer `kc_tracts10_shore_env_var' using driver `GPKG'
Writing layer `kc_tracts10_shore_env_var' to data source
  `../Maps/kc_tracts10_shore_env_var.gpkg' using driver `GPKG'
Writing 21436 features with 65 fields and geometry type Point.
```

Oppg. 4

I

Her sjekker vi områdevariablene fra WADOH ved hjelp av summary for både tracts10 og tracts10 shore.

```r
summary(kc_tracts10_env_data)
```

```
  GEO_ID_TRT          FEATURE_ID      TRACT_LBL           TRACT_STR
 Length:398         Min.   :10153   Length:398         Length:398
 Class :character   1st Qu.:25818   Class :character   Class :character
 Mode  :character   Median :44344   Mode  :character   Mode  :character
                    Mean   :36731
                    3rd Qu.:45226
                    Max.   :45837
```

```
   TRACT_INT         TRACT_FLT         TRACT_DEL           TRTLABEL_F       
 Min.   :   100   Min.   :   1.00   Length:398         Length:398        
 1st Qu.:  9625   1st Qu.:  96.25   Class :character   Class :character  
 Median : 24150   Median : 241.50   Mode  :character   Mode  :character  
 Mean   : 23022   Mean   : 230.22                                        
 3rd Qu.: 30076   3rd Qu.: 300.76                                        
 Max.   :990100   Max.   :9901.00                                        

  TRTLABEL_C         TRTLABEL_T         COUNTY_STR          COUNTY_INT
 Length:398         Length:398         Length:398         Min.   :33  
 Class :character   Class :character   Class :character   1st Qu.:33  
 Mode  :character   Mode  :character   Mode  :character   Median :33  
                                                          Mean   :33  
                                                          3rd Qu.:33  
                                                          Max.   :33  

  STATE_STR           STATE_INT    LEVEL_1            LEVEL_2         
 Length:398         Min.   :53   Length:398         Length:398        
 Class :character   1st Qu.:53   Class :character   Class :character  
 Mode  :character   Median :53   Mode  :character   Mode  :character  
                    Mean   :53                                        
                    3rd Qu.:53                                        
                    Max.   :53                                        

   LEVEL_3            TRACT_AREA          TRACT_PERI        LOGRECNO        
 Length:398         Min.   :2.466e+06   Min.   :  7060   Length:398        
 Class :character   1st Qu.:1.933e+07   1st Qu.: 20586   Class :character  
 Mode  :character   Median :3.362e+07   Median : 29573   Mode  :character  
                    Mean   :1.616e+08   Mean   : 44019                     
                    3rd Qu.:5.601e+07   3rd Qu.: 43667                     
                    Max.   :1.526e+10   Max.   :738820                     

   Shape_area          Shape_len           low                 mid        
 Min.   :2.466e+06   Min.   :  7060   Min.   :0.009298   Min.   :0.0000  
 1st Qu.:1.933e+07   1st Qu.: 20586   1st Qu.:0.053302   1st Qu.:0.2391  
 Median :3.362e+07   Median : 29573   Median :0.092424   Median :0.3339  
 Mean   :1.616e+08   Mean   : 44019   Mean   :0.125013   Mean   :0.3327  
 3rd Qu.:5.601e+07   3rd Qu.: 43667   3rd Qu.:0.166534   3rd Qu.:0.4261  
 Max.   :1.526e+10   Max.   :738820   Max.   :1.000000   Max.   :0.6790  
                                      NA's   :1          NA's   :1       
      high          EHD_percen       linguist_2      poverty_pe   
 Min.   :0.0000   Min.   :  1.00   Min.   : 0.45   Min.   : 1.97  
```

```
1st Qu.:0.4006   1st Qu.: 25.00   1st Qu.: 3.88   1st Qu.:10.53
Median :0.5637   Median : 50.00   Median : 8.72   Median :16.75
Mean   :0.5423   Mean   : 50.38   Mean   :10.62   Mean   :20.42
3rd Qu.:0.6955   3rd Qu.: 75.00   3rd Qu.:15.38   3rd Qu.:27.48
Max.   :0.8816   Max.   :100.00   Max.   :46.76   Max.   :75.48
NA's   :1        NA's   :1        NA's   :5       NA's   :1
  POC_percen      transporta      unemploy_2       housing_pe
Min.   : 7.54   Min.   :10.00   Min.   : 1.000   Min.   :15.14
1st Qu.:23.36   1st Qu.:18.00   1st Qu.: 3.350   1st Qu.:27.34
Median :36.29   Median :19.00   Median : 4.480   Median :32.26
Mean   :38.64   Mean   :18.97   Mean   : 5.099   Mean   :33.75
3rd Qu.:51.01   3rd Qu.:21.00   3rd Qu.: 6.460   3rd Qu.:39.13
Max.   :92.70   Max.   :26.00   Max.   :24.400   Max.   :81.89
NA's   :1       NA's   :1       NA's   :3        NA's   :1
  traffic_pe        diesel          ozone             PM25
Min.   : 0.00   Min.   : 0.14   Min.   :46.73   Min.   :3.787
1st Qu.: 0.00   1st Qu.: 6.65   1st Qu.:48.91   1st Qu.:5.642
Median : 3.60   Median :12.65   Median :49.78   Median :6.180
Mean   :16.07   Mean   :17.10   Mean   :50.62   Mean   :6.186
3rd Qu.:26.17   3rd Qu.:18.99   3rd Qu.:51.28   3rd Qu.:6.872
Max.   :97.75   Max.   :92.63   Max.   :62.89   Max.   :7.897
NA's   :1       NA's   :1       NA's   :1       NA's   :1
  toxic_rele         hazardous_         lead_perce       superfund
Min.   :   823.9   Min.   :0.02303   Min.   : 0.24   Min.   :0.03454
1st Qu.:  5180.9   1st Qu.:0.04168   1st Qu.: 6.46   1st Qu.:0.07358
Median : 10186.5   Median :0.05160   Median :13.79   Median :0.13133
Mean   : 19398.3   Mean   :0.08190   Mean   :17.08   Mean   :0.24645
3rd Qu.: 20058.1   3rd Qu.:0.09280   3rd Qu.:26.20   3rd Qu.:0.28436
Max.   :186434.6   Max.   :0.63781   Max.   :54.68   Max.   :1.46778
NA's   :1          NA's   :1         NA's   :1       NA's   :1
  facilities       wastewater         sen_pop_pe        socio_perc
Min.   :0.0523   Min.   :0.00e+00   Min.   :  1.00   Min.   :  1.00
1st Qu.:0.1612   1st Qu.:5.50e-06   1st Qu.: 25.00   1st Qu.: 25.00
Median :0.3652   Median :5.30e-04   Median : 50.00   Median : 50.00
Mean   :0.6046   Mean   :2.62e-02   Mean   : 50.38   Mean   : 50.38
3rd Qu.:0.9119   3rd Qu.:8.70e-03   3rd Qu.: 75.00   3rd Qu.: 75.00
Max.   :3.3682   Max.   :6.40e-01   Max.   :100.00   Max.   :100.00
NA's   :1        NA's   :1          NA's   :1        NA's   :1
         geometry
POLYGON      :398
epsg:2926    :  0
+proj=lcc ...:  0
```

```r
summary(kc_tracts10_shore_env_data)
```

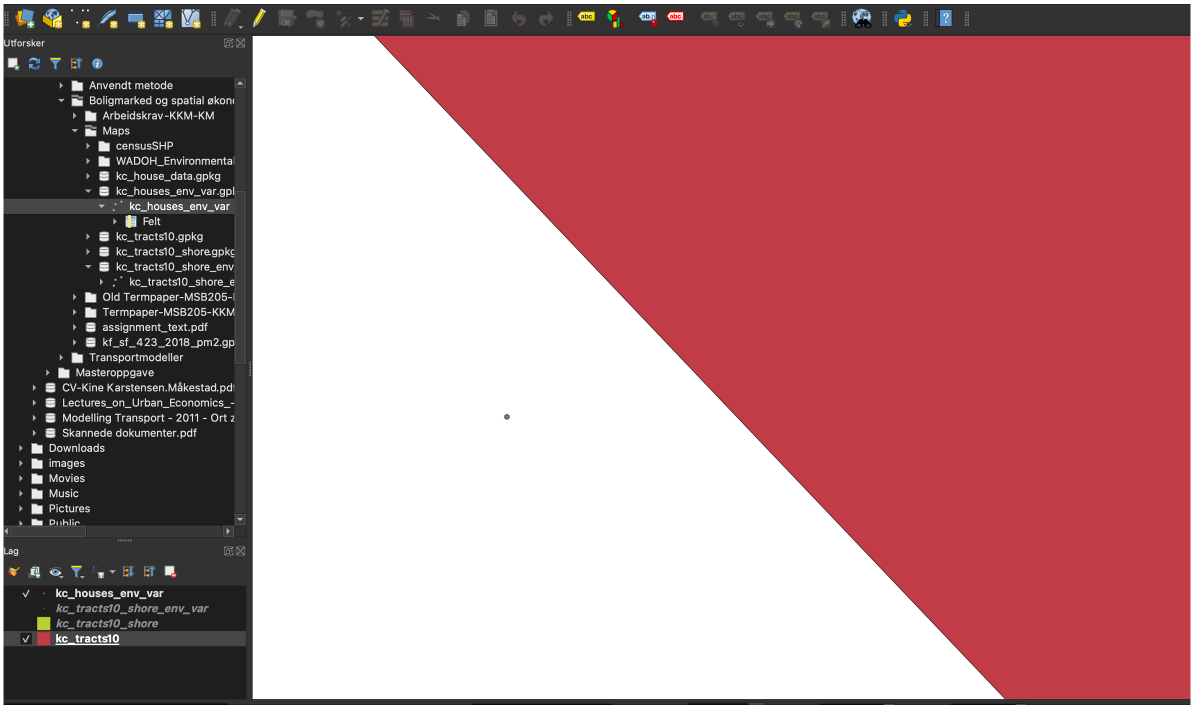
```
  GEO_ID_TRT           FEATURE_ID      TRACT_LBL          TRACT_STR
 Length:398         Min.   :10153   Length:398         Length:398
 Class :character   1st Qu.:27069   Class :character   Class :character
 Mode  :character   Median :44458   Mode  :character   Mode  :character
                    Mean   :36834
                    3rd Qu.:45197
                    Max.   :45838
   TRACT_INT         TRACT_FLT         TRACT_DEL          TRTLABEL_F
 Min.   :   100   Min.   :   1.00   Length:398         Length:398
 1st Qu.:  9625   1st Qu.:  96.25   Class :character   Class :character
 Median : 24150   Median : 241.50   Mode  :character   Mode  :character
 Mean   : 23022   Mean   : 230.22
 3rd Qu.: 30076   3rd Qu.: 300.76
 Max.   :990100   Max.   :9901.00
  TRTLABEL_C         TRTLABEL_T         COUNTY_STR          COUNTY_INT
 Length:398         Length:398         Length:398         Min.   :33
 Class :character   Class :character   Class :character   1st Qu.:33
 Mode  :character   Mode  :character   Mode  :character   Median :33
                                                          Mean   :33
                                                          3rd Qu.:33
                                                          Max.   :33
  STATE_STR           STATE_INT    LEVEL_1            LEVEL_2
 Length:398         Min.   :53   Length:398         Length:398
 Class :character   1st Qu.:53   Class :character   Class :character
 Mode  :character   Median :53   Mode  :character   Mode  :character
                    Mean   :53
                    3rd Qu.:53
                    Max.   :53
   LEVEL_3            TRACT_AREA          TRACT_PERI        LOGRECNO
 Length:398         Min.   :2.466e+06   Min.   :  7060   Length:398
 Class :character   1st Qu.:1.933e+07   1st Qu.: 20586   Class :character
```

```
Mode  :character   Median :3.362e+07   Median : 29573   Mode  :character
                   Mean   :1.616e+08   Mean   : 44023
                   3rd Qu.:5.601e+07   3rd Qu.: 43919
                   Max.   :1.526e+10   Max.   :738820

  Shape_area          Shape_len          low                 mid
Min.   :7.819e+05   Min.   :  7060   Min.   :0.009298   Min.   :0.0000
1st Qu.:1.794e+07   1st Qu.: 20297   1st Qu.:0.053302   1st Qu.:0.2391
Median :2.964e+07   Median : 28874   Median :0.092424   Median :0.3339
Mean   :1.504e+08   Mean   : 41303   Mean   :0.125013   Mean   :0.3327
3rd Qu.:5.020e+07   3rd Qu.: 40590   3rd Qu.:0.166534   3rd Qu.:0.4261
Max.   :1.526e+10   Max.   :738820   Max.   :1.000000   Max.   :0.6790
                                     NA's   :1          NA's   :1
     high          EHD_percen       linguist_2      poverty_pe
Min.   :0.0000   Min.   :  1.00   Min.   : 0.45   Min.   : 1.97
1st Qu.:0.4006   1st Qu.: 25.00   1st Qu.: 3.88   1st Qu.:10.53
Median :0.5637   Median : 50.00   Median : 8.72   Median :16.75
Mean   :0.5423   Mean   : 50.38   Mean   :10.62   Mean   :20.42
3rd Qu.:0.6955   3rd Qu.: 75.00   3rd Qu.:15.38   3rd Qu.:27.48
Max.   :0.8816   Max.   :100.00   Max.   :46.76   Max.   :75.48
NA's   :1        NA's   :1        NA's   :5       NA's   :1
  POC_percen      transporta      unemploy_2       housing_pe
Min.   : 7.54   Min.   :10.00   Min.   : 1.000   Min.   :15.14
1st Qu.:23.36   1st Qu.:18.00   1st Qu.: 3.350   1st Qu.:27.34
Median :36.29   Median :19.00   Median : 4.480   Median :32.26
Mean   :38.64   Mean   :18.97   Mean   : 5.099   Mean   :33.75
3rd Qu.:51.01   3rd Qu.:21.00   3rd Qu.: 6.460   3rd Qu.:39.13
Max.   :92.70   Max.   :26.00   Max.   :24.400   Max.   :81.89
NA's   :1       NA's   :1       NA's   :3        NA's   :1
  traffic_pe        diesel          ozone            PM25
Min.   : 0.00   Min.   : 0.14   Min.   :46.73   Min.   :3.787
1st Qu.: 0.00   1st Qu.: 6.65   1st Qu.:48.91   1st Qu.:5.642
Median : 3.60   Median :12.65   Median :49.78   Median :6.180
Mean   :16.07   Mean   :17.10   Mean   :50.62   Mean   :6.186
3rd Qu.:26.17   3rd Qu.:18.99   3rd Qu.:51.28   3rd Qu.:6.872
Max.   :97.75   Max.   :92.63   Max.   :62.89   Max.   :7.897
NA's   :1       NA's   :1       NA's   :1       NA's   :1
  toxic_rele         hazardous_         lead_perce       superfund
Min.   :   823.9   Min.   :0.02303   Min.   : 0.24   Min.   :0.03454
1st Qu.:  5180.9   1st Qu.:0.04168   1st Qu.: 6.46   1st Qu.:0.07358
Median : 10186.5   Median :0.05160   Median :13.79   Median :0.13133
Mean   : 19398.3   Mean   :0.08190   Mean   :17.08   Mean   :0.24645
3rd Qu.: 20058.1   3rd Qu.:0.09280   3rd Qu.:26.20   3rd Qu.:0.28436
```

```
 Max.   :186434.6   Max.   :0.63781   Max.   :54.68   Max.   :1.46778
 NA's   :1          NA's   :1         NA's   :1       NA's   :1
   facilities       wastewater         sen_pop_pe       socio_perc
 Min.   :0.0523   Min.   :0.00e+00   Min.   :  1.00   Min.   :  1.00
 1st Qu.:0.1612   1st Qu.:5.50e-06   1st Qu.: 25.00   1st Qu.: 25.00
 Median :0.3652   Median :5.30e-04   Median : 50.00   Median : 50.00
 Mean   :0.6046   Mean   :2.62e-02   Mean   : 50.38   Mean   : 50.38
 3rd Qu.:0.9119   3rd Qu.:8.70e-03   3rd Qu.: 75.00   3rd Qu.: 75.00
 Max.   :3.3682   Max.   :6.40e-01   Max.   :100.00   Max.   :100.00
 NA's   :1        NA's   :1          NA's   :1        NA's   :1
          geometry
 MULTIPOLYGON :398
 epsg:2926    :  0
 +proj=lcc ...:  0
```

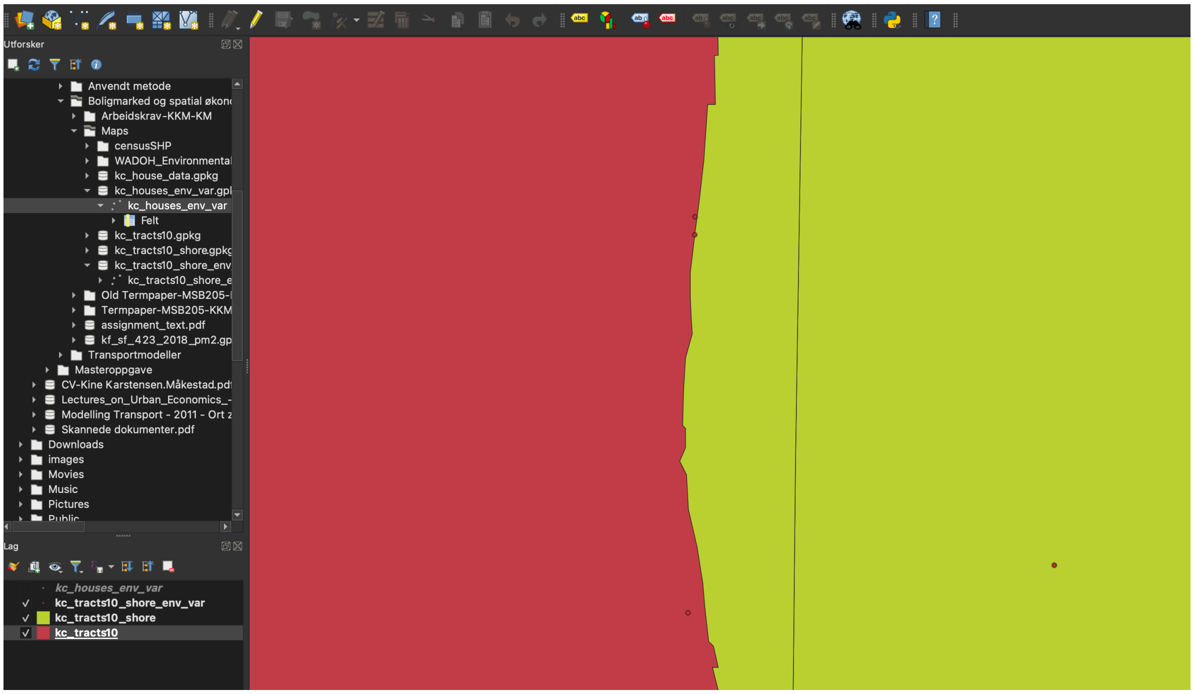
II

Når vi ser på disse dataene fra tracts10 både i R og i QGIS, så ser vi at kc_tracts10_env_data har en observasjon som går utenfor kommunegrensene.



Som en kan se på dette bilde er det den lille prikken som ligger utenfor kommunegrensen.

For den andre kc_tracts10_shore_env_data kan en se at det er ligger alle observasjonene innenfor kommunegrensene. Men nå vi la til shore-kartet kan en se at ved vannlinjen er det 25 observasjoner som ligger rett utenfor. Det er også grunnen til at det gir 25 NA.



iii

Her har vi droppet områdevariabelen med id 3518000180.

```r
kc_house_env_var <- arrange(kc_house_env_var, desc(id))
kc_house_env_var_omit <- kc_house_env_var[-c(11997),]
```

```r
st_write(kc_house_env_var_omit, "../Maps/kc_house_env_var_omit.gpkg", append = FALSE)
```

```
Deleting layer `kc_house_env_var_omit' using driver `GPKG'
Writing layer `kc_house_env_var_omit' to data source
  `../Maps/kc_house_env_var_omit.gpkg' using driver `GPKG'
Writing 21435 features with 65 fields and geometry type Point.
```

Nå skal vi lage en faktor-variabel av år og måned ut fra Date.

```r
kc_house_env_var_omit <- kc_house_env_var_omit %>%
  mutate(year_month = substr(date, start = 1, stop = 7))
```

Deretter skal vi slette gpkg filen fra husdataene og lagre den påp ny.

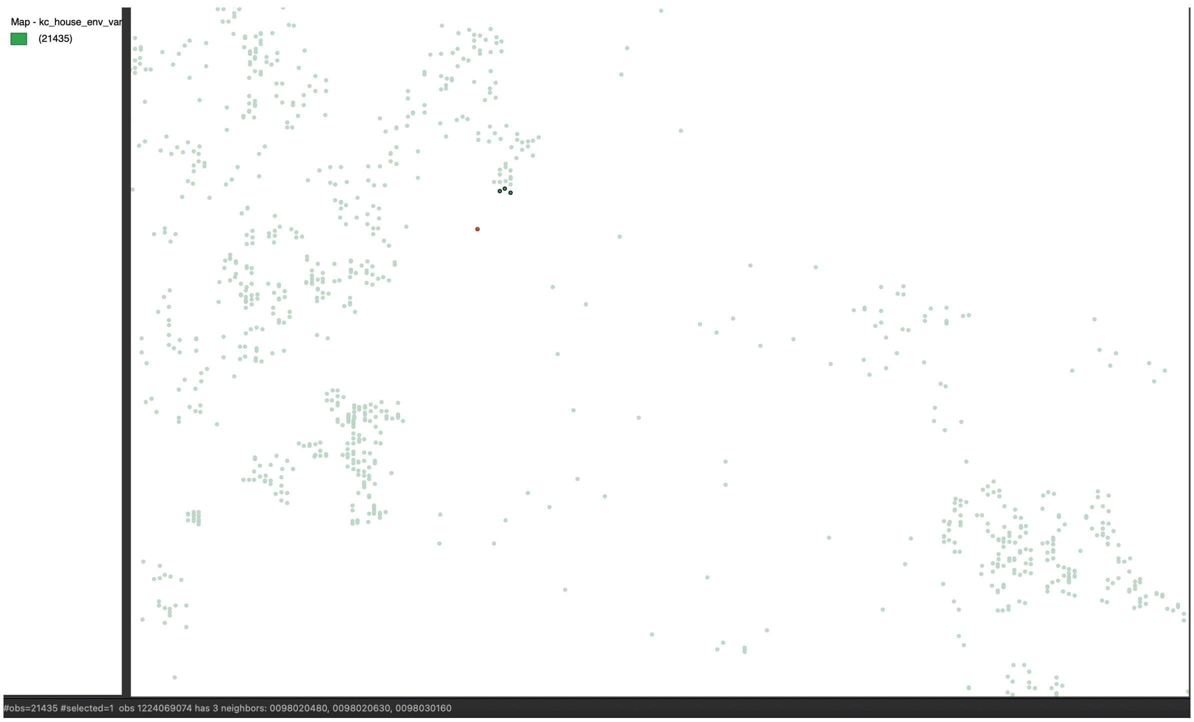
```r
st_write(kc_house_env_var_omit, "../Maps/kc_house_env_var_omit.gpkg", append = FALSE)
```

```
Deleting layer `kc_house_env_var_omit' using driver `GPKG'
Writing layer `kc_house_env_var_omit' to data source
  `../Maps/kc_house_env_var_omit.gpkg' using driver `GPKG'
Writing 21435 features with 66 fields and geometry type Point.
```

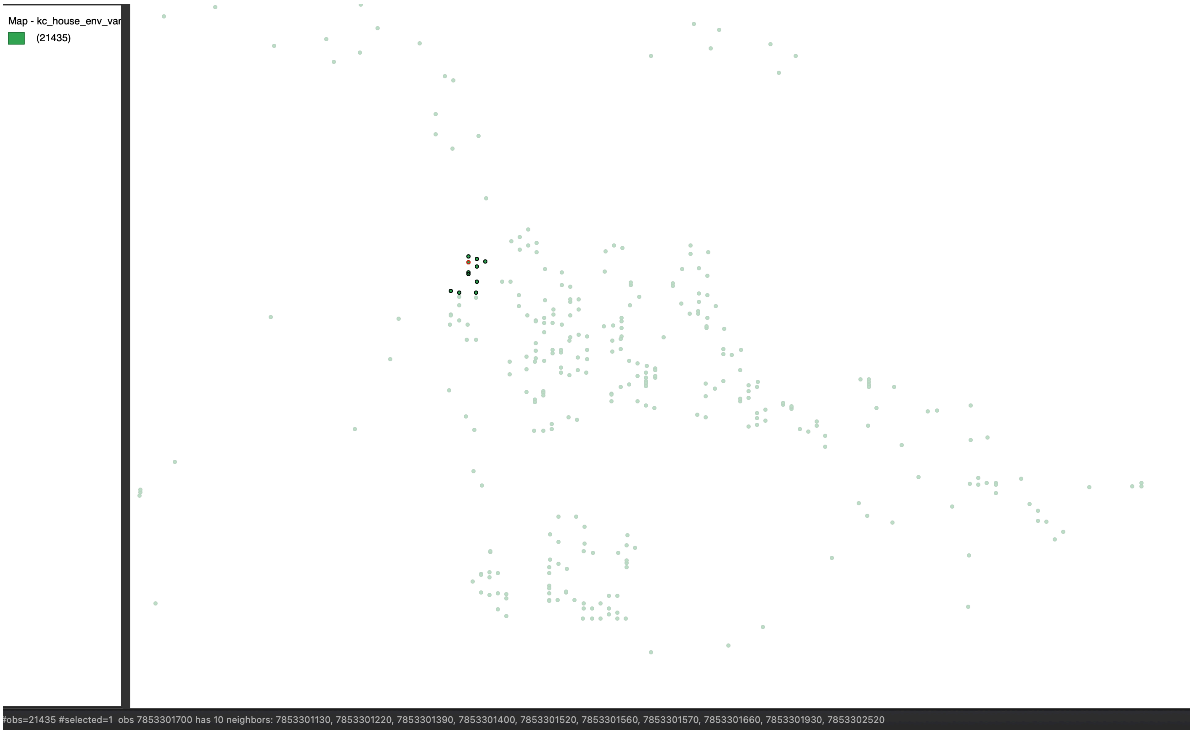
Oppg. 5

III

Her generer vi en vekt fil utfra 3 nærmeste naboer.



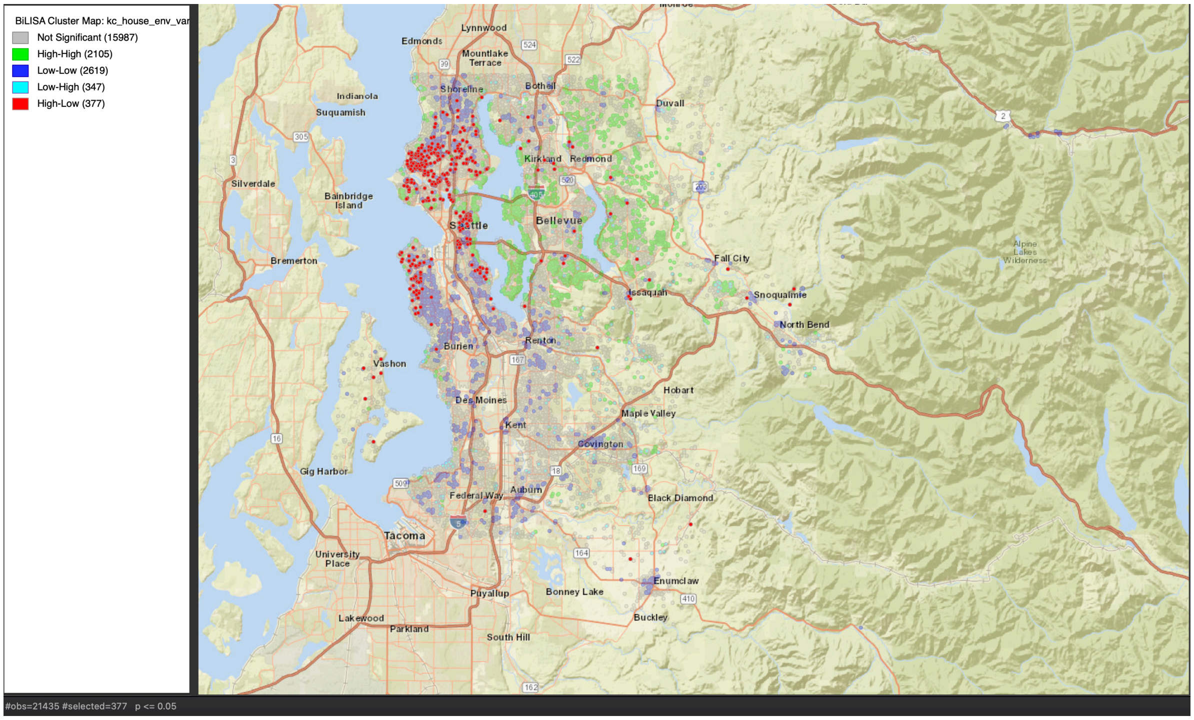
Nå skal vi generere en fil utfra en en plass med 10 nærmeste naboer.



IV

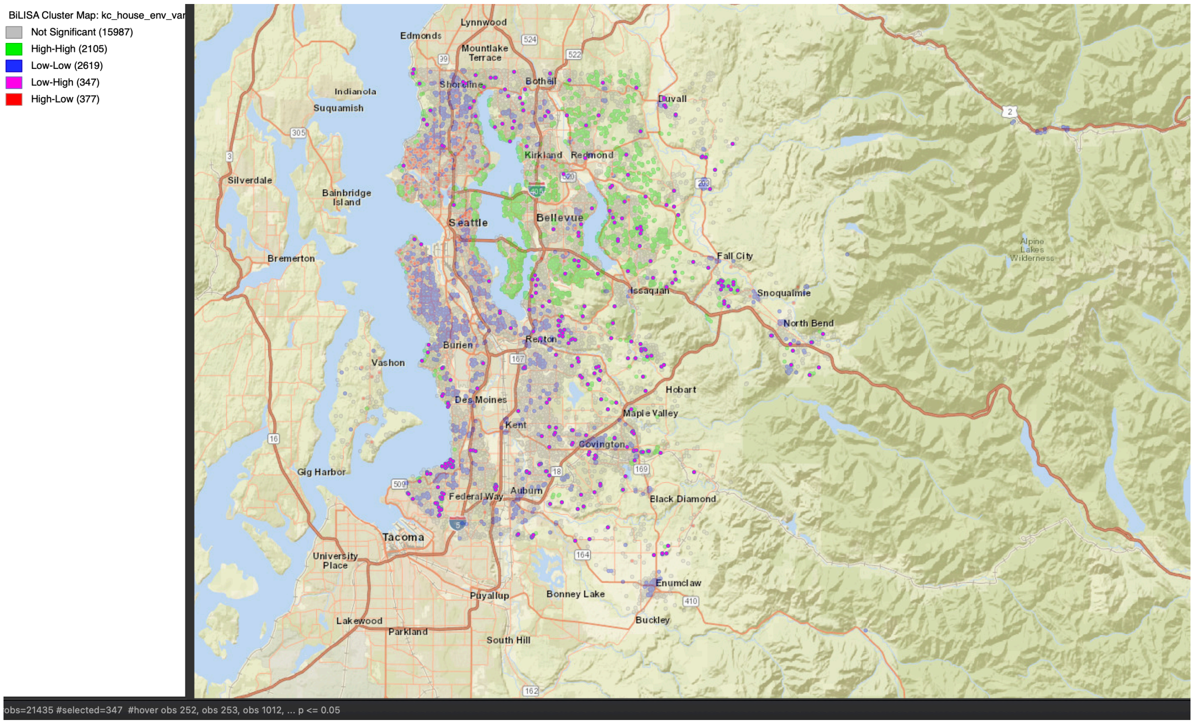
Nærmeste nabo 3:

Store og billige boliger.

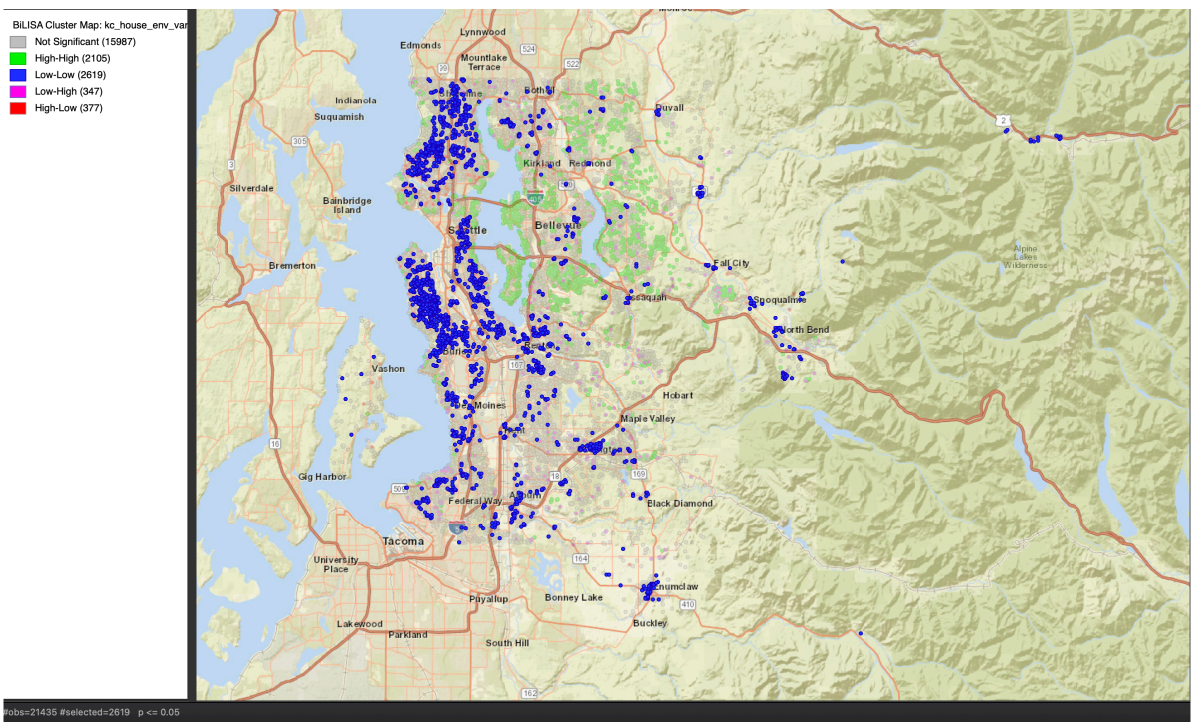


På bilde ovenfor ser man på de røde prikkene, og det er de plassene som er store og små boliger.

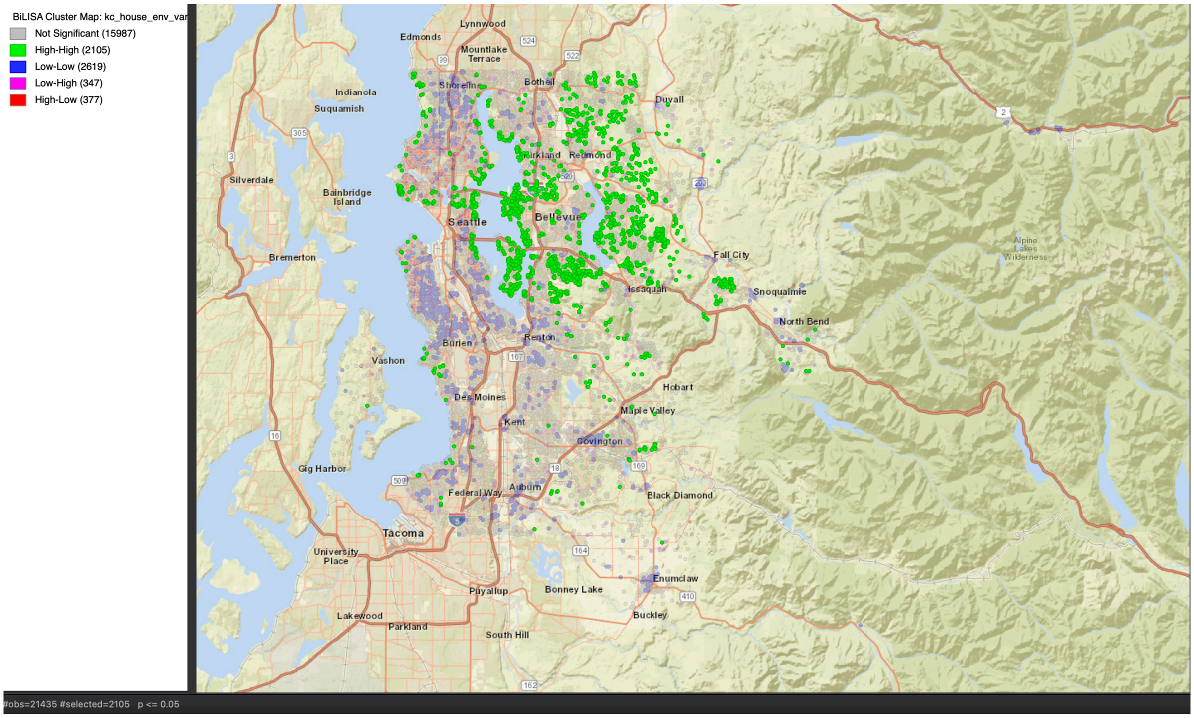
Neste er små og dyre boliger som er markert i rosa.



Deretter kommer de små og billige boligene som er markert i blå.

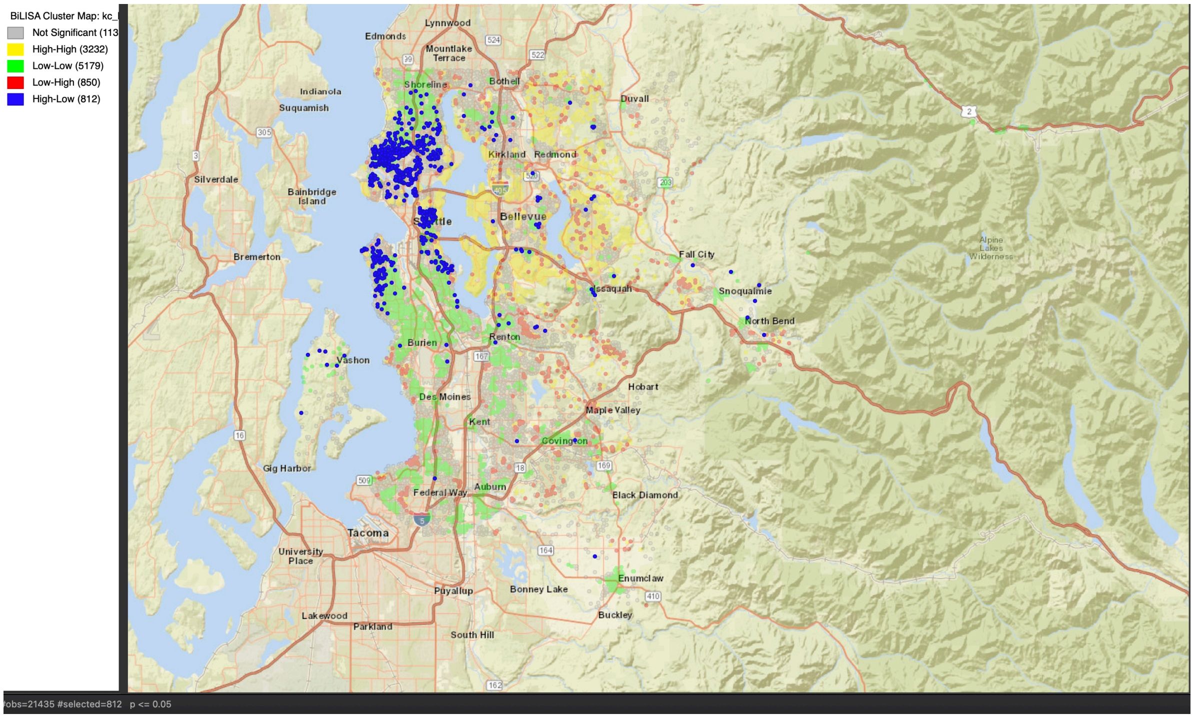


Til slutt er det de store og dyre boligene som er markert i grønn.

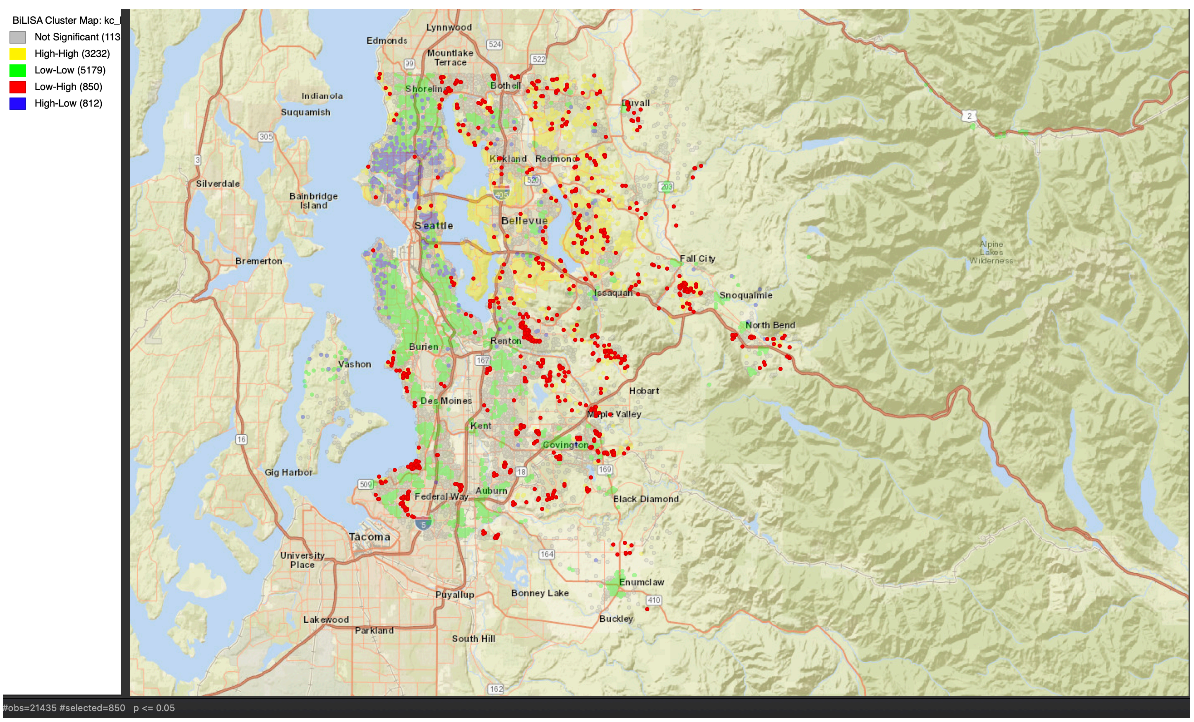


Nærmeste nabo 10:

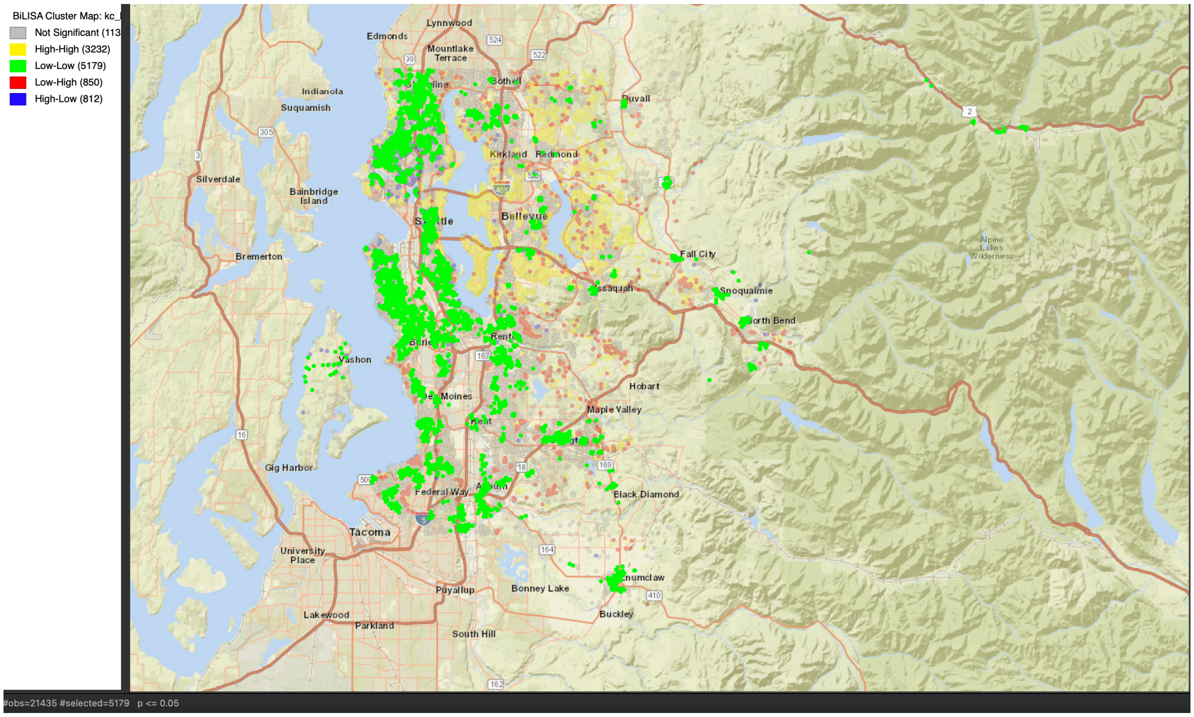
Første er store og billige boligene som en kan se på bilde nedenfor som er markert i blått.



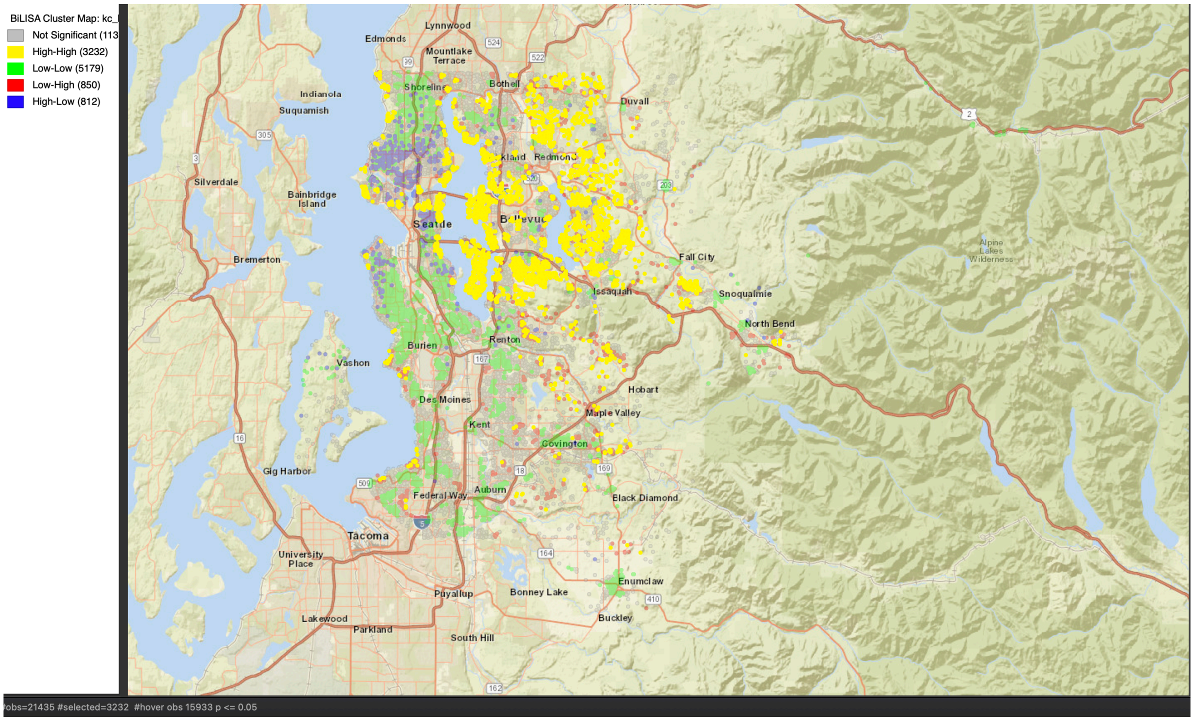
Så kommer små og dyre boliger, og det er markert i rødt.



Deretter kommer små og billige boliger, og dette er markert med grønt på kartet.



Til slutt kommer store og dyre boligene, og de er markert med gult på kartet nedenfor.



Oppg. 6

I Sammenfatning av funnene fra EDA

Etter å ha sett på pris i forhold til størrels epå boligen, kan en se at de små og billige boligene sentrer seg rundt Seattle sentrum. Da har vi sett i fordelingen på knear = 3 og knear = 10, der begge er nokså like. Ser en på de store og dyre boligene er de både med tre og ti naboer plasser rundt Bellevue og mer øst for Seattle. Under små og dyr boliger er mye sprendt rundt Seattle, mest utenfor sentrum området. Mens store og billige boliger er veldig sentrert i sentrum av Seattle, som de andre er det likt både med tre naboer to naboer.

Etter å ha sett på Moran´s I får en verdier på tre naboer = 0,399 og med ti naboer = 0.350. Ved positiv verdi vil det si at det klynger, og perfekt klynging er 1, der våre verdier viser til litt klyning, men også psredning. Noe en kan se i kartene ovenfor. Dette vil stemme overns med Moran´s I verdier som sier at perfekt tilnærming er 0.

II

1

Huskarakteristiska og tids-dummie

```r
mod1 <- "price ~ bedrooms + bathrooms + sqft_living + sqft_living15 + sqft_lot + sqft_lot1
```

2

Huskarateristika, dist_CBD, relevante tract variabler og tids-dummier.

```r
mod2 <- "price ~ bedrooms + bathrooms + sqft_living + sqft_living15 + sqft_lot + sqft_lot1
hazardous_ + lead_perce + socio_perc"
```

3

Huskarateristika, dist_CBD, EHD indeks og tids-dummier.

```r
mod3 <- "price ~ bedrooms + bathrooms + sqft_living + sqft_living15 + sqft_lot + sqft_lot1
```

Hedoniske modeller

```r
hedon1 <- lm(mod1, data = kc_house_env_var_omit)
hedon2 <- lm(mod2, data = kc_house_env_var_omit)
hedon3 <- lm(mod3, data = kc_house_env_var_omit)
```

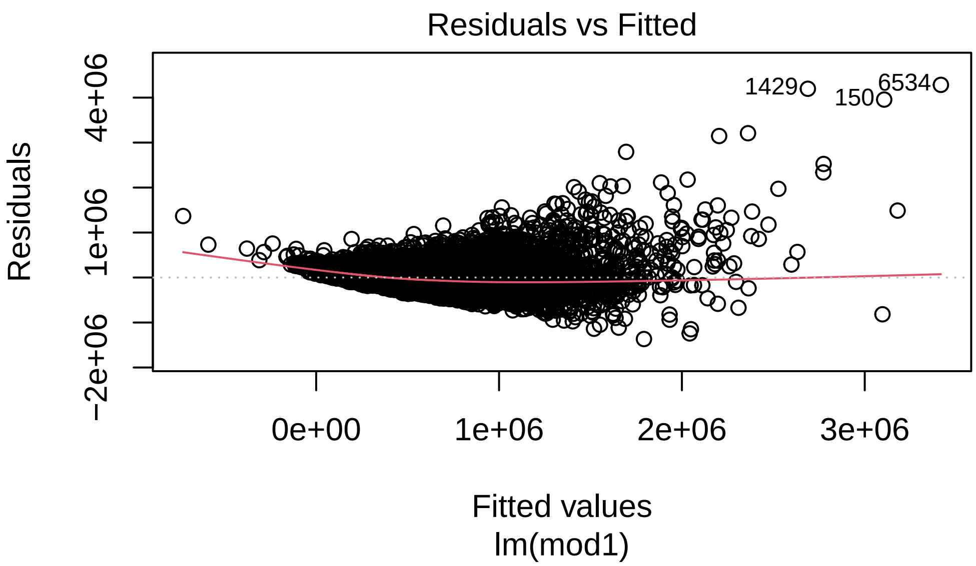
```r
huxreg("Hedon1" = hedon1, "Hedon2" = hedon2, "Hedon3" = hedon3, error_format = "[{statisti
```

iv

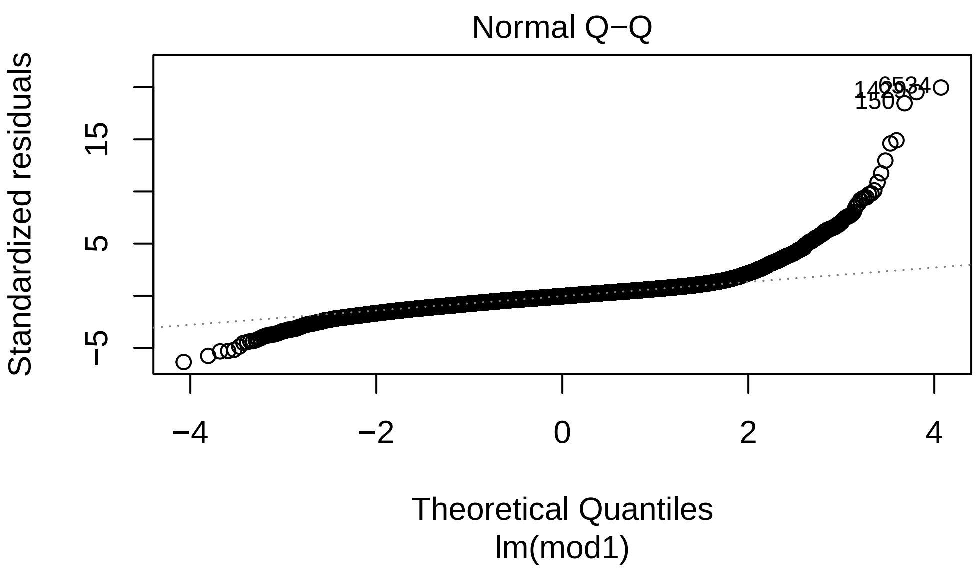
Etter å ha sett dataen i tre ulike hedoniske modeller, kan en se at mod2 ha en bedre marginal forklaringskraft i forhold til mod3. Grunnen til dette er fordi mod3 har mindre variabler og EHD-percen inneholder alle de miljømessige varibler fra mod2. En kan se alle de variablene en trenger i mod2.

Ser på plot(km_navnet)

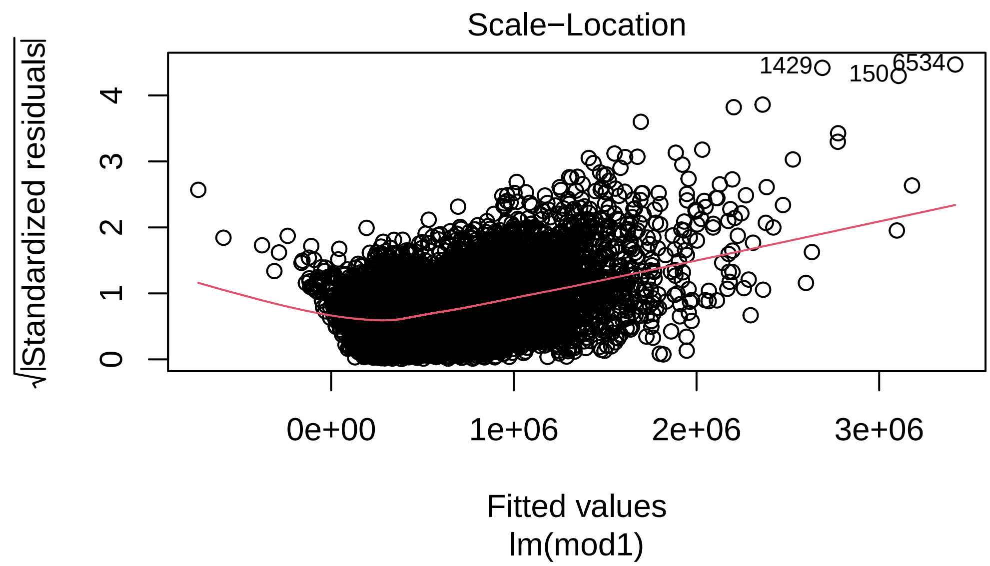
```r
hedon1 %>%
plot()
```

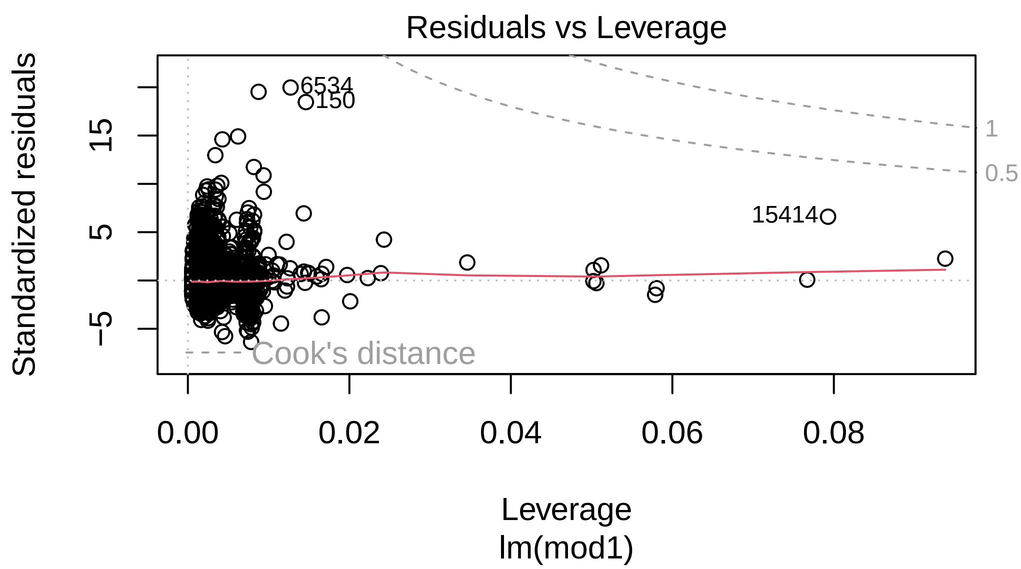


Residuals vs Fitted
1429
150
6534
Residuals
−2e+06
1e+06
4e+06
0e+00
1e+06
2e+06
3e+06
Fitted values
lm(mod1)

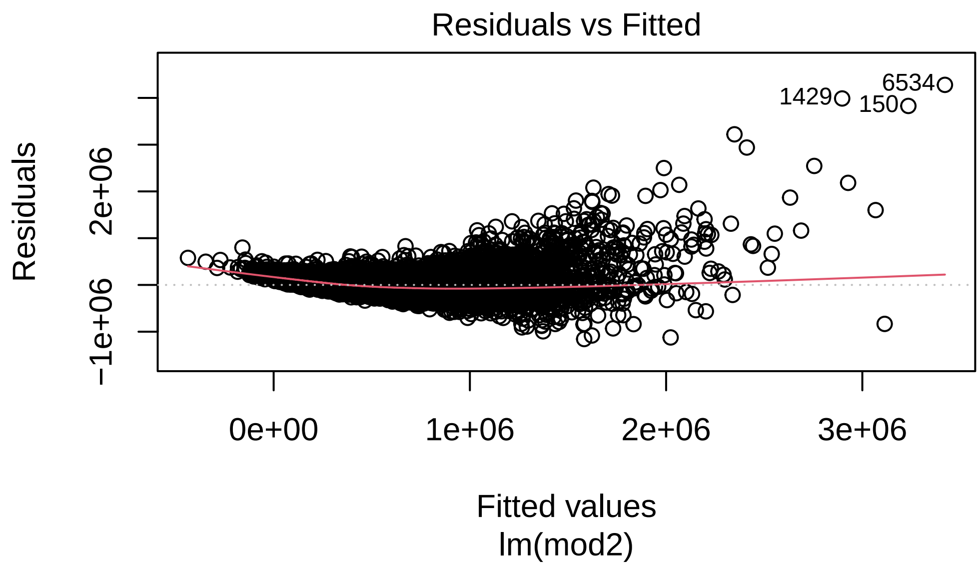


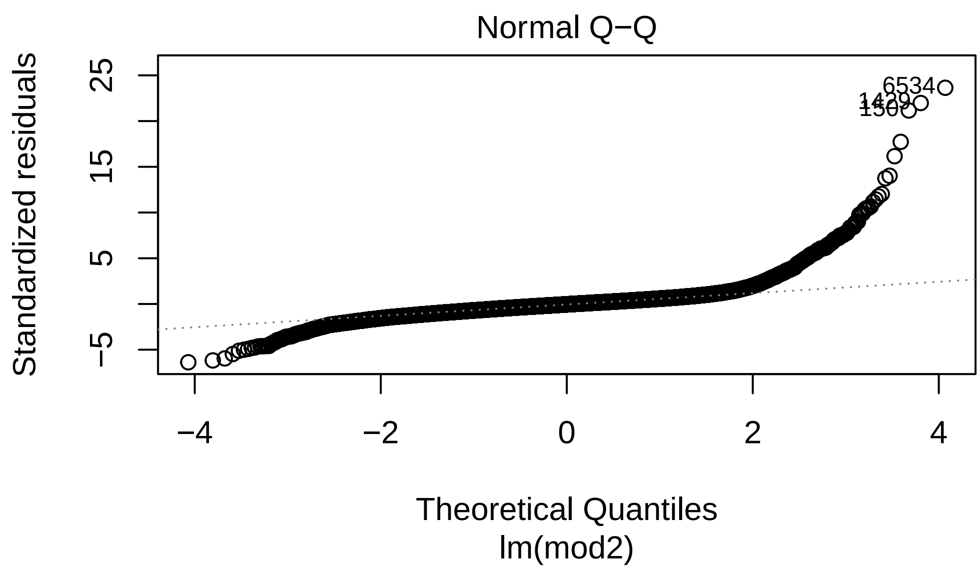
Normal Q−Q
150
Standardized residuals
−5
5
15
−4
−2
0
2
4
Theoretical Quantiles
lm(mod1)

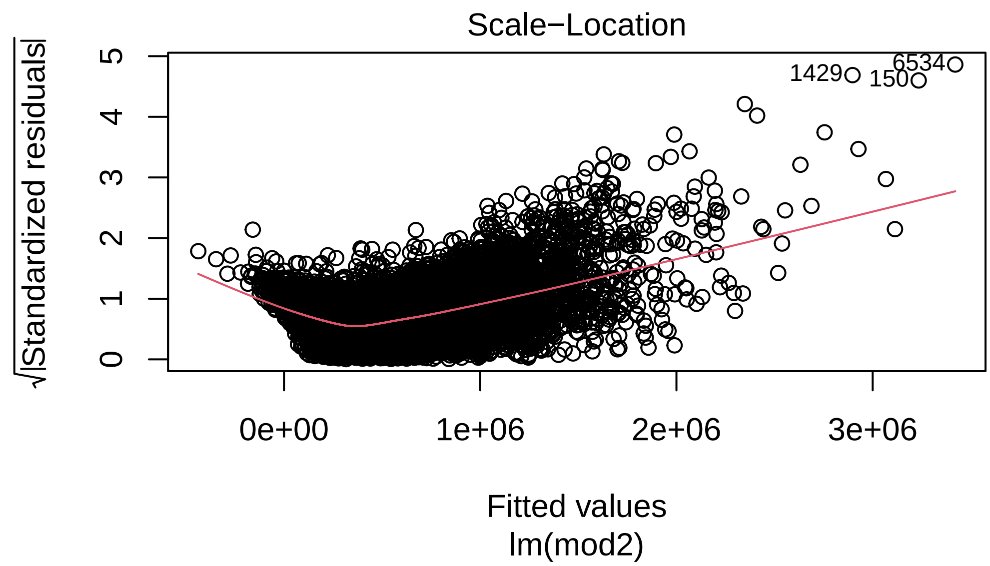

Scale−Location
√|Standardized residuals|
1429
150
6534
0
1
2
3
4
0e+00
1e+06
2e+06
3e+06
Fitted values
lm(mod1)

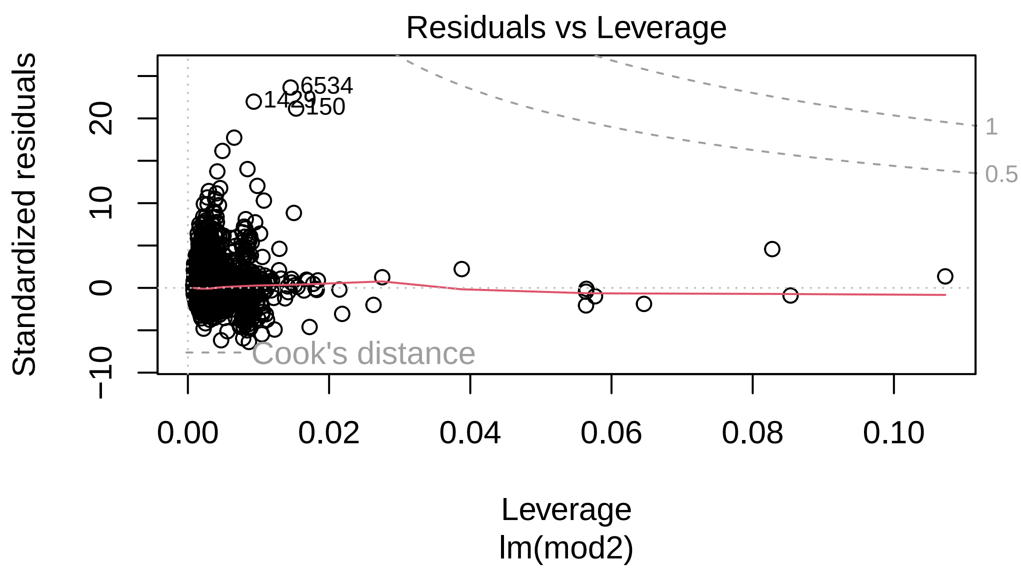

Residuals vs Leverage
Standardized residuals
6534
150
15414
1
0.5
Cook's distance
−5
5
15
0.00
0.02
0.04
0.06
0.08
Leverage
lm(mod1)

```
hedon2 %>%
  plot()
```

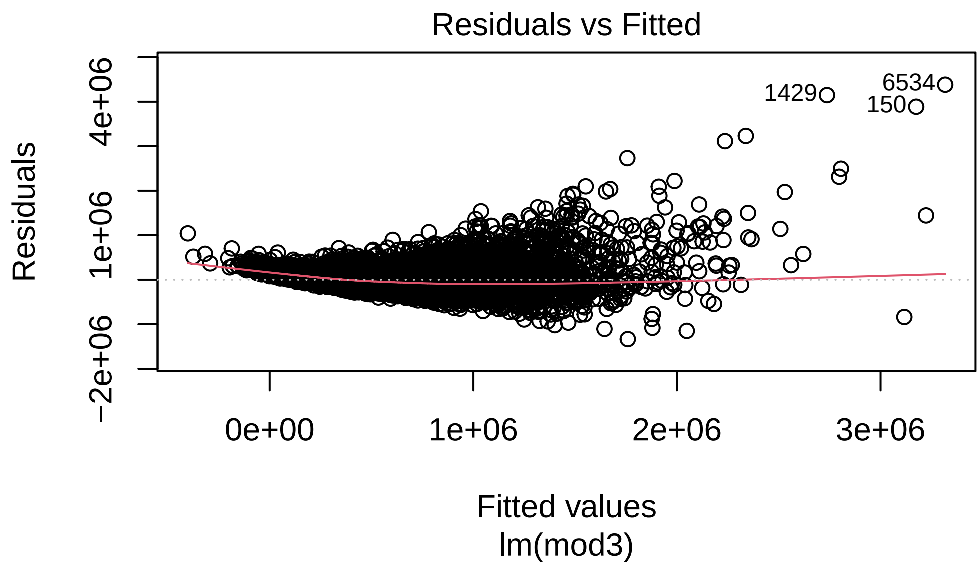

Residuals vs Fitted
Residuals
−1e+06
2e+06
0e+00
1e+06
2e+06
3e+06
Fitted values
lm(mod2)
1429
6534
150

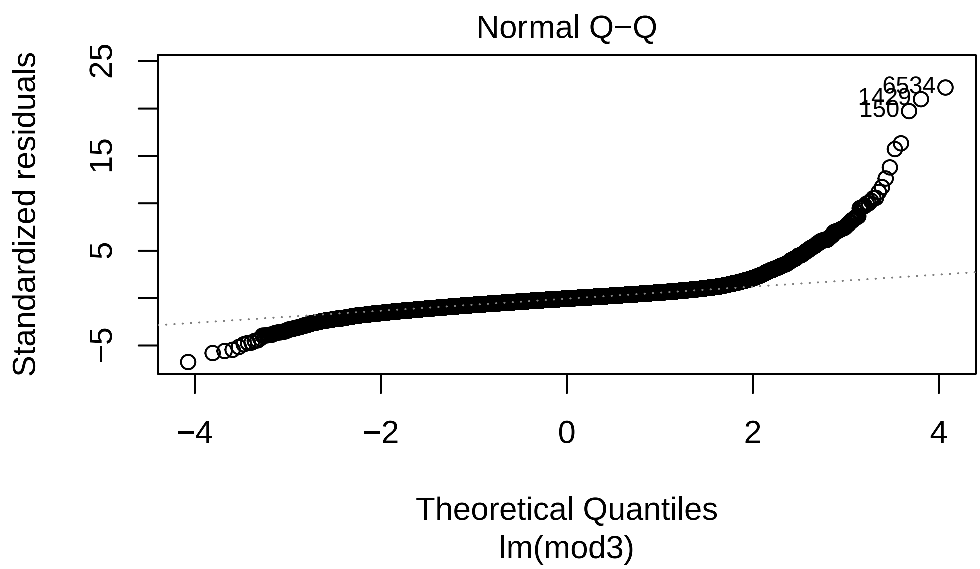

Normal Q−Q
Standardized residuals
−5
5
15
25
−4
−2
0
2
4
Theoretical Quantiles
lm(mod2)
6534
1429
150

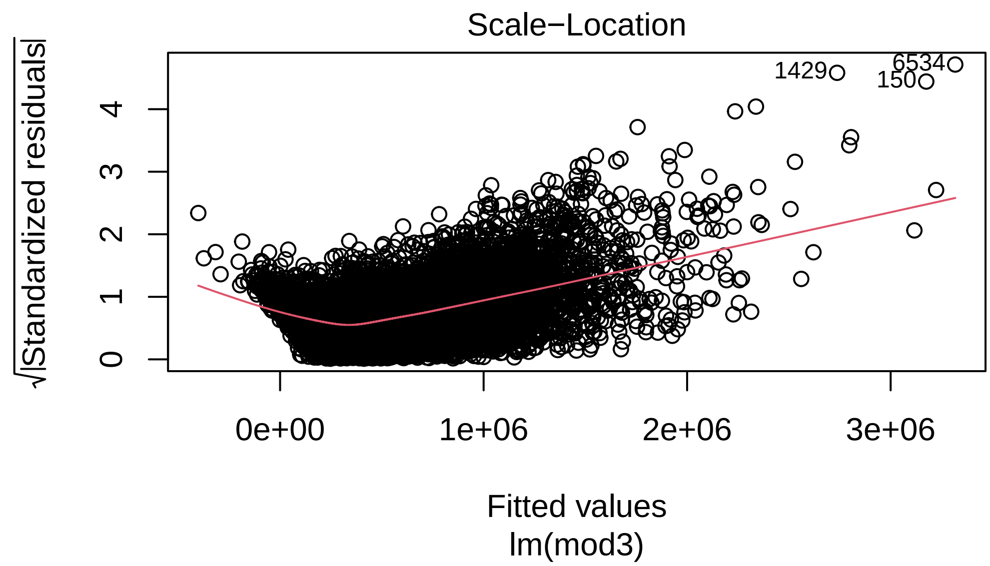

Scale−Location
√|Standardized residuals|
0
1
2
3
4
5
0e+00
1e+06
2e+06
3e+06
1429
150
6534
Fitted values
lm(mod2)

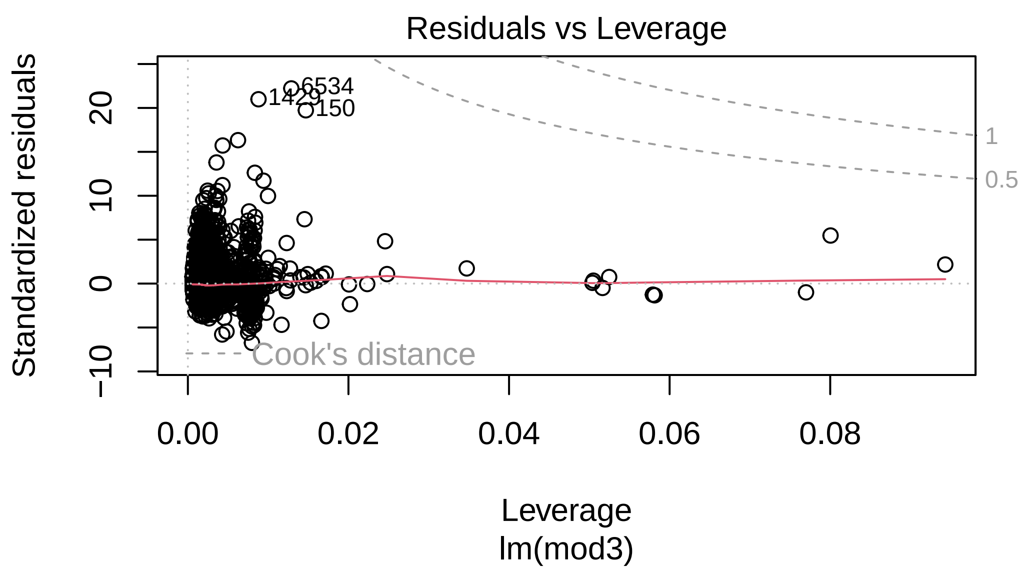

Residuals vs Leverage
Standardized residuals
−10
0
10
20
0.00
0.02
0.04
0.06
0.08
0.10
6534
1429
150
1
0.5
Cook's distance
Leverage
lm(mod2)

```
hedon3 %>%
  plot()
```


Residuals vs Fitted
Residuals
Fitted values
lm(mod3)
1429
6534
150
−2e+06
1e+06
4e+06
0e+00
1e+06
2e+06
3e+06


Normal Q−Q
Standardized residuals
Theoretical Quantiles
lm(mod3)
6534
1429
150
−5
5
15
25
−4
−2
0
2
4





Oppg. 7

Her bruker vi en simultan test på tids-dummiene for å avgjøre om vi trenger disse i modellen.

```
hedon1 %>%
    linearHypothesis(c("year_month2014-06=0", "year_month2014-07=0",
                       "year_month2014-08=0", "year_month2014-09=0",
                       "year_month2014-10=0", "year_month2014-11=0",
```

```r
                        "year_month2014-12=0", "year_month2015-01=0",
                        "year_month2015-02=0", "year_month2015-03=0",
                        "year_month2015-04=0", "year_month2015-05=0"),
                      white_adjust = hc3)
```

```r
hedon2 %>%
    linearHypothesis(c("year_month2014-06=0", "year_month2014-07=0",
                        "year_month2014-08=0", "year_month2014-09=0",
                        "year_month2014-10=0", "year_month2014-11=0",
                        "year_month2014-12=0", "year_month2015-01=0",
                        "year_month2015-02=0", "year_month2015-03=0",
                        "year_month2015-04=0", "year_month2015-05=0"),
                      white_adjust = hc4)
```

```r
hedon3 %>%
    linearHypothesis(c("year_month2014-06=0", "year_month2014-07=0",
                        "year_month2014-08=0", "year_month2014-09=0",
                        "year_month2014-10=0", "year_month2014-11=0",
                        "year_month2014-12=0", "year_month2015-01=0",
                        "year_month2015-02=0", "year_month2015-03=0",
                        "year_month2015-04=0", "year_month2015-05=0"),
                      white_adjust = hc1)
```

Når vi ser på disse testene ser vi på F-verdiene og p-verdiene for å se om vi skal forkaste H0. I følge testene vil vi da forkaste H0 da p-verdien er lavere signifikantnivået. Dette betyr da at vi vil kunne trenge disse tids-dummiene i modellen.

Oppg. 8

I, II, III og VI

Her laster vi inn gruppens utvalg for relevant gpkg fil.

```r
kc_house_data_1111 <- here("Data/kc_house_data_1111_Kine_og_Karoline.gpkg") %>%
  st_read() %>%
  st_transform(2926)
```

```
Reading layer `kc_house_data_1111_Kine_og_Karoline' from data source
  `/Users/kinemakestad/Documents/Master i sivilokonom/3. Semester/Boligmarked og spatial øko
  using driver `GPKG'
```

```
Simple feature collection with 1887 features and 51 fields
Geometry type: POINT
Dimension:     XY
Bounding box:  xmin: 1235740 ymin: 66643 xmax: 1423809 ymax: 287217.5
Projected CRS: NAD83(HARN) / Washington North (ftUS)
```

```
kc_house_data_1111 <- kc_house_data_1111 %>%
  mutate(
    dist_cbd = st_distance(cbd, ., by_element = TRUE),
    dist_cbd_km = set_units(dist_cbd, km),
    year_month = substr(date, start = 1, stop = 7)
  )
```

Gir ett nytt navn til low, mid og high.

```
kc_house_data_1111 <- kc_house_data_1111 %>% rename(low = inc_fam_low_per,
                                                    mid = inc_fam_med_per,
                                                    high = inc_fam_high_per)
```

```
hedon3_seed <- lm(mod3, data = kc_house_data_1111)
```

```
huxreg("Full" = hedon3, "Seed" = hedon3_seed, error_format = "[{statistic}]",
       note = "{stars}. T statistic in brackets.")
```

```
kc_house_data_1111_mat_nb <- knearneigh(kc_house_data_1111, k = 3)
kc_house_data_nb <- knn2nb(kc_house_data_1111_mat_nb)
kc_house_data_1111_W <- nb2listw(kc_house_data_nb, style = "W")
```

```
kc_house_data_1111_mat_nb10 <- knearneigh(kc_house_data_1111, k = 10)
kc_house_data_1111_nb10 <- knn2nb(kc_house_data_1111_mat_nb10)
kc_house_data_1111_W10 <- nb2listw(kc_house_data_1111_nb10)
```

Nå lager vi moran test av det nye datasettet som vi har lastet inn.

```
lm.morantest(hedon3_seed, kc_house_data_1111_W)
```

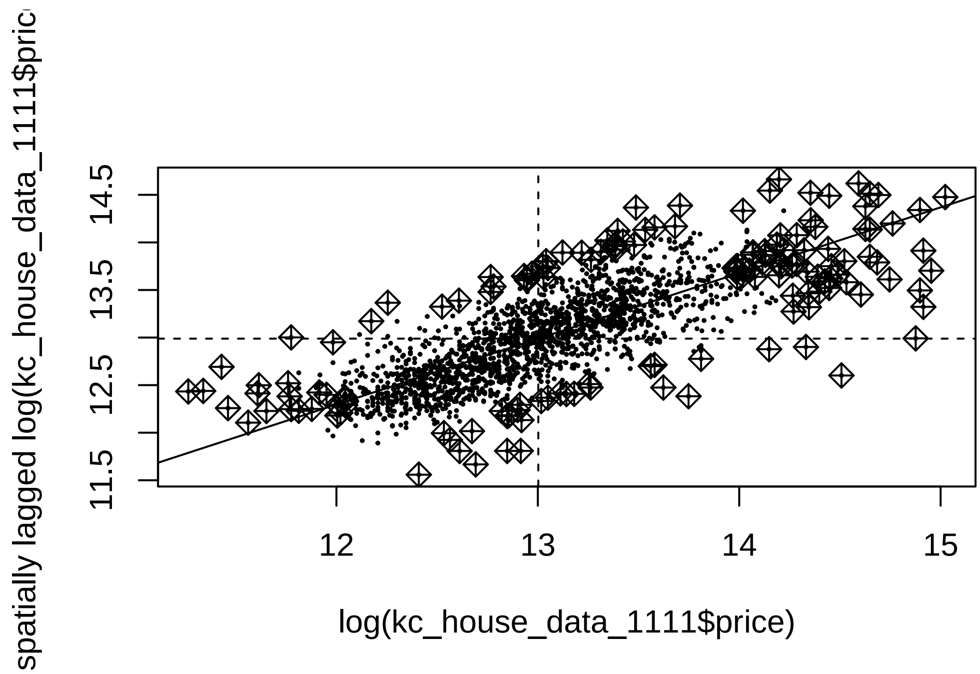
```

    Global Moran I for regression residuals

data:
```

```
model: lm(formula = mod3, data = kc_house_data_1111)
weights: kc_house_data_1111_W

Moran I statistic standard deviate = 20.481, p-value < 2.2e-16
alternative hypothesis: greater
sample estimates:
Observed Moran I      Expectation         Variance
    0.3553875351    -0.0033552144     0.0003067935
```

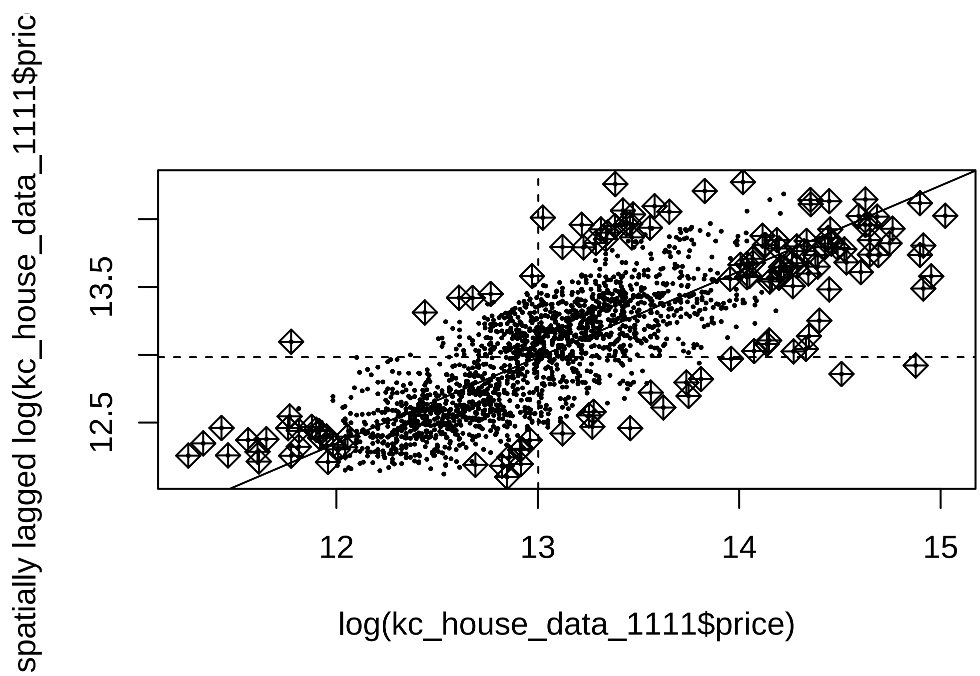
```
lm.morantest(hedon3_seed, kc_house_data_1111_W10)
```

```

    Global Moran I for regression residuals

data:
model: lm(formula = mod3, data = kc_house_data_1111)
weights: kc_house_data_1111_W10

Moran I statistic standard deviate = 29.534, p-value < 2.2e-16
alternative hypothesis: greater
sample estimates:
Observed Moran I      Expectation         Variance
    2.820889e-01    -2.746984e-03     9.301401e-05
```

```
moran.plot(log(kc_house_data_1111$price), listw = kc_house_data_1111_W, labels = FALSE, pc
```




```
moran.plot(log(kc_house_data_1111$price), listw = kc_house_data_1111_W10, labels = FALSE,
```



Deretter kjører vi lagranges multiplikatortest.

```r
kc_lagrange_3 <- lm.LMtests(hedon3_seed, kc_house_data_1111_W, test = "all")
kc_lagrange_3
```

```
    Lagrange multiplier diagnostics for spatial dependence

data:
model: lm(formula = mod3, data = kc_house_data_1111)
weights: kc_house_data_1111_W

LMerr = 408.22, df = 1, p-value < 2.2e-16


    Lagrange multiplier diagnostics for spatial dependence

data:
model: lm(formula = mod3, data = kc_house_data_1111)
weights: kc_house_data_1111_W

LMlag = 425.44, df = 1, p-value < 2.2e-16


    Lagrange multiplier diagnostics for spatial dependence

data:
model: lm(formula = mod3, data = kc_house_data_1111)
weights: kc_house_data_1111_W

RLMerr = 61.34, df = 1, p-value = 4.774e-15


    Lagrange multiplier diagnostics for spatial dependence

data:
model: lm(formula = mod3, data = kc_house_data_1111)
weights: kc_house_data_1111_W

RLMlag = 78.559, df = 1, p-value < 2.2e-16
```

```
	Lagrange multiplier diagnostics for spatial dependence

data:
model: lm(formula = mod3, data = kc_house_data_1111)
weights: kc_house_data_1111_W

SARMA = 486.78, df = 2, p-value < 2.2e-16
```

```r
kc_lagrange_10 <- lm.LMtests(hedon3_seed, kc_house_data_1111_W10, test = "all")
kc_lagrange_10
```

```
	Lagrange multiplier diagnostics for spatial dependence

data:
model: lm(formula = mod3, data = kc_house_data_1111)
weights: kc_house_data_1111_W10

LMerr = 828.59, df = 1, p-value < 2.2e-16


	Lagrange multiplier diagnostics for spatial dependence

data:
model: lm(formula = mod3, data = kc_house_data_1111)
weights: kc_house_data_1111_W10

LMlag = 674.6, df = 1, p-value < 2.2e-16


	Lagrange multiplier diagnostics for spatial dependence

data:
model: lm(formula = mod3, data = kc_house_data_1111)
weights: kc_house_data_1111_W10

RLMerr = 278.07, df = 1, p-value < 2.2e-16


	Lagrange multiplier diagnostics for spatial dependence
```

```
data:
model: lm(formula = mod3, data = kc_house_data_1111)
weights: kc_house_data_1111_W10

RLMlag = 124.07, df = 1, p-value < 2.2e-16


    Lagrange multiplier diagnostics for spatial dependence

data:
model: lm(formula = mod3, data = kc_house_data_1111)
weights: kc_house_data_1111_W10

SARMA = 952.66, df = 2, p-value < 2.2e-16
```

må skrive en liten konklusjon her.

IV

```
SDEM_seed <- errorsarlm(mod3, data = kc_house_data_1111, listw = kc_house_data_1111_W, Dur
```

```
Warning in errorsarlm(mod3, data = kc_house_data_1111, listw = kc_house_data_1111_W, : invers
  reciprocal condition number = 6.22569e-22 - using numerical Hessian.
```

```
SLX_seed <- lmSLX(mod3, data = kc_house_data_1111, listw = kc_house_data_1111_W, Durbin =
```

```
SEM_seed <- errorsarlm(mod3, data = kc_house_data_1111, listw = kc_house_data_1111_W, Durb
```

```
Warning in errorsarlm(mod3, data = kc_house_data_1111, listw = kc_house_data_1111_W, : invers
  reciprocal condition number = 4.26126e-22 - using numerical Hessian.
```

```
summary(impacts(SDEM_seed), zstats = TRUE)
```

```
Impact measures (SDEM, estimable, n):
                      Direct      Indirect         Total
bedrooms         -1.671756e+04 -1.771717e+04 -34434.733569
bathrooms         1.614290e+04  1.650382e+04  32646.722239
```

```
sqft_living        1.026975e+02  6.552541e+01    168.222885
sqft_living15      1.523337e+01  1.865162e+01     33.884991
sqft_lot           7.472955e-01  2.962511e-01      1.043547
sqft_lot15        -5.776350e-01 -1.400789e+00     -1.978424
sqft_above         6.489117e+01 -6.947738e+01     -4.586206
floors            -3.146776e+04  4.024927e+04   8781.509949
grade              8.010907e+04  4.748276e+04 127591.835172
yr_built          -1.057719e+03 -3.195723e+03  -4253.441823
yr_renovated       2.610505e+01  9.551686e+00     35.656740
waterfront         5.148195e+05 -2.400173e+05 274802.211278
condition          2.103244e+04 -1.350654e+04   7525.901083
view               4.253524e+04 -4.486853e+03  38048.390841
dist_cbd_km        1.096129e+04 -7.316344e+03   3644.942697
EHD_percen         8.387379e+02 -6.264486e+02    212.289295
low                2.600046e+04  3.352328e+05 361233.255808
high               1.102939e+05  3.200718e+05 430365.698788
year_month2014-06  1.203060e+03            NA   1203.059797
year_month2014-07  1.472893e+04            NA  14728.932484
year_month2014-08  2.763425e+04            NA  27634.248468
year_month2014-09  1.267803e+04            NA  12678.032919
year_month2014-10  9.369761e+03            NA   9369.760983
year_month2014-11  1.665905e+04            NA  16659.051265
year_month2014-12  2.579090e+03            NA   2579.090474
year_month2015-01  8.481748e+03            NA   8481.748032
year_month2015-02  2.985277e+04            NA  29852.774536
year_month2015-03  4.572246e+04            NA  45722.459142
year_month2015-04  7.389430e+04            NA  73894.295360
year_month2015-05  6.744666e+04            NA  67446.661508
========================================================
Standard errors:
                        Direct     Indirect        Total
bedrooms          5.474750e+03 1.088559e+04 1.370790e+04
bathrooms         9.011617e+03 1.824577e+04 2.330582e+04
sqft_living       1.239742e+01 2.452343e+01 3.195772e+01
sqft_living15     1.048318e+01 1.952842e+01 2.460402e+01
sqft_lot          1.793557e-01 3.952306e-01 4.891114e-01
sqft_lot15        3.086283e-01 6.251295e-01 7.570081e-01
sqft_above        1.245537e+01 2.480912e+01 3.205054e+01
floors            1.023776e+04 1.898160e+04 2.375632e+04
grade             6.072951e+03 1.198562e+04 1.502963e+04
yr_built          2.045448e+02 3.520614e+02 4.309271e+02
yr_renovated      1.000288e+01 2.041216e+01 2.685090e+01
waterfront        4.880346e+04 1.162636e+05 1.458740e+05
```

```
condition          6.508138e+03 1.319053e+04 1.686856e+04
view               6.544442e+03 1.296556e+04 1.639702e+04
dist_cbd_km        9.253968e+03 9.299094e+03 8.406290e+02
EHD_percen         4.672541e+02 5.956077e+02 4.913446e+02
low                1.073598e+05 1.470153e+05 1.372339e+05
high               7.769131e+04 1.045230e+05 9.415563e+04
year_month2014-06 1.662160e+04           NA 1.662160e+04
year_month2014-07 1.634411e+04           NA 1.634411e+04
year_month2014-08 1.667201e+04           NA 1.667201e+04
year_month2014-09 1.736447e+04           NA 1.736447e+04
year_month2014-10 1.711358e+04           NA 1.711358e+04
year_month2014-11 1.846831e+04           NA 1.846831e+04
year_month2014-12 1.815839e+04           NA 1.815839e+04
year_month2015-01 2.006516e+04           NA 2.006516e+04
year_month2015-02 1.892173e+04           NA 1.892173e+04
year_month2015-03 1.655718e+04           NA 1.655718e+04
year_month2015-04 1.591300e+04           NA 1.591300e+04
year_month2015-05 2.388122e+04           NA 2.388122e+04
========================================================
Z-values:
                       Direct   Indirect       Total
bedrooms          -3.05357556 -1.6275797 -2.51203509
bathrooms          1.79134380  0.9045284  1.40079682
sqft_living        8.28377544  2.6719506  5.26391931
sqft_living15      1.45312488  0.9551012  1.37721341
sqft_lot           4.16655543  0.7495651  2.13355601
sqft_lot15        -1.87162026 -2.2407978 -2.61347750
sqft_above         5.20989615 -2.8004771 -0.14309296
floors            -3.07369601  2.1204356  0.36964943
grade             13.19112819  3.9616441  8.48935416
yr_built          -5.17108558 -9.0771769 -9.87044364
yr_renovated       2.60975342  0.4679411  1.32795347
waterfront        10.54883176 -2.0644232  1.88383276
condition          3.23171347 -1.0239574  0.44614961
view               6.49944524 -0.3460592  2.32044581
dist_cbd_km        1.18449587 -0.7867804  4.33597065
EHD_percen         1.79503588 -1.0517804  0.43205786
low                0.24218048  2.2802582  2.63224532
high               1.41964290  3.0622132  4.57079093
year_month2014-06  0.07237931         NA  0.07237931
year_month2014-07  0.90117694         NA  0.90117694
year_month2014-08  1.65752384         NA  1.65752384
year_month2014-09  0.73011351         NA  0.73011351
```

```
year_month2014-10  0.54750450         NA  0.54750450
year_month2014-11  0.90203422         NA  0.90203422
year_month2014-12  0.14203301         NA  0.14203301
year_month2015-01  0.42271024         NA  0.42271024
year_month2015-02  1.57769767         NA  1.57769767
year_month2015-03  2.76148845         NA  2.76148845
year_month2015-04  4.64364420         NA  4.64364420
year_month2015-05  2.82425514         NA  2.82425514

p-values:
                  Direct     Indirect   Total
bedrooms          0.0022613  0.1036140  0.0120037
bathrooms         0.0732381  0.3657153  0.1612748
sqft_living       2.2204e-16 0.0075412  1.4102e-07
sqft_living15     0.1461891  0.3395265  0.1684463
sqft_lot          3.0924e-05 0.4535167  0.0328791
sqft_lot15        0.0612592  0.0250392  0.0089626
sqft_above        1.8895e-07 0.0051027  0.8862168
floors            0.0021142  0.0339693  0.7116437
grade             < 2.22e-16 7.4435e-05 < 2.22e-16
yr_built          2.3274e-07 < 2.22e-16 < 2.22e-16
yr_renovated      0.0090608  0.6398267  0.1841935
waterfront        < 2.22e-16 0.0389776  0.0595876
condition         0.0012305  0.3058554  0.6554892
view              8.0617e-11 0.7292982  0.0203168
dist_cbd_km       0.2362168  0.4314104  1.4512e-05
EHD_percen        0.0726480  0.2929003  0.6656994
low               0.8086403  0.0225924  0.0084823
high              0.1557117  0.0021971  4.8589e-06
year_month2014-06 0.9423001  NA         0.9423001
year_month2014-07 0.3674943  NA         0.3674943
year_month2014-08 0.0974136  NA         0.0974136
year_month2014-09 0.4653208  NA         0.4653208
year_month2014-10 0.5840322  NA         0.5840322
year_month2014-11 0.3670387  NA         0.3670387
year_month2014-12 0.8870539  NA         0.8870539
year_month2015-01 0.6725067  NA         0.6725067
year_month2015-02 0.1146351  NA         0.1146351
year_month2015-03 0.0057539  NA         0.0057539
year_month2015-04 3.4232e-06 NA         3.4232e-06
year_month2015-05 0.0047391  NA         0.0047391
```

```r
huxreg("SEM" = SEM_seed, "OLS" = hedon3_seed, error_format = "[{statistic}]", note = "{sta
```

```r
LR.Sarlm(SDEM_seed, SEM_seed)
```

```

    Likelihood ratio for spatial linear models

data:
Likelihood ratio = 173.68, df = 18, p-value < 2.2e-16
sample estimates:
Log likelihood of SDEM_seed  Log likelihood of SEM_seed
                  -25291.20                   -25378.04
```

```r
LR.Sarlm(SDEM_seed, SLX_seed)
```

```

    Likelihood ratio for spatial linear models

data:
Likelihood ratio = 382.62, df = 1, p-value < 2.2e-16
sample estimates:
Log likelihood of SDEM_seed  Log likelihood of SLX_seed
                  -25291.20                   -25482.51
```

Den beste modellen som vi ser utifra estimatene er *SDEM*. Deretter vil vi ta en kontrolltest mot *OLS-modellen*.

```r
LR1.Sarlm(SDEM_seed)
```

```

    Likelihood Ratio diagnostics for spatial dependence

data:
Likelihood ratio = 382.62, df = 1, p-value < 2.2e-16
sample estimates:
Log likelihood of spatial error model           Log likelihood of OLS fit y
                            -25291.20                             -25482.51
```

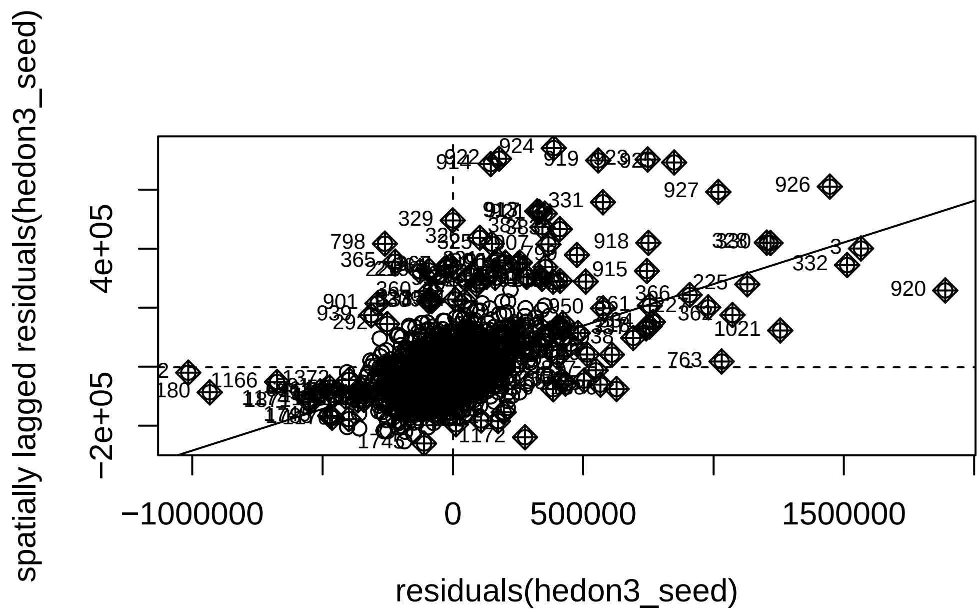
```r
Hausman.test(SEM_seed)
```

```
    Spatial Hausman test (asymptotic)

data:  NULL
Hausman test = 216.3, df = 31, p-value < 2.2e-16
```

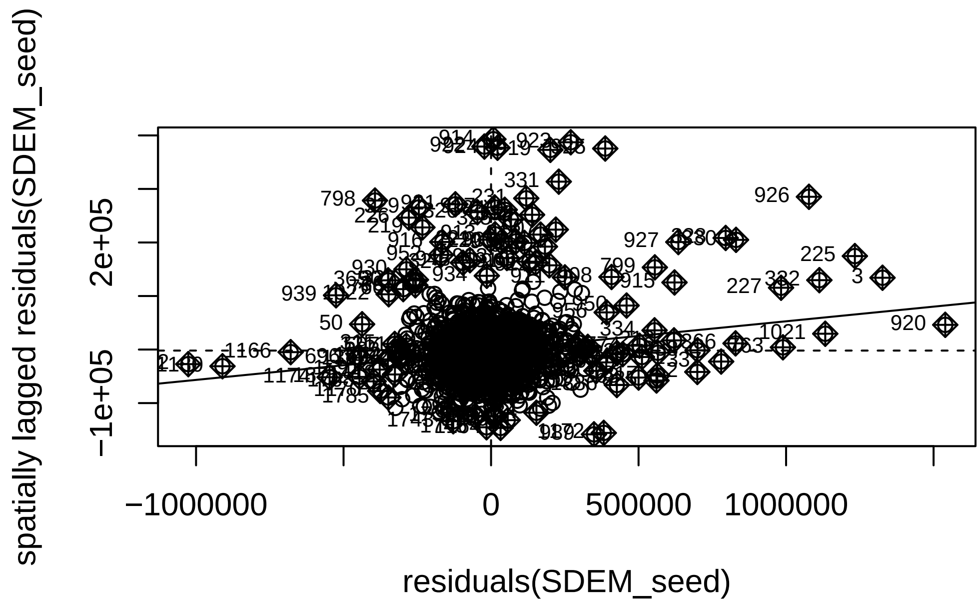
```r
bptest.Sarlm(SEM_seed,studentize = TRUE)
```

```
    studentized Breusch-Pagan test

data:
BP = 327.4, df = 30, p-value < 2.2e-16
```

```r
moran.plot(residuals(hedon3_seed), listw = kc_house_data_1111_W10)
```




```r
moran.plot(residuals(SDEM_seed), listw = kc_house_data_1111_W10)
```




```
moran.test(residuals(SDEM_seed), listw = kc_house_data_1111_W10)
```

```

	Moran I test under randomisation

data:  residuals(SDEM_seed)
weights: kc_house_data_1111_W10

Moran I statistic standard deviate = 5.687, p-value = 6.463e-09
alternative hypothesis: greater
sample estimates:
Moran I statistic       Expectation          Variance
     5.474007e-02     -5.302227e-04      9.445206e-05
```

Oppg. 9

```
set.seed(572)
kc_house_env_var_omit_2000 <- kc_house_env_var_omit[sample(1:nrow(kc_house_env_var_omit),
```

```
hedon3_2000 <- lm(mod3, data = kc_house_env_var_omit_2000)
```

```r
huxreg("Full" = hedon3, "2000 Seed" = hedon3_2000, "1111 Seed" = hedon3_seed,
    error_format = "[{statistic}]",
    note = "{stars}. T statistic in brackets.")
```

```r
kc_house_data_2000_mat_nb <- knearneigh(kc_house_env_var_omit_2000, k = 3)
```

```
Warning in knearneigh(kc_house_env_var_omit_2000, k = 3): knearneigh: identical
points found
```

```
Warning in knearneigh(kc_house_env_var_omit_2000, k = 3): knearneigh: kd_tree
not available for identical points
```

```r
kc_house_data_2000_nb <- knn2nb(kc_house_data_2000_mat_nb)
kc_house_data_2000_W <- nb2listw(kc_house_data_2000_nb, style = "W")
```

```r
kc_house_data_2000_mat_nb10 <- knearneigh(kc_house_env_var_omit_2000, k = 10)
```

```
Warning in knearneigh(kc_house_env_var_omit_2000, k = 10): knearneigh: identical
points found
```

```
Warning in knearneigh(kc_house_env_var_omit_2000, k = 10): knearneigh: kd_tree
not available for identical points
```

```r
kc_house_data_2000_nb10 <- knn2nb(kc_house_data_2000_mat_nb10)
kc_house_data_2000_W10 <- nb2listw(kc_house_data_2000_nb10, style = "W")
```

```r
lm.morantest(hedon3_2000, kc_house_data_2000_W)
```

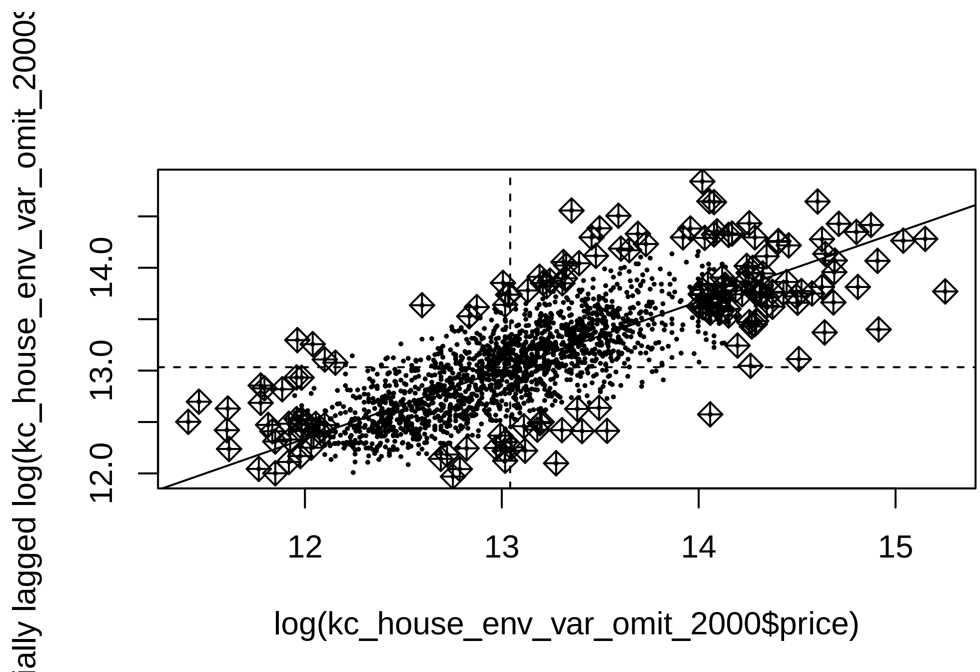
```

    Global Moran I for regression residuals

data:
model: lm(formula = mod3, data = kc_house_env_var_omit_2000)
weights: kc_house_data_2000_W

Moran I statistic standard deviate = 23.638, p-value < 2.2e-16
```

```
alternative hypothesis: greater
sample estimates:
Observed Moran I      Expectation         Variance
    0.3960742959    -0.0032840908     0.0002854216
```

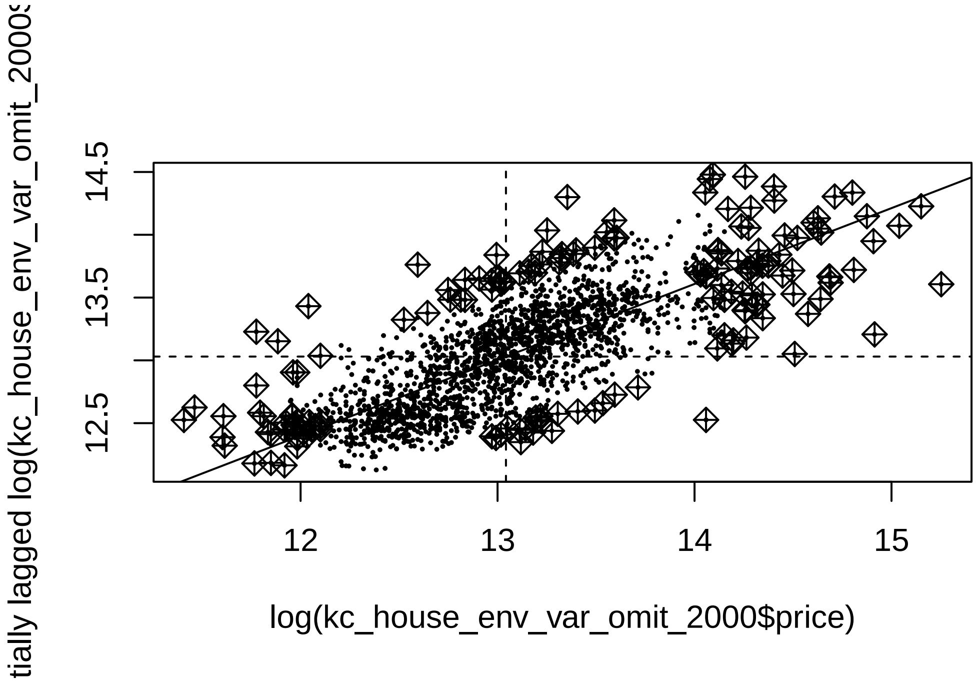
```
lm.morantest(hedon3_2000, kc_house_data_2000_W10)
```

```

    Global Moran I for regression residuals

data:
model: lm(formula = mod3, data = kc_house_env_var_omit_2000)
weights: kc_house_data_2000_W10

Moran I statistic standard deviate = 35.681, p-value < 2.2e-16
alternative hypothesis: greater
sample estimates:
Observed Moran I      Expectation         Variance
    3.295827e-01    -2.735180e-03     8.674189e-05
```

```
moran.plot(log(kc_house_env_var_omit_2000$price), listw= kc_house_data_2000_W, labels = FA
```



```r
moran.plot(log(kc_house_env_var_omit_2000$price), listw = kc_house_data_2000_W10, labels =
```




```r
kc_lagrange_3_2000 <- lm.LMtests(hedon3_2000, kc_house_data_2000_W, test = "all")
kc_lagrange_3_2000
```

```
    Lagrange multiplier diagnostics for spatial dependence

data:
model: lm(formula = mod3, data = kc_house_env_var_omit_2000)
weights: kc_house_data_2000_W

LMerr = 544.91, df = 1, p-value < 2.2e-16


    Lagrange multiplier diagnostics for spatial dependence

data:
model: lm(formula = mod3, data = kc_house_env_var_omit_2000)
weights: kc_house_data_2000_W
```

```
LMlag = 520.85, df = 1, p-value < 2.2e-16


    Lagrange multiplier diagnostics for spatial dependence

data:
model: lm(formula = mod3, data = kc_house_env_var_omit_2000)
weights: kc_house_data_2000_W

RLMerr = 106.1, df = 1, p-value < 2.2e-16


    Lagrange multiplier diagnostics for spatial dependence

data:
model: lm(formula = mod3, data = kc_house_env_var_omit_2000)
weights: kc_house_data_2000_W

RLMlag = 82.033, df = 1, p-value < 2.2e-16


    Lagrange multiplier diagnostics for spatial dependence

data:
model: lm(formula = mod3, data = kc_house_env_var_omit_2000)
weights: kc_house_data_2000_W

SARMA = 626.95, df = 2, p-value < 2.2e-16
```

```
kc_lagrange_10_2000 <- lm.LMtests(hedon3_2000, kc_house_data_2000_W10, test = "all")
kc_lagrange_10_2000
```

```
    Lagrange multiplier diagnostics for spatial dependence

data:
model: lm(formula = mod3, data = kc_house_env_var_omit_2000)
weights: kc_house_data_2000_W10

LMerr = 1211.9, df = 1, p-value < 2.2e-16
```

```
	Lagrange multiplier diagnostics for spatial dependence

data:
model: lm(formula = mod3, data = kc_house_env_var_omit_2000)
weights: kc_house_data_2000_W10

LMlag = 900.33, df = 1, p-value < 2.2e-16


	Lagrange multiplier diagnostics for spatial dependence

data:
model: lm(formula = mod3, data = kc_house_env_var_omit_2000)
weights: kc_house_data_2000_W10

RLMerr = 449.9, df = 1, p-value < 2.2e-16


	Lagrange multiplier diagnostics for spatial dependence

data:
model: lm(formula = mod3, data = kc_house_env_var_omit_2000)
weights: kc_house_data_2000_W10

RLMlag = 138.31, df = 1, p-value < 2.2e-16


	Lagrange multiplier diagnostics for spatial dependence

data:
model: lm(formula = mod3, data = kc_house_env_var_omit_2000)
weights: kc_house_data_2000_W10

SARMA = 1350.2, df = 2, p-value < 2.2e-16
```

```r
SDEM_2000 <- errorsarlm(mod3, data = kc_house_env_var_omit_2000, listw = kc_house_data_200
```

```
Warning in errorsarlm(mod3, data = kc_house_env_var_omit_2000, listw = kc_house_data_2000_W,
  reciprocal condition number = 5.62644e-22 - using numerical Hessian.
```

```r
SLX_2000 <- lmSLX(mod3, data = kc_house_env_var_omit_2000, listw = kc_house_data_2000_W, D
```

```r
SEM_2000 <- errorsarlm(mod3, data = kc_house_env_var_omit_2000, listw = kc_house_data_2000
```

```
Warning in errorsarlm(mod3, data = kc_house_env_var_omit_2000, listw = kc_house_data_2000_W,
  reciprocal condition number = 4.16073e-22 - using numerical Hessian.
```

```r
summary(impacts(SDEM_2000), zstats = TRUE)
```

```
Impact measures (SDEM, estimable, n):
                        Direct      Indirect         Total
bedrooms          -1.433858e+04 -1.453000e+03 -1.579158e+04
bathrooms          4.508600e+04  4.598831e+04  9.107431e+04
sqft_living        1.232876e+02  4.659072e+01  1.698784e+02
sqft_living15      1.550166e+00 -2.717099e+01 -2.562082e+01
sqft_lot          -1.100469e-01 -8.800375e-01 -9.900844e-01
sqft_lot15         1.352987e-01 -4.689139e-01 -3.336151e-01
sqft_above         4.858916e+01 -4.974018e+01 -1.151018e+00
floors            -2.791414e+04  9.576259e+03 -1.833788e+04
grade              7.428547e+04  4.701714e+04  1.213026e+05
yr_built          -1.510582e+03 -2.709645e+03 -4.220226e+03
yr_renovated       3.031278e+01  8.982580e+00  3.929536e+01
waterfront         4.032910e+05 -1.191041e+05  2.841869e+05
condition          2.617339e+04 -1.937950e+04  6.793888e+03
view               3.291681e+04 -2.873774e+02  3.262943e+04
dist_cbd_km       -3.523886e+02  3.938489e+03  3.586101e+03
EHD_percen        -3.045563e+01  5.373905e+02  5.069348e+02
low                3.123187e+05  9.240240e+04  4.047211e+05
high               2.807502e+05  2.438467e+05  5.245968e+05
year_month2014-06 -1.066114e+04            NA -1.066114e+04
year_month2014-07  1.005425e+04            NA  1.005425e+04
year_month2014-08 -1.143473e+04            NA -1.143473e+04
year_month2014-09 -3.257754e+04            NA -3.257754e+04
year_month2014-10 -1.502985e+04            NA -1.502985e+04
year_month2014-11 -1.660401e+04            NA -1.660401e+04
year_month2014-12 -2.673882e+04            NA -2.673882e+04
year_month2015-01 -3.716615e+04            NA -3.716615e+04
year_month2015-02 -2.049110e+04            NA -2.049110e+04
year_month2015-03  8.666831e+03            NA  8.666831e+03
year_month2015-04  1.438079e+04            NA  1.438079e+04
```

```
year_month2015-05  5.025999e+04            NA  5.025999e+04
=========================================================
Standard errors:
                        Direct     Indirect        Total
bedrooms          5.471072e+03 1.125229e+04 1.437770e+04
bathrooms         9.665067e+03 2.008819e+04 2.627238e+04
sqft_living       1.242004e+01 2.632203e+01 3.345099e+01
sqft_living15     9.968137e+00 1.866928e+01 2.361428e+01
sqft_lot          1.996143e-01 4.750435e-01 5.804732e-01
sqft_lot15        2.624552e-01 6.130887e-01 7.383382e-01
sqft_above        1.247512e+01 2.479700e+01 3.110061e+01
floors            1.078560e+04 2.051555e+04 2.571018e+04
grade             6.203348e+03 1.234202e+04 1.567325e+04
yr_built          2.101651e+02 3.730801e+02 4.567433e+02
yr_renovated      1.054753e+01 2.261678e+01 2.915110e+01
waterfront        5.070678e+04 1.027743e+05 1.271792e+05
condition         6.893713e+03 1.353349e+04 1.742068e+04
view              6.036025e+03 1.208407e+04 1.480239e+04
dist_cbd_km       4.098673e+03 4.181705e+03 8.444872e+02
EHD_percen        5.153117e+02 6.439454e+02 5.126662e+02
low               1.118786e+05 1.582119e+05 1.521279e+05
high              7.623650e+04 1.084715e+05 1.001498e+05
year_month2014-06 1.581512e+04           NA 1.581512e+04
year_month2014-07 1.557882e+04           NA 1.557882e+04
year_month2014-08 1.630042e+04           NA 1.630042e+04
year_month2014-09 1.686513e+04           NA 1.686513e+04
year_month2014-10 1.632059e+04           NA 1.632059e+04
year_month2014-11 1.766892e+04           NA 1.766892e+04
year_month2014-12 1.681080e+04           NA 1.681080e+04
year_month2015-01 2.012533e+04           NA 2.012533e+04
year_month2015-02 1.832330e+04           NA 1.832330e+04
year_month2015-03 1.635325e+04           NA 1.635325e+04
year_month2015-04 1.546198e+04           NA 1.546198e+04
year_month2015-05 2.349382e+04           NA 2.349382e+04
=========================================================
Z-values:
                       Direct   Indirect       Total
bedrooms          -2.62079844 -0.1291293 -1.09833790
bathrooms          4.66484085  2.2893203  3.46654176
sqft_living        9.92650738  1.7700272  5.07842483
sqft_living15      0.15551215 -1.4553852 -1.08497172
sqft_lot          -0.55129798 -1.8525408 -1.70565065
sqft_lot15         0.51551169 -0.7648385 -0.45184592
```

```
sqft_above          3.89488456 -2.0058947 -0.03700949
floors             -2.58809427  0.4667805 -0.71325377
grade              11.97506211  3.8095172  7.73946554
yr_built           -7.18759683 -7.2629033 -9.23982207
yr_renovated        2.87392309  0.3971644  1.34798910
waterfront          7.95339360 -1.1588893  2.23453997
condition           3.79670378 -1.4319664  0.38998980
view                5.45339199 -0.0237815  2.20433558
dist_cbd_km        -0.08597626  0.9418381  4.24648330
EHD_percen         -0.05910138  0.8345280  0.98882048
low                 2.79158683  0.5840418  2.66039960
high                3.68262147  2.2480260  5.23812049
year_month2014-06 -0.67411074         NA -0.67411074
year_month2014-07  0.64537944         NA  0.64537944
year_month2014-08 -0.70149906         NA -0.70149906
year_month2014-09 -1.93165127         NA -1.93165127
year_month2014-10 -0.92091333         NA -0.92091333
year_month2014-11 -0.93973000         NA -0.93973000
year_month2014-12 -1.59057416         NA -1.59057416
year_month2015-01 -1.84673537         NA -1.84673537
year_month2015-02 -1.11830833         NA -1.11830833
year_month2015-03  0.52997605         NA  0.52997605
year_month2015-04  0.93007417         NA  0.93007417
year_month2015-05  2.13928576         NA  2.13928576

p-values:
                  Direct     Indirect   Total
bedrooms          0.00877241 0.89725537 0.2720570
bathrooms         3.0886e-06 0.02206075 0.0005272
sqft_living       < 2.22e-16 0.07672261 3.8058e-07
sqft_living15     0.87641757 0.14556264 0.2779342
sqft_lot          0.58142942 0.06394820 0.0880731
sqft_lot15        0.60619550 0.44436772 0.6513800
sqft_above        9.8246e-05 0.04486748 0.9704774
floors            0.00965086 0.64065693 0.4756887
grade             < 2.22e-16 0.00013924 9.9920e-15
yr_built          6.5947e-13 3.7881e-13 < 2.22e-16
yr_renovated      0.00405408 0.69124620 0.1776619
waterfront        1.7764e-15 0.24650132 0.0254476
condition         0.00014663 0.15215343 0.6965441
view              4.9418e-08 0.98102689 0.0275008
dist_cbd_km       0.93148529 0.34627552 2.1715e-05
EHD_percen        0.95287136 0.40398350 0.3227510
```

```
low               0.00524503 0.55919216 0.0078048
high              0.00023085 0.02457453 1.6222e-07
year_month2014-06 0.50024091 NA         0.5002409
year_month2014-07 0.51868132 NA         0.5186813
year_month2014-08 0.48299162 NA         0.4829916
year_month2014-09 0.05340256 NA         0.0534026
year_month2014-10 0.35709568 NA         0.3570957
year_month2014-11 0.34735607 NA         0.3473561
year_month2014-12 0.11170544 NA         0.1117054
year_month2015-01 0.06478550 NA         0.0647855
year_month2015-02 0.26343533 NA         0.2634353
year_month2015-03 0.59612854 NA         0.5961285
year_month2015-04 0.35233268 NA         0.3523327
year_month2015-05 0.03241253 NA         0.0324125
```

```r
huxreg("SEM" = SEM_2000, "OLS" = hedon3_2000, error_format = "[{statistic}]", note = "{sta
```

```r
LR.Sarlm(SDEM_2000, SEM_2000)
```

```

    Likelihood ratio for spatial linear models

data:
Likelihood ratio = 125.52, df = 18, p-value < 2.2e-16
sample estimates:
Log likelihood of SDEM_2000  Log likelihood of SEM_2000
                  -26843.71                   -26906.47
```

```r
LR.Sarlm(SDEM_2000, SLX_2000)
```

```

    Likelihood ratio for spatial linear models

data:
Likelihood ratio = 451.32, df = 1, p-value < 2.2e-16
sample estimates:
Log likelihood of SDEM_2000  Log likelihood of SLX_2000
                  -26843.71                   -27069.36
```

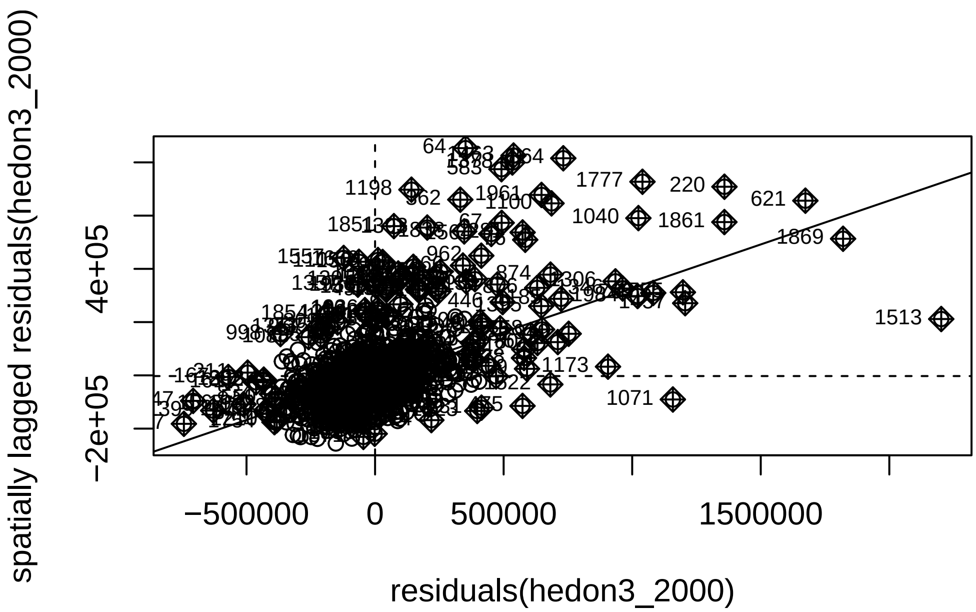
```r
Hausman.test(SEM_2000)
```

```
    Spatial Hausman test (asymptotic)

data:  NULL
Hausman test = 164.87, df = 31, p-value < 2.2e-16
```

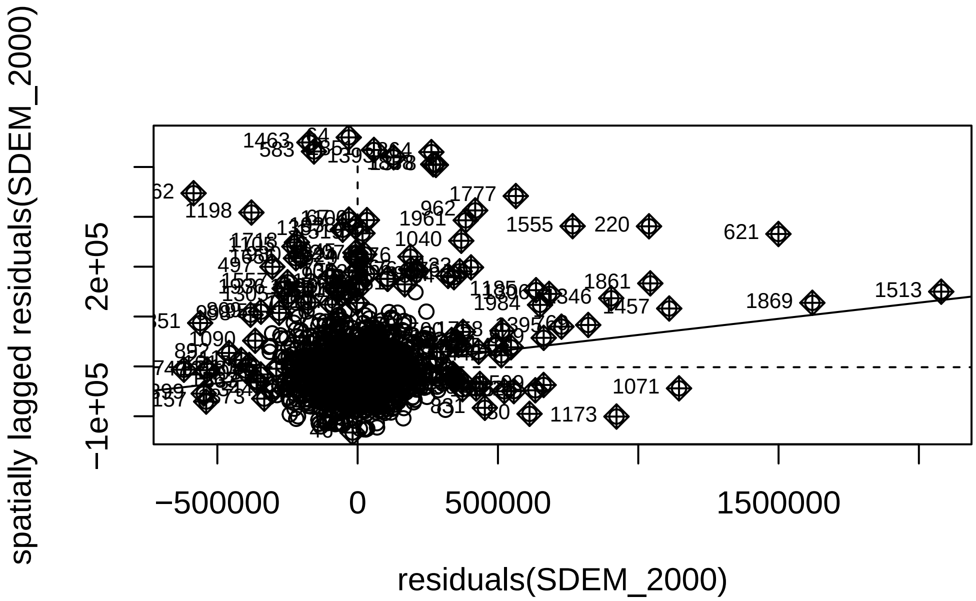
```r
bptest.Sarlm(SEM_2000, studentize = TRUE)
```

```
    studentized Breusch-Pagan test

data:
BP = 284.47, df = 30, p-value < 2.2e-16
```

```r
moran.plot(residuals(hedon3_2000), listw = kc_house_data_2000_W10)
```




```r
moran.plot(residuals(SDEM_2000), listw = kc_house_data_2000_W10)
```




```r
moran.test(residuals(SDEM_2000), listw = kc_house_data_2000_W10)
```

```
    Moran I test under randomisation

data:  residuals(SDEM_2000)
weights: kc_house_data_2000_W10

Moran I statistic standard deviate = 6.9648, p-value = 1.644e-12
alternative hypothesis: greater
sample estimates:
Moran I statistic       Expectation          Variance
     6.471588e-02     -5.002501e-04      8.767832e-05
```

Oppg. 10 Oppsummering

Etter å satt opp all data, og sammenlignet datasettet vi begynte med og de andre datasettene så ser vi at resultatene er ganske like. Videre ser vi at det ved å bruke et tilfeldig utvalg i motsetning til et definert utvalg er det ikke så stor forskjell.

Referanser

Bishop, Kelly C., Nicolai V. Kuminoff, H. Spencer Banzhaf, Kevin J. Boyle, Kathrine von Gravenitz, Jaren C. Pope, V. Kerry Smith, and Christopher D. Timmins. 2020. "Best Practices for Using Hedonic Property Value Models to Measure Willingness to Pay for Environmental Quality." *Review of Environmental Economics and Policy* 14 (2): 260–81. https://doi.org/10.1093/reep/reaa001.

	Hedon1	Hedon2	Hedon3
(Intercept)	6210567.524 ***	3372625.196 ***	3953788.747 ***
	[44.647]	[21.435]	[29.006]
bedrooms	-39596.759 ***	-26343.443 ***	-32336.935 ***
	[-19.444]	[-15.106]	[-17.223]
bathrooms	46467.945 ***	27539.039 ***	40542.973 ***
	[13.256]	[9.196]	[12.568]
sqft_living	167.682 ***	129.590 ***	157.120 ***
	[35.847]	[32.127]	[36.379]
sqft_living15	24.038 ***	33.258 ***	-5.904
	[6.662]	[10.233]	[-1.721]
sqft_lot	-0.003	0.211 ***	0.068
	[-0.061]	[4.598]	[1.433]
sqft_lot15	-0.556 ***	-0.178 *	-0.503 ***
	[-7.100]	[-2.564]	[-6.938]
sqft_above	-6.495	82.422 ***	25.189 ***
	[-1.426]	[20.251]	[5.938]
floors	26914.959 ***	-60197.734 ***	-4573.083
	[7.100]	[-17.189]	[-1.295]
grade	120002.514 ***	66185.300 ***	91928.479 ***
	[53.176]	[32.615]	[43.269]
yr_built	-3584.346 ***	-663.856 ***	-2604.253 ***
	[-50.313]	[-8.839]	[-38.234]
yr_renovated	10.616 **	31.350 ***	16.397 ***
	[2.706]	[9.305]	[4.537]
waterfront	579243.823 ***	614229.057 ***	613889.766 ***
	[31.092]	[38.668]	[35.803]
condition	20271.885 ***	30566.687 ***	25868.926 ***
	[8.057]	[14.046]	[11.091]
view	42889.032 ***	47818.802 ***	48886.142 ***
	[18.823]	[24.372]	[23.304]
year_month2014-06	3436.683	8506.371	6587.530

Res.Df	RSS	Df	Sum of Sq	F	Pr(>F)
2.14e+04	1e+15				
2.14e+04	9.97e+14	12	5.35e+12	9.57	7.55e-19

Res.Df	RSS	Df	Sum of Sq	F	Pr(>F)
2.12e+04	7.09e+14				
2.12e+04	7.03e+14	12	6e+12	15	4.97e-32

Res.Df	RSS	Df	Sum of Sq	F	Pr(>F)
2.14e+04	8.5e+14				
2.14e+04	8.43e+14	12	6.15e+12	13	4.17e-27

	Full	Seed
(Intercept)	3953788.747 ***	3899824.076 ***
	[29.006]	[9.035]
bedrooms	-32336.935 ***	-20996.271 ***
	[-17.223]	[-3.357]
bathrooms	40542.973 ***	30567.528 **
	[12.568]	[2.995]
sqft_living	157.120 ***	117.958 ***
	[36.379]	[8.552]
sqft_living15	-5.904	20.030
	[-1.721]	[1.766]
sqft_lot	0.068	0.670 **
	[1.433]	[3.227]
sqft_lot15	-0.503 ***	-1.211 ***
	[-6.938]	[-3.595]
sqft_above	25.189 ***	28.851 *
	[5.938]	[2.098]
floors	-4573.083	-4802.301
	[-1.295]	[-0.427]
grade	91928.479 ***	105715.611 ***
	[43.269]	[15.560]
yr_built	-2604.253 ***	-2582.541 ***
	[-38.234]	[-11.948]
yr_renovated	16.397 ***	15.421
	[4.537]	[1.388]
waterfront	613889.766 ***	496287.754 ***
	[35.803]	[9.121]
condition	25868.926 ***	16714.366 *
	[11.091]	[2.285]
view	48886.142 ***	41654.941 ***
	[23.304]	[5.699]
dist_cbd_km	3988.205 ***	4383.450 ***

	SEM	OLS
(Intercept)	757342.817	3899824.076 ***
	[1.862]	[9.035]
bedrooms	-16212.659 **	-20996.271 ***
	[-3.076]	[-3.357]
bathrooms	16593.606 *	30567.528 **
	[1.967]	[2.995]
sqft_living	93.458 ***	117.958 ***
	[8.272]	[8.552]
sqft_living15	15.370	20.030
	[1.534]	[1.766]
sqft_lot	0.725 ***	0.670 **
	[4.246]	[3.227]
sqft_lot15	-0.672 *	-1.211 ***
	[-2.252]	[-3.595]
sqft_above	67.812 ***	28.851 *
	[5.925]	[2.098]
floors	-29785.276 **	-4802.301
	[-2.976]	[-0.427]
grade	78822.963 ***	105715.611 ***
	[13.440]	[15.560]
yr_built	-852.789 ***	-2582.541 ***
	[-4.156]	[-11.948]
yr_renovated	19.861 *	15.421
	[2.255]	[1.388]
waterfront	559770.761 ***	496287.754 ***
	[12.824]	[9.121]
condition	21731.896 ***	16714.366 *
	[3.584]	[2.285]
view	45045.733 ***	41654.941 ***
	[7.209]	[5.699]
dist_cbd_km	6720.022 ***	4383.450 ***

	Full	2000 Seed	1111 Seed
(Intercept)	3953788.747 ***	4597674.068 ***	3899824.076 ***
	[29.006]	[10.281]	[9.035]
bedrooms	-32336.935 ***	-17554.778 **	-20996.271 ***
	[-17.223]	[-2.819]	[-3.357]
bathrooms	40542.973 ***	49559.854 ***	30567.528 **
	[12.568]	[4.617]	[2.995]
sqft_living	157.120 ***	148.307 ***	117.958 ***
	[36.379]	[10.555]	[8.552]
sqft_living15	-5.904	1.228	20.030
	[-1.721]	[0.113]	[1.766]
sqft_lot	0.068	-0.413	0.670 **
	[1.433]	[-1.821]	[3.227]
sqft_lot15	-0.503 ***	-0.387	-1.211 ***
	[-6.938]	[-1.329]	[-3.595]
sqft_above	25.189 ***	9.639	28.851 *
	[5.938]	[0.690]	[2.098]
floors	-4573.083	-1154.768	-4802.301
	[-1.295]	[-0.098]	[-0.427]
grade	91928.479 ***	98816.290 ***	105715.611 ***
	[43.269]	[14.163]	[15.560]
yr_built	-2604.253 ***	-2937.536 ***	-2582.541 ***
	[-38.234]	[-13.179]	[-11.948]
yr_renovated	16.397 ***	24.309 *	15.421
	[4.537]	[2.050]	[1.388]
waterfront	613889.766 ***	369162.768 ***	496287.754 ***
	[35.803]	[6.340]	[9.121]
condition	25868.926 ***	17448.457 *	16714.366 *
	[11.091]	[2.255]	[2.285]
view	48886.142 ***	34984.655 ***	41654.941 ***
	[23.304]	[5.113]	[5.699]
dist_cbd_km	3988.205 ***	3972.942 ***	4383.450 ***

	SEM	OLS
(Intercept)	1604203.897 ***	4597674.068 ***
	[3.880]	[10.281]
bedrooms	-15023.697 **	-17554.778 **
	[-2.983]	[-2.819]
bathrooms	32052.829 ***	49559.854 ***
	[3.804]	[4.617]
sqft_living	123.587 ***	148.307 ***
	[10.918]	[10.555]
sqft_living15	7.383	1.228
	[0.781]	[0.113]
sqft_lot	-0.080	-0.413
	[-0.434]	[-1.821]
sqft_lot15	-0.064	-0.387
	[-0.262]	[-1.329]
sqft_above	48.731 ***	9.639
	[4.106]	[0.690]
floors	-21051.815 *	-1154.768
	[-2.024]	[-0.098]
grade	69305.782 ***	98816.290 ***
	[11.919]	[14.163]
yr_built	-1287.842 ***	-2937.536 ***
	[-6.204]	[-13.179]
yr_renovated	25.546 **	24.309 *
	[2.739]	[2.050]
waterfront	421567.958 ***	369162.768 ***
	[8.601]	[6.340]
condition	26349.156 ***	17448.457 *
	[4.130]	[2.255]
view	33927.849 ***	34984.655 ***
	[5.740]	[5.113]
dist_cbd_km	5113.099 ***	3972.942 ***