

Termpaper

Sindre og Morten

“Data gitt med tillatelse fra King County” ^ (**KCGISC?**).

Oppgave 1

A

Hedoniske modeller blir tatt i bruk for å måle marginal «willingness to pay» (MWTP), denne blir tatt i bruk for å måle boligmarkedet sine miljømessige kvaliteter. Den hedoniske modellen ble fremstilt ved å ta i bruk et stort og avansert datasett fra boligmarkeder i store byer som har en avansert økonomi. Jo mer tilgang vi får på data til disse byområdene, desto bedre analyser får vi med bruk av en hedonisk modell.

Rosens første steg handler om å kunne definere et marked som vil tilfredsstille «law of one function». «Law of one function» er at hus som er identiske blir solgt for den samme prisen i et gitt marked. Steg to handler om data innsamling. Det forteller oss at det er met oppnålig med å ha et tilfeldig utvalg. Slike datainnsamlinger til en hedonisk modell som omhandler eiendomsverdi har som regel fokus på eneboliger.

C - Why could it be important to define a housing market as a single metropolitan (or travel to work) area and studying a relatively short period of time, when estimating

i. hedonic price function?

Bishop forklarer at i hedonisk modell innebærer det at markedet bør defineres slik at «loven om en prisfunksjon» oppfylles (Bishop mfl. 2020). Med en prisfunksjon menes det at identiske boliger vil selges til samme pris gjennom hele markedet. Likevel, de nøyaktige romlige og tidsmessige grensene som tilfredsstiller denne betingelsen kan variere på tvers av rom og tid ettersom informasjon, institusjoner og flyttekostnader endres (Bishop mfl. 2020). I praksis er det vanlig å definere markedet som et enkelt storbyområde (single metropolitan) over noen år. Flyttekostnader vil egentlig bryte med loven om en prisfunksjon, men for husholdninger som

flytter innenfor dette storbyområdet er det lite sannsynlig at disse kostnadene vil variere noe særlig.

Årsaken til det er at de fysiske kostnadene som vel som de økonomiske (f.eks lastebil-leie) ikke endrer seg på tvers av destinasjonssteder i hovedstadsområdet. De psykologiske kostnadene er også mer begrenset ved flytting fordi det tillates i lettere grad å opprettholde relasjonene til familie, venner og nabolag. Dette gjør at loven om én prisfunksjon opprettholdes mellom lokasjoner i et storbyområde gjennom arbitrasje. Altså, hvis to hus som er tilsvarende like i samme storbyområdet selges, så velger kjøperne naturligvis det rimeligste.

ii. Explain intuitively, by using an example, why it is important to avoid omitted variable bias when estimating a hedonic price model.

Gjennom empirien og teorien er det grunn til å tro av miljøfasiliteter er romlig korrelert på grunn av de naturlige trekken ved geografien som for eksempel fjell og hav, miljøtilbakemeldingseffekter (f.eks urbane varme øyer) og stemming på lokale felles goder. Dette potensialet for romlig korrelasjon har ført til utbredt bekymring for utelatt-variabel skjevhets (Bishop mfl. 2020). Det er først og fremst fordi det virker usannsynlig at forskere vil være i stand til å inkludere alle bekvemmeligheter som betyr noe for kjøpere. I tillegg vil uobserverte fasiliteter sannsynligvis være korrelert med tilbuddet av interesse, og dermed forårsake skjevhets. Dette kan forklares ved et eksempel: Hvis velstående og velutdannede boligkjøpere flytter til områder med bedre luftkvalitet og deretter stemmer for å øke offentlig skolefinansiering, vil estimatorer av MWTP for luftkvalitet være skjev oppover hvis skolekvalitet uteslutes fra modellen. Potensialet for denne typen oppførsel fra huseiere betyr at for at de resulterende estimatene skal være troværdige, må forskningsdesignet isolere eksogen variasjon i tilbuddet av interesse (Bishop mfl. 2020).

Oppgave 2

i.

Så over variablene og definisjonene på dem hos Kaggle.

ii.

Laster inn data

```
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.

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

iv.

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

v & vi.

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

vii.

koordinatene til Seattle:

Norske (wikipedia): - 47.60621, -122.33207

Engelske (wikipedia): - 47.609722, -122.333056

```
cbd <- st_sfc(st_point(c(-122.33207, 47.60621)), crs = 4326) %>%
  st_transform(2926)
```

viii.

```
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)
  )
```

Oppgave 3

```
kc_wadoh_map <- here("maps/WADOH_Environmental_Health_Disparities_Index_Calculated_for_King_County___wadoh_map.shp") %>%
  st_read() %>%
  st_transform(2926)

Reading layer `WADOH_Environmental_Health_Disparities_Index_Calculated_for_King_County___wadoh_map` from file
  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

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 concentration
    ozone,# ozone concentration
    PM25, # concentration 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 determinants, low best
)

acs_b19101_fam_inc <- read.dbf("../maps/censusSHP/acs_b19101_familyincome.dbf")
attach(acs_b19101_fam_inc)

acs_b19101_fam_inc <- acs_b19101_fam_inc %>%
  mutate(low = (E19101138 + E19101139 + E19101140 + E19101141 +
             E19101142 + E19101143)/E19101137) %>%
  mutate(mid = (E19101144 + E19101145 + E19101146 + E19101147 +
             E19101148 + E19101149)/E19101137) %>%
  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_ID_TRT")

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

Reading layer `tracts10' from data source
`/Users/Morten/Documents/Boligmarked/maps/censusSHP/tracts10.shp'
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)

```

```

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

Reading layer `tracts10_shore' from data source
  `/Users/Morten/Documents/Boligmarked/maps/censusSHP/tracts10_shore.shp'
  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)

kc_tracts10_env_data <- left_join(
  kc_tracts10, kc_wadoh_map_2,
  by = "GEO_ID_TRT"
)
kc_tracts10_shore_env_data <- left_join(
  kc_tracts10_shore, kc_wadoh_map_2,
  by = "GEO_ID_TRT"
)

kc_houses_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)

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.

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.

```

```
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.
```

```
st_write(kc_houses_env_var, ".../maps/kc_houses_env_var.gpkg", append = FALSE)
```

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

```
st_write(kc_tracts10_shore_env_var, ".../maps/kc_tracts10_shore_env_var.gpkg", append = FALSE)
```

```
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.
```

Oppgave 4

i.

```
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						

```
summary(kc_tracts10_shore_env_var)
```

	id	date	price
Length:	21436	Min. : 2014-05-02 00:00:00.00	Min. : 75000
Class :	character	1st Qu.: 2014-07-22 00:00:00.00	1st Qu.: 324866
Mode :	character	Median : 2014-10-17 00:00:00.00	Median : 450000
		Mean : 2014-10-29 17:30:02.34	Mean : 541650
		3rd Qu.: 2015-02-18 00:00:00.00	3rd Qu.: 645000
		Max. : 2015-05-27 00:00:00.00	Max. : 7700000
	bedrooms	bathrooms	sqft_living
Min. :	0.000	Min. : 0.000	Min. : 290
1st Qu.:	3.000	1st Qu.: 1.750	1st Qu.: 1430
Median :	3.000	Median : 2.250	Median : 1920
Mean :	3.372	Mean : 2.117	Mean : 2083
3rd Qu.:	4.000	3rd Qu.: 2.500	3rd Qu.: 2550
Max. :	33.000	Max. : 8.000	Max. : 13540
	sqft_lot		
Min. :	520	1st Qu.:	5040
Median :	7614	Mean :	15136
3rd Qu.:	10696	3rd Qu.:	10696
Max. :	1651359	Max. :	1651359
	floors	waterfront	view
Min. :	1.000	Min. : 0.000000	Min. : 0.0000
1st Qu.:	1.000	1st Qu.: 0.000000	1st Qu.: 0.0000
Median :	1.500	Median : 0.000000	Median : 0.0000
Mean :	1.496	Mean : 0.007604	Mean : 0.2351
3rd Qu.:	2.000	3rd Qu.: 0.000000	3rd Qu.: 0.0000
Max. :	3.500	Max. : 1.000000	Max. : 4.0000
	condition		
Min. :	1.00	1st Qu.:	3.00
Median :	3.00	Mean :	3.41
3rd Qu.:	4.00	3rd Qu.:	4.00
Max. :	5.00	Max. :	5.00
	grade	sqft_above	sqft_basement
Min. :	1.000	Min. : 290	Min. : 0.0
1st Qu.:	7.000	1st Qu.: 1200	1st Qu.: 0.0
Median :	7.000	Median : 1560	Median : 0.0
Mean :	7.662	Mean : 1791	Mean : 291.7
3rd Qu.:	8.000	3rd Qu.: 2220	3rd Qu.: 560.0
Max. :	13.000	Max. : 9410	Max. : 4820.0
	yr_built		
Min. :	1900	1st Qu.:	1952
Median :	1975	Mean :	1971
3rd Qu.:	1997	3rd Qu.:	1997
Max. :	2015	Max. :	2015
	sqft_living15	sqft_lot15	
Min. :	399	Min. : 651	Min. : 651
1st Qu.:	1490	1st Qu.: 5100	1st Qu.: 5100
Median :	1840	Median : 7620	Median : 7620
Mean :	1988	Mean : 12786	Mean : 12786
	zipcode		
Min. :	98001	Min. :	651
1st Qu.:	98033	1st Qu.: 5100	1st Qu.: 5100
Median :	98065	Median : 7620	Median : 7620
Mean :	98078	Mean : 12786	Mean : 12786

3rd Qu.:	0.00	3rd Qu.:	98117	3rd Qu.:	2370	3rd Qu.:	10087					
Max.	:	2015.00		Max.	:	98199	Max.	:	6210	Max.	:	871200

	geometry	dist_cbd	dist_cbd_km	GEO_ID_TRT
POINT	:21436	Min. : 3228	Min. : 0.9838	Length:21436
epsg:2926	: 0	1st Qu.: 32099	1st Qu.: 9.7837	Class :character
+proj=lcc	...: 0	Median : 54280	Median :16.5447	Mode :character
		Mean : 60638	Mean :18.4824	
		3rd Qu.: 83064	3rd Qu.:25.3178	
		Max. :253647	Max. :77.3117	

	FEATURE_ID	TRACT_LBL	TRACT_STR	TRACT_INT
Min.	:10153	Length:21436	Length:21436	Min. : 100
1st Qu.:	36346	Class :character	Class :character	1st Qu.:10402
Median :	44764	Mode :character	Mode :character	Median :24702
Mean :	38270			Mean :21224
3rd Qu.:	45279			3rd Qu.:31202
Max. :	45838			Max. :32800
NA's :	25			NA's :25
	TRACT_FLT	TRACT_DEL	TRTLABEL_F	TRTLABEL_C
Min. :	1.0	Length:21436	Length:21436	Length:21436
1st Qu.:	104.0	Class :character	Class :character	Class :character
Median :	247.0	Mode :character	Mode :character	Mode :character
Mean :	212.2			
3rd Qu.:	312.0			
Max. :	328.0			
NA's :	25			
	TRTLABEL_T	COUNTY_STR	COUNTY_INT	STATE_STR
Length:21436		Length:21436	Min. :33	Length:21436
Class :character		Class :character	1st Qu.:33	Class :character
Mode :character		Mode :character	Median :33	Mode :character
			Mean :33	
			3rd Qu.:33	
			Max. :33	
			NA's :25	
	STATE_INT	LEVEL_1	LEVEL_2	LEVEL_3
Min. :	53	Length:21436	Length:21436	Length:21436
1st Qu.:	53	Class :character	Class :character	Class :character
Median :	53	Mode :character	Mode :character	Mode :character
Mean :	53			
3rd Qu.:	53			
Max. :	53			
NA's :	25			

TRACT_AREA	TRACT_PERI	LOGRECNO	Shape_area
Min. : 2.792e+06	Min. : 8012	Length:21436	Min. : 2.792e+06
1st Qu.: 2.485e+07	1st Qu.: 23500	Class :character	1st Qu.: 2.281e+07
Median : 4.123e+07	Median : 32920	Mode :character	Median : 3.445e+07
Mean : 1.809e+08	Mean : 48212		Mean : 1.750e+08
3rd Qu.: 7.308e+07	3rd Qu.: 47962		3rd Qu.: 6.628e+07
Max. : 1.526e+10	Max. : 738820		Max. : 1.526e+10
NA's : 25	NA's : 25		NA's : 25
Shape_len	low	mid	high
Min. : 8012	Min. : 0.009298	Min. : 0.06768	Min. : 0.06129
1st Qu.: 23204	1st Qu.: 0.047091	1st Qu.: 0.21668	1st Qu.: 0.47602
Median : 31185	Median : 0.074766	Median : 0.30219	Median : 0.61143
Mean : 46861	Mean : 0.100082	Mean : 0.31115	Mean : 0.58877
3rd Qu.: 46624	3rd Qu.: 0.133557	3rd Qu.: 0.39313	3rd Qu.: 0.72987
Max. : 738820	Max. : 0.501433	Max. : 0.67904	Max. : 0.88162
NA's : 25	NA's : 25	NA's : 25	NA's : 25
EHD_percen	linguist_2	poverty_pe	POC_percen
Min. : 1.00	Min. : 0.450	Min. : 1.97	Min. : 7.54
1st Qu.: 19.00	1st Qu.: 3.120	1st Qu.: 8.93	1st Qu.: 21.13
Median : 41.00	Median : 7.000	Median : 13.60	Median : 33.26
Mean : 43.64	Mean : 9.003	Mean : 16.65	Mean : 35.33
3rd Qu.: 67.00	3rd Qu.: 12.730	3rd Qu.: 22.95	3rd Qu.: 46.34
Max. : 100.00	Max. : 40.350	Max. : 75.48	Max. : 92.70
NA's : 25	NA's : 220	NA's : 25	NA's : 25
transporta	unemploy_2	housing_pe	traffic_pe
Min. : 12.00	Min. : 1.000	Min. : 15.14	Min. : 0.00
1st Qu.: 18.00	1st Qu.: 3.230	1st Qu.: 25.64	1st Qu.: 0.00
Median : 20.00	Median : 4.310	Median : 30.46	Median : 0.10
Mean : 19.77	Mean : 4.775	Mean : 31.37	Mean : 11.52
3rd Qu.: 21.00	3rd Qu.: 6.050	3rd Qu.: 35.73	3rd Qu.: 19.14
Max. : 26.00	Max. : 13.620	Max. : 64.87	Max. : 84.98
NA's : 25	NA's : 102	NA's : 25	NA's : 25
diesel	ozone	PM25	toxic_rele
Min. : 0.14	Min. : 46.73	Min. : 3.787	Min. : 823.9
1st Qu.: 5.60	1st Qu.: 49.24	1st Qu.: 5.488	1st Qu.: 4143.8
Median : 10.16	Median : 49.97	Median : 6.044	Median : 8827.6
Mean : 13.68	Mean : 51.17	Mean : 6.002	Mean : 17251.1
3rd Qu.: 16.88	3rd Qu.: 52.32	3rd Qu.: 6.579	3rd Qu.: 17237.2
Max. : 92.63	Max. : 62.89	Max. : 7.897	Max. : 186434.6
NA's : 25	NA's : 25	NA's : 25	NA's : 25
hazardous_	lead_perce	superfund	facilities
Min. : 0.02303	Min. : 0.24	Min. : 0.03454	Min. : 0.0523
1st Qu.: 0.03985	1st Qu.: 5.34	1st Qu.: 0.06595	1st Qu.: 0.1420

Median :0.05160	Median :11.99	Median :0.11046	Median :0.2680
Mean :0.07409	Mean :16.60	Mean :0.21696	Mean :0.5248
3rd Qu.:0.07891	3rd Qu.:26.48	3rd Qu.:0.23841	3rd Qu.:0.7588
Max. :0.63781	Max. :54.68	Max. :1.46778	Max. :3.3682
NA's :25	NA's :25	NA's :25	NA's :25
wastewater	sen_pop_pe	socio_perc	
Min. :0.000000	Min. : 1.0	Min. : 1.00	
1st Qu.:0.000003	1st Qu.: 25.0	1st Qu.: 20.00	
Median :0.000290	Median : 48.0	Median : 43.00	
Mean :0.016168	Mean : 48.1	Mean : 44.51	
3rd Qu.:0.002900	3rd Qu.: 71.0	3rd Qu.: 67.00	
Max. :0.640000	Max. :100.0	Max. :100.00	
NA's :25	NA's :25	NA's :25	

ii.

Tract10 kartet har ikke havet med seg og får med det en NA observasjon som er utenfor countygrensen. **Tracts10_shore** har med havet på kartet, som gjør at det er flere obserasjoner som “havner” ut i havet og med det blir de til NA verdier (25stk).

I QGIS fant vi følgende obserasjoner ved å se på *tracts10*, *tracts10_shore* & *kc_houses_env_var*:

iii.

Dropper Obserasjonen 3518000180 ved å:

```
kc_houses_env_var <- arrange(kc_houses_env_var, desc(id))
kc_houses_env_var.omit <- kc_houses_env_var[-c(11997),]

st_write(kc_houses_env_var.omit, ".../maps/kc_houses_env_var.omit.gpkg", append = FALSE)
```

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

```
kc_houses_env_var.omit <- kc_houses_env_var.omit %>%
  mutate(
    year_month = substr(date, start = 1, stop = 7))
```

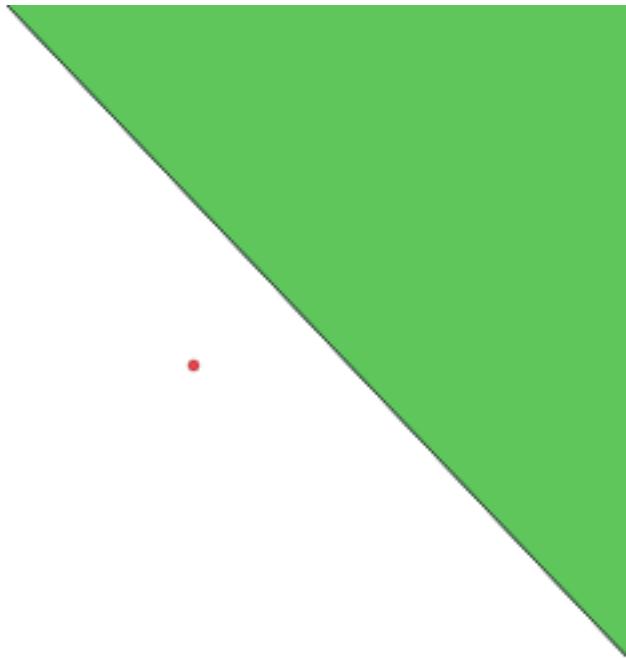


Figure 1: observasjon utenfor WA state

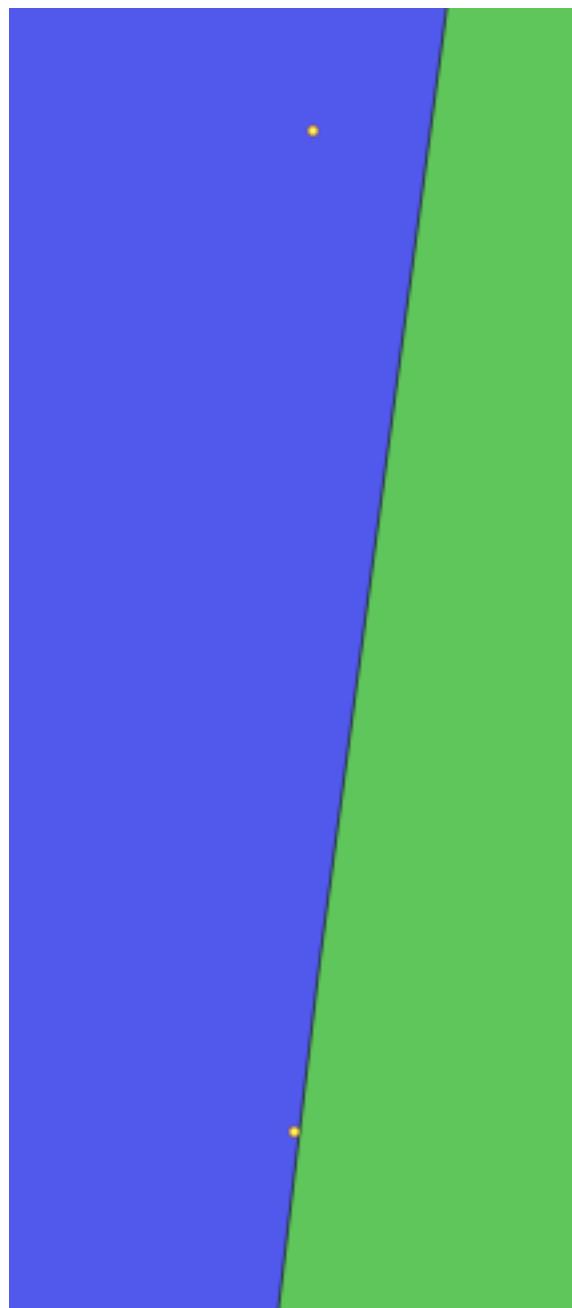


Figure 2: Observasjon utenfor kystlinjen.a



Figure 3: Observasjon utenfor kystlinjen.b

```
st_write(kc_houses_env_var OMIT, ".../maps/kc_houses_env_var OMIT.gpkg", append = FALSE)
```

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

Oppgave 5

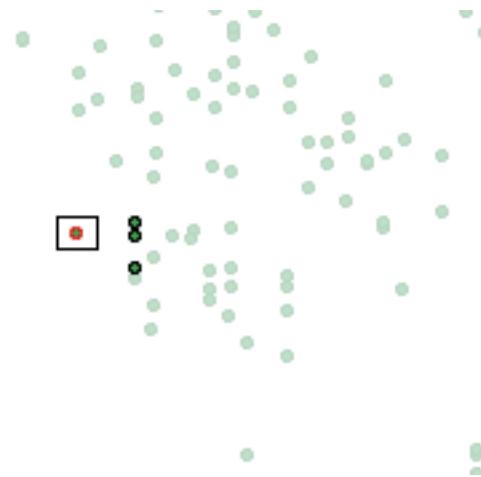


Figure 4: K-nearest 3

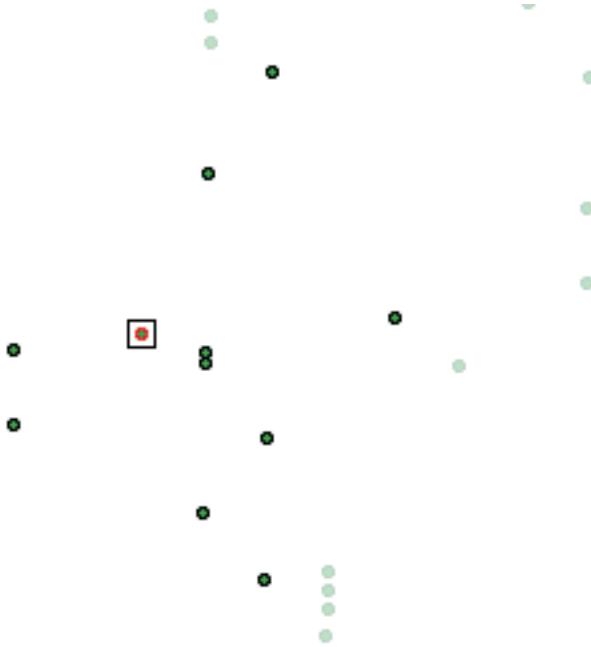


Figure 5: K-nearest 10

Oppgave 6

i)

Funn fra EDA

Vi ser at de store konsentrasjonene av store boliger til høye priser er i og rett rundt Seattle sentrum, vi kan også se at Mercer Island er veldig dyr plass og at vestsiden av Bellevue har store og dyre boliger.

Vi ser at store deler av små boliger til lave priser er i søre del King county.

Vi ser at de små og dyre boligene er rundt bykjernen, noe som gir mening med tanke på at det er mindre områder å bygge store boliger samt et populært område å bo.

Vi kan også se at de store og billige husene plasserer seg sammen med de små og billige boligene, dette sier oss at dette er et fattigere område hvor de lavlønnte bor.

Morans I verdien til K3 er på 0,398 og K10 sin er på 0,350. Når vi skulle regne ut Morgans I valgte vi å bruke price og sqft_living. Vi sammenlignet sqft_living og bedrooms, de så ganske like ut på kartet, men hadde forskjellige verdier på morans I. Verdien er finere med sqft_living fordi det er et tydligere mønster på at når sqft_living øker så blir det dyrere.

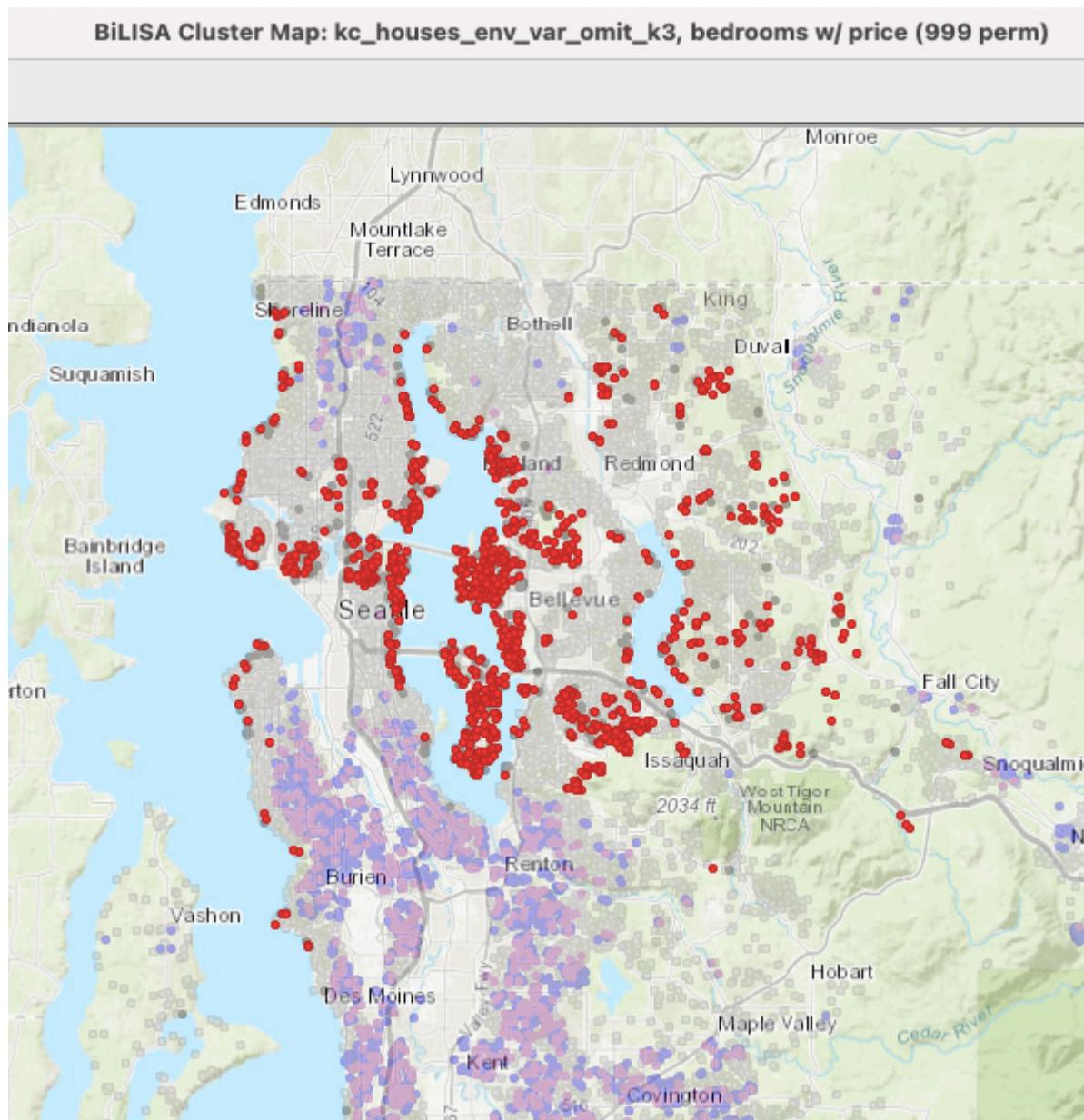


Figure 6: K3 - Store og dyre boliger

BiLISA Cluster Map: kc_houses_env_var_omit_k3, bedrooms w/ price (999 perm)

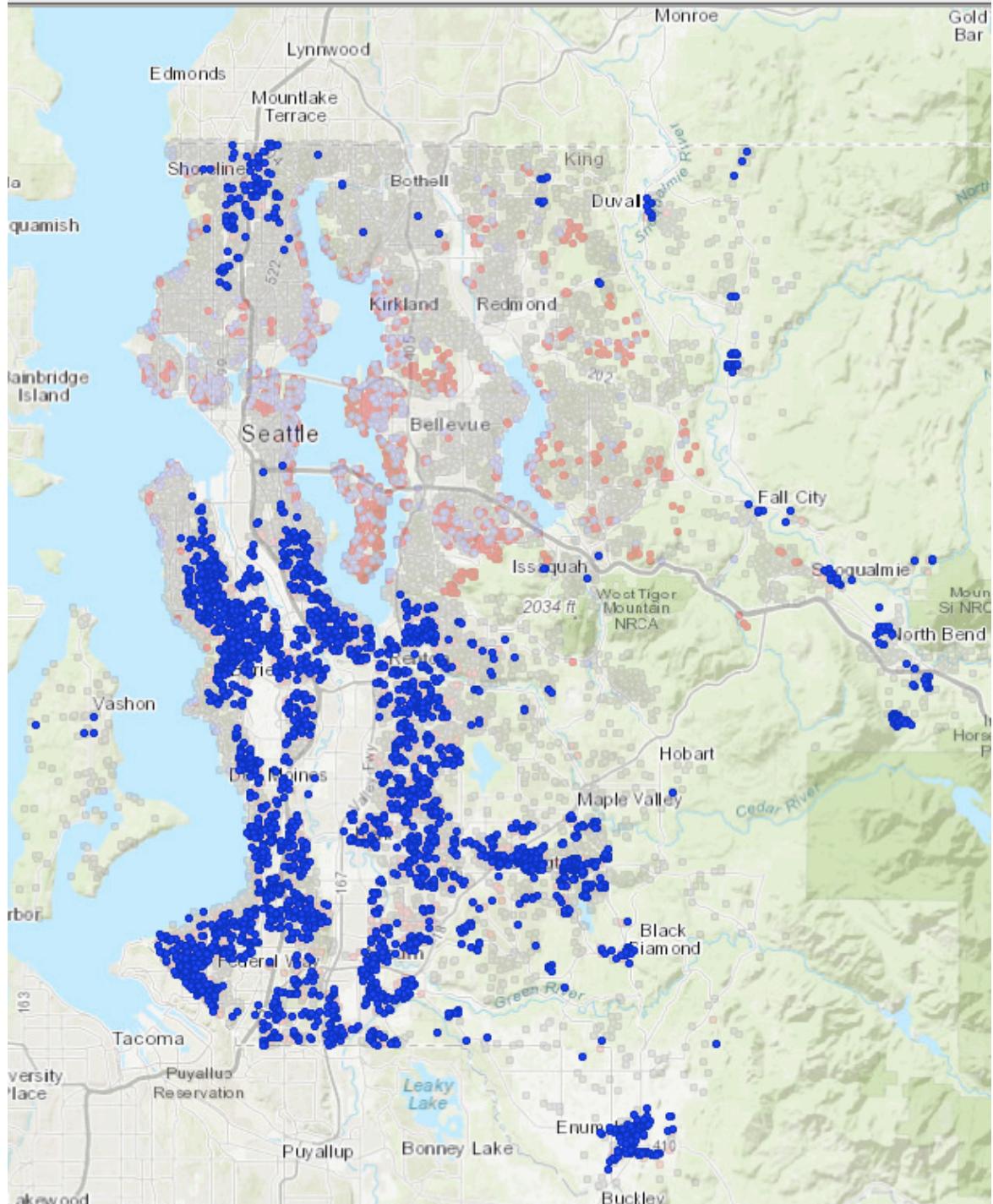


Figure 7: K3 - Små og billige boliger

BiLISA Cluster Map: kc_houses_env_var omit_k3, bedrooms w/ price (999 perm)

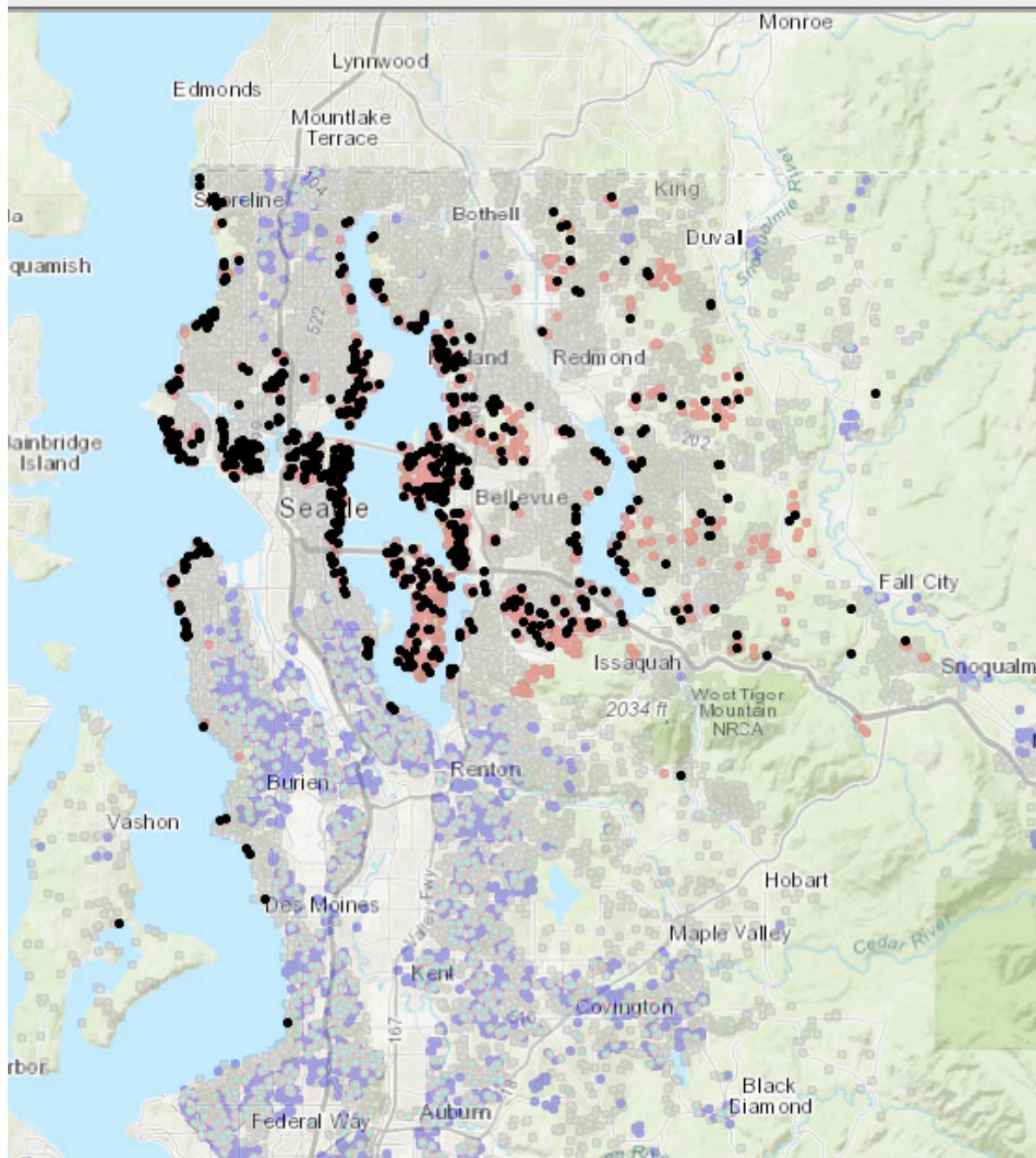


Figure 8: K3 - Små og dyre boliger

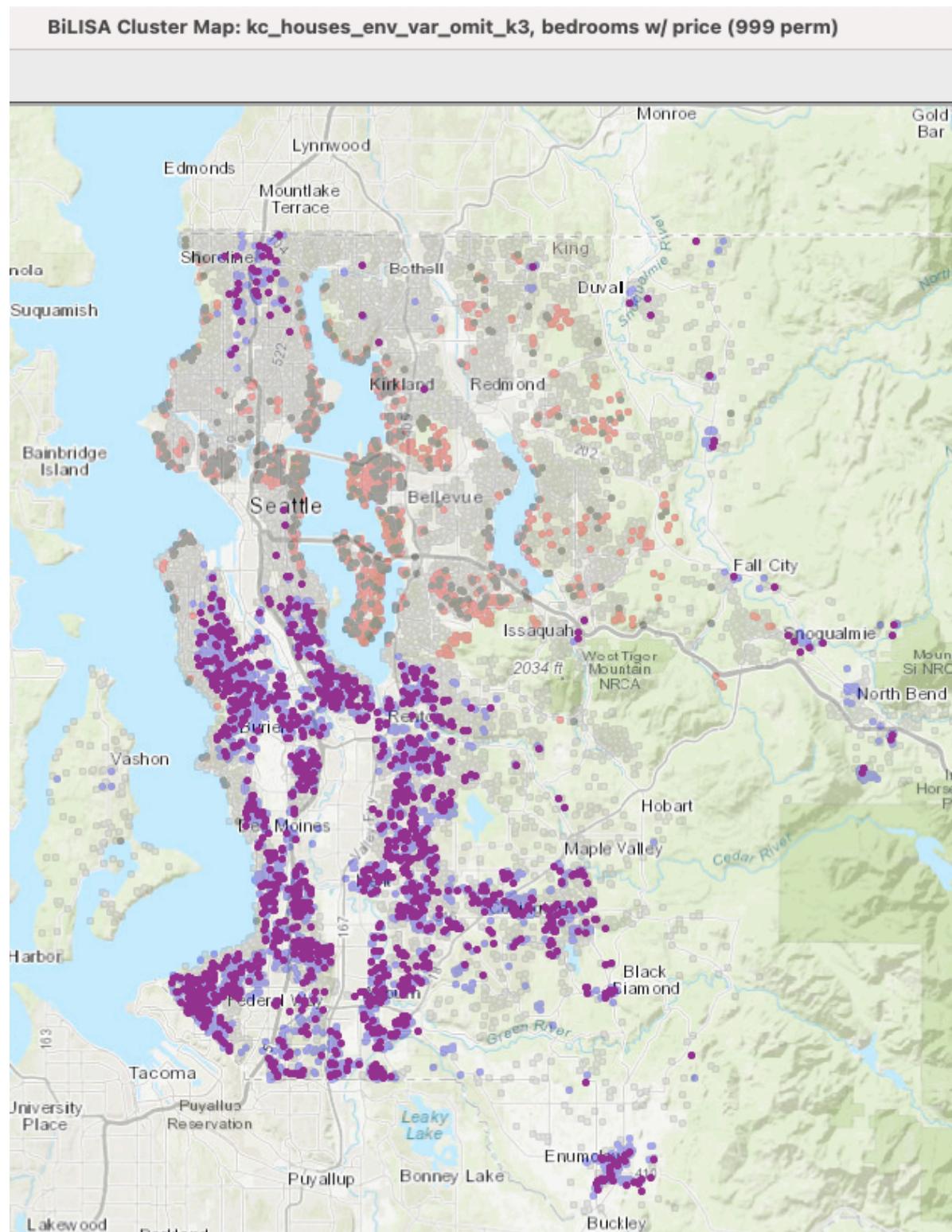


Figure 9: K3 - Store og billige boliger

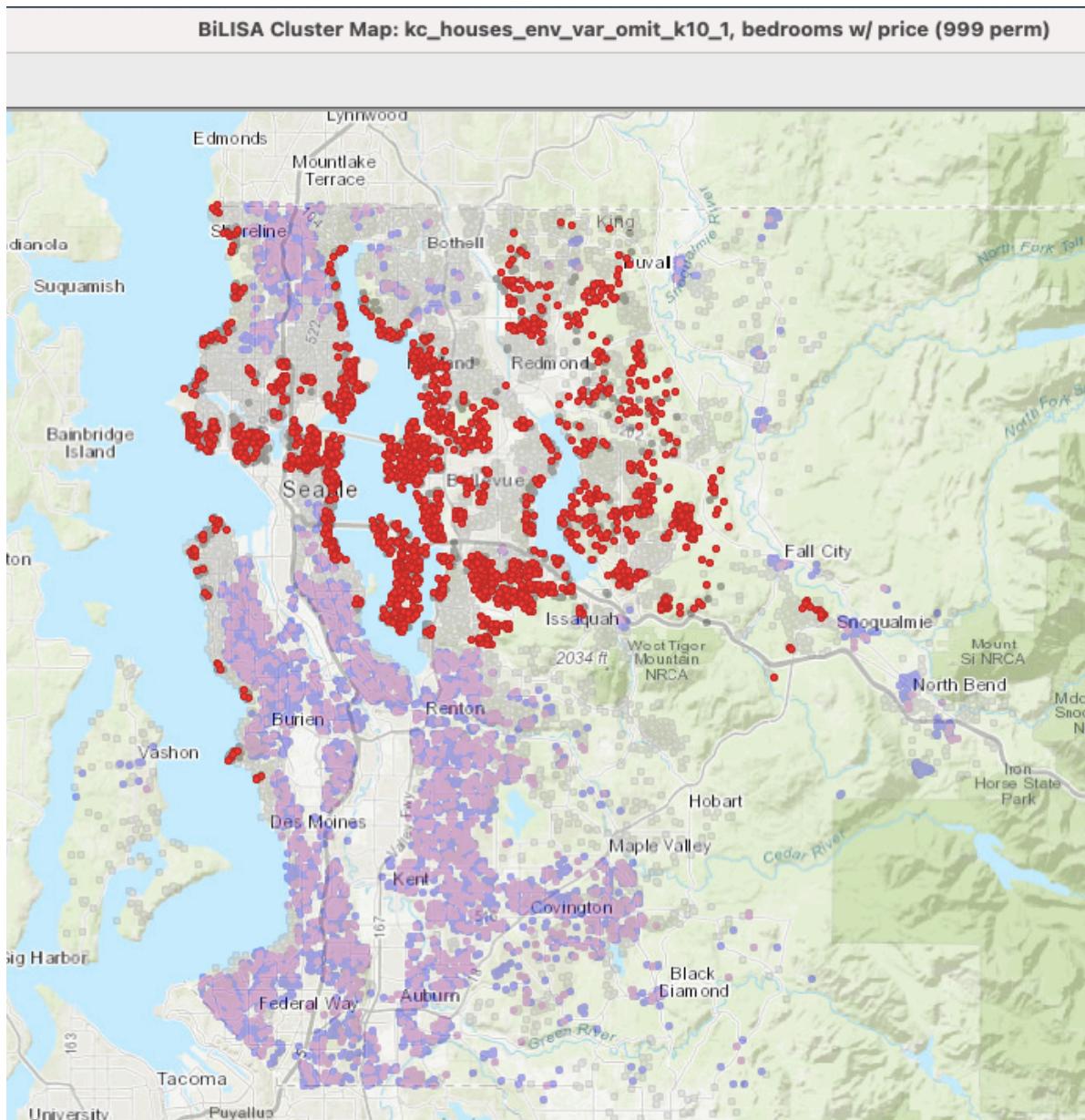


Figure 10: K10 - Store og dyre boliger

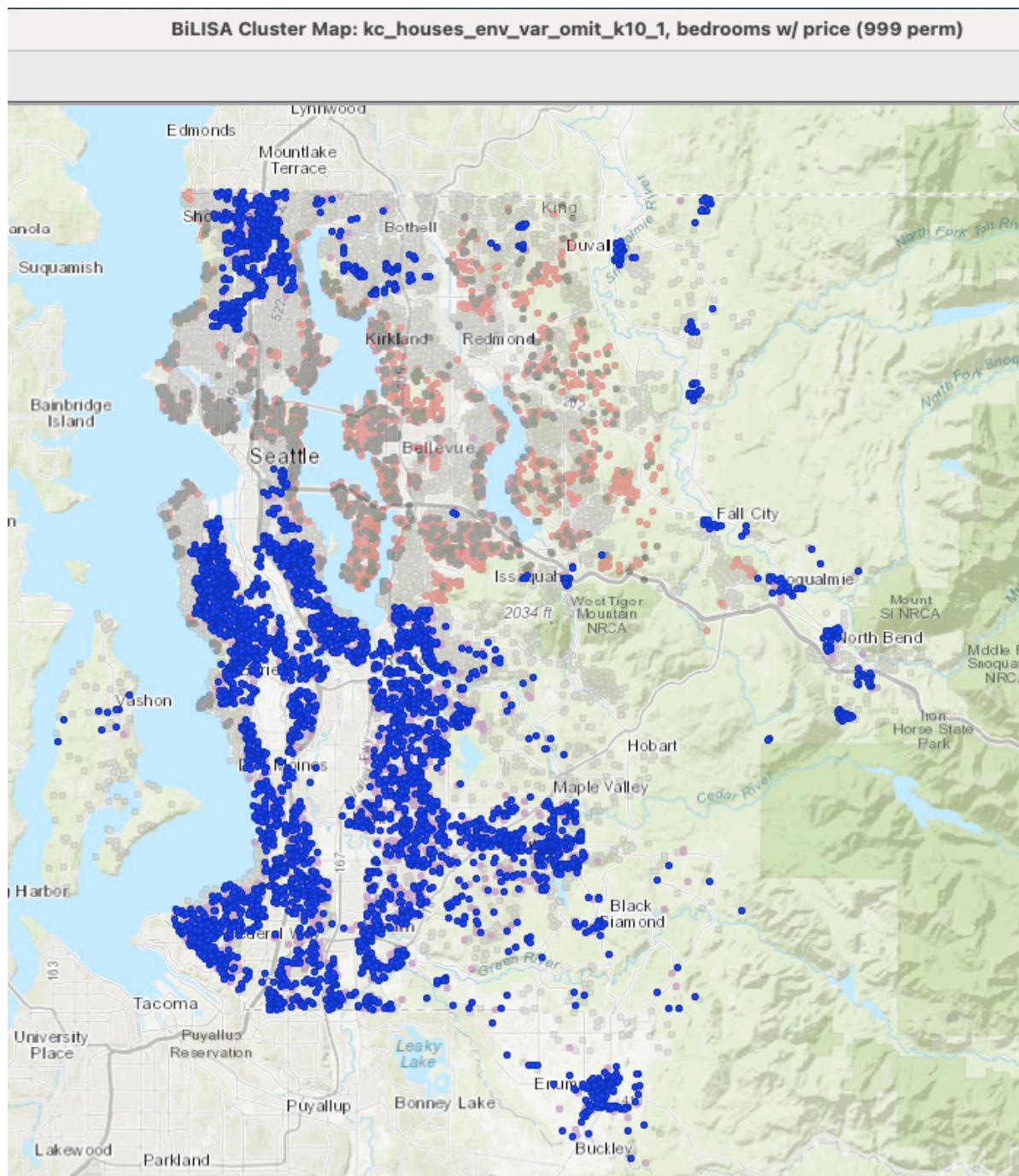


Figure 11: K10 - Små og billige boliger

BiLISA Cluster Map: kc_houses_env_var omit_k10_1, bedrooms w/ price (999 perm)

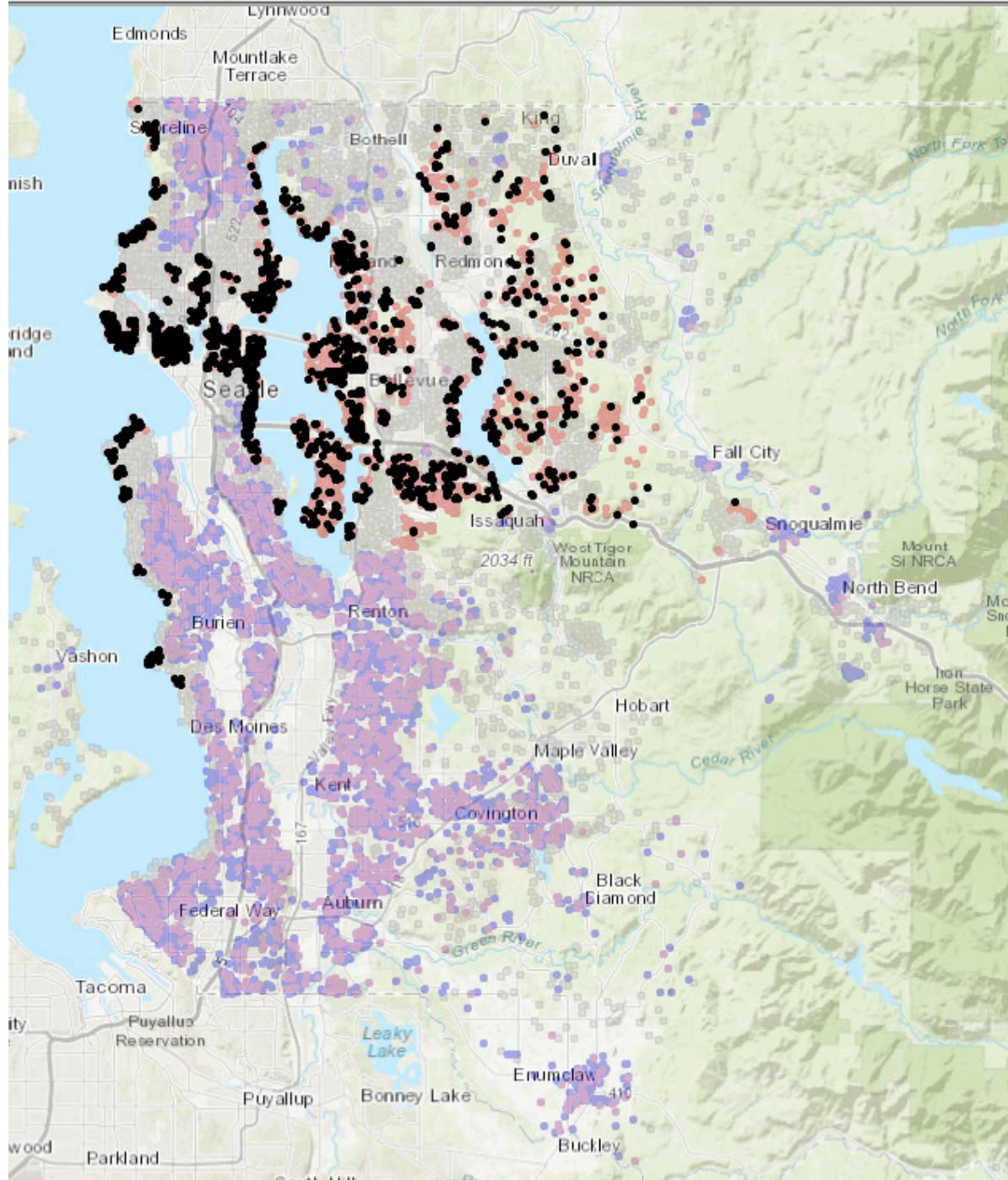


Figure 12: K10 - Små og dyre boliger

BiLISA Cluster Map: kc_houses_env_var omit_k10_1, bedrooms w/ price (999 perm)

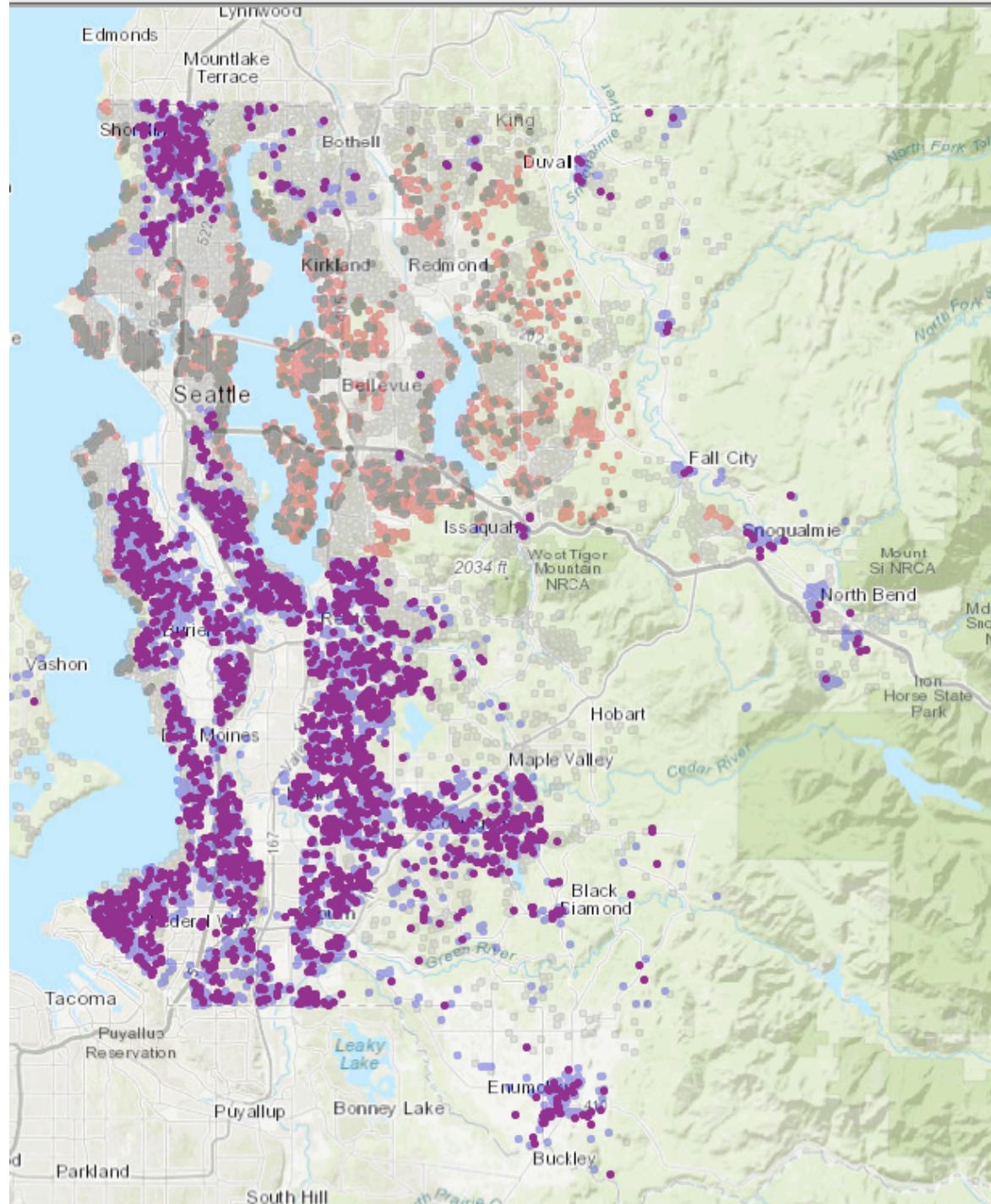


Figure 13: K10 - Store og billige boliger

Bivariate Moran's I (kc_houses_env_var_omit_k3): sqft_living and lagged price

Moran's I: 0.398

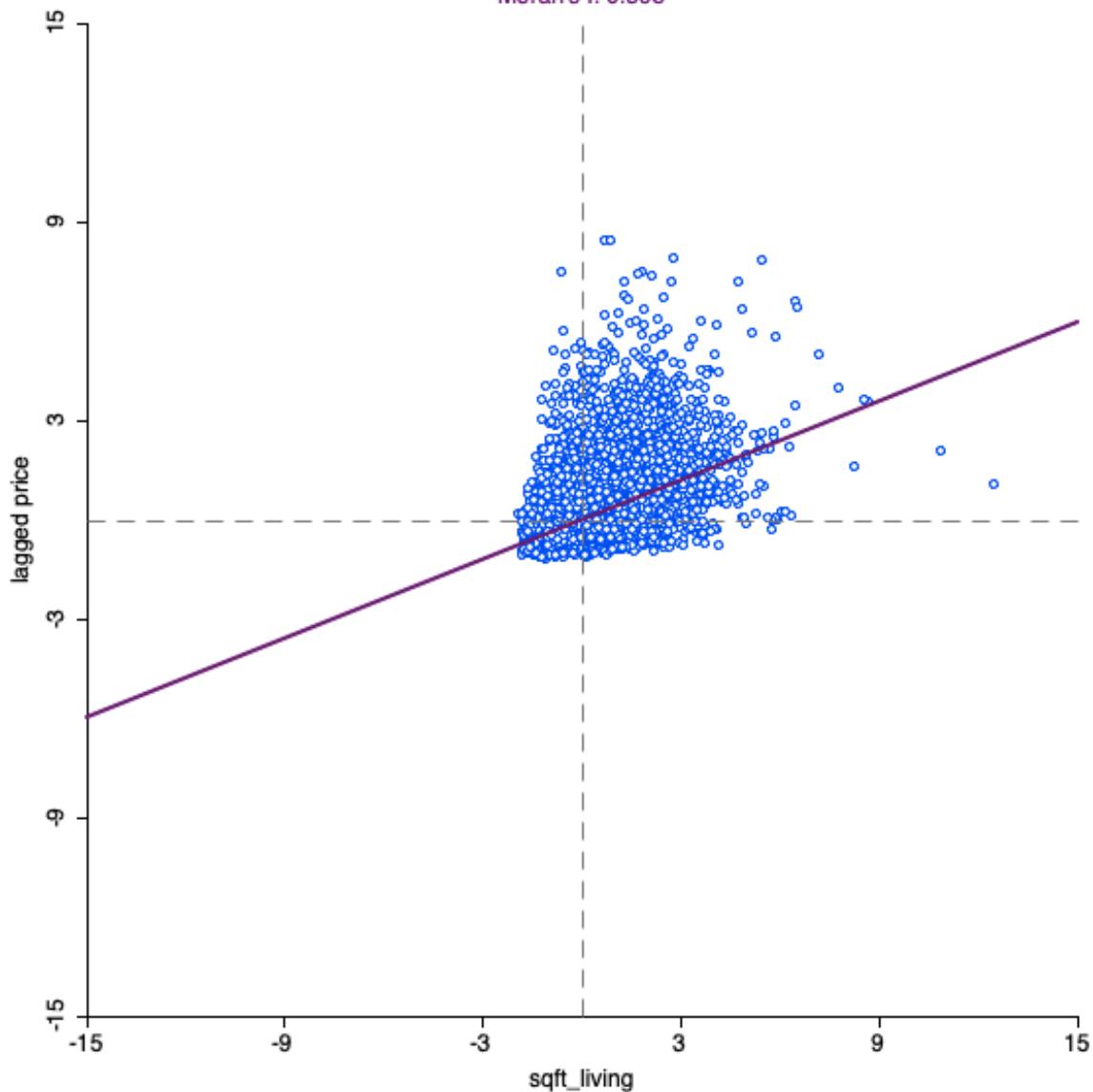


Figure 14: K3 - Bivariate Moran I

Bivariate Moran's I (kc_houses_env_var_omit_k10_1): sqft_living and lagged price

Moran's I: 0.350

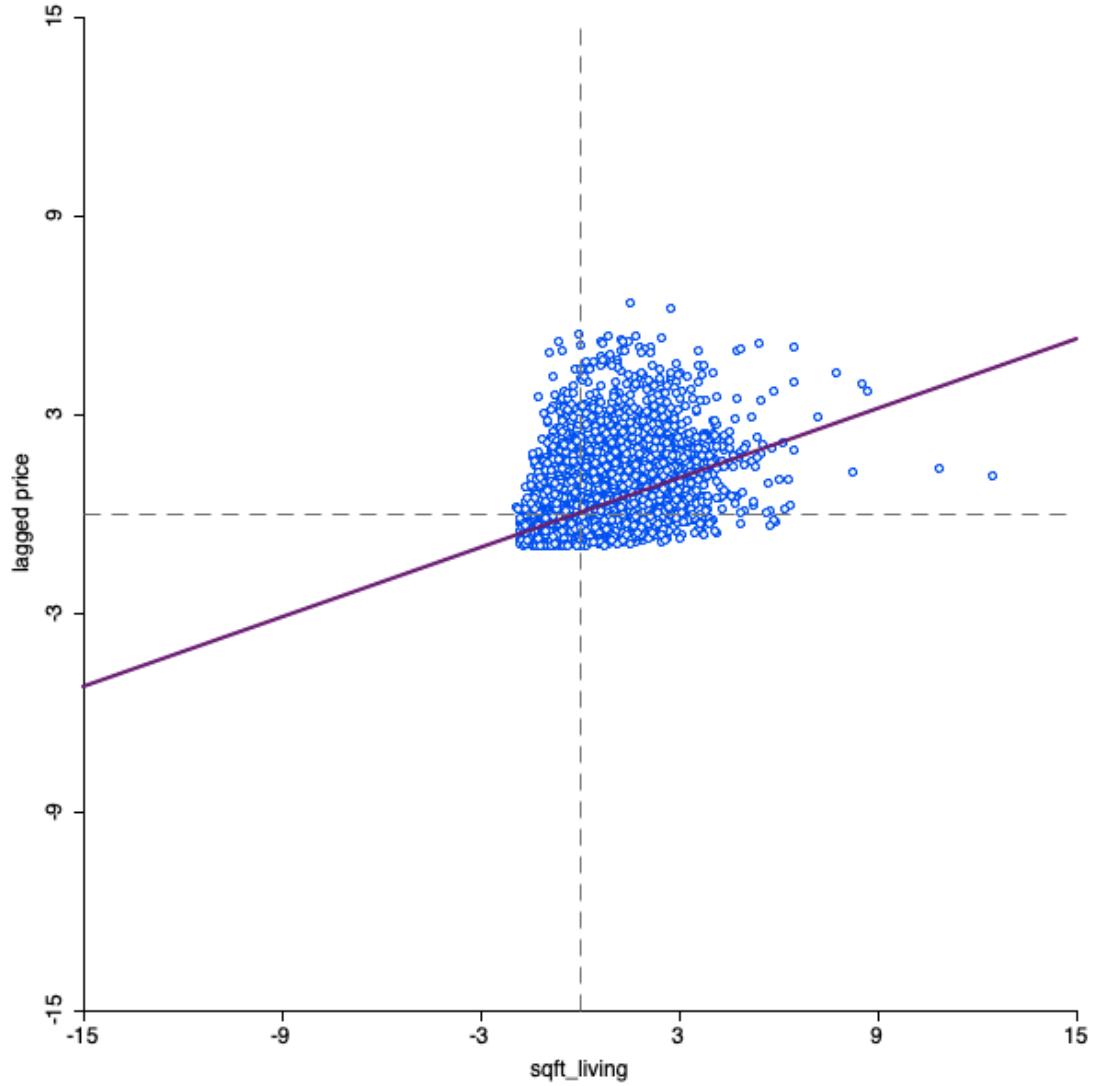


Figure 15: K10 - Bivariate Moran I

Verdier som er positive taler for klynging. 1 = perfekt klyning og 0 = perfekt tilfeldighet.

```
attach(kc_houses_env_var OMIT)
```

1. Huskarakteristika

```
mod1 <- "price ~ bedrooms + bathrooms + sqft_living + sqft_lot + sqft_above + floors + gra
```

2. Huskarakteristika + distanse til cbd + tracts_var

```
mod2 <- "price ~ bedrooms + bathrooms + sqft_living + sqft_lot + sqft_above + floors + gra
```

3. Huskarakteristika + distanse til cbd + EHD

```
mod3 <- "price ~ bedrooms + bathrooms + sqft_living + sqft_lot + sqft_above + floors + gra
```

```
hedon1 <- lm(mod1, data = kc_houses_env_var OMIT)
hedon2 <- lm(mod2, data = kc_houses_env_var OMIT)
hedon3 <- lm(mod3, data = kc_houses_env_var OMIT)
```

```
huxreg("Hedon1" = hedon1, "Hedon2" = hedon2, "Hedon3" = hedon3,
       error_format = "[{statistic}]",
       note = "{stars}. T statistic in brackets.")
```

(Kuminoff2010?) viser til at når en prisfunksjon skiftes over tiden, vil en modell som ignorerer tilfellet få resultater som er skjeve i estimatorer av helningen til prisfunksjonen, og også derfor estimatene av MWTP. Dette skjer fordi en standard DID-modellen kombinerer informasjon fra to hedoniske prisfunksjoner, altså beskrivelse av markedet før og etter tilfellet forklares i et estimat av MWTP (Bishop mfl. 2020).

Bishop forklarer videre at man kan møte på denne utfordringen ved å generalisere DID-modell ved å samhandle prisfunksjonsparametere med tidsperiode-dummy. Dette tillater en endring over tid på prisfunksjonsformen.

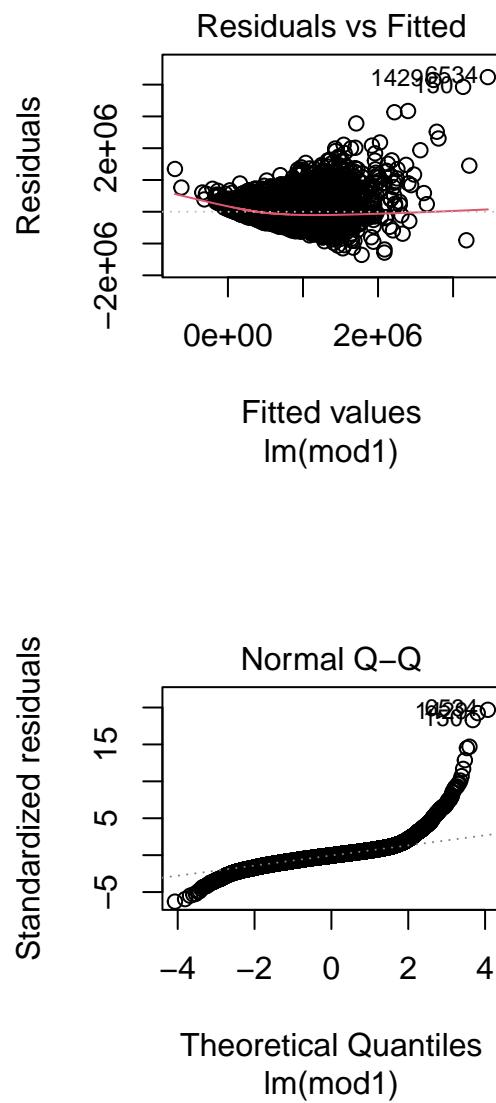
beskrivelse av huxreg

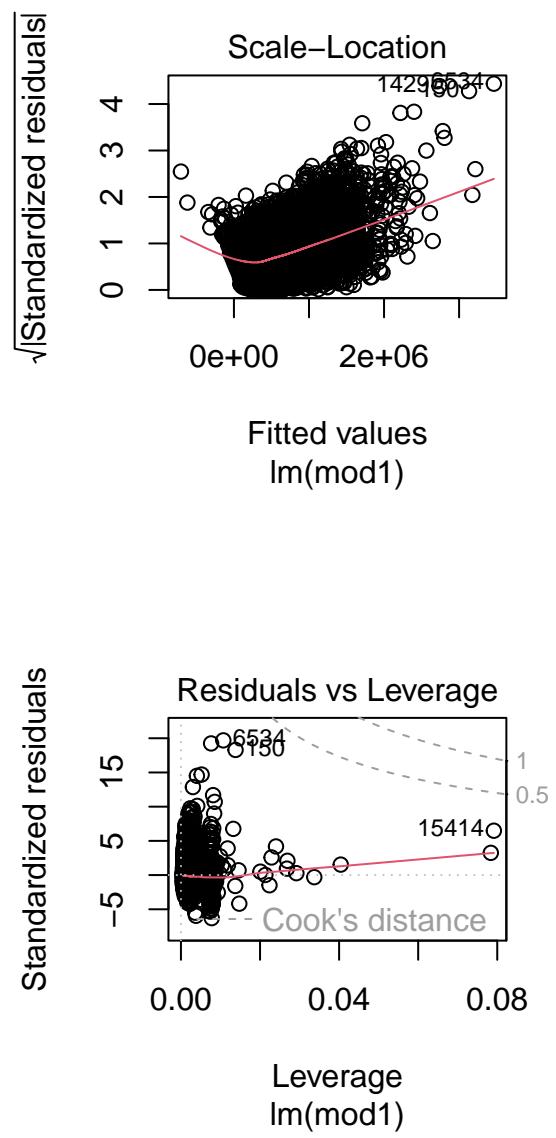
Vi kan se ut fra resultatene at forklaringskraften i mod2 er såvidt høyere enn forklaringskraften i mod3. Selvom mod2 har den sterkeste forklaringskraften så er den såpass marginal at vi velger

å gå for mod3 på grunn av færre variabler og med det lettere å arbeide med. I mod3 får vi samlet alle de miljømessige variablene i en variabel (EHD_percen).

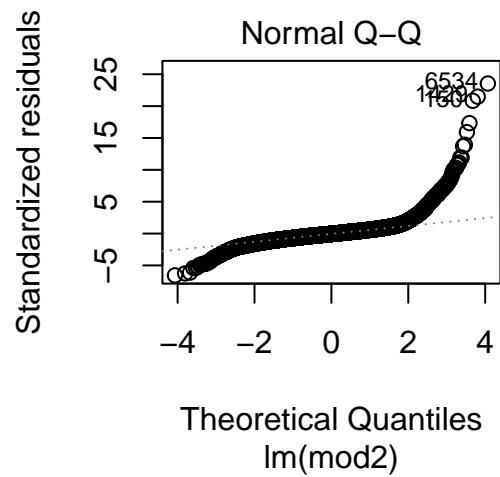
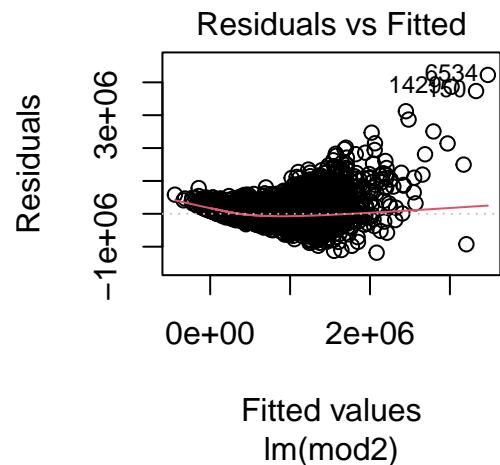
Plots

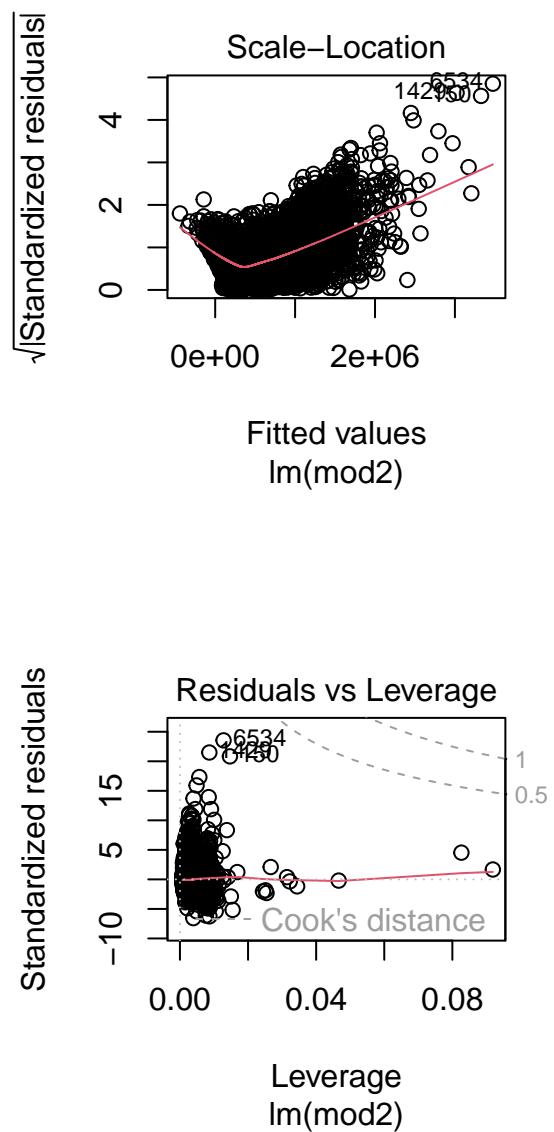
```
hedon1 %>%
  plot()
```



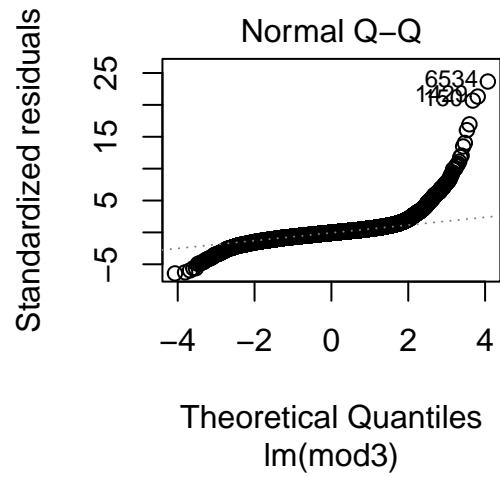
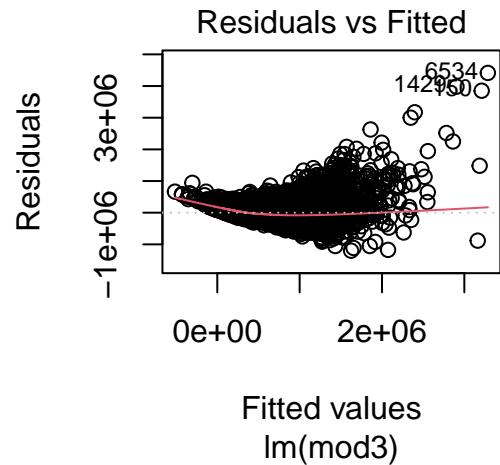


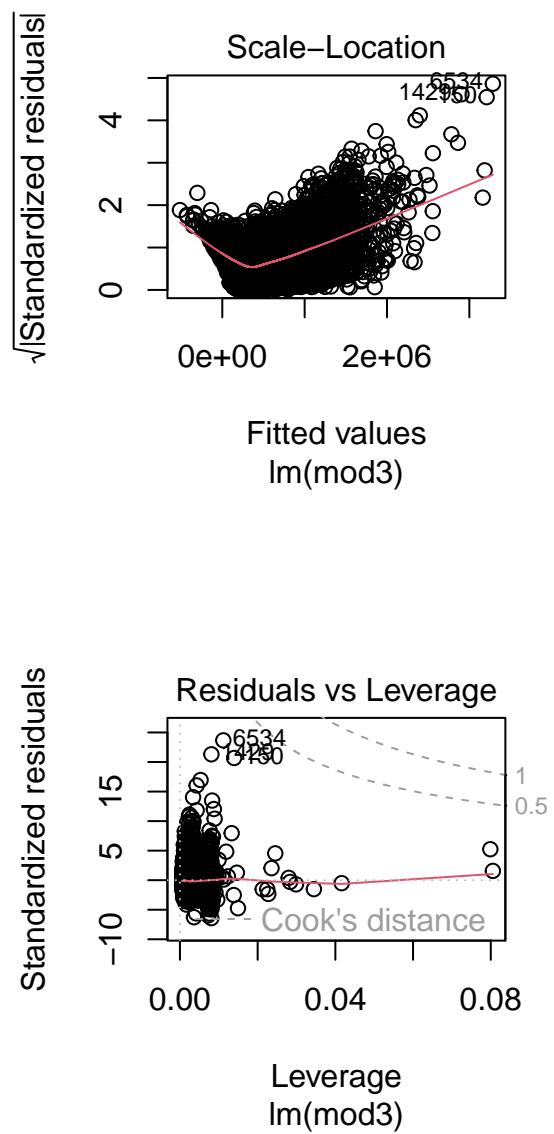
```
hedon2 %>%
  plot()
```





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





Oppgave 7

```
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",
    "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)

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)

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)

```

H_0 = Det er ikke forskjell mellom salgspris basert på salgstidspunktet.

Denne nullhypotesen kan vi forkaste på bakgrunn av signifikante F- og P-verdiene. Dette indikerer på at tidsdummyene vi bruker har en effekt, selvom de individuelt ikke er signifikante. Dette sier oss at det er forskjellige salgspriser ved forskjellige salgstidspunktet.

Oppgave 8

i.

```

kc_house_data_6666 <- here("maps/kc_house_data_6666_Sindre_og_Morten.gpkg") %>%
  st_read() %>%
  st_transform(2926)

```

Reading layer `kc_house_data_6666_Sindre_og_Morten' from data source

```

`/Users/Morten/Documents/Boligmarked/Termpaper_msbt_205_H22_S-M/maps/kc_house_data_6666_Sinc
using driver `GPKG'
Simple feature collection with 1887 features and 51 fields
Geometry type: POINT
Dimension:      XY
Bounding box:  xmin: 1226414 ymin: 72921.15 xmax: 1495965 ymax: 286273.8
Projected CRS: NAD83(HARN) / Washington North (ftUS)

```

```

kc_house_data_6666 <- kc_house_data_6666 %>%
  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)
  )

kc_house_data_6666 <- kc_house_data_6666 %>%
  rename(low = inc_fam_low_per,
         mid = inc_fam_med_per,
         high = inc_fam_high_per)

```

ii.

```

hedon3_seed <- lm(mod3, data = kc_house_data_6666)

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

kc_house_data_6666_mat_nb <- knearneigh(kc_house_data_6666, k = 3)
kc_house_data_6666_nb <- knn2nb(kc_house_data_6666_mat_nb)
kc_house_data_6666_W <- nb2listw(kc_house_data_6666_nb, style = "W")

kc_house_data_6666_mat_nb10 <- knearneigh(kc_house_data_6666, k = 10)
kc_house_data_6666_nb10 <- knn2nb(kc_house_data_6666_mat_nb10)
kc_house_data_6666_W10 <- nb2listw(kc_house_data_6666_nb10, style = "W")

lm.morantest(hedon3_seed, kc_house_data_6666_W)

```

```
Global Moran I for regression residuals
```

```
data:  
model: lm(formula = mod3, data = kc_house_data_6666)  
weights: kc_house_data_6666_W  
  
Moran I statistic standard deviate = 14.429, p-value < 2.2e-16  
alternative hypothesis: greater  
sample estimates:  
Observed Moran I      Expectation      Variance  
0.2493169179     -0.0031956991    0.0003062782
```

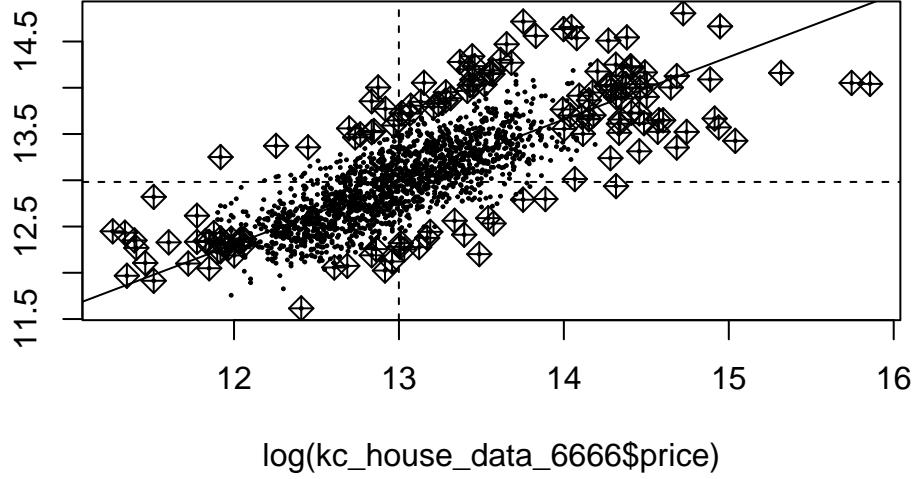
```
lm.morantest(hedon3_seed, kc_house_data_6666_W10)
```

```
Global Moran I for regression residuals
```

```
data:  
model: lm(formula = mod3, data = kc_house_data_6666)  
weights: kc_house_data_6666_W10  
  
Moran I statistic standard deviate = 23.341, p-value < 2.2e-16  
alternative hypothesis: greater  
sample estimates:  
Observed Moran I      Expectation      Variance  
2.224336e-01     -2.578743e-03    9.293224e-05
```

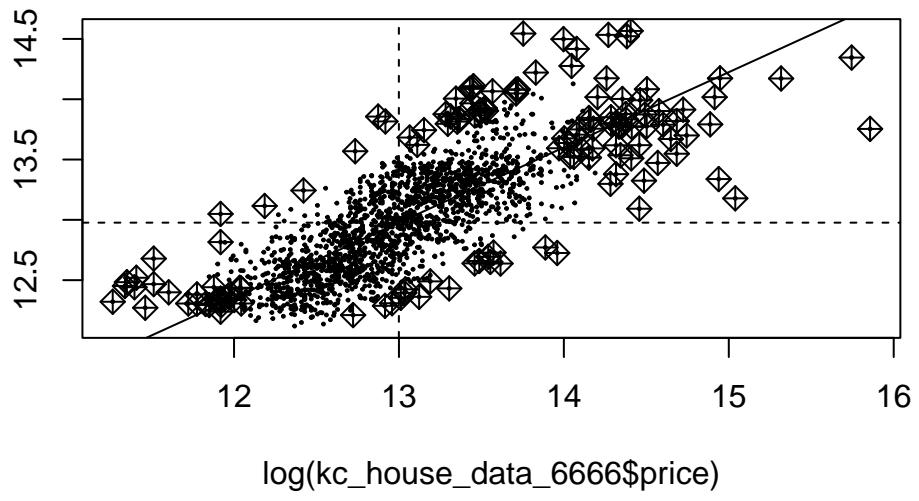
```
moran.plot(log(kc_house_data_6666$price), listw = kc_house_data_6666_W, labels = FALSE, po
```

spatially lagged log(kc_house_data_6666\$price)



```
moran.plot(log(kc_house_data_6666$price), listw = kc_house_data_6666_W10, labels = FALSE,
```

spatially lagged log(kc_house_data_6666\$price)



Ut i fra Global Morans I og plottene kan vi se at p-verdiene er signifikante som betyr vi kan forkaste $H_0 = \text{Ingen romlige effekt i residualene}$. Som betyr at vi har uforklarte spatial effects i residualene.

Vi kan se i plottene at de indikerer det samme. Vi ser at linjen har et positivt stigningstall. Dersom det ikke hadde vært noen effekt ville denne linjen vært mer vannrett ved den stippled linjen.

iv.

```
kc_lagrange_3 <- lm.LMtests(hedon3_seed, kc_house_data_6666_W,
                             test = "all")
kc_lagrange_3
```

Lagrange multiplier diagnostics for spatial dependence

```
data:
model: lm(formula = mod3, data = kc_house_data_6666)
weights: kc_house_data_6666_W

LMerr = 201.35, df = 1, p-value < 2.2e-16
```

Lagrange multiplier diagnostics for spatial dependence

```
data:
model: lm(formula = mod3, data = kc_house_data_6666)
weights: kc_house_data_6666_W

LMlag = 158.33, df = 1, p-value < 2.2e-16
```

Lagrange multiplier diagnostics for spatial dependence

```
data:
model: lm(formula = mod3, data = kc_house_data_6666)
weights: kc_house_data_6666_W

RLMerr = 58.563, df = 1, p-value = 1.965e-14
```

```
Lagrange multiplier diagnostics for spatial dependence
```

```
data:  
model: lm(formula = mod3, data = kc_house_data_6666)  
weights: kc_house_data_6666_W  
  
RLMlag = 15.537, df = 1, p-value = 8.091e-05
```

```
Lagrange multiplier diagnostics for spatial dependence
```

```
data:  
model: lm(formula = mod3, data = kc_house_data_6666)  
weights: kc_house_data_6666_W  
  
SARMA = 216.89, df = 2, p-value < 2.2e-16
```

```
kc_lagrange_10 <- lm.LMtests(hedon3_seed, kc_house_data_6666_W10,  
                               test = "all")  
kc_lagrange_10
```

```
Lagrange multiplier diagnostics for spatial dependence
```

```
data:  
model: lm(formula = mod3, data = kc_house_data_6666)  
weights: kc_house_data_6666_W10  
  
LMerr = 516.64, df = 1, p-value < 2.2e-16
```

```
Lagrange multiplier diagnostics for spatial dependence
```

```
data:  
model: lm(formula = mod3, data = kc_house_data_6666)  
weights: kc_house_data_6666_W10  
  
LMlag = 384.36, df = 1, p-value < 2.2e-16
```

```
Lagrange multiplier diagnostics for spatial dependence
```

```
data:  
model: lm(formula = mod3, data = kc_house_data_6666)  
weights: kc_house_data_6666_W10  
  
RLMerr = 195.48, df = 1, p-value < 2.2e-16
```

```
Lagrange multiplier diagnostics for spatial dependence
```

```
data:  
model: lm(formula = mod3, data = kc_house_data_6666)  
weights: kc_house_data_6666_W10  
  
RLMlag = 63.197, df = 1, p-value = 1.887e-15
```

```
Lagrange multiplier diagnostics for spatial dependence
```

```
data:  
model: lm(formula = mod3, data = kc_house_data_6666)  
weights: kc_house_data_6666_W10  
  
SARMA = 579.84, df = 2, p-value < 2.2e-16
```

Ved å ta en Langrangs multiplikatortest for $K10$ og $K3$ får vi at verdiene på LMerr og LMlag er signifikante. Når det kommer til RLMerr og RLMlag så kan vi se at disse også er signifikante. For å komme frem til hvilken vi velger av RLMerr og RLMlag så vil vi se på høyeste verdi. Vi ser at RLMerr har høyest verdi på både $K10$ og $K3$, med det kan vi si at det beste valget er en robust error modell (SEM).

v.

Her står vi ovenfor et lokalt fenomen. Veldig ofte benyttes det lokale effekter når det jobbes med boligdata og boliglitteratur ettersom det gir som oftest effekter lokalt og de nærmeste rundt seg (**LeSage2014?**).

```
SDEM_seed <- errorsarlm(mod3, data = kc_house_data_6666, listw = kc_house_data_6666_W, Dur
```

```
Warning in errorsarlm(mod3, data = kc_house_data_6666, listw = kc_house_data_6666_W, : invers
```

```
reciprocal condition number = 3.55444e-22 - using numerical Hessian.
```

```
SLX_seed <- lmSLX(mod3, data = kc_house_data_6666, listw = kc_house_data_6666_W, Durbin =  
  
SEM_seed <- errorsarlm(mod3, data = kc_house_data_6666,  
listw = kc_house_data_6666_W,  
Durbin = FALSE)
```

```
Warning in errorsarlm(mod3, data = kc_house_data_6666, listw = kc_house_data_6666_W, : inver  
reciprocal condition number = 2.98752e-22 - using numerical Hessian.
```

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

Impact measures (SDEM, estimable, n):

	Direct	Indirect	Total
bedrooms	-5.499710e+04	-4.901160e+03	-5.989826e+04
bathrooms	5.285257e+04	3.787409e+04	9.072666e+04
sqft_living	1.768634e+02	2.690962e+01	2.037730e+02
sqft_lot	2.596053e-01	-1.065106e-01	1.530947e-01
sqft_above	1.519736e+02	-2.785419e+01	1.241195e+02
floors	-9.031067e+04	7.605223e+04	-1.425845e+04
grade	1.495722e+04	2.822956e+04	4.318678e+04
yr_built	-6.312618e+02	-2.492492e+03	-3.123753e+03
yr_renovated	5.825219e+00	-4.537998e+00	1.287221e+00
waterfront	6.068063e+05	2.013749e+05	8.081812e+05
condition	2.929110e+04	3.150592e+04	6.079702e+04
view	6.410576e+04	-3.681377e+04	2.729200e+04
dist_cbd_km	-1.500263e+03	-7.429483e+03	-8.929746e+03
EHD_percen	3.381464e+02	-1.685720e+03	-1.347574e+03
low	3.285715e+03	2.381347e+05	2.414204e+05
high	5.472224e+04	1.967397e+05	2.514619e+05
year_month2014-06	1.398950e+04		NA 1.398950e+04
year_month2014-07	1.345523e+04		NA 1.345523e+04
year_month2014-08	6.830114e+03		NA 6.830114e+03
year_month2014-09	-9.264447e+03		NA -9.264447e+03
year_month2014-10	1.830891e+04		NA 1.830891e+04
year_month2014-11	6.708168e+03		NA 6.708168e+03
year_month2014-12	-1.113140e+04		NA -1.113140e+04
year_month2015-01	7.347229e+03		NA 7.347229e+03

year_month2015-02	-1.362163e+03	NA	-1.362163e+03
year_month2015-03	2.792870e+04	NA	2.792870e+04
year_month2015-04	4.870156e+04	NA	4.870156e+04
year_month2015-05	8.821624e+03	NA	8.821624e+03

Standard errors:

	Direct	Indirect	Total
bedrooms	6.729454e+03	1.268870e+04	1.580682e+04
bathrooms	1.123226e+04	2.174456e+04	2.736907e+04
sqft_living	1.552549e+01	2.977009e+01	3.726593e+01
sqft_lot	9.997138e-02	1.777446e-01	2.048172e-01
sqft_above	1.547574e+01	2.951845e+01	3.661721e+01
floors	1.256966e+04	2.178506e+04	2.596963e+04
grade	7.415280e+03	1.368050e+04	1.638300e+04
yr_built	2.628897e+02	4.397516e+02	5.037085e+02
yr_renovated	1.306494e+01	2.534776e+01	3.198411e+01
waterfront	7.008149e+04	1.673386e+05	2.023185e+05
condition	7.901669e+03	1.528025e+04	1.909289e+04
view	7.159577e+03	1.321062e+04	1.544535e+04
dist_cbd_km	4.798613e+03	4.979985e+03	1.010028e+03
EHD_percen	6.419941e+02	7.246967e+02	4.280279e+02
low	1.410026e+05	1.780698e+05	1.435120e+05
high	9.820909e+04	1.248818e+05	1.022919e+05
year_month2014-06	2.161892e+04	NA	2.161892e+04
year_month2014-07	2.178440e+04	NA	2.178440e+04
year_month2014-08	2.170979e+04	NA	2.170979e+04
year_month2014-09	2.215994e+04	NA	2.215994e+04
year_month2014-10	2.270132e+04	NA	2.270132e+04
year_month2014-11	2.352435e+04	NA	2.352435e+04
year_month2014-12	2.309511e+04	NA	2.309511e+04
year_month2015-01	2.662155e+04	NA	2.662155e+04
year_month2015-02	2.438627e+04	NA	2.438627e+04
year_month2015-03	2.163065e+04	NA	2.163065e+04
year_month2015-04	2.142501e+04	NA	2.142501e+04
year_month2015-05	2.991329e+04	NA	2.991329e+04

Z-values:

	Direct	Indirect	Total
bedrooms	-8.17259407	-0.3862618	-3.78939394
bathrooms	4.70542667	1.7417735	3.31493425
sqft_living	11.39180803	0.9039148	5.46807790
sqft_lot	2.59679594	-0.5992336	0.74747001
sqft_above	9.82012284	-0.9436200	3.38964798

floors	-7.18481396	3.4910273	-0.54904312
grade	2.01708067	2.0634886	2.63607280
yr_built	-2.40124161	-5.6679538	-6.20151051
yr_renovated	0.44586639	-0.1790295	0.04024564
waterfront	8.65858222	1.2033977	3.99459849
condition	3.70695155	2.0618718	3.18427474
view	8.95384811	-2.7866804	1.76700385
dist_cbd_km	-0.31264524	-1.4918686	-8.84109072
EHD_percen	0.52671262	-2.3261044	-3.14833177
low	0.02330252	1.3373115	1.68223180
high	0.55720137	1.5754063	2.45827818
year_month2014-06	0.64709529	NA	0.64709529
year_month2014-07	0.61765439	NA	0.61765439
year_month2014-08	0.31460980	NA	0.31460980
year_month2014-09	-0.41807190	NA	-0.41807190
year_month2014-10	0.80651331	NA	0.80651331
year_month2014-11	0.28515843	NA	0.28515843
year_month2014-12	-0.48198069	NA	-0.48198069
year_month2015-01	0.27598803	NA	0.27598803
year_month2015-02	-0.05585779	NA	-0.05585779
year_month2015-03	1.29116344	NA	1.29116344
year_month2015-04	2.27311699	NA	2.27311699
year_month2015-05	0.29490656	NA	0.29490656

p-values:

	Direct	Indirect	Total
bedrooms	2.2204e-16	0.69930282	0.00015102
bathrooms	2.5334e-06	0.08154809	0.00091665
sqft_living	< 2.22e-16	0.36604057	4.5494e-08
sqft_lot	0.00940978	0.54901709	0.45477990
sqft_above	< 2.22e-16	0.34536388	0.00069982
floors	6.7302e-13	0.00048117	0.58297586
grade	0.04368710	0.03906623	0.00838717
yr_built	0.01633954	1.4451e-08	5.5924e-10
yr_renovated	0.65569376	0.85791452	0.96789729
waterfront	< 2.22e-16	0.22882247	6.4804e-05
condition	0.00020977	0.03921995	0.00145117
view	< 2.22e-16	0.00532510	0.07722758
dist_cbd_km	0.75455020	0.13573360	< 2.22e-16
EHD_percen	0.59839317	0.02001298	0.00164205
low	0.98140896	0.18112098	0.09252390
high	0.57738985	0.11516269	0.01396050
year_month2014-06	0.51757028	NA	0.51757028

```

year_month2014-07 0.53680318 NA      0.53680318
year_month2014-08 0.75305794 NA      0.75305794
year_month2014-09 0.67589455 NA      0.67589455
year_month2014-10 0.41994694 NA      0.41994694
year_month2014-11 0.77552277 NA      0.77552277
year_month2014-12 0.62981967 NA      0.62981967
year_month2015-01 0.78255726 NA      0.78255726
year_month2015-02 0.95545510 NA      0.95545510
year_month2015-03 0.19664701 NA      0.19664701
year_month2015-04 0.02301913 NA      0.02301913
year_month2015-05 0.76806526 NA      0.76806526

```

```

huxreg("SEM" = SEM_seed, "OLS" = hedon3_seed,
       error_format = "[{statistic}]",
       note = "{stars}. T statistic in brackets.")

```

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

Likelihood ratio for spatial linear models

```

data:
Likelihood ratio = 95.981, df = 16, p-value = 1.952e-13
sample estimates:
Log likelihood of SDEM_seed Log likelihood of SEM_seed
-25696.93                  -25744.92

```

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

Likelihood ratio for spatial linear models

```

data:
Likelihood ratio = 157.78, df = 1, p-value < 2.2e-16
sample estimates:
Log likelihood of SDEM_seed Log likelihood of SLX_seed
-25696.93                  -25775.82

```

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

```
Likelihood Ratio diagnostics for spatial dependence
```

```
data:  
Likelihood ratio = 157.78, df = 1, p-value < 2.2e-16  
sample estimates:  
Log likelihood of spatial error model           Log likelihood of OLS fit y  
-25696.93                                     -25775.82
```

Ut fra resultatene i likelihood ratio test ser det ut til at SDEM er den beste modellen å anvende.

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

```
Spatial Hausman test (asymptotic)
```

```
data: NULL  
Hausman test = 91.271, df = 29, p-value = 2.317e-08
```

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

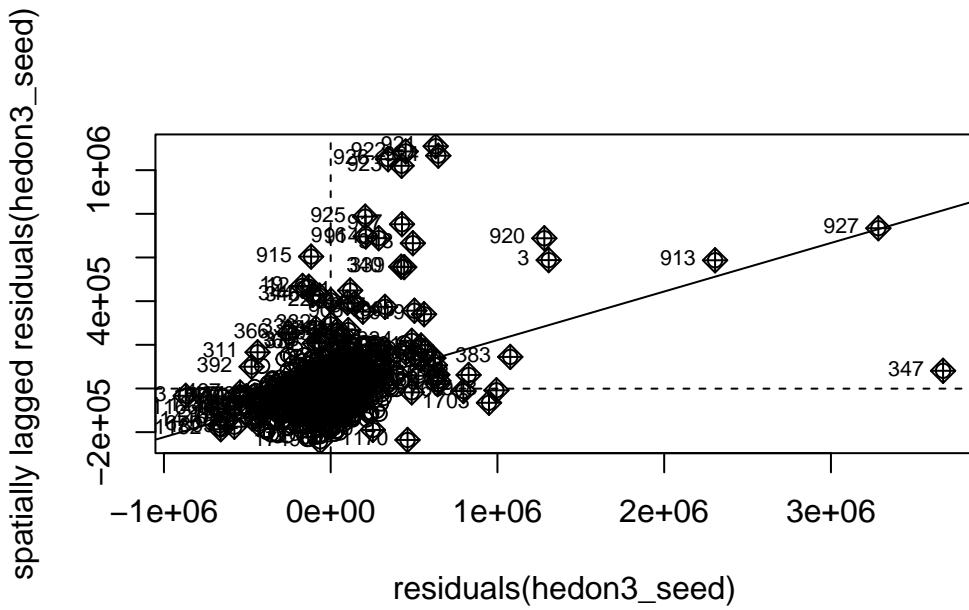
```
studentized Breusch-Pagan test
```

```
data:  
BP = 472.28, df = 28, p-value < 2.2e-16
```

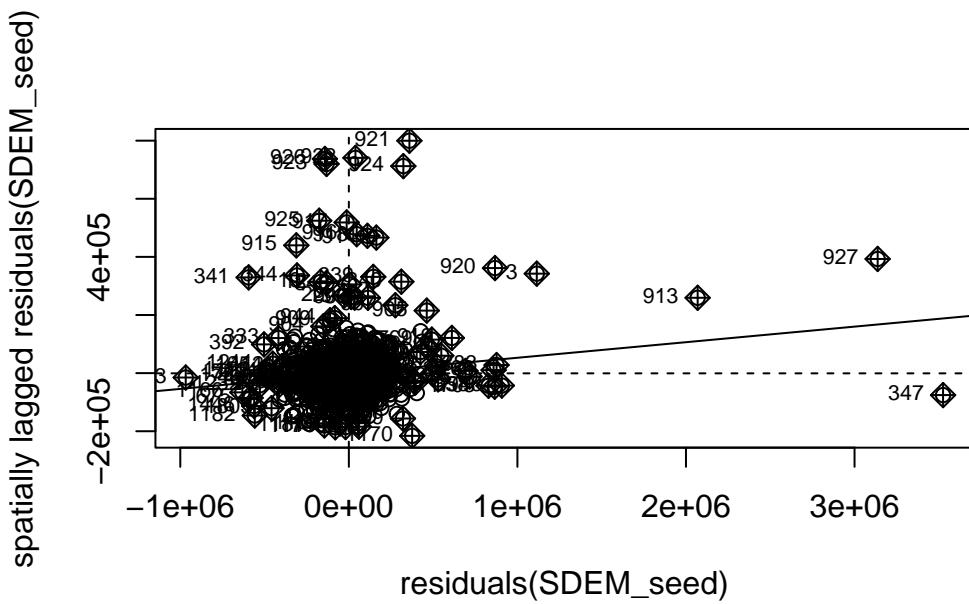
Vi kan forkaste $H_0 = \text{ingen heteroskedastisitet}$ og vi har heteroskedastisitet.

vi.

```
moran.plot(residuals(hedon3_seed), listw = kc_house_data_6666_W10)
```



```
moran.plot(residuals(SDEM_seed), listw = kc_house_data_6666_W10)
```



```
moran.test(residuals(SDEM_seed), listw = kc_house_data_6666_W10)
```

```
Moran I test under randomisation

data: residuals(SDEM_seed)
weights: kc_house_data_6666_W10

Moran I statistic standard deviate = 5.6986, p-value = 6.04e-09
alternative hypothesis: greater
sample estimates:
Moran I statistic      Expectation      Variance
5.358418e-02     -5.302227e-04    9.017628e-05
```

Vi ser at SDEM reduserer den romlige effekten i feilreddet. Likevel ser vi at Moran I ikke er signifikant.

Oppgave 9

```
set.seed(442)
kc_houses_env_var OMIT_2000 <- kc_houses_env_var OMIT[sample(1:nrow(
  kc_houses_env_var OMIT), 2000, replace = FALSE), ]

hedon3_2000 <- lm(mod3, data = kc_houses_env_var OMIT_2000)

huxreg("Full" = hedon3, "2000 Seed" = hedon3_2000, "6666 Seed" = hedon3_seed,
       error_format = "[{statistic}]",
       note = "{stars}. T statistic in brackets.")

kc_house_data_2000_mat_nb <- knearneigh(kc_houses_env_var OMIT_2000, k = 3)

Warning in knearneigh(kc_houses_env_var OMIT_2000, k = 3): knearneigh: identical
points found

Warning in knearneigh(kc_houses_env_var OMIT_2000, k = 3): knearneigh: kd_tree
not available for identical points
```

```
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")
kc_house_data_2000_mat_nb10 <- knearneigh(kc_houses_env_var OMIT_2000, k = 10)
```

```
Warning in knearneigh(kc_houses_env_var OMIT_2000, k = 10): knearneigh:
identical points found
```

```
Warning in knearneigh(kc_houses_env_var OMIT_2000, k = 10): knearneigh: kd_tree
not available for identical points
```

```
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")
```

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

```
Global Moran I for regression residuals
```

```
data:
model: lm(formula = mod3, data = kc_houses_env_var OMIT_2000)
weights: kc_house_data_2000_W

Moran I statistic standard deviate = 19.702, p-value < 2.2e-16
alternative hypothesis: greater
sample estimates:
Observed Moran I      Expectation      Variance
0.3293899740     -0.0029972881     0.0002846206
```

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

```
Global Moran I for regression residuals
```

```
data:
model: lm(formula = mod3, data = kc_houses_env_var OMIT_2000)
weights: kc_house_data_2000_W10

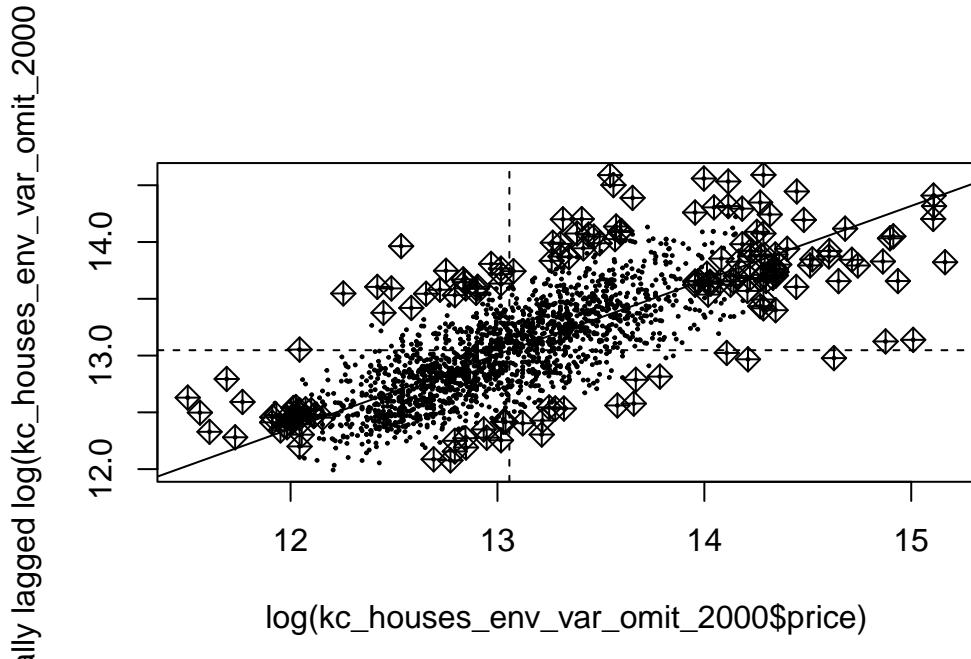
Moran I statistic standard deviate = 31.769, p-value < 2.2e-16
```

```
alternative hypothesis: greater
```

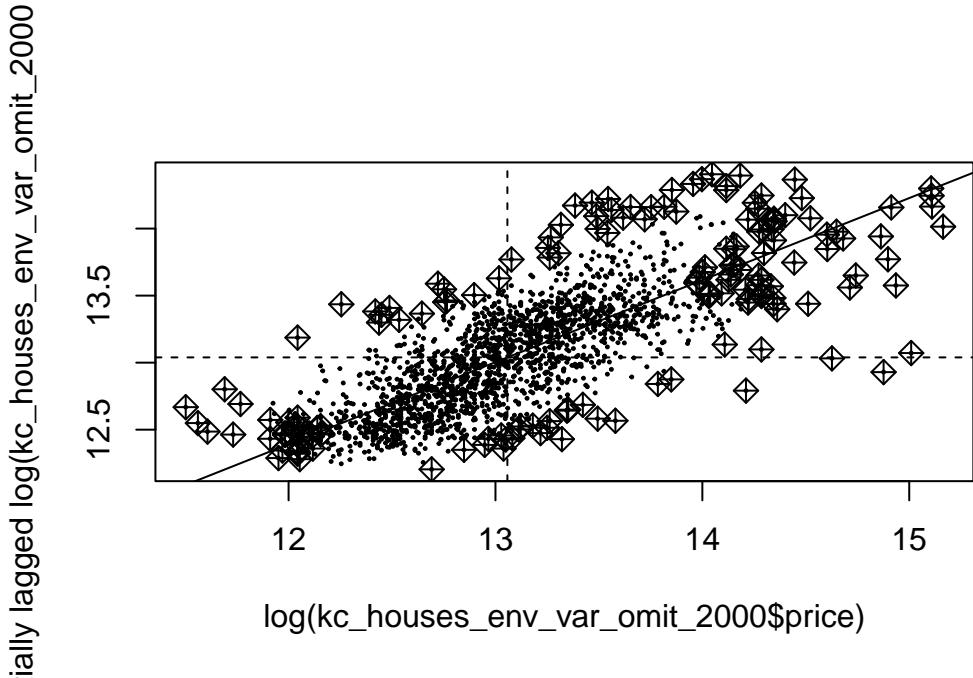
```
sample estimates:
```

Observed Moran I	Expectation	Variance
2.944919e-01	-2.498222e-03	8.739112e-05

```
moran.plot(log(kc_houses_env_var_omit_2000$price), listw = kc_house_data_2000_W, labels =
```



```
moran.plot(log(kc_houses_env_var_omit_2000$price), listw = kc_house_data_2000_W10, labels
```



```
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_houses_env_var_omit_2000)
weights: kc_house_data_2000_W

LMerr = 378.11, df = 1, p-value < 2.2e-16
```

Lagrange multiplier diagnostics for spatial dependence

```
data:
model: lm(formula = mod3, data = kc_houses_env_var_omit_2000)
weights: kc_house_data_2000_W

LMlag = 286.6, df = 1, p-value < 2.2e-16
```

```
Lagrange multiplier diagnostics for spatial dependence
```

```
data:  
model: lm(formula = mod3, data = kc_houses_env_var_omit_2000)  
weights: kc_house_data_2000_W  
  
RLMerr = 117.34, df = 1, p-value < 2.2e-16
```

```
Lagrange multiplier diagnostics for spatial dependence
```

```
data:  
model: lm(formula = mod3, data = kc_houses_env_var_omit_2000)  
weights: kc_house_data_2000_W  
  
RLMlag = 25.831, df = 1, p-value = 3.727e-07
```

```
Lagrange multiplier diagnostics for spatial dependence
```

```
data:  
model: lm(formula = mod3, data = kc_houses_env_var_omit_2000)  
weights: kc_house_data_2000_W  
  
SARMA = 403.94, 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_houses_env_var_omit_2000)  
weights: kc_house_data_2000_W10  
  
LMerr = 962.92, df = 1, p-value < 2.2e-16
```

Lagrange multiplier diagnostics for spatial dependence

```
data:  
model: lm(formula = mod3, data = kc_houses_env_var_omit_2000)  
weights: kc_house_data_2000_W10  
  
LMlag = 637.61, df = 1, p-value < 2.2e-16
```

Lagrange multiplier diagnostics for spatial dependence

```
data:  
model: lm(formula = mod3, data = kc_houses_env_var_omit_2000)  
weights: kc_house_data_2000_W10  
  
RIMerr = 414.65  df = 1  p-value < 2.2e-16
```

Lagrange multiplier diagnostics for spatial dependence

```
data:  
model: lm(formula = mod3, data = kc_houses_env_var_omit_2000)  
weights: kc_house_data_2000_W10  
  
RIMlag = 89.343  df = 1  p-value < 2.2e-16
```

Lagrange multiplier diagnostics for spatial dependence

```
data:  
model: lm(formula = mod3, data = kc_houses_env_var_omit_2000)  
weights: kc_house_data_2000_W10  
  
SABMA = 1052.3 df = 2 p-value < 2.2e-16
```

```
SDEM_2000 <- errorsarlm(mod3, data = kc_houses_env_var_omit_2000, listw = kc_house_data_2000,
                         Durbin = ss_formula(~bedrooms + bathrooms + sqft_living + sqft_living15 + sqft_above + sqft_basement + yearbuilt + yearremodadd + lat + long))
```

Warning in errorsarlm(mod3, data = kc_houses_env_var_omit_2000, listw = kc_house_data_2000_W:
reciprocal condition number = 6.1448e-22 - using numerical Hessian

```

SLX_2000 <- lmSLX(mod3, data = kc_houses_env_var OMIT_2000, listw = kc_house_data_2000_W,
SEM_2000 <- errorsarlm(mod3, data = kc_houses_env_var OMIT_2000, listw = kc_house_data_2000_W)

Warning in errorsarlm(mod3, data = kc_houses_env_var OMIT_2000, listw = kc_house_data_2000_W :
  reciprocal condition number = 5.33723e-22 - using numerical Hessian.

summary(impacts(SDEM_2000), zstats = TRUE)

Impact measures (SDEM, estimable, n):
          Direct      Indirect       Total
bedrooms     -3.075205e+04 -2.247129e+04 -5.322334e+04
bathrooms    3.423368e+04 -1.631009e+03  3.260267e+04
sqft_living   1.253952e+02  6.287994e+01  1.882752e+02
sqft_lot      4.806154e-02 -2.349830e-01 -1.869214e-01
sqft_above     7.523532e+01 -2.925434e+01  4.598098e+01
floors        -5.561932e+04  1.368998e+04 -4.192934e+04
grade          6.770711e+04  1.292796e+04  8.063507e+04
yr_built       -4.587992e+02 -4.960214e+02 -9.548207e+02
yr_renovated   3.521763e+01  4.774245e+01  8.296009e+01
waterfront     6.853044e+05 -1.543870e+05  5.309174e+05
condition      4.022586e+04  1.853773e+04  5.876358e+04
view           4.654533e+04 -6.562805e+03  3.998253e+04
dist_cbd_km    1.105390e+04 -2.022838e+04 -9.174482e+03
EHD_percen     -8.500003e+02  3.434215e+02 -5.065788e+02
low             1.264530e+05  2.114531e+05  3.379061e+05
high            6.278235e+04  3.924744e+05  4.552568e+05
year_month2014-06 1.931245e+04          NA  1.931245e+04
year_month2014-07 9.751026e+03          NA  9.751026e+03
year_month2014-08 1.329819e+04          NA  1.329819e+04
year_month2014-09 1.803634e+04          NA  1.803634e+04
year_month2014-10 3.838567e+04          NA  3.838567e+04
year_month2014-11 1.900585e+03          NA  1.900585e+03
year_month2014-12 1.728078e+04          NA  1.728078e+04
year_month2015-01 1.012296e+04          NA  1.012296e+04
year_month2015-02 4.369173e+04          NA  4.369173e+04
year_month2015-03 4.603373e+04          NA  4.603373e+04
year_month2015-04 4.546592e+04          NA  4.546592e+04
year_month2015-05 8.107049e+04          NA  8.107049e+04
=====
```

Standard errors:

	Direct	Indirect	Total
bedrooms	5.563631e+03	1.144840e+04	1.463721e+04
bathrooms	9.002426e+03	1.803797e+04	2.323683e+04
sqft_living	1.251559e+01	2.619389e+01	3.313395e+01
sqft_lot	1.005473e-01	2.101829e-01	2.497489e-01
sqft_above	1.226682e+01	2.480388e+01	3.095845e+01
floors	1.005275e+04	1.895356e+04	2.342522e+04
grade	5.946730e+03	1.129971e+04	1.400539e+04
yr_built	2.172062e+02	3.918882e+02	4.771913e+02
yr_renovated	1.048488e+01	2.203532e+01	2.802190e+01
waterfront	5.603553e+04	1.105735e+05	1.381409e+05
condition	6.551837e+03	1.304760e+04	1.653907e+04
view	5.954891e+03	1.155019e+04	1.395671e+04
dist_cbd_km	4.901740e+03	5.094681e+03	9.564346e+02
EHD_percen	4.873906e+02	5.801452e+02	4.107363e+02
low	1.200056e+05	1.684072e+05	1.586236e+05
high	8.406277e+04	1.104436e+05	9.504438e+04
year_month2014-06	1.636636e+04		NA 1.636636e+04
year_month2014-07	1.696122e+04		NA 1.696122e+04
year_month2014-08	1.721259e+04		NA 1.721259e+04
year_month2014-09	1.714392e+04		NA 1.714392e+04
year_month2014-10	1.728983e+04		NA 1.728983e+04
year_month2014-11	1.907930e+04		NA 1.907930e+04
year_month2014-12	1.828553e+04		NA 1.828553e+04
year_month2015-01	2.154223e+04		NA 2.154223e+04
year_month2015-02	1.929155e+04		NA 1.929155e+04
year_month2015-03	1.725704e+04		NA 1.725704e+04
year_month2015-04	1.657911e+04		NA 1.657911e+04
year_month2015-05	2.223002e+04		NA 2.223002e+04
=====			

Z-values:

	Direct	Indirect	Total
bedrooms	-5.52733453	-1.96283176	-3.63616684
bathrooms	3.80271642	-0.09042091	1.40306020
sqft_living	10.01911846	2.40055735	5.68224323
sqft_lot	0.47799948	-1.11799296	-0.74843747
sqft_above	6.13323569	-1.17942612	1.48524822
floors	-5.53274455	0.72229044	-1.78992272
grade	11.38560299	1.14409707	5.75743250
yr_built	-2.11227472	-1.26572182	-2.00091806
yr_renovated	3.35889751	2.16663287	2.96054469
waterfront	12.22981926	-1.39623862	3.84330269

condition	6.13963065	1.42077661	3.55301525
view	7.81632031	-0.56819866	2.86475344
dist_cbd_km	2.25509637	-3.97049017	-9.59237770
EHD_percen	-1.74398193	0.59195791	-1.23334298
low	1.05372541	1.25560636	2.13023869
high	0.74685078	3.55361898	4.78993875
year_month2014-06	1.18000880	NA	1.18000880
year_month2014-07	0.57490134	NA	0.57490134
year_month2014-08	0.77258517	NA	0.77258517
year_month2014-09	1.05205411	NA	1.05205411
year_month2014-10	2.22013019	NA	2.22013019
year_month2014-11	0.09961503	NA	0.09961503
year_month2014-12	0.94505242	NA	0.94505242
year_month2015-01	0.46991267	NA	0.46991267
year_month2015-02	2.26481197	NA	2.26481197
year_month2015-03	2.66753299	NA	2.66753299
year_month2015-04	2.74236228	NA	2.74236228
year_month2015-05	3.64689269	NA	3.64689269

p-values:

	Direct	Indirect	Total
bedrooms	3.2513e-08	0.04966573	0.00027673
bathrooms	0.00014312	0.92795274	0.16059889
sqft_living	< 2.22e-16	0.01637013	1.3294e-08
sqft_lot	0.63265057	0.26357000	0.45419633
sqft_above	8.6109e-10	0.23822854	0.13747808
floors	3.1526e-08	0.47011593	0.07346634
grade	< 2.22e-16	0.25258338	8.5403e-09
yr_built	0.03466289	0.20561269	0.04540122
yr_renovated	0.00078254	0.03026286	0.00307096
waterfront	< 2.22e-16	0.16264265	0.00012139
condition	8.2714e-10	0.15538171	0.00038084
view	5.3291e-15	0.56990008	0.00417334
dist_cbd_km	0.02412728	7.1725e-05	< 2.22e-16
EHD_percen	0.08116224	0.55387877	0.21744783
low	0.29200865	0.20925872	0.03315191
high	0.45515364	0.00037997	1.6683e-06
year_month2014-06	0.23799672	NA	0.23799672
year_month2014-07	0.56535802	NA	0.56535802
year_month2014-08	0.43976792	NA	0.43976792
year_month2014-09	0.29277473	NA	0.29277473
year_month2014-10	0.02640993	NA	0.02640993
year_month2014-11	0.92064996	NA	0.92064996

```

year_month2014-12 0.34463211 NA          0.34463211
year_month2015-01 0.63841741 NA          0.63841741
year_month2015-02 0.02352422 NA          0.02352422
year_month2015-03 0.00764104 NA          0.00764104
year_month2015-04 0.00609990 NA          0.00609990
year_month2015-05 0.00026543 NA          0.00026543

```

```

huxreg("SEM" = SEM_2000, "OLS" = hedon3_2000,
       error_format = "[{statistic}]",
       note = "{stars}. T statistic in brackets.")

```

Vi ser *SEM* har lavere verdi i AIC enn OLS, som indikerer på at vi må ta hensyn til romelig autokorrelasjon i feilreddet.

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

Likelihood ratio for spatial linear models

```

data:
Likelihood ratio = 80.083, df = 16, p-value = 1.608e-10
sample estimates:
Log likelihood of SDEM_2000 Log likelihood of SEM_2000
-26865.47                  -26905.51

```

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

Likelihood ratio for spatial linear models

```

data:
Likelihood ratio = 318.96, df = 1, p-value < 2.2e-16
sample estimates:
Log likelihood of SDEM_2000 Log likelihood of SLX_2000
-26865.47                  -27024.95

```

SDEM ser ut til å være en bedre modell enn SLX. Når vi sjekket autokorrelasjon i feilreddet så ble det enn autokorrelasjon i X-variablene.

```
LR1.Sarlm(SEM_2000)
```

Likelihood Ratio diagnostics for spatial dependence

```
data:  
Likelihood ratio = 318.96, df = 1, p-value < 2.2e-16  
sample estimates:  
Log likelihood of spatial error model           Log likelihood of OLS fit y  
-26865.47                                     -27024.95
```

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

Spatial Hausman test (asymptotic)

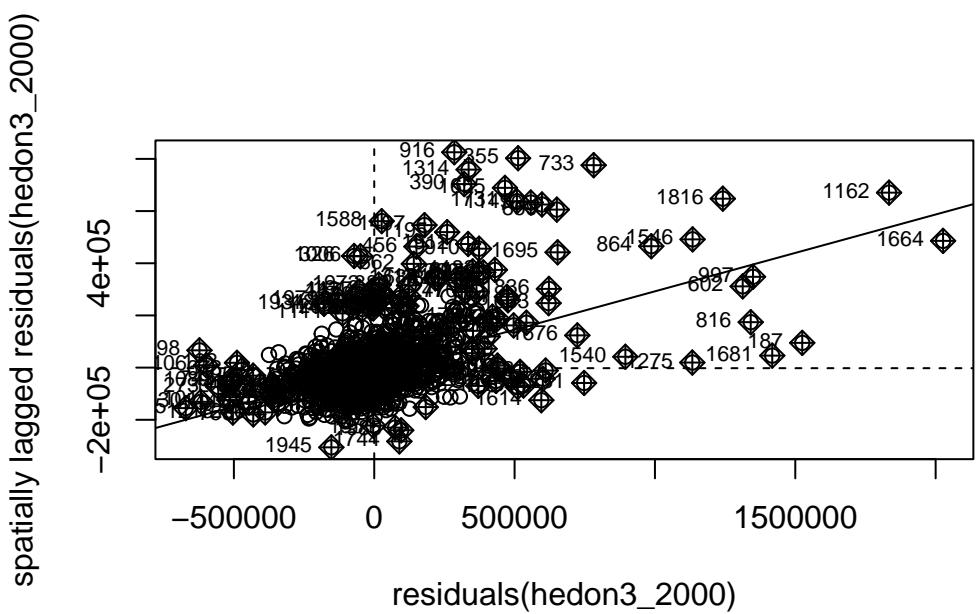
```
data: NULL  
Hausman test = 87.794, df = 29, p-value = 7.926e-08
```

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

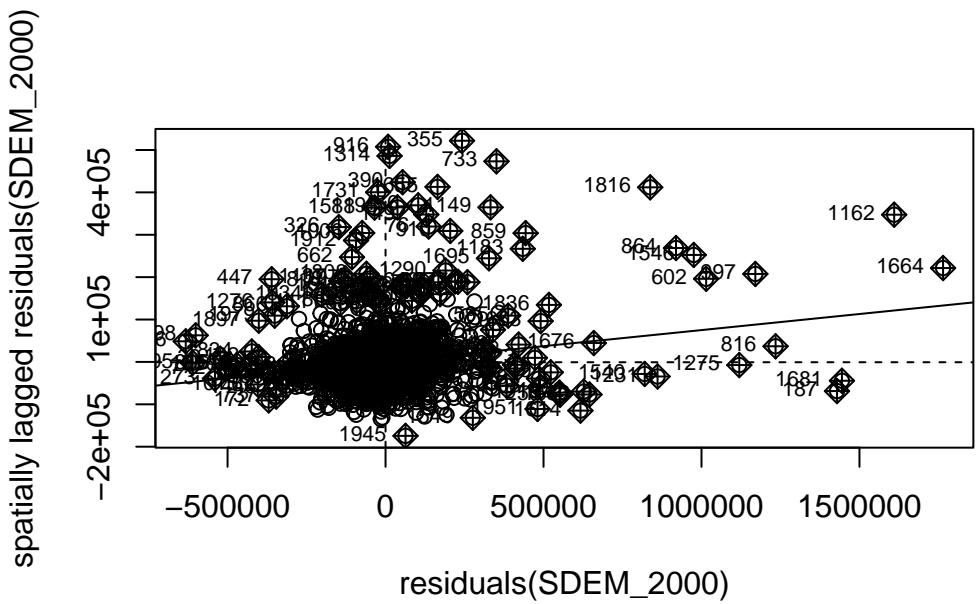
studentized Breusch-Pagan test

```
data:  
BP = 350.25, df = 28, p-value < 2.2e-16
```

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



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



Vi ser at SDEM tar noe av autokorrelasjonen i feilreddet.

```
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 = 8.1227, p-value = 2.28e-16
alternative hypothesis: greater
sample estimates:
Moran I statistic      Expectation      Variance
7.585637e-02    -5.002501e-04   8.836742e-05
```

oppsummering

Vi ser at SDEM fungerer og fjerner store deler av den romlige effekten som er i OLSEN. Vi ønsket i utgangspunktet å ta i bruk modell 2, men vi fikk ikke til beregningen der og gikk over til modell 3 på grunn av error som oppsto.

#Konklusjon

Vi ser ikke store forskjeller på resultatene på de forskjellige datasettene. SDEM er den beste modellen for alle tre datasettene.

References

- Bishop, Kelly C., Nicolai V. Kuminoff, H. Spencer Banzhaf, Kevin J. Boyle, Kathrine von Gravenitz, Jaren C. Pope, V. Kerry Smith, og 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)	6210939.368 *** [44.671]	753975.919 *** [4.710]	2061277.768 *** [15.431]
bedrooms	-39049.632 *** [-19.156]	-25328.756 *** [-14.673]	-28985.854 *** [-16.344]
bathrooms	46453.883 *** [13.238]	27307.508 *** [9.211]	32395.209 *** [10.627]
sqft_living	172.124 *** [37.305]	133.612 *** [34.160]	136.077 *** [33.757]
sqft_lot	-0.260 *** [-7.084]	0.170 *** [5.062]	0.145 *** [4.415]
sqft_above	-2.100 [-0.470]	95.399 *** [23.754]	72.713 *** [17.943]
floors	24564.665 *** [6.558]	-67284.965 *** [-19.390]	-37074.484 *** [-11.038]
grade	124528.853 *** [57.276]	68800.113 *** [35.485]	72917.129 *** [36.707]
yr_built	-3586.252 *** [-50.384]	-695.032 *** [-9.384]	-1277.559 *** [-18.400]
yr_renovated	9.073 * [2.310]	27.281 *** [8.189]	25.515 *** [7.470]
waterfront	574412.210 *** [30.786]	609996.992 *** [38.808]	605095.325 *** [37.367]
condition	19563.382 *** [7.763]	30535.045 *** [14.203]	31399.330 *** [14.269]
view	44765.891 *** [19.781]	49887.512 *** [26.008]	46864.744 *** [23.877]
year_month2014-06	3621.508 62 [0.520]	8905.815 [1.519]	7520.424 [1.244]
year_month2014-07	-273.639 [-0.039]	4115.231 [0.705]	3527.267 [0.586]
year_month2014-08	4935.598	10538.047	8170.200

Res.Df	RSS	Df	Sum of Sq	F	Pr(>F)
2.14e+04	1.01e+15				
2.14e+04	1e+15	12	5.32e+12	9.47	1.29e-18

Res.Df	RSS	Df	Sum of Sq	F	Pr(>F)
2.12e+04	6.96e+14				
2.12e+04	6.9e+14	12	5.84e+12	14.9	9.97e-32

Res.Df	RSS	Df	Sum of Sq	F	Pr(>F)
2.14e+04	7.59e+14				
2.14e+04	7.53e+14	12	5.97e+12	14.1	7.69e-30

	Full	seed
(Intercept)	2061277.768 *** [15.431]	2951606.497 *** [5.856]
bedrooms	-28985.854 *** [-16.344]	-56078.935 *** [-7.850]
bathrooms	32395.209 *** [10.627]	55620.735 *** [4.675]
sqft_living	136.077 *** [33.757]	180.269 *** [10.946]
sqft_lot	0.145 *** [4.415]	0.234 * [2.216]
sqft_above	72.713 *** [17.943]	135.065 *** [8.278]
floors	-37074.484 *** [-11.038]	-62111.658 *** [-4.828]
grade	72917.129 *** [36.707]	35767.297 *** [4.625]
yr_built	-1277.559 *** [-18.400]	-1605.519 *** [-6.107]
yr_renovated	25.515 *** [7.470]	4.871 [0.354]
waterfront	605095.325 *** [37.367]	535045.039 *** [7.286]
condition	31399.330 *** [14.269]	26646.594 ** [3.201]
view	46864.744 *** [23.877]	58143.104 *** [7.790]
dist_cbd_km	-9347.083 *** [-58.299]	-10371.011 *** [-16.627]
EHD_percen	-1174.710 *** [-15.393]	-1306.890 *** [-4.523]
low	165562.375 ***	172068.567

	SEM	OLS
(Intercept)	1517018.818 ** [2.978]	2951606.497 *** [5.856]
bedrooms	-56861.333 *** [-8.620]	-56078.935 *** [-7.850]
bathrooms	47921.645 *** [4.423]	55620.735 *** [4.675]
sqft_living	176.457 *** [11.696]	180.269 *** [10.946]
sqft_lot	0.278 ** [2.738]	0.234 * [2.216]
sqft_above	147.353 *** [9.745]	135.065 *** [8.278]
floors	-90036.310 *** [-7.151]	-62111.658 *** [-4.828]
grade	21291.047 ** [2.875]	35767.297 *** [4.625]
yr_built	-774.223 ** [-2.919]	-1605.519 *** [-6.107]
yr_renovated	2.695 [0.215]	4.871 [0.354]
waterfront	560430.090 *** [8.418]	535045.039 *** [7.286]
condition	20057.266 ** [2.622]	26646.594 ** [3.201]
view	69681.781 *** [9.622]	58143.104 *** [7.790]
dist_cbd_km	-1155 ₆₅ 471 *** [-14.198]	-10371.011 *** [-16.627]
EHD_percen	-1250.900 *** [-3.354]	-1306.890 *** [-4.523]
low	143735.222	172068.567

	Full	2000 Seed	6666 Seed
(Intercept)	2061277.768 *** [15.431]	1210875.723 ** [2.786]	2951606.497 *** [5.856]
bedrooms	-28985.854 *** [-16.344]	-32673.993 *** [-5.474]	-56078.935 *** [-7.850]
bathrooms	32395.209 *** [10.627]	37190.457 *** [3.854]	55620.735 *** [4.675]
sqft_living	136.077 *** [33.757]	129.767 *** [9.623]	180.269 *** [10.946]
sqft_lot	0.145 *** [4.415]	0.001 [0.005]	0.234 * [2.216]
sqft_above	72.713 *** [17.943]	69.394 *** [5.267]	135.065 *** [8.278]
floors	-37074.484 *** [-11.038]	-50159.047 *** [-4.702]	-62111.658 *** [-4.828]
grade	72917.129 *** [36.707]	77406.714 *** [12.162]	35767.297 *** [4.625]
yr_built	-1277.559 *** [-18.400]	-917.749 *** [-4.033]	-1605.519 *** [-6.107]
yr_renovated	25.515 *** [7.470]	32.312 ** [2.866]	4.871 [0.354]
waterfront	605095.325 *** [37.367]	651420.645 *** [10.743]	535045.039 *** [7.286]
condition	31399.330 *** [14.269]	41317.349 *** [5.855]	26646.594 ** [3.201]
view	46864.744 *** [23.877]	47129.031 *** [7.405]	58143.104 *** [7.790]
dist_cbd_km	-9347.083 *** 66 [-58.299]	-9361.880 *** [-17.962]	-10371.011 *** [-16.627]
EHD_percen	-1174.710 *** [-15.393]	-866.591 *** [-3.527]	-1306.890 *** [-4.523]
low	165562.375 ***	317959.816 ***	172068.567

	SEM	OLS
(Intercept)	493503.786 [1.191]	1210875.723 ** [2.786]
bedrooms	-27143.081 *** [-5.298]	-32673.993 *** [-5.474]
bathrooms	34985.810 *** [4.223]	37190.457 *** [3.854]
sqft_living	114.684 *** [9.876]	129.767 *** [9.623]
sqft_lot	0.071 [0.710]	0.001 [0.005]
sqft_above	80.237 *** [6.876]	69.394 *** [5.267]
floors	-58151.130 *** [-5.900]	-50159.047 *** [-4.702]
grade	68554.269 *** [11.793]	77406.714 *** [12.162]
yr_built	-449.658 * [-2.081]	-917.749 *** [-4.033]
yr_renovated	23.638 * [2.448]	32.312 ** [2.866]
waterfront	704814.899 *** [13.064]	651420.645 *** [10.743]
condition	35870.875 *** [5.806]	41317.349 *** [5.855]
view	50271.932 *** [8.552]	47129.031 *** [7.405]
dist_cbd_km	-10019.088 *** ₆₇ [-13.784]	-9361.880 *** [-17.962]
EHD_percen	-1250.388 *** [-3.767]	-866.591 *** [-3.527]
low	264677.218 *	317959.816 ***