Høgskulen på Vestlandet

MSB205 - Boligmarkedsøkonomi og anvendt spatial økonometri

Termpaper - House sales in King County, WA

Karl-Gunnar Severinsen   Silje Marie Danielsen   Sofie Brynjelsen

Innleveringsfrist: xx.11.2021

Innhold

1. Review of papers

Her kommer det noe tekst

2. House Sales in King County, WA

```r
kc_house_data <- read_csv("data/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.
```

```r
attach(kc_house_data)
```

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

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

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

Koordinater Seattle: 47.60621, -122.33207

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

```
## Geometry set for 1 feature
## Geometry type: POINT
## Dimension:      XY
## Bounding box:  xmin: 1270726 ymin: 224724.9 xmax: 1270726 ymax: 224724.9
## Projected CRS: NAD83(HARN) / Washington North (ftUS)
```

```
## POINT (1270726 224724.9)
```

```r
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)
  )
```

3. Family income

```r
kc_wadoh_map <- here("maps/WADOH_Environmental_Health_Disparities_Index_Calculated_for_K
      st_read() %>%
  st_transform(2926)
```

```
## Reading layer 'WADOH_Environmental_Health_Disparities_Index_Calculated_for_King_Count
##   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
```

```r
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
  )
```

```r
class(kc_wadoh_map)
```

```
## [1] "sf"         "data.frame"
```

```r
asc_b19101_fam_inc <- read.dbf("maps/censusSHP/acs_b19101_familyincome.dbf")
attach(asc_b19101_fam_inc)


asc_b19101_fam_inc <- asc_b19101_fam_inc %>%
    mutate(low = (E19101138 + E19101139 + E19101140 + E19101141 + E19101142 + E19101143)
    mutate(mid = (E19101144 + E19101145 + E19101146 + E19101147 + E19101148 + E19101149)
    mutate(high = (E19101150 + E19101151 + E19101152 + E19101153))


asc_b19101_fam_inc <- asc_b19101_fam_inc %>%
    select(GEOIDTRT, low, mid, high) %>%
    rename(GEO_ID_TRT = GEOIDTRT)


kc_wadoh_map_2 <- left_join(
    asc_b19101_fam_inc,
    st_drop_geometry(kc_wadoh_map),
    by = "GEO_ID_TRT")


class(kc_wadoh_map_2)


## [1] "data.frame"

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


## Reading layer 'tracts10' from data source
##   'C:\Users\bruker\Documents\Skole 2020\MSB205 - Termpaper\maps\censusSHP\tracts10.sh
##   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
##   'C:\Users\bruker\Documents\Skole 2020\MSB205 - Termpaper\maps\censusSHP\tracts10_sh
##   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)
```

```r
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"
  )
```

```r
class(kc_tracts10_env_data)
```

```
## [1] "sf"         "data.frame"
```

```
class(kc_tracts10_shore_env_data)
```

```
## [1] "sf"         "data.frame"
```

```
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.: 30077   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.   :     2466424   Min.   :  7060   Length:398
##  Class :character  1st Qu.:    19330041   1st Qu.: 20586   Class :character
##  Mode  :character  Median :    33617571   Median : 29573   Mode  :character
##                    Mean   :   161615457   Mean   : 44019
##                    3rd Qu.:    56011576   3rd Qu.: 43667
##                    Max.   :15258321253   Max.   :738820
##    Shape_area            Shape_len                geometry
##  Min.   :     2466424   Min.   :  7060   POLYGON       :398
##  1st Qu.:    19330041   1st Qu.: 20586   epsg:2926     :  0
##  Median :    33617571   Median : 29573   +proj=lcc ...:  0
##  Mean   :   161615457   Mean   : 44019
##  3rd Qu.:    56011576   3rd Qu.: 43667
##  Max.   :15258321253   Max.   :738820
```

```
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 = FALS
```

```
## 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.
```

4.

```
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.: 30077   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.    :     2466424   Min.    :  7060   Length:398
## Class :character   1st Qu.:    19330041   1st Qu.: 20586    Class :character
## Mode  :character   Median :    33617571   Median : 29573    Mode  :character
##                    Mean    :   161615457   Mean    : 44019
##                    3rd Qu.:    56011576   3rd Qu.: 43667
##                    Max.    :15258321253   Max.    :738820
##
##    Shape_area            Shape_len            low                 mid
## Min.    :     2466424   Min.    :  7060   Min.    :  0.0   Min.    :    0.0
## 1st Qu.:    19330041   1st Qu.: 20586    1st Qu.: 67.0   1st Qu.: 262.2
## Median :    33617571   Median : 29573    Median :111.0   Median : 399.5
## Mean    :   161615457   Mean    : 44019   Mean    :150.7   Mean    : 427.0
## 3rd Qu.:    56011576   3rd Qu.: 43667    3rd Qu.:202.2   3rd Qu.: 559.8
## Max.    :15258321253   Max.    :738820   Max.    :756.0   Max.    :1349.0
##
##      high            EHD_percen          linguist_2          poverty_pe
## Min.    :   0.0   Min.    :  1.00   Min.    : 0.45   Min.    : 1.97
## 1st Qu.: 427.0   1st Qu.: 25.00   1st Qu.: 3.88   1st Qu.:10.53
## Median : 665.5   Median : 50.00   Median : 8.72   Median :16.75
## Mean    : 731.8   Mean    : 50.38   Mean    :10.62   Mean    :20.42
## 3rd Qu.: 948.0   3rd Qu.: 75.00   3rd Qu.:15.38   3rd Qu.:27.48
## Max.    :3135.0   Max.    :100.00   Max.    :46.76   Max.    :75.48
##                   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.   :0.0000000    Min.   :  1.00    Min.   :  1.00
##    1st Qu.:0.1612    1st Qu.:0.0000055    1st Qu.: 25.00    1st Qu.: 25.00
##    Median :0.3652    Median :0.0005300    Median : 50.00    Median : 50.00
##    Mean   :0.6046    Mean   :0.0262015    Mean   : 50.38    Mean   : 50.38
##    3rd Qu.:0.9119    3rd Qu.:0.0087000    3rd Qu.: 75.00    3rd Qu.: 75.00
##    Max.   :3.3682    Max.   :0.6400000    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  Min.   :  75000
## Class :character  1st Qu.:2014-07-22 00:00:00  1st Qu.: 324866
## Mode  :character  Median :2014-10-17 00:00:00  Median : 450000
##                   Mean   :2014-10-29 17:30:02  Mean   : 541650
##                   3rd Qu.:2015-02-18 00:00:00  3rd Qu.: 645000
##                   Max.   :2015-05-27 00:00:00  Max.   :7700000
##
##     bedrooms        bathrooms      sqft_living      sqft_lot
## Min.   : 0.000   Min.   :0.000   Min.   :  290   Min.   :    520
## 1st Qu.: 3.000   1st Qu.:1.750   1st Qu.: 1430   1st Qu.:   5040
## Median : 3.000   Median :2.250   Median : 1920   Median :   7614
## Mean   : 3.372   Mean   :2.117   Mean   : 2083   Mean   :  15136
## 3rd Qu.: 4.000   3rd Qu.:2.500   3rd Qu.: 2550   3rd Qu.:  10696
## Max.   :33.000   Max.   :8.000   Max.   :13540   Max.   :1651359
##
##     floors        waterfront          view          condition
## Min.   :1.000   Min.   :0.000000   Min.   :0.0000   Min.   :1.00
## 1st Qu.:1.000   1st Qu.:0.000000   1st Qu.:0.0000   1st Qu.:3.00
```

14

```
##   Median :1.500   Median :0.000000   Median :0.0000   Median :3.00
##   Mean   :1.496   Mean   :0.007604   Mean   :0.2351   Mean   :3.41
##   3rd Qu.:2.000   3rd Qu.:0.000000   3rd Qu.:0.0000   3rd Qu.:4.00
##   Max.   :3.500   Max.   :1.000000   Max.   :4.0000   Max.   :5.00
##
##      grade         sqft_above    sqft_basement      yr_built
##   Min.   : 1.000   Min.   : 290   Min.   :   0.0   Min.   :1900
##   1st Qu.: 7.000   1st Qu.:1200   1st Qu.:   0.0   1st Qu.:1952
##   Median : 7.000   Median :1560   Median :   0.0   Median :1975
##   Mean   : 7.662   Mean   :1791   Mean   : 291.7   Mean   :1971
##   3rd Qu.: 8.000   3rd Qu.:2220   3rd Qu.: 560.0   3rd Qu.:1997
##   Max.   :13.000   Max.   :9410   Max.   :4820.0   Max.   :2015
##
##   yr_renovated        zipcode      sqft_living15    sqft_lot15
##   Min.   :   0.00   Min.   :98001   Min.   : 399   Min.   :   651
##   1st Qu.:   0.00   1st Qu.:98033   1st Qu.:1490   1st Qu.:  5100
##   Median :   0.00   Median :98065   Median :1840   Median :  7620
##   Mean   :  84.73   Mean   :98078   Mean   :1988   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
##  Min.   :     2791880  Min.   :   8012   Length:21436
##  1st Qu.:    24853588  1st Qu.:  23500   Class :character
##  Median :    41226319  Median :  32920   Mode  :character
##  Mean   :   180883201  Mean   :  48212
##  3rd Qu.:    73083589  3rd Qu.:  47962
##  Max.   :15258321253   Max.   : 738820
##  NA's   :25            NA's   :25
##   Shape_area           Shape_len            low             mid
##  Min.   :     2791880  Min.   :   8012   Min.   :  5.0   Min.   :  58.0
##  1st Qu.:    22811341  1st Qu.:  23204   1st Qu.: 67.0   1st Qu.: 309.0
##  Median :    34452366  Median :  31185   Median :102.0   Median : 422.0
##  Mean   :   175045321  Mean   :  46861   Mean   :144.6   Mean   : 462.8
##  3rd Qu.:    66278818  3rd Qu.:  46624   3rd Qu.:191.0   3rd Qu.: 590.0
##  Max.   :15258321253   Max.   : 738820   Max.   :756.0   Max.   :1349.0
##  NA's   :25            NA's   :25        NA's   :25      NA's   :25
##      high           EHD_percen       linguist_2       poverty_pe
##  Min.   :  17.0   Min.   :  1.00   Min.   : 0.450   Min.   : 1.97
##  1st Qu.: 573.0   1st Qu.: 19.00   1st Qu.: 3.120   1st Qu.: 8.93
##  Median : 813.0   Median : 41.00   Median : 7.000   Median :13.60
##  Mean   : 927.9   Mean   : 43.64   Mean   : 9.003   Mean   :16.65
##  3rd Qu.:1200.0   3rd Qu.: 67.00   3rd Qu.:12.730   3rd Qu.:22.95
##  Max.   :3135.0   Max.   :100.00   Max.   :40.350   Max.   :75.48
##  NA's   :25       NA's   :25       NA's   :220      NA's   :25
##   POC_percen       transporta       unemploy_2       housing_pe
##  Min.   : 7.54   Min.   :12.00   Min.   : 1.000   Min.   :15.14
##  1st Qu.:21.13   1st Qu.:18.00   1st Qu.: 3.230   1st Qu.:25.64
##  Median :33.26   Median :20.00   Median : 4.310   Median :30.46
##  Mean   :35.33   Mean   :19.77   Mean   : 4.775   Mean   :31.37
```

```
##  3rd Qu.:46.34   3rd Qu.:21.00   3rd Qu.: 6.050   3rd Qu.:35.73
##  Max.   :92.70   Max.   :26.00   Max.   :13.620   Max.   :64.87
##  NA's   :25      NA's   :25      NA's   :102      NA's   :25
##    traffic_pe       diesel           ozone            PM25
##  Min.   : 0.00   Min.   : 0.14   Min.   :46.73   Min.   :3.787
##  1st Qu.: 0.00   1st Qu.: 5.60   1st Qu.:49.24   1st Qu.:5.488
##  Median : 0.10   Median :10.16   Median :49.97   Median :6.044
##  Mean   :11.52   Mean   :13.68   Mean   :51.17   Mean   :6.002
##  3rd Qu.:19.14   3rd Qu.:16.88   3rd Qu.:52.32   3rd Qu.:6.579
##  Max.   :84.98   Max.   :92.63   Max.   :62.89   Max.   :7.897
##  NA's   :25      NA's   :25      NA's   :25      NA's   :25
##    toxic_rele       hazardous_       lead_perce       superfund
##  Min.   :   823.9   Min.   :0.02303   Min.   : 0.24   Min.   :0.03454
##  1st Qu.:  4143.8   1st Qu.:0.03985   1st Qu.: 5.34   1st Qu.:0.06595
##  Median :  8827.6   Median :0.05160   Median :11.99   Median :0.11046
##  Mean   : 17251.1   Mean   :0.07409   Mean   :16.60   Mean   :0.21696
##  3rd Qu.: 17237.2   3rd Qu.:0.07891   3rd Qu.:26.48   3rd Qu.:0.23841
##  Max.   :186434.6   Max.   :0.63781   Max.   :54.68   Max.   :1.46778
##  NA's   :25         NA's   :25        NA's   :25      NA's   :25
##    facilities       wastewater       sen_pop_pe       socio_perc
##  Min.   :0.0523   Min.   :0.000000   Min.   :  1.0   Min.   :  1.00
##  1st Qu.:0.1420   1st Qu.:0.000003   1st Qu.: 25.0   1st Qu.: 20.00
##  Median :0.2680   Median :0.000290   Median : 48.0   Median : 43.00
##  Mean   :0.5248   Mean   :0.016168   Mean   : 48.1   Mean   : 44.51
##  3rd Qu.:0.7588   3rd Qu.:0.002900   3rd Qu.: 71.0   3rd Qu.: 67.00
##  Max.   :3.3682   Max.   :0.640000   Max.   :100.0   Max.   :100.00
##  NA's   :25       NA's   :25         NA's   :25      NA's   :25
```

Ved å se på dataene sammen med kartene i QGIS, så ser vi at *kc___tracts10_env_data*
har én observasjons som faller utenfor kommunegrensene.
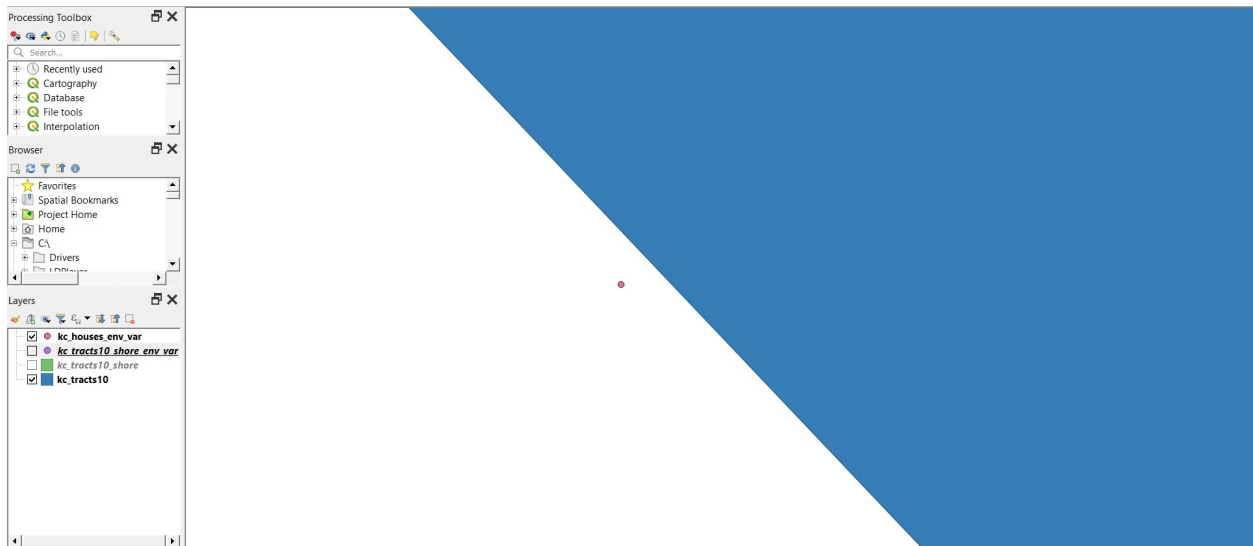
Figur 1: Observasjonen utenfor kommunegrensene.

Ved å studere *kc_tracts10_shore_env_data* ser vi at alle observasjonene ligger fortsatt innenfor kommunegrensene, men ved å legge til shore-kartet som viser vannlinjen ser vi at det er 25 observasjoner som ligger utenfor denne vannlinen, og derfor gir 25 NA.
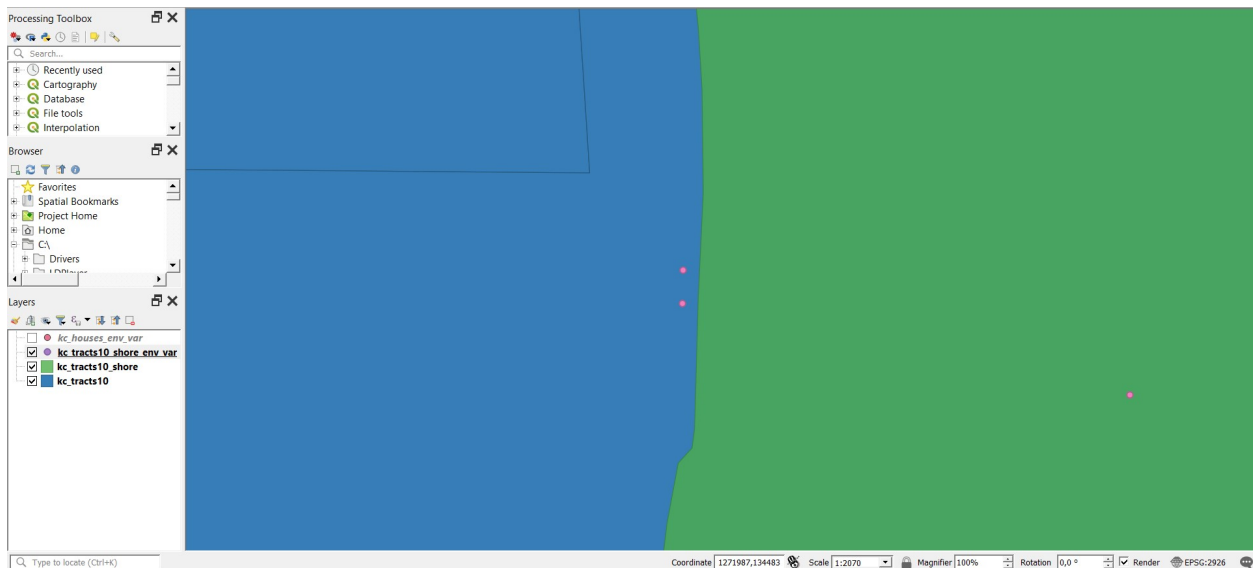
Figur 2: Observasjoner utenfor vannlinjen.

```
kc_houses_env_var <- arrange(kc_houses_env_var, desc(id))

kc_houses_env_var_omit <- kc_houses_env_var[-c(11997),]
```

```r
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.
```

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


# Alternativ måte, som ikke tar med "0" i mnd:
#     mutate(year_month = paste(year(date), month(date), sep = "-"))
```

```r
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.
```

5. Geoda-analyser

3 nærmeste naboer

Små og dyre boliger

Store og dyre boliger
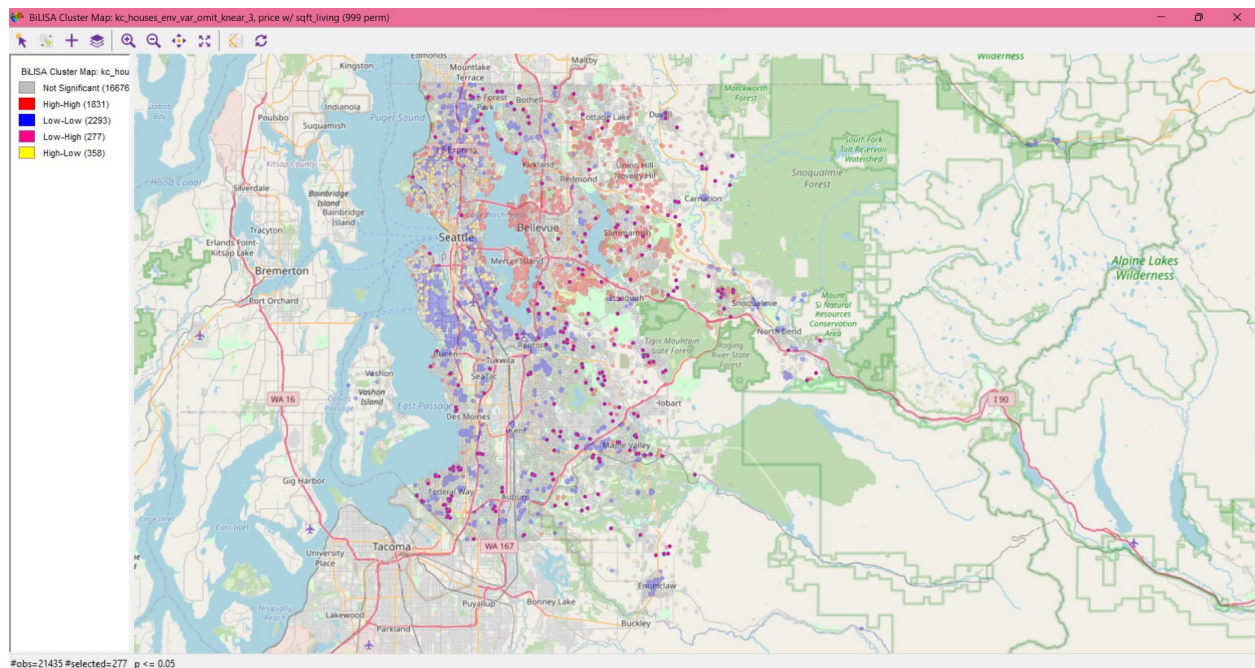
Store og billige boliger
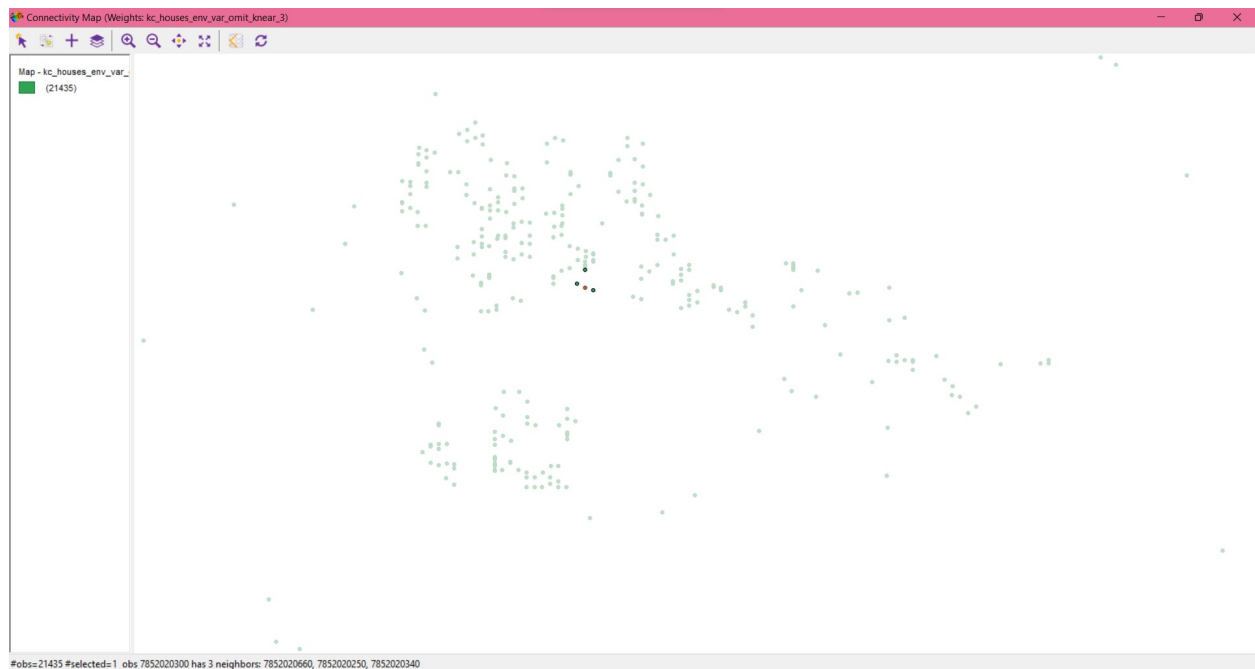
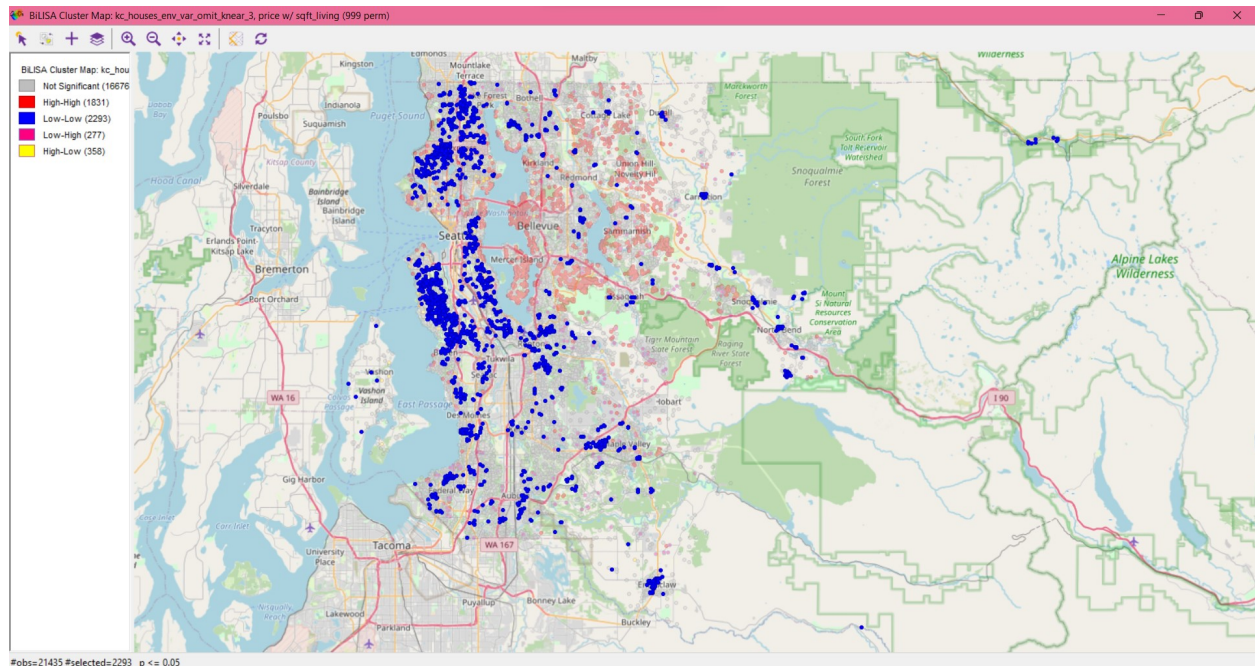Figur 3: Vektmatrise med 3 nærmeste naboer
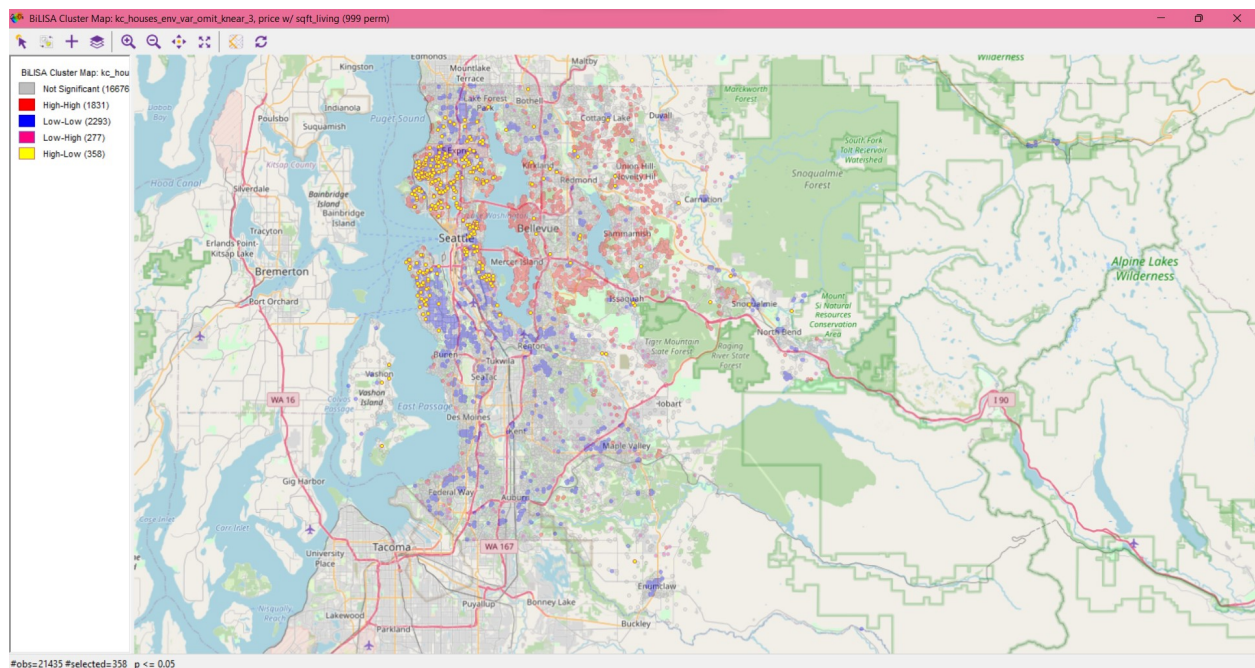
Figur 4: Små og dyre boliger markert i gule punkter.
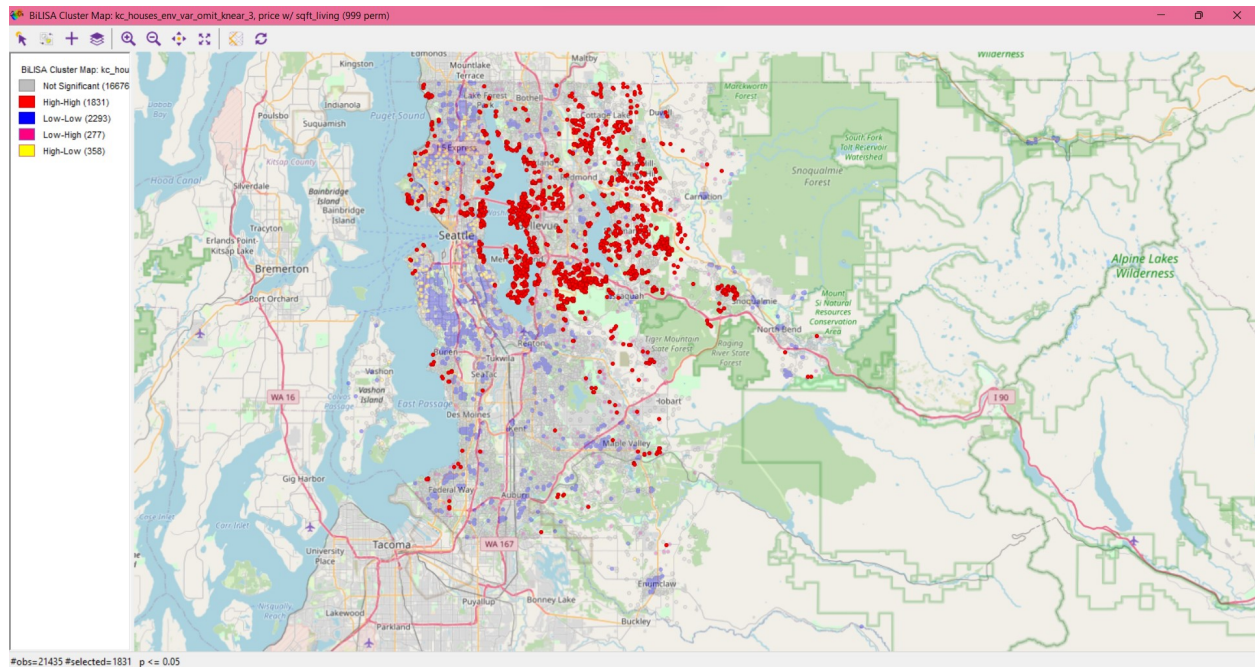
Figur 5: Store og dyre boliger markert i rød punkter.

Figur 6: Store og billige boliger markert i rosa punkter.

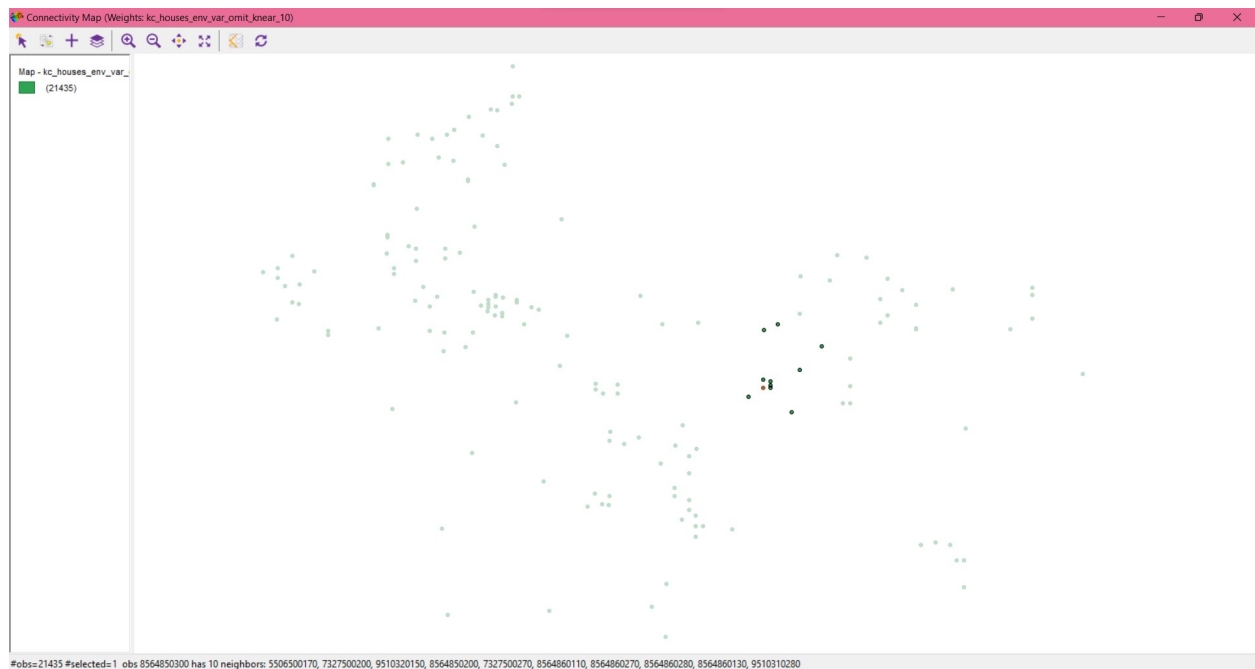Figur 7: Små og billige boliger markert i blå punkter.

Små og billige boliger

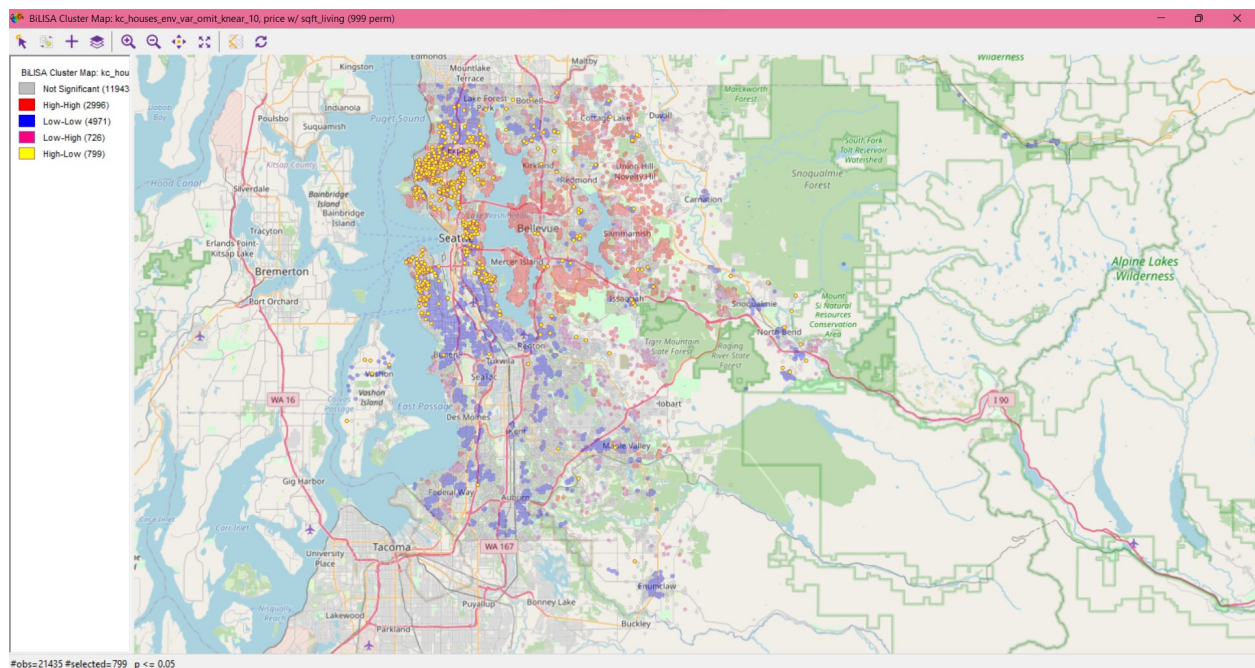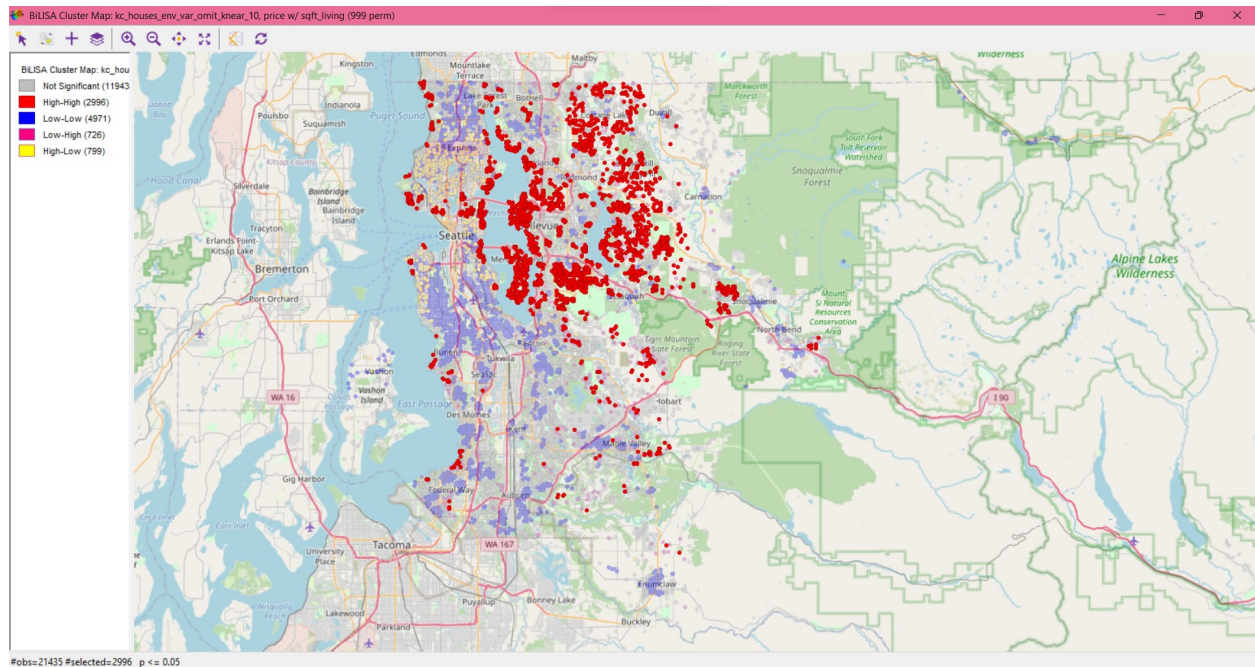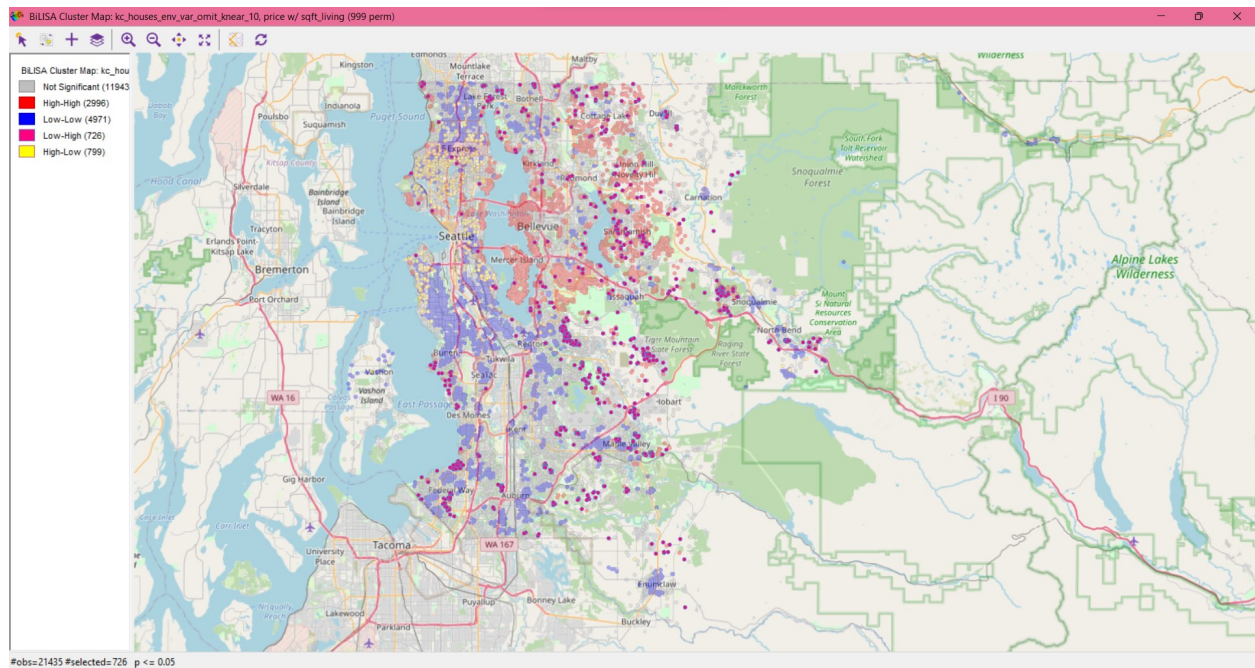## Nærmeste 10 naboer

Små og dyre boliger

Store og dyre boliger

Store og billige boliger

Små og billige boliger

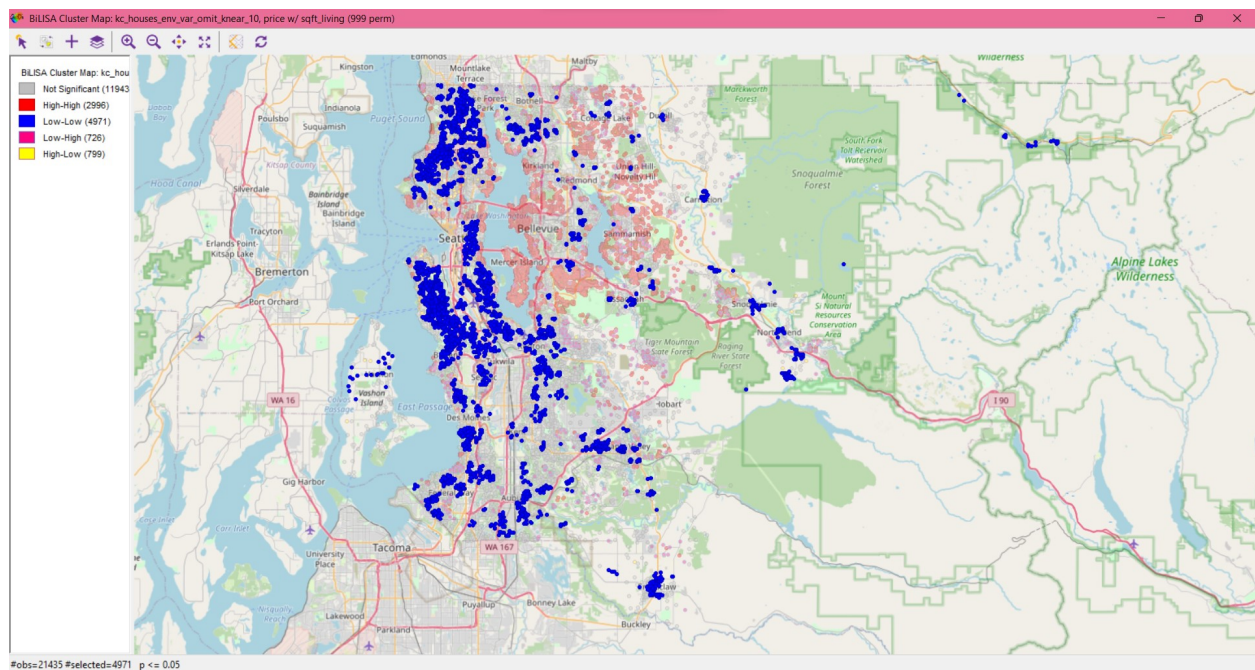Figur 8: Vektmatrise med 10 nærmeste naboer

Figur 9: Små og dyre boliger markert i gule punkter.

Figur 10: Store og dyre boliger markert i rød punkter.

Figur 11: Store og billige boliger markert i rosa punkter.

Figur 12: Små og billige boliger markert i blå punkter.