Arbeidskrav i MSB205

Thibiga Kuddyar Ingrid-Liv Morkken

Oppgave 1

Oppgave 2; House Sales in King County

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

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

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

```
dplyr::distinct(kc_house_data)
```

```
# A tibble: 21,436 x 21
   id         date                price bedro~1 bathr~2 sqft_~3 sqft_~4 floors
   <chr>      <dttm>              <dbl>   <dbl>   <dbl>   <dbl>   <dbl>  <dbl>
 1 9106000005 2015-05-27 00:00:00 1310000       4    2.25    3750    5000    2
 2 5101400871 2015-05-24 00:00:00  445500       2    1.75    1390    6670    1
 3 7923600250 2015-05-15 00:00:00  450000       5    2       1870    7344    1.5
 4 8730000270 2015-05-14 00:00:00  359000       2    2.75    1370    1140    2
 5 9178601660 2015-05-14 00:00:00 1695000       5    3       3320    5354    2
 6 1786200010 2015-05-14 00:00:00  456500       4    2.5     2580   11780    2
 7 1422700040 2015-05-14 00:00:00  183000       3    1       1170    7320    1
 8 6752600320 2015-05-14 00:00:00  360000       4    2.5     2020    7289    2
 9 4166600610 2015-05-14 00:00:00  335000       3    2       1410   44866    1
10 7129304540 2015-05-14 00:00:00  440000       5    2       1430    5600    1.5
# ... with 21,426 more rows, 13 more variables: waterfront <dbl>, view <dbl>,
#   condition <dbl>, grade <dbl>, sqft_above <dbl>, sqft_basement <dbl>,
```

```
#   yr_built <dbl>, yr_renovated <dbl>, zipcode <dbl>, lat <dbl>, long <dbl>,
#   sqft_living15 <dbl>, sqft_lot15 <dbl>, and abbreviated variable names
#   1: bedrooms, 2: bathrooms, 3: sqft_living, 4: sqft_lot
```

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

Koordinatene til Seattle hentet fra Wikipedia: 47°36'35N 122°1955V I desimaler: 47.609722, -122.333056

Kan hende ° dette vil skape problemer når det skal konventeres..

```
cbd <- st_sfc(st_point(c(-122.333056, 47.609722)), crs = 4326) %>%
  st_transform(2926)
cbd
```

```
Geometry set for 1 feature
Geometry type: POINT
Dimension:     XY
Bounding box:  xmin: 1270508 ymin: 226010.4 xmax: 1270508 ymax: 226010.4
Projected CRS: NAD83(HARN) / Washington North (ftUS)
```

```
POINT (1270508 226010.4)
```

```
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 WADOH Environmental Health Disparities Index County

```
kc_wadho_map <- here("../WADOH_Environmental_Health_Disparities_Index_Calculated_for_King_
  st_read() %>%
  st_transform(2926)
```

```
Reading layer `WADOH_Environmental_Health_Disparities_Index_Calculated_for_King_County___wado
  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_wadho_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
)
```

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

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

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

```r
kc_wadho_map_2 <- left_join(
  acs_b19101_fam_inc,
 st_drop_geometry(kc_wadho_map),
  by = "GEO_ID_TRT"
)
```

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

```
Reading layer `tracts10' from data source
  `C:\Users\thibi\Documents\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)
```

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

```
Reading layer `tracts10_shore' from data source
  `C:\Users\thibi\Documents\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)
```

```r
kc_tracts10_env_data <- left_join(
  kc_tracts10, kc_wadho_map_2,
  by = "GEO_ID_TRT"
)
```

```
kc_tracts10_shore_env_data <- left_join(
  kc_tracts10_shore,  kc_wadho_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, "../censusSHP/kc_house_data.gpkg", append=FALSE)
```

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

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

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

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

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

Spatial join