```r
library(tidyverse)
library(tmap)
library(geojsonio)
library(plotly)
library(rgdal)
library(broom)
library(mapview)
library(crosstalk)
library(sf)
library(sp)
library(spdep)
library(car)
library(fs)
library(janitor)
library(here)


LondonMSOA<-dir_info(here::here("GIS",
                                "statistical-gis-boundaries-london",
                                "ESRI"))%>%
  #$ means exact match
  dplyr::filter(str_detect(path,
                           "MSOA_2011_London_gen_MHW.shp$"))%>%
  dplyr::select(path)%>%
  pull()%>%
  #read in the file in
  st_read()
#check the data
qtm(LondonMSOA)

#read in data of MSOA area of London

LondonMSOAProfiles <- read_csv("GIS/msoa_data.csv",
                               na = c("", "NA", "n/a"),
                               locale = locale(encoding = 'Latin1'),
                               col_names = TRUE)
#check all of the columns have been read in correctly
Datatypelist <- LondonMSOAProfiles %>%
  summarise_all(class) %>%
  pivot_longer(everything(),
               names_to="All_variables",
               values_to="Variable_class")

Datatypelist
#merge boundaries and data
LonMSOAProfiles <- left_join(LondonMSOA,
            LondonMSOAProfiles,
            by = c("MSOA11CD" = "code"))


tmap_mode("view")
qtm(LonMSOAProfiles,
    fill = "average house price 2011",
    borders = NULL,
    fill.palette = "Reds")

q <- qplot(x = `unemployment rate`,
           y = `average house price 2011`,
           data=LonMSOAProfiles)
#plot with a regression line
```

```r
q + stat_smooth(method="lm", se=FALSE, size=1) +
  geom_jitter()
q <- qplot(x = `Average Household Income 2011`,
           y = `average house price 2011`,
           data=LonMSOAProfiles)
#plot with a regression line
q + stat_smooth(method="lm", se=FALSE, size=1) +
  geom_jitter()
#run the linear regression model and store its outputs in an object called
model1
Regressiondata<- LonMSOAProfiles%>%
  clean_names()%>%
  dplyr::select(average_house_price_2011,
                unemployment_rate)
Regressiondata1<- LonMSOAProfiles%>%
  clean_names()%>%
  dplyr::select(average_house_price_2011,
                average_household_income_2011)
#now model
model1 <- Regressiondata %>%
  lm(average_house_price_2011~unemployment_rate
     ,
     data=.)
model2 <- Regressiondata1 %>%
  lm(average_house_price_2011~average_household_income_2011
     ,
     data=.)
#show the summary of those outputs
summary(model1)
summary(model2)

Regressiondata2<- LonMSOAProfiles%>%
  clean_names()%>%
  dplyr::select(average_house_price_2011,
                unemployment_rate,
                average_household_income_2011)

model3 <- lm(average_house_price_2011 ~ unemployment_rate +
               average_household_income_2011, data = Regressiondata2)

#show the summary of those outputs
tidy(model3)
glance(model3)
#and for future use, write the residuals out
model_data3 <- model3 %>%
  augment(., Regressiondata2)

# also add them to the shapelayer
LonMSOAProfiles <- LonMSOAProfiles %>%
  mutate(model3resids = residuals(model3))

vif(model3)
#now plot the residuals
tmap_mode("view")


tm_shape(LonMSOAProfiles) +
  tm_polygons("model3resids",
              palette = "RdYlBu")
#calculate the centroids of all Wards in London
```

```r
coordsW <- LonMSOAProfiles%>%
  st_centroid()%>%
  st_geometry()
plot(coordsW)

LMSOA_nb <- LonMSOAProfiles %>%
  poly2nb(., queen=T)

#or nearest neighbours
knn_MSOAs <-coordsW %>%
  knearneigh(., k=4)
LMSOA_knn <- knn_MSOAs %>%
  knn2nb()

#plot them
plot(LMSOA_nb, st_geometry(coordsW), col="red")
plot(LMSOA_knn, st_geometry(coordsW), col="blue")
plot(LonMSOAProfiles)

LMSOA.knn_4_weight <- LMSOA_knn %>%
  nb2listw(., style="C")

Nearest_neighbour <- LonMSOAProfiles %>%
  st_drop_geometry()%>%
  dplyr::select(model3resids)%>%
  pull()%>%
  moran.test(., LMSOA.knn_4_weight)%>%
  tidy()
Nearest_neighbour

library(spatialreg)
sem_model1 <- errorsarlm(average_house_price_2011 ~ unemployment_rate +
                         average_household_income_2011,
                         data = Regressiondata2,
                         nb2listw(LMSOA_knn, style="C"),
                         method ="eigen")
sem_model1
```