

Thinking on Multiple Levels

Using the data from our last assignment,

1

Fit a linear model that predicts prices as having unique means in each year. Interpreting this model: 1. Which years are *not* statistically significant? why? 2. Are our predictions biased on average? How about for specific groups, such as for a given year or local authority? 3. Are its prediction errors of a consistent size over time? 4. **Challenge: is the prediction error for a LSOA in time t correlated with the prediction error in time $t + 1$?

```
library(sf)
library(tidyr)
weca = st_read('../data/weca.gpkg')

## Reading layer `weca' from data source `/Users/lw17329/Dropbox/work/teaching/gds/content/data/weca.gpkg'
## Simple feature collection with 543 features and 98 fields
## geometry type:  MULTIPOLYGON
## dimension:      XY
## bbox:           xmin: -2.718317 ymin: 51.2731 xmax: -2.252108 ymax: 51.67725
## CRS:            NA
wecatidy = tidyr::pivot_longer(weca, price_dec_1995:price_dec_2018, names_sep='_', names_to=c(NA, 'quarter'))

## Warning in val_cols[col_id] <- unname(as.list(data[cols])): number of items to
## replace is not a multiple of replacement length
```