

# The Assignment

- **DATA:** Available on Canvas under a folder named ‘Assignment’ within the Z402 Files. You have been provided House Price data from a region in the USA. Find the details in the next slide.
- **TASK:** The task is to estimate **house price** based on several explanatory variables like “lotSize”, “age”, “landValue”, etc. You may choose to use only some of the variables to explain house price if you like.
- **Your report** should be two pages (with font size at least 11) and include:
  - Summary of the Process of analysis
  - Rationale for the final model
  - Significance and some Diagnostics.
- Submission details on Canvas. Please do not submit your Excel Sheet (or Code) or an Appendix.

**NOTE:** This is an individual assignment, and you are expected to work on your own to complete the assignment.

# The Data

Data on house prices from a county in the USA. A data frame with 1728 observations on the following 16 variables.

1. **price** price (1000s of US dollars)
2. **lotSize** size of lot (square feet)
3. **age** age of house (years)
4. **landValue** value of land (1000s of US dollars)
5. **livingArea** living are (square feet)
6. **pctCollege** percent of neighbourhood that graduated college
7. **bedrooms** number of bedrooms
8. **firplaces** number of fireplaces
9. **bathrooms** number of bathrooms (half bathrooms have no shower or tub)
10. **rooms** number of rooms
11. **heating** type of heating system
12. **fuel** fuel used for heating
13. **sewer** type of sewer system
14. **waterfront** whether property includes waterfront
15. **newConstruction** whether the property is a new construction
16. **centralAir** whether the house has central air