STATS 3001 / STATS 4104 / STATS 7054
Statistical Modelling III
Practical 2 - Factors

Week 2

GOAL

This practical is intended to illustrate some of the properties of linear models involving factors and their implementation in R.

DATA

The dataset `loan.xlsx` - get it from MyUni - has the following variables.

Var	Description
`loan_amnt`	the amount of the loan in dollars
`term`	the term of the loan in months
`home_ownership`	home ownership status (rent/own/mortgage)
`annual_inc`	the annual income of the applicant in dollars

STEPS

- Read in the data
- Fit the model,

```
loan_amnt ~ home_ownership
```

- What is the reference category for `home_ownership`?
- Calculate the group means for each level of `home_ownership`. Show how these can be obtained from the `lm()` output.
- Redo the linear modelling using the zero sum constraint.
- Calculate the overall mean `loan_amnt`. How can you get this, and the group means from the new `lm()` output?
- Fit the models

```
loan_amnt ~ home_ownership + annual_inc
loan_amnt ~ home_ownership * annual_inc
```

- For each model, give the estimated regression line for each of the three groups.