

# STAT 118: Notes M

## Simple Linear Regression

Emily Malcolm-White

```
#LOAD PACKAGES
library(tidyverse)
library(scales)

#LOAD DATA
housingdata <- read.csv("https://raw.githubusercontent.com/abhiram-ds/housing_price_adv_re
```

The data comes from [Kaggle](#). It describing the sale of individual residential property in Ames, Iowa from 2006 to 2010.

## Motivation

```
housingdata %>%
  ggplot(aes(x=GrLivArea, y=SalePrice)) +
  geom_point(pch=16) +
  geom_smooth(method=lm, se=FALSE) +
  scale_y_continuous(labels = dollar_format(prefix="$")) +
  labs(x = "Above Ground Living Area (sqft)", y = "Sale Price") +
  theme_minimal()
```

`geom\_smooth()` using formula = 'y ~ x'