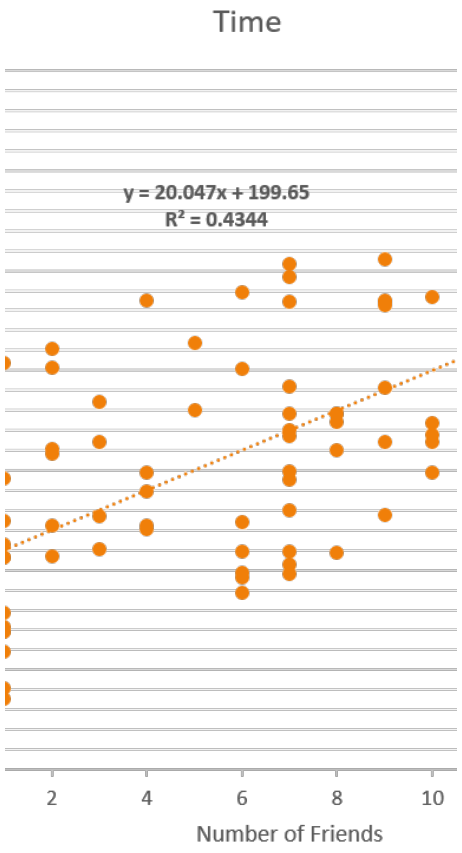
Evaluation Methods

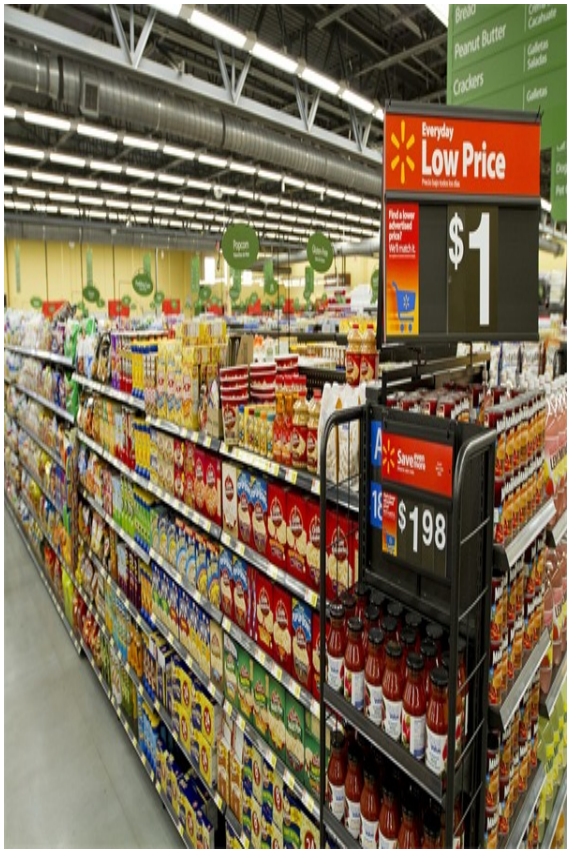
School of Information Studies
Syracuse University

Evaluation Methods



Agenda

- Intro
- R-squared
- P-values
- Model validation



R-Squared

$$R^2 = 1 - \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (y_i - \bar{y}_i)^2}$$

lm(formula = hardness ~ dens, data = hardness)

Residuals:

Min	1Q	Median	3Q	Max
-338.40	-96.98	-15.71	92.71	625.06

Coefficients:

	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	-1160.500	108.580	-10.69	2.07e-12	***
dens	57.507	2.279	25.24	< 2e-16	***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

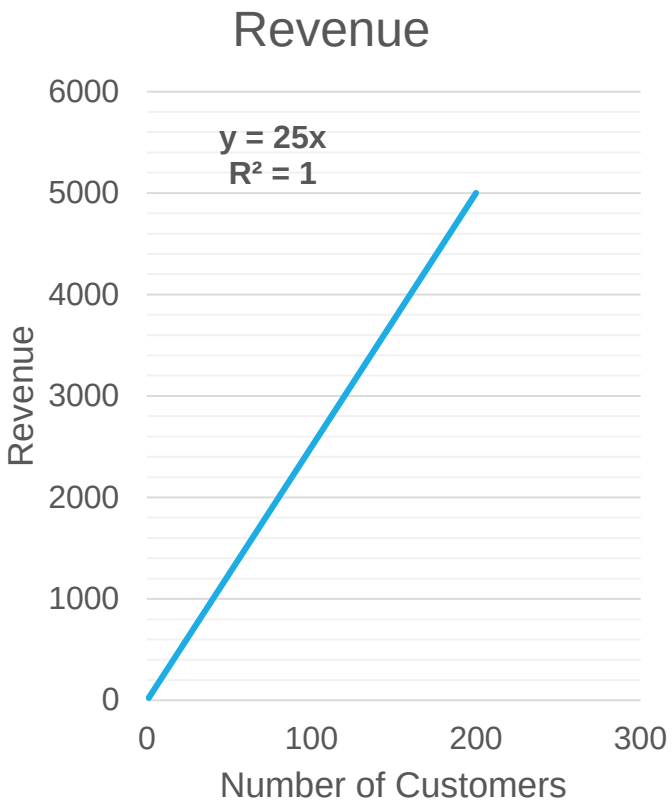
Residual standard error: 183.1 on 34 degrees of freedom

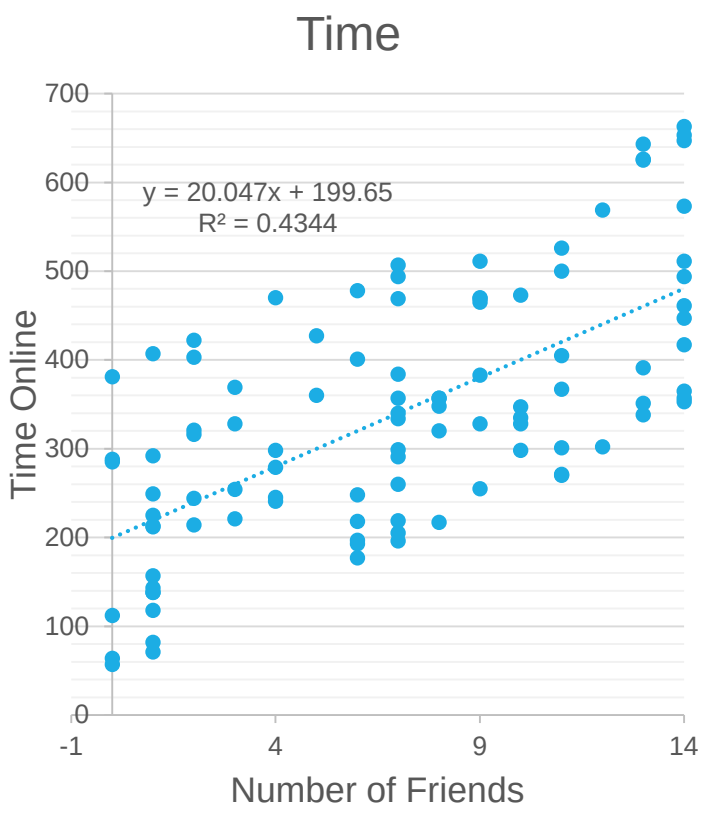
Multiple R-squared: 0.9493,

Adjusted R-squared: 0.9478

F-statistic: 637 on 1 and 34 DF, p-value: < 2.2e-16

R-Squared (cont.)





P-Values

- Low p-value?
  - Highly unlikely to occur randomly, therefore significant
- High p-value?
  - Coefficient might actually be zero, therefore consider removing from model

```
lm(formula = hardness ~ dens, data = hardness)

Residuals:
    Min      1Q  Median      3Q     Max
-338.40  -96.98  -15.71   92.71  625.06

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) -1160.500    108.580  -10.69 2.07e-12 ***
dens           57.507      2.279   25.24  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 183.1 on 34 degrees of freedom
Multiple R-squared:  0.9493,
Adjusted R-squared:  0.9478
F-statistic:   637 on 1 and 34 DF,  p-value: < 2.2e-16
```

Model Validation

- Collect new data
- Compare the results with:
  - Theoretical expectation (how much should a 0-bedroom house cost?)
  - Earlier empirical studies
  - Simulation (see GPA example from text)
- Split the original data with one portion for training and one for testing

Confusion Matrix

- True positive
- False positive
- Accuracy
- Precision
- Specificity