Simple Linear Regression Model

READING: 1.1

EXERCISES: CH.1: 1,5-8, 15, 17, 19

ASSIGNED: HW 3

PRODUCER: DR. MARIO

IMG CREDIT: ALEX RIEGERT-WATERS

Prerequisites for the Model

1. Single Quantitative Response Variable *Y*
2. Single Quantitative Predictor Variable *X*
3. Scatterplot of *Y* versus *X* (Book Does Side-by-Side Boxplots)
4. Evidence that a **Straight Line** is Reasonable for Modeling the Relationship between *Y* and *X*

Example: Honda Accords

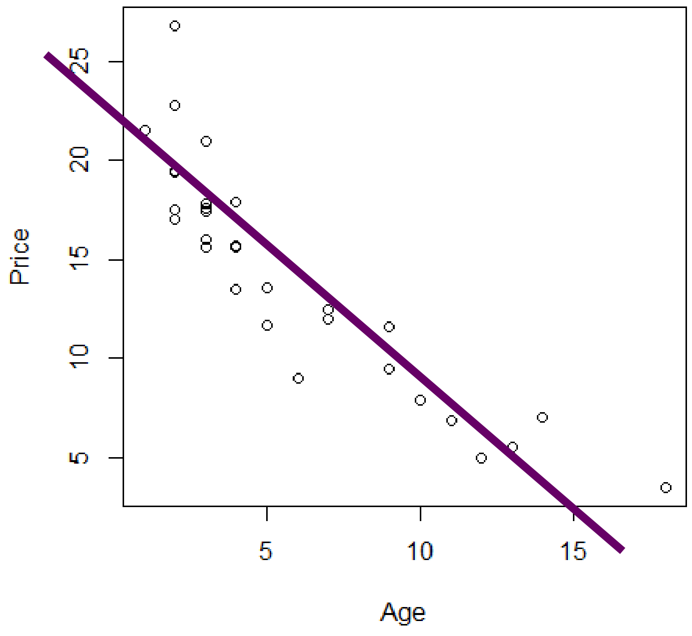
- Question: *Is there a linear relationship between the **age** of an Accord and the **price** of an Accord?*
- Both Variables are Quantitative
- Which One is Response?
- Scatter Plot of *price* **versus** *age* or *price* **on** *age* Shows Evidence of a **Linear** Relationship

Example: Honda Accord Price

```r
library(Stat2Data) #Package for Textbook
library(mosaic)

data("AccordPrice") #Puts dataset into Global Environment

plot(Price~Age,data=AccordPrice)
```



Simple Linear Regression Model

- General Form

$$Y = f(X) + \epsilon$$

$$= \boxed{\mu_Y} + \epsilon$$

Mean of Y Given X or E[Y|X]

- Simple Linear Regression

$$Y = \boxed{\beta_0 + \beta_1 X} + \epsilon$$

- Shape Depends on y-Intercept β_0 and Slope β_1

Fitting Model to Data

- Fitted (Estimated) Model

$$\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 X$$

- Residual for i^{th} Car

$$\hat{\epsilon}_i = Y_i - \hat{Y}_i = Y_i - (\hat{\beta}_0 + \hat{\beta}_1 X)$$

- Sum of Squared Errors (SSE)

$$\sum \hat{\epsilon}_i^2 = \sum (\hat{Y}_i - Y_i)^2$$

Fitting Model to Data

- Least Squares Regression: *Choose Estimates $\hat{\beta}_0$ and $\hat{\beta}_1$ such that SSE is **as small as possible***

$$\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 X$$

- Interpretation of $\hat{Y}$: ***Expected** value or **predicted** value of Y given a known value for **X***

- Interpretation of $\hat{\beta}_0$: *The predicted value of Y given that **X=0***

- Interpretation of $\hat{\beta}_1$: *The amount by which our expected value of Y would **change** if we **increase** X by **1 unit***

Centering Data

- Suppose we replace our Predictor X with $X - \bar{x}$

- Notice

$$
\begin{aligned}
Y &= \beta_0 + \beta_1(X - \bar{x}) + \epsilon \\
&= \beta_0 + \beta_1 X - \beta_1 \bar{x} + \epsilon \\
&= (\beta_0 - \beta_1 \bar{x}) + \beta_1 X + \epsilon \\
&= \beta_0^* + \beta_1 X + \epsilon \longrightarrow
\end{aligned}
$$

Slope Is Unaffected

Y-intercept Will Change

Thank You

Make Reasonable Decisions