Assessing Conditions

READING:	1.3
EXERCISES:	CH.1: 2, 10, 12, 22, 24, 26
ASSIGNED:	HW 3
PRODUCER:	DR. MARIO

Supplement for Lecture 5

- Download Zip Folder on Course Website for Supplement
- Unzip Folder and Open the Template.rmd File
- Stamp Data
  - *Price = Price of US Stamp (Cents)*
  - *Year = Year When Price of Stamp Changed*
  - Question: *Can we use a **linear regression** to understand how the **price** changes over **time**?*

Supplement for Lecture 5

- Example Based off Exercise 1.33 from Textbook
- Fit Straight Lines to Data With and Without the First Four Years
- Output From Summary Function
  - Estimated Parameters
  - Residual Standard Error
  - Degrees of Freedom

```
Call:
lm(formula = Price ~ Year, data = stamp)

Residuals:
    Min      1Q  Median      3Q     Max
-11.993  -7.001   1.788   5.447  17.273

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) -770.78108   99.39045  -7.755 7.28e-08 ***
Year           0.40080    0.05029   7.970 4.57e-08 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 8.03 on 23 degrees of freedom
Multiple R-squared:  0.7342,    Adjusted R-squared:  0.7226
F-statistic: 63.52 on 1 and 23 DF,  p-value: 4.572e-08
```

Supplement for Lecture 5

- Obtaining Fitted Values and Residuals from Models

```
mod1$residuals
mod1$fitted.values
```

- Saving Fitted Values and Residuals Into Dataset

```
stamp$fit=mod1$fitted.values
stamp$resid=mod1$residuals
```

Residual Versus Fit Plot

- Ideal:
  - Randomly Distributed Around 0 (Independence and Bias)
  - No Visible Patterns (Fit of Linear Model)
  - Constant Variance (Heteroscedasticity or Homoscedasticity)
  - Close to Zero (Good or Bad Fit)

- Observe Plots in Supplement for Lecture 5

- Remember: ***Positive Residual = Under Prediction***

Residual Versus Order

- Only Appropriate if the We Know the Chronological Order of Data
- Ideal:
  - Randomly Distributed Around 0 (Independence and Bias)
  - No Long Runs Above 0 or Below 0 (Positive Autocorrelation)
  - No Bouncing Up and Down (Negative Autocorrelation)
- Observe Plots in Supplement for Lecture 5

Histogram/Boxplot of Residuals

- Ideal:
  - Bell-shaped and Symmetric (Normal Distributed)
  - No Gaps (Few Outliers/Influential Points)
  - Centered Around 0 (Unbiased)

- Observe Plots in Supplement for Lecture 5

Normal Plots

- Ideal: Pattern is a Perfectly Straight Line (Aligns With Normal Dist.)
- Types:
  - Normal Quantile Plot (Observed Vs Expected Quantiles)
  - Normal Probability Plot (Observed Vs Cumulative Probability)
- Observe Plots in Supplement for Lecture 5

Thank You

Make Reasonable Decisions