Share A Graph

Ben Goldstone

9/27/23

Consumer Goods Data Import

Consumer Goods Dataset

```
ConsumerGood <- read_csv("~/CSVs/ConsumerGood.csv")
```

```
## Rows: 108 Columns: 4
```

```
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## dbl (4): rownames, distribution, share, price
```

```
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
head(ConsumerGood)
```

```
## # A tibble: 6 x 4
##   rownames distribution share price
##      <dbl>        <dbl> <dbl> <dbl>
## 1        1        0.905  2.74  106.
## 2        2        0.9    3.01  106.
## 3        3        0.988  2.20  107.
## 4        4        0.96   2.67  106.
## 5        5        0.954  2.87  106.
## 6        6        0.988  2.77  106.
```

Share and distribution of consumer goods

```
# Modify Dataset to include Ypred
ConsumerGood = ConsumerGood %>% mutate(Ypred=-0.8896+3.547*distribution)

# Linear Model
dist_share_model <- lm(share ~ distribution, data = ConsumerGood)
summary(dist_share_model)
```

```
##
## Call:
## lm(formula = share ~ distribution, data = ConsumerGood)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.03024 -0.35559 -0.04637  0.30933  1.69116
##
```

```
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   -0.8896     0.1718  -5.179 1.07e-06 ***
## distribution   3.5470     0.2250  15.764  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.5724 on 106 degrees of freedom
## Multiple R-squared:  0.701,  Adjusted R-squared:  0.6982
## F-statistic: 248.5 on 1 and 106 DF,  p-value: < 2.2e-16
```

```r
# scatter plot
gf_point(
  share ~ distribution,
  data = ConsumerGood,
  # sets color based on residual
  color = ifelse(residuals(dist_share_model) == 0,"green",
  # if point not on line set as red or blue
  ifelse(residuals(dist_share_model) > 0, "red", "blue")
  )) %>%
  # plots line of best fit
  gf_lm(color="gray50",alpha=0.5) %>%
  # plots dotted residual lines from point to linear model
  gf_linerange(Ypred+share~distribution,linetype="dotted",color="gray30")
```