Predictors of Productivity in a Work Setting

Daniel Pinedo

Psych 308c: Assignment 1

I have highlighted areas of feedback throughout the paper. Please come see a TA if you'd like to review your n

Predictors of Productivity in a Work Setting

Employees will likely be happier if they have autonomy over their time at work. However, it is crucial to determine in which manner this autonomy impacts productivity. A local company initiated a program to allow its employees to take a break from 1-60 minutes at any time during their work day. At the end of 30 days, the company wanted to know how this program worked by measuring productivity as well as the average length of their daily break, how much they enjoyed the break, and their overall desire to come to work. The purpose of this study was to determine how these three latter qualities predict employee productivity.

Method

The present study utilized a correlational design. Data collection methods included the use of an in-person survey administered at the end of the first month of the new break program.

Participants

Participants consisted of all 175 employees of the company. No demographic data was collected.

Measures

Each employee was assessed using the below measures.

Productivity. Productivity (Product) assessed the percentage of time meeting weekly goals using a 100-point scale, with higher scores indicating higher productivity.

Length. Length assessed the length of a break in minutes using a 60-point scale, with higher scores indicating greater amount of time.

Enjoyment. Enjoyment (Enjoy) assessed self-reported level of enjoyment of break using a 10 point scale, with higher scores indicating a greater perceived level of enjoyment.

Desire. Desire assessed self-reported level of desire to come to work using a 10-point scale, with higher scores indicating a greater perceived level of desire.

Planned Analysis

The present study planned to use correlation, simple regression, and multiple regression to assess the relationships between predictors, as well as predictors and the outcome variable.

Results

Data analysis can be found in Appendix A. Descriptive statistics can be found in Table 1. There were no missing data in the dataset and analysis continued with tests of assumptions. Descriptive statistics and inspection of histograms reveal that the data do not violate the assumption of univariate normality with the possible exception of the outcome variable (Product) being bimodal. Data appear to be normally distributed across all variables in the model, as evidenced by skew for all variables being below a threshold of ± 3.00 (Length = -0.07, Enjoy = -0.29, Desire = -0.33, Product = -0.15), and kurtosis below a threshold of ± 10.00 (Length = -0.43, Enjoy = 0.88, Desire = 0.60, Product = -0.39). Scatterplots were run to assess the assumption of homoscedasticity. This assumption does not appear to be violated for the variables of length of break, enjoyment of break, and desire to come to work as the variance across each variable appears to be stable based on visual inspection of the scatterplots. For example, there is less variance in the low end of enjoyment of break, and less variance at the high end of desire to come to work, but these differences do not appear to be significant. In addition, the variable of length of break appears to meet the assumption of homoscedasticity. Finally, the assumption of linearity appears to be met for all variables.

Both enjoyment of break and desire to come to work were significantly correlated with productivity (Table 2), therefore the relationship between the outcome (productivity) and these two predictors were assessed through regression analyses. Enjoyment of break explained 11% of the variance in intent to return, $R^2 = .11$, $F(1, 173) = 21.90$, $p < .001$ (Table 3). Adding the desire to go to work variable to the model explained 17% of the variance in productivity, $R^2 = .17$, $F(2, 172) = 17.4$, $p < .001$ (Table 4). Model comparison of the two models indicated that the model with both predictor variables of enjoyment of break and desire to go to work was significantly better than a model with only the predictor for enjoyment of break, $F(1, 172)$, $\Delta R^2 = .06$, $p < .001$.

Discussion

The purpose of the current project was to test predictors of work productivity. All employees ($N = 175$) were assessed for the length of their breaks, enjoyment of breaks, desire to go to work, and their performance. Correlation and regression analyses were used to determine whether these variables predicted productivity as time meeting weekly goals.

Correlation analyses demonstrated that length of break was not significantly related with performance and not included for further analysis. The remaining two predictor variables of enjoyment of break and desire to go to work were added to simple regression analyses. When including enjoyment and desire in the linear model both were significant predictors of performance (Table 4).

In summary, these results indicate that both enjoyment of break and desire to go to work were significant predictors of employee's productivity as time meeting weekly goals. However, when comparing the predictors, enjoyment is a better predictor of performance. In conclusion, if the company wants to increase employee time meeting weekly goals, they should continue to focus on promoting autonomy through flexible break times and thereby job satisfaction (e.g, desire to go to work). However, considering that the productivity variable appears to be nonnormal (i.e. bimodal), these results should also be reconsidered using alternative methods and analyses to confirm.

Table 1.

Descriptive Statistics of Predictors of Employee Productivity

	Length	Enjoy	Desire	Product
Mean	38.70	7.97	7.58	78.80
Median	38.5	8.00	7.60	79.00
SD	5.49	0.53	0.73	4.60
Min	24.90	6.00	5.20	67.00
Max	50.80	9.30	9.60	90.00
Skewness	-0.07	-0.29	-0.33	-0.15
Kurtosis	-0.43	0.88	0.60	-0.39

Table 2.

Correlation Matrix for Employee Productivity

	Length	Enjoy	Desire	Product
Length	1.00			
Enjoy	-0.03	1.00		
Desire	0.02	0.30*	1.00	
Product	0.03	0.34*	0.33*	1.00

Note. * $p < .001$

Table 3.

Regression of Enjoyment of Break onto Productivity

	β	Estimate	SE	*t*
Intercept		55.65	4.95	11.23*
Enjoyment	.34	2.91	0.62	4.68*

Note. * $p < .001$

Table 4.

Regression of Enjoyment of Break and Desire to Go to Work onto Productivity

	β	Estimate	SE	*t*
Intercept		48.77	5.22	9.35*
Enjoyment	.26	2.28	0.63	3.62*
Desire	.25	1.57	0.46	3.41*

Note. * $p < .001$

Appendix A

Statistical Analysis in R

Hypotheses

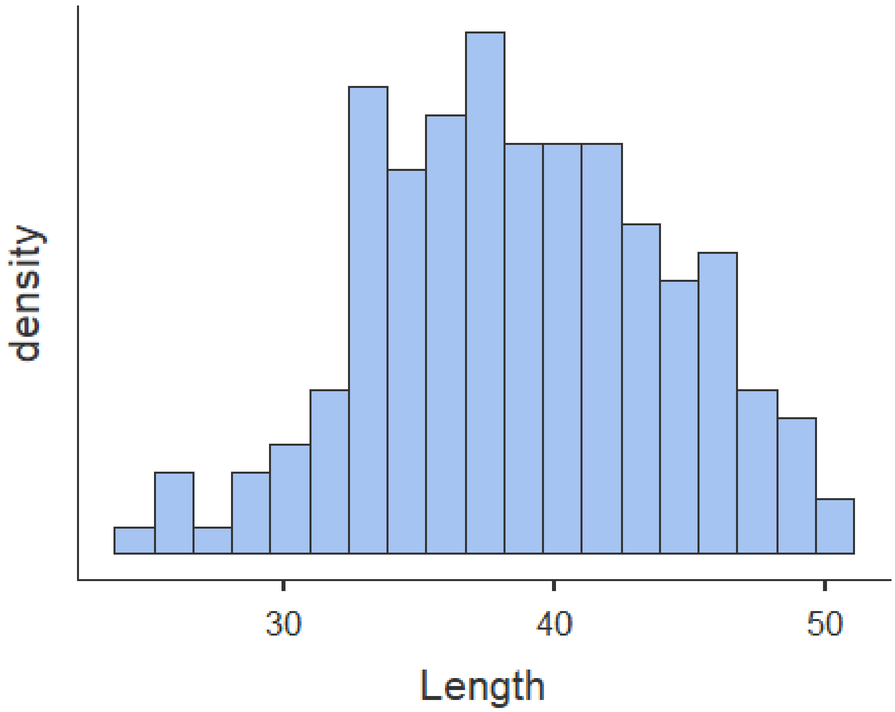
H_0: no relationship between variables

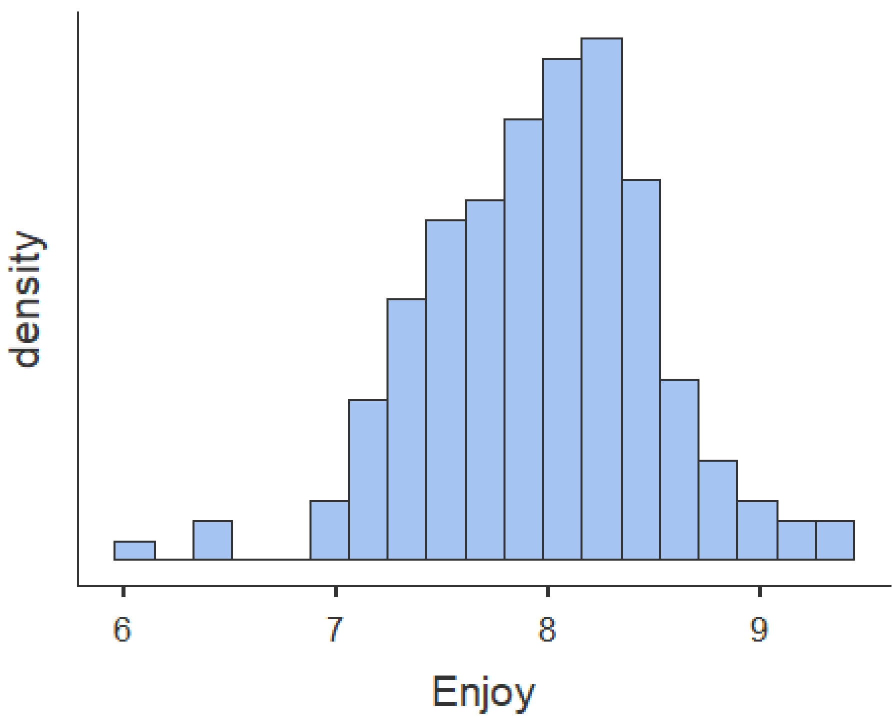
H_a: length, enjoy, and desire predict product

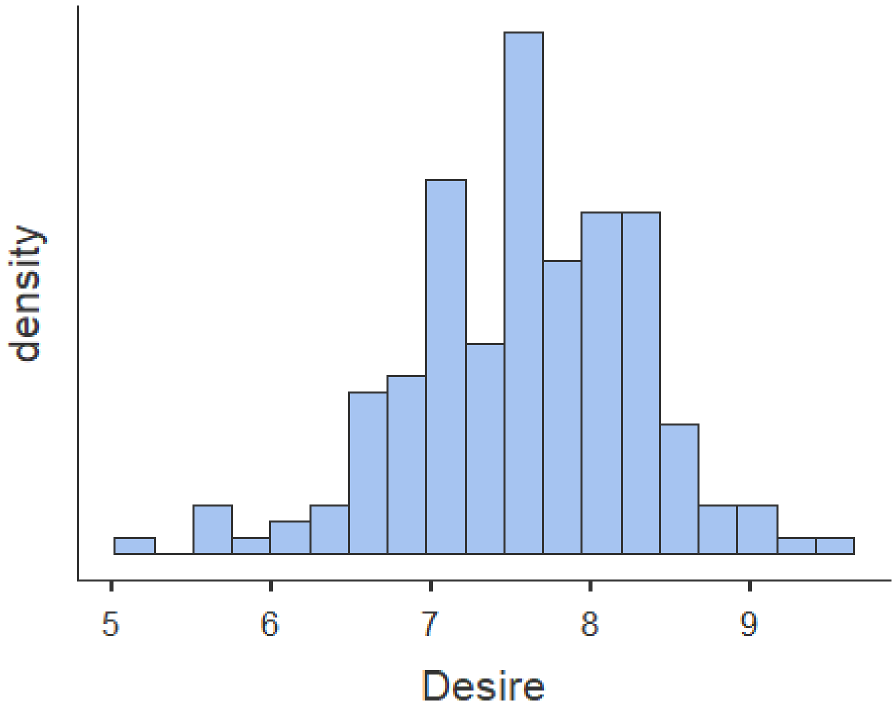
Descriptive Statistics and Assumptions

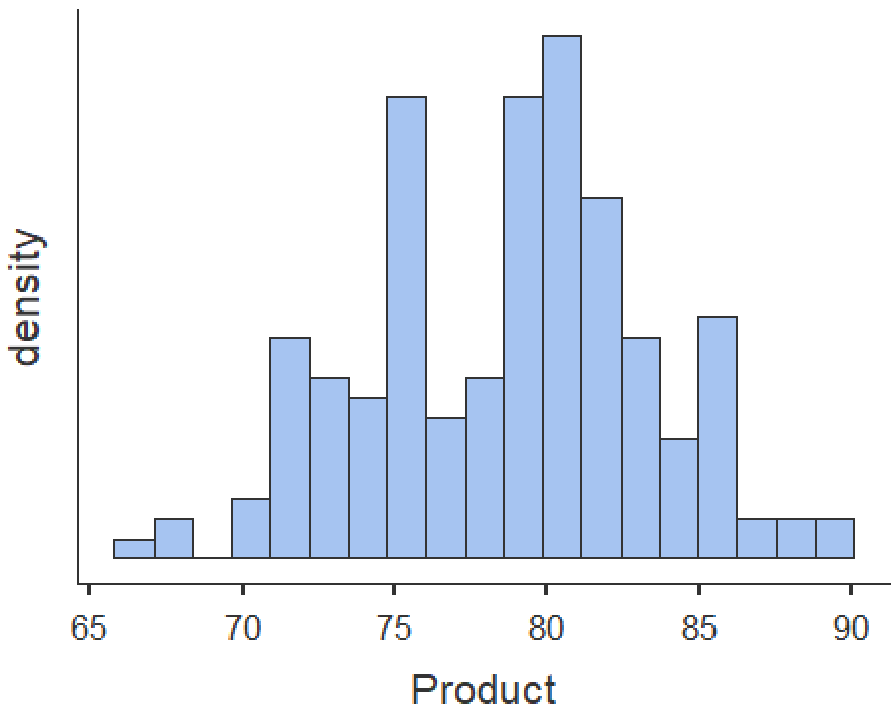
```
# Prerequisitites
  # 1. Variables are measured on the continuous level

# Assumptions
  # 1. Normal Distribution for X and Y (Product) [i.e. histogram, skew +-3, kurtosis +-10]
    # Distribition for Y appears to be bimodal, but otherwise normally distributed
    # Skew for Y is -0.15; Kurtosis for Y is -.38 ---> both pass
  # 2. Linear Relationship between X and Y
    # Visual inspection of scatterplot and prediction model line indicate a linear relationship
  # 3. Homoscedasticity
    # Visual inspection of scatterplots indicate homoscedasticity is true for all X/Y relationships
  # 4. [Examine residuals (e = Y - Y~predicted~) to understand 2 and 3]

# Descriptives [Assumption 1]
desc <- descriptives(data = dat,
                     vars = c('Length', 'Enjoy', 'Desire', 'Product'),
                     hist = TRUE,
                     sd = TRUE,
                     range = TRUE,
                     skew = TRUE,
                     kurt = TRUE)
desc

##
##  DESCRIPTIVES
##
```

```
## Descriptives
## ---------------------------------------------------------------
##                          Length    Enjoy    Desire    Product
## ---------------------------------------------------------------
##   N                          175      175      175        175
##   Missing                      0        0        0          0
##   Mean                      38.7     7.97     7.58       78.8
##   Median                    38.5     8.00     7.60         79
##   Standard deviation        5.49    0.531    0.727       4.60
##   Range                     25.9     3.30     4.40         23
##   Minimum                   24.9     6.00     5.20         67
##   Maximum                   50.8     9.30     9.60         90
##   Skewness               -0.0738   -0.288   -0.329     -0.150
##   Std. error skewness      0.184    0.184    0.184      0.184
##   Kurtosis                -0.428    0.876    0.598     -0.384
##   Std. error kurtosis      0.365    0.365    0.365      0.365
## ---------------------------------------------------------------
```

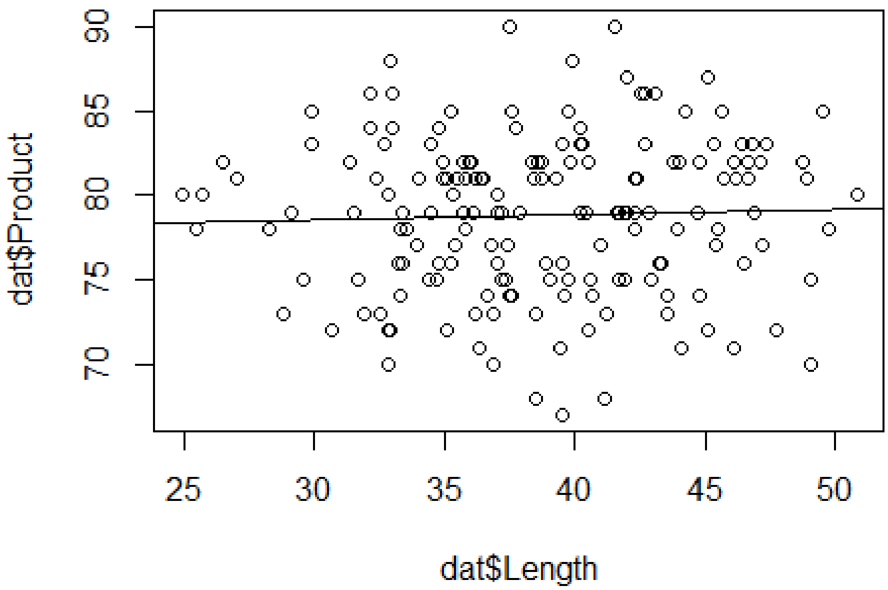

density
30
40
50
Length


density
6
7
8
9
Enjoy

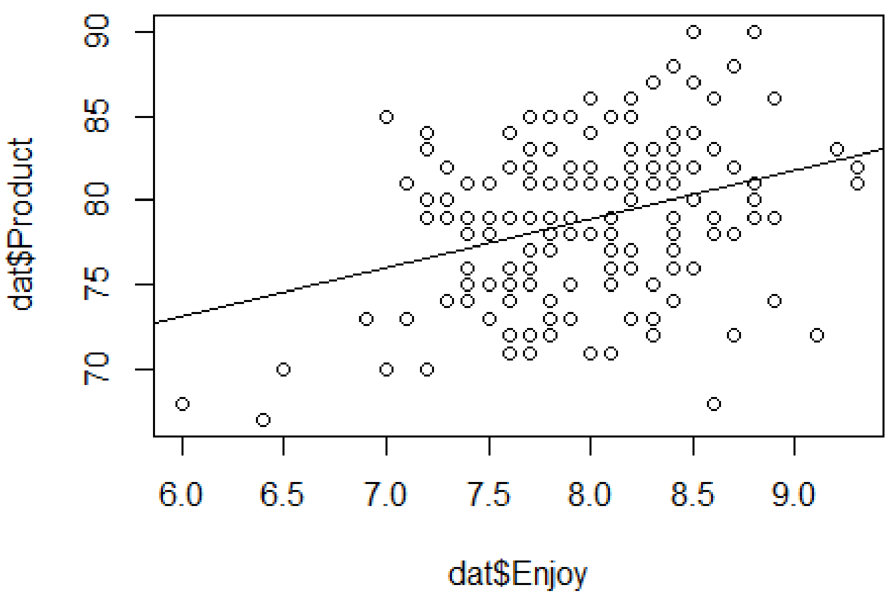




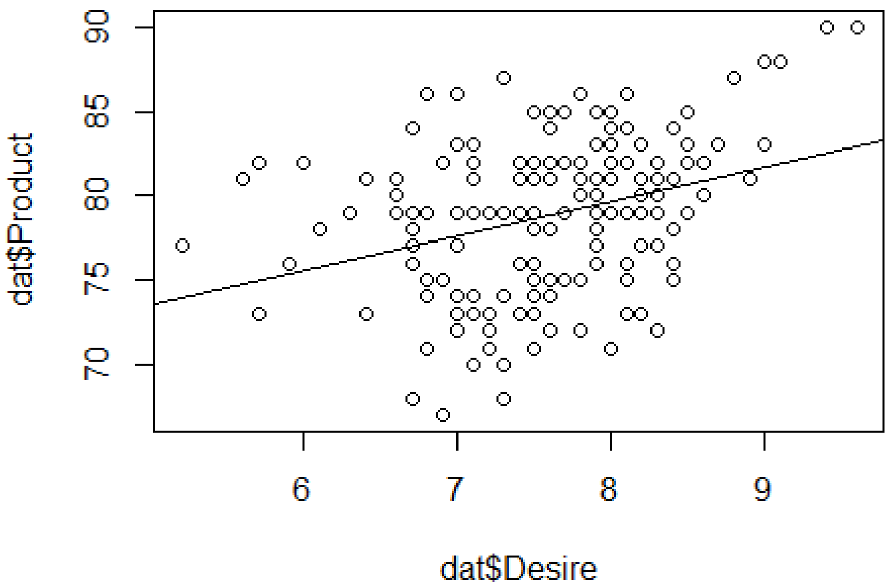

```
# Scatterplots
plot(dat$Length, dat$Product, abline(lm(dat$Product ~ dat$Length)))
```




```
plot(dat$Enjoy, dat$Product, abline(lm(dat$Product ~ dat$Enjoy)))
```



```
plot(dat$Desire, dat$Product, abline(lm(dat$Product ~ dat$Desire)))
```



Correlations

```
# Correlation
cortable <- corrMatrix(data = dat,
                  vars = c('Length', 'Enjoy', 'Desire', 'Product'),
                  flag = TRUE)
cortable

##
##  CORRELATION MATRIX
##
##  Correlation Matrix
##  -----------------------------------------------------------------
##                           Length    Enjoy     Desire    Product
##  -----------------------------------------------------------------
##    Length     Pearson's r       -    -0.025     0.015      0.034
##               p-value           -     0.744     0.843      0.659
```

```
##
##   Enjoy     Pearson's r             -     0.292     0.336
##             p-value             -     < .001    < .001
##
##   Desire    Pearson's r                       -     0.325
##             p-value                       -     < .001
##
##   Product   Pearson's r                               -
##             p-value                               -
## ------------------------------------------------------------
##   Note. * p < .05, ** p < .01, *** p < .001
```

Simple Regression

```
# Simple Regression Model 1
# Start with the simpler model first - Enjoy is most correlated with outcome variable (Product)
model1 <- linReg(data = dat,
           dep = 'Product', #outcome
           covs = c('Enjoy'), #predictors
           blocks = list(c('Enjoy')), #order - doesn't matter for simple regression as there is only
one variable
           modelTest = TRUE, #significance test on model [H0: R squared = 0]
           stdEst = TRUE) #standardized regression coefficient for individual variable [Stand.
Estimate]
model1 #print to screen

##
## LINEAR REGRESSION
##
## Model Fit Measures
## -------------------------------------------------------
##   Model   R       R²      F      df1   df2   p
## -------------------------------------------------------
##      1   0.336   0.113   21.9     1   173   < .001
## -------------------------------------------------------
```

```
##
##
##  MODEL SPECIFIC RESULTS
##
##  MODEL 1
##
##  Model Coefficients
##  ----------------------------------------------------------------------
##    Predictor    Estimate    SE      t       p        Stand. Estimate
##  ----------------------------------------------------------------------
##    Intercept      55.65    4.954    11.23    < .001
##    Enjoy           2.91    0.620     4.68    < .001             0.336
##  ----------------------------------------------------------------------

#This model is best fit for simple regression based on R squared and Beta Estimates

# ALTERNATIVE
model1.1<- lm(Product ~ Enjoy, data = dat)
summary(model1.1)

##
## Call:
## lm(formula = Product ~ Enjoy, data = dat)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -12.6459  -3.0646   0.3886   3.0980   9.6447
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  55.6501     4.9536  11.234  < 2e-16 ***
## Enjoy         2.9065     0.6204   4.685 5.65e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
## Residual standard error: 4.347 on 173 degrees of freedom
## Multiple R-squared:  0.1126, Adjusted R-squared:  0.1075
## F-statistic: 21.95 on 1 and 173 DF,  p-value: 5.649e-06
```

```
# Simple Regression Model 2
# Desire is second most correlated with outcome variable (Product)
model2 <- linReg(data = dat,
            dep = 'Product', #outcome
            covs = c('Desire'), #predictors
            blocks = list(c('Desire')), #order - doesn't matter foı simple regression as there is only
one variable
            modelTest = TRUE, #significance test on model [H0: R squared = 0]
            stdEst = TRUE) #standardized regression coefficient for individual variable
model2 #print to screen
```

```
##
##  LINEAR REGRESSION
##
##  Model Fit Measures
##  ----------------------------------------------------------
##    Model    R       R²      F      df1    df2    p
##  ----------------------------------------------------------
##        1    0.325    0.105    20.4     1    173    < .001
##  ----------------------------------------------------------
##
##
##  MODEL SPECIFIC RESULTS
##
##  MODEL 1
##
##  Model Coefficients
##  ----------------------------------------------------------------------
##    Predictor    Estimate    SE      t       p        Stand. Estimate
```

```
## ----------------------------------------------------------------------
##    Intercept      63.24    3.465    18.25    < .001
##    Desire          2.05    0.455     4.51    < .001             0.325
## ----------------------------------------------------------------------
```

```
# ALTERNATIVE
model2.1<- lm(Product ~ Desire, data = dat)
summary(model2.1)
```

```
##
## Call:
## lm(formula = Product ~ Desire, data = dat)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -10.4027  -3.5318   0.5179   3.1867   8.8026
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  63.2376     3.4648  18.251  < 2e-16 ***
## Desire        2.0529     0.4548   4.514 1.17e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.364 on 173 degrees of freedom
## Multiple R-squared:  0.1054, Adjusted R-squared:  0.1002
## F-statistic: 20.37 on 1 and 173 DF,  p-value: 1.174e-05
```

Multiple Regression

```
# Multiple regression test #A

modelA <- linReg(data = dat,
             dep = 'Product', #outcome
             covs = c('Enjoy', 'Desire'), #predictors
             blocks = list(c('Enjoy', 'Desire')), #order matters here if separate blocks of variables
```

```
are provided
            modelTest = TRUE,
            stdEst = TRUE,
            ciStdEst = TRUE,
            r2Adj = TRUE)
modelA

##
## LINEAR REGRESSION
##
## Model Fit Measures
## ---------------------------------------------------------------------------
##   Model    R        R²       Adjusted R²    F       df1    df2    p
## ---------------------------------------------------------------------------
##       1   0.411    0.169          0.159    17.4       2    172    < .001
## ---------------------------------------------------------------------------
##
##
## MODEL SPECIFIC RESULTS
##
## MODEL 1
##
## Model Coefficients
## -------------------------------------------------------------------------------------------
##   Predictor    Estimate    SE       t       p          Stand. Estimate    Lower    Upper
## -------------------------------------------------------------------------------------------
##   Intercept      48.77    5.215    9.35    < .001
##   Enjoy           2.28    0.630    3.62    < .001              0.263    0.120    0.407
##   Desire          1.57    0.460    3.41    < .001              0.248    0.104    0.391
## -------------------------------------------------------------------------------------------

# ALTERNATIVE
modelA.1<- lm(Product ~ Enjoy + Desire, data = dat)
summary(modelA.1)
```

```
##
## Call:
## lm(formula = Product ~ Enjoy + Desire, data = dat)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -11.8048  -3.0613   0.4341   3.0130   8.3458
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  48.7712     5.2153   9.351  < 2e-16 ***
## Enjoy         2.2791     0.6298   3.619 0.000389 ***
## Desire        1.5662     0.4598   3.406 0.000820 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.219 on 172 degrees of freedom
## Multiple R-squared:  0.1687, Adjusted R-squared:  0.159
## F-statistic: 17.45 on 2 and 172 DF,  p-value: 1.262e-07
```

```
# Multiple regression test #B

modelB <- linReg(data = dat,
            dep = 'Product', #outcome
            covs = c('Enjoy', 'Desire', 'Length'), #predictors
            blocks = list(c('Enjoy', 'Desire', 'Length')), #order matters here if separate blocks of variables are provided
            modelTest = TRUE,
            stdEst = TRUE,
            ciStdEst = TRUE,
            r2Adj = TRUE)
modelB
```

```
##
## LINEAR REGRESSION
##
## Model Fit Measures
## -----------------------------------------------------------------------
##   Model    R        R²       Adjusted R²    F       df1    df2    p
## -----------------------------------------------------------------------
##       1    0.412    0.170         0.155    11.7      3    171    < .001
## -----------------------------------------------------------------------
##
##
## MODEL SPECIFIC RESULTS
##
## MODEL 1
##
## Model Coefficients
## ---------------------------------------------------------------------------------------
##   Predictor    Estimate    SE        t        p         Stand. Estimate    Lower    Upper
## ---------------------------------------------------------------------------------------
##   Intercept    47.5536    5.7223    8.310    < .001
##   Enjoy         2.2892    0.6314    3.626    < .001            0.2643     0.120    0.408
##   Desire        1.5606    0.4609    3.386    < .001            0.2467     0.103    0.391
##   Length        0.0305    0.0584    0.523     0.602            0.0364    -0.101    0.174
## ---------------------------------------------------------------------------------------
```

```
# ALTERNATIVE
modelB.1<- lm(Product ~ Enjoy + Desire + Length, data = dat)
summary(modelB.1)
```

```
##
## Call:
## lm(formula = Product ~ Enjoy + Desire + Length, data = dat)
##
## Residuals:
```

```
##      Min       1Q   Median       3Q      Max
## -11.8072  -3.0748   0.3791   2.9392   8.5144
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 47.55359    5.72230   8.310 2.84e-14 ***
## Enjoy        2.28918    0.63140   3.626  0.00038 ***
## Desire       1.56058    0.46091   3.386  0.00088 ***
## Length       0.03050    0.05837   0.523  0.60195
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.228 on 171 degrees of freedom
## Multiple R-squared:   0.17,  Adjusted R-squared:  0.1554
## F-statistic: 11.67 on 3 and 171 DF,  p-value: 5.36e-07
```

Model Comparison

```
# Hierarchical regression with model comparison (significance of R squared change)
# 2 models plus comparison of them for final homework should be presented

# Comparison Model 1
  # Model B: Product ~ Enjoy + Desire + Length
  # Model A: Product ~ Enjoy + Desire

compare1 <- linReg(data = dat,
                   dep = 'Product',
                   covs = c('Enjoy', 'Desire', 'Length'),
                   blocks = list(
                     list('Enjoy', 'Desire'), #Model A
                     list('Length')), #Model B
                   modelTest = TRUE,
                   r2Adj = TRUE,
                   stdEst = TRUE,
```

```
            ciStdEst = TRUE)
compare1

##
## LINEAR REGRESSION
##
## Model Fit Measures
## ------------------------------------------------------------------------
##  Model   R      R²      Adjusted R²   F     df1   df2   p
## ------------------------------------------------------------------------
##      1   0.411   0.169        0.159   17.4     2   172   < .001
##      2   0.412   0.170        0.155   11.7     3   171   < .001
## ------------------------------------------------------------------------
##
##
## Model Comparisons
## ----------------------------------------------------------------
##  Model      Model   <U+0394>R²      F      df1   df2   p
## ----------------------------------------------------------------
##      1   -      2   0.00133   0.273     1   171    0.602
## ----------------------------------------------------------------
##
##
## MODEL SPECIFIC RESULTS
##
## MODEL 1
##
## Model Coefficients
## ------------------------------------------------------------------------------------------
##  Predictor   Estimate   SE     t     p       Stand. Estimate   Lower   Upper
## ------------------------------------------------------------------------------------------
##  Intercept     48.77   5.215   9.35   < .001
##  Enjoy         2.28   0.630   3.62   < .001            0.263   0.120   0.407
##  Desire        1.57   0.460   3.41   < .001            0.248   0.104   0.391
```

```
## ----------------------------------------------------------------------------------------
##
##
## MODEL 2
##
## Model Coefficients
## -------------------------------------------------------------------------------------------
##   Predictor   Estimate   SE       t      p        Stand. Estimate   Lower    Upper
## -------------------------------------------------------------------------------------------
##   Intercept   47.5536   5.7223   8.310   < .001
##   Enjoy        2.2892   0.6314   3.626   < .001        0.2643    0.120    0.408
##   Desire       1.5606   0.4609   3.386   < .001        0.2467    0.103    0.391
##   Length       0.0305   0.0584   0.523    0.602        0.0364   -0.101    0.174
## -------------------------------------------------------------------------------------------

# ALTERNATIVE
stats::anova(modelB.1, modelA.1)

## Analysis of Variance Table
##
## Model 1: Product ~ Enjoy + Desire + Length
## Model 2: Product ~ Enjoy + Desire
##   Res.Df    RSS Df Sum of Sq      F Pr(>F)
## 1    171 3057.3
## 2    172 3062.2 -1   -4.8823 0.2731 0.6019

# Both statistical tests yield no significant difference between models B and A

# Comparison Model 2
  # Model A: Product ~ Enjoy + Desire
  # Model 1: Product ~ Enjoy


compare2 <- linReg(data = dat,
                   dep = 'Product',
                   covs = c('Enjoy', 'Desire'),
```

```
          blocks = list(
            list('Enjoy'), #Model 1
            list('Desire')), #Model A
          modelTest = TRUE,
          r2Adj = TRUE,
          stdEst = TRUE,
          ciStdEst = TRUE)
compare2
```

```
##
##  LINEAR REGRESSION
##
##  Model Fit Measures
##  -------------------------------------------------------------------------
##    Model    R        R²       Adjusted R²    F      df1    df2    p
##  -------------------------------------------------------------------------
##        1    0.336    0.113          0.107    21.9      1    173    < .001
##        2    0.411    0.169          0.159    17.4      2    172    < .001
##  -------------------------------------------------------------------------
##
##
##  Model Comparisons
##  ----------------------------------------------------------------
##    Model        Model    <U+0394>R²      F      df1    df2    p
##  ----------------------------------------------------------------
##        1    -        2    0.0561    11.6      1    172    < .001
##  ----------------------------------------------------------------
##
##
##  MODEL SPECIFIC RESULTS
##
##  MODEL 1
##
##  Model Coefficients
```

```
## ------------------------------------------------------------------------------------
##   Predictor   Estimate   SE     t      p       Stand. Estimate   Lower   Upper
## ------------------------------------------------------------------------------------
##   Intercept     55.65   4.954   11.23   < .001
##   Enjoy          2.91   0.620    4.68   < .001           0.336
## ------------------------------------------------------------------------------------
##
##
## MODEL 2
##
## Model Coefficients
## -----------------------------------------------------------------------------------
##   Predictor   Estimate   SE     t      p       Stand. Estimate   Lower   Upper
## -----------------------------------------------------------------------------------
##   Intercept     48.77   5.215   9.35   < .001
##   Enjoy          2.28   0.630   3.62   < .001           0.263   0.120   0.407
##   Desire         1.57   0.460   3.41   < .001           0.248   0.104   0.391
## -----------------------------------------------------------------------------------
```

```r
# ALTERNATIVE
stats::anova(modelA.1, model1.1)
```

```
## Analysis of Variance Table
##
## Model 1: Product ~ Enjoy + Desire
## Model 2: Product ~ Enjoy
##   Res.Df    RSS Df Sum of Sq      F  Pr(>F)
## 1    172 3062.2
## 2    173 3268.7 -1   -206.56 11.602 0.00082 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
# Both statistical tests yield significant difference between models A and 1
# These model comparisons yield that model A is the best fit for outcome variable Product
```

Interpretation

```
# Interpret
```

Visualization

```
# plotting a multiple regression model based on:
  # Model A: Product ~ Enjoy + Desire (from lm command of model created 'modelA.1')



# create predicted values from three predictors and save in object
model_p <- ggpredict(modelA.1, terms = c('Enjoy', 'Desire'), full.data = TRUE,  pretty = FALSE)

# plot predicted line
plot <- ggplot(model_p, aes(x, predicted)) +
    geom_smooth(method = "lm", se = FALSE, fullrange=TRUE) + xlab("Score") +
ggtitle("Plot of Model Predicting Productivity") + ylab("Weekly Goal Percentage") +
    geom_point() + theme_minimal()


plot
```