Cloverleaf Regression Analysis - Revisited

Danny Moncada (monca016), MSBA

```
## Load data file
cloverleaf <- read.csv("PaidSearch.csv", header = T)

## Just some random regression models on the de-duplicated dataset no real rhyme or reason as to which
## No variables have been altered as of yet, just that some were investigated as factors instead of int

m1 <- lm(cloverleaf$clicks ~ cloverleaf$adQuality + cloverleaf$adrank + cloverleaf$numberofwords + as.f
summary(m1)
```

```
##
## Call:
## lm(formula = cloverleaf$clicks ~ cloverleaf$adQuality + cloverleaf$adrank +
##     cloverleaf$numberofwords + as.factor(cloverleaf$retailer) +
##     as.factor(cloverleaf$brandname))
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1556.4  -695.4  -197.6   143.9  7639.8
##
## Coefficients:
##                                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)                       -516.0035   165.5822  -3.116  0.00188 **
## cloverleaf$adQuality                81.4649    17.5617   4.639 3.97e-06 ***
## cloverleaf$adrank                   -0.4798     5.2492  -0.091  0.92719
## cloverleaf$numberofwords           215.5021    49.3604   4.366 1.40e-05 ***
## as.factor(cloverleaf$retailer)1    536.0624   131.8060   4.067 5.13e-05 ***
## as.factor(cloverleaf$brandname)1 -136.8608   152.6547  -0.897  0.37018
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1646 on 1004 degrees of freedom
## Multiple R-squared:  0.1043, Adjusted R-squared:  0.09989
## F-statistic: 23.39 on 5 and 1004 DF,  p-value: < 2.2e-16
```

```
m2 <- lm(cloverleaf$revenue ~ cloverleaf$bidprice + cloverleaf$adrank + cloverleaf$adQuality + cloverlea
summary(m2)
```

```
##
## Call:
## lm(formula = cloverleaf$revenue ~ cloverleaf$bidprice + cloverleaf$adrank +
##     cloverleaf$adQuality + cloverleaf$numberofwords + as.factor(cloverleaf$retailer) +
##     as.factor(cloverleaf$brandname))
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1732.7  -776.9  -209.2   197.1  8917.8
##
## Coefficients:
```

```
##                                  Estimate Std. Error t value Pr(>|t|)
## (Intercept)                      -544.681    184.921  -2.945   0.0033 **
## cloverleaf$bidprice              -237.603    462.727  -0.513   0.6077
## cloverleaf$adrank                  -1.211      5.788  -0.209   0.8344
## cloverleaf$adQuality               87.527     19.149   4.571 5.46e-06 ***
## cloverleaf$numberofwords          238.939     54.699   4.368 1.38e-05 ***
## as.factor(cloverleaf$retailer)1   584.567    146.220   3.998 6.86e-05 ***
## as.factor(cloverleaf$brandname)1 -137.042    166.328  -0.824   0.4102
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1792 on 1003 degrees of freedom
## Multiple R-squared:  0.1069, Adjusted R-squared:  0.1016
## F-statistic: 20.02 on 6 and 1003 DF,  p-value: < 2.2e-16
```

```
m3 <- lm(cloverleaf$clickthroughrate ~ cloverleaf$adQuality + cloverleaf$adrank + cloverleaf$numberofwo
summary(m3)
```

```
##
## Call:
## lm(formula = cloverleaf$clickthroughrate ~ cloverleaf$adQuality +
##     cloverleaf$adrank + cloverleaf$numberofwords + as.factor(cloverleaf$retailer) +
##     as.factor(cloverleaf$brandname))
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.51091 -0.04920  0.02223  0.05785  0.82454
##
## Coefficients:
##                                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)                      0.0541875  0.0137431   3.943 8.61e-05
## cloverleaf$adQuality             0.0269127  0.0014576  18.464  < 2e-16
## cloverleaf$adrank               -0.0018852  0.0004357  -4.327 1.66e-05
## cloverleaf$numberofwords        -0.0319009  0.0040968  -7.787 1.71e-14
## as.factor(cloverleaf$retailer)1  0.1786126  0.0109397  16.327  < 2e-16
## as.factor(cloverleaf$brandname)1 0.0135382  0.0126701   1.069    0.286
##
## (Intercept)                      ***
## cloverleaf$adQuality             ***
## cloverleaf$adrank                ***
## cloverleaf$numberofwords         ***
## as.factor(cloverleaf$retailer)1  ***
## as.factor(cloverleaf$brandname)1
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1366 on 1004 degrees of freedom
## Multiple R-squared:  0.6468, Adjusted R-squared:  0.6451
## F-statistic: 367.8 on 5 and 1004 DF,  p-value: < 2.2e-16
```

```
m4 <- lm(cloverleaf$conversionrate ~ cloverleaf$adQuality + cloverleaf$adrank + cloverleaf$numberofword
summary(m4)
```

```
##
## Call:
## lm(formula = cloverleaf$conversionrate ~ cloverleaf$adQuality +
##     cloverleaf$adrank + cloverleaf$numberofwords + as.factor(cloverleaf$retailer) +
##     as.factor(cloverleaf$brandname))
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.16236 -0.06103 -0.01885  0.00587  0.98519
##
## Coefficients:
##                                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)                     -0.0770086  0.0157003  -4.905 1.09e-06
## cloverleaf$adQuality             0.0043400  0.0016652   2.606 0.009287
## cloverleaf$adrank                0.0034504  0.0004977   6.932 7.40e-12
## cloverleaf$numberofwords         0.0176582  0.0046803   3.773 0.000171
## as.factor(cloverleaf$retailer)1  0.0242595  0.0124977   1.941 0.052524
## as.factor(cloverleaf$brandname)1 0.0313461  0.0144746   2.166 0.030577
##
## (Intercept)                     ***
## cloverleaf$adQuality            **
## cloverleaf$adrank               ***
## cloverleaf$numberofwords        ***
## as.factor(cloverleaf$retailer)1 .
## as.factor(cloverleaf$brandname)1 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1561 on 1004 degrees of freedom
## Multiple R-squared:  0.07261,    Adjusted R-squared:  0.06799
## F-statistic: 15.72 on 5 and 1004 DF,  p-value: 6.339e-15
```

```
## Does ad rank have an effect on revenue?

m5 <- lm(cloverleaf$revenue ~ as.factor(cloverleaf$adrank))
summary(m5)
```

```
##
## Call:
## lm(formula = cloverleaf$revenue ~ as.factor(cloverleaf$adrank))
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -872.3 -872.3 -134.2    0.0 9509.8
##
## Coefficients:
##                                  Estimate Std. Error t value Pr(>|t|)
## (Intercept)                     1.774e-12  1.796e+02   0.000    1.000
## as.factor(cloverleaf$adrank)1   8.723e+02  1.964e+02   4.441 9.96e-06 ***
## as.factor(cloverleaf$adrank)2   1.342e+02  3.202e+02   0.419    0.675
## as.factor(cloverleaf$adrank)3   4.552e+01  3.816e+02   0.119    0.905
## as.factor(cloverleaf$adrank)4   2.462e+02  3.816e+02   0.645    0.519
## as.factor(cloverleaf$adrank)5   7.724e+01  3.865e+02   0.200    0.842
## as.factor(cloverleaf$adrank)6   9.998e+00  4.770e+02   0.021    0.983
```

```
## as.factor(cloverleaf$adrank)7   3.657e-14  7.310e+02   0.000    1.000
## as.factor(cloverleaf$adrank)8   3.891e-14  7.861e+02   0.000    1.000
## as.factor(cloverleaf$adrank)9   3.041e-14  9.544e+02   0.000    1.000
## as.factor(cloverleaf$adrank)10  2.695e+01  4.092e+02   0.066    0.947
## as.factor(cloverleaf$adrank)11 -4.116e-14  1.883e+03   0.000    1.000
## as.factor(cloverleaf$adrank)12  2.245e-14  1.097e+03   0.000    1.000
## as.factor(cloverleaf$adrank)13  4.332e-14  6.867e+02   0.000    1.000
## as.factor(cloverleaf$adrank)14  1.804e-14  1.338e+03   0.000    1.000
## as.factor(cloverleaf$adrank)15 -2.964e-14  1.338e+03   0.000    1.000
## as.factor(cloverleaf$adrank)16  1.799e+02  7.861e+02   0.229    0.819
## as.factor(cloverleaf$adrank)17  1.050e+02  6.867e+02   0.153    0.879
## as.factor(cloverleaf$adrank)18  2.252e-14  9.544e+02   0.000    1.000
## as.factor(cloverleaf$adrank)19  2.232e+01  7.861e+02   0.028    0.977
## as.factor(cloverleaf$adrank)20  7.695e+01  4.302e+02   0.179    0.858
## as.factor(cloverleaf$adrank)21  1.677e-13  1.338e+03   0.000    1.000
## as.factor(cloverleaf$adrank)22  1.659e-13  1.338e+03   0.000    1.000
## as.factor(cloverleaf$adrank)23 -9.987e-13  8.574e+02   0.000    1.000
## as.factor(cloverleaf$adrank)24  4.574e-13  5.501e+02   0.000    1.000
## as.factor(cloverleaf$adrank)25  9.579e-13  1.338e+03   0.000    1.000
## as.factor(cloverleaf$adrank)27 -1.724e-13  1.338e+03   0.000    1.000
## as.factor(cloverleaf$adrank)30 -1.704e-13  1.338e+03   0.000    1.000
## as.factor(cloverleaf$adrank)39 -1.212e-13  4.382e+02   0.000    1.000
## as.factor(cloverleaf$adrank)44  1.508e-13  1.883e+03   0.000    1.000
## as.factor(cloverleaf$adrank)49 -2.108e-14  6.867e+02   0.000    1.000
## as.factor(cloverleaf$adrank)50  7.043e-14  6.194e+02   0.000    1.000
## as.factor(cloverleaf$adrank)63  5.395e-14  5.931e+02   0.000    1.000
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1875 on 977 degrees of freedom
## Multiple R-squared:  0.04753,    Adjusted R-squared:  0.01634
## F-statistic: 1.524 on 32 and 977 DF,  p-value: 0.03224
```

```
## Too many levels of rank. Preliminary inspection of ad rank suggests that maybe we should inspect ads
## This creates an additional variable that assigns a 1 to all ad campaigns with ad rank 1, and a 0 to

adrank_1 <- ifelse(cloverleaf$adrank == 1,1,0)
adrank_1
```

```
##    [1] 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1
##   [35] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
##   [69] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
##  [103] 1 1 1 1 1 1 1 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
##  [137] 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
##  [171] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
##  [205] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 0 0 0 0 1 1 1 1 1 1 0
##  [239] 0 0 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
##  [273] 1 1 1 1 1 1 1 1 0 0 0 0 0 0 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
##  [307] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
##  [341] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
##  [375] 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 0 0 0 0 0 0 1 1 1 1 1 1 1 1 1 1 1 1
##  [409] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 0 1 0 1 1 0 0 0 0 0 0 0 0 0 0 0 0
##  [443] 0 0 0 0 1 1 1 0 0 0 0 1 1 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
##  [477] 0 0 1 0 0 1 0 0 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 0 0 1 1 1 0 0 1 0 0
```

```
##  [511] 0 0 0 0 0 0 0 1 1 0 0 1 0 1 0 0 0 0 0 0 0 1 1 1 1 1 1 1 1 1 1 1 1
##  [545] 1 1 1 1 1 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
##  [579] 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
##  [613] 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
##  [647] 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0
##  [681] 0 0 0 0 0 0 0 0 0 1 1 1 0 0 0 0 0 0 0 0 0 1 1 1 1 1 1 1 0 0 0 0 0
##  [715] 1 0 0 1 1 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 1 1 1 1 0 1 1 1 1 1 1 1 0 0 0
##  [749] 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0
##  [783] 0 0 0 0 0 0 0 1 0 0 1 1 1 1 1 1 1 0 0 0 0 0 0 0 0 0 0 1 1 1 0 0 0 0
##  [817] 0 0 1 1 1 1 1 1 1 1 1 1 0 1 1 0 0 0 1 1 1 1 1 1 1 0 0 0 0 0 0 0 0 0
##  [851] 0 0 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1
##  [885] 1 1 1 1 1 1 1 1 1 1 1 1 0 0 0 1 1 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0
##  [919] 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
##  [953] 1 0 0 1 1 1 1 0 0 1 1 1 1 1 1 1 1 0 0 0 0 0 1 0 0 0 1 1 1 1 1 1 1 1
##  [987] 1 1 1 1 1 1 1 1 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
```

```
## Let's try this again

m6 <- lm(cloverleaf$revenue ~ as.factor(adrank_1))
summary(m6)
```

```
##
## Call:
## lm(formula = cloverleaf$revenue ~ as.factor(adrank_1))
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -872.3 -872.3  -50.1  -50.1 9509.8
##
## Coefficients:
##                       Estimate Std. Error t value Pr(>|t|)
## (Intercept)              50.06      86.55   0.578    0.563
## as.factor(adrank_1)1    822.22     116.76   7.042 3.51e-12 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1846 on 1008 degrees of freedom
## Multiple R-squared:  0.04689,    Adjusted R-squared:  0.04594
## F-statistic: 49.59 on 1 and 1008 DF,  p-value: 3.505e-12
```

```
## Ad rank does appear to have an effect on revenue
## Or does it...?

# Removing some of the variables we determined were misleading (adQuality, brand name)

m7 <- lm(cloverleaf$clicks ~ as.factor(adrank_1) + as.factor(cloverleaf$numberofwords) + as.factor(clove
summary(m7)
```

```
##
## Call:
## lm(formula = cloverleaf$clicks ~ as.factor(adrank_1) + as.factor(cloverleaf$numberofwords) +
##     as.factor(cloverleaf$retailer))
##
```

```
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1660.0  -506.7  -136.6   261.7  7314.0
##
## Coefficients:
##                                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)                          -1515.6      222.3  -6.818 1.59e-11
## as.factor(adrank_1)1                   501.0      110.7   4.524 6.78e-06
## as.factor(cloverleaf$numberofwords)1  1612.7      241.6   6.675 4.07e-11
## as.factor(cloverleaf$numberofwords)2  1253.9      228.2   5.495 4.95e-08
## as.factor(cloverleaf$numberofwords)3  1521.4      190.9   7.968 4.38e-15
## as.factor(cloverleaf$numberofwords)4   321.1      315.4   1.018    0.309
## as.factor(cloverleaf$numberofwords)5  1016.6     1621.9   0.627    0.531
## as.factor(cloverleaf$retailer)1       1153.3      146.9   7.849 1.07e-14
##
## (Intercept)                         ***
## as.factor(adrank_1)1                ***
## as.factor(cloverleaf$numberofwords)1 ***
## as.factor(cloverleaf$numberofwords)2 ***
## as.factor(cloverleaf$numberofwords)3 ***
## as.factor(cloverleaf$numberofwords)4
## as.factor(cloverleaf$numberofwords)5
## as.factor(cloverleaf$retailer)1     ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1607 on 1002 degrees of freedom
## Multiple R-squared:  0.1482, Adjusted R-squared:  0.1423
## F-statistic: 24.91 on 7 and 1002 DF,  p-value: < 2.2e-16
```

```
m8 <- lm(cloverleaf$revenue ~ as.factor(adrank_1) + as.factor(cloverleaf$numberofwords) + as.factor(clo
summary(m8)
```

```
##
## Call:
## lm(formula = cloverleaf$revenue ~ as.factor(adrank_1) + as.factor(cloverleaf$numberofwords) +
##     as.factor(cloverleaf$retailer))
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1824.5  -550.3   -82.8   266.7  8557.6
##
## Coefficients:
##                                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)                          -1654.8      241.9  -6.840 1.38e-11
## as.factor(adrank_1)1                   556.2      120.5   4.615 4.44e-06
## as.factor(cloverleaf$numberofwords)1  1737.6      263.0   6.608 6.32e-11
## as.factor(cloverleaf$numberofwords)2  1388.1      248.4   5.589 2.95e-08
## as.factor(cloverleaf$numberofwords)3  1648.9      207.8   7.933 5.68e-15
## as.factor(cloverleaf$numberofwords)4   328.4      343.3   0.957    0.339
## as.factor(cloverleaf$numberofwords)5  1098.6     1765.4   0.622    0.534
## as.factor(cloverleaf$retailer)1       1274.2      159.9   7.967 4.39e-15
##
## (Intercept)                         ***
```

```
## as.factor(adrank_1)1                 ***
## as.factor(cloverleaf$numberofwords)1 ***
## as.factor(cloverleaf$numberofwords)2 ***
## as.factor(cloverleaf$numberofwords)3 ***
## as.factor(cloverleaf$numberofwords)4
## as.factor(cloverleaf$numberofwords)5
## as.factor(cloverleaf$retailer)1      ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1749 on 1002 degrees of freedom
## Multiple R-squared:  0.1499, Adjusted R-squared:  0.144
## F-statistic: 25.24 on 7 and 1002 DF,  p-value: < 2.2e-16
```

```
m9 <- lm(cloverleaf$clickthroughrate ~ as.factor(adrank_1) + as.factor(cloverleaf$numberofwords) + as.fa
summary(m9)
```

```
##
## Call:
## lm(formula = cloverleaf$clickthroughrate ~ as.factor(adrank_1) +
##     as.factor(cloverleaf$numberofwords) + as.factor(cloverleaf$retailer))
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.42564 -0.10271  0.02878  0.07644  0.80785
##
## Coefficients:
##                                      Estimate Std. Error t value Pr(>|t|)
## (Intercept)                          0.042615   0.017581   2.424   0.0155
## as.factor(adrank_1)1                 0.234417   0.008758  26.767  < 2e-16
## as.factor(cloverleaf$numberofwords)1 -0.005255   0.019108  -0.275   0.7834
## as.factor(cloverleaf$numberofwords)2 -0.072965   0.018048  -4.043 5.68e-05
## as.factor(cloverleaf$numberofwords)3 -0.084887   0.015103  -5.621 2.47e-08
## as.factor(cloverleaf$numberofwords)4 -0.165391   0.024947  -6.630 5.49e-11
## as.factor(cloverleaf$numberofwords)5  0.722968   0.128281   5.636 2.26e-08
## as.factor(cloverleaf$retailer)1       0.233493   0.011622  20.091  < 2e-16
##
## (Intercept)                          *
## as.factor(adrank_1)1                 ***
## as.factor(cloverleaf$numberofwords)1
## as.factor(cloverleaf$numberofwords)2 ***
## as.factor(cloverleaf$numberofwords)3 ***
## as.factor(cloverleaf$numberofwords)4 ***
## as.factor(cloverleaf$numberofwords)5 ***
## as.factor(cloverleaf$retailer)1      ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1271 on 1002 degrees of freedom
## Multiple R-squared:  0.695,  Adjusted R-squared:  0.6929
## F-statistic: 326.2 on 7 and 1002 DF,  p-value: < 2.2e-16
```

```r
m10 <- lm(cloverleaf$conversionrate ~ as.factor(adrank_1) + as.factor(cloverleaf$numberofwords) + as.fa
summary(m10)
```

```
##
## Call:
## lm(formula = cloverleaf$conversionrate ~ as.factor(adrank_1) +
##     as.factor(cloverleaf$numberofwords) + as.factor(cloverleaf$retailer))
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.13456 -0.06004 -0.03300 -0.00343  0.99657
##
## Coefficients:
##                                      Estimate Std. Error t value Pr(>|t|)
## (Intercept)                          0.041034   0.021624   1.898   0.0580
## as.factor(adrank_1)1                -0.074516   0.010772  -6.918 8.17e-12
## as.factor(cloverleaf$numberofwords)1 -0.008037   0.023502  -0.342   0.7324
## as.factor(cloverleaf$numberofwords)2 -0.001199   0.022198  -0.054   0.9569
## as.factor(cloverleaf$numberofwords)3  0.036910   0.018576   1.987   0.0472
## as.factor(cloverleaf$numberofwords)4  0.037241   0.030684   1.214   0.2251
## as.factor(cloverleaf$numberofwords)5  0.033482   0.157779   0.212   0.8320
## as.factor(cloverleaf$retailer)1       0.056611   0.014294   3.960 8.01e-05
##
## (Intercept)                          .
## as.factor(adrank_1)1                 ***
## as.factor(cloverleaf$numberofwords)1
## as.factor(cloverleaf$numberofwords)2
## as.factor(cloverleaf$numberofwords)3 *
## as.factor(cloverleaf$numberofwords)4
## as.factor(cloverleaf$numberofwords)5
## as.factor(cloverleaf$retailer)1      ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1563 on 1002 degrees of freedom
## Multiple R-squared:  0.07161,    Adjusted R-squared:  0.06512
## F-statistic: 11.04 on 7 and 1002 DF,  p-value: 1.702e-13
```

```r
## Subsets the dataset into two distinct sets
## Also creates unique ad rank variables for each subset with a 1 for campaings with ad in position 1,

cloverleaf_retailer <- cloverleaf[which(cloverleaf$retailer== 1), ]
cloverleaf_nonretailer <- cloverleaf[which(cloverleaf$retailer== 0), ]
adrank_1 <- ifelse(cloverleaf_retailer$adrank == 1,1,0)
adrank_2 <- ifelse(cloverleaf_nonretailer$adrank == 1,1,0)

## The same models investigated with the new subsets

m11 <- lm(cloverleaf_retailer$clicks ~ as.factor(adrank_1) + as.factor(cloverleaf_retailer$numberofword
summary(m11)
```

```
##
## Call:
```

```
## lm(formula = cloverleaf_retailer$clicks ~ as.factor(adrank_1) +
##     as.factor(cloverleaf_retailer$numberofwords) + as.factor(cloverleaf_retailer$brandname))
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2099.1 -1997.1  -307.7   -62.4  6874.9
##
## Coefficients:
##                                            Estimate Std. Error t value
## (Intercept)                                -1465.61     647.17  -2.265
## as.factor(adrank_1)1                        1413.64     366.16   3.861
## as.factor(cloverleaf_retailer$numberofwords)2   70.74     560.18   0.126
## as.factor(cloverleaf_retailer$numberofwords)3 1805.29     309.93   5.825
## as.factor(cloverleaf_retailer$numberofwords)4 -196.35     574.44  -0.342
## as.factor(cloverleaf_retailer$brandname)1     345.77     647.98   0.534
##                                            Pr(>|t|)
## (Intercept)                                0.024124 *
## as.factor(adrank_1)1                       0.000134 ***
## as.factor(cloverleaf_retailer$numberofwords)2 0.899574
## as.factor(cloverleaf_retailer$numberofwords)3 1.26e-08 ***
## as.factor(cloverleaf_retailer$numberofwords)4 0.732689
## as.factor(cloverleaf_retailer$brandname)1     0.593934
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2542 on 362 degrees of freedom
## Multiple R-squared:  0.1409, Adjusted R-squared:  0.1291
## F-statistic: 11.88 on 5 and 362 DF,  p-value: 1.175e-10
```

```
m12 <- lm(cloverleaf_retailer$revenue ~ as.factor(adrank_1) + as.factor(cloverleaf_retailer$numberofwor
summary(m12)
```

```
##
## Call:
## lm(formula = cloverleaf_retailer$revenue ~ as.factor(adrank_1) +
##     as.factor(cloverleaf_retailer$numberofwords) + as.factor(cloverleaf_retailer$brandname))
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2279.4 -2279.4  -337.1    59.4  8102.7
##
## Coefficients:
##                                            Estimate Std. Error t value
## (Intercept)                                 -1605.3      711.1  -2.257
## as.factor(adrank_1)1                         1484.4      402.3   3.689
## as.factor(cloverleaf_retailer$numberofwords)2   223.7      615.5   0.363
## as.factor(cloverleaf_retailer$numberofwords)3  1946.0      340.6   5.714
## as.factor(cloverleaf_retailer$numberofwords)4  -230.1      631.2  -0.365
## as.factor(cloverleaf_retailer$brandname)1        454.4      712.0   0.638
##                                            Pr(>|t|)
## (Intercept)                                0.024574 *
## as.factor(adrank_1)1                       0.000259 ***
## as.factor(cloverleaf_retailer$numberofwords)2 0.716454
## as.factor(cloverleaf_retailer$numberofwords)3  2.3e-08 ***
```

```
## as.factor(cloverleaf_retailer$numberofwords)4 0.715644
## as.factor(cloverleaf_retailer$brandname)1     0.523735
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2793 on 362 degrees of freedom
## Multiple R-squared:  0.134,  Adjusted R-squared:  0.1221
## F-statistic:  11.2 on 5 and 362 DF,  p-value: 4.682e-10
```

```
m13 <- lm(cloverleaf_retailer$clickthroughrate ~ as.factor(adrank_1) + as.factor(cloverleaf_retailer$nu
summary(m13)
```

```
##
## Call:
## lm(formula = cloverleaf_retailer$clickthroughrate ~ as.factor(adrank_1) +
##     as.factor(cloverleaf_retailer$numberofwords) + as.factor(cloverleaf_retailer$brandname))
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.47719 -0.10150  0.02281  0.07120  0.45467
##
## Coefficients:
##                                              Estimate Std. Error t value
## (Intercept)                                   0.05651    0.03189   1.772
## as.factor(adrank_1)1                          0.43914    0.01804  24.341
## as.factor(cloverleaf_retailer$numberofwords)2 -0.05762    0.02760  -2.088
## as.factor(cloverleaf_retailer$numberofwords)3 -0.06814    0.01527  -4.462
## as.factor(cloverleaf_retailer$numberofwords)4 -0.20141    0.02830  -7.116
## as.factor(cloverleaf_retailer$brandname)1      0.04968    0.03193   1.556
##                                              Pr(>|t|)
## (Intercept)                                    0.0772 .
## as.factor(adrank_1)1                          < 2e-16 ***
## as.factor(cloverleaf_retailer$numberofwords)2   0.0375 *
## as.factor(cloverleaf_retailer$numberofwords)3 1.08e-05 ***
## as.factor(cloverleaf_retailer$numberofwords)4 6.00e-12 ***
## as.factor(cloverleaf_retailer$brandname)1       0.1206
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1252 on 362 degrees of freedom
## Multiple R-squared:  0.6882, Adjusted R-squared:  0.6839
## F-statistic: 159.8 on 5 and 362 DF,  p-value: < 2.2e-16
```

```
m14 <- lm(cloverleaf_retailer$conversionrate ~ as.factor(adrank_1) + as.factor(cloverleaf_retailer$numbe
summary(m14)
```

```
##
## Call:
## lm(formula = cloverleaf_retailer$conversionrate ~ as.factor(adrank_1) +
##     as.factor(cloverleaf_retailer$numberofwords) + as.factor(cloverleaf_retailer$brandname))
##
## Residuals:
##      Min       1Q   Median       3Q      Max
```

```
## -0.37732 -0.02016 -0.00823  0.02492  0.66777
##
## Coefficients:
##                                              Estimate Std. Error t value
## (Intercept)                                   0.01264    0.04849   0.261
## as.factor(adrank_1)1                         -0.35715    0.02743 -13.019
## as.factor(cloverleaf_retailer$numberofwords)2  0.02398    0.04197   0.571
## as.factor(cloverleaf_retailer$numberofwords)3  0.04508    0.02322   1.942
## as.factor(cloverleaf_retailer$numberofwords)4  0.07709    0.04304   1.791
## as.factor(cloverleaf_retailer$brandname)1      0.31960    0.04855   6.583
##                                              Pr(>|t|)
## (Intercept)                                    0.7945
## as.factor(adrank_1)1                          < 2e-16 ***
## as.factor(cloverleaf_retailer$numberofwords)2   0.5682
## as.factor(cloverleaf_retailer$numberofwords)3   0.0530 .
## as.factor(cloverleaf_retailer$numberofwords)4   0.0741 .
## as.factor(cloverleaf_retailer$brandname)1     1.62e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1904 on 362 degrees of freedom
## Multiple R-squared:  0.3311, Adjusted R-squared:  0.3218
## F-statistic: 35.83 on 5 and 362 DF,  p-value: < 2.2e-16
```

```
m15 <- lm(cloverleaf_nonretailer$clicks ~ as.factor(adrank_2) + as.factor(cloverleaf_nonretailer$number
summary(m15)
```

```
##
## Call:
## lm(formula = cloverleaf_nonretailer$clicks ~ as.factor(adrank_2) +
##     as.factor(cloverleaf_nonretailer$numberofwords) + as.factor(cloverleaf_nonretailer$brandname))
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -432.39 -174.10  -38.12   39.78 2053.61
##
## Coefficients:
##                                                 Estimate Std. Error
## (Intercept)                                       180.27      30.58
## as.factor(adrank_2)1                              140.99      36.10
## as.factor(cloverleaf_nonretailer$numberofwords)2 -256.29      37.01
## as.factor(cloverleaf_nonretailer$numberofwords)3 -334.19      40.86
## as.factor(cloverleaf_nonretailer$numberofwords)4 -314.53     119.63
## as.factor(cloverleaf_nonretailer$numberofwords)5 -433.39     388.05
## as.factor(cloverleaf_nonretailer$brandname)1       114.14      37.28
##                                                 t value Pr(>|t|)
## (Intercept)                                       5.896 6.06e-09 ***
## as.factor(adrank_2)1                              3.906 0.000104 ***
## as.factor(cloverleaf_nonretailer$numberofwords)2 -6.924 1.08e-11 ***
## as.factor(cloverleaf_nonretailer$numberofwords)3 -8.179 1.56e-15 ***
## as.factor(cloverleaf_nonretailer$numberofwords)4 -2.629 0.008764 **
## as.factor(cloverleaf_nonretailer$numberofwords)5 -1.117 0.264480
## as.factor(cloverleaf_nonretailer$brandname)1       3.061 0.002297 **
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 386.5 on 635 degrees of freedom
## Multiple R-squared:  0.1499, Adjusted R-squared:  0.1418
## F-statistic: 18.65 on 6 and 635 DF,  p-value: < 2.2e-16
```

```
m16 <- lm(cloverleaf_nonretailer$revenue ~ as.factor(adrank_2) + as.factor(cloverleaf_nonretailer$numbe
summary(m16)
```

```
##
## Call:
## lm(formula = cloverleaf_nonretailer$revenue ~ as.factor(adrank_2) +
##     as.factor(cloverleaf_nonretailer$numberofwords) + as.factor(cloverleaf_nonretailer$brandname))
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -469.38 -169.51  -40.30   79.27 1416.33
##
## Coefficients:
##                                                 Estimate Std. Error
## (Intercept)                                       169.51      27.36
## as.factor(adrank_2)1                              180.30      32.30
## as.factor(cloverleaf_nonretailer$numberofwords)2 -248.78      33.12
## as.factor(cloverleaf_nonretailer$numberofwords)3 -337.68      36.56
## as.factor(cloverleaf_nonretailer$numberofwords)4 -323.37     107.05
## as.factor(cloverleaf_nonretailer$numberofwords)5 -469.38     347.24
## as.factor(cloverleaf_nonretailer$brandname)1       119.57      33.36
##                                                 t value Pr(>|t|)
## (Intercept)                                       6.195 1.05e-09 ***
## as.factor(adrank_2)1                              5.582 3.53e-08 ***
## as.factor(cloverleaf_nonretailer$numberofwords)2 -7.511 1.99e-13 ***
## as.factor(cloverleaf_nonretailer$numberofwords)3 -9.236  < 2e-16 ***
## as.factor(cloverleaf_nonretailer$numberofwords)4 -3.021 0.002622 **
## as.factor(cloverleaf_nonretailer$numberofwords)5 -1.352 0.176932
## as.factor(cloverleaf_nonretailer$brandname)1       3.584 0.000364 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 345.9 on 635 degrees of freedom
## Multiple R-squared:  0.2022, Adjusted R-squared:  0.1946
## F-statistic: 26.82 on 6 and 635 DF,  p-value: < 2.2e-16
```

```
m17 <- lm(cloverleaf_nonretailer$clickthroughrate ~ as.factor(adrank_2) + as.factor(cloverleaf_nonretail
summary(m17)
```

```
##
## Call:
## lm(formula = cloverleaf_nonretailer$clickthroughrate ~ as.factor(adrank_2) +
##     as.factor(cloverleaf_nonretailer$numberofwords) + as.factor(cloverleaf_nonretailer$brandname))
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.22247 -0.05501 -0.02123  0.05453  0.84735
```

```
##
## Coefficients:
##                                                 Estimate Std. Error
## (Intercept)                                     0.055013   0.008010
## as.factor(adrank_2)1                            0.129978   0.009456
## as.factor(cloverleaf_nonretailer$numberofwords)2 -0.081194   0.009696
## as.factor(cloverleaf_nonretailer$numberofwords)3 -0.079757   0.010703
## as.factor(cloverleaf_nonretailer$numberofwords)4 -0.130084   0.031337
## as.factor(cloverleaf_nonretailer$numberofwords)5  0.767596   0.101652
## as.factor(cloverleaf_nonretailer$brandname)1      0.047413   0.009767
##                                                 t value Pr(>|t|)
## (Intercept)                                       6.868 1.55e-11 ***
## as.factor(adrank_2)1                             13.745  < 2e-16 ***
## as.factor(cloverleaf_nonretailer$numberofwords)2  -8.374 3.56e-16 ***
## as.factor(cloverleaf_nonretailer$numberofwords)3  -7.452 3.03e-13 ***
## as.factor(cloverleaf_nonretailer$numberofwords)4  -4.151 3.76e-05 ***
## as.factor(cloverleaf_nonretailer$numberofwords)5   7.551 1.51e-13 ***
## as.factor(cloverleaf_nonretailer$brandname)1       4.855 1.52e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1013 on 635 degrees of freedom
## Multiple R-squared:  0.4489, Adjusted R-squared:  0.4437
## F-statistic: 86.21 on 6 and 635 DF,  p-value: < 2.2e-16
```

```
m18 <- lm(cloverleaf_nonretailer$conversionrate ~ as.factor(adrank_2) + as.factor(cloverleaf_nonretailer
summary(m18)
```

```
##
## Call:
## lm(formula = cloverleaf_nonretailer$conversionrate ~ as.factor(adrank_2) +
##     as.factor(cloverleaf_nonretailer$numberofwords) + as.factor(cloverleaf_nonretailer$brandname))
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.09840 -0.02378 -0.00909 -0.00498  0.98130
##
## Coefficients:
##                                                 Estimate Std. Error
## (Intercept)                                     0.004984   0.007774
## as.factor(adrank_2)1                            0.011803   0.009178
## as.factor(cloverleaf_nonretailer$numberofwords)2  0.013713   0.009411
## as.factor(cloverleaf_nonretailer$numberofwords)3  0.028399   0.010389
## as.factor(cloverleaf_nonretailer$numberofwords)4  0.091217   0.030416
## as.factor(cloverleaf_nonretailer$numberofwords)5 -0.007184   0.098663
## as.factor(cloverleaf_nonretailer$brandname)1     -0.009604   0.009479
##                                                 t value Pr(>|t|)
## (Intercept)                                       0.641  0.52165
## as.factor(adrank_2)1                              1.286  0.19890
## as.factor(cloverleaf_nonretailer$numberofwords)2   1.457  0.14557
## as.factor(cloverleaf_nonretailer$numberofwords)3   2.734  0.00644 **
## as.factor(cloverleaf_nonretailer$numberofwords)4   2.999  0.00281 **
## as.factor(cloverleaf_nonretailer$numberofwords)5  -0.073  0.94198
## as.factor(cloverleaf_nonretailer$brandname)1      -1.013  0.31139
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.09828 on 635 degrees of freedom
## Multiple R-squared:  0.02462,    Adjusted R-squared:  0.0154
## F-statistic: 2.671 on 6 and 635 DF,  p-value: 0.01442
```

```
## Does brand name matter when retailer is mentioned?
## It appears so, but it is most likely caused by the variability explained by another variable (shown
```

```
m19 <- lm(cloverleaf_retailer$clickthroughrate ~ as.factor(cloverleaf_retailer$brandname))
summary(m19)
```

```
##
## Call:
## lm(formula = cloverleaf_retailer$clickthroughrate ~ as.factor(cloverleaf_retailer$brandname))
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.41799 -0.06305  0.08201  0.13040  0.58201
##
## Coefficients:
##                                          Estimate Std. Error t value
## (Intercept)                               0.04302    0.04610   0.933
## as.factor(cloverleaf_retailer$brandname)1  0.37497    0.04740   7.911
##                                          Pr(>|t|)
## (Intercept)                                 0.351
## as.factor(cloverleaf_retailer$brandname)1 3.07e-14 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2061 on 366 degrees of freedom
## Multiple R-squared:  0.146,  Adjusted R-squared:  0.1437
## F-statistic: 62.58 on 1 and 366 DF,  p-value: 3.069e-14
```

```
## Does brand name matter when retailer is not mentioned?
## It appears so, but it is most likely caused by the variability explained by another variable (shown
```

```
m20 <- lm(cloverleaf_nonretailer$clickthroughrate ~ as.factor(cloverleaf_nonretailer$brandname))
summary(m20)
```

```
##
## Call:
## lm(formula = cloverleaf_nonretailer$clickthroughrate ~ as.factor(cloverleaf_nonretailer$brandname))
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.12053 -0.08053 -0.02490  0.01787  0.87947
##
## Coefficients:
##                                      Estimate Std. Error t value
## (Intercept)                          0.024897   0.008138   3.059
```

```
## as.factor(cloverleaf_nonretailer$brandname)1 0.095630   0.010362   9.229
##                                                Pr(>|t|)
## (Intercept)                                     0.00231 **
## as.factor(cloverleaf_nonretailer$brandname)1  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1276 on 640 degrees of freedom
## Multiple R-squared:  0.1175, Adjusted R-squared:  0.1161
## F-statistic: 85.17 on 1 and 640 DF,  p-value: < 2.2e-16
```

```
## More inspection of brandname?

m21 <- lm(cloverleaf_retailer$conversionrate ~ as.factor(cloverleaf_retailer$brandname))
summary(m21)
```

```
##
## Call:
## lm(formula = cloverleaf_retailer$conversionrate ~ as.factor(cloverleaf_retailer$brandname))
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.06540 -0.06540 -0.06540 -0.05044  0.93460
##
## Coefficients:
##                                             Estimate Std. Error t value
## (Intercept)                                 0.003125   0.051684   0.060
## as.factor(cloverleaf_retailer$brandname)1 0.062277   0.053149   1.172
##                                             Pr(>|t|)
## (Intercept)                                    0.952
## as.factor(cloverleaf_retailer$brandname)1    0.242
##
## Residual standard error: 0.2311 on 366 degrees of freedom
## Multiple R-squared:  0.003737,   Adjusted R-squared:  0.001015
## F-statistic: 1.373 on 1 and 366 DF,  p-value: 0.2421
```

```
m22 <- lm(cloverleaf_nonretailer$conversionrate ~ as.factor(cloverleaf_nonretailer$brandname))
summary(m22)
```

```
##
## Call:
## lm(formula = cloverleaf_nonretailer$conversionrate ~ as.factor(cloverleaf_nonretailer$brandname))
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.01876 -0.01876 -0.01704 -0.01704  0.98296
##
## Coefficients:
##                                                Estimate Std. Error t value
## (Intercept)                                    0.017039   0.006320   2.696
## as.factor(cloverleaf_nonretailer$brandname)1 0.001723   0.008047   0.214
##                                                Pr(>|t|)
## (Intercept)                                      0.0072 **
```

```
## as.factor(cloverleaf_nonretailer$brandname)1   0.8306
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.09912 on 640 degrees of freedom
## Multiple R-squared:  7.16e-05,   Adjusted R-squared:  -0.001491
## F-statistic: 0.04583 on 1 and 640 DF,  p-value: 0.8306
```

```
## Future considerations:
## Time series, landing page, log landing page
```