Emotion features

R. Cassano-Coleman

2025-01-13

This script looks at the correlation between emotional feature differences and similarity ratings in the expanded corpus.
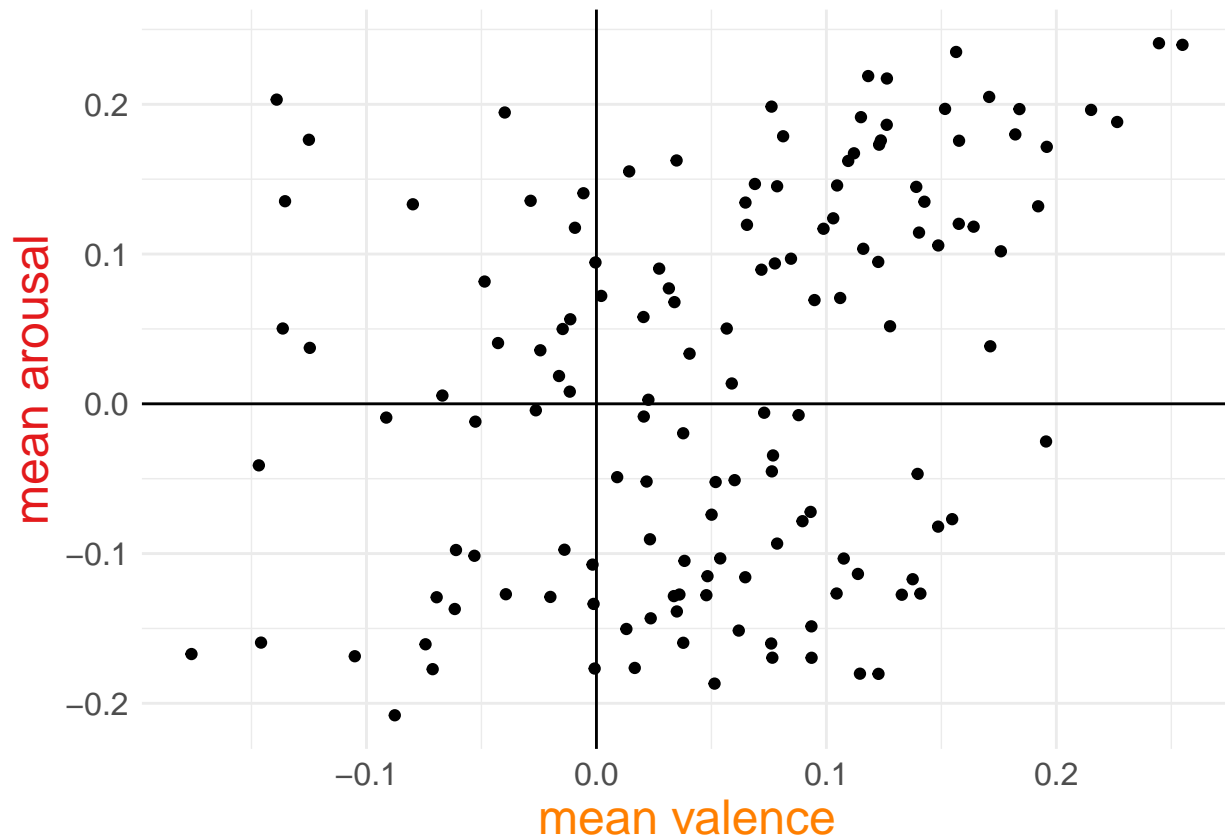
First, load the features and feature differences.

Keep only emotional features.

Colors:

- Valence: orange #FF7F00
- Arousal: red #E31A1C

Is there a relationship between valence and arousal in this dataset?
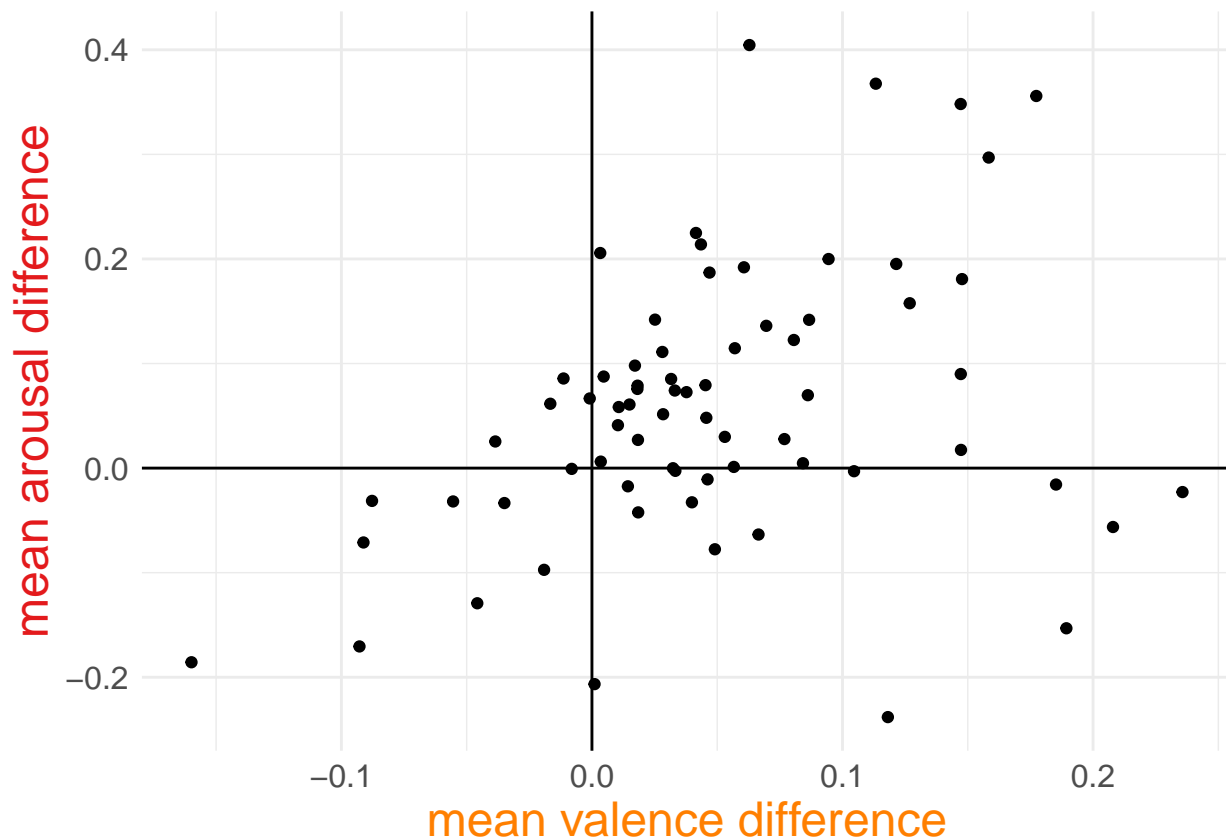


```
## 
## Call:
## lm(formula = mean_arousal ~ mean_valence, data = emot_feat)
## 
## Residuals:
##      Min      1Q   Median      3Q      Max
## -0.23110 -0.11233  0.02644  0.10062  0.28276
## 
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -0.01035    0.01196  -0.865    0.388
## mean_valence  0.49825    0.11515   4.327 2.88e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 
## Residual standard error: 0.1217 on 138 degrees of freedom
## Multiple R-squared:  0.1195, Adjusted R-squared:  0.1131
## F-statistic: 18.72 on 1 and 138 DF,  p-value: 2.883e-05
```

What's the correlation between valence and arousal?

```
## [1] 0.3456295
```

Is there a correlation between valence difference and arousal difference?

This means that a majority of the originals are more positive (+ valence difference) and higher arousal (+ arousal difference) than covers. This makes sense, as covers tend to be more laid back than originals.

```
##
## Call:
## lm(formula = mean_arousal_difference ~ mean_valence_difference,
##     data = emot_feat_diffs)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.33896 -0.06582  0.00642  0.06016  0.33785
##
## Coefficients:
##                         Estimate Std. Error t value Pr(>|t|)
## (Intercept)              0.02787    0.01771   1.574  0.12013
## mean_valence_difference  0.61781    0.20245   3.052  0.00324 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1238 on 68 degrees of freedom
## Multiple R-squared:  0.1205, Adjusted R-squared:  0.1075
## F-statistic: 9.313 on 1 and 68 DF,  p-value: 0.003244

## [1] 0.3470747
```

The correlation between differences is about the same as between the actual valence and arousal (r = 0.34).

Do emotional feature differences predict similarity?

```
##
## Call:
## lm(formula = mean_sim ~ mean_valence_difference, data = emot_feat_diffs_scaled)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.6144 -0.5870 -0.1789  0.6614  1.8885
##
## Coefficients:
##                           Estimate Std. Error t value Pr(>|t|)
## (Intercept)             -4.363e-16  1.004e-01   0.000        1
## mean_valence_difference -5.519e-01  1.011e-01  -5.457 7.35e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.84 on 68 degrees of freedom
## Multiple R-squared:  0.3045, Adjusted R-squared:  0.2943
## F-statistic: 29.78 on 1 and 68 DF,  p-value: 7.348e-07

##
## Call:
## lm(formula = mean_sim ~ mean_arousal_difference, data = emot_feat_diffs_scaled)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.91540 -0.48391 -0.00222  0.78581  1.78578
##
## Coefficients:
##                           Estimate Std. Error t value Pr(>|t|)
## (Intercept)             -3.633e-16  1.075e-01    0.00        1
## mean_arousal_difference -4.504e-01  1.083e-01   -4.16 9.13e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.8994 on 68 degrees of freedom
## Multiple R-squared:  0.2029, Adjusted R-squared:  0.1912
## F-statistic: 17.31 on 1 and 68 DF,  p-value: 9.131e-05

##
## Call:
## lm(formula = mean_sim ~ mean_valence_difference + mean_arousal_difference,
##     data = emot_feat_diffs_scaled)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.5696 -0.5582  0.1276  0.5399  1.6711
##
## Coefficients:
##                           Estimate Std. Error t value Pr(>|t|)
## (Intercept)             -3.975e-16  9.439e-02   0.000  1.00000
## mean_valence_difference -4.584e-01  9.959e-02  -4.603 1.91e-05 ***
## mean_arousal_difference -3.139e-01  9.959e-02  -3.152  0.00242 **
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.7898 on 67 degrees of freedom
## Multiple R-squared:  0.3944, Adjusted R-squared:  0.3763
## F-statistic: 21.81 on 2 and 67 DF,  p-value: 5.058e-08

##
## Call:
## lm(formula = mean_sim ~ mean_valence_difference * mean_arousal_difference,
##     data = emot_feat_diffs_scaled)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.5637 -0.5600  0.1167  0.5410  1.6625
##
## Coefficients:
##                                                  Estimate Std. Error t value
## (Intercept)                                     -0.004788   0.099450  -0.048
## mean_valence_difference                         -0.457419   0.100494  -4.552
## mean_arousal_difference                         -0.322292   0.112450  -2.866
## mean_valence_difference:mean_arousal_difference  0.016315   0.099276   0.164
##                                                  Pr(>|t|)
## (Intercept)                                       0.96174
## mean_valence_difference                          2.35e-05 ***
## mean_arousal_difference                           0.00557 **
## mean_valence_difference:mean_arousal_difference   0.86997
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.7956 on 66 degrees of freedom
## Multiple R-squared:  0.3946, Adjusted R-squared:  0.3671
## F-statistic: 14.34 on 3 and 66 DF,  p-value: 2.699e-07
```

Is this interaction worth including in the model?

```
## Analysis of Variance Table
##
## Model 1: mean_sim ~ mean_valence_difference + mean_arousal_difference
## Model 2: mean_sim ~ mean_valence_difference * mean_arousal_difference
##   Res.Df    RSS Df Sum of Sq     F Pr(>F)
## 1     67 41.789
## 2     66 41.771  1  0.017093 0.027   0.87
```

Not worth it!

Plot valence and arousal differences with similarity as color.