Correlation

Pradeep Paladugula

2020-03-22

*Title*: Big Data Analytics Services for Enhancing Business Intelligence

*Abstract*: This article examines how to use big data analytics services to enhance business intelligence (BI). More specifically, this article proposes an ontology of big data analytics and presents a big data analytics service-oriented architecture (BASOA), and then applies BASOA to BI, where our surveyed data analysis shows that the proposed BASOA is viable for enhancing BI and enterprise information systems. This article also explores temporality, expectability, and relativity as the characteristics of intelligence in BI. These characteristics are what customers and decision makers expect from BI in terms of systems, products, and services of organizations. The proposed approach in this article might facilitate the research and development of business analytics, big data analytics, and BI as well as big data science and big data computing.

# Dataset:

data <- read.csv('07\_data.csv')

- Gender of the participant surveyed on these topics  
- Temporality: an average score of the rated ability to adapt to change over time 1 (not changing) to 7 (changing a lot)  
- Expectability: a rated degree of satisfaction with the BI  
- Relativity: average score rating of how much better one system is than another in BI 1 (not very good) to 7 (very good)  
- Positive emotion: how positive participants felt about BI (higher scores are more positive, ranges from 1 to 7).

summary(data)

## gender temporality expectability relativity positive   
## men :150 Min. :1.737 Min. :0.000 Min. :-2.301 Min. :-0.9128   
## women:150 1st Qu.:2.823 1st Qu.:2.000 1st Qu.: 2.439 1st Qu.: 2.0175   
## Median :3.581 Median :3.000 Median : 3.564 Median : 3.0780   
## Mean :3.532 Mean :3.643 Mean : 3.569 Mean : 3.1446   
## 3rd Qu.:4.225 3rd Qu.:5.000 3rd Qu.: 4.731 3rd Qu.: 4.2614   
## Max. :5.184 Max. :9.000 Max. :10.508 Max. : 8.2126   
## NA's :9 NA's :9 NA's :9 NA's :9

# Data Screening:

## Accuracy:

a. Include output that indicates if the data are or are not accurate.  
b. If the data are not accurate, delete the inaccurate scores.  
c. Include a summary that shows that you fixed the inaccurate scores.

data$temporality[data$temporality < 1 | data$temporality > 7] = NA  
data$relativity[data$relativity < 1 | data$relativity > 7] = NA  
data$positive[data$positive < 1 | data$positive > 7] = NA  
summary(data)

## gender temporality expectability relativity positive   
## men :150 Min. :1.737 Min. :0.000 Min. :1.065 Min. :1.053   
## women:150 1st Qu.:2.823 1st Qu.:2.000 1st Qu.:2.625 1st Qu.:2.337   
## Median :3.581 Median :3.000 Median :3.618 Median :3.250   
## Mean :3.532 Mean :3.643 Mean :3.650 Mean :3.423   
## 3rd Qu.:4.225 3rd Qu.:5.000 3rd Qu.:4.676 3rd Qu.:4.396   
## Max. :5.184 Max. :9.000 Max. :6.952 Max. :6.918   
## NA's :9 NA's :9 NA's :44 NA's :41

## Missing:

a. Since any accuracy errors will create more than 5% missing data, exclude all data pairwise for the rest of the analyses.

cleanData <- na.omit(data)  
summary(cleanData)

## gender temporality expectability relativity positive   
## men :107 Min. :1.737 Min. :0.00 Min. :1.065 Min. :1.053   
## women:100 1st Qu.:2.865 1st Qu.:2.00 1st Qu.:2.634 1st Qu.:2.284   
## Median :3.522 Median :3.00 Median :3.591 Median :3.193   
## Mean :3.511 Mean :3.57 Mean :3.661 Mean :3.395   
## 3rd Qu.:4.188 3rd Qu.:5.00 3rd Qu.:4.684 3rd Qu.:4.378   
## Max. :5.102 Max. :8.00 Max. :6.952 Max. :6.918

## Outliers:

a. Include a summary of your mahal scores.  
b. What are the df for your Mahalanobis cutoff?  
c. What is the cut off score for your Mahalanobis measure?  
d. How many outliers did you have?

noMiss <- cleanData[,-c(1)]  
mahalScores = mahalanobis(noMiss, colMeans(noMiss, na.rm = TRUE), cov(noMiss, use = "pairwise.complete.obs"))  
mahalScores

## 3 4 5 6 7 8 9   
## 4.9995676 7.7528981 3.4928848 2.8482000 2.1121758 9.0196110 2.7785428   
## 12 13 16 17 18 19 20   
## 4.0151669 9.6081393 8.1755024 5.2475559 11.2900110 2.9232378 3.3700619   
## 22 24 28 29 30 31 32   
## 9.4295298 7.9294934 2.4968146 3.6749320 4.2856946 8.3638500 2.2227511   
## 34 37 38 39 40 42 43   
## 6.7993885 1.3100589 0.3517478 8.6130777 9.9905082 3.6773014 6.1931893   
## 44 45 46 49 50 51 52   
## 3.4494180 3.5773830 0.8675374 0.3334311 3.6147266 9.3815495 3.3335206   
## 53 55 56 59 60 62 63   
## 2.4320038 3.2579986 1.8985941 2.3810604 5.9115020 8.6528501 0.5204975   
## 64 65 67 68 69 70 71   
## 2.8870370 3.7149859 1.9060051 1.7955915 6.8169888 0.8373976 5.8860736   
## 73 74 75 76 77 78 81   
## 5.9871045 3.7436519 1.2377745 3.6873824 3.6379470 6.1606333 1.7554671   
## 82 83 85 86 88 90 97   
## 6.5803737 8.4432518 4.3106671 1.7894347 5.7204345 4.1863734 2.4839408   
## 98 100 101 102 104 105 106   
## 4.4697295 5.1250418 2.4749193 2.2301882 4.0613730 4.1151160 4.6082655   
## 107 108 110 112 113 114 116   
## 2.7759276 4.1529765 8.1095076 3.5660510 3.8097153 4.6284445 1.8070772   
## 118 119 121 122 124 126 127   
## 10.5342902 5.7619392 4.1070443 3.4844882 4.5348788 3.9246492 5.5076890   
## 129 130 131 132 133 135 137   
## 2.4983191 4.4431081 5.7777735 2.3261948 1.2202746 3.2217760 4.6863715   
## 139 142 143 144 146 147 148   
## 5.8057381 5.4417299 1.6583857 1.9234562 2.3883633 8.5028548 2.3913462   
## 149 150 151 152 153 154 155   
## 1.9439762 1.7205863 6.3168498 1.4229752 5.7672969 1.7062730 6.9954663   
## 157 158 159 160 162 163 164   
## 6.3092421 1.5238410 8.5513133 2.9414767 1.4604493 1.5367560 4.6798505   
## 166 167 171 172 173 174 175   
## 1.2078748 10.1958360 2.6983768 0.1628773 0.9232773 1.9758751 5.3312994   
## 178 179 180 181 182 183 184   
## 2.3860644 1.9442042 3.7431914 2.6740206 3.0777747 5.3991064 7.5768331   
## 187 188 190 191 194 195 196   
## 3.1434331 1.2793994 3.9667099 7.4172474 2.7536696 3.2561614 5.7222224   
## 198 199 201 202 203 204 206   
## 2.2749307 4.7663118 3.2205634 5.6399895 2.9426663 1.4221475 4.9542651   
## 207 208 209 210 212 213 216   
## 0.5290791 2.3579975 5.5376682 2.6545010 5.6929771 2.7118265 1.0298209   
## 217 218 220 221 222 223 224   
## 5.0186512 3.0412746 1.1129561 6.9554058 10.5638923 4.1655858 2.1790288   
## 225 226 227 229 230 231 232   
## 2.2113229 2.1417774 1.8637116 3.6151409 4.9296574 0.9738669 3.5685284   
## 233 234 235 236 239 240 241   
## 1.5043300 2.7375489 3.8930982 4.8382047 3.5167003 3.5008299 3.9180309   
## 244 246 248 249 250 251 252   
## 3.6101413 4.3154217 2.6978315 2.9742802 5.5079217 7.6045754 1.3583994   
## 255 256 257 258 259 260 261   
## 3.3696235 3.5574126 1.7051520 3.7657469 2.5077374 4.7538616 3.6087123   
## 262 263 265 268 269 270 271   
## 2.3318945 3.8043452 2.2711964 2.9097699 3.2509074 2.4744362 3.1190776   
## 274 275 276 278 279 280 282   
## 4.1625853 4.7214993 5.0006462 5.1647829 0.4777734 1.4532629 1.8331003   
## 283 286 287 288 289 290 292   
## 0.7790014 6.9994368 6.0580823 6.6685918 3.3198625 1.2496108 4.6216365   
## 294 295 296 298   
## 2.4496549 4.7525092 6.5243098 2.3416289

cutOff <- qchisq(1-.001, ncol(noMiss))  
cutOff

## [1] 18.46683

summary(mahalScores < cutOff)

## Mode TRUE   
## logical 207

noOut <- subset(cleanData, mahalScores < cutOff)  
noOut

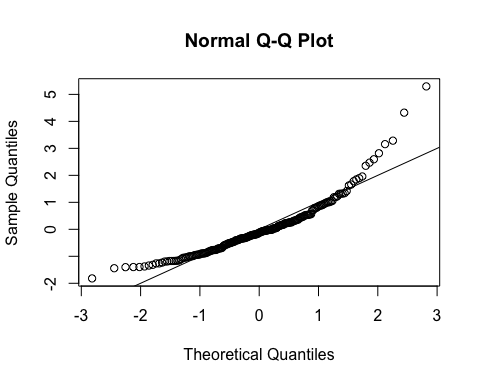
## gender temporality expectability relativity positive  
## 3 women 4.586402 3 2.572868 1.223659  
## 4 women 4.331763 1 1.347142 1.795185  
## 5 women 4.259832 2 2.446816 2.202066  
## 6 women 4.660960 5 3.775775 2.552891  
## 7 women 4.036579 5 5.454099 3.019725  
## 8 women 4.530757 0 4.198125 2.614516  
## 9 women 3.208714 5 2.967409 4.861072  
## 12 women 4.165093 6 4.306817 4.648106  
## 13 women 3.238961 0 4.171403 5.985619  
## 16 women 4.571096 5 1.432909 1.774882  
## 17 women 3.982522 7 3.970333 2.305769  
## 18 women 4.938608 0 4.389296 2.700692  
## 19 women 4.549474 5 3.377292 2.378240  
## 20 women 4.369895 5 4.693000 1.380504  
## 22 women 5.102181 7 3.414720 1.606853  
## 24 women 4.386829 2 6.226048 4.115312  
## 28 women 4.172433 3 5.044878 3.939312  
## 29 women 3.956321 5 3.369465 1.156475  
## 30 women 4.711761 2 3.904243 2.516802  
## 31 women 4.453575 3 6.823700 1.903434  
## 32 women 3.902674 3 5.126074 2.473248  
## 34 women 3.774439 5 6.890952 1.996212  
## 37 women 3.756984 5 4.735555 2.660313  
## 38 women 3.580643 4 3.427767 2.766761  
## 39 women 4.880203 5 1.708481 3.702972  
## 40 women 4.497580 8 5.794069 1.332978  
## 42 women 3.357322 3 2.929358 1.053080  
## 43 women 4.489371 1 4.012190 4.108021  
## 44 women 4.510224 6 4.307027 3.368739  
## 45 women 4.107657 6 4.572920 1.766861  
## 46 women 3.532016 5 4.294086 3.482421  
## 49 women 3.697138 3 3.931719 3.546228  
## 50 women 4.429213 2 3.470736 1.941459  
## 51 women 4.501909 0 2.048414 4.014267  
## 52 women 3.827140 2 5.053673 2.672276  
## 53 women 4.017592 6 4.237871 3.008700  
## 55 women 4.575270 3 5.000347 2.568674  
## 56 women 4.351336 4 5.061722 2.861471  
## 59 women 3.942428 2 3.759540 2.082552  
## 60 women 3.677005 2 5.420026 5.265431  
## 62 women 4.814168 6 2.093971 1.992358  
## 63 women 3.923024 3 3.638681 3.072078  
## 64 women 4.230702 3 5.229934 2.386120  
## 65 women 4.304208 6 4.446851 4.264415  
## 67 women 4.161915 5 3.176782 3.078182  
## 68 women 3.965989 2 3.154534 3.717381  
## 69 women 3.625076 2 3.567371 6.496277  
## 70 women 3.687233 5 4.262741 3.502236  
## 71 women 4.847632 3 5.849361 3.078863  
## 73 women 4.478682 1 3.302405 1.780330  
## 74 women 4.379877 3 3.479904 4.891106  
## 75 women 4.196348 3 3.940425 2.645100  
## 76 women 4.850326 3 4.403256 2.836303  
## 77 women 4.534421 3 2.627935 1.929732  
## 78 women 3.751175 7 2.983053 3.529162  
## 81 women 3.664787 5 4.769536 4.388264  
## 82 women 4.825495 6 2.832920 3.352723  
## 83 women 4.796455 4 1.601453 4.412049  
## 85 women 4.317156 3 2.636163 1.119419  
## 86 women 3.994235 4 3.611095 1.663199  
## 88 women 3.611191 5 1.334881 4.056140  
## 90 women 3.944988 6 4.291311 4.961690  
## 97 women 4.219107 3 2.589417 3.997488  
## 98 women 4.269762 6 5.605873 1.819293  
## 100 women 4.450653 4 3.438885 5.382597  
## 101 women 3.930155 6 4.812154 3.020127  
## 102 women 4.064349 3 4.853484 2.164687  
## 104 women 3.854926 1 4.018477 2.433384  
## 105 women 3.769718 4 5.294288 5.446013  
## 106 women 3.722240 6 6.140299 3.965083  
## 107 women 4.241797 6 4.772485 3.192741  
## 108 women 4.328969 2 3.879007 4.646513  
## 110 women 4.743356 1 1.771300 2.589870  
## 112 women 4.602453 6 4.553060 2.942572  
## 113 women 3.747071 3 5.952156 3.062766  
## 114 women 4.458603 4 6.146362 2.246621  
## 116 women 3.699054 4 5.403625 3.818490  
## 118 women 4.406552 8 6.952331 2.437226  
## 119 women 4.179697 6 2.852232 1.400704  
## 121 women 4.052246 2 1.668263 2.358009  
## 122 women 3.520630 4 5.317935 5.238084  
## 124 women 3.803345 1 1.731494 2.586268  
## 126 women 3.637311 1 2.632513 2.028701  
## 127 women 4.940459 5 4.825002 4.469485  
## 129 women 4.533000 4 5.122298 2.694319  
## 130 women 3.915179 6 6.203989 3.148667  
## 131 women 4.437800 2 5.324121 1.846499  
## 132 women 4.201867 2 3.033992 3.108088  
## 133 women 3.370932 4 4.925695 2.921602  
## 135 women 3.837700 5 4.943528 1.438993  
## 137 women 3.927346 6 5.796859 1.865389  
## 139 women 4.386522 2 1.414181 3.492809  
## 142 women 4.307427 7 5.113893 2.161609  
## 143 women 4.320685 4 4.713258 3.690872  
## 144 women 4.210081 5 3.455504 3.674238  
## 146 women 4.409023 3 4.286366 3.954879  
## 147 women 3.927122 8 4.289237 2.033583  
## 148 women 4.238844 3 2.375924 2.846238  
## 149 women 3.507742 5 2.726591 3.761112  
## 150 women 3.713487 3 2.452957 2.200834  
## 151 men 2.336727 2 5.294903 3.772522  
## 152 men 2.726785 3 3.069251 4.496115  
## 153 men 3.106870 3 1.520510 1.180436  
## 154 men 2.534931 3 2.857508 3.847856  
## 155 men 3.815084 2 6.271164 2.500474  
## 157 men 1.989499 4 4.668152 2.936548  
## 158 men 2.690661 3 3.093465 4.529629  
## 159 men 2.793906 6 5.175555 6.302868  
## 160 men 3.100414 2 2.054146 2.179032  
## 162 men 3.225676 5 4.675207 3.505386  
## 163 men 3.396440 4 2.356302 2.766210  
## 164 men 2.237072 3 2.507146 2.242218  
## 166 men 2.956398 4 3.387720 2.719957  
## 167 men 2.128099 0 2.504722 1.623457  
## 171 men 3.409561 3 3.849424 5.523582  
## 172 men 3.521942 3 3.334153 3.498738  
## 173 men 3.591058 4 4.358191 2.316429  
## 174 men 3.074163 4 3.794276 5.188459  
## 175 men 2.786789 3 1.484051 5.677937  
## 178 men 3.068317 4 5.371776 3.528376  
## 179 men 3.556170 5 5.361251 3.249898  
## 180 men 2.392952 2 3.695500 5.093878  
## 181 men 2.816046 2 2.237057 2.545074  
## 182 men 3.168144 3 3.860453 5.712322  
## 183 men 3.300704 3 1.168180 1.607464  
## 184 men 2.061124 6 3.112401 3.968311  
## 187 men 3.444325 5 5.115539 1.845404  
## 188 men 3.050001 4 4.664999 3.232131  
## 190 men 3.991960 3 2.512196 5.232465  
## 191 men 2.612402 1 1.197637 1.666395  
## 194 men 2.816316 2 1.739012 4.215709  
## 195 men 3.221441 3 3.527149 5.813175  
## 196 men 2.238792 4 5.359369 3.461063  
## 198 men 3.326545 3 2.307952 1.969035  
## 199 men 2.728919 2 1.065340 4.761410  
## 201 men 3.009388 2 1.754372 2.372983  
## 202 men 1.822843 3 2.431565 3.109280  
## 203 men 2.365260 3 4.251906 3.829847  
## 204 men 4.018225 4 3.624222 4.416546  
## 206 men 2.559531 2 1.322553 2.396667  
## 207 men 3.069821 4 3.408699 3.415337  
## 208 men 2.932070 3 2.510986 2.133501  
## 209 men 2.900932 4 5.919196 2.226729  
## 210 men 2.408588 3 2.702978 2.996351  
## 212 men 2.599196 3 1.478144 1.765139  
## 213 men 2.893652 5 4.982687 3.327594  
## 216 men 2.828903 4 3.584666 3.443160  
## 217 men 2.693676 5 3.019242 1.746532  
## 218 men 2.653215 3 1.616944 4.095464  
## 220 men 3.415689 3 2.693594 2.419690  
## 221 men 3.034004 3 2.945140 6.917850  
## 222 men 1.737005 2 1.166226 1.729102  
## 223 men 3.330804 3 4.492552 5.908982  
## 224 men 3.434845 3 3.564123 5.313472  
## 225 men 2.500352 3 3.361036 2.953067  
## 226 men 2.861364 4 2.884833 4.930289  
## 227 men 3.477868 4 5.004304 2.261938  
## 229 men 2.673674 2 4.475379 2.856397  
## 230 men 2.579059 5 4.132823 5.673863  
## 231 men 2.976619 3 3.077592 4.450620  
## 232 men 2.319537 4 4.320963 3.221564  
## 233 men 2.581890 3 3.097525 3.995890  
## 234 men 2.708071 2 1.763934 3.899858  
## 235 men 2.775622 2 3.807071 5.604616  
## 236 men 2.898672 1 1.443664 2.470409  
## 239 men 2.870655 2 4.725485 4.632785  
## 240 men 2.422825 5 3.636333 3.527521  
## 241 men 2.480026 2 2.362665 5.384554  
## 244 men 2.623366 4 3.999557 2.006255  
## 246 men 2.217544 2 3.182469 2.660570  
## 248 men 2.755291 3 4.807695 3.077961  
## 249 men 3.532593 6 3.262747 3.423234  
## 250 men 2.499757 5 4.448056 1.972448  
## 251 men 3.204959 4 1.550434 6.255282  
## 252 men 3.356991 3 3.129086 2.057760  
## 255 men 2.907395 4 5.579760 4.059044  
## 256 men 3.627211 5 2.330340 4.558613  
## 257 men 2.889013 4 2.690931 4.436595  
## 258 men 3.388204 4 3.511552 5.936271  
## 259 men 2.567475 4 2.650998 3.075368  
## 260 men 3.449674 5 3.424650 5.920239  
## 261 men 2.757023 2 2.867790 1.880720  
## 262 men 2.476768 2 3.108310 3.861350  
## 263 men 2.836130 4 1.534532 4.294841  
## 265 men 2.868961 5 3.992758 4.494895  
## 268 men 3.143223 1 3.542089 3.165685  
## 269 men 3.085682 2 4.202921 5.218222  
## 270 men 3.355050 3 2.218999 4.954981  
## 271 men 2.911627 2 1.446835 4.038551  
## 274 men 2.905948 2 3.868879 5.760789  
## 275 men 2.905350 2 2.164807 5.869336  
## 276 men 2.703809 3 1.410999 2.005636  
## 278 men 3.325057 2 4.459036 1.151065  
## 279 men 3.472604 4 3.744902 4.233559  
## 280 men 3.817198 5 3.475110 4.035084  
## 282 men 2.519300 3 3.268659 4.368698  
## 283 men 3.481826 3 2.510573 3.576451  
## 286 men 3.621432 4 6.878332 2.128341  
## 287 men 2.672530 3 4.852248 6.190165  
## 288 men 2.375195 2 2.614499 6.507269  
## 289 men 2.885393 4 2.232491 5.147775  
## 290 men 3.179726 4 3.085760 4.639073  
## 292 men 2.566394 5 3.591400 5.539623  
## 294 men 2.794594 4 4.190815 5.107869  
## 295 men 2.281292 2 4.191726 5.108889  
## 296 men 2.648135 3 5.229720 1.622311  
## 298 men 2.773299 4 2.159914 3.692774

# Assumptions:

## Linearity:

a. Include a picture that shows how you might assess multivariate linearity.  
b. Do you think you've met the assumption for linearity?

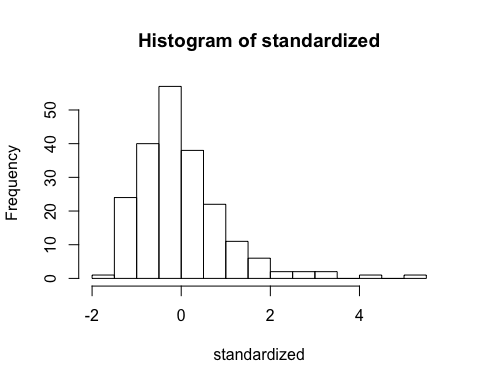
random = rchisq(nrow(noOut), 7)  
fake = lm(random~., data = noOut)  
standardized = rstudent(fake)  
fitvalues = scale(fake$fitted.values)  
qqnorm(standardized)  
abline(0,1)



## Normality:

a. Include a picture that shows how you might assess multivariate normality.  
b. Do you think you've met the assumption for normality?   
  
 From below picture it is we can confirm that the assumption has met the normality

hist(standardized, breaks=15)



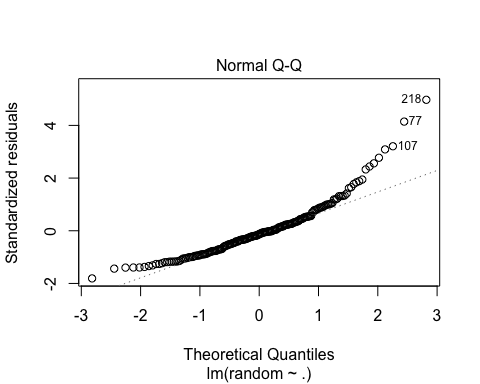
library(moments)  
skewness(noOut[,-c(1)], na.rm=TRUE)

## temporality expectability relativity positive   
## -0.03251117 0.27131476 0.11609442 0.42983894

kurtosis(noOut[,-c(1)], na.rm=TRUE)

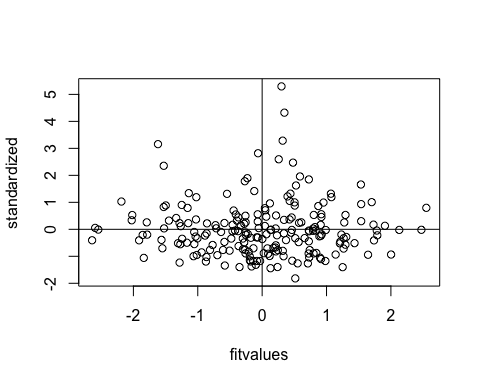
## temporality expectability relativity positive   
## 2.001327 2.894151 2.399474 2.331996

plot(fake, 2)

 ## Homogeneity and Homoscedasticity:

a. Include a picture that shows how you might assess multivariate homogeneity.  
b. Do you think you've met the assumption for homogeneity?  
c. Do you think you've met the assumption for homoscedasticity?  
  
 Both the assumption of homogeneity and homoscedasticity are met from looking at graph.

{plot(fitvalues, standardized)   
 abline(0,0, v = 0)}

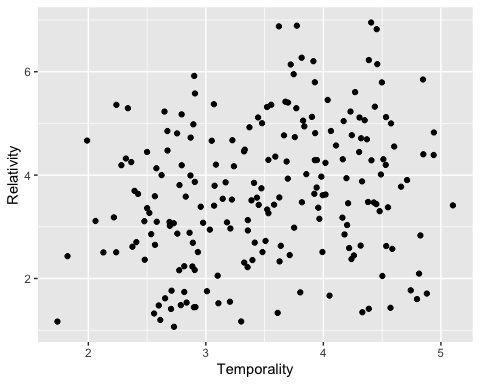


# Hypothesis Testing / Graphs:

Create a scatter plot of temporality and relativity.

a. Be sure to check x/y axis labels and length.  
b. What type of relationship do these two variables appear to have?  
  
 Positive relationship.

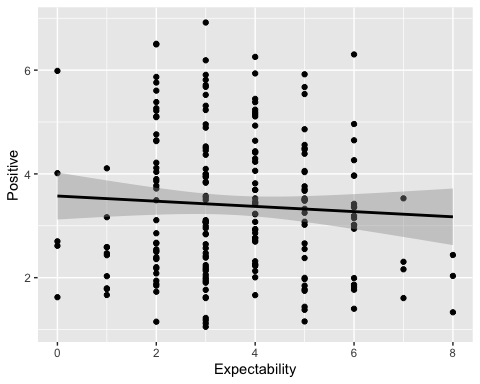
library(ggplot2)  
scatter = ggplot(cleanData, aes(temporality, relativity))  
scatter + geom\_point() + xlab("Temporality") + ylab("Relativity")



Create a scatter plot of expectability and positive emotion.

a. Include a linear line on the graph.   
b. Be sure to check x/y axis labels and length.  
c. What type of relationship do these two variables appear to have?  
  
 Expectability and Positive have no relationship.

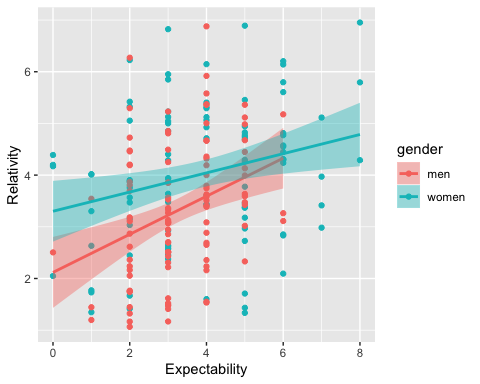
scatter2 = ggplot(cleanData, aes(expectability, positive))  
scatter2 + geom\_point() + geom\_smooth(method = "lm", color = "black") + xlab("Expectability") + ylab("Positive")



Create a scatter plot of expectability and relativity, grouping by gender.

a. Include a linear line on the graph.   
b. Be sure to check x/y axis labels and length.  
c. What type of relationship do these two variables appear to have for each group?  
  
 Men and women has postive relationship based on the below graph.

scatter3 = ggplot(cleanData, aes(expectability, relativity, color = gender))  
scatter3 + geom\_point() + geom\_smooth(method = "lm", aes(fill = gender)) + xlab("Expectability") + ylab("Relativity")



Include a correlation table of all of the variables (cor).

a. Include the output for Pearson.  
b. Include the output for Spearman.  
c. Include the output for Kendall.  
d. Which correlation was the strongest?  
e. For the correlations with gender, would point biserial or biserial be more appropriate? Why?

cor(noMiss, use="pairwise.complete.obs", method = "pearson")

## temporality expectability relativity positive  
## temporality 1.0000000 0.20209959 0.22754039 -0.25275929  
## expectability 0.2020996 1.00000000 0.32131190 -0.05948884  
## relativity 0.2275404 0.32131190 1.00000000 -0.03781882  
## positive -0.2527593 -0.05948884 -0.03781882 1.00000000

cor(noMiss, use="pairwise.complete.obs", method = "spearman")

## temporality expectability relativity positive  
## temporality 1.0000000 0.18839790 0.22410733 -0.25005006  
## expectability 0.1883979 1.00000000 0.31390642 -0.02067472  
## relativity 0.2241073 0.31390642 1.00000000 -0.01305998  
## positive -0.2500501 -0.02067472 -0.01305998 1.00000000

cor(noMiss, use="pairwise.complete.obs", method = "kendall")

## temporality expectability relativity positive  
## temporality 1.0000000 0.14863742 0.1507903 -0.16429811  
## expectability 0.1486374 1.00000000 0.2273461 -0.01704321  
## relativity 0.1507903 0.22734605 1.0000000 -0.01561840  
## positive -0.1642981 -0.01704321 -0.0156184 1.00000000

Calculate confidence interval for temporality and relativity.

cor.test(cleanData$temporality, cleanData$relativity, method = "pearson")

##   
## Pearson's product-moment correlation  
##   
## data: cleanData$temporality and cleanData$relativity  
## t = 3.3456, df = 205, p-value = 0.0009763  
## alternative hypothesis: true correlation is not equal to 0  
## 95 percent confidence interval:  
## 0.09408996 0.35295823  
## sample estimates:  
## cor   
## 0.2275404

Calculate the difference in correlations for 1) temporality and expectbility and 2) temporality and positive emotion.

a. Include the output from the test through Pearson's test.  
b. Is there a significant difference in their correlations?  
  
 Yes

cor.test(cleanData$temporality, cleanData$expectability, method = "pearson")

##   
## Pearson's product-moment correlation  
##   
## data: cleanData$temporality and cleanData$expectability  
## t = 2.9546, df = 205, p-value = 0.003497  
## alternative hypothesis: true correlation is not equal to 0  
## 95 percent confidence interval:  
## 0.06759244 0.32939147  
## sample estimates:  
## cor   
## 0.2020996

cor.test(cleanData$temporality, cleanData$positive, method = "pearson")

##   
## Pearson's product-moment correlation  
##   
## data: cleanData$temporality and cleanData$positive  
## t = -3.7404, df = 205, p-value = 0.0002384  
## alternative hypothesis: true correlation is not equal to 0  
## 95 percent confidence interval:  
## -0.3761634 -0.1205443  
## sample estimates:  
## cor   
## -0.2527593

Calculate the difference in correlations for gender on temporality and relativity.

a. Include the output from the test.  
b. Is there a significant difference in their correlations?  
  
 NO

library(cocor)  
women <- subset(cleanData, gender = "women")  
men <- subset(cleanData, gender = "men")  
subsetList <- list(women, men)  
cocor(~temporality + relativity | temporality + relativity, data = subsetList)

##   
## Results of a comparison of two correlations based on independent groups  
##   
## Comparison between r1.jk (temporality, relativity) = 0.2275 and r2.hm (temporality, relativity) = 0.2275  
## Difference: r1.jk - r2.hm = 0  
## Data: subsetList: j = temporality, k = relativity, h = temporality, m = relativity  
## Group sizes: n1 = 207, n2 = 207  
## Null hypothesis: r1.jk is equal to r2.hm  
## Alternative hypothesis: r1.jk is not equal to r2.hm (two-sided)  
## Alpha: 0.05  
##   
## fisher1925: Fisher's z (1925)  
## z = 0.0000, p-value = 1.0000  
## Null hypothesis retained  
##   
## zou2007: Zou's (2007) confidence interval  
## 95% confidence interval for r1.jk - r2.hm: -0.1831 0.1831  
## Null hypothesis retained (Interval includes 0)

Calculate the partial and semipartial correlations for all variables, and include the output. a. Are any of the correlations significant after controlling for all other relationships?

library(ppcor)

## Loading required package: MASS

pcor(cleanData[ , -c(1)], method = "pearson")

## $estimate  
## temporality expectability relativity positive  
## temporality 1.0000000 0.13170144 0.17590386 -0.24664382  
## expectability 0.1317014 1.00000000 0.28896009 -0.01557263  
## relativity 0.1759039 0.28896009 1.00000000 0.02450912  
## positive -0.2466438 -0.01557263 0.02450912 1.00000000  
##   
## $p.value  
## temporality expectability relativity positive  
## temporality 0.0000000000 5.978787e-02 1.164039e-02 0.0003637952  
## expectability 0.0597878683 0.000000e+00 2.646039e-05 0.8246126598  
## relativity 0.0116403917 2.646039e-05 0.000000e+00 0.7272215415  
## positive 0.0003637952 8.246127e-01 7.272215e-01 0.0000000000  
##   
## $statistic  
## temporality expectability relativity positive  
## temporality 0.000000 1.8929453 2.5459421 -3.6261593  
## expectability 1.892945 0.0000000 4.3005021 -0.2219028  
## relativity 2.545942 4.3005021 0.0000000 0.3493062  
## positive -3.626159 -0.2219028 0.3493062 0.0000000  
##   
## $n  
## [1] 207  
##   
## $gp  
## [1] 2  
##   
## $method  
## [1] "pearson"

# Theory:

- What are we using as our model for understanding the data in a correlational analysis?  
   
   
- How might we determine model fit?  
 The Primary tool to determine model fit using graphical residual analysis.  
   
- What is the difference between correlation and covariance?  
 Covariance: Indicates the direction of the linear relationship.  
 Correlation: Measures both the strength and direction of the linear relationship between two variables.  
   
- What is the difference between R and r?  
 r: Correlation r determines how well two variables are correlated with each other. Basically it is defined in numerical range between -1 to +1, Correlation value more nearest to 1 in either direction says two strong coefficient of determination.  
 R: R square is literally the square of correlation between x and y.  
   
- When would I want to use a nonparametric correlation over Pearson's correlation?  
   
   
- What is the distinction between semi-partial and partial correlations?   
 semi-partial correlations is same as partail correlations, But variation is:  
 Partial Correlations: Partial correlation measures the strength of a relationship between two variables, while controlling for the effect of one or more other variables.  
 Semi-partial correltions: the semi partial correlation statistic can tell us the particular part of variance, that a particular independent variable explains. It explains how one specific independent variable affects the dependent variable, while other variables are controlled for to prevent them getting in the way.