Gene Expression Data Analysis and Visualization
410.671
HW #3

1.) Load the golub data training set in the multtest library. Also load Biobase and annotate libraries, if they are not loaded with the multtest library. Remember that the golub data training set is in the multtest library, so see the help file for information on this data set (2.5 pts)

```
library(multtest)
library(Biobase)
library(annotate)
data(golub)
```

2.) Cast the matrix to a data frame and label the gene names as numbers (e.g. “g1”,”g2”,etc). (2.5 pts)

```
golub.df <- as.data.frame(golub)
rownames(golub.df) <- paste("g", rownames(golub.df), sep="")
```

3.) Get the sample labels (see lecture notes) and set the sample labels to the data frame. (2.5 pts)

```
colnames(golub.df) <- golub.cl
```

4.) Use the t-test function in the lecture #7 notes and modify it to "wilcox.test" instead of "t.test". Change the "$p.value" argument to "$statistic". Assign the following arguments to the function: (2.5 pts)
exact=F
alternative="two.sided"
correct=T
Run the function on all of the genes in the dataset and save it as "original.wmw.run"

```
wilcox.test.all.genes <- function(x,s1,s2){
  x1 <- x[s1]
  x2 <- x[s2]
  x1 <- as.numeric(x1)
  x2 <- as.numeric(x2)
  wmw.out <- wilcox.test(x1, x2, exact=F, alternative="two.sided", correct=T)
  out <- as.numeric(wmw.out$statistic)
  return(out)
}

original.wmw.run <- apply(golub.df, 1, wilcox.test.all.genes, s1=c(1:27), s2=c(28:38))
```

5.) Now write a for loop to iterate 500 times, where in each iteration, the columns of the data frame are shuffled (class labels mixed up), the WMW test is calculated on all of the genes, and the maximum test statistic (W) is saved in a list. (5 pts)

**Hints:**
Use sample() to sample the number of columns
Get the maximum test statistic across all genes with max()

```
> length(golub.df[colnames(golub.df)==0])
[1] 27
> length(golub.df[colnames(golub.df)==1])
[1] 11

max.test.stats <- c()
for (i in 1:500){
  x1 <- sample(colnames(golub.df),27)
  x2 <- sample(colnames(golub.df), 11)
  wmw.run <- apply(golub.df, 1, wilcox.test.all.genes, s1=x1, s2=x2)
  max.test.stats <- c(max.test.stats, max(wmw.run))
}
```

6.) Once you have the list of maximum test statistics, get the 95% value test statistic. Subset the original.wmw.run list of values with only those that have a higher test statistic than the 95% value that you calculated. Print the gene names and test statistics out. (5 pts)

```
test.stat.95 <- quantile(as.numeric(max.test.stats), 0.95)
original.95 <- original.wmw.run[original.wmw.run > test.stat.95]
> original.95
  g9  g10  g17  g23  g25  g39  g47  g48  g51  g54  g55  g59
 188  191  204  249  210  212  201  193  211  186  251  204
 g63  g66  g74  g77  g78  g79  g81  g83  g84  g89  g94  g96
 198  258  234  220  227  211  244  188  225  195  187  274
 g98 g102 g104 g106 g109 g112 g117 g120 g121 g124 g126 g127
 188  224  236  215  188  209  232  185  225  199  262  256
g128 g135 g138 g142 g148 g149 g151 g153 g154 g156 g158 g159
 216  244  233  216  213  205  208  235  197  196  246  191
g163 g167 g171 g172 g174 g177 g178 g182 g184 g187 g188 g189
 206  215  223  253  262  230  236  261  244  232  232  216
g192 g193 g194 g195 g201 g202 g204 g205 g206 g207 g210 g214
 241  220  215  228  188  255  263  228  221  246  202  211
g218 g220 g225 g226 g227 g228 g230 g232 g233 g235 g237 g239
 200  240  202  244  205  233  230  250  188  220  237  247
g240 g241 g242 g244 g246 g248 g250 g253 g254 g258 g259 g265
 206  202  186  236  254  244  189  266  213  233  266  188
g270 g271 g277 g282 g283 g285 g286 g288 g289 g290 g291 g295
 232  227  185  242  271  226  260  238  213  225  221  205
```

g296 g297 g298 g304 g307 g309 g312 g313 g314 g316 g318 g320
185 257 234 243 252 245 189 248 246 234 210 241
g323 g325 g327 g328 g329 g330 g331 g332 g333 g334 g335 g336
262 228 195 202 275 245 212 186 188 221 254 215
g337 g338 g342 g344 g345 g347 g357 g359 g360 g362 g363 g364
232 188 201 259 274 187 247 235 202 200 230 215
g365 g366 g368 g369 g370 g373 g376 g377 g385 g389 g391 g394
212 190 215 209 205 189 244 229 228 227 245 290
g395 g398 g399 g402 g407 g421 g422 g423 g426 g435 g441 g445
250 235 247 209 240 211 270 207 237 210 197 225
g446 g447 g449 g453 g454 g455 g460 g462 g464 g472 g478 g479
198 202 185 242 213 226 195 259 191 185 237 248
g486 g489 g490 g493 g494 g500 g504 g514 g515 g516 g520 g522
187 250 248 191 251 207 212 194 264 210 208 257
g523 g527 g529 g538 g542 g543 g544 g546 g551 g555 g557 g559
283 194 229 211 238 201 224 274 234 193 221 222
g560 g561 g563 g564 g567 g571 g573 g579 g590 g591 g603 g605
262 281 255 196 219 237 209 195 245 185 195 213
g610 g611 g615 g620 g621 g627 g634 g640 g645 g648 g651 g652
218 227 207 239 269 236 243 192 229 271 189 193
g655 g661 g663 g671 g675 g678 g679 g685 g686 g688 g689 g692
192 187 192 199 230 190 201 232 234 258 217 185
g693 g695 g696 g698 g701 g702 g703 g704 g710 g711 g713 g715
199 251 221 241 240 186 285 273 199 226 266 240
g717 g718 g720 g725 g726 g730 g733 g738 g742 g746 g753 g758
283 210 233 252 238 214 220 271 218 277 223 206
g763 g764 g770 g772 g776 g777 g785 g793 g797 g800 g801 g804
248 241 186 204 199 201 256 224 217 220 263 193
g807 g811 g828 g835 g837 g838 g839 g842 g846 g849 g852 g856
248 249 186 272 210 283 272 229 242 272 216 191
g858 g862 g864 g866 g867 g871 g877 g882 g885 g886 g889 g890
235 259 225 269 190 195 236 187 233 220 201 224
g892 g893 g906 g909 g912 g922 g924 g934 g936 g940 g949 g952
204 226 223 201 196 272 185 191 199 222 193 218
g958 g962 g963 g965 g966 g967 g970 g971 g974 g977 g978 g984
190 217 250 222 221 213 217 242 210 224 190 283
g985 g990 g997 g1003 g1006 g1007 g1011 g1014 g1016 g1018 g1019 g1020
213 209 232 201 275 202 243 259 237 197 254 195
g1021 g1025 g1027 g1030 g1037 g1040 g1042 g1045 g1047 g1057 g1058 g1060
214 244 220 244 281 206 284 270 191 228 217 264
g1061 g1070 g1072 g1075 g1078 g1081 g1083 g1086 g1094 g1101 g1105 g1108
224 260 191 223 222 249 185 278 244 258 215 191
g1109 g1110 g1117 g1126 g1133 g1140 g1145 g1146 g1161 g1162 g1163 g1167
239 242 214 216 211 200 262 193 245 261 207 199
g1173 g1174 g1175 g1176 g1177 g1179 g1181 g1182 g1186 g1192 g1193 g1197
195 240 239 216 185 199 227 197 200 196 234 189

g1198 g1199 g1202 g1203 g1206 g1221 g1222 g1223 g1225 g1228 g1231 g1238
226 203 256 210 237 225 193 205 254 208 221 223
g1245 g1252 g1253 g1254 g1255 g1258 g1262 g1270 g1271 g1273 g1278 g1282
245 202 255 185 220 196 206 186 268 209 218 203
g1285 g1288 g1289 g1290 g1292 g1293 g1295 g1298 g1310 g1316 g1317 g1318
208 235 237 232 194 255 218 229 221 255 229 207
g1324 g1327 g1329 g1330 g1334 g1337 g1340 g1341 g1347 g1348 g1351 g1352
191 267 238 208 266 244 202 228 223 241 207 222
g1353 g1354 g1355 g1359 g1360 g1364 g1368 g1369 g1372 g1377 g1381 g1382
192 212 205 244 233 220 279 209 198 209 248 200
g1388 g1390 g1392 g1399 g1410 g1421 g1425 g1426 g1428 g1435 g1437 g1440
220 234 195 195 197 207 224 215 224 208 187 203
g1441 g1445 g1453 g1455 g1456 g1459 g1466 g1468 g1469 g1470 g1474 g1478
224 263 254 266 261 250 198 245 202 243 250 202
g1480 g1481 g1484 g1489 g1491 g1501 g1502 g1505 g1508 g1509 g1513 g1515
242 216 237 242 242 233 189 202 215 215 241 214
g1519 g1524 g1531 g1538 g1542 g1543 g1547 g1552 g1554 g1557 g1559 g1560
218 282 226 185 266 210 243 189 193 209 260 188
g1561 g1562 g1564 g1565 g1566 g1568 g1571 g1579 g1585 g1586 g1594 g1598
195 202 246 208 210 191 226 245 267 190 238 275
g1604 g1605 g1606 g1608 g1610 g1613 g1616 g1619 g1620 g1621 g1629 g1631
218 207 187 190 249 220 258 205 202 236 245 222
g1633 g1636 g1638 g1640 g1641 g1642 g1644 g1648 g1649 g1653 g1658 g1660
219 218 262 265 224 265 197 234 246 259 214 192
g1661 g1663 g1666 g1668 g1671 g1679 g1681 g1684 g1688 g1690 g1691 g1695
196 206 191 222 254 215 192 223 192 200 266 189
g1696 g1698 g1699 g1701 g1702 g1707 g1709 g1712 g1714 g1718 g1719 g1723
250 215 231 235 215 209 198 217 189 229 217 256
g1724 g1725 g1726 g1731 g1732 g1733 g1739 g1741 g1742 g1744 g1745 g1752
196 190 207 187 264 198 188 196 227 190 218 225
g1756 g1765 g1766 g1770 g1773 g1775 g1777 g1779 g1786 g1787 g1791 g1793
191 194 252 221 247 199 230 189 209 245 200 216
g1795 g1804 g1805 g1806 g1807 g1809 g1811 g1812 g1813 g1817 g1818 g1821
227 187 249 189 241 231 284 193 214 272 205 222
g1827 g1828 g1832 g1834 g1845 g1851 g1855 g1856 g1857 g1859 g1860 g1866
217 193 199 290 205 231 229 263 211 201 210 209
g1869 g1873 g1876 g1880 g1881 g1882 g1883 g1887 g1890 g1900 g1903 g1908
272 234 216 208 220 260 281 227 213 196 253 216
g1909 g1916 g1917 g1919 g1920 g1921 g1926 g1929 g1931 g1932 g1934 g1938
275 273 201 209 274 190 248 191 207 229 230 231
g1939 g1943 g1944 g1945 g1947 g1948 g1949 g1951 g1955 g1956 g1959 g1962
268 239 249 223 254 252 200 190 240 191 272 227
g1963 g1969 g1971 g1973 g1978 g1979 g1980 g1981 g1985 g1986 g1991 g1992
250 227 188 221 271 225 229 238 232 210 221 203
g1993 g1995 g2000 g2002 g2006 g2012 g2018 g2019 g2020 g2027 g2032 g2039
246 287 195 283 191 206 199 222 265 190 239 214

g2040 g2045 g2050 g2052 g2061 g2063 g2065 g2067 g2069 g2072 g2073 g2076
201 200 224 240 241 210 237 189 217 219 201 186
g2077 g2079 g2080 g2083 g2087 g2088 g2092 g2096 g2097 g2105 g2107 g2108
215 244 231 196 247 185 187 236 186 252 185 196
g2110 g2118 g2122 g2132 g2143 g2146 g2149 g2150 g2155 g2156 g2162 g2164
249 228 270 238 217 218 230 196 205 216 192 189
g2169 g2170 g2179 g2180 g2182 g2183 g2188 g2193 g2196 g2202 g2207 g2208
205 196 267 259 226 207 210 185 189 220 202 248
g2210 g2213 g2216 g2217 g2218 g2223 g2224 g2225 g2228 g2231 g2234 g2235
219 252 249 190 210 208 193 225 189 207 191 252
g2236 g2241 g2242 g2244 g2245 g2250 g2251 g2253 g2254 g2259 g2261 g2262
242 199 187 246 232 203 222 194 229 225 224 247
g2265 g2266 g2267 g2276 g2279 g2285 g2289 g2290 g2291 g2293 g2294 g2297
259 272 202 245 228 244 274 213 199 236 229 257
g2299 g2302 g2303 g2306 g2307 g2310 g2313 g2318 g2319 g2320 g2327 g2330
197 254 220 232 260 187 265 239 185 194 200 241
g2331 g2341 g2343 g2345 g2347 g2348 g2352 g2355 g2356 g2359 g2364 g2365
194 187 239 232 250 231 214 237 263 200 218 243
g2368 g2373 g2374 g2376 g2378 g2386 g2388 g2390 g2395 g2397 g2402 g2408
196 210 242 220 193 288 187 193 204 215 262 203
g2410 g2418 g2419 g2422 g2424 g2429 g2430 g2432 g2433 g2438 g2442 g2443
256 279 222 235 216 225 259 214 215 262 190 216
g2444 g2447 g2449 g2451 g2458 g2459 g2464 g2466 g2472 g2473 g2478 g2482
227 233 227 191 199 244 213 259 190 187 187 202
g2485 g2486 g2487 g2489 g2496 g2497 g2498 g2506 g2512 g2516 g2517 g2526
189 191 197 288 235 207 217 240 213 187 205 219
g2532 g2538 g2543 g2554 g2557 g2558 g2567 g2568 g2573 g2576 g2581 g2582
217 240 207 209 232 227 195 195 233 242 215 194
g2584 g2592 g2593 g2594 g2595 g2599 g2602 g2610 g2616 g2627 g2629 g2638
185 229 258 215 216 214 228 207 270 253 212 217
g2642 g2645 g2653 g2657 g2660 g2668 g2673 g2683 g2686 g2693 g2702 g2706
240 272 214 186 194 230 250 209 253 244 279 185
g2708 g2716 g2720 g2727 g2729 g2736 g2747 g2751 g2753 g2755 g2770 g2776
193 187 189 211 215 261 194 187 260 203 238 213
g2783 g2784 g2786 g2787 g2790 g2794 g2798 g2801 g2802 g2803 g2804 g2815
217 197 246 222 223 228 229 274 224 247 228 223
g2817 g2818 g2828 g2829 g2833 g2841 g2842 g2844 g2845 g2851 g2860 g2861
220 199 195 273 193 193 231 236 189 281 272 192
g2863 g2867 g2871 g2876 g2879 g2888 g2889 g2890 g2894 g2902 g2903 g2907
202 201 203 211 272 229 232 190 216 236 242 232
g2925 g2939 g2947 g2950 g2955 g2959 g2961 g2973 g2982 g2983 g2984 g2985
212 291 219 258 267 211 203 203 194 222 204 252
g2988 g2997 g3000 g3001 g3006 g3024 g3033 g3037 g3046
230 190 190 195 211 189 207 198 276
g9 g10 g17 g23 g25 g39 g47 g48 g51 g54 g55 g59 g63 g66
188 191 204 249 210 212 201 193 211 186 251 204 198 258

g74 g77 g78 g79 g81 g83 g84 g89 g94 g96 g98 g102 g104 g106
234 220 227 211 244 188 225 195 187 274 188 224 236 215
g109 g112 g117 g120 g121 g124 g126 g127 g128 g135 g138 g142 g148 g149
188 209 232 185 225 199 262 256 216 244 233 216 213 205
g151 g153 g154 g156 g158 g159 g163 g167 g171 g172 g174 g177 g178 g182
208 235 197 196 246 191 206 215 223 253 262 230 236 261
g184 g187 g188 g189 g192 g193 g194 g195 g201 g202 g204 g205 g206 g207
244 232 232 216 241 220 215 228 188 255 263 228 221 246
g210 g214 g218 g220 g225 g226 g227 g228 g230 g232 g233 g235 g237 g239
202 211 200 240 202 244 205 233 230 250 188 220 237 247
g240 g241 g242 g244 g246 g248 g250 g253 g254 g258 g259 g265 g270 g271
206 202 186 236 254 244 189 266 213 233 266 188 232 227
g277 g282 g283 g285 g286 g288 g289 g290 g291 g295 g296 g297 g298 g304
185 242 271 226 260 238 213 225 221 205 185 257 234 243
g307 g309 g312 g313 g314 g316 g318 g320 g323 g325 g327 g328 g329 g330
252 245 189 248 246 234 210 241 262 228 195 202 275 245
g331 g332 g333 g334 g335 g336 g337 g338 g342 g344 g345 g347 g357 g359
212 186 188 221 254 215 232 188 201 259 274 187 247 235
g360 g362 g363 g364 g365 g366 g368 g369 g370 g373 g376 g377 g385 g389
202 200 230 215 212 190 215 209 205 189 244 229 228 227
g391 g394 g395 g398 g399 g402 g407 g421 g422 g423 g426 g435 g441 g445
245 290 250 235 247 209 240 211 270 207 237 210 197 225
g446 g447 g449 g453 g454 g455 g460 g462 g464 g472 g478 g479 g486 g489
198 202 185 242 213 226 195 259 191 185 237 248 187 250
g490 g493 g494 g500 g504 g514 g515 g516 g520 g522 g523 g527 g529 g538
248 191 251 207 212 194 264 210 208 257 283 194 229 211
g542 g543 g544 g546 g551 g555 g557 g559 g560 g561 g563 g564 g567 g571
238 201 224 274 234 193 221 222 262 281 255 196 219 237
g573 g579 g590 g591 g603 g605 g610 g611 g615 g620 g621 g627 g634 g640
209 195 245 185 195 213 218 227 207 239 269 236 243 192
g645 g648 g651 g652 g655 g661 g663 g671 g675 g678 g679 g685 g686 g688
229 271 189 193 192 187 192 199 230 190 201 232 234 258
g689 g692 g693 g695 g696 g698 g701 g702 g703 g704 g710 g711 g713 g715
217 185 199 251 221 241 240 186 285 273 199 226 266 240
g717 g718 g720 g725 g726 g730 g733 g738 g742 g746 g753 g758 g763 g764
283 210 233 252 238 214 220 271 218 277 223 206 248 241
g770 g772 g776 g777 g785 g793 g797 g800 g801 g804 g807 g811 g828 g835
186 204 199 201 256 224 217 220 263 193 248 249 186 272
g837 g838 g839 g842 g846 g849 g852 g856 g858 g862 g864 g866 g867 g871
210 283 272 229 242 272 216 191 235 259 225 269 190 195
g877 g882 g885 g886 g889 g890 g892 g893 g906 g909 g912 g922 g924 g934
236 187 233 220 201 224 204 226 223 201 196 272 185 191
g936 g940 g949 g952 g958 g962 g963 g965 g966 g967 g970 g971 g974 g977
199 222 193 218 190 217 250 222 221 213 217 242 210 224
g978 g984 g985 g990 g997 g1003 g1006 g1007 g1011 g1014 g1016 g1018 g1019
g1020

190 283 213 209 232 201 275 202 243 259 237 197 254 195
g1021 g1025 g1027 g1030 g1037 g1040 g1042 g1045 g1047 g1057 g1058 g1060 g1061 g1070
214 244 220 244 281 206 284 270 191 228 217 264 224 260
g1072 g1075 g1078 g1081 g1083 g1086 g1094 g1101 g1105 g1108 g1109 g1110 g1117 g1126
191 223 222 249 185 278 244 258 215 191 239 242 214 216
g1133 g1140 g1145 g1146 g1161 g1162 g1163 g1167 g1173 g1174 g1175 g1176 g1177 g1179
211 200 262 193 245 261 207 199 195 240 239 216 185 199
g1181 g1182 g1186 g1192 g1193 g1197 g1198 g1199 g1202 g1203 g1206 g1221 g1222 g1223
227 197 200 196 234 189 226 203 256 210 237 225 193 205
g1225 g1228 g1231 g1238 g1245 g1252 g1253 g1254 g1255 g1258 g1262 g1270 g1271 g1273
254 208 221 223 245 202 255 185 220 196 206 186 268 209
g1278 g1282 g1285 g1288 g1289 g1290 g1292 g1293 g1295 g1298 g1310 g1316 g1317 g1318
218 203 208 235 237 232 194 255 218 229 221 255 229 207
g1324 g1327 g1329 g1330 g1334 g1337 g1340 g1341 g1347 g1348 g1351 g1352 g1353 g1354
191 267 238 208 266 244 202 228 223 241 207 222 192 212
g1355 g1359 g1360 g1364 g1368 g1369 g1372 g1377 g1381 g1382 g1388 g1390 g1392 g1399
205 244 233 220 279 209 198 209 248 200 220 234 195 195
g1410 g1421 g1425 g1426 g1428 g1435 g1437 g1440 g1441 g1445 g1453 g1455 g1456 g1459
197 207 224 215 224 208 187 203 224 263 254 266 261 250
g1466 g1468 g1469 g1470 g1474 g1478 g1480 g1481 g1484 g1489 g1491 g1501 g1502 g1505
198 245 202 243 250 202 242 216 237 242 242 233 189 202
g1508 g1509 g1513 g1515 g1519 g1524 g1531 g1538 g1542 g1543 g1547 g1552 g1554 g1557
215 215 241 214 218 282 226 185 266 210 243 189 193 209
g1559 g1560 g1561 g1562 g1564 g1565 g1566 g1568 g1571 g1579 g1585 g1586 g1594 g1598
260 188 195 202 246 208 210 191 226 245 267 190 238 275
g1604 g1605 g1606 g1608 g1610 g1613 g1616 g1619 g1620 g1621 g1629 g1631 g1633 g1636
218 207 187 190 249 220 258 205 202 236 245 222 219 218
g1638 g1640 g1641 g1642 g1644 g1648 g1649 g1653 g1658 g1660 g1661 g1663 g1666 g1668
262 265 224 265 197 234 246 259 214 192 196 206 191 222
g1671 g1679 g1681 g1684 g1688 g1690 g1691 g1695 g1696 g1698 g1699 g1701 g1702 g1707
254 215 192 223 192 200 266 189 250 215 231 235 215 209

g1709 g1712 g1714 g1718 g1719 g1723 g1724 g1725 g1726 g1731 g1732 g1733 g1739 g1741
198 217 189 229 217 256 196 190 207 187 264 198 188 196
g1742 g1744 g1745 g1752 g1756 g1765 g1766 g1770 g1773 g1775 g1777 g1779 g1786 g1787
227 190 218 225 191 194 252 221 247 199 230 189 209 245
g1791 g1793 g1795 g1804 g1805 g1806 g1807 g1809 g1811 g1812 g1813 g1817 g1818 g1821
200 216 227 187 249 189 241 231 284 193 214 272 205 222
g1827 g1828 g1832 g1834 g1845 g1851 g1855 g1856 g1857 g1859 g1860 g1866 g1869 g1873
217 193 199 290 205 231 229 263 211 201 210 209 272 234
g1876 g1880 g1881 g1882 g1883 g1887 g1890 g1900 g1903 g1908 g1909 g1916 g1917 g1919
216 208 220 260 281 227 213 196 253 216 275 273 201 209
g1920 g1921 g1926 g1929 g1931 g1932 g1934 g1938 g1939 g1943 g1944 g1945 g1947 g1948
274 190 248 191 207 229 230 231 268 239 249 223 254 252
g1949 g1951 g1955 g1956 g1959 g1962 g1963 g1969 g1971 g1973 g1978 g1979 g1980 g1981
200 190 240 191 272 227 250 227 188 221 271 225 229 238
g1985 g1986 g1991 g1992 g1993 g1995 g2000 g2002 g2006 g2012 g2018 g2019 g2020 g2027
232 210 221 203 246 287 195 283 191 206 199 222 265 190
g2032 g2039 g2040 g2045 g2050 g2052 g2061 g2063 g2065 g2067 g2069 g2072 g2073 g2076
239 214 201 200 224 240 241 210 237 189 217 219 201 186
g2077 g2079 g2080 g2083 g2087 g2088 g2092 g2096 g2097 g2105 g2107 g2108 g2110 g2118
215 244 231 196 247 185 187 236 186 252 185 196 249 228
g2122 g2132 g2143 g2146 g2149 g2150 g2155 g2156 g2162 g2164 g2169 g2170 g2179 g2180
270 238 217 218 230 196 205 216 192 189 205 196 267 259
g2182 g2183 g2188 g2193 g2196 g2202 g2207 g2208 g2210 g2213 g2216 g2217 g2218 g2223
226 207 210 185 189 220 202 248 219 252 249 190 210 208
g2224 g2225 g2228 g2231 g2234 g2235 g2236 g2241 g2242 g2244 g2245 g2250 g2251 g2253
193 225 189 207 191 252 242 199 187 246 232 203 222 194
g2254 g2259 g2261 g2262 g2265 g2266 g2267 g2276 g2279 g2285 g2289 g2290 g2291 g2293
229 225 224 247 259 272 202 245 228 244 274 213 199 236
g2294 g2297 g2299 g2302 g2303 g2306 g2307 g2310 g2313 g2318 g2319 g2320 g2327 g2330
229 257 197 254 220 232 260 187 265 239 185 194 200 241

```
g2331 g2341 g2343 g2345 g2347 g2348 g2352 g2355 g2356 g2359 g2364 g2365 g2368
g2373
 194   187   239   232   250   231   214   237   263   200   218   243   196   210
g2374 g2376 g2378 g2386 g2388 g2390 g2395 g2397 g2402 g2408 g2410 g2418 g2419
g2422
 242   220   193   288   187   193   204   215   262   203   256   279   222   235
g2424 g2429 g2430 g2432 g2433 g2438 g2442 g2443 g2444 g2447 g2449 g2451 g2458
g2459
 216   225   259   214   215   262   190   216   227   233   227   191   199   244
g2464 g2466 g2472 g2473 g2478 g2482 g2485 g2486 g2487 g2489 g2496 g2497 g2498
g2506
 213   259   190   187   187   202   189   191   197   288   235   207   217   240
g2512 g2516 g2517 g2526 g2532 g2538 g2543 g2554 g2557 g2558 g2567 g2568 g2573
g2576
 213   187   205   219   217   240   207   209   232   227   195   195   233   242
g2581 g2582 g2584 g2592 g2593 g2594 g2595 g2599 g2602 g2610 g2616 g2627 g2629
g2638
 215   194   185   229   258   215   216   214   228   207   270   253   212   217
g2642 g2645 g2653 g2657 g2660 g2668 g2673 g2683 g2686 g2693 g2702 g2706 g2708
g2716
 240   272   214   186   194   230   250   209   253   244   279   185   193   187
g2720 g2727 g2729 g2736 g2747 g2751 g2753 g2755 g2770 g2776 g2783 g2784 g2786
g2787
 189   211   215   261   194   187   260   203   238   213   217   197   246   222
g2790 g2794 g2798 g2801 g2802 g2803 g2804 g2815 g2817 g2818 g2828 g2829 g2833
g2841
 223   228   229   274   224   247   228   223   220   199   195   273   193   193
g2842 g2844 g2845 g2851 g2860 g2861 g2863 g2867 g2871 g2876 g2879 g2888 g2889
g2890
 231   236   189   281   272   192   202   201   203   211   272   229   232   190
g2894 g2902 g2903 g2907 g2925 g2939 g2947 g2950 g2955 g2959 g2961 g2973 g2982
g2983
 216   236   242   232   212   291   219   258   267   211   203   203   194   222
g2984 g2985 g2988 g2997 g3000 g3001 g3006 g3024 g3033 g3037 g3046
 204   252   230   190   190   195   211   189   207   198   276
```

7.) Now we want to compare these results to those using the empirical Bayes method in the limma package. Load this library and calculate p-values for the same dataset using the eBayes() function. (5 pts)

```
library(limma)
design <- cbind(Grp1=1, Grp2v1=c(rep(1,27),
            rep(0,11)))
fit <- lmFit(golub.df,design)
fit <- eBayes(fit)
```

```
#p-values
p.values <- fit$p.value[,2]
```

8.) Sort the empirical Bayes p-values and acquire the lowest *n* p-values, where *n* is defined as the number of significant test statistics that you found in problem 6. Intersect the gene names for your two methods and report how many are in common between the two differential expression methods, when choosing the top *n* genes from each set. (2.5 pts)

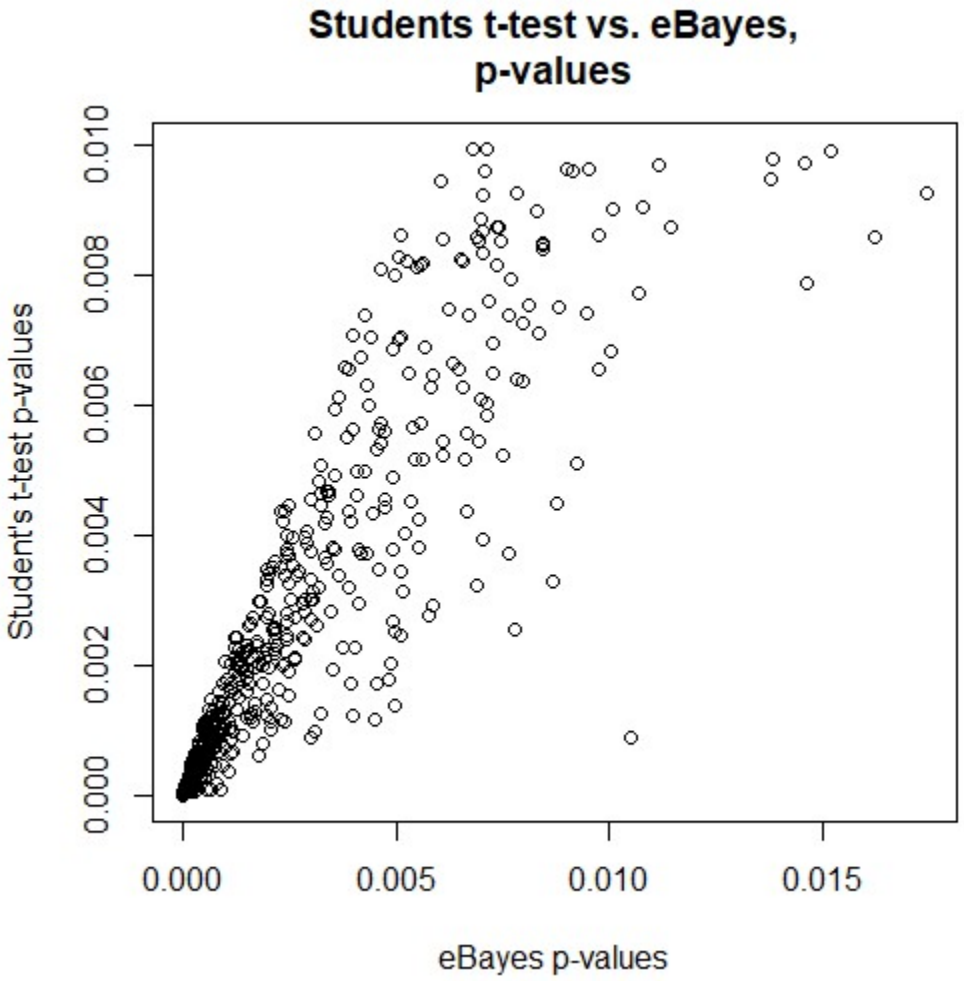
```
p.values.sorted <- sort(p.values)
n.lowest.p <- p.values.sorted[1:length(original.95)]
interesct.genes <- intersect(names(original.95),names(n.lowest.p))
> length(interesct.genes)
[1] 492
```

9.) Finally, compare the results from a Student's t-test with the empirical Bayes method. To do this, first calculate a two sample (two-tailed) Student's t-test on all genes. Make sure that you are running a Student's t-test and not a Welch's t-test. Then extract only those genes with a p-value less than 0.01 from this test. Plot the gene p-values<0.01 for the Student's t-test vs. the same genes in the empirical Bayes method. Make sure to label the axes and title appropriately. (7.5 pts)

```
t.test.all.genes <- function(x,s1,s2){
  x1 <- x[s1]
  x2 <- x[s2]
  x1 <- as.numeric(x1)
  x2 <- as.numeric(x2)
  t.out <- t.test(x1, x2, alternative="two.sided", var.equal=T )
  out <- as.numeric(t.out$p.value)
  return(out)
}

pv <- apply(golub.df, 1, t.test.all.genes, s1=c(1:27),s2=c(28:38))
pv.01 <- pv[pv < 0.01]

plot(p.values[names(pv.01)],pv.01, main="Students t-test vs. eBayes,\np-values",
    xlab="eBayes p-values", ylab="Student's t-test p-values")
```


Students t-test vs. eBayes,
p-values
Student's t-test p-values
0.000
0.002
0.004
0.006
0.008
0.010
0.000
0.005
0.010
0.015
eBayes p-values