Activity 9 - Aryan Khandelwal

…

**Question 1**: Load up SURVEY10 using data(SURVEY10). Making student profiles is valuable from a marketing perspective so that advertisments can be targeted to the right set of people. To make a profile, associations must be analyzed.

1. Run cor\_matrix on the first 7 columns of the data frame to obtain a correlation matrix (there are so many variables doing cor\_matrix on the entirety of the data frame results in too many values to look at).

data(SURVEY10)#load in data using data() command  
cor\_matrix(SURVEY10[,1:7])#cor\_matrix on only the first 7 columns of SURVEY10, e.g. of SURVEY10[,1:7]

## Height Weight DesiredWeight GPA TxtPerDay MinPerDayFaceBook  
## Height 1.000 0.601 0.774 -0.138 -0.085 -0.082  
## Weight 0.601 1.000 0.871 -0.215 -0.030 0.022  
## DesiredWeight 0.774 0.871 1.000 -0.212 -0.073 -0.064  
## GPA -0.138 -0.215 -0.212 1.000 0.072 -0.027  
## TxtPerDay -0.085 -0.030 -0.073 0.072 1.000 0.109  
## MinPerDayFaceBook -0.082 0.022 -0.064 -0.027 0.109 1.000

1. Now run all\_correlations on the entirety of the dataframe to get the pairwise correlations between every quantitative variable. Add the argument sorted="strength" to have this list sorted by strength (from most negative to most positive). What pair of quantities has the strongest positive, strongest negative, and strongest overall relationship?

all\_correlations(SURVEY10,sorted="strength")

## var1 var2 correlation pval  
## 1 DesiredWeight NumBodyPiercings -0.588011431 3.123079e-66  
## 2 Height NumBodyPiercings -0.581874637 1.415521e-64  
## 3 Weight NumBodyPiercings -0.440424465 1.571993e-34  
## 4 NumBodyPiercings PercMoreAthleticThan -0.264786905 1.115451e-12  
## 5 NumBodyPiercings WeeklyHrsVideoGame -0.257427414 4.824701e-12  
## 6 NumBodyPiercings PercFunnierThan -0.246268218 4.076588e-11  
## 7 NumBodyPiercings PercMoreAttractiveThan -0.223745463 2.218836e-09  
## 8 Weight GPA -0.214698664 9.857550e-09  
## 9 DesiredWeight GPA -0.212174164 1.477443e-08  
## 10 NumBodyPiercings PercMoreIntelligentThan -0.201370289 7.894653e-08  
## 11 GPA PercentDateable -0.140393901 1.963759e-04  
## 12 Height GPA -0.137646918 2.621714e-04  
## 13 TxtPerDay PercMoreIntelligentThan -0.136561419 2.934565e-04  
## 14 Height NumTattoos -0.110066194 3.572351e-03  
## 15 NumBodyPiercings PercentDateable -0.096963721 1.031627e-02  
## 16 TxtPerDay PercFunnierThan -0.089300927 1.820185e-02  
## 17 Height TxtPerDay -0.084561819 2.537136e-02  
## 18 Height MinPerDayFaceBook -0.082193276 2.979055e-02  
## 19 GPA OwnAttractiveness -0.080982150 3.229498e-02  
## 20 MinPerDayFaceBook PercMoreIntelligentThan -0.078829486 3.719155e-02  
## 21 NumBodyPiercings OwnAttractiveness -0.077846977 3.962803e-02  
## 22 DesiredWeight TxtPerDay -0.072822623 5.429963e-02  
## 23 GPA NumTattoos -0.070025086 6.426685e-02  
## 24 DesiredWeight MinPerDayFaceBook -0.064472414 8.851521e-02  
## 25 GPA WeeklyHrsVideoGame -0.064185847 8.994345e-02  
## 26 GPA PercFunnierThan -0.063677498 9.252254e-02  
## 27 MinPerDayFaceBook PercMoreAthleticThan -0.062591100 9.823294e-02  
## 28 DesiredWeight NumTattoos -0.059258761 1.175147e-01  
## 29 TxtPerDay WeeklyHrsVideoGame -0.056908874 1.328089e-01  
## 30 TxtPerDay PercMoreAttractiveThan -0.051914439 1.703707e-01  
## 31 GPA PercMoreAttractiveThan -0.042028059 2.671434e-01  
## 32 PercentDateable PercMoreIntelligentThan -0.038684528 3.071043e-01  
## 33 NumTattoos DistanceMovedToSchool -0.038465652 3.098513e-01  
## 34 NumTattoos PercMoreIntelligentThan -0.034824876 3.579113e-01  
## 35 WeeklyHrsVideoGame NumPhoneContacts -0.034002156 3.693880e-01  
## 36 PercentDateable OwnAttractiveness -0.031128683 4.112333e-01  
## 37 Weight TxtPerDay -0.030450056 4.215110e-01  
## 38 GPA MinPerDayFaceBook -0.026650097 4.817720e-01  
## 39 Weight NumPhoneContacts -0.024653539 5.152159e-01  
## 40 PercentDateable PercMoreAthleticThan -0.022101644 5.596495e-01  
## 41 Height NumPhoneContacts -0.019108335 6.140227e-01  
## 42 DistanceMovedToSchool NumPhoneContacts -0.017688344 6.406038e-01  
## 43 TxtPerDay PercMoreAthleticThan -0.017477194 6.445971e-01  
## 44 NumBodyPiercings DistanceMovedToSchool -0.013874380 7.142305e-01  
## 45 WeeklyHrsVideoGame DistanceMovedToSchool -0.012612616 7.392276e-01  
## 46 Weight NumTattoos -0.011058382 7.703983e-01  
## 47 GPA PercMoreAthleticThan -0.010727704 7.770802e-01  
## 48 TxtPerDay PercentDateable -0.008327162 8.260516e-01  
## 49 NumTattoos PercMoreAthleticThan -0.007794828 8.370107e-01  
## 50 DesiredWeight NumPhoneContacts -0.007253297 8.481915e-01  
## 51 NumTattoos PercentDateable -0.006852183 8.564931e-01  
## 52 TxtPerDay DistanceMovedToSchool -0.005687164 8.806914e-01  
## 53 NumTattoos PercFunnierThan -0.003177723 9.331642e-01  
## 54 NumTattoos NumPhoneContacts -0.002916540 9.386463e-01  
## 55 NumTattoos WeeklyHrsVideoGame -0.000176690 9.962794e-01  
## 56 DistanceMovedToSchool PercMoreAttractiveThan 0.003058173 9.356731e-01  
## 57 MinPerDayFaceBook PercMoreAttractiveThan 0.003262598 9.313834e-01  
## 58 NumPhoneContacts PercMoreIntelligentThan 0.006191043 8.702104e-01  
## 59 PercentDateable PercMoreAttractiveThan 0.006981413 8.538167e-01  
## 60 GPA DistanceMovedToSchool 0.007460941 8.439006e-01  
## 61 MinPerDayFaceBook OwnAttractiveness 0.014069801 7.103854e-01  
## 62 MinPerDayFaceBook DistanceMovedToSchool 0.014713172 6.977788e-01  
## 63 DistanceMovedToSchool PercFunnierThan 0.020250350 5.930042e-01  
## 64 Weight MinPerDayFaceBook 0.021725587 5.663509e-01  
## 65 MinPerDayFaceBook PercFunnierThan 0.022598470 5.508553e-01  
## 66 PercentDateable NumPhoneContacts 0.030852989 4.153906e-01  
## 67 DistanceMovedToSchool PercentDateable 0.031063142 4.122194e-01  
## 68 DistanceMovedToSchool OwnAttractiveness 0.031134968 4.111388e-01  
## 69 MinPerDayFaceBook WeeklyHrsVideoGame 0.032074445 3.971602e-01  
## 70 WeeklyHrsVideoGame PercentDateable 0.032822699 3.862342e-01  
## 71 NumBodyPiercings NumPhoneContacts 0.033019102 3.833969e-01  
## 72 NumTattoos PercMoreAttractiveThan 0.033340746 3.787781e-01  
## 73 Weight DistanceMovedToSchool 0.035085475 3.543233e-01  
## 74 NumTattoos OwnAttractiveness 0.038812001 3.055119e-01  
## 75 MinPerDayFaceBook NumTattoos 0.040288102 2.874706e-01  
## 76 DesiredWeight DistanceMovedToSchool 0.041760912 2.701986e-01  
## 77 DistanceMovedToSchool PercMoreIntelligentThan 0.044470064 2.403136e-01  
## 78 Weight OwnAttractiveness 0.054818937 1.476642e-01  
## 79 PercentDateable PercFunnierThan 0.058448340 1.226254e-01  
## 80 DistanceMovedToSchool PercMoreAthleticThan 0.060845084 1.079952e-01  
## 81 TxtPerDay NumTattoos 0.060968850 1.072789e-01  
## 82 Height DistanceMovedToSchool 0.061224490 1.058112e-01  
## 83 WeeklyHrsVideoGame OwnAttractiveness 0.063794716 9.192264e-02  
## 84 NumPhoneContacts OwnAttractiveness 0.063974684 9.100768e-02  
## 85 TxtPerDay OwnAttractiveness 0.065338104 8.431114e-02  
## 86 GPA NumPhoneContacts 0.066584731 7.854192e-02  
## 87 NumPhoneContacts PercMoreAttractiveThan 0.069048891 6.808146e-02  
## 88 GPA TxtPerDay 0.072355284 5.586902e-02  
## 89 NumPhoneContacts PercFunnierThan 0.074828642 4.797413e-02  
## 90 MinPerDayFaceBook NumPhoneContacts 0.075423816 4.622075e-02  
## 91 Weight PercMoreAttractiveThan 0.079316920 3.603079e-02  
## 92 WeeklyHrsVideoGame PercMoreAttractiveThan 0.085580126 2.365281e-02  
## 93 WeeklyHrsVideoGame PercMoreAthleticThan 0.089514669 1.792512e-02  
## 94 TxtPerDay MinPerDayFaceBook 0.109063200 3.890153e-03  
## 95 Height PercentDateable 0.110207434 3.529544e-03  
## 96 Weight PercMoreIntelligentThan 0.112148964 2.986468e-03  
## 97 MinPerDayFaceBook NumBodyPiercings 0.112694137 2.848294e-03  
## 98 WeeklyHrsVideoGame PercFunnierThan 0.113452641 2.665733e-03  
## 99 NumPhoneContacts PercMoreAthleticThan 0.118630498 1.678478e-03  
## 100 WeeklyHrsVideoGame PercMoreIntelligentThan 0.126148703 8.300076e-04  
## 101 MinPerDayFaceBook PercentDateable 0.131573324 4.874455e-04  
## 102 Height OwnAttractiveness 0.136518301 2.947685e-04  
## 103 GPA NumBodyPiercings 0.137175004 2.753682e-04  
## 104 Weight PercentDateable 0.142315987 1.599216e-04  
## 105 DesiredWeight PercentDateable 0.143563381 1.397736e-04  
## 106 DesiredWeight OwnAttractiveness 0.150842079 6.232041e-05  
## 107 GPA PercMoreIntelligentThan 0.153376239 4.662854e-05  
## 108 Weight PercMoreAthleticThan 0.158044122 2.700022e-05  
## 109 Height PercMoreIntelligentThan 0.175034280 3.234068e-06  
## 110 DesiredWeight PercMoreIntelligentThan 0.182631601 1.168992e-06  
## 111 TxtPerDay NumBodyPiercings 0.199076257 1.113902e-07  
## 112 DesiredWeight PercMoreAttractiveThan 0.215293769 8.954180e-09  
## 113 TxtPerDay NumPhoneContacts 0.218417070 5.382270e-09  
## 114 Weight PercFunnierThan 0.235014552 3.161374e-10  
## 115 NumTattoos NumBodyPiercings 0.241237667 1.031569e-10  
## 116 Height PercMoreAttractiveThan 0.251945326 1.394446e-11  
## 117 Weight WeeklyHrsVideoGame 0.254080813 9.250101e-12  
## 118 PercMoreIntelligentThan OwnAttractiveness 0.268063572 5.726106e-13  
## 119 Height PercFunnierThan 0.277096003 8.686943e-14  
## 120 DesiredWeight WeeklyHrsVideoGame 0.279535830 5.156825e-14  
## 121 Height WeeklyHrsVideoGame 0.290835040 4.304442e-15  
## 122 DesiredWeight PercFunnierThan 0.302398223 3.013519e-16  
## 123 Height PercMoreAthleticThan 0.306022553 1.277214e-16  
## 124 DesiredWeight PercMoreAthleticThan 0.317913852 7.016113e-18  
## 125 PercFunnierThan OwnAttractiveness 0.322333552 2.307142e-18  
## 126 PercMoreAthleticThan OwnAttractiveness 0.380665636 1.595159e-25  
## 127 PercMoreIntelligentThan PercMoreAthleticThan 0.387199357 2.018012e-26  
## 128 PercMoreAthleticThan PercFunnierThan 0.411011691 7.192889e-30  
## 129 PercMoreIntelligentThan PercFunnierThan 0.430716109 6.106518e-33  
## 130 PercMoreAttractiveThan PercFunnierThan 0.523020812 2.454992e-50  
## 131 PercMoreAttractiveThan PercMoreIntelligentThan 0.524007071 1.492597e-50  
## 132 PercMoreAttractiveThan PercMoreAthleticThan 0.536880100 1.936289e-53  
## 133 PercMoreAttractiveThan OwnAttractiveness 0.559191232 9.503662e-59  
## 134 Height Weight 0.601042971 7.200545e-70  
## 135 Height DesiredWeight 0.774149426 1.481493e-140  
## 136 Weight DesiredWeight 0.871371589 9.805054e-218

*Response*: The most negative is DesiredWeight and NumBodyPiercings, the most positive is between Weights and DesiredWeight. The strongest is Weights and DesiredWeight because they have the greatest absolute value.

1. Which quantity has the strongest correlation with GPA. Run all\_correlations adding the argument interest="GPA" to find out.

all\_correlations(SURVEY10,interest='GPA',sorted="magnitude")[1,]

## var1 var2 correlation pval  
## 1 Weight GPA -0.2146987 9.85755e-09

*Response*: Weight has the strongest correlation with GPA.

1. The analysis you have done so far has assumed that Pearson’s correlation coefficient provides a suitable measure of the strength of the relationship. This may not be the case.

* Rerun the all\_correlations command from part b, but add the argument type="spearman" so that Spearmans’ rank correlations are considered. You can ignore the There were 50 or more warnings... message.

#all\_correlations with argument type="spearman"  
all\_correlations(SURVEY10,type="spearman",sorted="strength")

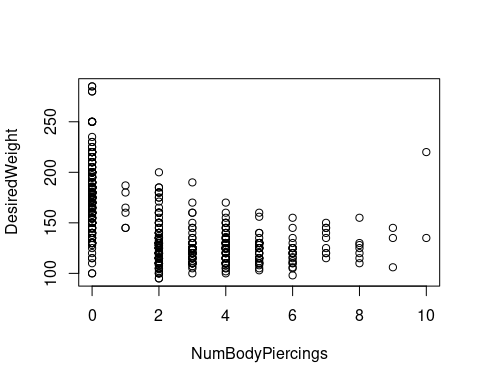
## var1 var2 correlation pval  
## 1 DesiredWeight NumBodyPiercings -0.702552581 4.516960e-105  
## 2 Height NumBodyPiercings -0.676238865 1.392226e-94  
## 3 Weight NumBodyPiercings -0.572657506 3.745198e-62  
## 4 NumBodyPiercings WeeklyHrsVideoGame -0.401618597 1.780525e-28  
## 5 NumBodyPiercings PercMoreAthleticThan -0.307267978 9.483049e-17  
## 6 DesiredWeight GPA -0.274943297 1.370355e-13  
## 7 NumBodyPiercings PercFunnierThan -0.273889134 1.710567e-13  
## 8 Weight GPA -0.270213510 3.678613e-13  
## 9 NumBodyPiercings PercMoreAttractiveThan -0.249602476 2.178074e-11  
## 10 NumBodyPiercings PercMoreIntelligentThan -0.225404079 1.676249e-09  
## 11 Height GPA -0.200948320 8.413300e-08  
## 12 GPA OwnAttractiveness -0.179237948 1.851492e-06  
## 13 GPA WeeklyHrsVideoGame -0.162213557 1.635508e-05  
## 14 GPA PercFunnierThan -0.132360345 4.504576e-04  
## 15 NumBodyPiercings OwnAttractiveness -0.127863973 7.029835e-04  
## 16 DesiredWeight TxtPerDay -0.118195232 1.746258e-03  
## 17 GPA PercentDateable -0.114511567 2.428689e-03  
## 18 Height NumTattoos -0.113263898 2.710135e-03  
## 19 GPA PercMoreAttractiveThan -0.111468398 3.167510e-03  
## 20 Weight TxtPerDay -0.107165894 4.562264e-03  
## 21 TxtPerDay PercMoreIntelligentThan -0.103461887 6.184355e-03  
## 22 Height TxtPerDay -0.100041741 8.123973e-03  
## 23 NumBodyPiercings PercentDateable -0.099743873 8.316250e-03  
## 24 DesiredWeight MinPerDayFaceBook -0.085230077 2.423185e-02  
## 25 GPA NumTattoos -0.084518691 2.544647e-02  
## 26 Height MinPerDayFaceBook -0.082406521 2.936730e-02  
## 27 MinPerDayFaceBook PercMoreAthleticThan -0.071220169 5.983840e-02  
## 28 TxtPerDay PercFunnierThan -0.064286474 8.943983e-02  
## 29 WeeklyHrsVideoGame NumPhoneContacts -0.063855883 9.161084e-02  
## 30 GPA PercMoreAthleticThan -0.063777296 9.201160e-02  
## 31 DesiredWeight NumTattoos -0.059984113 1.130834e-01  
## 32 PercentDateable PercMoreIntelligentThan -0.053887677 1.546783e-01  
## 33 MinPerDayFaceBook PercMoreIntelligentThan -0.038177531 3.134920e-01  
## 34 Weight MinPerDayFaceBook -0.037995806 3.158027e-01  
## 35 PercentDateable OwnAttractiveness -0.034693931 3.597229e-01  
## 36 NumTattoos PercMoreAthleticThan -0.034342931 3.646069e-01  
## 37 TxtPerDay WeeklyHrsVideoGame -0.032809359 3.864274e-01  
## 38 NumTattoos PercMoreIntelligentThan -0.030508842 4.206148e-01  
## 39 NumTattoos WeeklyHrsVideoGame -0.025619031 4.988945e-01  
## 40 Weight NumTattoos -0.025477858 5.012637e-01  
## 41 PercentDateable PercMoreAthleticThan -0.022939124 5.448649e-01  
## 42 NumTattoos NumPhoneContacts -0.021271621 5.744915e-01  
## 43 PercentDateable PercMoreAttractiveThan -0.019171776 6.128465e-01  
## 44 NumTattoos PercFunnierThan -0.018772595 6.202637e-01  
## 45 TxtPerDay PercMoreAttractiveThan -0.018206512 6.308483e-01  
## 46 NumTattoos PercentDateable -0.016418107 6.647792e-01  
## 47 DesiredWeight NumPhoneContacts -0.015787857 6.769067e-01  
## 48 WeeklyHrsVideoGame DistanceMovedToSchool -0.014906787 6.940010e-01  
## 49 Weight NumPhoneContacts -0.014358169 7.047249e-01  
## 50 TxtPerDay PercentDateable -0.013989705 7.119605e-01  
## 51 NumTattoos DistanceMovedToSchool -0.012068565 7.500931e-01  
## 52 GPA MinPerDayFaceBook -0.011965528 7.521566e-01  
## 53 MinPerDayFaceBook PercMoreAttractiveThan -0.006969246 8.540687e-01  
## 54 NumPhoneContacts PercMoreIntelligentThan -0.005826430 8.777924e-01  
## 55 Height NumPhoneContacts -0.004543441 9.045561e-01  
## 56 MinPerDayFaceBook OwnAttractiveness -0.003546151 9.254366e-01  
## 57 DistanceMovedToSchool OwnAttractiveness 0.002450190 9.484417e-01  
## 58 NumTattoos PercMoreAttractiveThan 0.003612665 9.240422e-01  
## 59 TxtPerDay PercMoreAthleticThan 0.005250869 8.897835e-01  
## 60 GPA DistanceMovedToSchool 0.013927220 7.131901e-01  
## 61 MinPerDayFaceBook NumTattoos 0.015035646 6.914910e-01  
## 62 MinPerDayFaceBook PercFunnierThan 0.018311593 6.288777e-01  
## 63 DesiredWeight DistanceMovedToSchool 0.021808843 5.648640e-01  
## 64 TxtPerDay OwnAttractiveness 0.022540141 5.518843e-01  
## 65 PercentDateable NumPhoneContacts 0.025417948 5.022710e-01  
## 66 NumBodyPiercings DistanceMovedToSchool 0.025648867 4.983945e-01  
## 67 Height DistanceMovedToSchool 0.028126269 4.578236e-01  
## 68 Weight DistanceMovedToSchool 0.029578160 4.349339e-01  
## 69 DistanceMovedToSchool PercMoreAttractiveThan 0.031873445 4.001266e-01  
## 70 GPA NumPhoneContacts 0.035267582 3.518294e-01  
## 71 WeeklyHrsVideoGame PercentDateable 0.039190958 3.008102e-01  
## 72 NumTattoos OwnAttractiveness 0.039692112 2.946667e-01  
## 73 NumPhoneContacts PercMoreAttractiveThan 0.041099475 2.778656e-01  
## 74 DistanceMovedToSchool PercMoreIntelligentThan 0.041862355 2.690356e-01  
## 75 NumPhoneContacts PercFunnierThan 0.041990935 2.675665e-01  
## 76 PercentDateable PercFunnierThan 0.045138723 2.333096e-01  
## 77 MinPerDayFaceBook WeeklyHrsVideoGame 0.046138309 2.231106e-01  
## 78 DistanceMovedToSchool PercFunnierThan 0.046268938 2.218017e-01  
## 79 NumBodyPiercings NumPhoneContacts 0.057068443 1.317239e-01  
## 80 NumPhoneContacts OwnAttractiveness 0.063420757 9.384744e-02  
## 81 DistanceMovedToSchool PercentDateable 0.065794736 8.215941e-02  
## 82 TxtPerDay NumTattoos 0.066665014 7.818165e-02  
## 83 DistanceMovedToSchool PercMoreAthleticThan 0.080356492 3.365782e-02  
## 84 WeeklyHrsVideoGame OwnAttractiveness 0.097170770 1.015379e-02  
## 85 Weight OwnAttractiveness 0.098266167 9.331476e-03  
## 86 NumPhoneContacts PercMoreAthleticThan 0.101030088 7.514111e-03  
## 87 Height PercentDateable 0.101433389 7.277255e-03  
## 88 WeeklyHrsVideoGame PercMoreAttractiveThan 0.104924444 5.490379e-03  
## 89 DistanceMovedToSchool NumPhoneContacts 0.108567687 4.056430e-03  
## 90 WeeklyHrsVideoGame PercMoreIntelligentThan 0.109670956 3.694658e-03  
## 91 Weight PercentDateable 0.111446901 3.173388e-03  
## 92 DesiredWeight PercentDateable 0.113770758 2.592394e-03  
## 93 Weight PercMoreAttractiveThan 0.115225029 2.280008e-03  
## 94 Weight PercMoreIntelligentThan 0.120283891 1.442439e-03  
## 95 MinPerDayFaceBook DistanceMovedToSchool 0.122183781 1.209110e-03  
## 96 TxtPerDay DistanceMovedToSchool 0.123844656 1.034180e-03  
## 97 WeeklyHrsVideoGame PercMoreAthleticThan 0.126444899 8.066558e-04  
## 98 MinPerDayFaceBook NumBodyPiercings 0.127702795 7.141034e-04  
## 99 GPA TxtPerDay 0.136739549 2.880940e-04  
## 100 MinPerDayFaceBook NumPhoneContacts 0.141983019 1.657437e-04  
## 101 MinPerDayFaceBook PercentDateable 0.155612288 3.596121e-05  
## 102 Height PercMoreIntelligentThan 0.169509979 6.598752e-06  
## 103 Height OwnAttractiveness 0.171025710 5.438191e-06  
## 104 DesiredWeight OwnAttractiveness 0.177928300 2.205959e-06  
## 105 GPA NumBodyPiercings 0.179249709 1.848571e-06  
## 106 DesiredWeight PercMoreIntelligentThan 0.179380779 1.816316e-06  
## 107 WeeklyHrsVideoGame PercFunnierThan 0.181606503 1.344403e-06  
## 108 NumTattoos NumBodyPiercings 0.192009743 3.136532e-07  
## 109 Weight PercMoreAthleticThan 0.204492179 4.909588e-08  
## 110 GPA PercMoreIntelligentThan 0.216320600 7.580818e-09  
## 111 DesiredWeight PercMoreAttractiveThan 0.220485316 3.825999e-09  
## 112 TxtPerDay MinPerDayFaceBook 0.227281065 1.217224e-09  
## 113 PercMoreIntelligentThan OwnAttractiveness 0.229317612 8.574471e-10  
## 114 TxtPerDay NumBodyPiercings 0.230885583 6.532384e-10  
## 115 Weight PercFunnierThan 0.241177031 1.043046e-10  
## 116 Height PercMoreAttractiveThan 0.263636674 1.406588e-12  
## 117 Height PercFunnierThan 0.280694945 4.017852e-14  
## 118 DesiredWeight PercFunnierThan 0.308294815 7.410757e-17  
## 119 TxtPerDay NumPhoneContacts 0.311909108 3.087270e-17  
## 120 Height PercMoreAthleticThan 0.320457243 3.707769e-18  
## 121 DesiredWeight PercMoreAthleticThan 0.325059002 1.151342e-18  
## 122 PercFunnierThan OwnAttractiveness 0.326705553 7.539489e-19  
## 123 Weight WeeklyHrsVideoGame 0.356942453 1.977612e-22  
## 124 Height WeeklyHrsVideoGame 0.365287694 1.726906e-23  
## 125 PercMoreIntelligentThan PercMoreAthleticThan 0.365328254 1.706262e-23  
## 126 PercMoreAthleticThan OwnAttractiveness 0.377081823 4.861135e-25  
## 127 PercMoreAthleticThan PercFunnierThan 0.388648571 1.267682e-26  
## 128 PercMoreIntelligentThan PercFunnierThan 0.389260395 1.041033e-26  
## 129 DesiredWeight WeeklyHrsVideoGame 0.409808354 1.091264e-29  
## 130 PercMoreAttractiveThan PercMoreIntelligentThan 0.481216979 8.490530e-42  
## 131 PercMoreAttractiveThan PercFunnierThan 0.499189864 2.524450e-45  
## 132 PercMoreAttractiveThan PercMoreAthleticThan 0.530512805 5.376017e-52  
## 133 PercMoreAttractiveThan OwnAttractiveness 0.593722449 8.337345e-68  
## 134 Height Weight 0.674692072 5.325269e-94  
## 135 Height DesiredWeight 0.814235864 9.059118e-167  
## 136 Weight DesiredWeight 0.908992504 4.170041e-267

* What pair of variables have the strong negative association?

*Response*: DesiredWeight and NumBodyPiercings have the strongest negative associations.

* Make a scatterplot of their relationship (same sort of syntax as what you use for associate: plot(y~x,data=DATA) and determine if you should be using Pearson or Spearman to gauge the strength.

plot(DesiredWeight~NumBodyPiercings, data = SURVEY10)



*Response*: Pearson should be used to gauge the strenght.

* The association is highly statistically significant, but there is obviously not a cause-effect physical law underlying the relationship. Why do you think we see such a strong relationship (i.e., what might be a lurking variable)?

*Response*: Gender is the lurking variable.

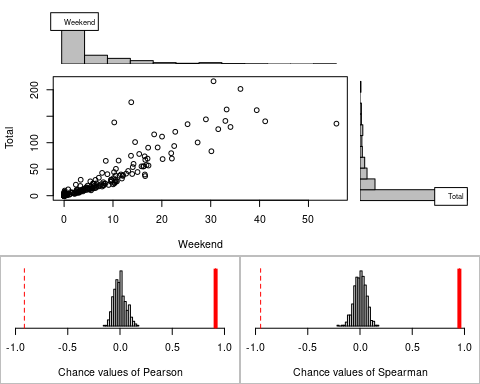
**Question 2**: Load up MOVIE using data(MOVIE). This shows the opening weekend and total domestic gross of movies from about a decade ago.

data(MOVIE)

1. Run associate() as if you were predicting Total from Weekend. Which measure (Pearson vs Spearman) is appropriate for testing the significance of the relationship and why?

associate(Total~Weekend, data = MOVIE)

## Association between Weekend (numerical) and Total (numerical)  
## using 309 complete cases



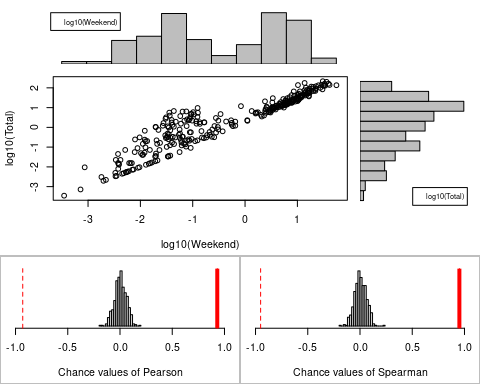
## Permutation procedure:  
## Value Estimated p-value  
## Pearson's r 0.9153056 0  
## Spearman's rank correlation 0.9511769 0  
## With 500 permutations, we are 95% confident that:  
## the p-value of Pearson's correlation (r) is between 0 and 0.007   
## the p-value of Spearman's rank correlation is between 0 and 0.007   
## Note: If 0.05 is in this range, increase the permutations= argument.  
##   
##   
##   
## Advice: If stream of points is well described by an ellipse, use Pearson's r.  
## Otherwise, as long as stream is monotonic, use Spearman's rank correlation  
## or try logs, e.g. associate( log10(y)~log10(x) )

*Response:* Pearson correlation should be used because the data points are not heteroscedastic and don’t has extreme outliers.

1. Run associate() again but this time use log10(Total) as your y variable and log10(Weekend) as your x variable. The scatterplot should look somewhat “nicer”. Notice how Spearmans’ rank correlation is the same value as before (this is one advantage of Spearman’s rank correlation, it is the same value if you look at monotonic transformations of the original values). Has Pearson’s correlation increased or decreased?

associate(log10(Total)~log10(Weekend), data = MOVIE)

## Association between log10(Weekend) (numerical) and log10(Total) (numerical)  
## using 309 complete cases



## Permutation procedure:  
## Value Estimated p-value  
## Pearson's r 0.9309400 0  
## Spearman's rank correlation 0.9511769 0  
## With 500 permutations, we are 95% confident that:  
## the p-value of Pearson's correlation (r) is between 0 and 0.007   
## the p-value of Spearman's rank correlation is between 0 and 0.007   
## Note: If 0.05 is in this range, increase the permutations= argument.  
##   
##   
##   
## Advice: If stream of points is well described by an ellipse, use Pearson's r.  
## Otherwise, as long as stream is monotonic, use Spearman's rank correlation  
## or try logs, e.g. associate( log10(y)~log10(x) )

*Response:* Pearson’s correlation has increased from 0.91 to 0.93

**Question 3**: Load EX2.CENSUS using data(EX2.CENSUS).

data(EX2.CENSUS)

1. Report the five variables most strongly correlated with ResponseRate.

all\_correlations(EX2.CENSUS,interest='ResponseRate',sorted="magnitude")[1:5,]

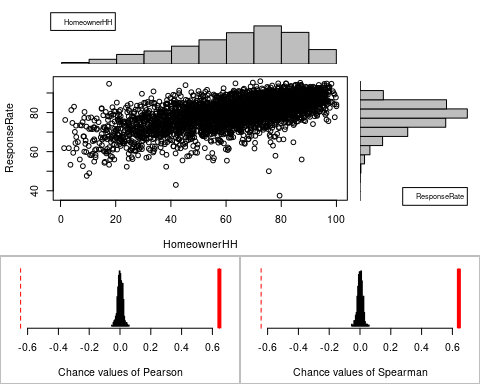
## var1 var2 correlation pval  
## 1 ResponseRate HomeownerHH 0.6454007 0.000000e+00  
## 2 ResponseRate RentingHH -0.6066960 0.000000e+00  
## 3 ResponseRate Whites 0.5608395 5.274055e-292  
## 4 ResponseRate FemaleHH -0.5560803 4.473113e-286  
## 5 ResponseRate MarriedHH 0.5535108 6.480532e-283

*Response:* HomeownerHH, RentingHH, Whites, FemaleHH, MarriedHH are the most strongly correlated with ResponseRate

1. Is the association between ResponseRate and HomeownerHH statistically significant?

associate(ResponseRate~HomeownerHH, data = EX2.CENSUS)

## Association between HomeownerHH (numerical) and ResponseRate (numerical)  
## using 3534 complete cases



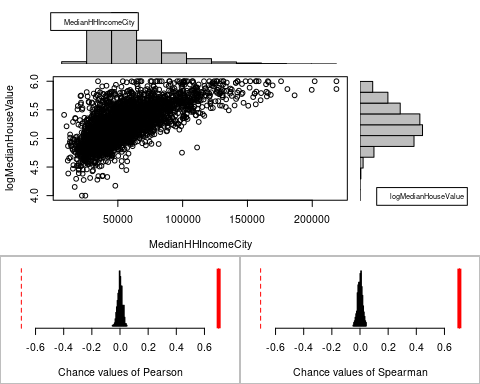
## Permutation procedure:  
## Value Estimated p-value  
## Pearson's r 0.6454007 0  
## Spearman's rank correlation 0.6421986 0  
## With 500 permutations, we are 95% confident that:  
## the p-value of Pearson's correlation (r) is between 0 and 0.007   
## the p-value of Spearman's rank correlation is between 0 and 0.007   
## Note: If 0.05 is in this range, increase the permutations= argument.  
##   
##   
##   
## Advice: If stream of points is well described by an ellipse, use Pearson's r.  
## Otherwise, as long as stream is monotonic, use Spearman's rank correlation  
## or try logs, e.g. associate( log10(y)~log10(x) )

*Response:* The association is statistically significant.

1. Describe the direction and form of the relationship between logMedianHouseValue and MedianHHIncomeCity (the x-variable). Does it have unusual features? Which measure should be used to gauge the strength?

associate(logMedianHouseValue~MedianHHIncomeCity, data = EX2.CENSUS)

## Association between MedianHHIncomeCity (numerical) and logMedianHouseValue (numerical)  
## using 3534 complete cases



## Permutation procedure:  
## Value Estimated p-value  
## Pearson's r 0.7009088 0  
## Spearman's rank correlation 0.7060014 0  
## With 500 permutations, we are 95% confident that:  
## the p-value of Pearson's correlation (r) is between 0 and 0.007   
## the p-value of Spearman's rank correlation is between 0 and 0.007   
## Note: If 0.05 is in this range, increase the permutations= argument.  
##   
##   
##   
## Advice: If stream of points is well described by an ellipse, use Pearson's r.  
## Otherwise, as long as stream is monotonic, use Spearman's rank correlation  
## or try logs, e.g. associate( log10(y)~log10(x) )

*Response:* Spearman association should be used, because it has heteroscedastic, and has outliers.

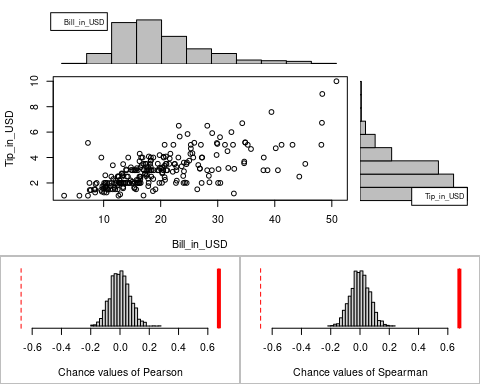
**Question 4**: Examine the association between bill and tip amounts in the EX2.TIPS data frame. Which measure is appropriate for measuring the strength of the relationship (Pearson vs. Spearman). Is the relationship statistically significant. Does the choice of Pearson vs. Spearman affect this conclusion?

data("EX2.TIPS")  
names(EX2.TIPS)

## [1] "Tip.Percentage" "Bill\_in\_USD" "Tip\_in\_USD" "Gender" "Smoker"   
## [6] "Weekday" "Day\_Night" "Size\_of\_Party"

associate(Tip\_in\_USD~Bill\_in\_USD, data = EX2.TIPS, permutations = 2000)

## Association between Bill\_in\_USD (numerical) and Tip\_in\_USD (numerical)  
## using 244 complete cases



## Permutation procedure:  
## Value Estimated p-value  
## Pearson's r 0.6757341 0  
## Spearman's rank correlation 0.6789681 0  
## With 2000 permutations, we are 95% confident that:  
## the p-value of Pearson's correlation (r) is between 0 and 0.002   
## the p-value of Spearman's rank correlation is between 0 and 0.002   
## Note: If 0.05 is in this range, increase the permutations= argument.  
##   
##   
##   
## Advice: If stream of points is well described by an ellipse, use Pearson's r.  
## Otherwise, as long as stream is monotonic, use Spearman's rank correlation  
## or try logs, e.g. associate( log10(y)~log10(x) )

*Response:*Spearman should be used because there is some heteroscedasticity. The test is statisically significant because the p value is below 0.05. The choice whether Spearman or Pearson did not effect the significany of the test because they were both lower than 0.05.