Factor Analysis

Jon\_Minton

Monday, March 16, 2015

# Dimensions of Diversity

This section will look at the relationship between various forms of diversity within Greater Glasgow.

## Warning: package 'rgl' was built under R version 3.1.3

Now to explore some of the differences between types of diversity. Using a pairs plot

# Inner join (will automatically cut to Greater Glasgow) while renaming vars  
  
div\_all <- demo\_div %>%  
 inner\_join(dtype\_div) %>%  
 inner\_join(eth\_div) %>%  
 inner\_join(sec\_div) %>%  
 inner\_join(ten\_div) %>%  
 inner\_join(spac\_div) %>%  
 inner\_join(band\_div) %>%  
 select(-X, -social, -rented, -owned)

## Joining by: "datazone"  
## Joining by: "datazone"

## Warning: joining factors with different levels, coercing to character  
## vector

## Joining by: "datazone"

## Warning: joining character vector and factor, coercing into character  
## vector

## Joining by: "datazone"

## Warning: joining character vector and factor, coercing into character  
## vector

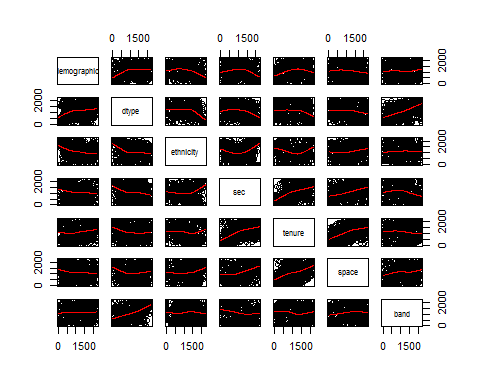
## Joining by: "datazone"

## Warning: joining character vector and factor, coercing into character  
## vector

## Joining by: "datazone"

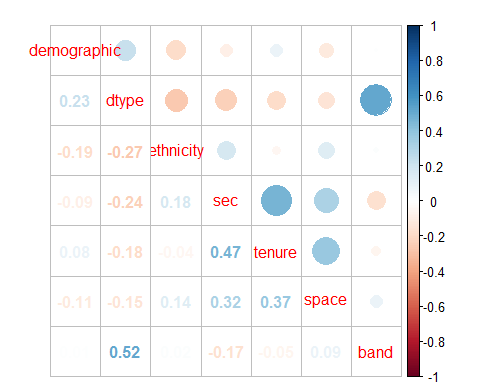
## Warning: joining character vector and factor, coercing into character  
## vector

d\_rankdiv\_all <- div\_all %>%  
 mutate\_each(funs(min\_rank), -datazone)  
  
pairs(sample\_n(d\_rankdiv\_all[,-1], 500), panel=panel.smooth)

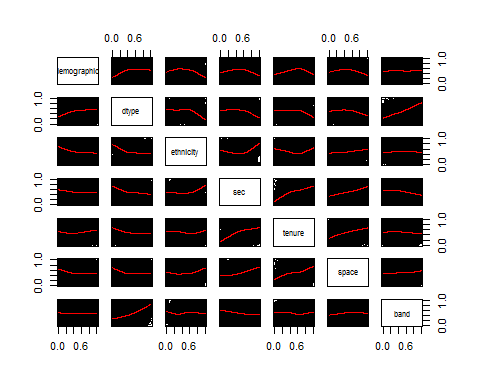


Now to look at correlations.

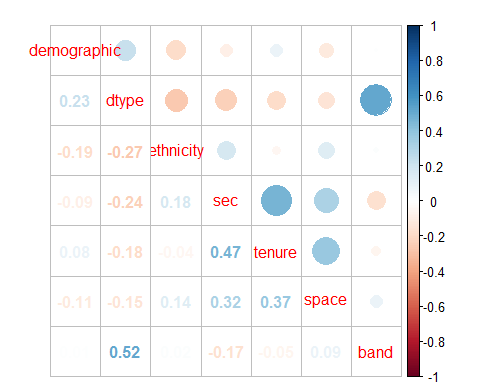
cor(d\_rankdiv\_all[,-1], method="spearman") %>%   
 corrplot.mixed()



d\_propdiv\_all <- div\_all %>%  
 mutate\_each(funs(percent\_rank), -datazone)  
  
pairs(d\_propdiv\_all[,-1], panel=panel.smooth)



cor(d\_propdiv\_all[,-1], method="spearman") %>%   
 corrplot.mixed()



# No difference

# factor analysis

## Basic example using factanal function

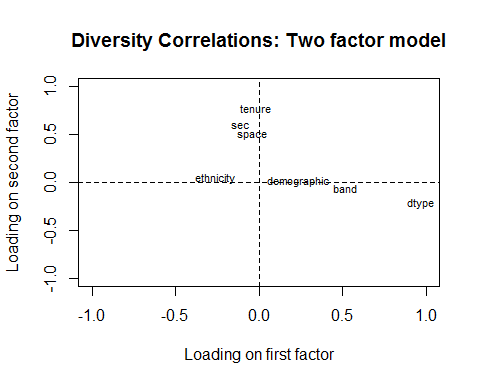
# One factor model  
m1 <- factanal(  
 ~ demographic + dtype + ethnicity + sec + tenure + space + band,   
 data=d\_propdiv\_all,   
 factors=1  
 )  
  
m1$loadings

##   
## Loadings:  
## Factor1  
## demographic -0.120   
## dtype -0.406   
## ethnicity 0.201   
## sec 0.720   
## tenure 0.602   
## space 0.468   
## band -0.230   
##   
## Factor1  
## SS loadings 1.373  
## Proportion Var 0.196

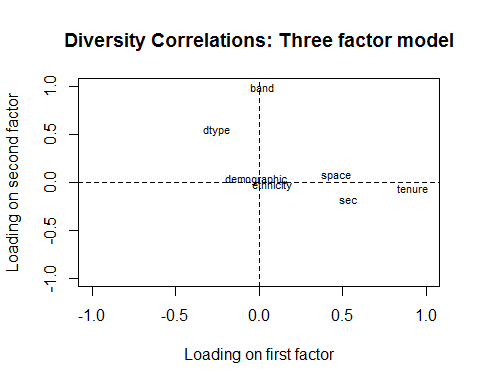
# Two factor model  
  
m2 <- factanal(  
 ~ demographic + dtype + ethnicity + sec + tenure + space + band,   
 data=d\_propdiv\_all,   
 factors=2  
 )  
  
m2$loadings

##   
## Loadings:  
## Factor1 Factor2  
## demographic 0.237   
## dtype 0.974 -0.217   
## ethnicity -0.263   
## sec -0.115 0.597   
## tenure 0.778   
## space 0.491   
## band 0.517   
##   
## Factor1 Factor2  
## SS loadings 1.356 1.256  
## Proportion Var 0.194 0.179  
## Cumulative Var 0.194 0.373

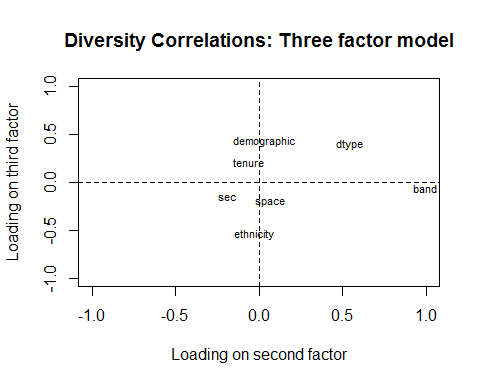
# Three factor model  
  
m3 <- factanal(  
 ~ demographic + dtype + ethnicity + sec + tenure + space + band,   
 data=d\_propdiv\_all,   
 factors=3  
 )  
  
  
# Plots  
load2 <- m2$loadings  
  
plot(load2, type="n", xlim=c(-1, 1), ylim=c(-1,1),   
 xlab="Loading on first factor", ylab="Loading on second factor",   
 main="Diversity Correlations: Two factor model")   
text(load2, labels=names(d\_propdiv\_all)[-1], cex=0.7)   
abline(v=0, lty="dashed"); abline(h=0, lty="dashed")



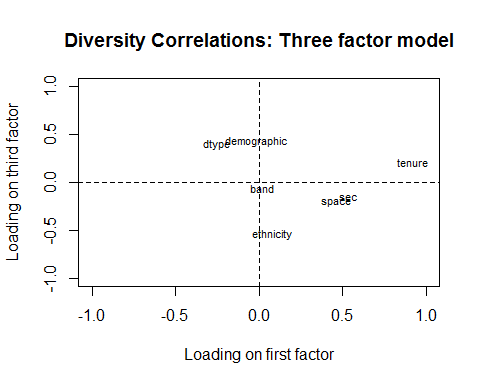
load3\_12 <- m3$loadings[,1:2]  
plot(load3\_12, type="n", xlim=c(-1, 1), ylim=c(-1,1),   
 xlab="Loading on first factor", ylab="Loading on second factor",   
 main="Diversity Correlations: Three factor model")   
text(load3\_12, labels=names(d\_propdiv\_all)[-1], cex=0.7)   
abline(v=0, lty="dashed"); abline(h=0, lty="dashed")



load3\_23 <- m3$loadings[,2:3]  
plot(load3\_23, type="n", xlim=c(-1, 1), ylim=c(-1,1),   
 xlab="Loading on second factor", ylab="Loading on third factor",   
 main="Diversity Correlations: Three factor model")   
text(load3\_23, labels=names(d\_propdiv\_all)[-1], cex=0.7)   
abline(v=0, lty="dashed"); abline(h=0, lty="dashed")

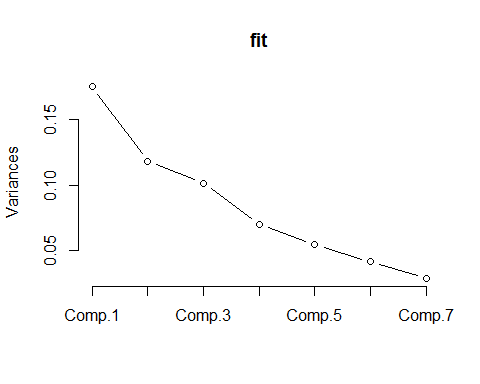


load3\_13 <- m3$loadings[,c(1,3)]  
plot(load3\_13, type="n", xlim=c(-1, 1), ylim=c(-1,1),   
 xlab="Loading on first factor", ylab="Loading on third factor",   
 main="Diversity Correlations: Three factor model")   
text(load3\_13, labels=names(d\_propdiv\_all)[-1], cex=0.7)   
abline(v=0, lty="dashed"); abline(h=0, lty="dashed")

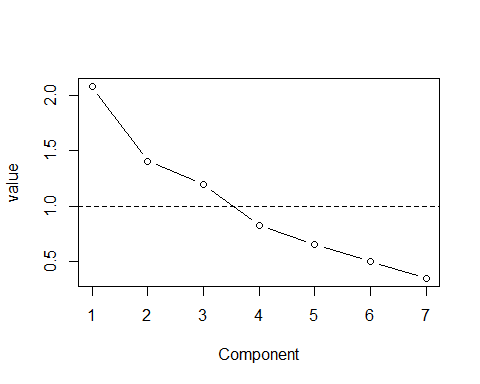


## Examples using princomp function

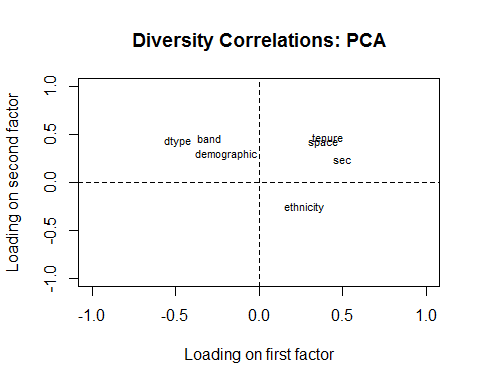
fit <- princomp(  
 ~ demographic + dtype + ethnicity + sec + tenure + space + band,   
 data=d\_propdiv\_all  
 )  
  
plot(fit, type="lines")



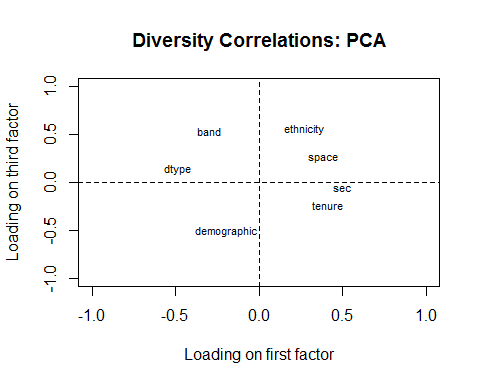
ev <- eigen(cor(d\_propdiv\_all[,c("demographic", "dtype", "ethnicity", "sec", "tenure", "space", "band")]))$values  
  
plot(ev, type="b", xlab="Component", ylab="value"); abline(h=1, lty="dashed")



plot(fit$loadings[,c(1,2)], type="n", xlim=c(-1, 1), ylim=c(-1,1),   
 xlab="Loading on first factor", ylab="Loading on second factor",   
 main="Diversity Correlations: PCA")   
text(fit$loadings[,c(1,2)], labels=names(d\_propdiv\_all)[-1], cex=0.7)   
abline(v=0, lty="dashed"); abline(h=0, lty="dashed")



plot(fit$loadings[,c(1,3)], type="n", xlim=c(-1, 1), ylim=c(-1,1),   
 xlab="Loading on first factor", ylab="Loading on third factor",   
 main="Diversity Correlations: PCA")   
text(fit$loadings[,c(1,3)], labels=names(d\_propdiv\_all)[-1], cex=0.7)   
abline(v=0, lty="dashed"); abline(h=0, lty="dashed")



plot(fit$loadings[,c(2,3)], type="n", xlim=c(-1, 1), ylim=c(-1,1),   
 xlab="Loading on second factor", ylab="Loading on third factor",   
 main="Diversity Correlations: PCA")   
text(fit$loadings[,c(2,3)], labels=names(d\_propdiv\_all)[-1], cex=0.7)   
abline(v=0, lty="dashed"); abline(h=0, lty="dashed")

