Homework 1

library(tidyverse)

## Loading tidyverse: ggplot2  
## Loading tidyverse: tibble  
## Loading tidyverse: tidyr  
## Loading tidyverse: readr  
## Loading tidyverse: purrr  
## Loading tidyverse: dplyr

## Conflicts with tidy packages ----------------------------------------------

## filter(): dplyr, stats  
## lag(): dplyr, stats

df <- as\_tibble(read.csv("~/Development-Economics/data/index\_test\_data.csv"))  
  
  
# rownames(df) <- df$X  
  
df

## # A tibble: 6 x 5  
## City Pollution.index Cost.of.living Language Misc.Amenities  
## <fctr> <int> <int> <int> <int>  
## 1 Shanghai 80 30 9 44  
## 2 Mexico DF 73 17 22 29  
## 3 Panama City 16 20 22 32  
## 4 Hanoi 28 14 18 25  
## 5 Singapore 44 70 70 71  
## 6 Chittagong 46 11 27 14

normalize<-function(x){  
 (x - min(x)) / (max(x) - min(x))  
 }  
  
norm.df <- df %>%  
 mutate\_if(is.numeric, funs(normalize))  
  
norm.df

## # A tibble: 6 x 5  
## City Pollution.index Cost.of.living Language Misc.Amenities  
## <fctr> <dbl> <dbl> <dbl> <dbl>  
## 1 Shanghai 1.000000 0.32203390 0.0000000 0.5263158  
## 2 Mexico DF 0.890625 0.10169492 0.2131148 0.2631579  
## 3 Panama City 0.000000 0.15254237 0.2131148 0.3157895  
## 4 Hanoi 0.187500 0.05084746 0.1475410 0.1929825  
## 5 Singapore 0.437500 1.00000000 1.0000000 1.0000000  
## 6 Chittagong 0.468750 0.00000000 0.2950820 0.0000000

pi.weight <- 0.25  
col.weight <- 0.2  
lang.weight <- 0.15  
amen.weight <- 0.4  
  
sum(pi.weight, col.weight, lang.weight, amen.weight)

## [1] 1

weighted.norm.df <- norm.df %>%  
 mutate(Pollution.index = Pollution.index \* pi.weight) %>%  
 mutate(Cost.of.living = Cost.of.living \* col.weight) %>%  
 mutate(Language = Language \* lang.weight) %>%  
 mutate(Misc.Amenities = Misc.Amenities \* amen.weight)  
  
weighted.norm.df

## # A tibble: 6 x 5  
## City Pollution.index Cost.of.living Language Misc.Amenities  
## <fctr> <dbl> <dbl> <dbl> <dbl>  
## 1 Shanghai 0.2500000 0.06440678 0.00000000 0.21052632  
## 2 Mexico DF 0.2226562 0.02033898 0.03196721 0.10526316  
## 3 Panama City 0.0000000 0.03050847 0.03196721 0.12631579  
## 4 Hanoi 0.0468750 0.01016949 0.02213115 0.07719298  
## 5 Singapore 0.1093750 0.20000000 0.15000000 0.40000000  
## 6 Chittagong 0.1171875 0.00000000 0.04426230 0.00000000

df

## # A tibble: 6 x 5  
## City Pollution.index Cost.of.living Language Misc.Amenities  
## <fctr> <int> <int> <int> <int>  
## 1 Shanghai 80 30 9 44  
## 2 Mexico DF 73 17 22 29  
## 3 Panama City 16 20 22 32  
## 4 Hanoi 28 14 18 25  
## 5 Singapore 44 70 70 71  
## 6 Chittagong 46 11 27 14