# appdensitycalculation

Data Science Toolbox Discussion Exercise

## Barangay Density Data

### Data Manipulation

#Import CSV Files  
population = read.csv("population.csv")  
regionarea = read.csv("regionarea.csv")  
  
#Load dplyr library  
library("dplyr")

##   
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':  
##   
## filter, lag

## The following objects are masked from 'package:base':  
##   
## intersect, setdiff, setequal, union

#Count the Number of Barangay per Region  
NumofBrgy= count(population,Region)  
BarData = left\_join(population, regionarea, by=c("Region"))  
BarData = left\_join(BarData, NumofBrgy, by=c("Region"))

Explanation: We first got the number of barangays per region using *count* function. We used *leftjoin* to add the columns Area (from regionarea.csv) and Number of Barangay per Region (from count function).

### Barangay Area Calculation

Formula:

Code Snippet:

#Divide the area by the number of barangays in the region  
BarData$BrgyArea = BarData$Area / BarData$n

Explanation: In order to get the area per barangay, the total area of the region is divided by the total number of barangays in the region.

### Barangay Population Density Calculation

Formula:

Code Snippet:

#get the population density  
BarData$Density = BarData$Population / BarData$BrgyArea

Explanation: The barangay population density was calculated by dividing the barangay population by the barangay area calculated previously.

### Getting the Top 5 Population Densities for Barangays in the Philippines

#outputs the top 5 barangays in the Philippines  
Top5BrgyPH = slice\_max(BarData,Density,n=5)  
Top5BrgyPH

## Region Province CityProvince Barangay  
## 1 NATIONAL CAPITAL REGION NATIONAL CAPITAL REGION CALOOCAN CITY Barangay 176  
## 2 NATIONAL CAPITAL REGION NATIONAL CAPITAL REGION QUEZON CITY Commonwealth  
## 3 NATIONAL CAPITAL REGION NATIONAL CAPITAL REGION QUEZON CITY Batasan Hills  
## 4 NATIONAL CAPITAL REGION NATIONAL CAPITAL REGION CITY OF PASIG Pinagbuhatan  
## 5 NATIONAL CAPITAL REGION NATIONAL CAPITAL REGION QUEZON CITY Payatas  
## Population HouseholdPopulation Nhouseholds Area n BrgyArea Density  
## 1 246515 246439 55086 619.54 1706 0.3631536 678817.5  
## 2 198285 198235 45432 619.54 1706 0.3631536 546008.7  
## 3 161409 161352 34587 619.54 1706 0.3631536 444464.9  
## 4 151979 150800 33720 619.54 1706 0.3631536 418497.9  
## 5 130333 130277 28152 619.54 1706 0.3631536 358892.2

Explanation: *slice\_max* was used to get the 5 highest densities in the Philippines.

Note: We also determined the Top 5 barangays per Region and per City.

## City Density Data

### Data Manipulation

#Import CSV Files  
population = read.csv("population.csv")  
regionarea = read.csv("regionarea.csv")  
  
#Load dplyr library  
library("dplyr")  
  
#Count the Number of Cities per Region  
NumofCities= count(population,Region, CityProvince)  
NumOfCity = count(NumofCities, Region)  
  
#get the total population per city  
TotalPerCity = aggregate(Population ~ CityProvince, population, sum)  
  
#left join the count of the number of cities per region and the total population per city  
BarData = left\_join(population, regionarea, by=c("Region"))  
BarData = left\_join(BarData, NumOfCity, by=c("Region"))  
BarData = left\_join(BarData, TotalPerCity, by=c("CityProvince"))

Explanation: We first got the number of cities per region using the *count* function. We then calculated the total population per city using the *aggregate* command. Lastly, we used *leftjoin* to add the columns, Area (from regionarea.csv), Number of City (from count command), and Total Population (from aggregate command)

### City Area Calculation

Formula:

Code Snippet:

BarData$CityArea = BarData$Area / BarData$n

Explanation: In order to get the area per city, the total area of the region is divided by the total number of cities in the region.

### City Population Density Calculation

Formula:

Code Snippet:

BarData$Density = BarData$Population.y / BarData$CityArea

Explanation: The city population density was calculated by dividing the city population by the city area calculated previously.

### Getting the Top 5 Population Densities for Cities in the Philippines

#filter it to have just the city, region and density value  
TotalPerCity1 = aggregate(Density ~ CityProvince + Region ,BarData, mean)  
  
#output csv file with top 5 city population densities in the PH  
Top5CitiesPH = slice\_max(TotalPerCity1,Density,n=5)  
Top5CitiesPH

## CityProvince Region Density  
## 1 QUEZON CITY NATIONAL CAPITAL REGION 80566.18  
## 2 CITY OF MANILA NATIONAL CAPITAL REGION 48846.75  
## 3 CALOOCAN CITY NATIONAL CAPITAL REGION 43463.90  
## 4 TAGUIG CITY NATIONAL CAPITAL REGION 22086.64  
## 5 CITY OF PASIG NATIONAL CAPITAL REGION 20725.22

Explanation: *slice\_max* was used to get the 5 highest densities in the Philippines.

Note: We also determined the Top 5 cities per Region and per City.