Exploratory Analysis Project

Constantin Raether

2/19/2019

INTRODUCTION Greenhouse Gas is an opensource data set owned by the City of Los Angeles, California. The dataset describes greenhouse gas emission sources by industrial sectors and subsector within city limits and how much each sector emitted 4 different greenhouse gases within the year 2013. For my project, the variables I will explore are different sectors that are sources of greenhouse gas emissions and C02 emissions and CH4 emissions by each sector. A major initiative by current Los Angeles mayor Eric Garecetti is to reduce greenhouse gas emissions within city limits and put Los Angeles on a path of zero emissions. The initiative will need to identify sectors that emit greenhouse gases and create plans and agendas to reduce emissions within those sectors. This data can be found at The Mayor of Los Angeles’ Office of Sutstainablity: [linked phrase](https://data.lacity.org/A-Livable-and-Sustainable-City/Community-Wide-Greenhouse-Gas-Emissions/y3m9-i8tg)

library(ggplot2)  
library(knitr)  
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

library(readxl)  
GreenhouseGas <- read\_excel("/Users/coco/Documents/math130/GreenhouseGas.xlsx", sheet=1, col\_names=TRUE)

UNIVARIATE

table(GreenhouseGas$Sector)

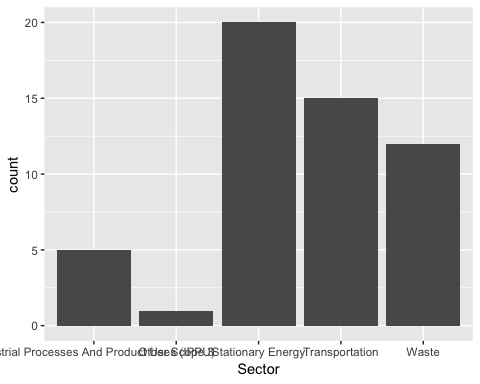
##   
## Industrial Processes And Product Uses (IPPU)   
## 5   
## Other Scope 3   
## 1   
## Stationary Energy   
## 20   
## Transportation   
## 15   
## Waste   
## 12

table(GreenhouseGas$'Sub-Sector')

##   
## Agriculture, forestry and fishing activities   
## 3   
## Aviation   
## 3   
## Biological treatment of waste   
## 3   
## Commercial and institutional buildings and facilities   
## 3   
## Energy industries   
## 4   
## Fugitive emissions from mining, processing, storage, and transportation of coal   
## 1   
## Incineration and open burning   
## 3   
## Manufacturing industries and construction   
## 3   
## Non-specified sources   
## 3   
## Off-road transportation   
## 3   
## On-road transportation   
## 3   
## Railways   
## 3   
## Residential buildings   
## 3   
## Solid waste disposal   
## 3   
## Wastewater treatment and discharge   
## 3   
## Waterborne navigation   
## 3

There were a total of 5 sectors that were measured for greenhouse gas contributions in 2013. Each sector had subsectors that further describe the exact source of greenhouse gas emissions.

ggplot(GreenhouseGas, aes(x=Sector)) + geom\_bar()



The bargraph shows the distribution and counts of the subsectors of each sector that was counted for emissions within city limits of Los Angeles in 2013. Stationary Energy and Transportation are significant sources of all greenhouse gases within city limits.

summary(GreenhouseGas$CO2)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 0 0 0 699072 126655 9155397

A summary of the distribution of CO2 (carbon dioxide) across all sectors.

summary(GreenhouseGas$CH4)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 0 0 0 11904 391 556802

A summary of the distribution of CH4 (methane) across all sectors.

select(GreenhouseGas, Sector, CO2, CH4)

## # A tibble: 53 x 3  
## Sector CO2 CH4  
## <chr> <dbl> <dbl>  
## 1 Stationary Energy 2689158 1419  
## 2 Stationary Energy 3314853 2545  
## 3 Stationary Energy 397782 305  
## 4 Stationary Energy 1415354 747  
## 5 Stationary Energy 5291479 4062  
## 6 Stationary Energy 634977 487  
## 7 Stationary Energy 2996486 1905  
## 8 Stationary Energy 814089 625  
## 9 Stationary Energy 97691 75  
## 10 Stationary Energy 784229 314  
## # … with 43 more rows

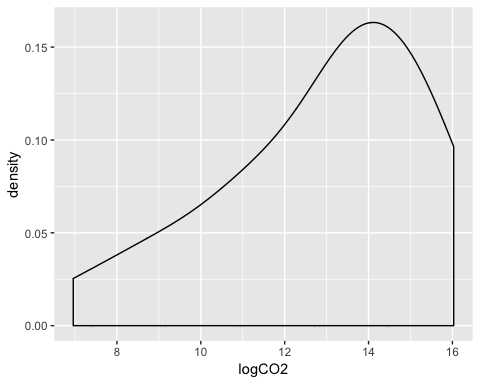
a <- mutate(GreenhouseGas, logCO2 = log(CO2), logCH4 = log(CH4))  
select(a, logCO2, logCH4)

## # A tibble: 53 x 2  
## logCO2 logCH4  
## <dbl> <dbl>  
## 1 14.8 7.26  
## 2 15.0 7.84  
## 3 12.9 5.72  
## 4 14.2 6.62  
## 5 15.5 8.31  
## 6 13.4 6.19  
## 7 14.9 7.55  
## 8 13.6 6.44  
## 9 11.5 4.32  
## 10 13.6 5.75  
## # … with 43 more rows

Now that we know how the distribution of greenhouse gases is highly variable across all sectors, I natural logged both greenhouse gases to give a more clear picture of the averages. By taking the natural log the distribution of greenhouse gas emissions across all sectors, we are able to find a more even distribution of the average of all CO2 and CH4 emissions.

ggplot(a, aes(x=logCO2)) + geom\_density()

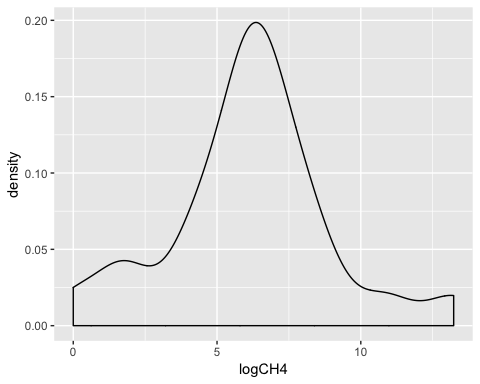
## Warning: Removed 33 rows containing non-finite values (stat\_density).



The density plot below show the average magnitude of CO2 across all sectors and that it ranges from 0 digits of recorded emissions to 16 digits of recorded emissions.

ggplot(a, aes(x=logCH4)) + geom\_density()

## Warning: Removed 30 rows containing non-finite values (stat\_density).



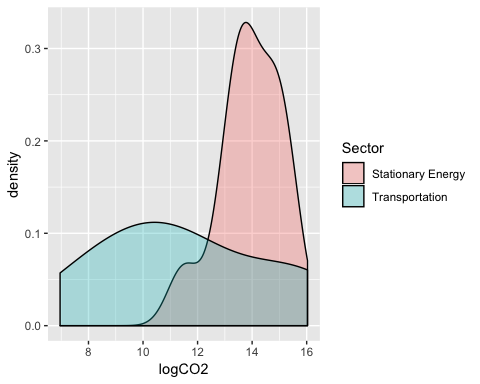
The density plot below show the average magnitude of CH4 across all sectors and that it ranges from 0 digits of recorded emissions to 12 digits of recorded CH4 emissions.

BIVARIATE

The univariate density plots gave us an average of both CO2 and CH4 emissions from all 5 sectors. But do all sectors emit both CO2 and CH4? Or are they only emitted by one or several sectors? I will use bivariate graphs to show which sectors emit either greenhouse gas and which sector emits more than others. This will give us a better understanding of which sectors we should focus on to reduce greenhouse gas emissions within city limits.

ggplot(a, aes(x=logCO2, fill=Sector)) + geom\_density(alpha=.3)

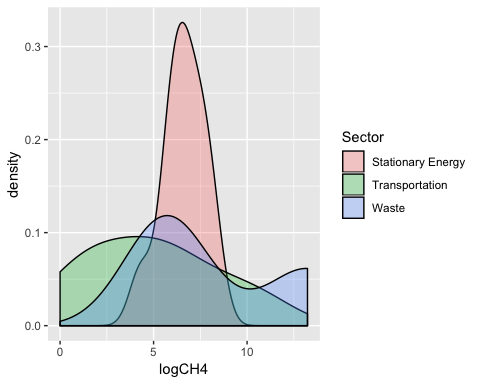
## Warning: Removed 33 rows containing non-finite values (stat\_density).



CO2 emissions is restricted to the sectors of Stationary Energy and Transportation. Stationary Energy consist of sub-sectors including residental buildings and construction industries and manufacturing. In order to reduce a significant source of CO2 emissions within city limits, efforts must be made to reduce CO2 emissions from residental sources but also the construction industry.

ggplot(a, aes(x=logCH4, fill=Sector)) + geom\_density(alpha=.3)

## Warning: Removed 30 rows containing non-finite values (stat\_density).



Stationary Energy is also the main source of CH4 emissions within city limits and the subsectors of residental buildings and the construction industry are the main culprits. Surpisingly, the waste sector is not the main source of CH4 emissions even with solid waste disposal being a significant source of methane worldwide.

In conclusion, in order to significantly reduce pernicious CO2 and CH4 gas emissions within city limits, Los Angeles will need to address the sectors Stationary Energy, Transportation, and Waste Disposal. How the city will address these issues with an increasing population, more cars on city roads, and more solid waste to dispose of is something that every politician, especially the mayors office, in the city will have to confront.