Industry GHG

This memo records the result from industry\_GHG.r.

* Manufacutre emission distribution (384 ind)

######### Part A. preparation  
library(dplyr)  
library(ggplot2)  
library(tidyr)  
#(i)load GHG data  
GHG\_IO=read.csv(file="GHGIO.csv",header=T, as.is=T)  
GHG\_IO[is.na(GHG\_IO)]=0  
GHG\_IO\_ind=GHG\_IO[,colnames(GHG\_IO)!="HE"]  
GHG\_IO\_F=GHG\_IO[,colnames(GHG\_IO)=="HE"]  
  
#load index mapping  
sector\_ind=read.csv(file="indcode\_20161223.csv",header=T, as.is=T)  
indname=read.csv(file="indname.csv",header=T, as.is=T)

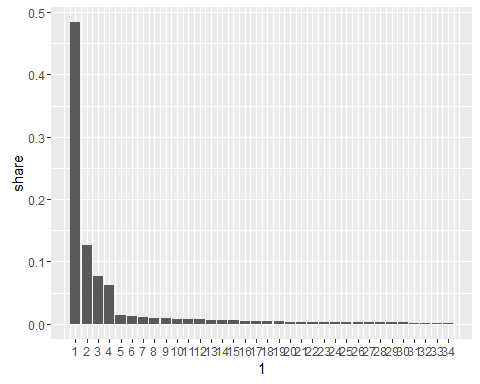
To select energy intensive industry, we obtain share of emission by each manufacture among total emission from manufactutre. So we excludes AFF (1-25), Mining(26-34), Refined Oil (99-108), Service (281-286,302-384), Construction (287-301) on IO table. Then we aggregate emissions by industries, and optain emission share for each industry. Then we select industries with largest emission that covers 90% of total manufacture emissions.

# obtain share and cumulative share (distribution function) of ghg from manufacture (35:98,109:271)  
GHG\_ind\_dist=GHG\_IO\_ind %>% gather(key="ind",value,c(35:98,109:271)) %>% group\_by(ind) %>% summarise(ind\_ghg=sum(value)) %>% mutate(share=ind\_ghg/sum(ind\_ghg)) %>% arrange(desc(share)) %>% mutate(dist=cumsum(share))  
  
#add names  
indname$xcode=paste("X",indname$basecode,sep="")  
GHG\_ind\_dist.N=merge(GHG\_ind\_dist,indname,by.x="ind",by.y="xcode",all.x=T)  
  
GHG\_ind\_dist.N=GHG\_ind\_dist.N[order(GHG\_ind\_dist.N$share,decreasing=T),]  
  
# keep top emitters covering 90% of emission  
Top\_ind=GHG\_ind\_dist.N %>% filter(dist<0.9) %>% arrange(desc(share)) %>% select(ind\_ghg,share,dist,basename)  
print(Top\_ind)

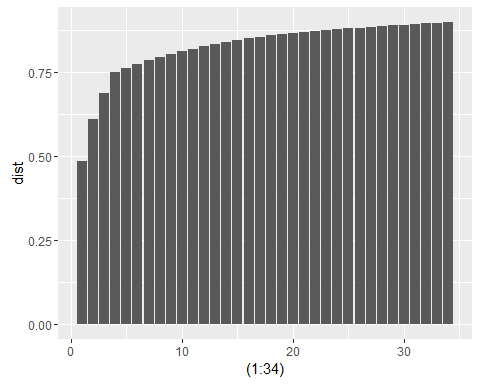
## ind\_ghg share dist basename  
## 1 72111376.2 0.484288357 0.4842884 선철  
## 2 18713514.5 0.125676941 0.6099653 지방족 기초유분  
## 3 11364981.3 0.076325379 0.6862907 방향족 기초유분  
## 4 9219385.1 0.061915901 0.7482066 시멘트  
## 5 2026555.8 0.013610021 0.7618166 기초무기화합물  
## 6 1846020.1 0.012397573 0.7742142 조강  
## 7 1533543.5 0.010299030 0.7845132 제조임가공서비스  
## 8 1368609.1 0.009191357 0.7937046 LCD 평판 디스플레이  
## 9 1336970.4 0.008978878 0.8026834 합금철  
## 10 1223687.2 0.008218086 0.8109015 석회 및 석고제품  
## 11 1186364.8 0.007967434 0.8188690 아스콘 및 아스팔트 제품  
## 12 1155165.1 0.007757902 0.8266269 석유화학중간제품  
## 13 998348.5 0.006704747 0.8333316 금속처리  
## 14 922339.9 0.006194286 0.8395259 전자기기용 유리제품  
## 15 889216.0 0.005971831 0.8454977 금속처리 가공품  
## 16 688842.9 0.004626158 0.8501239 자동차 부분품  
## 17 665187.0 0.004467288 0.8545912 승용차  
## 18 571569.9 0.003838571 0.8584297 레미콘  
## 19 539672.8 0.003624355 0.8620541 기타 화학제품  
## 20 502633.1 0.003375603 0.8654297 연 및 아연 제련, 정련 및 합금제품  
## 21 439575.3 0.002952117 0.8683818 석탄화합물  
## 22 411893.3 0.002766209 0.8711480 조립용 플라스틱제품  
## 23 406978.9 0.002733205 0.8738812 기타 기초유기화합물  
## 24 396898.2 0.002665504 0.8765467 집적회로  
## 25 381932.8 0.002564999 0.8791117 도료  
## 26 374325.6 0.002513910 0.8816256 합성수지  
## 27 338845.1 0.002275629 0.8839013 금속 주물  
## 28 336252.9 0.002258220 0.8861595 콘크리트제품  
## 29 329773.3 0.002214704 0.8883742 산업용(전자기기 제외) 유리제품  
## 30 326017.2 0.002189479 0.8905637 기타 비철금속 1차제품  
## 31 299773.4 0.002013230 0.8925769 의약품  
## 32 297397.7 0.001997275 0.8945742 냉간압연강재  
## 33 294254.6 0.001976166 0.8965503 이동전화기  
## 34 291090.9 0.001954920 0.8985053 봉제의류

The pdf and cdf of industrial emission is given as follows.

Top\_ind %>%ggplot(aes(x=(1:34),y=share))+geom\_bar(stat="identity")+scale\_x\_continuous(breaks=c(1:34),labels(Top\_ind$basename))



Top\_ind %>%ggplot(aes(x=(1:34),y=dist))+geom\_bar(stat="identity")



* Manufacutre emission distribution (49 ind)

#(ii)preparing index  
#row\_ind=sector\_ind[,1:2]  
col\_ind=sector\_ind[,3:4]  
sec\_group=sector\_ind[,5:8]  
#sec\_BR=sector\_ind[,c(5,6,15,16)]  
  
#IOind\_model=row\_ind[(1:384),]  
sec\_group=sec\_group[!is.na(sec\_group[,1]),]  
#sec\_BR=sec\_BR[!is.na(sec\_BR[,1]),]  
sectorname=sec\_group[,1:2]

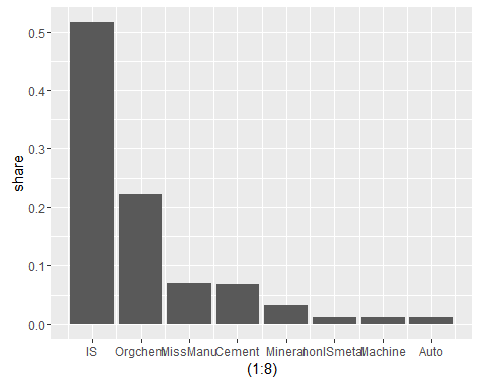
We aggregated industry emission of 384 industries according to our standard model sector definition. Then we select sectors with largest emission that covers 95% of total manufacture emissions.

GHG\_ind\_core=GHG\_IO\_ind %>% gather(key="ind",value,c(35:98,109:271)) %>% group\_by(ind) %>% summarise(ind\_ghg=sum(value))  
  
col\_ind$xcode=paste("X",col\_ind$basecode\_column,sep="")  
  
GHG\_sector=merge(GHG\_ind\_core,col\_ind,by.x="ind",by.y="xcode",all.x=T)  
GHG\_sector\_dist=GHG\_sector %>% group\_by(sector\_column) %>% summarize(sector\_ghg=sum(ind\_ghg))%>% mutate(share=sector\_ghg/sum(sector\_ghg)) %>% arrange(desc(sector\_ghg)) %>% mutate(dist=cumsum(share))  
sectorname=sec\_group[,1:2]  
  
GHG\_sector\_dist.N=merge(GHG\_sector\_dist,sectorname, by.x="sector\_column",by.y="sector\_ind",all.x=T,sort=F)  
  
Top\_sector=GHG\_sector\_dist.N %>% filter(dist<0.95) %>% arrange(desc(share))   
print(Top\_sector)

## sector\_column sector\_ghg share dist sector\_name  
## 1 10 76921198 0.51659034 0.5165903 IS  
## 2 12 33158859 0.22268954 0.7392799 Orgchem  
## 3 25 10452199 0.07019528 0.8094752 MissManu  
## 4 11 10127208 0.06801269 0.8774879 Cement  
## 5 15 4763822 0.03199306 0.9094809 Mineral  
## 6 16 1784716 0.01198586 0.9214668 nonISmetal  
## 7 17 1780003 0.01195421 0.9334210 Machine  
## 8 22 1613792 0.01083796 0.9442590 Auto

The pdf and cdf of industrial emission is given as follows. Notice that only the top 4 industries has larger than 5% share, and they take up 88% of industry emission. (It would be sufficient to have separate model for three. IS/Orgchem/Cement)

Top\_sector%>%ggplot(aes(x=(1:8),y=share))+geom\_bar(stat="identity")+scale\_x\_continuous(breaks=(1:8),labels=(Top\_sector$sector\_name))



Top\_sector%>%ggplot(aes(x=(1:8),y=dist))+geom\_bar(stat="identity")+scale\_x\_continuous(breaks=(1:8),labels=(Top\_sector$sector\_name))

