

Estudio de emisiones de Carbono

Grupo Ladef

Julio 2021

Este estudio está destinado a documentar las etapas de análisis y visualización de datos de las emisiones de carbono, producidas entre 2009 y 2011 tomadas de la Universidad de Maine. Este proyecto fue el resultado del trabajo de Mark Anderssen, Mario Teisl y Eva Manandhar de la Facultad de Economía de la Universidad de Maine. Este trabajo ha sido modificado por grupo Ladef para la cátedra Técnicas y Herramientas Modernas I de la Universidad Nacional de Cuyo.

Cargar primer conjunto de datos de las emisiones de CO2 SSI

```
d <- data.frame(read.csv("e:/CO2.csv"))
```

Preparar la primer tabla de datos para analizar

```
levels(d$Discip) <- c("na", "Natural", "Social")
d$Name<- as.factor(d$Name)
d$Cat<- as.factor(d$Cat)
d$Discip<- as.factor(d$Discip)
d$Type <- as.factor(d$Type)
levels(d$Type) <- c("Visitor", "Full", "Assoc.", "Assist.",
  "Post Doc.", "PhD Student", "MS Student", "Admin.", "other")
d$Type2 <- d$Type
levels(d$Type2) <- c("Visitor/other", "Professor", "Professor",
  "Professor", "Student/Post Doc", "Student/Post Doc",
  "Student/Post Doc", "Administrator", "Visitor/other")
levels(d$Cat) <- c("Student", "Faculty/Staff")
d$Type3 <- d$Type2
levels(d$Type3) <- c("Admin/Visitor/other", "Professor/Student",
  "Professor/Student", "Admin/Visitor/other")
```

VALIDACIÓN - deberían ser 98 observaciones de 10 variables

```
str(d)
```

```
## 'data.frame': 98 obs. of 10 variables:
## $ Rank_ann : int 1 2 3 4 5 6 7 8 9 10 ...
## $ Rank_tot : int 11 23 25 26 1 2 7 3 12 14 ...
## $ Name : Factor w/ 98 levels "G01","G02","G03",...: 80 15 87 26 53 67 22 41 59 27 ...
## $ CO2annual: int 2561 1720 1698 1611 1596 1589 1560 1506 1261 1248 ...
## $ CO2total : int 2561 1720 1698 1611 4787 4768 3119 4518 2523 2496 ...
## $ Cat : Factor w/ 2 levels "Student","Faculty/Staff": 2 1 2 1 2 2 1 2 2 1 ...
## $ Type : Factor w/ 9 levels "Visitor","Full",...: 6 6 1 6 2 5 6 4 4 6 ...
## $ Discip : Factor w/ 3 levels ".", "1", "2": 2 3 1 3 3 3 2 3 2 3 ...
## $ Type2 : Factor w/ 4 levels "Visitor/other",...: 3 3 1 3 2 3 3 2 2 3 ...
## $ Type3 : Factor w/ 2 levels "Admin/Visitor/other",...: 2 2 1 2 2 2 2 2 2 2 ...
```

Cargar segundo conjunto de datos de las emisiones de CO2 SSI

```
carbonTW <- data.frame(read.csv("e:/carbonTW.csv"))
```

Preparar la segunda tabla de datos

```
carbonTW <- carbonTW[!is.na(carbonTW$month), ]
carbonTW$monthfact <- as.factor(carbonTW$month)
levels(carbonTW$monthfact) <- c("Jan", "Feb", "Mar", "Apr", "May",
  "Jun", "Jul", "Aug", "Sep", "Oct", "Nov", "Dec")

attach(carbonTW)
carbonTW$season <- "NA"
carbonTW$season[month < 6] <- "Spring"
carbonTW$season[month >= 6 & month < 9] <- "Summer"
carbonTW$season[month >= 9] <- "Fall"
carbonTW$fiscal <- "NA"
carbonTW$fiscal[year == 2009] <- "f2010"
carbonTW$fiscal[year == 2010 & month < 7] <- "f2010"
carbonTW$fiscal[year == 2010 & month >= 7] <- "f2011"
carbonTW$fiscal[year == 2011 & month < 7] <- "f2011"
carbonTW$fiscal[year == 2011 & month >= 7] <- "f2011"
carbonTW$fiscal[year == 2012] <- "f2012"
carbonTW$fiscal <- as.factor(carbonTW$fiscal)
carbonTW$allmonths <- "NA"
carbonTW$allmonths <- year - 2009 + month/12
detach(carbonTW)

carbonTW$Name<- as.factor(carbonTW$Name)
carbonTW$perstype<- as.factor(carbonTW$perstype)
carbonTW$striptype<- as.factor(carbonTW$striptype)
carbonTW$dissem <- carbonTW$striptype
carbonTW$dissem = factor(carbonTW$dissem, levels(carbonTW$dissem)[c(1, 2, 4, 3)])
levels(carbonTW$dissem) <- c("Research", "Dissemination",
  "Research", "other")
```

VALIDACIÓN - deberían ser 403 observaciones de 17 variables

```
str(carbonTW)
```

```
## 'data.frame':    403 obs. of  17 variables:
## $ num           : int  1 2 3 4 5 6 7 8 9 11 ...
## $ TripID        : int  40 48 41 49 42 50 99 1 23 25 ...
## $ year          : int  2009 2009 2009 2009 2009 2009 2009 2009 2009 2009 ...
## $ month         : int  9 9 10 10 10 10 11 11 11 11 ...
## $ fiscal        : Factor w/ 3 levels "f2010","f2011",...: 1 1 1 1 1 1 1 1 1 1 ...
## $ total_car_miles: num  526 526 228 228 278 ...
## $ total_car_CO2  : num  196.9 196.9 85.4 85.4 104.1 ...
## $ total_air_CO2  : num  NA NA NA NA NA ...
## $ Total_CO2_kg   : num  196.9 196.9 85.4 85.4 104.1 ...
## $ Name          : Factor w/ 82 levels "G01","G02","G03",...: 19 19 19 19 19 19 56 49 35 2 ...
## $ perstype      : Factor w/ 2 levels "staff","student": 2 2 2 2 2 2 1 1 2 2 ...
## $ persrank      : int  6 6 6 6 6 6 0 0 6 6 ...
## $ triptype      : Factor w/ 4 levels "admin","confpres",...: 3 3 3 3 3 3 1 1 3 3 ...
## $ monthfact     : Factor w/ 12 levels "Jan","Feb","Mar",...: 9 9 10 10 10 10 11 11 11 11 ...
## $ season        : chr  "Fall" "Fall" "Fall" "Fall" ...
## $ allmonths     : num  0.75 0.75 0.833 0.833 0.833 ...
```

```
## $ dissem      : Factor w/ 3 levels "Research","Dissemination",...: 3 3 3 3 3 3 1 1 3 3 ...
```

Calcular un resumen simple de las estadísticas

```
summary(d$CO2annual)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      14.0  155.0   466.5   558.6   800.0  2561.0
```

```
length(d$CO2annual[d$CO2annual > 1000])
```

```
## [1] 15
```

```
summary(carbonTW$Total_CO2_kg)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      7.49  64.02  104.82  249.42  326.23 1617.63
```

```
ddply(carbonTW, .(fiscal), summarize, car_trips =
  NROW(!is.na(total_car_miles)),
  car_miles = sum(!is.na(total_car_miles)), car_CO2 =
  sum(!is.na(total_car_CO2)),
  air_trips = NROW(total_air_CO2), air_CO2 =
  sum(!is.na(total_air_CO2)), total_CO2 = sum(Total_CO2_kg),
  .inform = TRUE, .drop = TRUE)
```

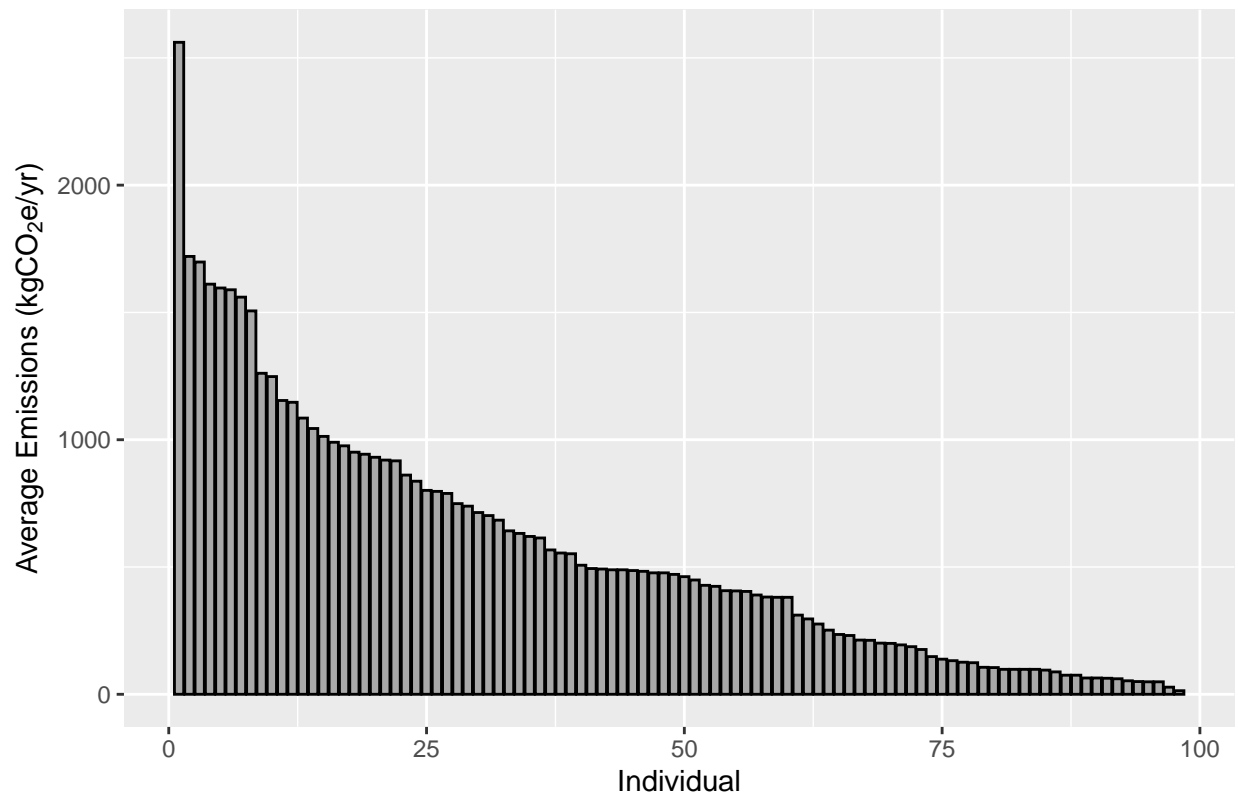
```
##      fiscal car_trips car_miles car_CO2 air_trips air_CO2 total_CO2
## 1 f2010         115         96      96      115      23  22498.60
## 2 f2011         201        140     139      201      67  58483.63
## 3 f2012          87         67      67       87      20  19534.17
```

Gráficos

Promedios anuales por individuo

```
ggplot(d, aes(x = Rank_ann, y = CO2annual)) + geom_bar(colour = "black", fill = "darkgrey",
  stat = "identity") + labs(y = expression("Promedio de Emisiones (kgCO" [2] *
  "e/yr)"), x = "Individual") + ggtitle("Emisiones Individuales") +
  theme(legend.position = "none")
```

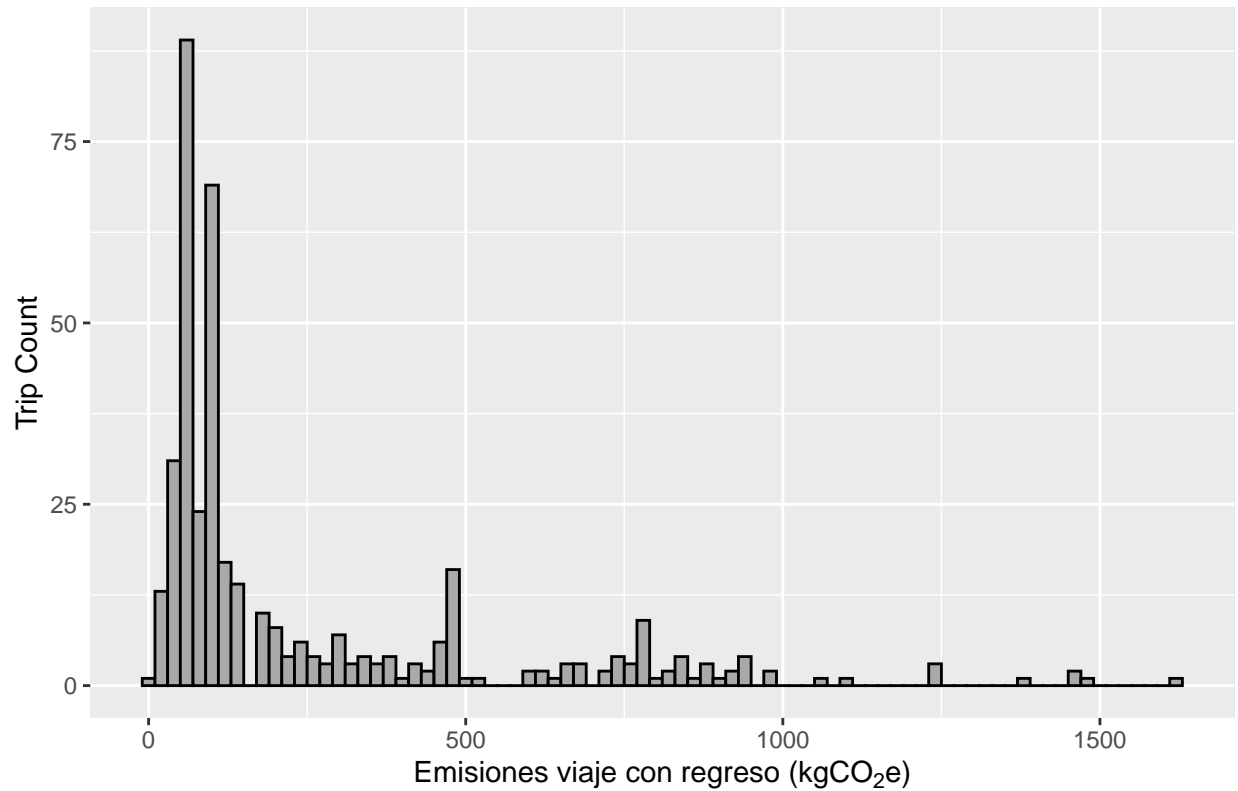
Individual Emissions



Histograma de Emisiones de viaje

```
ggplot(carbonTW, aes(Total_CO2_kg)) + geom_histogram(colour = "black",  
  fill = "darkgrey", binwidth = 20) + labs(x = expression(  
  "Emisiones viaje con regreso (kgCO"  $2$  " * "e)" ),  
  y = "Trip Count") + ggtitle("Frecuencia de viajes ida/vuelta")
```

Frecuencia de viajes ida/vuelta



```
ggplot(carbonTW, aes(x = dissem, y = Total_CO2_kg)) +
  geom_jitter(aes(colour = dissem), position =
    position_jitter(width = 0.3), alpha = 0.85, size = 3) +
  geom_boxplot(alpha = 0, outlier.colour = "red",
    outlier.shape = NA) + labs(x = " muestreo poblacion",
    y = expression("Emisiones Totales en (kgC0"[2] * "e)")) +
  ggtitle("Emisiones por motivo de viaje") +
  theme(legend.position = "none")
```

Emisiones por motivo de viaje

