

Análisis Estratificado de Consumo de Carne

Nestor Torrech (801-17-4120) | Miguel Torres (801-19-8755) | Daniela Vicente (801-17-7932)

5/24/2022

Resumen

En este estudio se busca analizar la correlación que tiene el consumo de cuatro tipos de carne con emisiones de CO₂, PIB, y con Pobreza extrema. Para esto se consiguió una base de datos que recopila información de consumo de cuatro tipos de carne. Esta base de datos luego se muestreo utilizando el método estratificado en donde cada estrato era una de las carnes. Después esto se unió a otras tablas con las variables de CO₂, PIB y Pobreza extrema que se mencionaron anteriormente. En conclusión, los resultados esperados probaron ser parcialmente correctos.

Las carnes tuvieron alta correlación con emisiones de CO₂ anuales. Sin embargo, con emisiones per cápita, solo dos de las carnes probaron tener buena correlación con esta variable. En el caso de PIB, el estudio probó inconcluso debido a la alta cantidad de valores que tuvieron que ser imputados. Finalmente, con la pobreza extrema, se encontró que dos tipos de carne tenían una correlación leve y negativa con la pobreza extrema.

Introducción

El consumo de carne y el producto interno bruto (PIB) tienen un vínculo; cuando un país crece y su PIB aumenta, las personas pueden permitirse comprar más productos cárnicos. Como resultado, el consumo de carne está fuertemente influenciado por el crecimiento económico. La carne, que tradicionalmente es una fuente más cara de calorías y proteínas, se vuelve más asequible a medida que aumentan los ingresos. Por consiguiente, el consumo de carne aumenta cuando un país crece y su producto interno bruto aumenta porque las personas pueden permitirse comprar más productos cárnicos. No obstante, cuando los consumidores se vuelven conscientes de cómo comer carne puede afectar negativamente su salud y el medio ambiente, los niveles de consumo de carne comienzan a estancarse y luego caen después de que el PIB alcanza un nivel particular. Existe un vínculo entre el PIB real per cápita y el consumo de carne per cápita; debido a que cuanto más bajo es el nivel de ingresos de un país, menos carne consume, y cuanto más alto es el nivel de ingresos, más carne consume. (Nozaki, 2016) Por lo que, el consumo de carne per cápita aumenta junto con el crecimiento del PIB real per cápita. Esto podría deberse a que, a diferencia de otros alimentos básicos, la carne no es un producto esencial y su consumo se ve fácilmente afectado por el nivel de ingresos. El consumo de carne, por otro lado, tiene un impacto pernicioso para el medio ambiente ya que, contribuye a la liberación de gases de efecto invernadero como el metano y el CO₂, entre otros. Por lo tanto, estos gases coadyuva al calentamiento global y al cambio climático. A continuación, se presenta un estudio sobre las posibles relaciones entre el consumo de diversos tipos de carne con el producto interno bruto y/o la emisión de dióxido de carbono. Por consiguiente, se definió como hipótesis para propósitos de esta investigación la siguiente: el consumo de carnes cuentan con correlaciones altas positivas con emisiones de CO₂ y con crecimiento económico, y en el caso de pobreza extrema, el consumo de carnes tendrá una correlación negativa.

Revisión Literaria

El consumo de carne se ha desplazado hacia las aves. Esto refleja el precio más barato del pollo en comparación con otras carnes en los países en desarrollo de bajos ingresos, mientras que demuestra una mayor preferencia por las carnes blancas en los países de altos ingresos, que son más fáciles de cocinar y se consideran una opción alimenticia más saludable. En 2030, se prevé que la carne de ave representa el 41% del total de proteínas derivadas de fuentes cárnicas en todo el mundo, 2 puntos porcentuales más que en el período de referencia. Otros productos cárnicos tienen participaciones mundiales más pequeñas: carne de res (20%), carne de cerdo (34%) y carne de ovina (5%) (OECD-FAO Agricultural Outlook, 2021).

De acuerdo con un estudio exploratorio y mixto Impacto de la renta y la privación social en el consumo de carne en México, entre 2007 y 2016 el consumo per cápita de carne bovina en México disminuyó 2.1 por ciento anual, de 18 a 14.8 kg. Mientras tanto, el consumo per cápita de pollo y cerdo muestra un patrón

diferente, creciendo a una tasa anual de 2 y 3.3%, respectivamente, en el mismo período (Sanabria et.al., 2018). El objetivo de este estudio, sin embargo, fue observar los efectos de la privación económica y social en el consumo de carne. Luego de completar un cuestionario y calcular una muestra de 429 consumidores, los consumidores fueron clasificados en uno de cuatro grupos económicos (pobreza multidimensional moderada o extrema, vulnerable por carencia social, vulnerable por ingreso y no multidimensional pobre y no vulnerable). También, se identificaron tres agrupaciones económicas: consumo de carne de res, consumo de carne de cerdo y consumo de carne de pollo. Los resultados de las pruebas T y F revelan que la variación en el consumo de carne de res, cerdo y pollo depende del ingreso obtenido por las tres categorías económicas estudiadas y la cantidad de carencias sociales que enfrenta cada uno de ellos. La variación en el consumo de carne se da principalmente entre el grupo de consumidores vulnerables por carencia social y los consumidores del grupo de pobres multidimensionales extremos; es decir, la variación en el consumo de carne se da entre los menos pobres y los más pobres del grupo de consumidores encuestados. Esto demuestra que los hogares consumidores consumen más carne de pollo porque es la más barata. La carne de res se consume en segundo lugar porque su precio ha subido, lo que la convierte en un bien de difícil acceso. El cerdo es la carne menos popular, debido a su alto costo y la percepción entre los consumidores de que es la menos saludable de las tres. Puede afirmarse que el crecimiento del precio general de la carne de res en el mercado ha impactado a la población mexicana. Además, las amas de casa optan por la carne de pollo porque es la carne más fácilmente disponible y adaptable.

Por otra parte, según un estudio realizado por Bereznicka y Pawlonka (2018), investigaron y analizaron la asociación entre el consumo de carne y los ingresos. Por lo que definieron dos hipótesis. La primera hipótesis fue que entre los encuestados polacos, el consumo de carne y productos cárnicos aumentó junto con el ingreso familiar por miembro. La segunda hipótesis fue que entre los encuestados austriacos, el consumo de carne y productos cárnicos aumentó junto con el ingreso familiar por miembro. La categoría más pobre de consumidores austriacos tiene un ingreso de alrededor de EUR 316 por miembro del hogar, según la investigación. Mientras que el valor de ingreso más bajo por miembro del hogar para clientes polacos es de alrededor de EUR 160. De acuerdo con la hipótesis proporcionada por H1, el consumo de carne y productos cárnicos aumenta a la par que aumenta el ingreso familiar por miembro. Sin embargo, a medida que aumentan los ingresos, también lo hace el consumo, lo que da como resultado un crecimiento inferior al proporcional en el consumo de carne y productos animales. Para los consumidores austriacos, sin embargo, se descubrió que, de acuerdo con la hipótesis H2, un aumento en el ingreso familiar va acompañado de un aumento en el consumo de carne y productos cárnicos, pero solo cuando el ingreso familiar no supera el 170% del ingreso promedio por persona. miembro del hogar en Austria (Bereznicka & Pawlonka, 2018). Más aún, según este estudio los ingresos son un factor significativo del consumo de carne y productos cárnicos en las economías emergentes. La carne, por otro lado, se considera un bien ordinario, lo que tiene como consecuencia una disminución de la utilidad marginal.

Según Masuda y Goldsmith (2010), el crecimiento del ingreso per cápita de China, así como el crecimiento de la población y la urbanización respaldan el aumento del consumo de carne per cápita y en total. Se empleó un modelo vectorial de corrección de errores para estimar las ecuaciones de cointegración o las elasticidades ingreso a largo plazo. Donde proporcionaron las elasticidades de ingreso estimadas y el ingreso per cápita para China, las cantidades de consumo de carne de China se proyectan utilizando una forma recursiva para 2030. Los resultados mostraron que, con respecto al aumento de ingresos, el consumo de carne de cerdo es inelástico, carne de cordero y cabra y carne de ave tienen una elasticidad unitaria, mientras que el consumo de carne bovina es elástico. Además, como el PIB crece anualmente al 4.9 % en 2010-20 y al 2.8 % en 2020 hasta 2030, el consumo total de carne aumenta al 3.2 % y al 2.0 %, respectivamente, y alcanza los 136 mil toneladas métricas en 2030. Además, con un crecimiento anual del PIB del 4.9 % de 2010 a 2020 y del 2.8 % de 2020 a 2030, el consumo total de carne aumenta un 3.2 % y un 2 %, respectivamente, alcanzando los 136 millones de toneladas métricas en 2030 (Masuda & Goldsmith, 2010).

Por otra parte, de acuerdo con otro artículo donde la hipótesis era que el cambio en el PIB per cápita estaba asociado con cambios en el consumo de diferentes tipos de carne, se construyó un modelo de regresión lineal de cambios en el PIB per cápita en el tiempo (años) para cada país. Luego, se determinó la asociación entre el cambio en el PIB per cápita y el cambio en el consumo de carne per cápita utilizando modelos de regresión lineal. Dicho esto, se observó una tendencia lineal decreciente en la carne de res como proporción del consumo total de carne en la mayoría de los países estudiados (23 de 35). En contraste con la contribución

proporcional decreciente de la carne de res y cerdo al consumo mundial total de carne, la contribución de las aves de corral aumentó del 33% al 43.4% entre 2000 y 2019. Por otro lado, se observó una reducción en la contribución de la carne de ovino al consumo mundial. consumo total de carne donde para el año 2000 el aporte del ovino fue del 5.6% en el 2000 y en el 2019 que fue del 5.1% (Whitton et.al., 2021). El aumento del PIB se asoció con cambios en el consumo total de carne en muchos países, pero no en ciertos países de ingresos altos. Estas tendencias probablemente reflejan el apetito del consumidor y la industrialización tecnológica histórica. El PIB más alto entre los países del Grupo 1, en los que el consumo de carne se relaciona positivamente con el cambio en el PIB, fue Japón con \$40,247. Esto sugiere que con un PIB per cápita superior a los \$40,000, el crecimiento del PIB ya no es un motor del crecimiento del consumo de carne (Whitton et.al., 2021).

Un estudio que analizó cómo el consumo de carne afecta las emisiones de gases de efecto invernadero al tiempo que tiene en cuenta el crecimiento económico y el uso de energía. Los hallazgos empíricos muestran que todos los tipos de consumo de carne aumentan las emisiones de gases de efecto invernadero a largo plazo. Sin embargo, la relación entre el consumo de carne y las emisiones de gases de efecto invernadero parece no ser lineal. Las pruebas para una conexión en forma de U muestran claramente que las emisiones de metano y óxido nitroso del aumento del consumo de carne en los Estados Unidos se pueden reducir con la formulación de políticas adecuadas y la implementación de tecnología relevante. Cabe mencionar, que el consumo de carne de res resultó tener un mayor impacto adverso en el medio ambiente, ya que libera una mayor cantidad de emisión de dióxido de carbono que los otros tipos de carne (Shafiulla et.al. 2021). El crecimiento económico está impulsado por el consumo total de carne y el consumo de carne de res. Esto indica que cuando aumenta el consumo total de carne y el consumo de carne de res per cápita, también lo hace el PIB real per cápita (Shafiullah et.al., 2021).

Metodologia

Primeramente, conseguimos los datos de consumo de carne que se estará analizando.

```
library(tidyverse)
```

```
## -- Attaching packages ----- tidyverse 1.3.1 --
```

```
## v ggplot2 3.3.5      v purrr  0.3.4
## v tibble  3.1.4      v dplyr  1.0.7
## v tidyr   1.1.4      v stringr 1.4.0
## v readr   2.1.1      v forcats 0.5.1
```

```
## Warning: package 'stringr' was built under R version 4.1.3
```

```
## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()     masks stats::lag()
```

```
library(ggplot2)
library(latexpdf)
```

```
dat_meat <-
```

```
read_csv("C:\\Users\\nesto\\Documents\\School\\ESTA 5502\\Project Meat Temp\\meat_consumption_worldwide
```

```
## Rows: 13760 Columns: 5
```

```
## -- Column specification -----
## Delimiter: ","
## chr (3): LOCATION, SUBJECT, MEASURE
## dbl (2): TIME, Value
```

```
##
## i Use 'spec()' to retrieve the full column specification for this data.
## i Specify the column types or set 'show_col_types = FALSE' to quiet this message.
```

```
head(dat_meat)
```

```
## # A tibble: 6 x 5
##   LOCATION SUBJECT MEASURE  TIME Value
##   <chr>    <chr>    <chr>  <dbl> <dbl>
## 1 AUS      BEEF      KG_CAP  1991  27.7
## 2 AUS      BEEF      KG_CAP  1992  26.2
## 3 AUS      BEEF      KG_CAP  1993  26.2
## 4 AUS      BEEF      KG_CAP  1994  25.5
## 5 AUS      BEEF      KG_CAP  1995  25.3
## 6 AUS      BEEF      KG_CAP  1996  27.3
```

```
tail(dat_meat)
```

```
## # A tibble: 6 x 5
##   LOCATION SUBJECT MEASURE      TIME Value
##   <chr>      <chr>   <chr>    <dbl> <dbl>
## 1 MYS        SHEEP    THND_TONNE 2021  43.4
## 2 MYS        SHEEP    THND_TONNE 2022  44.2
## 3 MYS        SHEEP    THND_TONNE 2023  44.9
## 4 MYS        SHEEP    THND_TONNE 2024  45.6
## 5 MYS        SHEEP    THND_TONNE 2025  46.3
## 6 MYS        SHEEP    THND_TONNE 2026  47.0
```

Como podemos ver, nuestros datos están compuesto por una serie de tiempo que sigue la tasa de consumo de carne de varios países desde 1990 a 2026. Sin embargo, hay unos cambios que tenemos que hacer.

1. Tenemos que redondear los datos de Value
2. Queremos que todos los países en el análisis tengan datos de 1990 a 2018, lo cual compone los datos históricos de esta tabla
3. En MEASURE, existen dos tipos (KG_CAP y THND_TONNE). Para propósitos de estandarización, nos quedaremos con solamente una de estas medidas. Durante el transcurso de esta fase exploratoria, determinaremos cual queremos

Limpieza

```
clean_meat <- dat_meat %>%
  mutate(Value = round(Value,2)) %>%
  group_by(LOCATION, SUBJECT) %>%
  filter(TIME %in% 1990:2018, TIME != is.na(TIME), LOCATION != "BRICS",
         LOCATION != "EU28", LOCATION != "OECD" , LOCATION != "WLD", LOCATION != "SSA") %>%
  arrange(LOCATION)
list("Observaciones" = dim(clean_meat)[1])
```

```
## $Observaciones
## [1] 9686
```

Explicacion

Utilizando la funciones de Tidyverse, pudimos resolver los primeros dos asuntos que teníamos con la tabla. Ahora solo tenemos observaciones que ocurrieron entre 1990 a 2018 y también redondeamos los datos de la tasa de consumo. Esto nos deja con 9,686 observaciones al momento.

Se debe notar que también se eliminaron de 'LOCATION' lo que son las clasificaciones BRICS, EU28, OECD, WLD y SSA ya que dichas clasificaciones se componen de multiples países

KG_CAP o THND_TONNE?

Si se acuerdan, al principio nosotros discutimos que estos datos tenian dos metricas para medir el consumo de carne. KG_CAP y THND_TONNE. Para poder trabajar con los datos de manera adecuada, tendremos que quedarnos con solamente una de las dos medidas.

Para tomar una decisión informada, exploremos los datos un poco.

Cantidad de observaciones agrupado a base de las medidas

```
clean_meat %>%
  group_by(MEASURE) %>%
  count()
```

```
## # A tibble: 2 x 2
## # Groups:   MEASURE [2]
##   MEASURE      n
##   <chr>      <int>
## 1 KG_CAP      4713
## 2 THND_TONNE  4973
```

Explicación Como vemos, en términos de distribución de datos THND_TONNE tiene un poquito más de observaciones que KG_CAP. Sin embargo, no es tanta la diferencia. Pero, para ir a la segura, miremos a los casos en donde solo existe THND_TONNE.

ANTI-JOIN

Primero, separamos las tablas a base de SUBJECT via nuevas variables.

```
clean_meat_kg <- data.frame(clean_meat %>% filter(MEASURE == "KG_CAP"))
clean_meat_ton <- data.frame(clean_meat %>% filter(MEASURE == "THND_TONNE"))
```

```
head(anti_join(clean_meat_ton, clean_meat_kg, by = c("LOCATION", "TIME")))
)
```

```
##   LOCATION SUBJECT    MEASURE TIME Value
## 1     CHE     BEEF THND_TONNE 1990   173
## 2     CHE     BEEF THND_TONNE 1991   178
## 3     CHE     BEEF THND_TONNE 1992   176
## 4     CHE     BEEF THND_TONNE 1993   171
## 5     CHE     BEEF THND_TONNE 1994   157
## 6     CHE     BEEF THND_TONNE 1995   156
```

Se recomienda correrlo sin "Head" para ver el resultado mas a fondo

```
# anti_join(clean_meat_ton, clean_meat_kg, by = c("LOCATION", "TIME"))
```

Usando este método, vemos que hay 260 datos de THND_TONNE que se NO tiene un match con la tabla de KG_CAP

1. UKR no tiene datos en KG_CAP durante el principio de los 90s
2. NOR no tiene datos en KG_CAP

Verifiquemos si hay datos en clean_meat_kg que no están en clean_meat_ton

```
anti_join(clean_meat_kg, clean_meat_ton, by = c("LOCATION", "TIME"))
```

```
## [1] LOCATION SUBJECT MEASURE TIME      Value
## <0 rows> (or 0-length row.names)
```

Como podemos ver en los resultados, todas las entradas en `clean_meat_kg` tienen su contraparte en `clean_meat_ton`. Tomando esto en cuenta, nos encontramos en la situación en donde la única medida que tiene entradas independientes es **THND_TONNE**

Decision

Con estos factores en mente, vamos a decidir quedarnos con la medida de **THND_TONNE** para contener la mayor parte de los datos posibles.

```
head(clean_meat_ton)
```

```
##   LOCATION SUBJECT    MEASURE TIME  Value
## 1     ARG POULTRY THND_TONNE 1990 337.82
## 2     ARG POULTRY THND_TONNE 1991 393.04
## 3     ARG POULTRY THND_TONNE 1992 586.10
## 4     ARG POULTRY THND_TONNE 1993 714.00
## 5     ARG POULTRY THND_TONNE 1994 726.60
## 6     ARG POULTRY THND_TONNE 1995 803.22
```

Preparando los Estratos

Ya que tenemos la medida particular que estaremos utilizando, ahora vamos a preparar variables para cada tipo de carne

Pero antes de eso, verifiquemos los números de cada tipo de carne a ver si está distribuido bien

```
clean_meat_ton %>% group_by(SUBJECT) %>% count()
```

```
## # A tibble: 4 x 2
## # Groups:   SUBJECT [4]
##   SUBJECT      n
##   <chr>   <int>
## 1 BEEF     1243
## 2 PIG      1242
## 3 POULTRY  1244
## 4 SHEEP    1244
```

```
meat_poultry_ton <- clean_meat_ton %>% filter(SUBJECT == "POULTRY")
meat_beef_ton <- clean_meat_ton %>% filter(SUBJECT == "BEEF")
meat_pig_ton <- clean_meat_ton %>% filter(SUBJECT == "PIG")
meat_sheep_ton <- clean_meat_ton %>% filter(SUBJECT == "SHEEP")
```

Ahora que tenemos nuestros estratos definidos, tenemos que buscar el tamaño de las muestras para cada estrato. Para esto primero sacaremos la muestra para los datos completos


```

tamano_mu_exacto <- function(N,s,e,C) #Usar cuando tengo el N
{prob=(C+1)/2
z0=round(qnorm(prob),2)
num=N*(z0*s)^2
den=N*e^2+(z0*s)^2
n=num/den
return(ceiling(n))
}

n_meat <- tamano_mu_exacto(dim(clean_meat_ton)[1],
sd(clean_meat_ton$Value),125, 0.95 )
n_meat

```

```
## [1] 2118
```

Para conseguir los estratos, nuestro equipo determinó que los métodos de muestreo que nos interesaban eran el de Varianza Mínima, el Proporcional y el uniforme. Para poder llegar a una decisión informada sobre cual escogeríamos, decidimos probar varios métodos.

Varianza Mínima

```

sig1=sd(meat_poultry_ton$Value)
sig2=sd(meat_beef_ton$Value)
sig3=sd(meat_pig_ton$Value)
sig4=sd(meat_sheep_ton$Value)
sigj=c(sig1,sig2,sig3,sig4)

N1=dim(meat_poultry_ton)[1]
N2=dim(meat_beef_ton)[1]
N3=dim(meat_pig_ton)[1]
N4=dim(meat_sheep_ton)[1]
Nj=c(N1,N2,N3,N4)

den1=sum(Nj*sigj)

nj_vm=round(n_meat*((Nj*sigj)/den1))
list("Sample Poultry" = nj_vm[1], "Sample Beef" = nj_vm[2],
"Sample Pig" = nj_vm[3], "Sample Sheep" = nj_vm[4])

## $'Sample Poultry'
## [1] 527
##
## $'Sample Beef'
## [1] 367
##
## $'Sample Pig'
## [1] 1137
##
## $'Sample Sheep'
## [1] 88

```

Método Proporcional

```
proportional_assignment <- function(n,Nj,N){
  fraction = Nj/N
  answer = n * fraction
  return(answer)
}

n_poultry <- proportional_assignment(n_meat, dim(meat_poultry_ton)[1],
                                   dim(clean_meat_ton)[1])

n_beef <- proportional_assignment(n_meat,dim(meat_beef_ton)[1],
                                 dim(clean_meat_ton)[1])

n_pig <- proportional_assignment(n_meat,dim(meat_pig_ton)[1],
                                dim(clean_meat_ton)[1])

n_sheep <- proportional_assignment(n_meat,dim(meat_sheep_ton)[1],
                                  dim(clean_meat_ton)[1])

list("Sample Poultry" = n_poultry, "Sample Beef" = n_beef,
     "Sample Pig" = n_pig, "Sample Sheep" = n_sheep)

## $'Sample Poultry'
## [1] 529.8194
##
## $'Sample Beef'
## [1] 529.3935
##
## $'Sample Pig'
## [1] 528.9676
##
## $'Sample Sheep'
## [1] 529.8194
```

Método Uniforme (ESTA ES LA QUE SE ESCOGIÓ)

```
uniform_sample <- function(k,n){
  answer <- n/k
  return(answer)
}

n_poultry <- uniform_sample(4,n_meat)
n_beef <- uniform_sample(4,n_meat)
n_pig <- uniform_sample(4,n_meat)
n_sheep <- uniform_sample(4,n_meat)

list("Sample Poultry" = n_poultry, "Sample Beef" = n_beef,
     "Sample Pig" = n_pig, "Sample Sheep" = n_sheep)

## $'Sample Poultry'
```

```
## [1] 529.5
##
## $'Sample Beef'
## [1] 529.5
##
## $'Sample Pig'
## [1] 529.5
##
## $'Sample Sheep'
## [1] 529.5
```

Posiciones

Considerando que deseamos hacer un análisis de correlaciones, la muestra a base de Varianza Mínima hacía poco sentido. Esto es debido a la diferencia radical entre cada estrato en términos de la cantidad de datos, lo cual llevaría a ciertas carnes teniendo mas representacion a través del tiempo que otras. El método proporcional, aunque resulto con estratos bastante similares, se tuvo que rechazar por razones similares.

El método uniforme por su naturaleza produce muestras de tamaño igual. Lo cual lo hace el método ideal para los propósitos de esta investigación.

Utilizando metodos habilitados por el paquete tidyverse, podemos asegurarnos que la muestra tenga una cantidad relativamente equitativa para cada pais, tiempo, y sujeto. Esto veremos en el siguiente “chunk”

```
set.seed(1234)

muestra_keys <- distinct(clean_meat_ton[c(1,4)]) %>%
  slice_sample(n = 530, replace = FALSE)

muestra_beef <- clean_meat_ton[, -3] %>%
  filter(SUBJECT == "BEEF") %>%
  inner_join(muestra_keys, by = c("LOCATION", "TIME"))

muestra_poultry <- clean_meat_ton[, -3] %>%
  filter(SUBJECT == "POULTRY") %>%
  inner_join(muestra_keys, by = c("LOCATION", "TIME"))

muestra_sheep <- clean_meat_ton[, -3] %>%
  filter(SUBJECT == "SHEEP") %>%
  inner_join(muestra_keys, by = c("LOCATION", "TIME"))

muestra_pig <- clean_meat_ton[, -3] %>%
  filter(SUBJECT == "PIG") %>%
  inner_join(muestra_keys, by = c("LOCATION", "TIME"))

muestra_meat <- rbind(muestra_beef, muestra_pig,
  muestra_poultry, muestra_sheep) %>%
  distinct()

head(muestra_meat)
```

```
##   LOCATION SUBJECT TIME   Value
```

```
## 1      ARG      BEEF 1990 2588.92
## 2      ARG      BEEF 1993 2507.87
## 3      ARG      BEEF 1999 2383.91
## 4      ARG      BEEF 2000 2385.11
## 5      ARG      BEEF 2002 2180.40
## 6      ARG      BEEF 2008 2706.54
```

Ahora tendremos que verificar que los datos esten distribuidos de manera equitativa. De esto no ser el caso, se tendra que decidir si vamos a manejar los datos perdidos mediante un proceso de borrarlos, y si los reemplazaremos con imputacion.

```
Time <-muestra_meat %>%
  count(TIME, sort = TRUE) %>%
  mutate(percent = 100 * n / sum(n))

Meat <-muestra_meat %>%
  count(SUBJECT, sort = TRUE) %>%
  mutate(percent = 100 * n / sum(n))

Country <-muestra_meat %>%
  count(LOCATION, sort = TRUE) %>%
  mutate(percent = 100 * n / sum(n))

list(Time, Meat, Country)
```

```
## [[1]]
##      TIME  n  percent
## 1  2004  96  4.539007
## 2  2016  96  4.539007
## 3  2018  96  4.539007
## 4  1997  92  4.349882
## 5  1993  84  3.971631
## 6  2007  84  3.971631
## 7  1990  80  3.782506
## 8  1996  80  3.782506
## 9  2000  80  3.782506
## 10 2002  80  3.782506
## 11 2008  80  3.782506
## 12 1994  76  3.593381
## 13 2017  76  3.593381
## 14 2010  72  3.404255
## 15 2014  72  3.404255
## 16 2015  72  3.404255
## 17 1991  68  3.215130
## 18 1998  68  3.215130
## 19 2001  68  3.215130
## 20 2005  68  3.215130
## 21 2012  68  3.215130
## 22 2013  68  3.215130
## 23 1992  67  3.167849
## 24 1999  64  3.026005
## 25 2003  60  2.836879
## 26 2009  56  2.647754
```

```

## 27 2006 52 2.458629
## 28 1995 48 2.269504
## 29 2011 44 2.080378
##
## [[2]]
##   SUBJECT    n  percent
## 1     BEEF 529 25.01182
## 2  POULTRY 529 25.01182
## 3     SHEEP 529 25.01182
## 4       PIG 528 24.96454
##
## [[3]]
##   LOCATION    n  percent
## 1         CHE  68 3.215130
## 2         RUS  63 2.978723
## 3         KAZ  60 2.836879
## 4         MEX  60 2.836879
## 5         MOZ  60 2.836879
## 6         PAK  60 2.836879
## 7         PRY  60 2.836879
## 8         SAU  60 2.836879
## 9         SDN  60 2.836879
## 10        VNM  60 2.836879
## 11        BGD  56 2.647754
## 12        CAN  56 2.647754
## 13        CHL  56 2.647754
## 14        JPN  56 2.647754
## 15        KOR  56 2.647754
## 16        PHL  56 2.647754
## 17        CHN  55 2.600473
## 18        AUS  54 2.553191
## 19        DZA  52 2.458629
## 20        IRN  52 2.458629
## 21        PER  52 2.458629
## 22        GHA  48 2.269504
## 23        IND  48 2.269504
## 24        MYS  48 2.269504
## 25        NZL  48 2.269504
## 26        URY  48 2.269504
## 27        ZMB  48 2.269504
## 28        EGY  44 2.080378
## 29        ETH  44 2.080378
## 30        HTI  44 2.080378
## 31        THA  44 2.080378
## 32        ZAF  44 2.080378
## 33        COL  40 1.891253
## 34        IDN  40 1.891253
## 35        NGA  40 1.891253
## 36        NOR  40 1.891253
## 37        TZA  40 1.891253
## 38        BRA  36 1.702128
## 39        TUR  36 1.702128
## 40        USA  36 1.702128
## 41        ISR  32 1.513002

```

```
## 42      UKR 28 1.323877
## 43      ARG 27 1.276596
```

Como podemos ver, los porcentajes de distribucion no son perfectos para los casos que traimo. Aunque son relativamente cerca por su mayor parte, esto nos quiere decir que para ciertas fechas, paises y carnes no hay datos disponibles en nuestra muestra. Para arreglar este asunto, primero veremos que se pueda hacer mediante la imputacion.

Para esto estaremos usando la libreria MICE, cual contiene codigo especializado en la imputacion de datos. Algo que sin duda sera de mucho uso en esta situacion. Tambien, tendremos que unir los datos en un formato en se tenga una columna para cada valor de cada carne particular.

```
library(mice)

## Warning: package 'mice' was built under R version 4.1.3

##
## Attaching package: 'mice'

## The following object is masked from 'package:stats':
##
##      filter

## The following objects are masked from 'package:base':
##
##      cbind, rbind

muestra_meat_joined <- muestra_beef %>%
full_join(muestra_poultry, by = c("LOCATION", "TIME"), suffix = c("_beef", "_poultry")) %>%
full_join(muestra_sheep, by = c("LOCATION", "TIME"), suffix = c("_beef", "_sheep")) %>%
full_join(muestra_pig, by = c("LOCATION", "TIME"), suffix = c("_beef", "_pig")) %>%
  select(-c("SUBJECT_beef", "SUBJECT_poultry", "SUBJECT_beef_beef", "SUBJECT_pig" )) %>%
  mutate(Beef = Value_beef, Poultry = Value_poultry, Sheep = Value_beef_beef,
         Pig = Value_pig, .keep = "unused") %>%
  distinct()

list("NA" = sum(is.na(muestra_meat_joined)), "Observations" = dim(muestra_meat_joined)[1])

## $'NA'
## [1] 5
##
## $Observations
## [1] 530
```

Antes de proceder con el analisis como tal, vemos que tenemos **cinco** valores que no estan disponibles. Considerando que los valores que se deben imputar no son tantos, tenemos bastante flexibilidad a la hora de escoger nuestro metodo. En teoria, pudiessmos ir con el metodo de imputacion de media o mediana.

Sin embargo, considerando la naturaleza de los datos—en donde hay ciertos paises que tienen un consumo **drasticamente** mas alto que otros, dicha medida pudiese resultar en datos no ideales. Por esto, al momento favorecemos los metodos de regresion.

```
columns <- colnames(muestra_meat_joined)

reg <- mice(muestra_meat_joined[,names(muestra_meat_joined)
  %in% columns],m = 1,
  maxit = 1,method="norm.predict",print=F)
```

```
## Warning: Number of logged events: 1
```

```
meat_joined_imputed <- complete(reg)
all.equal(muestra_meat_joined, meat_joined_imputed)
```

```
## [1] "Component \"Beef\": 'is.NA' value mismatch: 0 in current 1 in target"
## [2] "Component \"Poultry\": 'is.NA' value mismatch: 0 in current 1 in target"
## [3] "Component \"Sheep\": 'is.NA' value mismatch: 0 in current 1 in target"
## [4] "Component \"Pig\": 'is.NA' value mismatch: 0 in current 2 in target"
```

Mirando los resultados de la imputacion, vemos que la unica diferencia entre la muestra original y la imputada en terminos de dato es que la imputada no tiene los valores NA. Lo cual es algo bueno, ya que eso fue lo que estabamos buscando.

Considerando la cantidad pequena de los datos imputados, estos no se graficaran para los propositos de brevedad. Ahora que tenemos nuestros datos imputados, podemos proceder con nuestro análisis de la relacion que tiene el consumo de nuestras cuatro carnes a las variables socio-economicas que hemos recogido

Emisiones co2 Annual

Primero, conseguiremos nuestros datos sobre emisiones de co2 anual por país. Luego de esto, realizaremos una unión entre nuestras dos tablas. Es muy probable que, no todas las observaciones encajen bien con las tablas y que no matcheen completamente. En dichos casos, realizaremos un proceso de imputacion de la misma manera que hicimos anteriormente.

De la misma manera que las uniones que realizamos anteriormente, aqui se unira a base de lugar y periodo.

```
dat_emissions <-
read_csv("C:\\Users\\nesto\\Documents\\School\\ESTA 5502\\Project Meat Temp\\annual-co2-emissions-per-c
```

```
## Rows: 23949 Columns: 4
```

```
## -- Column specification -----
## Delimiter: ","
## chr (2): Entity, Code
## dbl (2): Year, Annual CO2 emissions
```

```
##
## i Use 'spec()' to retrieve the full column specification for this data.
## i Specify the column types or set 'show_col_types = FALSE' to quiet this message.
```

```
head(dat_emissions)
```

```
## # A tibble: 6 x 4
##   Entity      Code   Year 'Annual CO2 emissions'
##   <chr>      <chr> <dbl>          <dbl>
## 1 Afghanistan AFG    1949          14656
## 2 Afghanistan AFG    1950          84272
## 3 Afghanistan AFG    1951          91600
## 4 Afghanistan AFG    1952          91600
## 5 Afghanistan AFG    1953         106256
## 6 Afghanistan AFG    1954         106256
```

```
muestra_meat_emissions <- meat_joined_imputed %>%
  left_join(dat_emissions[, -1], by = c("LOCATION" = "Code", "TIME" = "Year")) %>%
  mutate(Emissions = (`Annual CO2 emissions`^(1/4)), .keep = "unused")

sum(is.na(muestra_meat_emissions))
```

```
## [1] 0
```

Wow! Sorprendentemente, resulto que todas las observaciones en esta categoria de analisis cubrieron todos los datos en su totalidad. Perfecto. Con eso hecho, hagamos las graficas de correlacion.

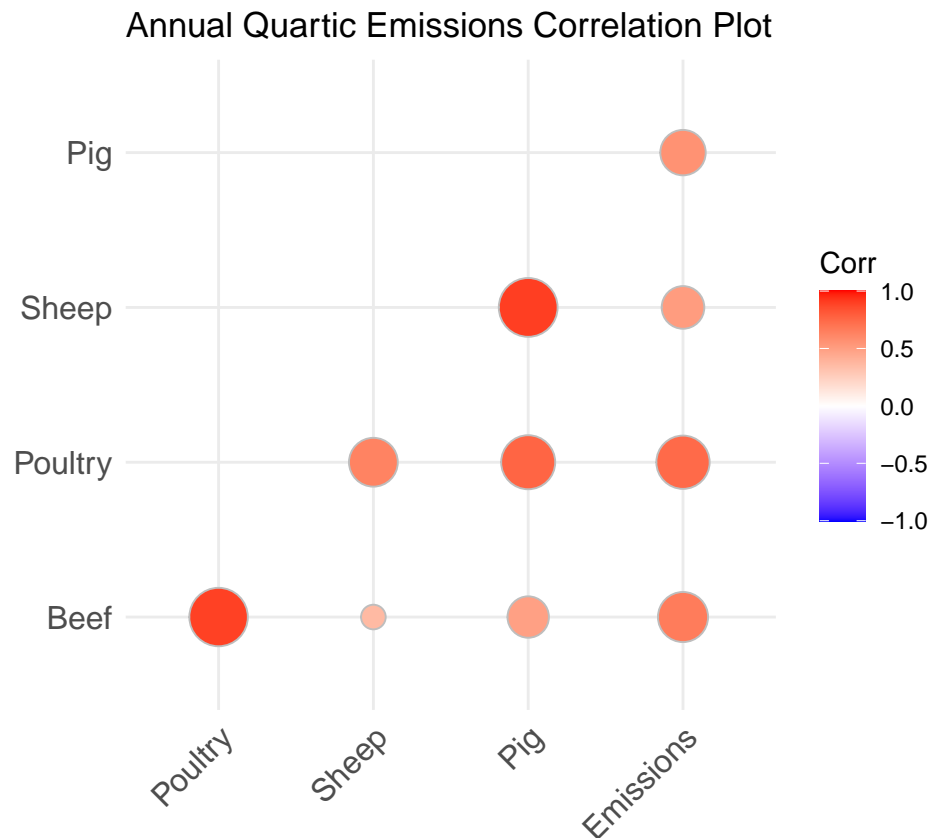
```
library(ggcorrplot)
```

```
## Warning: package 'ggcorrplot' was built under R version 4.1.3
```

```
pmats <- cor_pmat(muestra_meat_emissions[-c(1:2)])

ggcorrplot(cor(muestra_meat_emissions[-c(1:2)]), method = "circle", hc.order = FALSE,
  type = "lower", p.mat = pmats) + labs(title = "Annual Quartic Emissions Correlation Plot" )
```

```
## Warning: 'guides(<scale> = FALSE)' is deprecated. Please use 'guides(<scale> =
## "none")' instead.
```

```
cor(muestra_meat_emissions[-c(1:2)])
```

```
##           Beef  Poultry   Sheep    Pig Emissions
## Beef      1.0000000 0.8917957 0.3553088 0.4864581 0.6625453
## Poultry   0.8917957 1.0000000 0.6269883 0.7572536 0.7369883
## Sheep     0.3553088 0.6269883 1.0000000 0.9010693 0.5110594
## Pig       0.4864581 0.7572536 0.9010693 1.0000000 0.5558588
## Emissions 0.6625453 0.7369883 0.5110594 0.5558588 1.0000000
```

Con las emisiones anuales, vemos que **todas** las carnes cuentan con un nivel de significancia en terminos de su correlacion con las emisiones. En este caso, **Pollo** viene siendo el que tiene la correlacion mas alta

Emisiones co2 per Capita

Aqui tenemos otros datos de co2, pero al nivel per capita

```
dat_emissions_cap <-
read_csv("C:\\Users\\nesto\\Documents\\School\\ESTA 5502\\Project Meat Temp\\co-emissions-per-capita.csv")

## Rows: 23307 Columns: 4

## -- Column specification -----
## Delimiter: ","
## chr (2): Entity, Code
## dbl (2): Year, Annual CO2 emissions (per capita)
```

```
##
## i Use 'spec()' to retrieve the full column specification for this data.
## i Specify the column types or set 'show_col_types = FALSE' to quiet this message.
```

```
head(dat_emissions_cap)
```

```
## # A tibble: 6 x 4
##   Entity      Code  Year 'Annual CO2 emissions (per capita)'
##   <chr>      <chr> <dbl>                                <dbl>
## 1 Afghanistan AFG    1949                                0.0019
## 2 Afghanistan AFG    1950                                0.0109
## 3 Afghanistan AFG    1951                                0.0117
## 4 Afghanistan AFG    1952                                0.0115
## 5 Afghanistan AFG    1953                                0.0132
## 6 Afghanistan AFG    1954                                0.013
```

```
muestra_meat_emissions_cap <- meat_joined_imputed %>%
  left_join(dat_emissions_cap[, -1], by = c("LOCATION" = "Code", "TIME" = "Year")) %>%
  mutate(emissions_cap = `Annual CO2 emissions (per capita)`, .keep = "unused")

sum(is.na(muestra_meat_emissions))
```

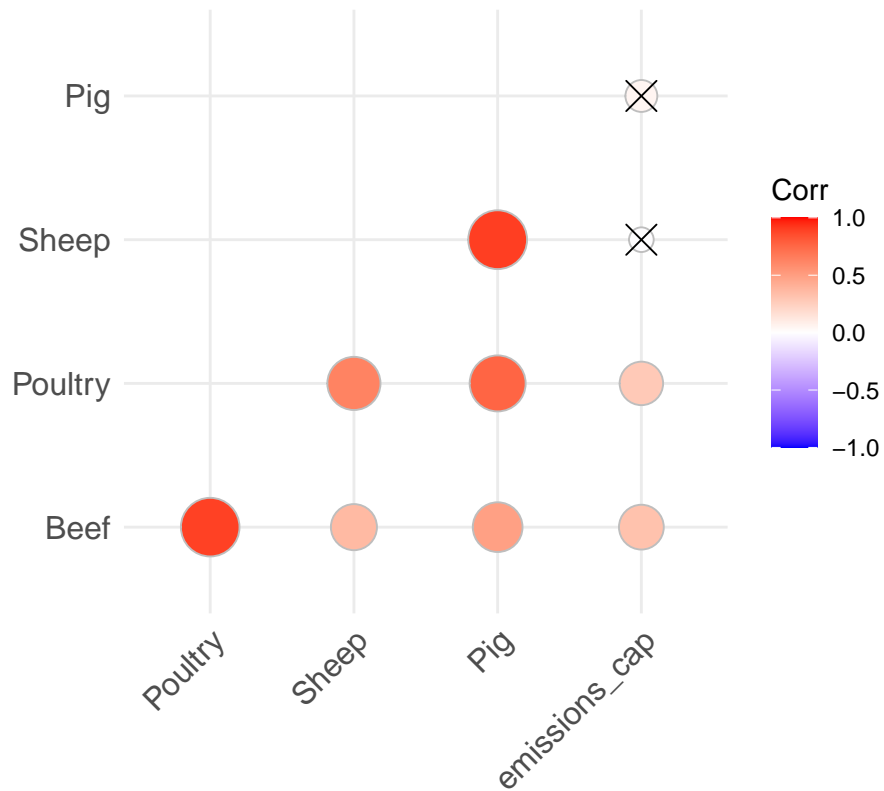
```
## [1] 0
```

```
pmats <- cor_pmat(muestra_meat_emissions_cap[-c(1:2)])

ggcorrplot(cor(muestra_meat_emissions_cap[-c(1:2)]), method = "circle", hc.order = FALSE,
  type = "lower", p.mat = pmats) + labs(title = "Quartic Emissions per Capita Correlation Plot")
```

```
## Warning: 'guides(<scale> = FALSE)' is deprecated. Please use 'guides(<scale> =
## "none")' instead.
```

Quartic Emissions per Capita Correlation Plot



```
cor(muestra_meat_emissions_cap[-c(1:2)])
```

```
##           Beef  Poultry      Sheep      Pig emissions_cap
## Beef      1.000000 0.8917957 0.35530878 0.48645812 0.31780657
## Poultry   0.8917957 1.0000000 0.62698831 0.75725363 0.27613210
## Sheep     0.3553088 0.6269883 1.00000000 0.90106933 -0.01332035
## Pig       0.4864581 0.7572536 0.90106933 1.00000000 0.05444262
## emissions_cap 0.3178066 0.2761321 -0.01332035 0.05444262 1.00000000
```

Aqui vemos que, al nivel per capita, cerdo y oveja no son significativos en terminos de su correlacion con emisiones. Esto se debe contrastar con, por ejemplo, lo que viene siendo las emisiones anuales—en donde todas las carnes eran significativas. Sin embargo, vale resaltar que **Beef** y **Poultry** siguen teniendo las correlaciones mas altas

GDP

Crecimiento economico es la proxima variable de interes con cual queremos realizar nuestro analisis.

```
raw_gdp <-
read_csv("C:\\Users\\nesto\\Documents\\School\\ESTA 5502\\Project Meat Temp\\gdp_spending.csv")
```

```
## Rows: 4447 Columns: 17
```

```
## -- Column specification -----
## Delimiter: ","
## chr (13): LOCATION, Country, VARIABLE, Variable, FREQUENCY, Frequency, TIME,...
## dbl (2): PowerCode Code, Value
## lgl (2): Flag Codes, Flags

##
## i Use 'spec()' to retrieve the full column specification for this data.
## i Specify the column types or set 'show_col_types = FALSE' to quiet this message.
```

```
head(raw_gdp)
```

```
## # A tibble: 6 x 17
##   LOCATION Country VARIABLE Variable FREQUENCY Frequency TIME Time 'Unit Code'
##   <chr>      <chr>   <chr>    <chr>    <chr>      <chr>    <chr> <chr> <chr>
## 1 AUS      Austr~ GDPV_AN~ Gross d~ A        Annual    2000  2000  PC
## 2 AUS      Austr~ GDPV_AN~ Gross d~ A        Annual    2001  2001  PC
## 3 AUS      Austr~ GDPV_AN~ Gross d~ A        Annual    2002  2002  PC
## 4 AUS      Austr~ GDPV_AN~ Gross d~ A        Annual    2003  2003  PC
## 5 AUS      Austr~ GDPV_AN~ Gross d~ A        Annual    2004  2004  PC
## 6 AUS      Austr~ GDPV_AN~ Gross d~ A        Annual    2005  2005  PC
## # ... with 8 more variables: Unit <chr>, PowerCode Code <dbl>, PowerCode <chr>,
## #   Reference Period Code <chr>, Reference Period <chr>, Value <dbl>,
## #   Flag Codes <lgl>, Flags <lgl>
```

```
dat_gdp <- raw_gdp[-c(2,3,4,5, 7,9,10,11,12,13,14,16,17)] %>%
  mutate(LOCATION = LOCATION, TIME = as.integer(Time), .keep = "unused") %>%
  filter(Frequency == "Annual")      # Aquí limitamos el analisis a lo anual
```

```
## Warning in mask$eval_all_mutate(quo): NAs introduced by coercion
```

```
# Join
muestra_meat_gdp <- meat_joined_imputed %>%
  left_join(dat_gdp[-2], by = c("LOCATION" = "LOCATION", "TIME" = "TIME")) %>%
  mutate(gdp = Value, .keep = "unused")

head(muestra_meat_gdp)
```

```
##   LOCATION TIME   Beef Poultry   Sheep   Pig      gdp
## 1      ARG 1990 2588.92  337.82 28.48706 153.63      NA
## 2      ARG 1993 2507.87  714.00 72.12000 211.26      NA
## 3      ARG 1999 2383.91  959.58 55.36000 287.45      NA
## 4      ARG 2000 2385.11  932.89 59.17000 283.10 -0.7889986
## 5      ARG 2002 2180.40  697.74 58.38000 214.72 -10.8944846
## 6      ARG 2008 2706.54 1338.96 50.15000 304.85  4.0572331
```

Ahora, vamos a verificar cuantos valores NA existen dentro de nuestra nueva tabla.

```
sum(is.na(muestra_meat_gdp$gdp))
```

```
## [1] 379
```

Vemos aqui que contamos con **379** valores perdido. En este caso, eso vendria siendo la mayoria de los datos de esta base de datos en particular. Tomando esto en mente, deseamos resaltar que los resultados que se encuentren en el caso de la variable **gdp** seran puramente hipoteticos debido a esta situacion.

Pero, saliendo de eso, considerando la cantidad grande de valores perdidos, entedemos que el modelo adecuado para encontrar a estos valores es uno de regresion. Aqui en el caso de **gdp** contamos con una ventaja: **Los unicos NA vienen de la columna gdp**. Esto quiere decir que a la hora de construir una regresion, tendremos el resto de las variables en los datos disponible como predictores.

```
sum(is.na(muestra_meat_gdp))
```

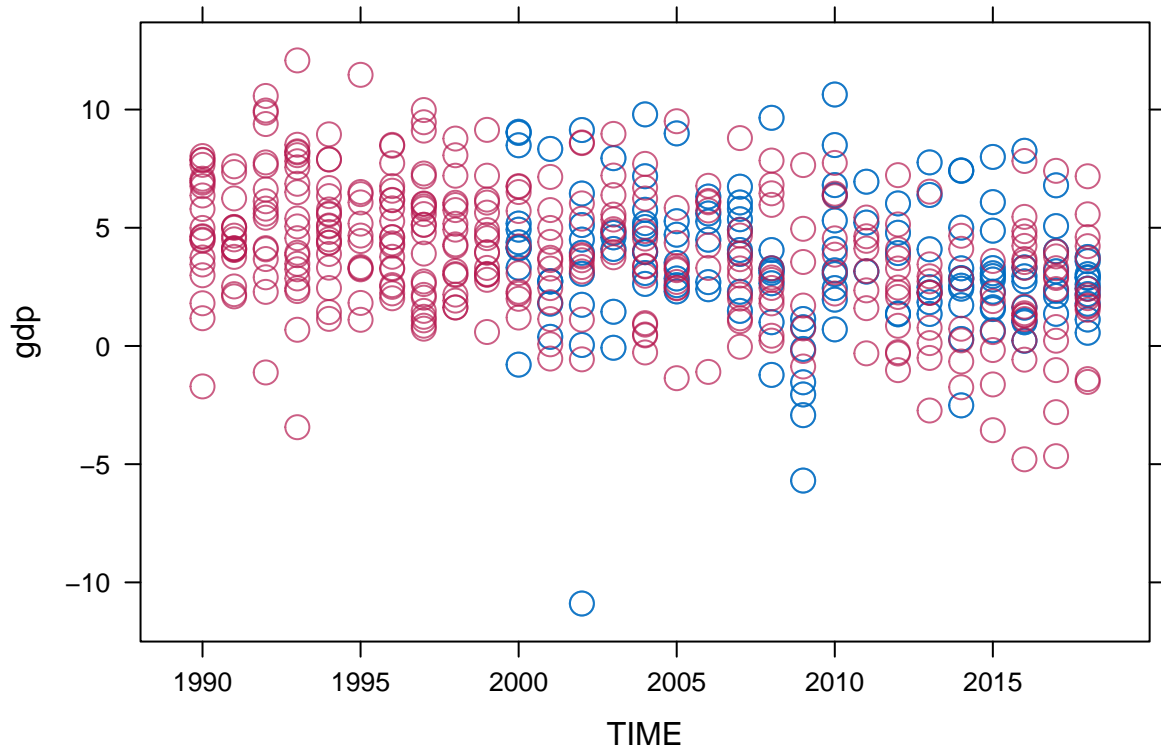
```
## [1] 379
```

```
columns_gdp=colnames(muestra_meat_gdp)
rege <- mice(muestra_meat_gdp[,names(muestra_meat_gdp) %in% columns_gdp],m = 1,
            maxit = 1,method="norm.nob",print=F)
```

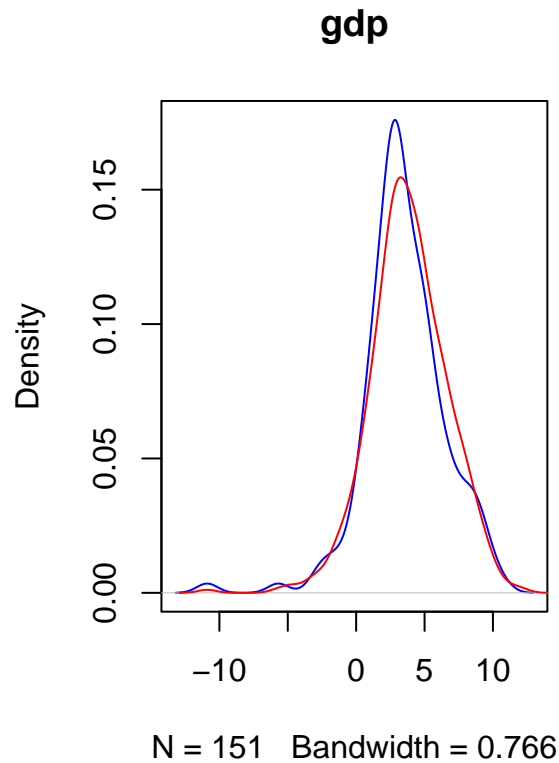
```
## Warning: Number of logged events: 1
```

```
imputed_meat_gdp <- mice::complete(rege)
```

```
xyplot(rege,gdp~TIME,cex=1.5)
```



```
par(mfrow=c(1,2))
plot(density(muestra_meat_gdp$gdp,na.rm = T),col="blue",main="gdp")
lines(density(imputed_meat_gdp$gdp),col="red")
```

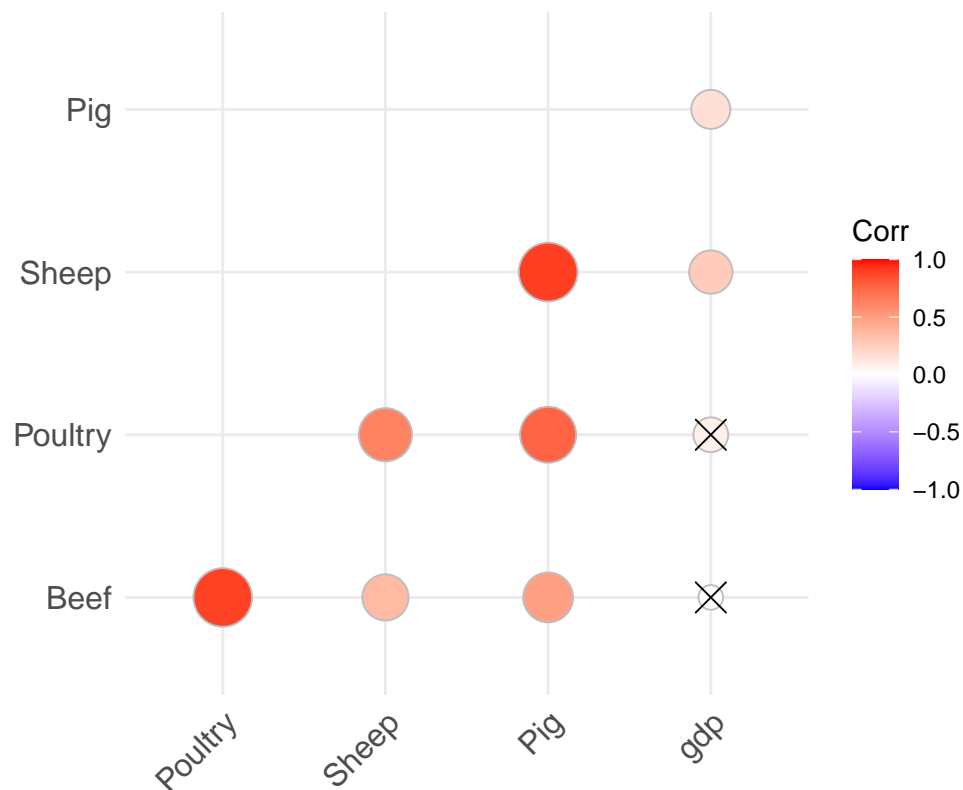


```
pmats <- cor_pmat(imputed_meat_gdp[-c(1:2)])

ggcorrplot(cor(imputed_meat_gdp[-c(1:2)]),method ="circle", hc.order = FALSE,
type = "lower",p.mat = pmats) + labs(title = "Annual Gross Domestic Product Growth Correlation Plot")
```

```
## Warning: 'guides(<scale> = FALSE)' is deprecated. Please use 'guides(<scale> =
## "none")' instead.
```

Annual Gross Domestic Product Growth Correlation Plot



```
cor(imputed_meat_gdp[-c(1:2)])
```

```
##           Beef  Poultry  Sheep  Pig  gdp
## Beef      1.00000000 0.89179569 0.3553088 0.4864581 -0.001046906
## Poultry    0.891795688 1.00000000 0.6269883 0.7572536 0.083459393
## Sheep      0.355308776 0.62698831 1.0000000 0.9010693 0.274433493
## Pig        0.486458125 0.75725363 0.9010693 1.0000000 0.157362179
## gdp       -0.001046906 0.08345939 0.2744335 0.1573622 1.000000000
```

Viendo aqui, los resultados demuestran que las carnes con mas altas correlaciones con el gdp son **Pig** y **Sheep**. Curiosamente, nosotros esperabamos que Poultry y Beef sean las que tenian la mayor significancia. Sin embargo, como se puede ver en la grafica, esto no es el caso.

Tanto poultry como beef son insignificantas a base de los pvalues que tenemos en la grafica.

No queremos llegar a demasiadas conclusiones sobre esto resultados, aunque nos alla dado eso. Otra vez resaltamos que la gran mayoría de los datos en gdp fueron **imputados** para el proposito de analisis. Esto hace que los resultados que recibimos en esta seccion se queden en un nivel **hipotetico**. Se tendria que mirar mas a fondos, con mas datos disponibles, si los resultados que sacamos aqui se pueden replicar en una base de datos completa.

Extreme Poverty

Ahora vamos a la pobreza extrema. Se ha dicho tradicionalmente que la carne es una senal de que una poblacion esta creciendo en termino socioeconomicos. Aqui podremos por lo menos ver (o empezar una discussion) sobre si es cierto o no

```
poverty <-
read_csv("C:\\Users\\nesto\\Documents\\School\\ESTA 5502\\Project Meat Temp\\share-of-population-in-ext.
```

```
## Rows: 6953 Columns: 4
```

```
## -- Column specification -----
## Delimiter: ","
## chr (2): Entity, Code
## dbl (2): Year, 1.90_per_day_share_of_population_below_poverty_line
```

```
##
## i Use 'spec()' to retrieve the full column specification for this data.
## i Specify the column types or set 'show_col_types = FALSE' to quiet this message.
```

```
head(poverty)
```

```
## # A tibble: 6 x 4
##   Entity Code   Year '1.90_per_day_share_of_population_below_poverty_line'
##   <chr>   <chr> <dbl>                                <dbl>
## 1 Albania ALB   1981                                0.465
## 2 Albania ALB   1982                                0.408
## 3 Albania ALB   1983                                0.465
## 4 Albania ALB   1984                                0.535
## 5 Albania ALB   1985                                0.535
## 6 Albania ALB   1986                                0.465
```

```
dat_poverty <- poverty[,-1] %>%
mutate(Extreme_Poverty_PCT =
round(`1.90_per_day_share_of_population_below_poverty_line`, 2), .keep = "unused" )
```

```
# Join
```

```
muestra_meat_poverty <- meat_joined_imputed %>%
  left_join(dat_poverty, by = c("LOCATION" = "Code", "TIME" = "Year"))
```

```
head(muestra_meat_poverty)
```

```
##   LOCATION TIME   Beef Poultry   Sheep   Pig Extreme_Poverty_PCT
## 1      ARG 1990 2588.92  337.82 28.48706 153.63          NA
## 2      ARG 1993 2507.87  714.00 72.12000 211.26          NA
## 3      ARG 1999 2383.91  959.58 55.36000 287.45          NA
## 4      ARG 2000 2385.11  932.89 59.17000 283.10          NA
## 5      ARG 2002 2180.40  697.74 58.38000 214.72          NA
## 6      ARG 2008 2706.54 1338.96 50.15000 304.85          NA
```

```
sum(is.na(muestra_meat_poverty$Extreme_Poverty_PCT))
```

```
## [1] 34
```


En este caso contamos con solamente **34** de los datos siendo NA. Esto es buena, ya que no es la mayoría de los datos. Sin embargo, esto no excluye a estos datos de la imputación. Siguiendo otra vez con el procedimiento que hemos estado utilizando hasta ahora, haremos una imputación mediante métodos regresivos hacia los datos.

```
sum(is.na(muestra_meat_poverty))
```

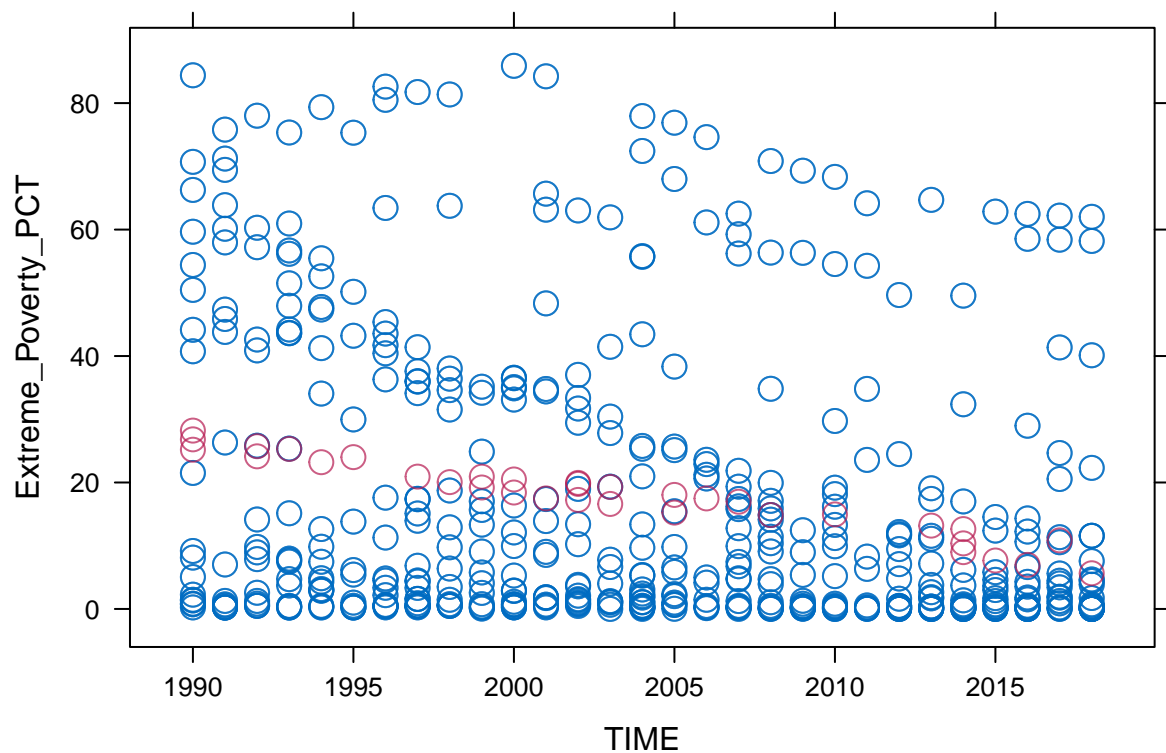
```
## [1] 34
```

```
columns_poverty=colnames(muestra_meat_poverty)
reg <- mice(muestra_meat_poverty[,names(muestra_meat_poverty) %in% columns_poverty],m = 1,
           maxit = 1,method="norm.predict",print=F)
```

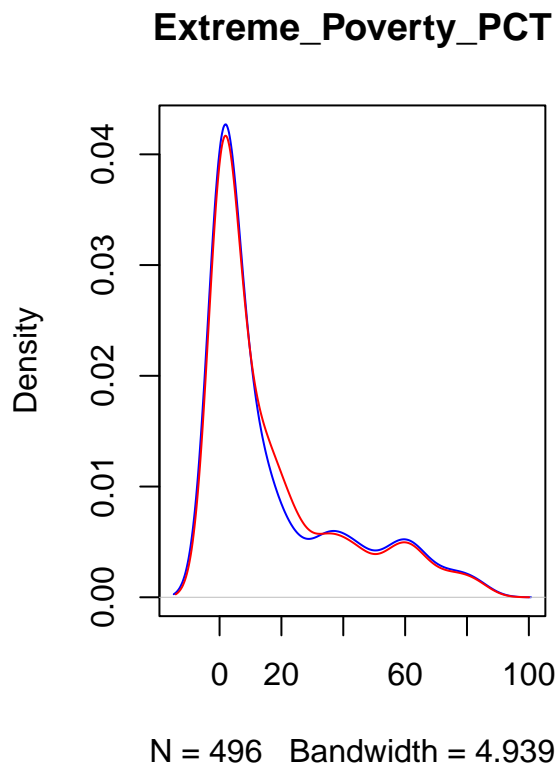
```
## Warning: Number of logged events: 1
```

```
imputed_meat_poverty <- mice::complete(reg)
```

```
xyplot(reg,Extreme_Poverty_PCT~TIME,cex=1.5)
```

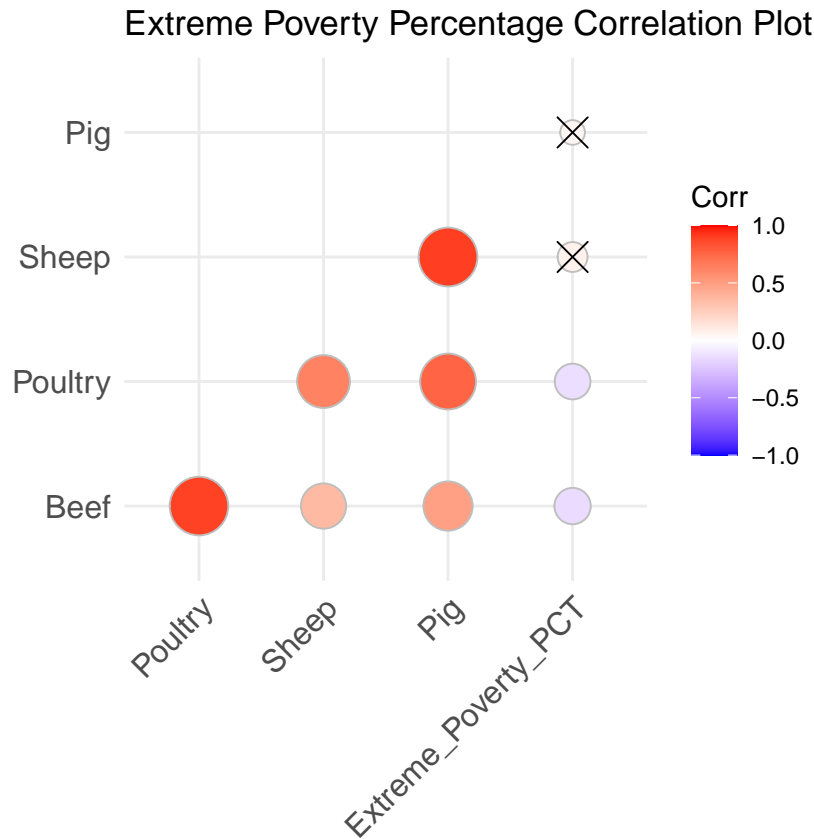


```
par(mfrow=c(1,2))
plot(density(muestra_meat_poverty$Extreme_Poverty_PCT,na.rm = T),col="blue",main="Extreme_Poverty_PCT")
lines(density(imputed_meat_poverty$Extreme_Poverty_PCT),col="red")
```



Vemos aqui que los datos imputados mas o menos siguen la tendencia de los datos originales. Esto es tanto en el xyplot como en la grafica de densidad que compara las imputaciones con los datos originales.

```
pmats <- cor_pmat(imputed_meat_poverty[-c(1:2)])  
  
ggcorrplot(cor(imputed_meat_poverty[-c(1:2)]),method = "circle", hc.order = FALSE,  
            type = "lower",p.mat = pmats) + labs(title = "Extreme Poverty Percentage Correlation Plot")  
  
## Warning: 'guides(<scale> = FALSE)' is deprecated. Please use 'guides(<scale> =  
## "none")' instead.
```



```
cor(imputed_meat_poverty[-c(1:2)])
```

```
##           Beef      Poultry      Sheep      Pig
## Beef      1.0000000  0.8917957  0.35530878  0.48645812
## Poultry   0.8917957  1.0000000  0.62698831  0.75725363
## Sheep     0.3553088  0.6269883  1.00000000  0.90106933
## Pig       0.4864581  0.7572536  0.90106933  1.00000000
## Extreme_Poverty_PCT -0.1531845 -0.1386908  0.06913077  0.04609289
##           Extreme_Poverty_PCT
## Beef      -0.15318446
## Poultry   -0.13869076
## Sheep      0.06913077
## Pig        0.04609289
## Extreme_Poverty_PCT  1.00000000
```

Ahora vemos que, entre todas las carnes, las únicas con una correlación con el porcentaje de pobreza extrema son **Poultry** y **Beef**. Pig y Sheep parecen ser carnes que no tienen una correlación con la pobreza extrema.

Si se debe hacer una salvedad. Esta viene siendo que las correlaciones de Poultry y Beef no son extremadamente altas. Sin embargo, nosotros pensamos que si vale la pena resaltar a esta. Ya que, en la vida real, no se suelen encontrar correlaciones extremadamente alta. También se debe considerar la complejidad de lo que es la pobreza extrema. Sin duda hay un número inmenso de variables que contribuyen a este fenómeno desafortunado, pero debido a limitaciones de tiempo no se puede entrar a estas muy a fondo.

Conclusion

En conclusion, algunas de nuestras hipotesis probaron cierto para el estudio. Las correlaciones de los tipos de carne con las emisiones de CO2 probaron ser bastante altos y positivos. En el caso de las emisiones per capita, esto solo fue parcialmente cierto en donde solamente dos de las cuatro carnes probaron tener una correlacion decente.

En el caso del gdp, es inconcluso. Esto se debe a que los procedimientos de joins y muestreo en nuestra situacion llevo a que la mayoria de los datos sean imputables. Es por esto que las correlaciones que se obtuvieron en este, aunque interesante, solo pueden quedarse como puramente hipotesis. En el caso de la pobreza extrema, si se encontro que algunas de las carnes tenian correlaciones negativas con la pobreza extrema. Lo que parcialmente valida conocimiento previo sobre la relacion negativa con el consumo de carne y la pobreza extrema

Bibliografia

1. OECD-FAO Agricultural Outlook 2021-2030. (n.d.). Retrieved May 25, 2022, from <https://www.fao.org/3/cb5332en/Meat.pdf>
2. Impact of income and social deprivation on meat ... - scielo.org.mx. (n.d.). Retrieved May 25, 2022, from <http://www.scielo.org.mx/pdf/remexca/v9n6/2007-0934-remexca-9-06-1245-en.pdf>
3. Meat consumption as an indicator of economic well-being — case study of ... (n.d.). Retrieved May 25, 2022, from https://www.researchgate.net/publication/328901259_MEAT_CONSUMPTION_AS_AN_INDICATOR_OF_ECONOMIC_WELL-BEING_-_CASE_STUDY_OF_A_DEVELOPED_AND_DEVELOPING_ECONOMY
4. Whitton, C., Bogueva, D., Marinova, D., & Phillips, C. J. C. (2021, December 6). Are we approaching peak meat consumption? analysis of meat consumption from 2000 to 2019 in 35 countries and its relationship to gross domestic product. MDPI. Retrieved May 24, 2022, from <https://www.mdpi.com/2076-2615/11/12/3466>
5. The OECD creates better policies for better lives. read the OECD's main figures to find out more about their work. OECD. (n.d.). Retrieved May 24, 2022, from <https://www.oecd.org/about/>
6. Kenton, W. (2021, May 19). What is the International Poverty Line? Investopedia. Retrieved May 24, 2022, from <https://www.investopedia.com/terms/i/international-poverty-line.asp>
7. China's meat consumption: An income elasticity analysis and long-term ... (n.d.). Retrieved May 25, 2022, from https://www.researchgate.net/publication/254383740_China's_Meat_Consumption_An_Income_Elasticity_Analysis_and_Long-Term_Projections
8. Does meat consumption exacerbate greenhouse gas emissions? evidence ... (n.d.). Retrieved May 25, 2022, from https://www.researchgate.net/publication/344780826_Does_Meat_Consumption_Exacerbate_Greenhouse_Gas_Emissions_Evidence_from_US_Data