



República
del Ecuador



Cuentas Satélite de Educación 2007-2019

Informe de procesamiento de
indicadores territoriales 2019

Enero, 2021



Buenas cifras,
mejores vidas



Buenas cifras,
mejores vidas

Dirección responsable del contenido:

Dirección de Estadísticas Económicas

Elaborado por:

Henry Valdiviezo

Contacto:

inec@inec.gob.ec

www.ecuadorencifras.gob.ec

(02) 2234 164 · (02) 2235 890 · (02) 2526 072



Contenido

1.	Introducción	4
2.	Objetivos:	4
3.	Desarrollo.....	4
3.1.	Aspectos Conceptuales.....	4
3.2.	Insumos	5
3.3.	Código de programación	6
3.4.	Estructura base de datos VAB provincial procesada.....	16
3.5.	Resultados: mapas provinciales de la producción, consumo intermedio y VAB de la educación	17
4.	Bibliografía	20

Contenido de tablas

Gráfico 1: Distribución de la producción de educación y población a nivel provincial 2019.....	19
Gráfico 2: Distribución de la producción de las actividades enseñanza superior a nivel provincial 2019.....	19



1. Introducción

El Instituto Nacional de Estadísticas y Censos (INEC) dentro de su programación estadística plurianual y en el marco del Sistema Nacional de Planificación viene elaborando las Cuentas Satélite de Salud (CSE), como una herramienta de cuantificación económica de los flujos de oferta y demanda de servicios de educación en el país.

En esta línea, la Dirección de Estadísticas Económicas (DECON), a través de la Unidad de Gestión de Análisis de Síntesis (GASIN), planificó la ejecución del proyecto de actualización de las Cuentas Satélite de Educación (CSE) periodo 2018-2019, a partir del segundo trimestre del 2021 con la finalidad de publicar los resultados en diciembre de 2021.

Durante la etapa de planificación del proyecto, se establecieron objetivos de mejora para este producto estadístico, entre ellas, dos aspectos relevantes que se desarrollan en el presente documento. El 1ero, se refiere a la generación de información estadística con cobertura provincial para los principales agregados económicos de oferta de las CSE (producción, consumo intermedio y valor agregado bruto VAB); y el 2do, como parte del proceso transversal de innovación y automatización de la producción estadística del INEC, comprende el desarrollo de código de programación para automatizar la construcción de las bases de datos de indicadores territoriales.

Bajo estas premisas, el presente documento describe el proceso automatizado para la construcción de la base de datos del VAB a nivel provincial periodo 2018-2019, mediante el uso de sintaxis con el software de uso libre "R".

2. Objetivos:

El presente documento tiene como finalidad los siguientes puntos:

- Construir la base de datos del VAB con desagregación provincial mediante sintaxis en el programa R.
- Generar mapas provinciales sobre el comportamiento de la Producción, CI y VAB de la enseñanza.

3. Desarrollo

3.1. Aspectos Conceptuales

Conforme al marco de buenas prácticas estadísticas, el proceso inicia con la revisión y análisis de metodologías, recomendaciones y prácticas internacionales aplicadas al sistema educativo como la Clasificación Internacional Normalizada de la Educación, 2011, la Metodología de Cuentas Nacionales de la Educación, 2016, el "Manual on regional accounts methods" de Eurostat (2013) y otros documentos técnicos sobre medición regional del PIB, como el Sistema de Cuentas Nacionales de México y Ecuador.



A continuación, el proceso inicia con la obtención de indicadores monetarios de producción, gasto, ingreso, salario, etc. a nivel de cada una de las industrias de la educación y desagregado por provincia; los datos monetarios se complementan información de indicadores físicos como el número de docentes, número de alumnos o beneficiarios según niveles educativos y por cada provincia.

En función de la información disponible, para el cálculo de un agregado económico como la producción o valor agregado bruto a nivel territorial, se combinan métodos ascendentes y descendentes, los cuales son utilizados en función de la existencia de información directa para la medición del fenómeno económico a nivel territorial (método ascendente); mientras, en los casos con ausencia de información que permita la medición directa a nivel territorial, se utiliza información clave y relacionada como los indicadores físicos (docentes y alumnos) para la distribución de valores del fenómeno económico a nivel territorial.

Desde esta perspectiva, la información presupuestaria a nivel de micro datos para instituciones educativas y distritos educativos permitió construir las estadísticas territoriales desde un enfoque “ascendente”, mientras la información de indicadores físicos como los alumnos y docentes permitió construir la información estadística territorial desde un enfoque “descendente”.

Al final, los valores provinciales se reconcilian con el agregado a nivel nacional, para que lo provincial sea consistente con la información de las industrias (niveles educativos) a escala nacional.

A continuación, se describe el proceso para construir la base de datos del VAB provincial, mediante una exploración a la sintaxis construida con el software de uso libre “R”.

3.2. Insumos

Para la construcción de esta sintaxis se requieren los siguientes insumos:

1. Base de datos unificada: esta base de datos incluye información económica primaria de los sectores institucionales del gobierno general (GG), sociedades no financieras (SNF), instituciones sin fines de lucro (ISFLH) y hogares productores a nivel provincial. No obstante, la información provincial no está disponible para todos los sectores o subsectores, especialmente del sector privado; por lo cual, para estos segmentos, los valores de producción provincial son estimados bajo un método descendente que toma a la distribución provincial de beneficiarios como variable clave de asignación provincial.
2. Matrices de distribución económica: se incluyen matrices previamente distribuidas en los valores económicos a nivel provincial para los



programas educativos del MINEDUC, coordinaciones zonales y hogares productores.

3. Matrices de distribución de beneficiarios: estas matrices se generan para todas las industrias a nivel desagregado de las CSE. Por lo cual, se aplican siempre que la información primaria de la base unificada no tenga información provincial suficiente o simplemente carezca de ella.
4. Nomenclaturas de las CSE: clasificador de nomenclaturas que correlaciona los sectores institucionales con las industrias de enseñanza (niveles educativos).
5. Clasificador geográfico: corresponde al clasificador de codificación provincial del INEC.
6. VAB, Producción y Consumo intermedio a nivel nacional y por industrias: esta información se obtiene con el procesamiento de las bases de datos de Economía Total y VAB Nacional.

Con estos insumos se desarrolla el código de programación que homologa e integra los insumos de distribución, y posteriormente genera el cálculo de las variables que conforman la base de datos del VAB provincial.

3.3. Código de programación

A continuación, se observa en detalle el proceso que contiene la sintaxis elaborada para la construcción de la base de datos de Economía Total.

El procesamiento inicia con el tratamiento de la base unificada (microdato a nivel provincial) y matrices con valores económicos previamente distribuidas según niveles de educación (industrias) y a nivel provincial.

```
library(tidyverse)
library(tidyr)
library(openxlsx)
library(reshape2)
library(mice)
library(cartography)
library(sf)
library(mapsf)

#### PROCESAR VAB TERRITORIAL ####

area_trabajo<-"C:/Users/henry/Desktop/ASIN/PROCESAMIENTO/CSE2019"
setwd(area_trabajo)

ini_territ = 2018
fin<-2019

### 1 Funciones de distribución y consolidación de datos
# se preparan funciones para consolidar datos, distribuir y ajustar valores

source("Fun_CS_v01.r")
```



```
### 2. Construir y homologar matrices de distribución con valores económicos

#Clasificador geográfico - provincial
clasifi_prov = read.xlsx("5_Matriz_Territ_CSE_2018_19_F.xlsx",sheet = "Clasif_Geog",
                        rows = c(5:29))

# Matriz a nivel provincial
mtd2 = read.xlsx("3.1_CSE_MatrizDistrib_2015-19_b.xlsx",sheet = 1,startRow = 1)
mtd2$Total = NULL

mtd2 = melt(mtd2,id= c("ejercicio","descripcion","id_registro"),value.name = "distrib",
           na.rm = T)

mtd2 = mtd2 %>% filter(distrib>0)
#"1_BUCSE2014-19_Sdist.csv"
bdp<- read.csv2("1_BUCSE2015-19_Sdistv2xu.csv",sep = ",", dec = ".",stringsAsFactors =F)
bdp<- bdp[bdp$ejercicio>2014,]

table(is.na(bdp$cod_provincia))
if (class(bdp$cod_provincia)=="integer" ){
  bdp = bdp %>% mutate(cod_provincia=ifelse(is.na(cod_provincia),99,cod_provincia),
                    cod_provincia = ifelse(cod_provincia<10,paste0("0",cod_provincia),
                    as.character(cod_provincia)))
}
table(is.na(bdp$cod_provincia))
table((bdp$cod_provincia))
nrow(bdp)

#### S14
bdp1 = bdp %>% filter(substr(codigo_N6,1,6)=="S14.01" & ejercicio>2017)
mtd1 = read.xlsx("5_Matriz_Territ_CSE_2018_19_F.xlsx",sheet = "S14",startRow = 6)

mtd1 = mtd1 %>% select(ejercicio,codigo_N6,cod_provincia,descr_provincia,valor)
bdp1 = bdp1 %>% select(ejercicio,tipo,codigo_N1,codigo_N6,descr_codigo_N6,cod_CN,devengado)
sum(bdp1$devengado)
bdp1 = bdp1 %>% left_join(mtd1)

bdp1 = bdp1 %>%
  group_by(ejercicio,codigo_N1,codigo_N6,descr_codigo_N6,tipo,cod_CN) %>%
  nest() %>%
  mutate(mod_obj = map(data, ~ participa(., "cod_provincia","valor",
    "devengado",nuevo_df = 2),id="ver")) %>%
  select(ejercicio,codigo_N1,codigo_N6,descr_codigo_N6,tipo,cod_CN,mod_obj) %>%
  unnest(mod_obj)
sum(bdp1$devengado_dist,na.rm = T)
bdp1$devengado = bdp1$devengado_dist
bdp1 = bdp1 %>% select(ejercicio,tipo,codigo_N1,codigo_N6,descr_codigo_N6,cod_CN,
                    cod_provincia,devengado)
bdp1 = bdp1 %>% left_join(clasifi_prov[,1:2])
bdp1$fuente = "BBD unificada hogares"

bdp = bdp %>% filter(!(codigo_N6 %in% unique(bdp1$codigo_N6) & ejercicio>2017))
bdp = bind_rows(bdp,bdp1)
sum(bdp$devengado,na.rm = T)
```



```
#### universidades distribución por sedes
sum(bdp$devengado,na.rm = T)

bdp1 = bdp %>% filter(id_registro %in% c("1034","1031","1029","1044","1041","1042",
                                         "1038","1002","1079","1012") & ejercici
o>2017)

mtd1 = bind_rows(read.xlsx("5_Matriz_Territ_CSE_2018_19_F.xlsx",sheet = "C6.01.0
1",startRow = 6),
                 read.xlsx("5_Matriz_Territ_CSE_2018_19_F.xlsx",sheet = "C6.02.0
1",startRow = 6))
mtd1 = mtd1 %>% filter(id_registro %in% c("1034","1031","1029","1044","1041","10
42",
                                         "1038","1002","1079","1012"))

# UTPL se gestiona desde información salarial de las sedes, si los centros tienen
más de 5
# trabajadores se considera la producción
mtd1 = mtd1 %>% mutate(beneficiarios = ifelse(id_registro=="1031",
                                              ifelse(round(docentes)>4,round(d
ocentes),0),beneficiarios))
# Espoch se reduce el costo de la planta administrativa de la sede central y el re
sto se distribuye
# según alumnos, el porcentaje identificado es 24%.
mtd1 = mtd1 %>%
  mutate(beneficiarios = ifelse(id_registro=="1002" & cod_provincia!="06",
                                beneficiarios*(1-0.24),beneficiarios))
mtd1 = mtd1 %>% select(ejercicio:beneficiarios)

bdp1 = bdp1 %>% group_by(ejercicio,id_registro,tipos,codigo_N1,codigo_N6,descr_co
digo_N6,cod_CN) %>%
  summarise(devengado=sum(devengado,na.rm = T))

sum(bdp1$devengado)
bdp1 = bdp1 %>% left_join(mtd1)

bdp1 = bdp1 %>%
  group_by(ejercicio,id_registro,codigo_N1,codigo_N6,descr_codigo_N6,tipos,cod_CN
) %>%
  nest() %>%
  mutate(mod_obj = map(data, ~ participa(., "cod_provincia","beneficiarios",
                                         "devengado",nuevo_df = 2))) %>%
  select(ejercicio,id_registro,codigo_N1,codigo_N6,descr_codigo_N6,tipos,cod_CN,m
od_obj) %>%
  unnest(mod_obj)

sum(bdp1$devengado_dist,na.rm = T)
bdp1$devengado = bdp1$devengado_dist
bdp1 = bdp1 %>% select(ejercicio,id_registro,tipos,codigo_N1,codigo_N6,descr_codi
go_N6,cod_CN,
                    cod_provincia,devengado)

bdp1 = bdp1 %>% left_join(clasifi_prov[,1:2])
mtd1 = mtd1 %>% group_by(id_registro,descr_entid_inst) %>% summarise(n())
bdp1 = bdp1 %>% left_join(mtd1[,1:2])
bdp1$fuente = "BBD universidades sede distribuida"

bdp = bdp %>% filter(!(id_registro %in% c("1034","1031","1029","1044","1041","10
42",
                                         "1038","1002","1079","1012") & ejercici
o>2017))
```




```
bdp = bind_rows(bdp,bdp1)
sum(bdp$devengado,na.rm = T)

# Distribución de zonales en distritos
mtd3 = read.xlsx("3.1_CSE_MatrizDistrib_2015-19_b.xlsx",sheet = 2,startRow = 1)

bdp2 = bdp %>% filter(id_registro %in% unique(mtd3$id_registro) & ejercicio>=in
i_territ)
bdp = bdp %>% filter(!(id_registro %in% unique(mtd3$id_registro) & ejercicio>=i
ni_territ))

bdp2 = bdp2 %>% group_by(ejercicio,tipo,codigo_N6,id_registro,cod_entid_inst,
cod_provincia,cod_CN) %>% summarise(devengado=round(sum
(devengado,na.rm = T)))
# Valor que se imputa a las coordinaciones zonales, similar a distritales = 0,10
661957
bdp2 = bdp2 %>% group_by(ejercicio,tipo,codigo_N6,id_registro,cod_entid_inst,
cod_provincia,cod_CN) %>% summarise(devengado=round(sum
(devengado,na.rm = T)))
bdp2 = bdp2 %>% mutate(devengado1 = devengado - round(devengado*(1-0.10661957)),
devengado = devengado - devengado1)
zonales = bdp2 %>% select(-devengado,devengado = devengado1)
zonales$fuente = "Distribuido desde BBD unificada zonales"

bdp2 = bdp2 %>% select(-devengado1)

bdp2 = bdp2 %>% mutate(descripcion = ifelse(cod_CN %in% c("P.2","P.51"),"Alumnos
","Docentes"))
sum(bdp2$devengado,na.rm = T)

bdp2 = merge(as.data.frame(bdp2),mtd3,by = c("ejercicio","descripcion","id_regis
tro"),all.x = T)
bdp2 = bdp2 %>%
mutate(id_registro = Cod_Distrito)

bdp2 = bdp2 %>%
group_by(ejercicio,descripcion,tipo,codigo_N6,cod_entid_inst,cod_provincia,cod
_CN) %>%
nest() %>%
mutate(mod_obj = map(data, ~ participa(., "id_registro", "Porcentaje",
"devengado",nuevo_df = 2),id="ver")) %>
%
select(ejercicio,descripcion,tipo,codigo_N6,cod_entid_inst,cod_provincia,cod_C
N,mod_obj) %>%
unnest(mod_obj)
bdp2$devengado = bdp2$devengado_dist
bdp2$codigo_N6 = "S13.01.01.02.01.02"
sum(bdp2$devengado,na.rm = T)
bdp2$cod_provincia = bdp2$cod_provincia %>% replace_na(99)
table(is.na(bdp2$cod_provincia))
bdp2$fuente = "Distribuido desde BBD unificada zonales"

zonales = bind_rows(zonales,bdp2)
zonales = zonales %>% select(-c("descripcion","Porcentaje","clave_dist","devenga
do_dist"))
sum(zonales$devengado)
bdp = bind_rows(bdp,zonales)

#####
# S13 (gobierno), S11 (SNF) y S15 (ISFLH)
bdp2 = bdp %>% filter(id_registro %in% unique(mtd2$id_registro) & ejercicio>=in
i_territ)
```



```
bdp = bdp %>% filter(!(id_registro %in% unique(mtd2$id_registro) & ejercicio>=i
ni_territ))
nrow(bdp)+nrow(bdp2)

bdp2 = bdp2 %>% select(ejercicio, tipo, codigo_N1, codigo_N6, descr_codigo_N6, id_reg
istro,
                      cod_entid_inst, cod_provincia, descr_provincia, cod_CN, deven
gado)

bdp2 = bdp2 %>% group_by(ejercicio, tipo, codigo_N6, id_registro, cod_entid_inst,
                      cod_provincia, cod_CN) %>% summarise(devengado=round(sum
(devengado, na.rm = T)))

bdp2 = bdp2 %>% mutate(descripcion = ifelse(cod_CN %in% c("P.2", "P.51"), "Alumnos
", "Docentes"))
sum(bdp2$devengado, na.rm = T)

bdp2 = merge(as.data.frame(bdp2), mtd2, by = c("ejercicio", "descripcion", "id_regis
tro"), all.x = T)
bdp2 = bdp2 %>%
  mutate(id_registro = ifelse(id_registro %in% c("S11.01.0X", "S15.01.02", "S15.01
.04"),
                      cod_entid_inst, id_registro))

table(is.na(bdp2$cod_provincia))

bdp2 = mutate(bdp2, id_grupo = paste( sep = ":", as.character(ejercicio),
                      codigo_N6, descripcion, id_registro,
                      cod_provincia, as.character(tipo), cod_CN),
              v_dist = round(devengado * distrib))

# ajustar diferencias por distribución
Sys.time()
bdp2 = bdp2 %>%
  group_by(id_grupo) %>%
  nest() %>%
  mutate(mod_obj = map(data, ~ ajusta_dist(., "devengado", "v_dist", id="ver"))) %>
%
  select(id_grupo, mod_obj) %>%
  unnest(mod_obj)
Sys.time()

sum(bdp2$v_dist, na.rm = T)

#View(bdp2)
bdp2$id_grupo = NULL
bdp2$codigo_N6 = bdp2$variable
bdp2$codigo_N1 = substr(bdp2$codigo_N6, 1, 3)
bdp2$devengado = bdp2$v_dist
bdp2$variable = NULL
bdp2$v_dist = NULL
bdp2$distrib = NULL
bdp2$descripcion = NULL
bdp2$fuente = "ValorDistribuidoNiveles"
bdp = bind_rows(bdp, bdp2)

# incluir matriz de programas educativos

bdp1 = read.xlsx("6_Progr_MINEDUC_2015-19.xlsx", sheet = 1, startRow = 1)
bdp1 = pivot_longer(bdp1, S13.01.02.01.02.01:S13.01.04.02.03.01, names_to = "codig
o_N6",
                    values_to = "devengado")
bdp1 = bdp1 %>% select(ejercicio, descr_programa=programa, id_registro, cod_provinc
```



```
ia,
                                descr_provincia,codigo_N6,devengado)
bdp1 = bdp1 %>% mutate(codigo_N1 = substr(codigo_N6,1,3),cod_CN = "P.2",
                        fuente = "Programas educativos MINEDUC")
bdp = bind_rows(bdp,bdp1)

setwd(direccion0)
write.csv(bdp,"bdg_distr_prov_niveles.csv")
setwd(area_trabajo)

### 2 Cargar VAB, Prod, CI Nacional
VAB_N = VAB %>% filter(PRODUCTOS %in% c("P.1","P.2","B.1b") & EJERCICIO>=ini_territ ) %>%
group_by(ejercicio=EJERCICIO,codigo_industria,PRODUCTOS) %>%
summarise(Corriente=sum(Corriente,na.rm = T),Constante=sum(Constante,na.rm = T))
%>%
gather(key = "Tipo",value="Total",Corriente,Constante) %>%
pivot_wider(names_from = c( PRODUCTOS,Tipo),values_from = Total)
VAB_N$ejercicio = as.numeric(VAB_N$ejercicio)

### 3.1 Procesar Matrices de distribución desde microdato

bdp2 = bdp %>% filter(cod_provincia %in% c( "01","02","03","04","05","06","07","08",
                                           "09",
                                           "10","11","12","13","14","15","16","17",
                                           "18",
                                           "19","20","21","22","23","24" ))

bdp2 = bdp2 %>% filter((cod_CN=="P.11" | (tipo ==2 & cod_CN %in% c("D.11","D.121",
"D.122","D.29","P.2")))) %>% select(ejercicio,codigo_N6,cod_provincia,tipo,cod_CN,devengado) %>%
left_join(clasif[,c("codigo_N6","codigo_N1","codigo_industria","descr_codigo_industria")],
          by="codigo_N6")

bdp2 = as.data.frame(bdp2)
glimpse(bdp2)

bdp2 = bdp2 %>%
group_by(ejercicio,codigo_industria,descr_codigo_industria,codigo_N1,codigo_N6,cod_provincia,tipo) %>%
summarise(devengado=sum(devengado,na.rm = T)) %>%
pivot_wider(names_from = tipo,values_from=devengado,names_prefix="dev") %>%
mutate(produccion=ifelse(codigo_N1 %in% c("S11","S14"),dev1,dev2)) %>%
select(-starts_with("dev"))
```

Procesadas las bases de datos con información económica a nivel provincial, el siguiente paso se orienta a procesar una base de datos con claves de distribución basada en datos físicos de beneficiarios y docentes en cada industria de la educación y a nivel provincial.

Al finalizar el proceso se tiene una matriz consolidada de distribución, que recoge las claves de distribución basada en 1) valores económicos provinciales obtenidos desde el microdato, 2) valores de alumnos (beneficiarios) distribuidos según industria y provincia, 3) Valores de docentes distribuidos al igual que los alumnos.

Como se mencionó previamente, los valores económicos registran un nivel de cobertura alto para el sector público, mientras para el sector privado se observó



falta de cobertura de información en varias provincias; por lo cual el sector privado en gran medida se estima mediante las claves de alumnos o docentes a nivel provincial.

```
# 3.2 Generación de matrices distribuidas
setwd(area_trabajo)
#Datos de estudiantes y docentes a nivel territorial y por industria

wb = "5_Matriz_Territ_CSE_2018_19_F.xlsx"
nombres = getSheetNames(wb)
vhojas = nombres[c(3:25)]

MDist = consolidaHojas(wb,vhojas,finicio=6)
unique(MDist$id_hoja)
MDist = MDist %>% filter(id_hoja!="C6.01.01_2")

MDist = drop_na(MDist,id_hoja)

MDist = MDist %>% group_by(ejercicio,cod_provincia,codigo_industria,
  descr_codigo_industria) %>% summarise(beneficiarios = sum(beneficiarios,na.rm =
T),
  docentes = sum(docentes,na.rm = T),produccion = sum(produccion,na.rm = T))

Alumnos_prov = MDist %>% filter(substr(codigo_industria,1,1) %in% c("2","3","4","
5","6")) %>%
  select(1:5)
clasif_ind = clasif %>%
  group_by(codigo_industria,descr_codigo_industria,
    codigo_industria_agg,descr_codigo_industria_agg) %>% summarise(n()) %>%
  select(-5)

Alumnos_prov = Alumnos_prov %>% left_join(clasif_ind)

# Imputación valores faltantes
# Datos de producción a nivel territorial
temp1 = expand.grid(ejercicio=2018:2019,codigo_industria=c("08.01.0","09.01.0","1.
01.01",
  "10.01.0","11.01.0","12.01.0","13.01.0","2.01.01","2.01.02","2.02.02","3.01.01",
  "3.02.01","4.01.01","4.01.02","4.02.01","4.02.02","6.01.01","6.02.01","7.01.01",
  "7.02.01","2.02.01","5.0
1.01","5.02.01"),
  cod_provincia=c("01","02","03","04","05","06","07","08","09","1
0","11","12","13",
  "14","15","16","17","18","19","20","21","22","2
3","24"))

temp = bdp2 %>% group_by(ejercicio,cod_provincia,codigo_industria) %>%
  summarise(valor = sum(produccion,na.rm = T)) %>%
  filter(ejercicio>2017)

temp = temp %>% right_join(temp1) %>% left_join(MDist[,c("ejercicio","cod_provinci
a",
  "codigo_industria","produccion","beneficiarios","docentes")])

temp = clasif %>% group_by(codigo_industria,descr_codigo_industria,producto) %>%
  summarise(n()) %>% filter(codigo_industria!="99999") %>%
  select(codigo_industria,descr_codigo_industria,producto) %>% right_join(temp) %>
%
  left_join(clasifi_prov[,1:2])

MDist = temp %>% filter(!(producto=="Característico" & !(codigo_industria %in% c(
"1.01.01","2.02.01",
```



```
      "5.01.01", "5.02.01", "6.01.01", "6.02.01"))))

temp = temp %>%
  filter(producto=="Característico" & !(codigo_industria %in% c("1.01.01", "2.02.01",
    "5.01.01", "5.02.01", "6.01.01", "6.02.01")))

# Imputación valores faltantes

temp = temp %>% mutate(valor = ifelse(valor==0 & (beneficiarios>0 | docentes > 0),
NA,valor),
  beneficiarios = ifelse(beneficiarios==0 & docentes > 5,NA,b
eneficiarios),
  docentes = ifelse(docentes>5 & beneficiarios > 0,NA,docente
s))

sapply(temp, function(x) sum(is.na(x)))

cor(temp$valor,temp$beneficiarios,use = "complete.obs")
cor(temp$valor,temp$docentes,use = "complete.obs")

# Ejecutamos una imputación múltiple: 1) Generamos imputaciones 2) Estimamos coefi
cientes puntuales
# 3) Agrupamos los estimadores 4) Predecimos valores
# Generamos imputaciones múltiples m=100

init = mice(temp, maxit=0)
meth = init$method
predM = init$predictorMatrix
predM[, c("valor")] = 0
meth[c("codigo_industria", "descr_codigo_industria", "producto",
  "ejercicio", "cod_provincia", "produccion", "descr_provincia")] = ""
temp1 = mice(temp, method=meth, predictorMatrix=predM, m=40, seed = 2021)
summary(temp1)

# Estimación del modelo predictivo en cada imputación
lm_valor=with(temp1, lm(valor~-1+beneficiarios+docentes))
lm_benefi=with(temp1, lm(beneficiarios~-1+docentes))
lm_docent=with(temp1, lm(docentes~-1+beneficiarios))

lm_valor=pool(lm_valor)
lm_benefi=pool(lm_benefi)
lm_docent=pool(lm_docent)
summary(lm_valor)
summary(lm_benefi)
summary(lm_docent)

# Opción 2 de imputación, predecir valores con el modelo ajustado
x1 = lm_benefi$pooled$estimate[1]
x2 = lm_docent$pooled$estimate[1]
x3 = lm_valor$pooled$estimate[1:2] #1 beneficiarios 2 docentes

temp = temp %>% mutate(beneficiarios = ifelse(is.na(beneficiarios), docentes*x1, ben
eficiarios),
  docentes = ifelse(is.na(docentes), beneficiarios*x2, docentes
),
  valor = ifelse(is.na(valor), beneficiarios*x3[1]+docentes*x3
[2], valor))
temp$produccion = temp$valor

sapply(temp, function(x) sum(is.na(x)))
```



```
par(mfrow=c(1,2))
plot(round(temp$beneficiarios/1000),round(temp$valor/1000))
plot(round(temp$docentes),round(temp$valor/1000))

dev.off()

# Se trabajará con la opción 2 de estimación mediante regresión de imputación por
# ser robusta

# Organización de base de datos de distribución final

MDist = MDist %>% mutate( valor = ifelse(is.na( valor),0,valor), produccion = valo
r )
MDist = bind_rows(MDist,temp)
MDist$valor = NULL
MDist = MDist %>% mutate( beneficiarios = ifelse(is.na( beneficiarios),0,beneficia
rios),
                        docentes = ifelse(is.na( docentes),0,docentes))
MDist = MDist %>% pivot_longer(cols=c(beneficiarios,docentes,produccion),names_to
= "clave",
                              values_to = "valor")
MDist = MDist %>% left_join(clasifi_prov)

#MDist = bind_rows(MDist,temp)
MDist = MDist %>% group_by(ejercicio,cod_region,descr_region,cod_provincia,descr_
provincia,
                          producto,codigo_industria,descr_codigo_industria,clave) %>%
  summarise(valor = sum(valor,na.rm = T))

MDist = full_join(MDist,VAB_N,by=c("ejercicio","codigo_industria"))
MDist = MDist %>% mutate(prioridad = ifelse(clave=="produccion",1,0))

MDist = MDist %>% group_by(ejercicio,codigo_industria,clave) %>%
  mutate(valor_ag = sum(valor,na.rm = T),cobert=round(valor_ag/P.1_Corriente*1000)
)
# Si el valor de producción de la información territorial de base cubre a menos el
75%
# del valor real de la producción a nivel nacional, se toma como fuente de distrib
ución
# en cada sub-industria de la educación del sector público, caso contrario se toma
la clave
# de número de beneficiarios.

MDist = MDist %>% mutate(prioridad = ifelse(prioridad==1 & cobert>=0.60,1,0))
MDist = MDist %>% mutate(prioridad = ifelse(prioridad==1 & codigo_industria=="1.01
.01",0,prioridad))
# Se trabaja con el dato de empleo en desarrollo infantil público, al no existir d
esagregación provincial
MDist = MDist %>% mutate(prioridad = ifelse(prioridad==1 & codigo_industria=="2.01
.01",0,prioridad))
MDist = MDist %>% mutate(prioridad = ifelse(prioridad==1 & codigo_industria=="2.01
.02",0,prioridad))
MDist = MDist %>% mutate(prioridad = ifelse(prioridad==1 & codigo_industria=="2.02
.01",0,prioridad))
#MDist = MDist %>% mutate(prioridad = ifelse(prioridad==1 & codigo_industria=="2.0
2.02",0,prioridad))
# si bien existe buena cobertura en valores de producción Los ITTs privados, existe
n muchas provincias
# sin datos (13) por lo cual se trabajará con datos de empleo desde la base del di
rectorio de empresas
MDist = MDist %>% mutate(prioridad = ifelse(prioridad==1 & codigo_industria=="3.01
.01",0,prioridad))
MDist = MDist %>% mutate(prioridad = ifelse(prioridad==1 & codigo_industria=="4.01
```



```
.01",0,prioridad))
MDist = MDist %>% mutate(prioridad = ifelse(prioridad==1 & codigo_industria=="4.01
.02",0,prioridad))
MDist = MDist %>% mutate(prioridad = ifelse(prioridad==1 & codigo_industria=="5.01
.01",0,prioridad))
MDist = MDist %>% mutate(prioridad = ifelse(prioridad==1 & codigo_industria=="5.02
.01",0,prioridad))
MDist = MDist %>% mutate(prioridad = ifelse(prioridad==1 & codigo_industria=="6.02
.01",1,prioridad))
MDist = MDist %>% mutate(prioridad = ifelse(prioridad==1 & codigo_industria=="6.01
.01",1,prioridad))
MDist = MDist %>% mutate(prioridad = ifelse(prioridad==1 & codigo_industria=="7.01
.01",0,prioridad))
MDist = MDist %>% mutate(prioridad = ifelse(prioridad==1 & codigo_industria=="7.02
.01",0,prioridad))

#identifico las sub-industria que no tienen clave de distribución (producción) vál
ida
temp = MDist %>% filter( prioridad==0 )
# Filtro matriz de claves basadas en la producción donde prioridad == 1
MDist = MDist %>% filter(prioridad==1)

temp = temp %>% filter(!(codigo_industria %in% unique(MDist$codigo_industria)))

unique(temp$codigo_industria)
head(temp)

## Identificar matriz de claves basadas en los beneficiarios como prioridad == 2
# Se toma como siguiente variable clave de prioridad al dato de beneficiarios, no
obstante debido
# a que se trata de registros nuevos, se trabaja con datos consolidados entre el 2
018 y 2019
# y ese valor se aplicará en la distribución de cada año.
#ajuste para 5.01.01 ITTs privados se trabaja con el dato de empleo, se reemplaza
en beneficiarios
# para que se ejecute el proceso
temp = temp %>%
  mutate(clave = ifelse(codigo_industria=="5.01.01" & clave=="beneficiarios","doce
ntes1",
                        ifelse(codigo_industria=="5.01.01" & clave=="docentes","be
neficiarios",clave)))

if (nrow(temp) > 0) {
  temp = temp %>% filter(clave=="beneficiarios") %>% group_by(codigo_industria,cod
_provincia) %>%
  mutate(valor =sum(valor,na.rm = T))
  temp$prioridad = 2

  MDist = bind_rows(MDist,temp )
}

MDist$cobert = NULL
MDist$valor_ag = NULL
```

Una vez que se tienen las matrices de distribución basadas en datos monetarios y físicos, se generan los agregados a nivel territorial y se estructura la base de datos del VAB provincial.

```
### 4 Calcular y estructurar La base de datos del VAB provincial
temp = NULL
for (h in c( "P.1_Corriente", "P.2_Corriente", "B.1b_Corriente")){
  VAB_Prov = MDist %>%
```




```
group_by(ejercicio,prioridad,codigo_industria,descr_codigo_industria) %>%  
nest() %>%  
mutate(mod_obj = map(data, ~ participa(., "cod_provincia", "valor",  
                                     h,nuevo_df = 2,dec = 10L),id="ver")) %>%  
select(ejercicio,prioridad,codigo_industria,descr_codigo_industria,mod_obj) %>%  
unnest(mod_obj)  
VAB_Prov = VAB_Prov %>% left_join(clasifi_prov)  
VAB_Prov = clasif %>% group_by(codigo_industria,producto,sector) %>%  
summarise(n()) %>% filter(codigo_industria!="99999") %>%  
select(codigo_industria,producto,sector) %>% right_join(VAB_Prov)  
val_dist = paste0(h,"_dist")  
VAB_Prov = VAB_Prov %>% mutate(indicador = h) %>%  
rename(agregado = h,agregado_dist = val_dist)  
temp = bind_rows(VAB_Prov,temp)  
}  
VAB_Prov = temp %>% filter(producto=="Característico")  
View(VAB_Prov)  
  
VAB_Prov = clasif %>% group_by(codigo_industria,codigo_industria_agg,descr_codigo_  
industria_agg) %>%  
summarise(n()) %>% filter(codigo_industria!="99999") %>% right_join(VAB_Prov)
```

El archivo de sintaxis (script completo) que automatiza la construcción de la base de datos del VAB provincial y genera los tabulados se encuentra ubicado en:

"R:\CGTPE\DECON\AS\CS_MPE_2021\CSE_2018-
19\5_Proc\5.7_Finali_archiv_dat\5.7.2_Compil_prod_ant\1_Tabulados\7_Indic
a_Territ_CSE\0_Arch_Procesam"

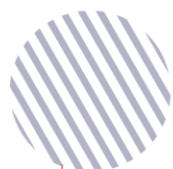
Nombre del archivo: CSE_Prov2019

3.4. Estructura base de datos VAB provincial procesada

Una vez que se ejecuta la sintaxis se obtiene una BDD con una estructura de 17 variables, como se muestra en la siguiente tabla:

Tabla 1: Descripción de variables de la base datos de VAB provincial.

Nro.	Variables	Descripción de variables
1	Ejercicio	Ejercicio o año de la información en la base de datos.
2	Sector	Sector al que pertenece la institución (público o privado).
3	Producto	Tipo de producto de las CSE (característico o conexo).
4	Codigo_industria_agg	Código de industria de las CSE a nivel agregado
5	Descr_codigo_industria_agg	Descripción del código de industria de las CSE a nivel agregado.
6	Codigo_industria	Código de la industria de las CSE.



7	Descr_codigo_industria	Descripción del código de la industria de las CSE.
8	Cod_region	Código de las regiones naturales del Ecuador.
9	Descr_region	Descripción del código de las regiones naturales del Ecuador.
10	Cod_provincia	Código de la provincia según división política administrativa.
11	Descr_provincia	Descripción del código de la provincia según división política administrativa.
12	Indicador	Tipo de agregado económico (P.1 = producción, P.2 = consumo intermedio y B.1b=VAB).
13	Agregado	Valor del agregado económico a nivel nacional.
14	Prioridad	Indica el tipo de variable usada como clave de distribución (1: variable monetaria y 2: Variable física).
15	Valor	Valor absoluto de la clave de distribución.
16	Clave_dist	Es el porcentaje que asigna la clave de distribución para cada provincia.
17	Agregado_dist	Valor del agregado económico a nivel provincial.

La base de datos finalizada se archiva en la siguiente dirección:
R:\CGTPE\DECON\AS\CS_MPE_2021\CSE_2018-
19\5_Proc\5.7_Finali_archiv_dat\5.7.1_Compil_bas_dat\3_Resultados\RESULTA
DOS_22f

Nombre del Archivo: "VAB_Prov.xlsx"

3.5. Resultados: mapas provinciales de la producción, consumo intermedio y VAB de la educación

A manera de ilustración, a continuación, se muestra la sintaxis para generar el mapa provincial del valor de la producción de los servicios de educación a nivel provincial, el cual también incluye información de la población en cada provincia.

```
### 5 Generar Mapas de resultados
## Lectura de mapas

temp=paste0(area_trabajo, "/SHP")
setwd(temp)
MapEcu <- st_read("nxprovincias.shp")
```



```
### Generación de mapas temáticos y escritura NIVEL 2
setwd(area_trabajo)
poblacion = read.xlsx("5_Matriz_Territ_CSE_2018_19_F.xlsx",sheet = "Poblacion",rows = c( 5:29))
x2<- loadWorkbook ("7_Indicadores_Territoriales_CSE_2007-19.xlsx")
hojas_x2 = x2$sheet_names

setwd(dir_terri)
# Figura 1.1 Producción y población

prod_prov = VAB_Prov %>% filter(ejercicio==2019 & indicador=="P.1_Corriente") %>%
  group_by(ejercicio,cod_provincia) %>%
  summarise(produccion=round(sum(agregado_dist,na.rm = T))) %>%
  left_join(poblacion[,c(1,4)]) %>% arrange(cod_provincia)
prod_prov$Poblac_2019 = prod_prov$Poblac_2019 / 1000

Map1 = left_join(MapEcu,prod_prov,by=c("DPA_PROVIN"="cod_provincia"))
Map2 = Map1[Map1$DPA_PROVIN=="20",]
Map1 = Map1[ Map1$DPA_PROVIN!="20",]

rangos1 = round(quantile(prod_prov$produccion,probs = c(0,0.2,0.4,0.6,0.8,1),na.rm=T))
#rangos2 = round(quantile(prod_prov$Poblac_2019,probs = c(0,0.25,0.5,0.85,1)))
rangos2 = c(100,500,1000,3000,max(prod_prov$Poblac_2019))

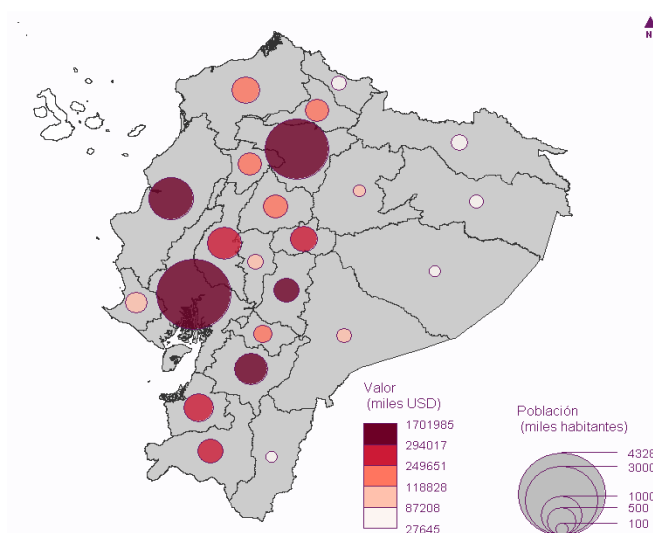
archivo_png = paste0("CSE_", "_PROD", "_1.1.png")
tema = "candy"
png(archivo_png,pointsize=16,
    width = 750, height = 680)
mf_theme(tema)
Map1 %>%
  mf_map() %>%
  mf_prop_choro(c("Poblac_2019","produccion"),leg_pos = c("n","n"),alpha = 0.8,inches = 0.6,pal="Reds")
mf_legend_p(pos = "bottomright",val = rangos2,title = "Población \n (millones habitantes)",cex = 2,inches = 0.7,
            col = "gray",val_rnd = 0,val_cex = 0.9,title_cex = 1,frame = F)
mf_legend_c(pos = c(859832.5,9650022),val = rangos1,title = "Valor \n (millones USD)",cex = 1.47,
            pal = "Reds",val_rnd = 0,val_cex = 0.9,title_cex = 1,frame = F)
mf_inset_on(x = Map2,cex = .21, pos = "topleft")
mf_map(Map2,alpha = 0.9,col = "white",frame.plot = T )
mf_inset_off()
mf_arrow(pos = "topright")

dev.off()
```



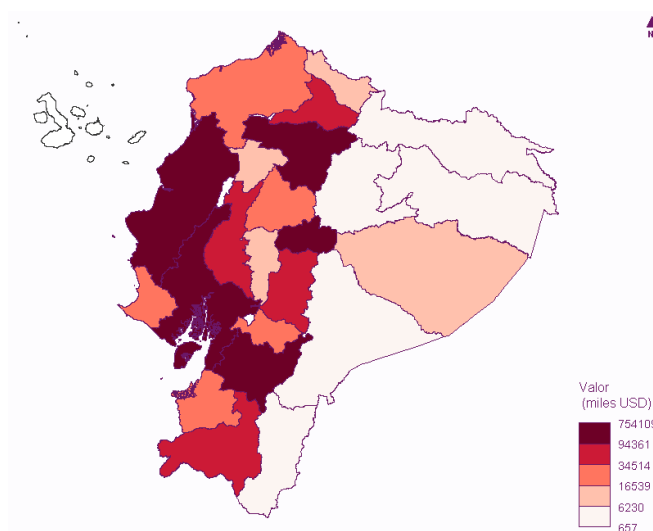
Mediante la ejecución de la sintaxis, es posible obtener los gráficos que se muestran a continuación:

Gráfico 1: Distribución de la producción de educación y población a nivel provincial 2019.



Fuente: INEC

Gráfico 2: Distribución de la producción de las actividades enseñanza superior a nivel provincial 2019.



Fuente: INEC



4. Bibliografía

- UNESCO. (2013). Clasificación Internacional Normalizada de la Educación 2011. CINE 2011. Montreal, Canadá.
- UNESCO. (2016). Methodology of National Education Accounts. Montreal, Canadá.
- INEC. (2012). Clasificación Nacional Central de Productos. Quito, Ecuador.
- INEGI. (2004). Producto Interno Bruto por entidad Federativa, Metodología. México.
- European Commission. (2013). Manual on regional accounts methods" de Eurostat, Luxemburgo, Unión Europea.
- ONU (2016). Sistema de Cuentas Nacionales 2008, New York, Estados Unidos.

FIRMAS DE RESPONSABILIDAD
ELABORADO POR:
Responsable de la Unidad de Gestión de Análisis de Síntesis
Nombre: Henry Valdiviezo



@ecuadorencifras



@ecuadorencifras



@InecEcuador



t.me/ecuadorencifras



INEC/Ecuador



INECEcuador



INEC Ecuador