

# Informe 1: Encuesta Protección Social de Hogares

Matías Deneken

Investigación Alejandra Inostroza

## Introducción

Para los propositos solicitados, se realizó un adecuación de la Encuesta Protección Social (EPS) específicamente las olas del 2004, 2006, 2009, 2015 y 2016. Toda la información se encuentre en **Dropxbox** y **GitHub**

Una primera etapa buscó renombrar las variables de acuerdo a lo solicitado **aquí**. Esto con el fin de darle un seguimiento longitudinal a las variables señaladas en el documento.

Esto se encuentran en diferentes scripts:

- Para la bbdd del 2004 se puede encontrar su procesamiento pinchando **aquí**
- Para la bbdd del 2006 se puede encontrar su procesamiento pinchando **aquí**
- Para la bbdd del 2009 se puede encontrar su procesamiento pinchando **aquí**
- Para la bbdd del 2015 se puede encontrar su procesamiento pinchando **aquí**
- Para la bbdd del 2020 se puede encontrar su procesamiento pinchando **aquí**

Una segunda etapa contrastó que tipos de variables se encontraban de forma similar en cada una de las BBDD.

A partir del procesamiento realizado por mí, y de acuerdo a la información proporcionada por el procesamiento de la asistente anterior y mis correcciones, esto quedaría así:

Variable	2004	2006	2009	2015
folio_n	Sí	Sí	Sí	No
region	Sí	Sí	Sí	Sí
fact_exp	Sí	Sí	No	No
parentesco	Sí	Sí	Sí	Sí
sexo	Sí	Sí	Sí	Sí
edad	Sí	Sí	Sí	Sí
ecivil	Sí	Sí	Sí	Sí
asiste	Sí	Sí	No	Sí
neduc	Sí	Sí	No	No
fineduc	No	No	No	Sí
ncarreras	Sí	Sí	Sí	Sí
prueba	Sí	No	No	No
ingotros	Sí	Sí	No	No
ingnlabe	Sí	Sí	No	No
vivienda	Sí	Sí	Sí	Sí

Variable	2004	2006	2009	2015
montoviv	Sí	Sí	Sí	Sí
ple	Sí	Sí	No	No
pse	Sí	Sí	Sí	Sí
disc	Sí	Sí	No	No
disca	Sí	Sí	No	No
discph	No	No	No	Sí*
ncasado	Sí	Sí	Sí	Sí
hogar	Sí	No	Sí	Sí
npareja	Sí	Sí	No	No
edadpareja	Sí	Sí	Sí	Sí
edadconv	Sí	Sí	Sí	Sí
vivepareja	Sí	Sí	No	No
nhijos	Sí	Sí	Sí	Sí
sexoh	Sí	Sí	Sí	Sí*
anoh	Sí	Sí	Sí	Sí*
hcasa	Sí	Sí	Sí	Sí*
nwork	Sí	Sí	Sí	Sí
anoiw	Sí	Sí	Sí	Sí
anotw	Sí	Sí	Sí	Sí
situationw	Sí	Sí	Sí	No
tipow	Sí	Sí	Sí	Sí
catw	Sí	Sí	Sí	Sí
contrato	Sí	Sí	Sí	Sí
inglab	Sí	Sí	Sí	Sí
hworke	Sí	Sí	Sí	Sí
inact	Sí	Sí	No	No
programa	Sí	Sí	Sí	Sí
npersona	Sí	Sí	Sí	Sí
perteneceh	Sí	Sí	Sí	Sí
edadi	Sí	Sí	Sí	Sí
trabajoi	Sí	Sí	Sí	Sí
contratoi	Sí	No	No	No
inglabi	Sí	Sí	Sí	Sí
hworki	Sí	No	No	No
ingnlabi	Sí	Sí	Sí	Sí
factor_EPS	No	No	No	Sí

A lo largo de los años estudiados, varias variables han sido consistentemente usadas, modificadas o eliminadas, reflejando cambios en la estructuración de datos o en las necesidades de información.

### Variables Consistentes

Algunas variables han aparecido de manera constante a lo largo de todos los años analizados, como **parentesco**, **ecivil**, **vivienda**, y **ncasado**, lo que indica su importancia sostenida en el estudio.

## Eliminaciones y Cambios

- **Eliminaciones Notables:** Variables como **fact\_exp**, **prueba**, **ingotros**, **ple**, y **disc** fueron eliminadas en años posteriores, posiblemente debido a cambios en la metodología de recopilación de datos o prioridades de investigación.
- **Renombramientos:** Se observaron cambios de nombre en algunas variables para clarificar su propósito o reflejar mejor la información que representan, como **disca** a **discpa** y **contrato** a **contratow**.

## Adiciones y Desagregaciones

- **Nuevas Variables en 2015:** Se introdujeron nuevas variables como **fineduc**, **factor\_EPS**, y series desagregadas como **discph1-5**, indicando un enfoque en mayor detalle y especificidad en los datos recogidos.
- **Detalles Añadidos:** La aparición de múltiples variables desagregadas en 2015 como **sexoh1-5** y **anoh1-5** muestra un esfuerzo por capturar más detalles y variabilidad dentro de las categorías existentes.

Este análisis revela una evolución en la recopilación y presentación de datos a lo largo del tiempo, con un esfuerzo por adaptar el conjunto de datos a nuevas necesidades analíticas y metodológicas, manteniendo algunas variables fundamentales mientras se introducen nuevas para una comprensión más profunda de los temas estudiados.

Las bbdd pueden encontrarse **aquí**.

Fueron guardas en formato R y Stata.

```
saveRDS(eps_04, "eps_04.rds")
saveRDS(eps_06, "eps_06.rds")
saveRDS(eps_09, "eps_09.rds")
saveRDS(eps_15, "eps_15.rds")
saveRDS(eps_20_cont, "eps_20_cont.rds")
saveRDS(eps_20_pres, "eps_20_pres.rds")
saveRDS(eps_20_rent, "eps_20_rent.rds")

# Guardar como archivos DTA
write_dta(eps_04, "eps_04.dta")
write_dta(eps_06, "eps_06.dta")
write_dta(eps_09, "eps_09.dta")
write_dta(eps_15, "eps_15.dta")
write_dta(eps_20_cont, "eps_20_cont.dta")
write_dta(eps_20_pres, "eps_20_pres.dta")
write_dta(eps_20_rent, "eps_20_rent.dta")
```

## Proyecciones.

Dada las características de la encuesta, en que el folio se extiende a lo largo de las preguntas, es necesario reacomodar las variables para que puedan ser utilizada de forma longitudinal y análisis de curso de vida. Es decir, depurar las bases de datos de forma conjunta.