

Práctica Dirigida 4

Mg. Sc. J. Eduardo Gamboa U.

Los archivos que se utilizarán en esta práctica dirigida corresponden a la [Encuesta Nacional de Hogares \(ENAH\)](#), llevada a cabo por el INEI de manera trimestral desde hace más de dos décadas. Esta encuesta está dividida por módulos, cada uno de los cuales aborda una temática en particular.

Pregunta 1

El archivo `datos1.txt` contiene datos acerca del módulo 85 (Gobernabilidad, Democracia y Transparencia) correspondientes al último trimestre del año 2024.

Leer el archivo utilizando las funciones `read.table` y `read_delim`, y almacenarlo en data frames de nombres `datos1_a` y `datos1_b`, respectivamente. Luego, visualizar las 5 primeras filas de cada data frame.

```
datos1_a = read.table(file = 'datos1.txt', header = TRUE)
str(datos1_a)
```

```
'data.frame':  8374 obs. of  9 variables:
 $ AÑO      : int  2024 2024 2024 2024 2024 2024 2024 2024 2024 2024 ...
 $ MES      : int  10 10 10 10 10 10 10 10 10 10 ...
 $ CONGLOME: int  15010 15010 15010 15010 15010 15010 15010 15015 15015 15015 ...
 $ VIVIENDA: int   5 10 15 20 25 35 2 15 30 68 ...
 $ HOGAR    : int  11 11 11 11 11 11 11 11 11 11 ...
 $ CODPERSONO: int   1 1 4 2 1 1 1 1 2 2 ...
 $ UBIGEO   : int  10101 10101 10101 10101 10101 10101 10101 10101 10101 10101 ...
 $ DOMINIO  : int   4 4 4 4 4 4 4 4 4 4 ...
 $ P1.01    : int   2 5 1 1 1 1 2 3 9 1 ...
```

```
head(datos1_a, 5)
```

	AÑO	MES	CONGLOME	VIVIENDA	HOGAR	CODPERSONO	UBIGEO	DOMINIO	P1.01
1	2024	10	15010	5	11	1	10101	4	2
2	2024	10	15010	10	11	1	10101	4	5
3	2024	10	15010	15	11	4	10101	4	1
4	2024	10	15010	20	11	2	10101	4	1
5	2024	10	15010	25	11	1	10101	4	1

```
library(readr)
datos1_b = read_delim(file = 'datos1.txt', )
```

Rows: 8374 Columns: 9

-- Column specification -----

Delimiter: " "

dbl (9): AÑO, MES, CONGLOME, VIVIENDA, HOGAR, CODPERSONO, UBIGEO, DOMINIO, P1.01

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.

```
str(datos1_b)
```

```
spc_tbl_ [8,374 x 9] (S3: spec_tbl_df/tbl_df/tbl/data.frame)
 $ AÑO      : num [1:8374] 2024 2024 2024 2024 2024 ...
 $ MES      : num [1:8374] 10 10 10 10 10 10 10 10 10 10 ...
 $ CONGLOME: num [1:8374] 15010 15010 15010 15010 15010 ...
 $ VIVIENDA: num [1:8374] 5 10 15 20 25 35 2 15 30 68 ...
 $ HOGAR    : num [1:8374] 11 11 11 11 11 11 11 11 11 11 ...
 $ CODPERSONO: num [1:8374] 1 1 4 2 1 1 1 1 2 2 ...
 $ UBIGEO   : num [1:8374] 10101 10101 10101 10101 10101 ...
 $ DOMINIO  : num [1:8374] 4 4 4 4 4 4 4 4 4 4 ...
 $ P1.01    : num [1:8374] 2 5 1 1 1 1 2 3 9 1 ...
- attr(*, "spec")=
 .. cols(
 ..   AÑO = col_double(),
 ..   MES = col_double(),
 ..   CONGLOME = col_double(),
 ..   VIVIENDA = col_double(),
 ..   HOGAR = col_double(),
 ..   CODPERSONO = col_double(),
```

```

..  UBIGEO = col_double(),
..  DOMINIO = col_double(),
..  P1.01 = col_double()
.. )
- attr(*, "problems")=<externalptr>

```

```
head(datos1_b,5)
```

```
# A tibble: 5 x 9
```

	AÑO	MES	CONGLOME	VIVIENDA	HOGAR	CODPERSONO	UBIGEO	DOMINIO	P1.01
	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>
1	2024	10	15010	5	11	1	10101	4	2
2	2024	10	15010	10	11	1	10101	4	5
3	2024	10	15010	15	11	4	10101	4	1
4	2024	10	15010	20	11	2	10101	4	1
5	2024	10	15010	25	11	1	10101	4	1

Pregunta 2

El archivo datos2.txt contiene datos acerca del módulo 85 (Gobernabilidad, Democracia y Transparencia) correspondientes al último trimestre del año 2024, pero con información añadida al inicio del archivo.

Leer las 10 primeras filas de la tabla de datos, utilizando las funciones `read.table` y `read_delim`, y almacenarlo en data frames de nombres `datos2_a` y `datos2_b`, respectivamente.

```
datos2_a = read.table(file = 'datos2.txt', skip = 5, nrow = 10, header = TRUE)
datos2_a
```

	AÑO	MES	CONGLOME	VIVIENDA	HOGAR	CODPERSONO	UBIGEO	DOMINIO	P1.01
1	2024	10	15010	5	11	1	10101	4	2
2	2024	10	15010	10	11	1	10101	4	5
3	2024	10	15010	15	11	4	10101	4	1
4	2024	10	15010	20	11	2	10101	4	1
5	2024	10	15010	25	11	1	10101	4	1
6	2024	10	15010	35	11	1	10101	4	1
7	2024	10	15015	2	11	1	10101	4	2
8	2024	10	15015	15	11	1	10101	4	3
9	2024	10	15015	30	11	2	10101	4	9
10	2024	10	15015	68	11	2	10101	4	1

```
datos2_b = read_delim(file = "datos2.txt", skip = 5, n_max = 10)
```

Rows: 10 Columns: 9

-- Column specification -----

Delimiter: " "

dbl (9): AÑO, MES, CONGLOME, VIVIENDA, HOGAR, CODPERSO, UBIGEO, DOMINIO, P1.01

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.

```
datos2_b
```

A tibble: 10 x 9

	AÑO	MES	CONGLOME	VIVIENDA	HOGAR	CODPERSO	UBIGEO	DOMINIO	P1.01
	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>
1	2024	10	15010	5	11	1	10101	4	2
2	2024	10	15010	10	11	1	10101	4	5
3	2024	10	15010	15	11	4	10101	4	1
4	2024	10	15010	20	11	2	10101	4	1
5	2024	10	15010	25	11	1	10101	4	1
6	2024	10	15010	35	11	1	10101	4	1
7	2024	10	15015	2	11	1	10101	4	2
8	2024	10	15015	15	11	1	10101	4	3
9	2024	10	15015	30	11	2	10101	4	9
10	2024	10	15015	68	11	2	10101	4	1

Pregunta 3

El archivo datos3.txt contiene datos acerca del módulo 85 (Gobernabilidad, Democracia y Transparencia) correspondientes al último trimestre del año 2024, pero con un distinto separador de columnas.

Leer el archivo utilizando las funciones read.table, read_delim, y almacenarlo en data frames de nombres datos3_a, datos3_b y datos3_c, respectivamente. Luego, visualizar las 5 primeras filas de cada data frame.

```
datos3_a = read.table(file = 'datos3.txt', header = TRUE)
head(datos3_a,5)
```

	AÑO	MES	CONGLOME	VIVIENDA	HOGAR	CODPERSONO	UBIGEO	DOMINIO	P1.01
1	2024	10	15010	5	11	1	10101	4	2
2	2024	10	15010	10	11	1	10101	4	5
3	2024	10	15010	15	11	4	10101	4	1
4	2024	10	15010	20	11	2	10101	4	1
5	2024	10	15010	25	11	1	10101	4	1

```
datos3_b = read_delim(file = "datos3.txt")
```

Rows: 8374 Columns: 9

-- Column specification -----

Delimiter: "\t"

dbl (9): AÑO, MES, CONGLOME, VIVIENDA, HOGAR, CODPERSONO, UBIGEO, DOMINIO, P1.01

- 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(datos3_b,5)
```

A tibble: 5 x 9

	AÑO	MES	CONGLOME	VIVIENDA	HOGAR	CODPERSONO	UBIGEO	DOMINIO	P1.01
	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>
1	2024	10	15010	5	11	1	10101	4	2
2	2024	10	15010	10	11	1	10101	4	5
3	2024	10	15010	15	11	4	10101	4	1
4	2024	10	15010	20	11	2	10101	4	1
5	2024	10	15010	25	11	1	10101	4	1

```
datos3_c = read_tsv(file = "datos3.txt")
```

```
Rows: 8374 Columns: 9
```

```
-- Column specification -----
```

```
Delimiter: "\t"
```

```
dbl (9): AÑO, MES, CONGLOME, VIVIENDA, HOGAR, CODPERSONA, UBIGEO, DOMINIO, P1.01
```

```
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(datos3_c,5)
```

```
# A tibble: 5 x 9
```

	AÑO	MES	CONGLOME	VIVIENDA	HOGAR	CODPERSONA	UBIGEO	DOMINIO	P1.01
	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>
1	2024	10	15010	5	11	1	10101	4	2
2	2024	10	15010	10	11	1	10101	4	5
3	2024	10	15010	15	11	4	10101	4	1
4	2024	10	15010	20	11	2	10101	4	1
5	2024	10	15010	25	11	1	10101	4	1