

# Clase 9:

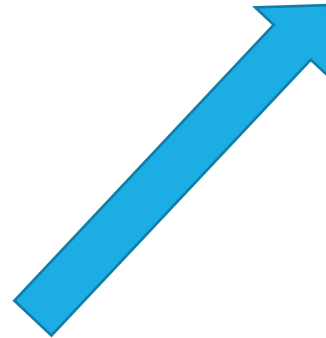
## Operaciones con datos: Reestructurar y contraer bases de datos

The logo for STATA, featuring the word "STATA" in a bold, blue, sans-serif font.

# Operaciones con datos: Ejemplo 1

## Serie histórica de la tasa de cambio del dólar

	fecha	mesaño	tasa
1	27nov1991	1991m11	693.32
2	28nov1991	1991m11	693.99
3	29nov1991	1991m11	694.7
4	30nov1991	1991m11	694.7
5	01dec1991	1991m12	643.42
6	02dec1991	1991m12	643.42
7	03dec1991	1991m12	639.22
8	04dec1991	1991m12	635.7
9	05dec1991	1991m12	631.51
10	06dec1991	1991m12	627.16
11	07dec1991	1991m12	638.06
12	08dec1991	1991m12	638.06
13	09dec1991	1991m12	638.06
14	10dec1991	1991m12	622.92
15	11dec1991	1991m12	627.46
16	12dic1991	1991m12	633.06



1. Contraer

`collapse (mean) tasa, by(mesaño)`

## Tasa promedio mensual

	mesaño	tasa
1	1991m11	694.1775
2	1991m12	630.41258
3	1992m1	644.05806
4	1992m2	635.67172
5	1992m3	640.50194
6	1992m4	649.955
7	1992m5	659.77129
8	1992m6	677.16533
9	1992m7	704.77226
10	1992m8	694.17452
11	1992m9	697.50333
12	1992m10	707.80484

# Operaciones con datos: Ejemplo 2

## Conteo de líderes sociales asesinados

`reshape wide frecuencia, i(departamento) j(año)`

	departamento	frecuen~2016	frecuen~2017	frecuen~2018	frecuen~2019
1	ANTIOQUIA	19	21	30	10
2	ARAUCA	2	1	9	2
3	ATLANTICO	3	2	2	.
4	BOGOTA	1	2	3	.
5	BOLIVAR	1	1	2	5
6	BOYACA	.	.	1	.
7	CALDAS	.	4	1	.
8	CAQUETA	4	3	10	8
9	CASANARE	2	1	.	2
10	CAUCA	47	30	33	23

2. Cambiar la estructura

`reshape long frecuencia, i(departamento) j(año)`

	departamento	año	frecuencia
1	ANTIOQUIA	2016	19
2	ANTIOQUIA	2017	21
3	ANTIOQUIA	2018	30
4	ANTIOQUIA	2019	10
5	ARAUCA	2016	2
6	ARAUCA	2017	1
7	ARAUCA	2018	9
8	ARAUCA	2019	2
9	ATLANTICO	2016	3
10	ATLANTICO	2017	2
11	ATLANTICO	2018	2
12	BOGOTA	2016	1
13	BOGOTA	2017	2
14	BOGOTA	2018	3

`collapse (mean) promedio=frecuencia ///  
(max) maximo=frecuencia ///  
(min) minimo= frecuencia, ///  
by(departamento)`

	departamento	promedio	maximo	minimo
1	ANTIOQUIA	20	30	10
2	ARAUCA	3.5	9	1
3	ATLANTICO	2.3333333	3	2
4	BOGOTA	2	3	1
5	BOLIVAR	2.25	5	1
6	BOYACA	1	1	1
7	CALDAS	2.5	4	1
8	CAQUETA	6.25	10	3

1. Contraer  
Otra vez

# Referencias R

- Del libro R para principiantes
  - [Importar Excel en R](#)
- Curso de manejo de datos de [DataNovia](#)
  - [Renombrar columnas de dataframes con dplyr](#)
  - [Manejo de duplicados con dplyr](#)
- DataCamp
  - Sintaxis de tubo ([Pipe operator](#))
- Curso [R4ULPGC](#): Introducción a R
  - [Librerías Tidyverse](#)
  - [Manejo de fechas](#)
- [R para usuarios de Stata](#)
  - [Manejo de bases en R](#)
  - [Unir y reestructurar](#)
- [Datatofish](#): Tutoriales de ciencia de datos
  - [R](#)
- StackOverflow
  - [Manejo de fechas](#)

# Referencias Python

- Tutoriales de Datacamp
  - [Tutorial de Jupyter Notebook](#)
  - [Tutorial de importación de Excel](#)
- [Datatofish](#): Tutoriales de ciencia de datos
  - [Python: Leer Excel](#)
- Comparación Python y otros:
  - Stata: [Daniel M. Sullivan](#)
  - [Documentación](#) de [Pandas](#): [Stata](#) y [R](#)
- Otros
  - [Manejar con archivos en Python](#)
  - [Eliminar filas de un \*DataFrame\* bajo una condición aplicada a una columna](#)
  - [Eliminar filas/columnas con valores faltantes \(bases de la sintaxis de pandas\)](#)