
DMAIC

**Teorema del límite
central (CLT)**

AUTOPISTA

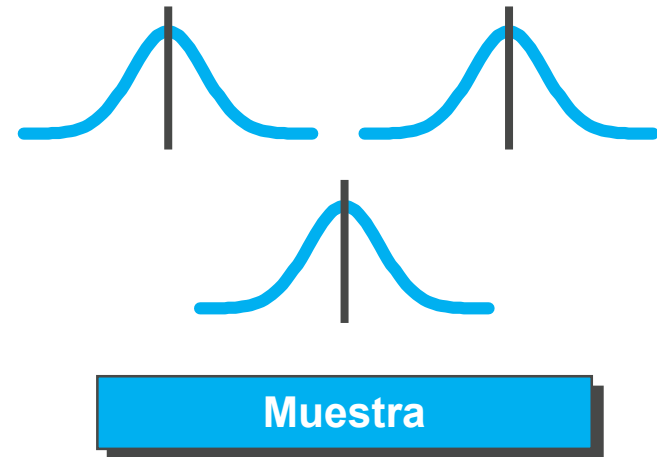
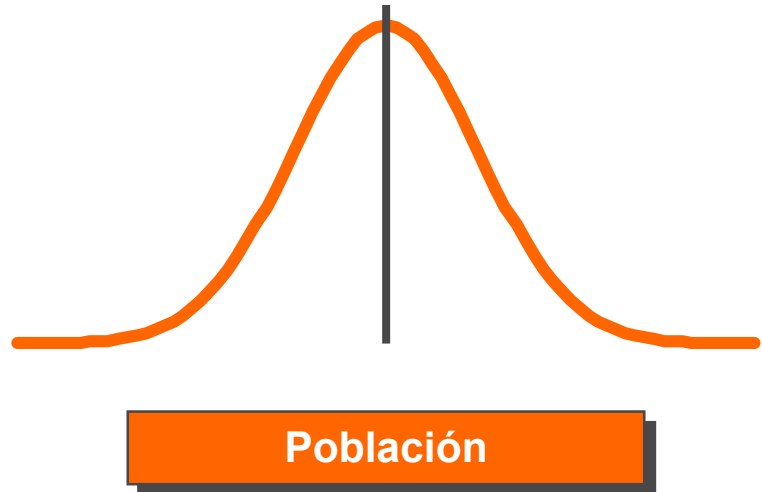
Teorema del límite central (CLT)



- Objetivo
 - Entender el significado del CLT.
 - Conocer el riesgo de las decisiones basadas en estadística
- Aplicaciones
 - Evaluación de datos.
 - Reducción de la varianza: Optimización de procesos.

¿Qué significa Teorema del límite central?

- Imagina que tienes una población.
- A esta población le tomas unas muestras y calculas la media. Si continúas tomando muestras y calculando la media puedes hacer una gráfica de la distribución de las medias.
- El CLT dice la distribución de las medias tiende a una distribución normal sin importar la distribución de la población.
- NOTA: Tomando grandes muestras **NO** significa que se tiene una distribución normal. Mucha gente se confunde con esta descripción.



Ejemplo del CLT



Worksheet 1 ***

↓	C1
1	39
2	25
3	26
4	33
5	31
6	27
7	57
8	11
9	12
10	12
11	28
12	33
13	46
14	26
15	13

1000 datos
Min 10
Max 60

Ejemplo del CLT



Worksheet 1 ***

	C1
1	39
2	25
3	26
4	33
5	31
6	27
7	57
8	11
9	12
10	12
11	28
12	33
13	46
14	26
15	13

Sample From Columns

C1

Number of rows to sample: 5

From columns:
C1

Store samples in:
C2

Repetimos 4x

Sample with replacement

Select Help OK Cancel

Ejemplo del CLT



The screenshot shows the 'Sample From Columns' dialog box in Excel. The dialog is open over a worksheet named 'Worksheet 1 ***'. The dialog has the following settings:

- Number of rows to sample: 5
- From columns: C1
- Store samples in: C2
- Sample with replacement:

Buttons: Select, Help, OK, Cancel.

A blue text overlay 'Repetimos 4x' is positioned over the 'Store samples in:' field.

The background shows two parts of the worksheet:

- Worksheet 1 *** (left):

	C1
1	39
2	25
3	26
4	33
5	31
6	27
7	57
8	11
9	12
10	12
11	28
12	33
13	46
14	26
15	13
- Worksheet 1 *** (right):

	C1	C2	C3	C4	C5	C6
1	39	53	14	46	30	54
2	25	36	18	56	33	16
3	26	24	43	45	16	23
4	33	14	30	13	23	26
5	31	51	52	18	14	50
6	27					

Ejemplo del CLT



The screenshot shows the Minitab software interface. In the background, a worksheet titled 'Worksheet 1 ***' contains a table with 6 rows and 6 columns (C1 to C6). The data values are as follows:

	C1	C2	C3	C4	C5	C6
1	39	53	14	46	30	54
2	25	36	18	56	33	16
3	26	24	43	45	16	23
4	33	14	30	13	23	26
5	31	51	52	18	14	50
6	27					

In the foreground, the 'Sample From Columns' dialog box is open. It shows 'C1' selected in the 'From columns:' list, 'Number of rows to sample:' set to 5, and 'Store samples in:' set to 'C2'. A blue text annotation 'Repetimos 4x' is overlaid on the dialog.

Below the dialog box, the 'Display Descriptive Statistics' dialog box is open. The 'Variables:' list contains 'C2-C6'. The 'By variables (optional):' list is empty. The 'replacement' checkbox is checked. The 'OK' and 'Cancel' buttons are visible.

Ejemplo del CLT



Worksheet 1 ***

↓	C1
1	39
2	25
3	26
4	33
5	31
6	27

Sample From Columns

Number of rows to sample: 5

From columns: C1

Store samples in: C2

Worksheet 1 ***

↓	C1	C2	C3	C4	C5	C6
1	39	53	14	46	30	54
2	25	36	18	56	33	16
3	26	24	43	45	16	23
4	33	14	30	13	23	26
5	31	51	52	18	14	50
6	27					

Repetimos 4x

Descriptive Statistics: C2; C3; C4; C5; C6

Variable	N	N*	Mean	SE Mean	StDev	Minimum	Q1	Median	Q3	Maximum
C2	5	0	35,60	7,55	16,89	14,00	19,00	36,00	52,00	53,00
C3	5	0	31,40	7,22	16,15	14,00	16,00	30,00	47,50	52,00
C4	5	0	35,60	8,47	18,93	13,00	15,50	45,00	51,00	56,00
C5	5	0	23,20	3,73	8,35	14,00	15,00	23,00	31,50	33,00
C6	5	0	33,80	7,63	17,06	16,00	19,50	26,00	52,00	54,00

Display Descriptive Statistics

Variables: C2-C6

By variables (optional):

Select Statistics... Graphs... Help OK Cancel

Ejemplo del CLT



Worksheet 1 ***

↓	C1
1	39
2	25
3	26
4	33
5	31
6	27

Sample From Columns

C1

Number of rows to sample: 5

From columns:
C1

Store samples in:
C2

Worksheet 1 ***

↓	C1	C2	C3	C4	C5	C6
1	39	53	14	46	30	54
2	25	36	18	56	33	16
3	26	24	43	45	16	23
4	33	14	30	13	23	26
5	31	51	52	18	14	50
6	27					

Repetimos 4x

Display Descriptive Statistics

Descriptive Statistics: C2; C3; C4; C5; C6

Variable	N	N*	Mean	SE Mean	StDev	Minimum	Q1	Median	Q3	Maximum
C2	5	0	35,60	7,55	16,89	14,00	19,00	36,00	52,00	53,00
C3	5	0	31,40	7,22	16,15	14,00	16,00	30,00	47,50	52,00
C4	5	0	35,60	8,47	18,93	13,00	15,50	45,00	51,00	56,00
C5	5	0	23,20	3,73	8,35	14,00	15,00	23,00	31,50	33,00
C6	5	0	33,80	7,63	17,06	16,00	19,50	26,00	52,00	54,00

Variables:
C2-C6

By variables (optional):

Select Statistics... Graphs... Help OK Cancel

Rango de la media = 12,4
Rango de la StDev = 10,58

Ejemplo del CLT – con más muestras (30)



Tomamos 30 muestras, 5 veces

Worksheet 1 ***		
↓	C1	C2
1	39	
2	25	
3	26	
4	33	
5	31	
6	27	
7	57	
8	11	
9	12	
10	12	
11	28	
12	33	
13	46	
14	26	
15	13	

Ejemplo del CLT – con más muestras (30)



Tomamos 30 muestras, 5 veces

Descriptive Statistics: C8; C9; C10; C11; C12

Variable	N	N*	Mean	SE Mean	StDev	Minimum	Q1	Median	Q3	Maximum
C8	30	0	36,77	2,59	14,20	13,00	24,75	36,00	51,00	59,00
C9	30	0	32,57	2,47	13,52	12,00	20,75	33,50	42,00	59,00
C10	30	0	37,37	2,87	15,70	10,00	22,75	42,50	52,00	59,00
C11	30	0	33,30	2,59	14,21	10,00	22,00	34,50	46,00	56,00
C12	30	0	39,17	2,41	13,21	13,00	29,50	39,50	49,00	59,00

Worksheet 1 ***		
↓	C1	C2
1	39	
2	25	
3	26	
4	33	
5	31	
6	27	
7	57	
8	11	
9	12	
10	12	
11	28	
12	33	
13	46	
14	26	
15	13	

Ejemplo del CLT – con más muestras (30)



Tomamos 30 muestras, 5 veces

Descriptive Statistics: C8; C9; C10; C11; C12

Variable	N	N*	Mean	SE Mean	StDev	Minimum	Q1	Median	Q3	Maximum
C8	30	0	36,77	2,59	14,20	13,00	24,75	36,00	51,00	59,00
C9	30	0	32,57	2,47	13,52	12,00	20,75	33,50	42,00	59,00
C10	30	0	37,37	2,87	15,70	10,00	22,75	42,50	52,00	59,00
C11	30	0	33,30	2,59	14,21	10,00	22,00	34,50	46,00	56,00
C12	30	0	39,17	2,41	13,21	13,00	29,50	39,50	49,00	59,00

Rango de la media = 6,6
Rango de la StDev = 2,49

Rango de la media = 12,4
Rango de la StDev = 10,58

↓	C1	C2
1	39	
2	25	
3	26	
4	33	
5	31	
6	27	
7	57	
8	11	
9	12	
10	12	
11	28	
12	33	
13	46	
14	26	
15	13	

Ejemplo del CLT – con más muestras (30)



Tomamos 30 muestras, 5 veces

Descriptive Statistics: C8; C9; C10; C11; C12

Variable	N	N*	Mean	SE Mean	StDev	Minimum	Q1	Median	Q3	Maximum
C8	30	0	36,77	2,59	14,20	13,00	24,75	36,00	51,00	59,00
C9	30	0	32,57	2,47	13,52	12,00	20,75	33,50	42,00	59,00
C10	30	0	37,37	2,87	15,70	10,00	22,75	42,50	52,00	59,00
C11	30	0	33,30	2,59	14,21	10,00	22,00	34,50	46,00	56,00
C12	30	0	39,17	2,41	13,21	13,00	29,50	39,50	49,00	59,00

Rango de la media = 6,6
Rango de la StDev = 2,49

Rango de la media = 12,4
Rango de la StDev = 10,58

Mientras más muestras se tengan, menor es el error estándar.

$$S_{\bar{x}} = \frac{S_x}{\sqrt{n}}$$

Ejemplo del CLT – calculando la media (5)



Worksheet 1 ***

↓	C1	C2	C3	C4	C5	C6
1	39	53	14	46	30	54
2	25	36	18	56	33	16
3	26	24	43	45	16	23
4	33	14	30	13	23	26
5	31	51	52	18	14	50
6	27					

Ejemplo del CLT – calculando la media (5)



Worksheet 1 ***						
↓	C1	C2	C3	C4	C5	C6
1	39	53	14	46	30	54
2	25	36	18	56	33	16
3	26	24	43	45	16	23
4	33	14	30	13	23	26
5	31	51	52	18	14	50
6	27					

C26
Media5
39,4
31,8
30,2
21,2
37,0

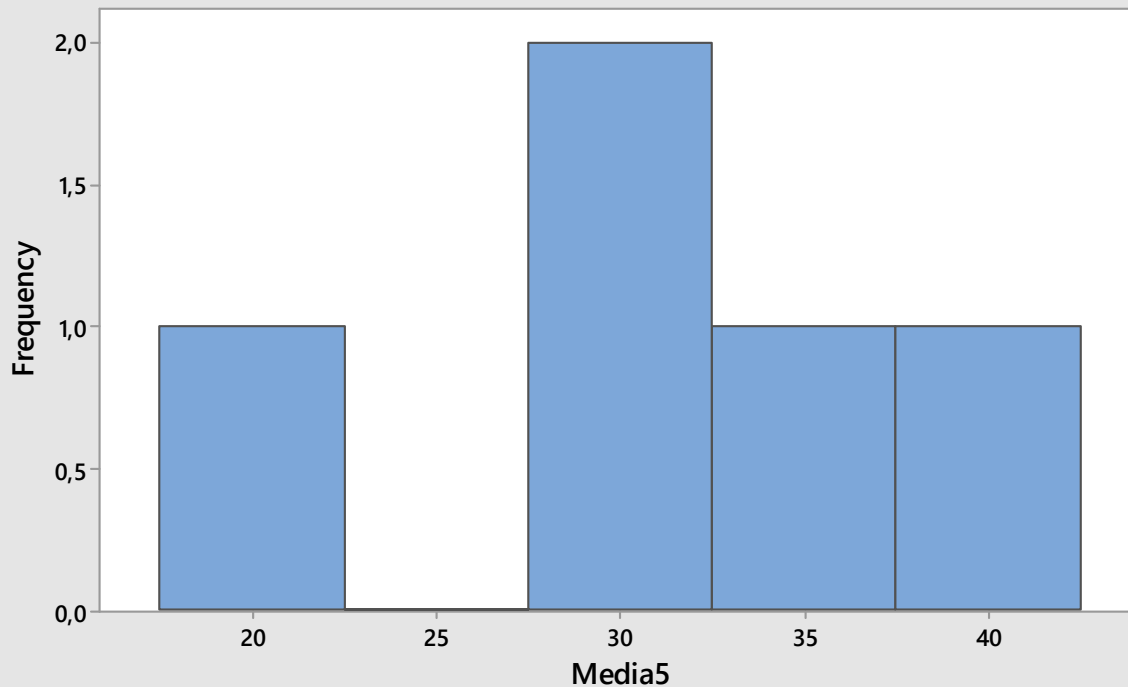
Ejemplo del CLT – calculando la media (5)



Worksheet 1 ***						
↓	C1	C2	C3	C4	C5	C6
1	39	53	14	46	30	54
2	25	36	18	56	33	16
3	26	24	43	45	16	23

C26
Media5
39,4
31,8
30,2
21,2

Histogram of Media5



Ejemplo del CLT – calculando la media (30)



Worksheet 1 ***

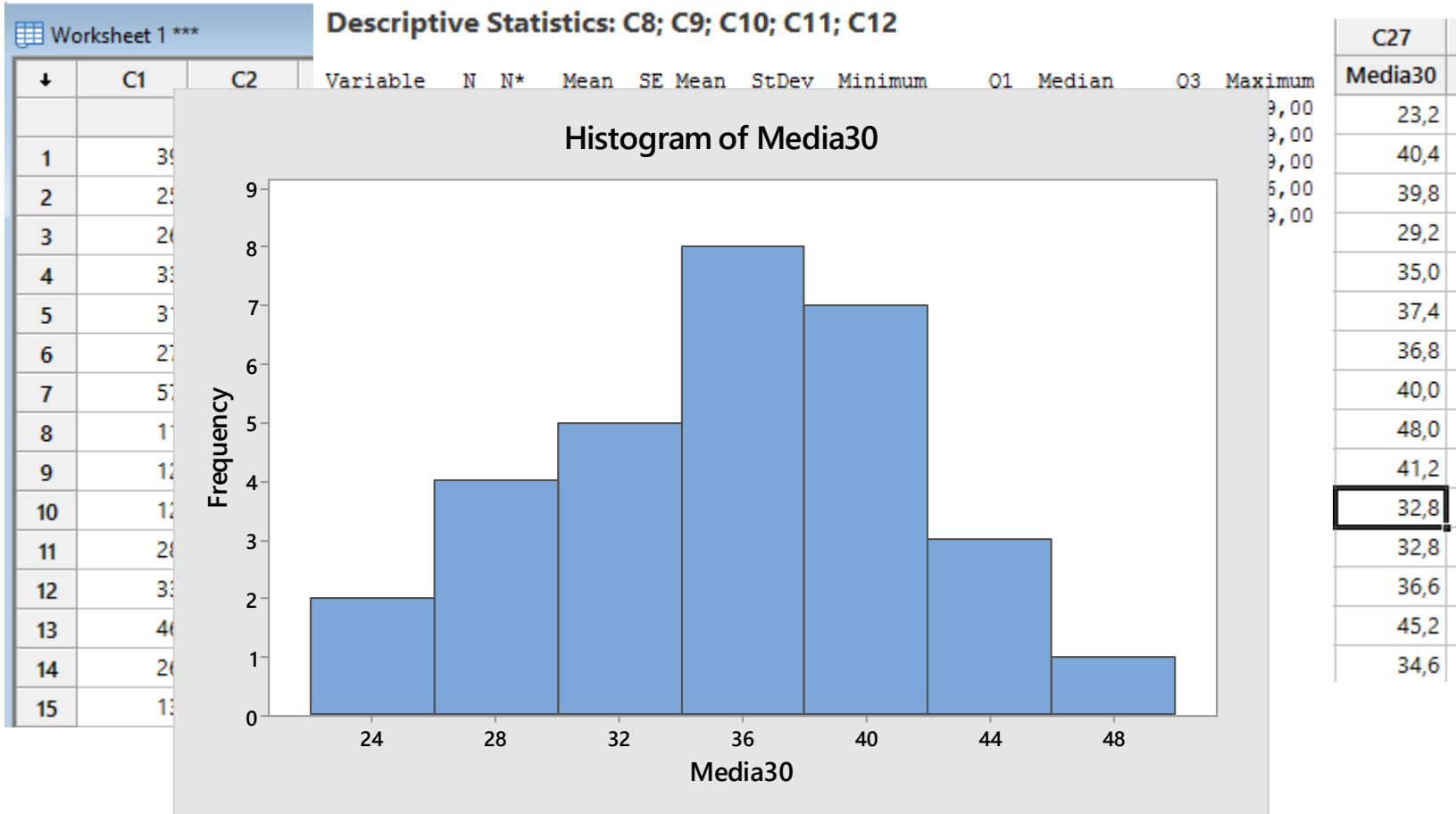
↓	C1	C2
1	39	
2	25	
3	26	
4	33	
5	31	
6	27	
7	57	
8	11	
9	12	
10	12	
11	28	
12	33	
13	46	
14	26	
15	13	

Ejemplo del CLT – calculando la media (30)

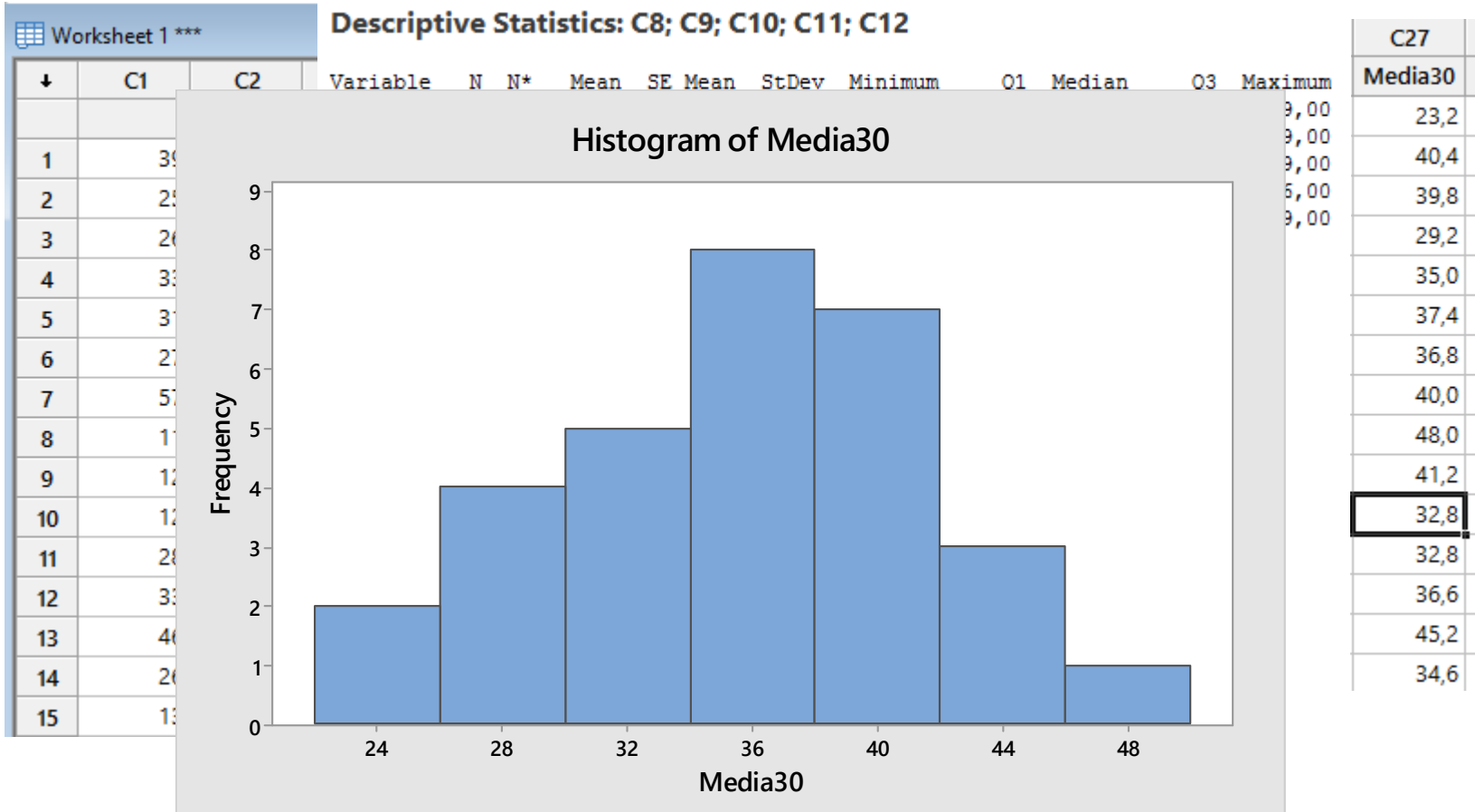


Worksheet 1 ***			Descriptive Statistics: C8; C9; C10; C11; C12											C27
↓	C1	C2	Variable	N	N*	Mean	SE Mean	StDev	Minimum	Q1	Median	Q3	Maximum	Media30
			C8	30	0	36,77	2,59	14,20	13,00	24,75	36,00	51,00	59,00	23,2
1	39		C9	30	0	32,57	2,47	13,52	12,00	20,75	33,50	42,00	59,00	40,4
2	25		C10	30	0	37,37	2,87	15,70	10,00	22,75	42,50	52,00	59,00	39,8
3	26		C11	30	0	33,30	2,59	14,21	10,00	22,00	34,50	46,00	56,00	29,2
4	33		C12	30	0	39,17	2,41	13,21	13,00	29,50	39,50	49,00	59,00	35,0
5	31													37,4
6	27													36,8
7	57													40,0
8	11													48,0
9	12													41,2
10	12													32,8
11	28													32,8
12	33													36,6
13	46													45,2
14	26													34,6
15	13													

Ejemplo del CLT – calculando la media (30)



Ejemplo del CLT – calculando la media (30)



Las medias de las muestras tienden a una distribución normal

- Mientras más muestras, menor es la variación. Se recomienda tener 30 muestras para datos variables y 50 para datos atributo.
- El CLT ayuda a entender el riesgo/error que se tiene en una muestra:
 - 95% de los datos normalmente distribuidos están a +/- 2 desviaciones estándar de la media. Por tal motivo, se tiene una probabilidad del 95% que la media de la muestra esté a 2 errores estándar de la media de la población.

$$\bar{x} \pm 2 * \frac{s}{\sqrt{n}}$$

Ejemplo de aplicación



- Se tiene una muestra de 20 piezas. El tiempo de fabricación promedio es de 45 minutos (media) con una StDev de 10 minutos
- Los 45 minutos es una media estimada del valor real de la población.
- Usando el CLT hay un 95% de probabilidad de que la media real se encuentre entre 40,5 y 49,5 minutos. Este rango se le conoce como intervalo de confianza.

$$\bar{x} \pm 2 * \frac{s}{\sqrt{n}} \quad 45 + 2 * \frac{10}{\sqrt{20}} = 49,5 \quad 45 - 2 * \frac{10}{\sqrt{20}} = 40,5$$

- Si hay una propuesta de un cambio en el proceso en reducir la fabricación en 4 minutos, ¿Lo harían?
- Respuesta en la descripción del video.