

# Statistiques à 2 variables ... et calculatrice graphique



## Bienvenue !

### Congrès SBPMef 2015

Equipe **CASIO** *Education*

Françoise Delpérée – [fdelperee@hotmail.com](mailto:fdelperee@hotmail.com)  
Virginie Loward – [virginie.loward@skynet.be](mailto:virginie.loward@skynet.be)

# Statistiques à 2 variables & calculatrice graphique



15:30 Introduction



15:35 Calcul et affichage de la droite de régression

15:45 Construction intuitive de la droite

16:15 Construction intuitive de la mesure de la qualité de l'ajustement

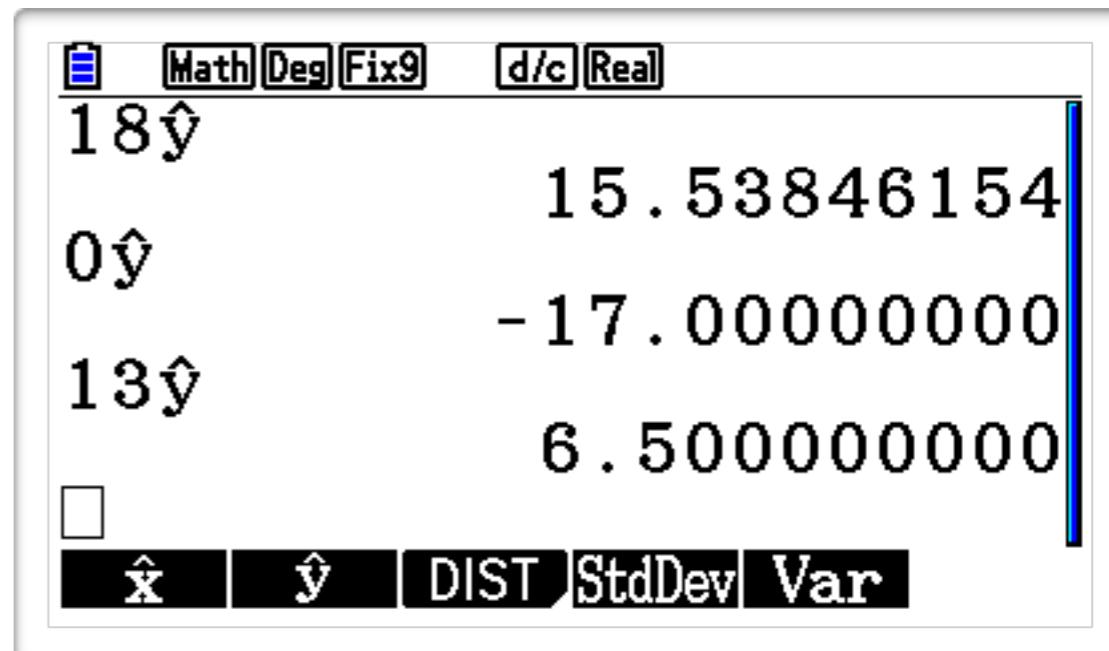
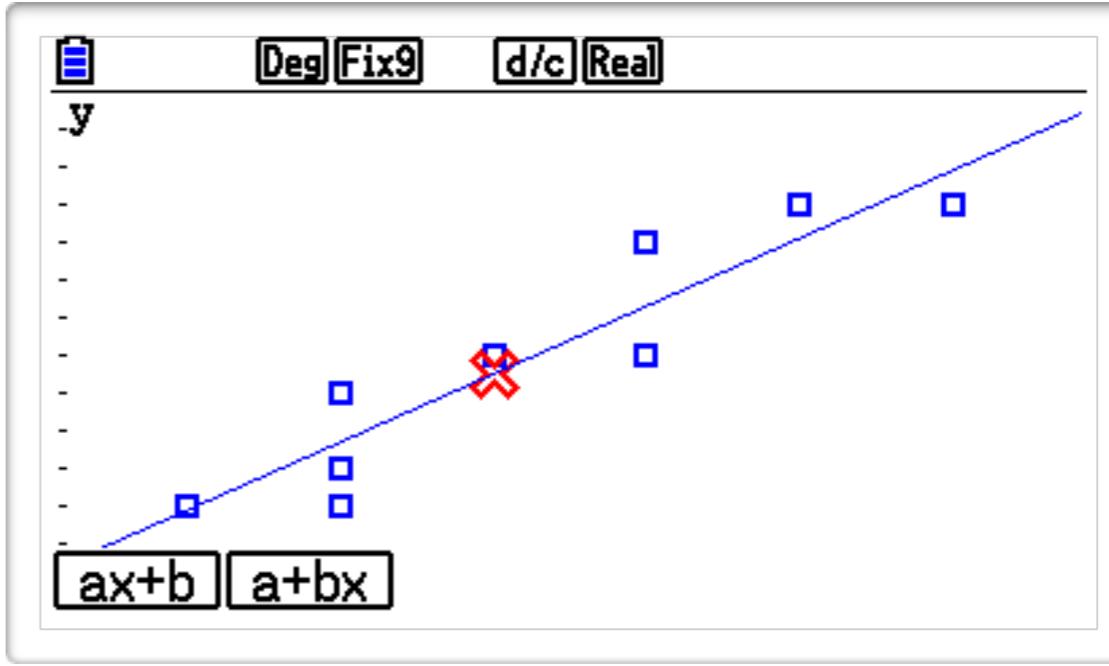
16:30 Réponses aux questions

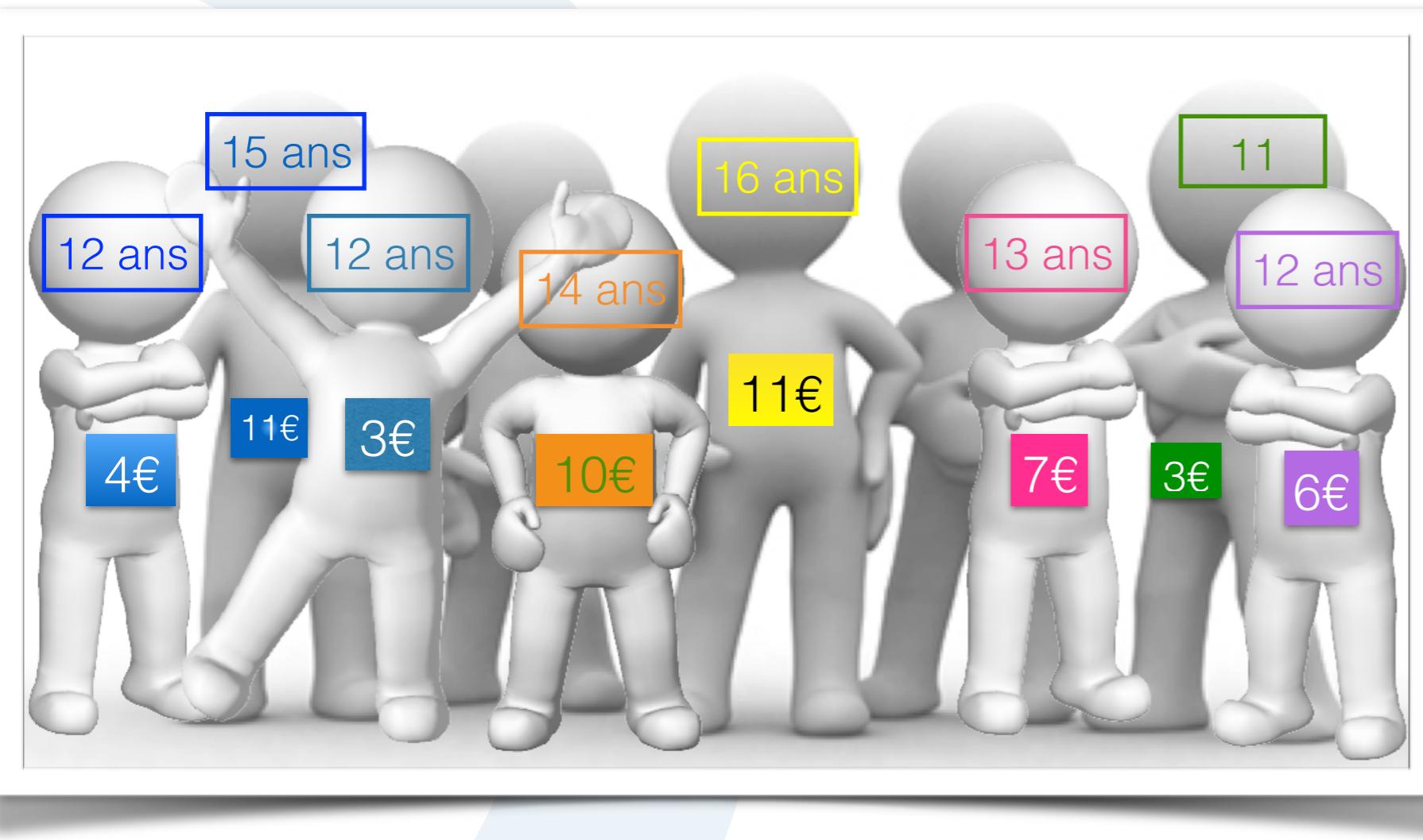
# CALCUL ET AFFICHAGE DE LA DROITE



# CALCUL ET AFFICHAGE DE LA DROITE

- Superposition des nuage de points, point moyen et droite de régression;
- Prédiction d'une ordonnée théorique pour une abscisse donnée.





Sur une année, on a noté, pour 10 adolescents, leur âge et leur montant hebdomadaire moyen d'argent de poche, exprimé en euros.

Comment afficher le nuage de point, l'élève moyen et la droite de régression qui prédit l'argent de poche que recevrait un adolescent ?

Combien recevrait théoriquement un ado de 13 ans, un jeune de 18 ans et un nouveau-né ?

Exercice inspiré de *Eléments de Statistique*.

*DROESBEKE, Editions de l'Université de Bruxelles, Bruxelles 1992.*



Allumer la calculatrice

Appuyer sur **AC/ON**



Allumer la calculatrice

Appuyer sur **AC/ON**

Démarrer le menu STAT

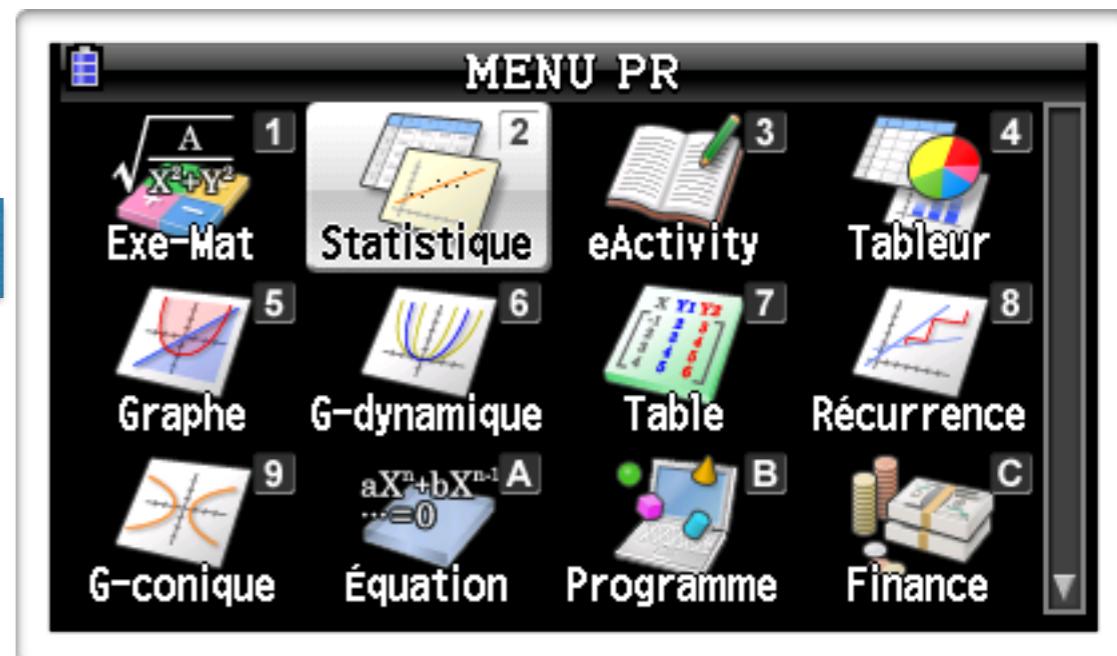
Appuyer sur **MENU**

Se déplacer **▼**

& Valider avec **EXE**

ou

Appuyer sur **2**





Allumer la calculatrice

Appuyer sur **AC/ON**

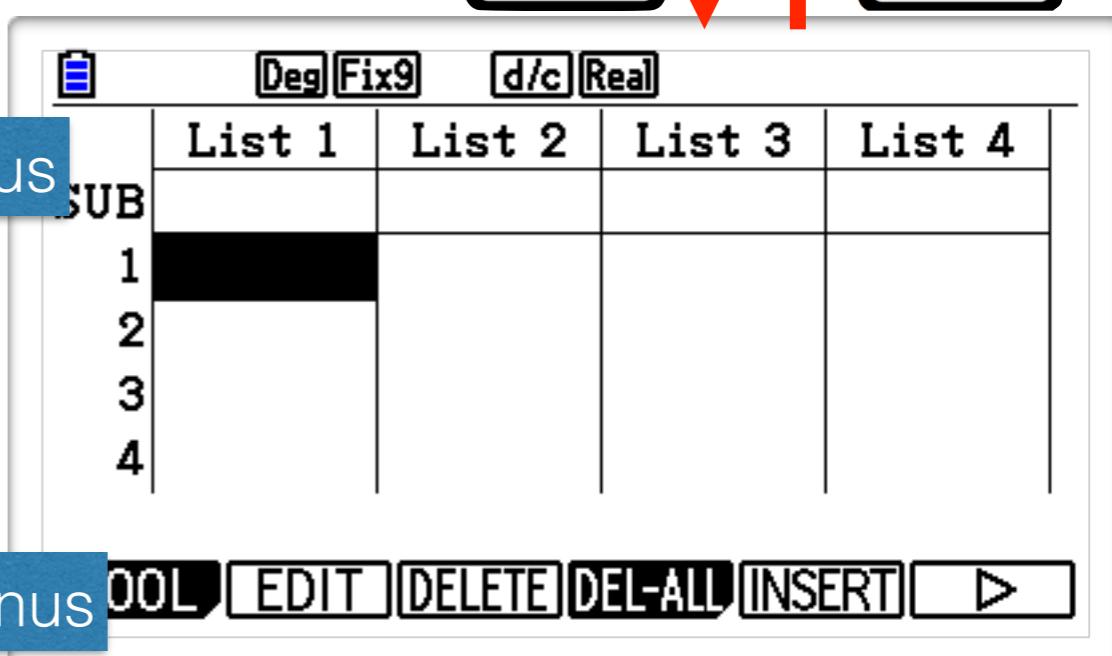
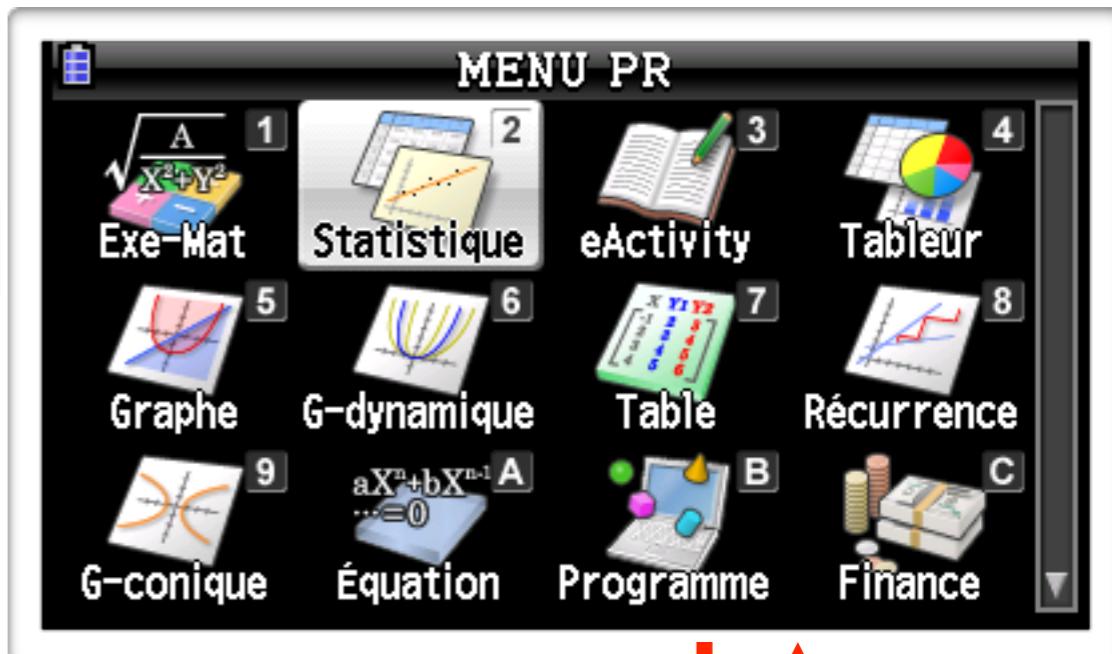
Démarrer le menu STAT

Appuyer sur **MENU**

Se déplacer & Valider avec **EXE**

ou

Appuyer sur **2**



Remonter dans les menus

Appuyer sur **EXIT**

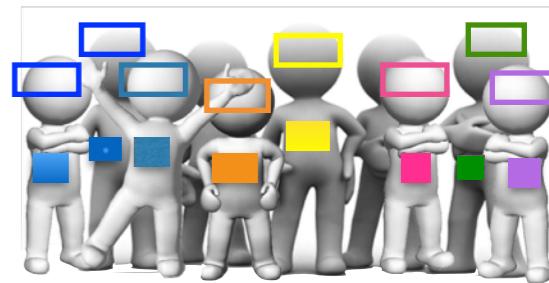
Descendre dans les menus



	List 1	List 2	List 3	List 4
SUB	ARGENT	AGE		
1		4	12	
2		3	12	
3		11	15	
4		10	14	

TOOL EDIT DELETE DEL-ALL INSERT >

GRAPH CALC TEST INTR DIST >



$i$	$x_i$	$y_i$
1	12	4
2	12	3
3	15	11
4	14	10
5	16	11
6	14	7
7	12	6
8	13	7
9	11	3
10	11	3

## Généraux

SHIFT MENU

Stat Wind	:	Auto
Resid List	:	List7
List File	:	File1
Sub Name	:	On
Frac Result	:	d/c
Func Type	:	Y=
Graph Func	:	On
Auto		Manual

F1

EXIT

		Deg	Fix9	d/c	Real
SUB	AGE	List 1	List 2	List 3	List 4
1	12		4		
2	12		3		
3	15		11		
4	14		10		

GRAPH1 GRAPH2 GRAPH3 SELECT

SET

## Particuliers

F6

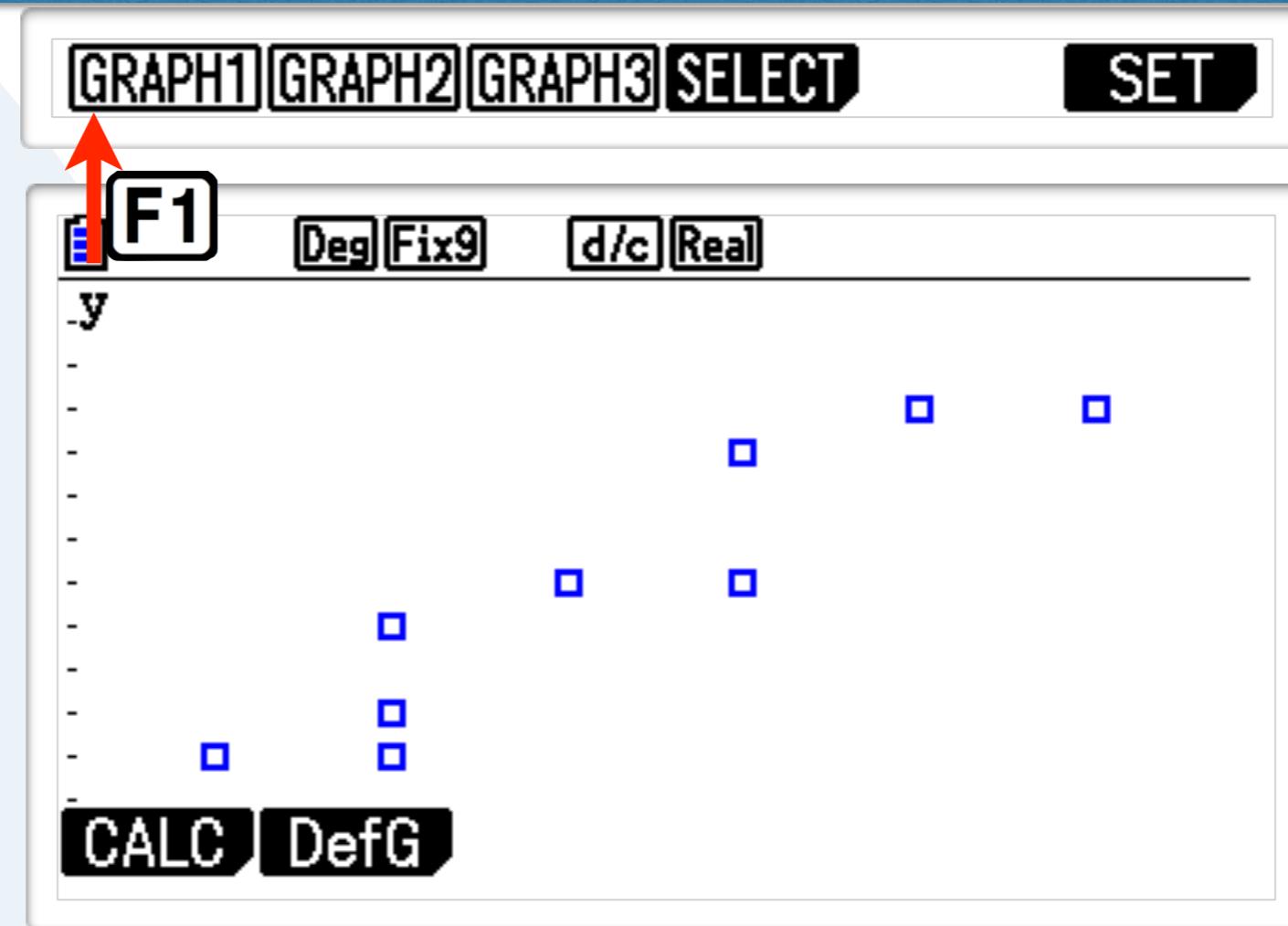
		Deg	Fix9	d/c	Real
StatGraph1					
Graph Type	:	Scatter			
XList	:	List1			
YList	:	List2			
Frequency	:	1			
Mark Type	:	□			
Color Link	:	Off			

Scatter xyLine NPPlot Pie

▶

EXIT

LIST

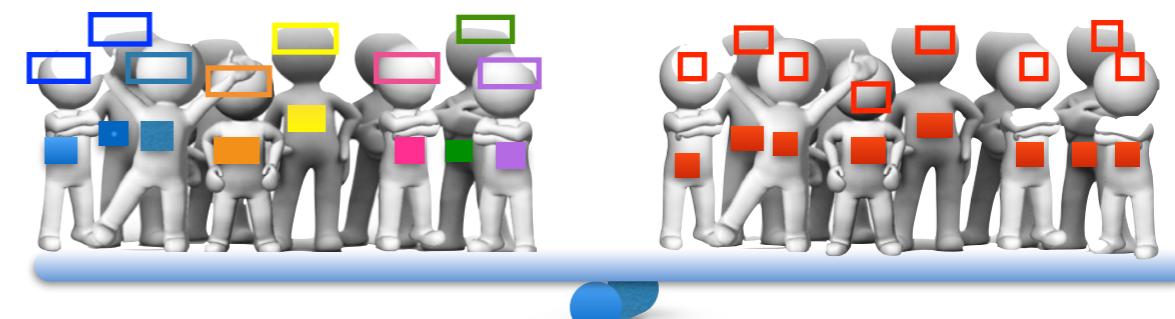
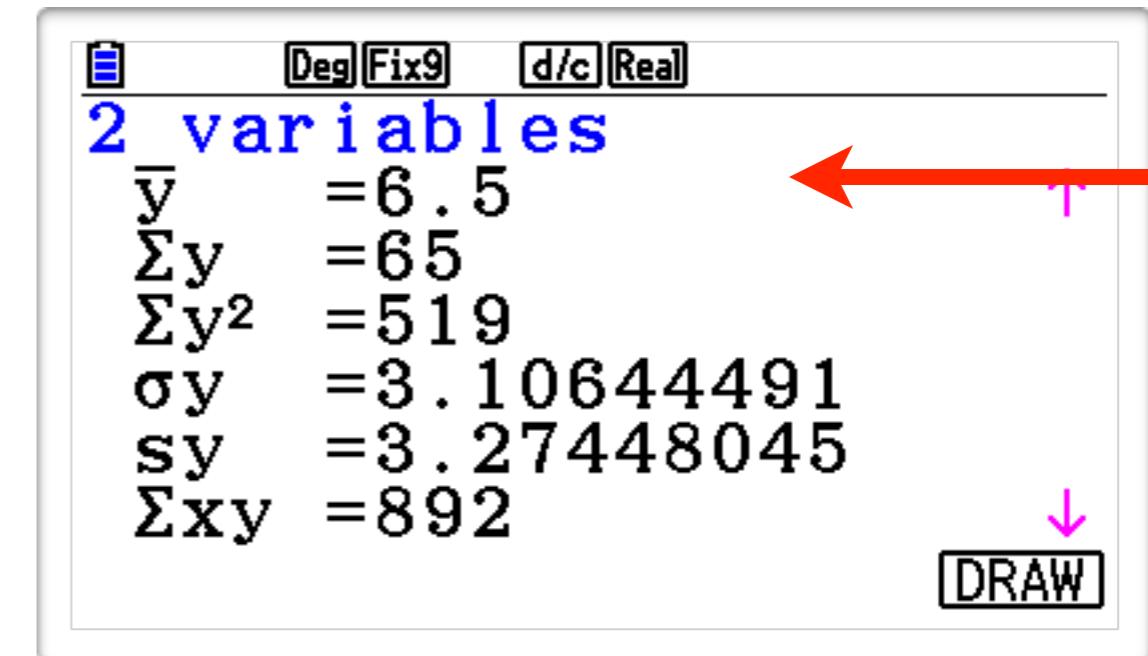
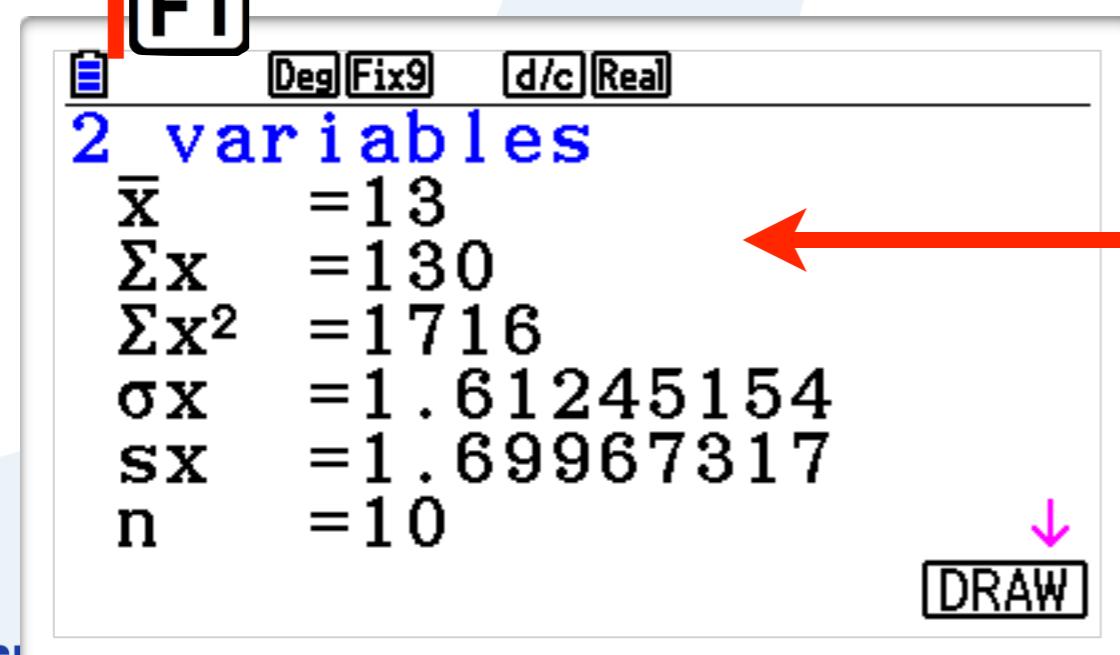
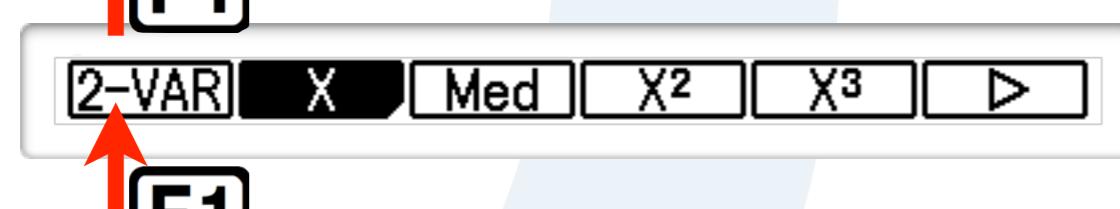
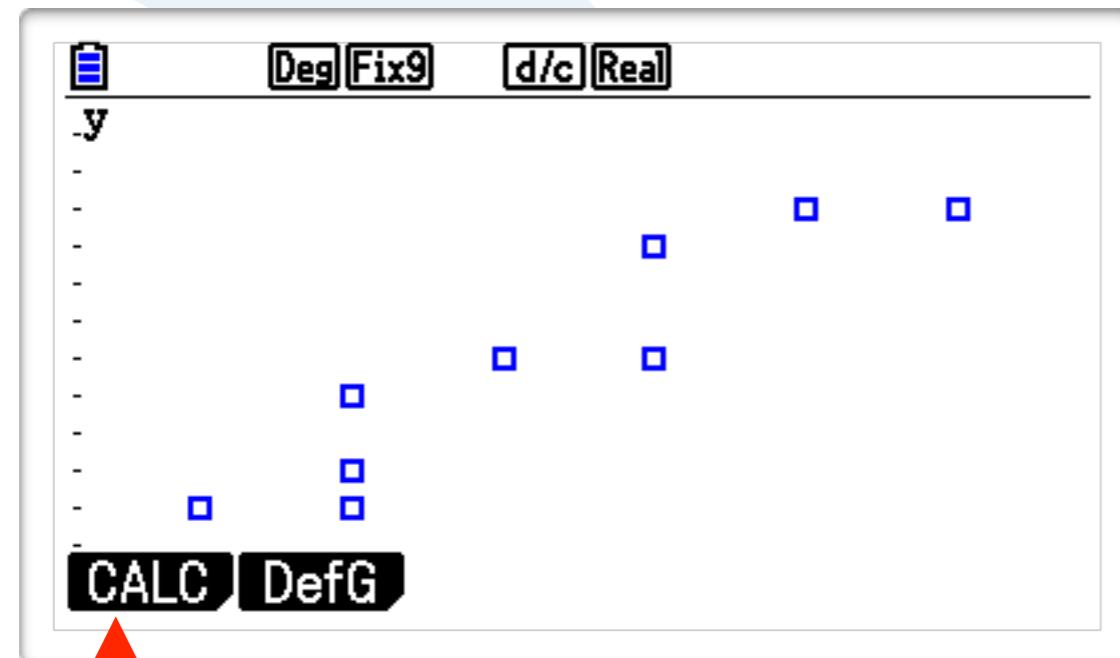


Le nuage est concentré.

Le nuage semble avoir une orientation globale.

Les adolescents les plus âgés reçoivent davantage d'argent que leurs cadets.

Une droite qui traduirait cette orientation serait de pente positive.



« L'élève moyen a 13 ans et reçoit 6,5 € ».





Calculator screen showing a table of data:

	List 1	List 2	List 3	List 4
SUB	AGE	ARGENT		
1	12	4	13	6.5
2	12	3		
3	15	11		
4	14	10		

Below the table are buttons: GRAPH1, GRAPH2, GRAPH3, SELECT, and SET. A red arrow points from the '13' in the table to the '13' in the 'List 3' column.

Calculator screen showing a scatter plot configuration:

```

StatGraph1
Graph Type : Scatter
XList : List1
YList : List2
Frequency : 1
Mark Type : □
Color Link : Off

```

Below the configuration are buttons: GRAPH1, GRAPH2, GRAPH3. A pink arrow points from the '13' in the table to the '13' in the 'List 3' column.

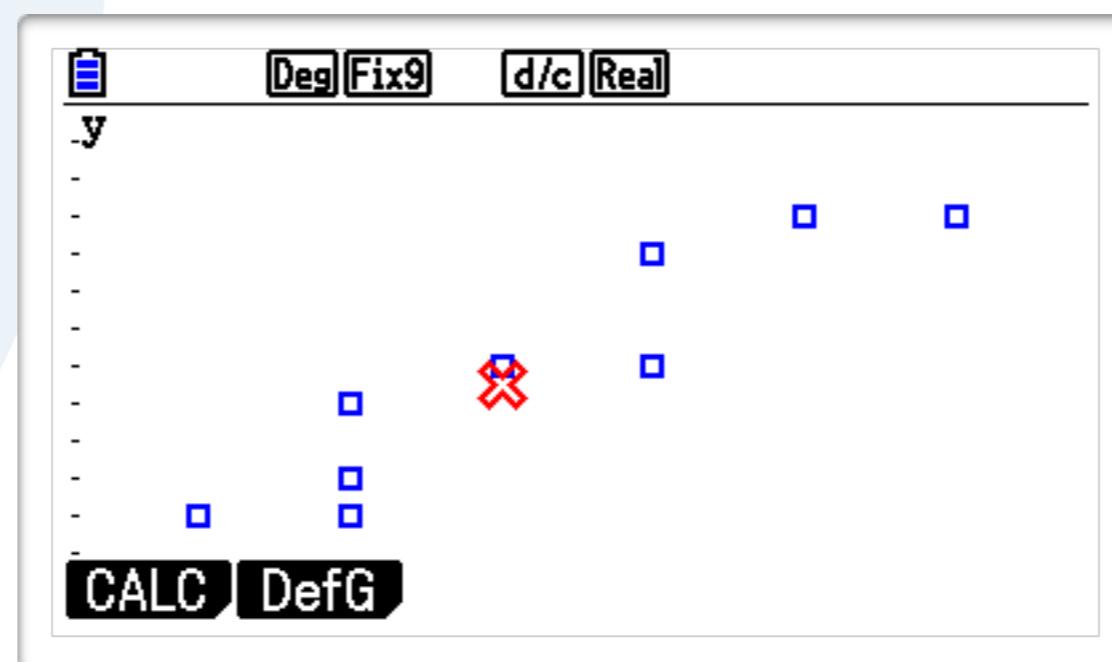
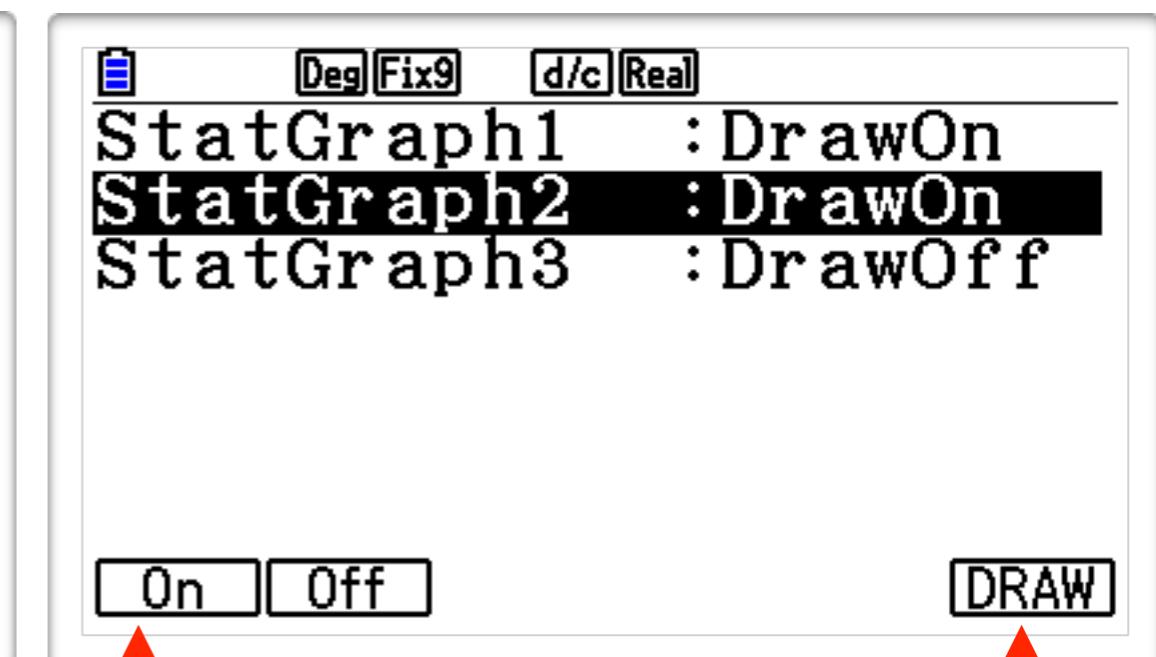
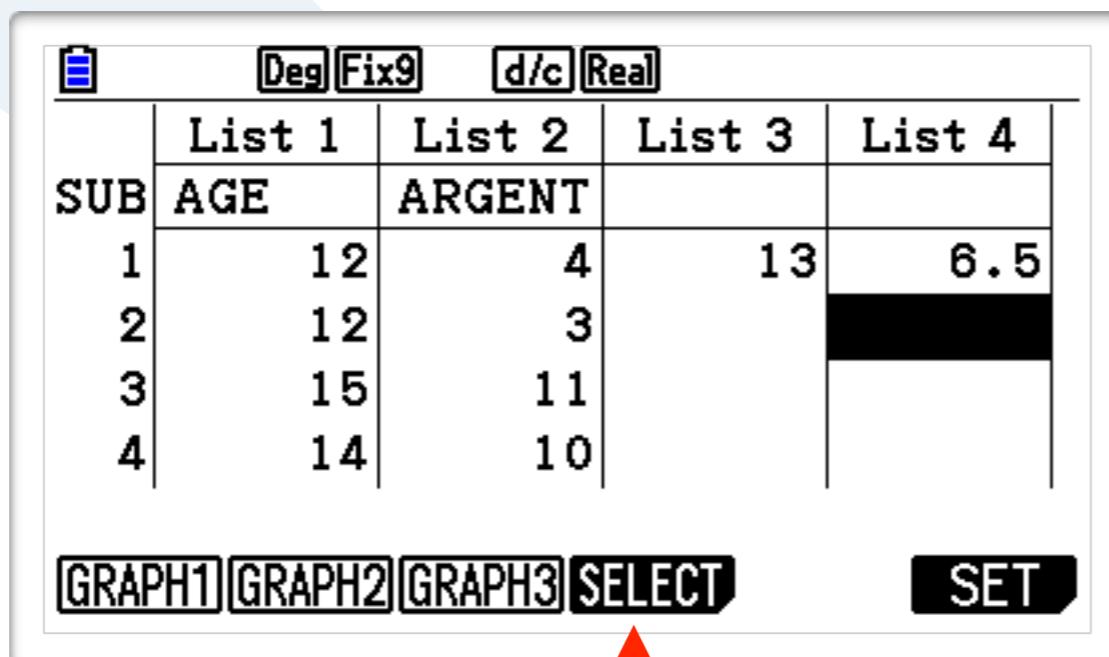
Calculator screen showing a scatter plot configuration:

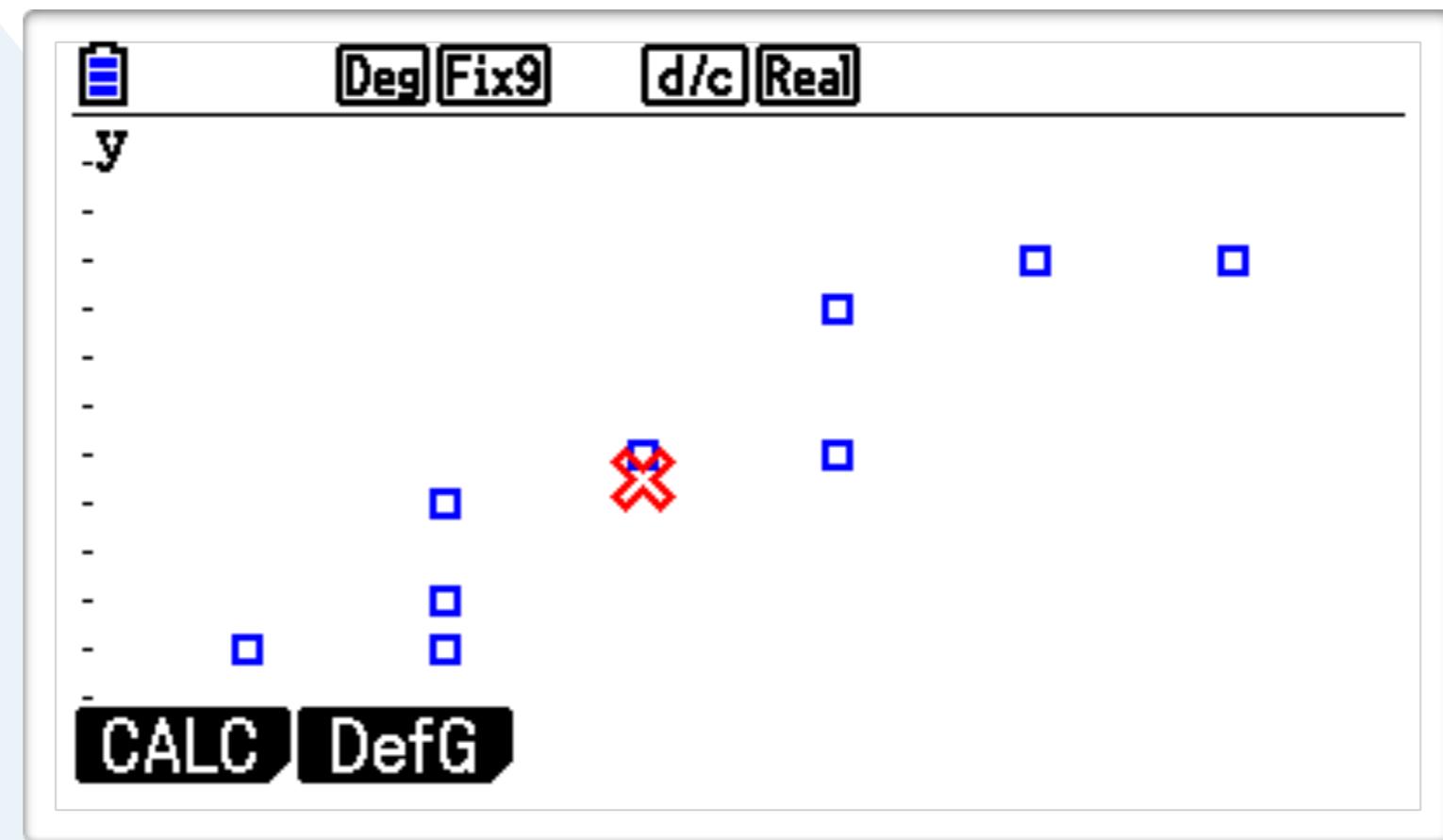
```

StatGraph2
Graph Type : Scatter
XList : List3
YList : List4
Frequency : 1
Mark Type : ∗
Color Link : Off

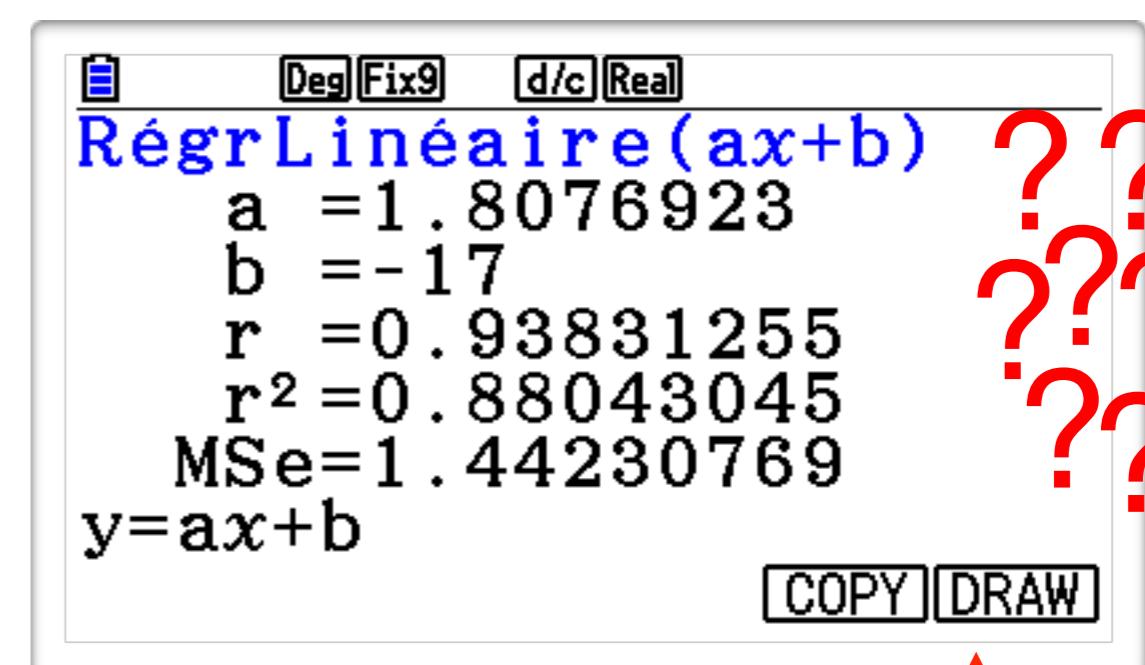
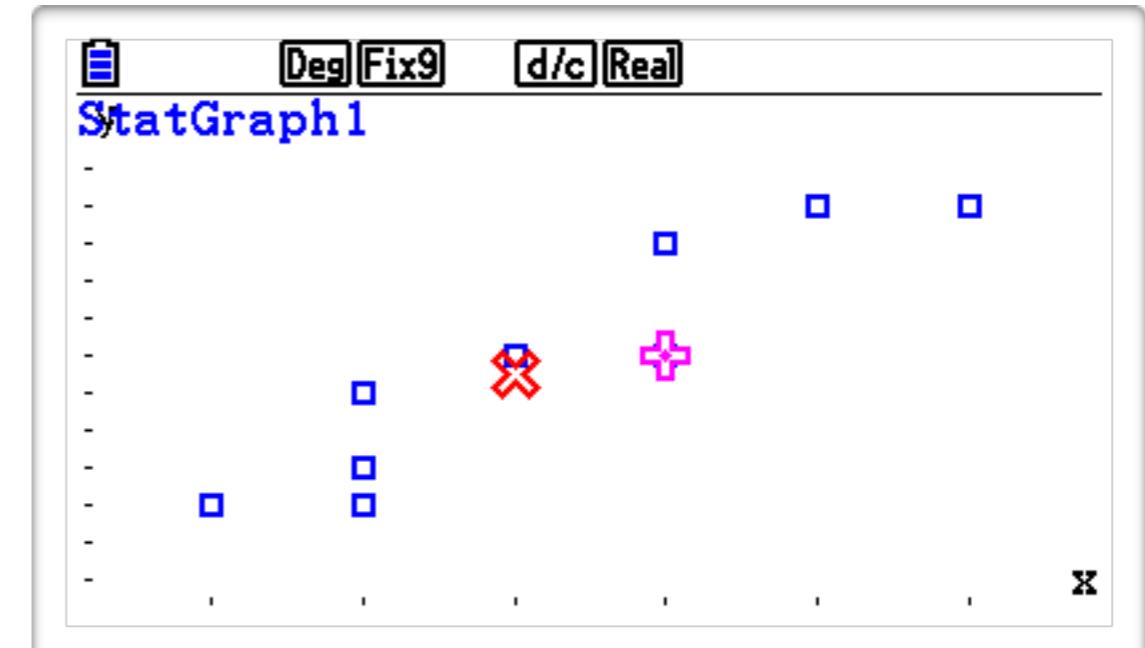
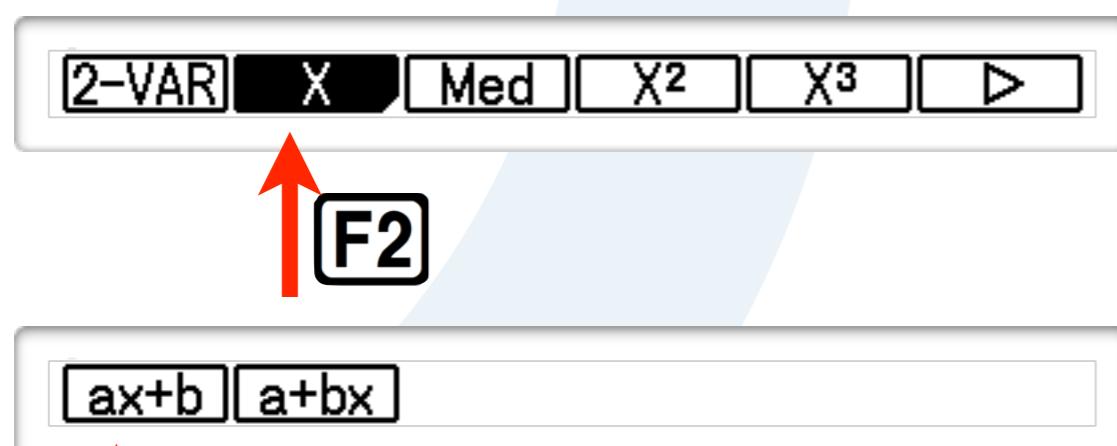
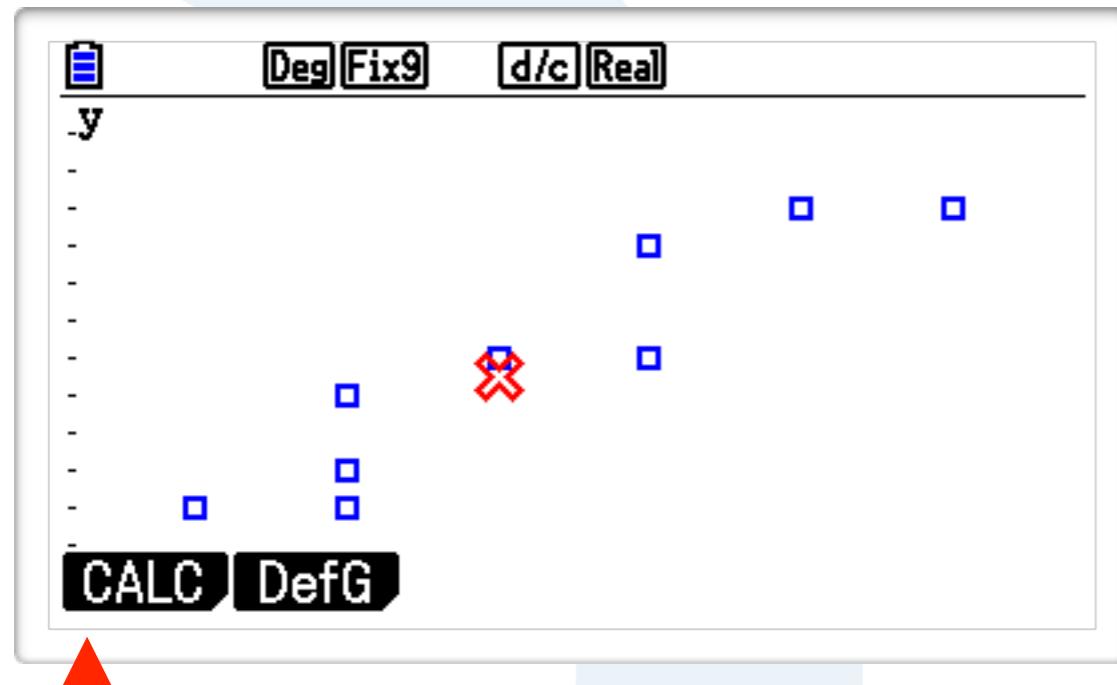
```

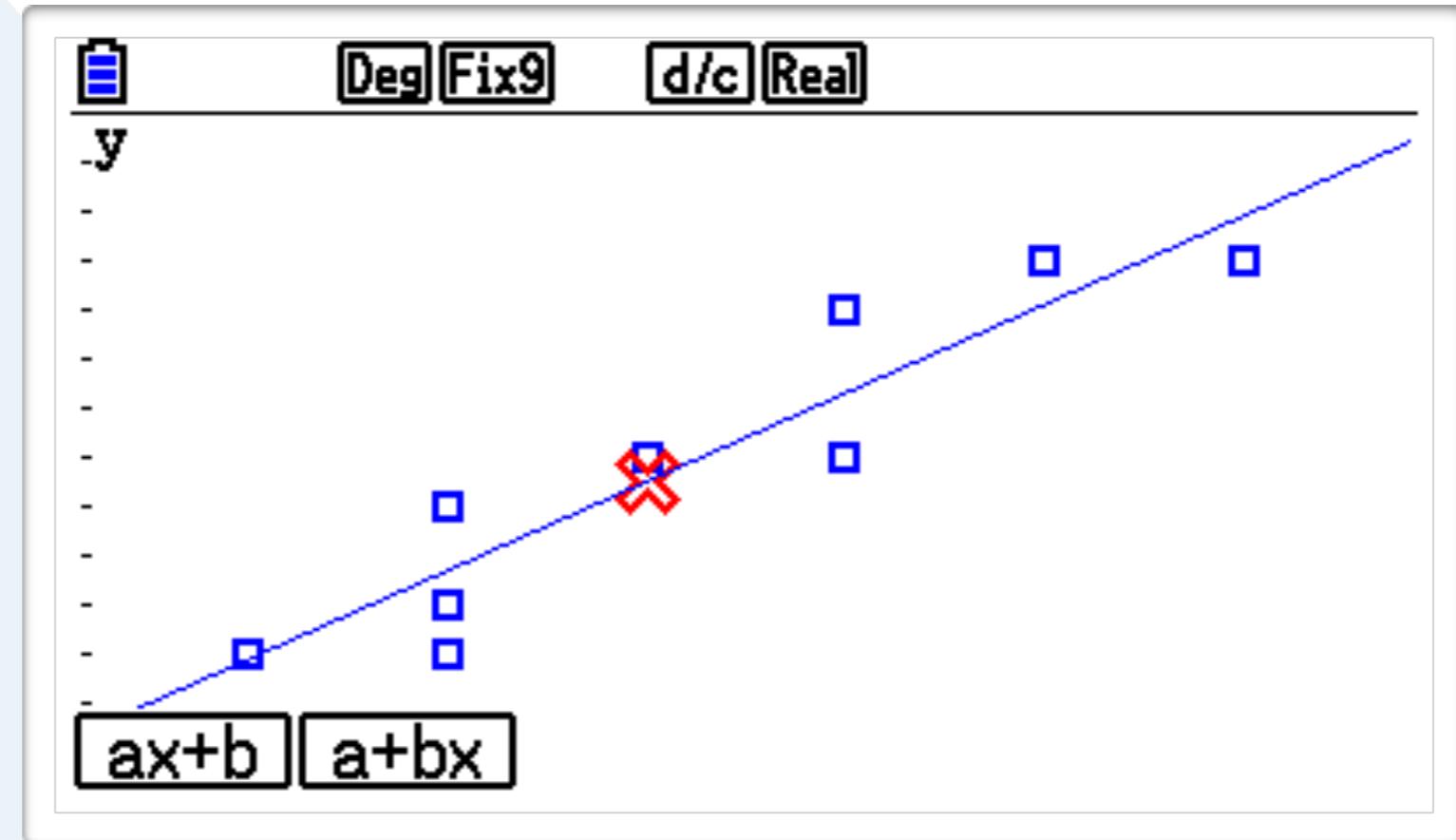
Below the configuration are buttons: X&Y, OnlyX, OnlyY, Off. Red arrows point from the '13' in the table to the '13' in the 'List 3' column and the '6.5' in the 'List 4' column. A green arrow points from the '13' in the table to the '13' in the 'List 3' column.





Le point moyen se trouve au milieu du nuage de points.  
L' « élève moyen », qui a 13 ans et reçoit 6,5 € d'argent de poche,  
ne correspond à aucun élève en particulier.





Le point moyen se trouve sur la droite de régression.

La droite traduit bien l'orientation du nuage.

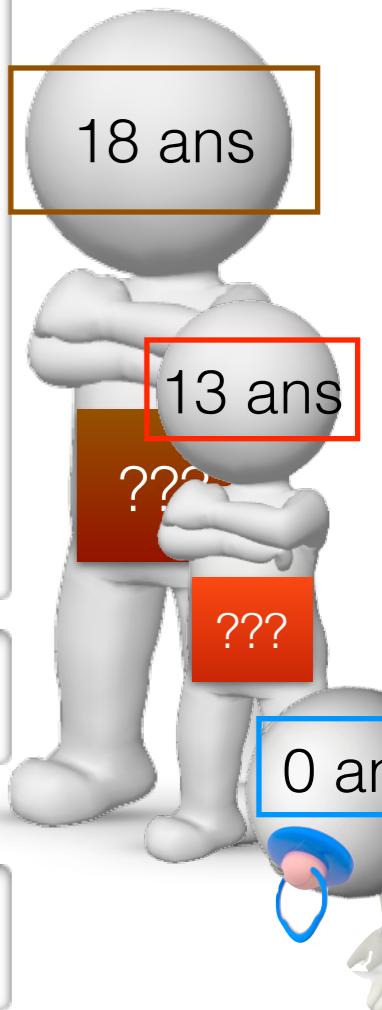
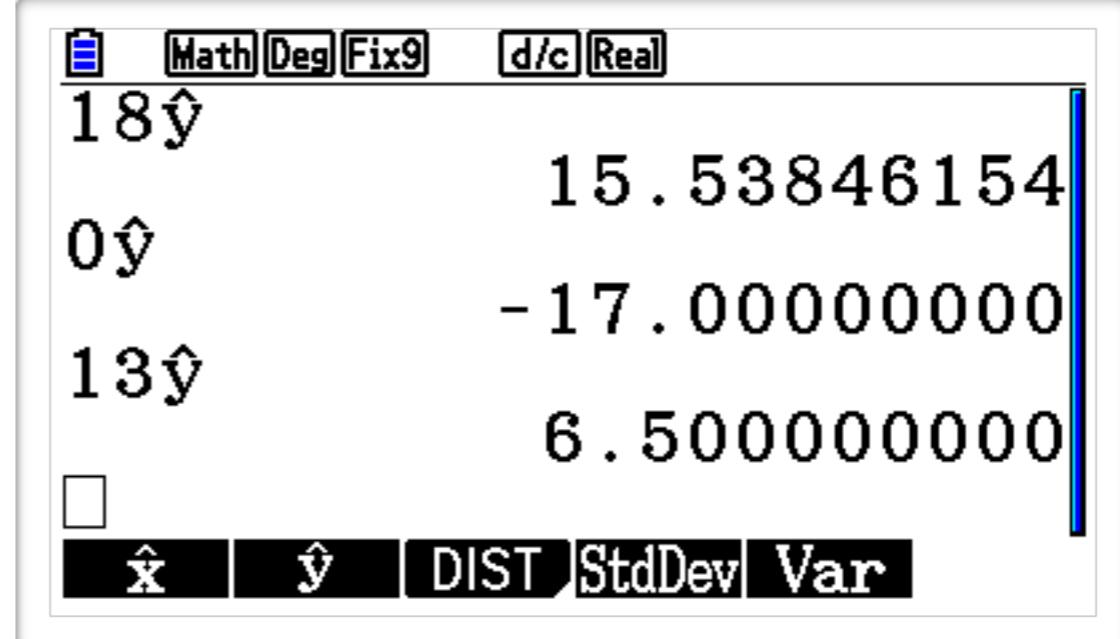
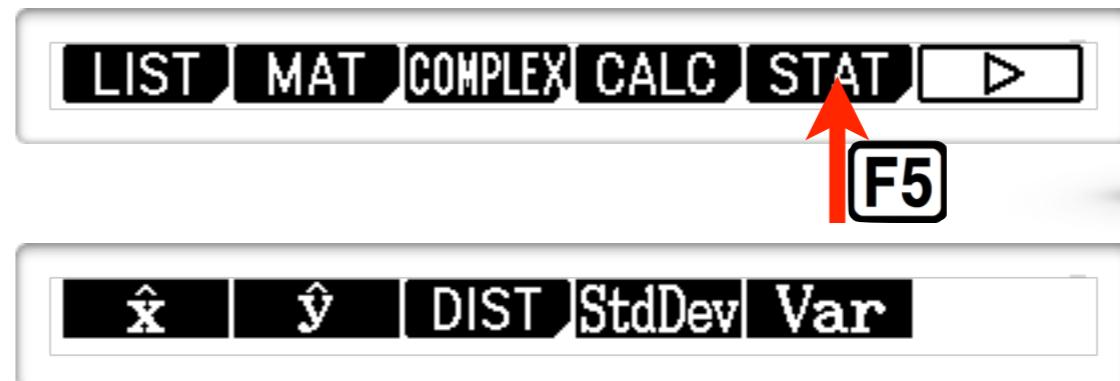
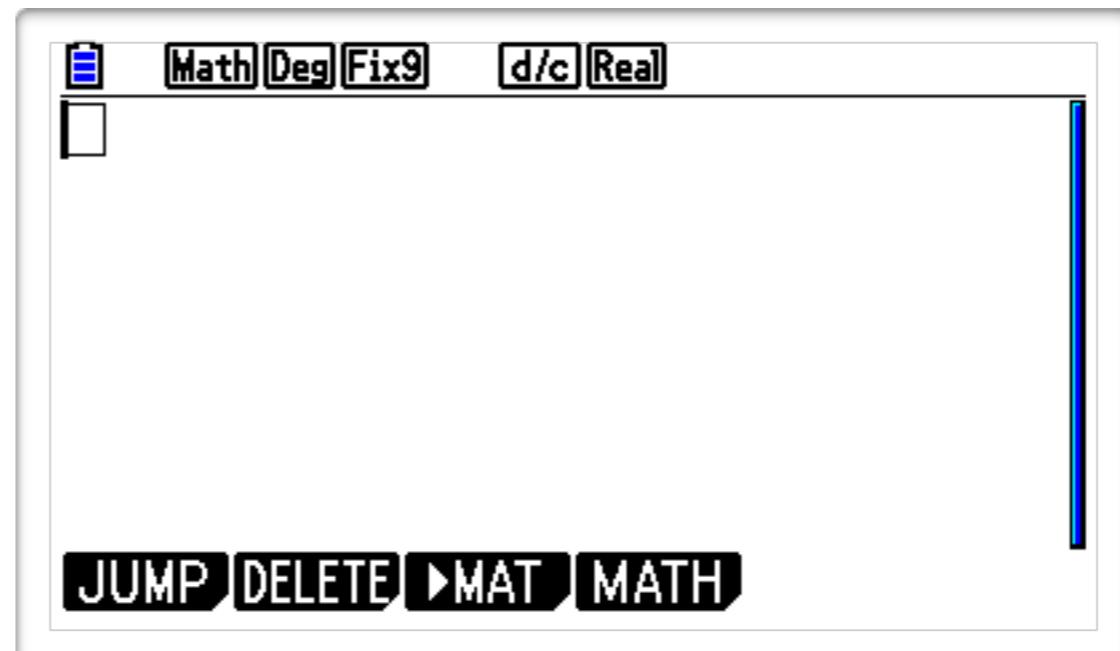
La droite permet de prédire avec plus ou moins de précision  
l'argent de poche pour un âge donné.

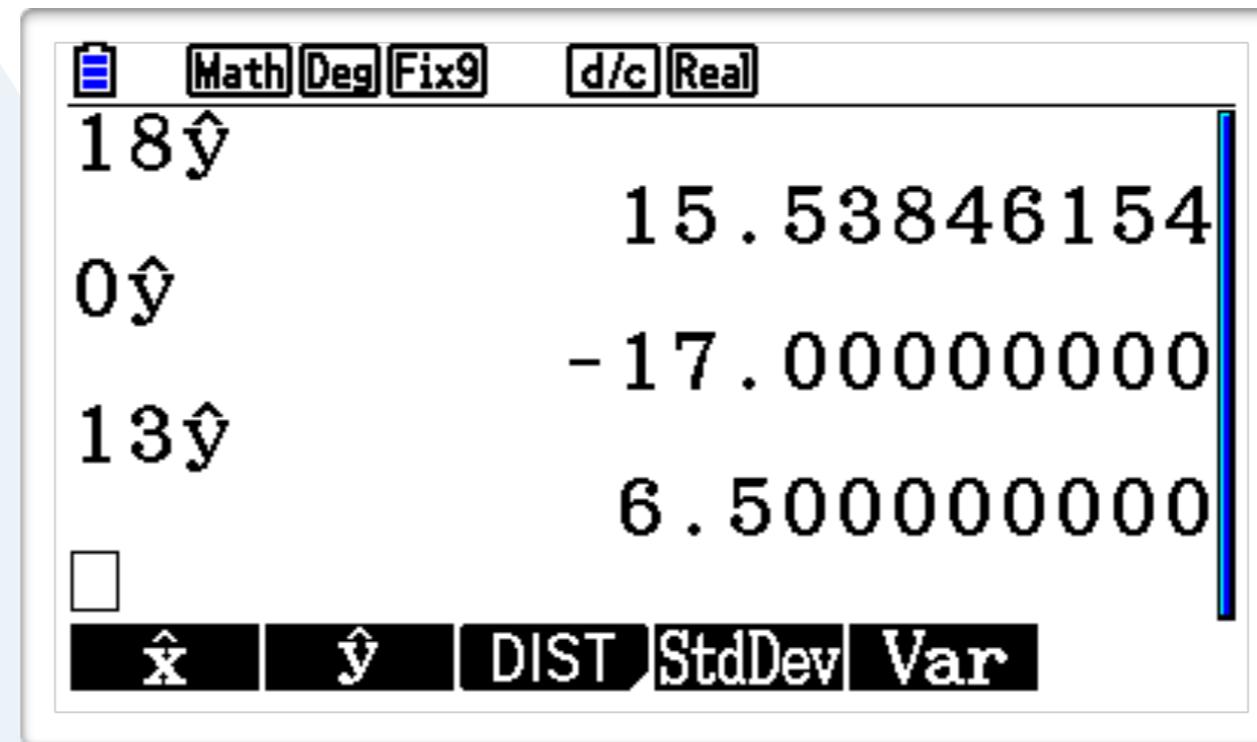


MENU

1

OPTN





Un jeune de 18 ans recevrait 15,5 €.

Un nouveau-né devrait 17 € à ses parents !!!

Le résultat prédit n'est raisonnable qu'aux alentours du domaine de variation (ici entre 11 et 16 ans).

Le point moyen appartient à la droite de régression.



# CONSTRUCTION INTUITIVE DE LA DROITE

# CONSTRUCTION INTUITIVE DE LA DROITE

- Equation

$$y = \bar{y} + a(x - \bar{x})$$

- Covariance

$$s_{xy} = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})$$

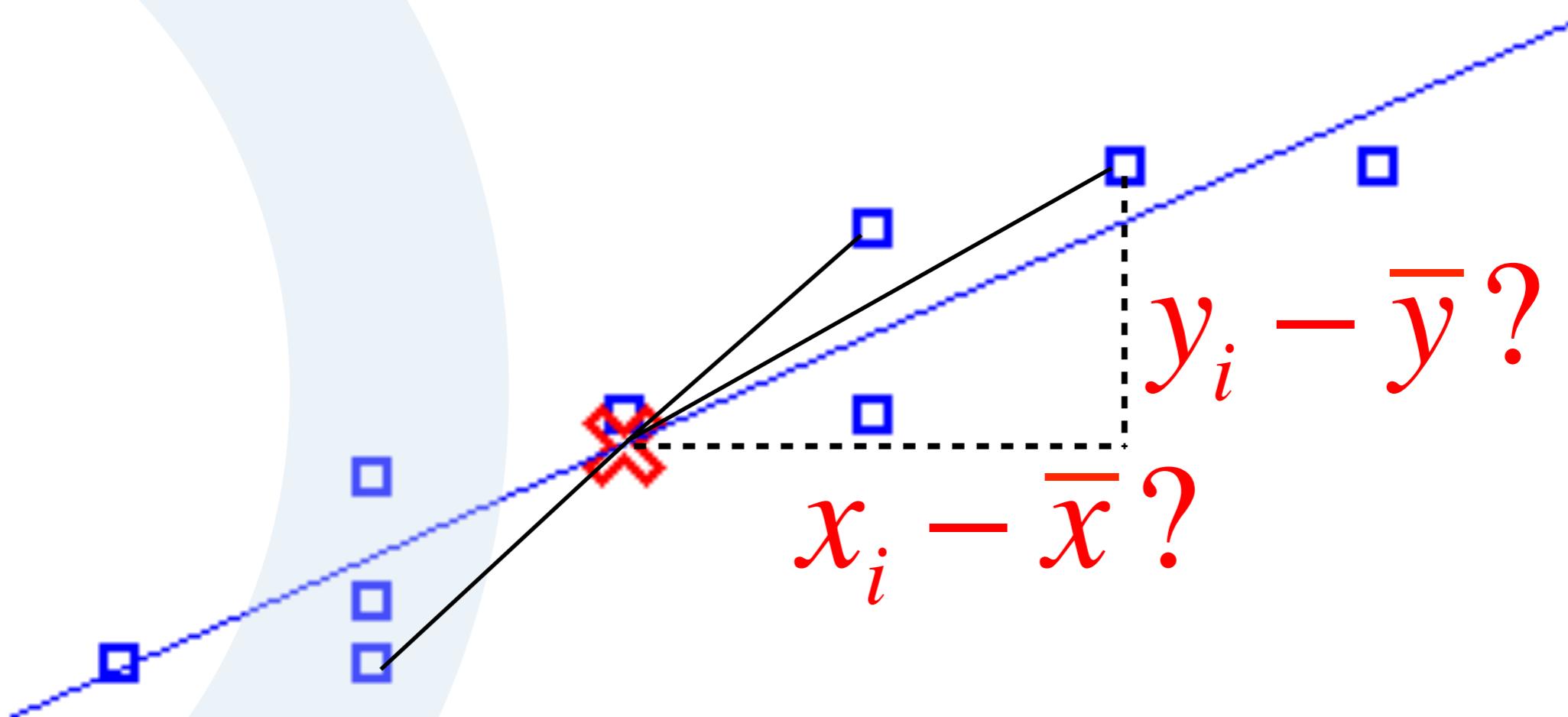
- Pente

$$\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})$$

$$a = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2}$$



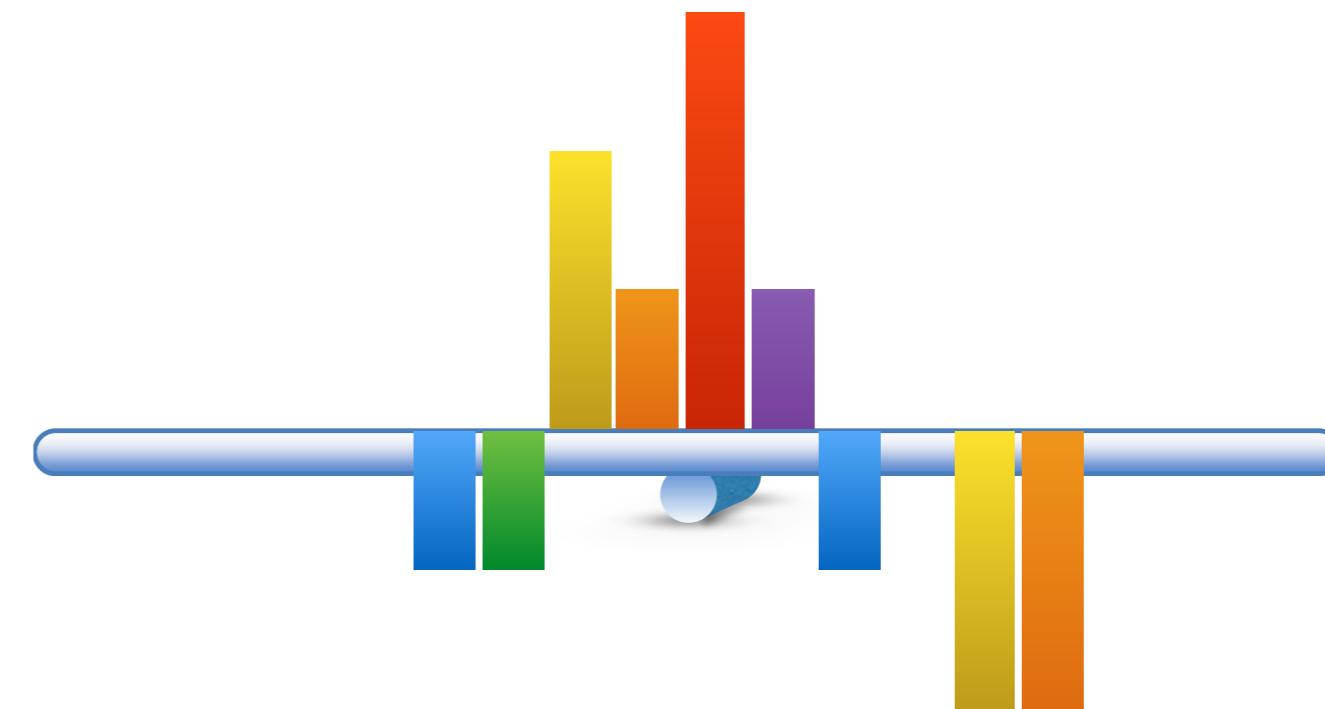
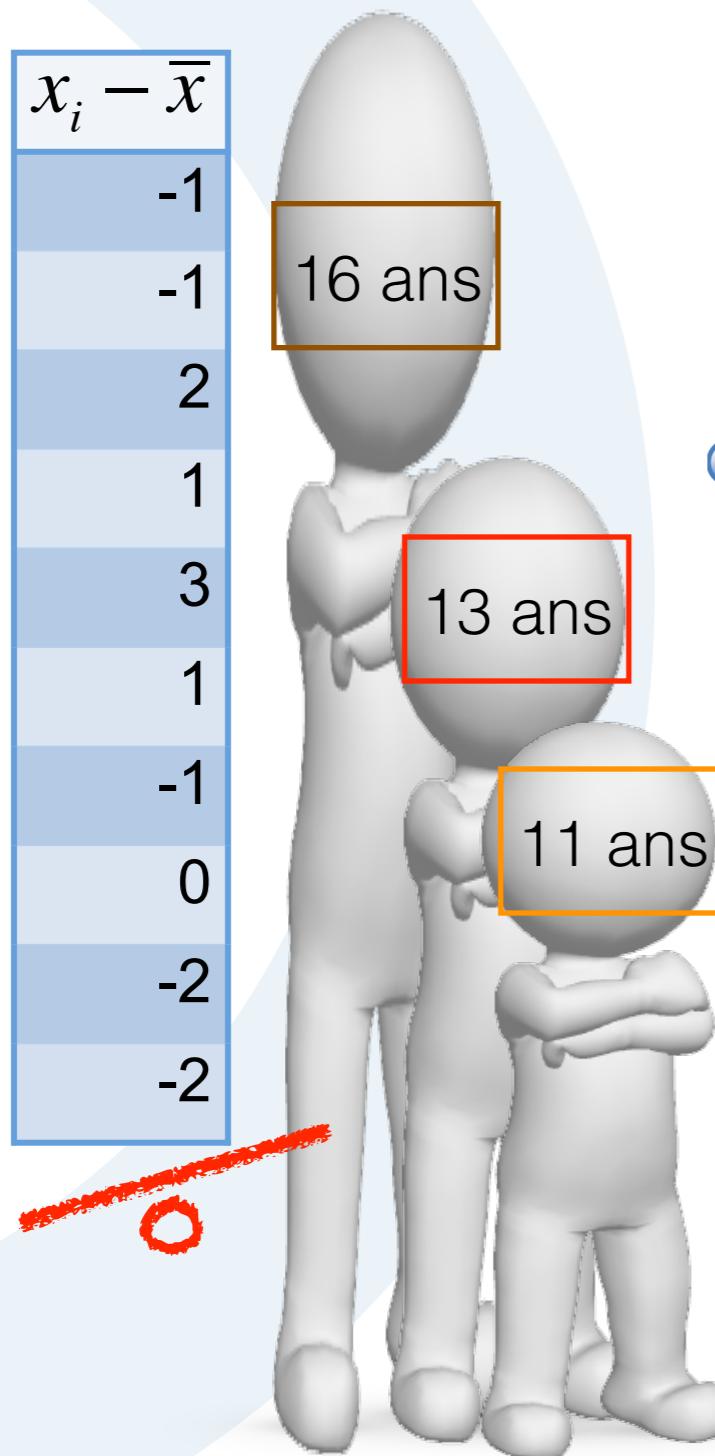
Dans le nuage, le point moyen est un point particulier.



Comment les points du nuage varient-ils par rapport à lui?

Où se situent-ils par rapport à lui?

$i$	$x_i$	$x_i - \bar{x}$
1	12	-1
2	12	-1
3	15	2
4	14	1
5	16	3
6	14	1
7	12	-1
8	13	0
9	11	-2
10	11	-2



$$\frac{\sum_{i=1}^n (x_i - \bar{x})}{n} = 0$$

« La moyenne des déviations est nulle »

$$x_i - \bar{x}$$

Deg Fix9 d/c Real

**2**

	List 1	List 2	List 3	List 4
SUB	AGE	ARGENT		
1	12	4		
2	12	3		
3	15	11		
4	14	10		

TOOL EDIT DELETE DEL-ALL INSERT >

**2**

	List 1	List 2	List 3	List 4
SUB	AGE	ARGENT		
1	12	4		
2	12	3		
3	15	11		
4	14	10		

List 1-13

OPTN

**F1**

LIST	COMPLEX	CALC	HYPERBL	PROB	>
List	Lst→Mat	Dim	Fill(	Seq	>

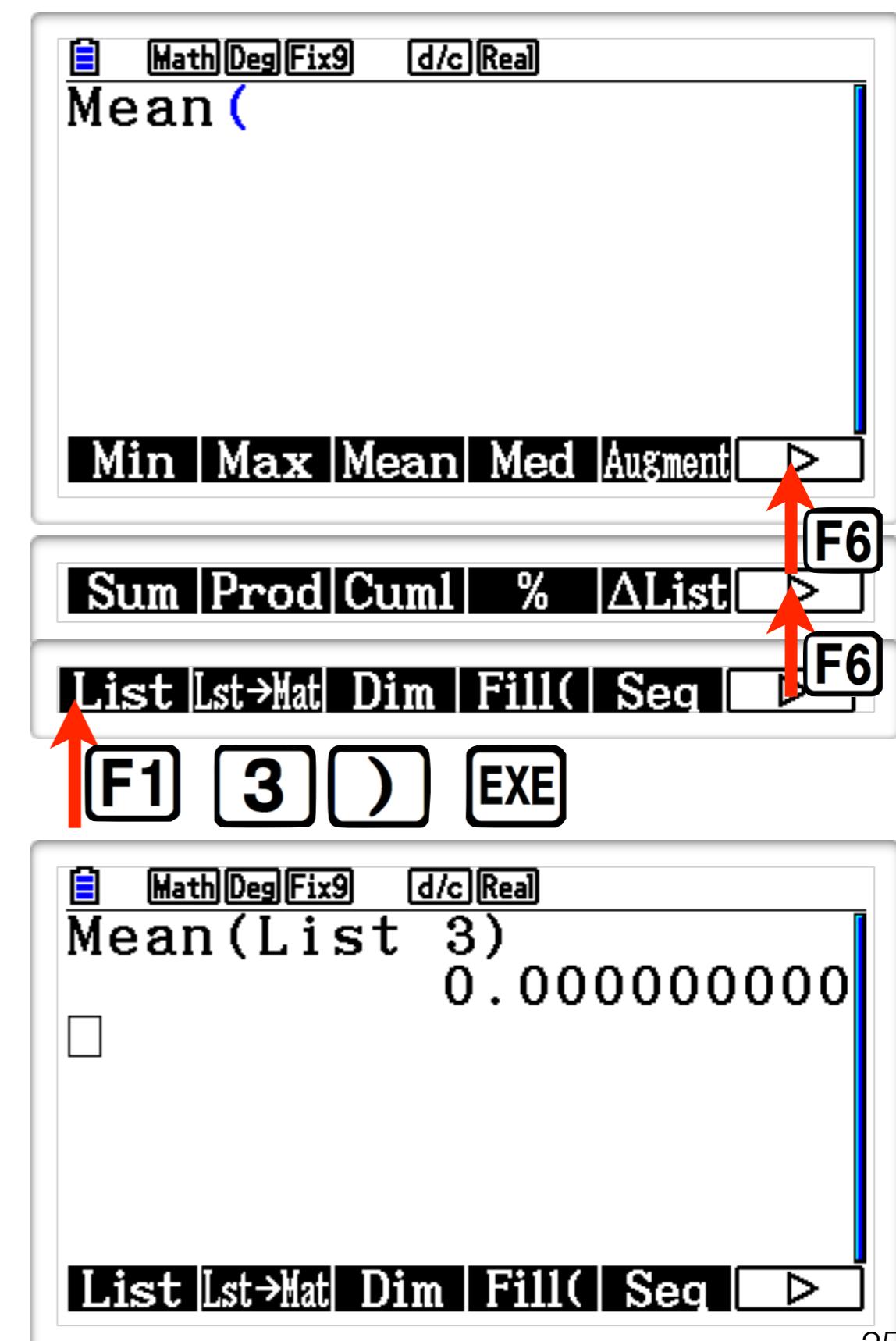
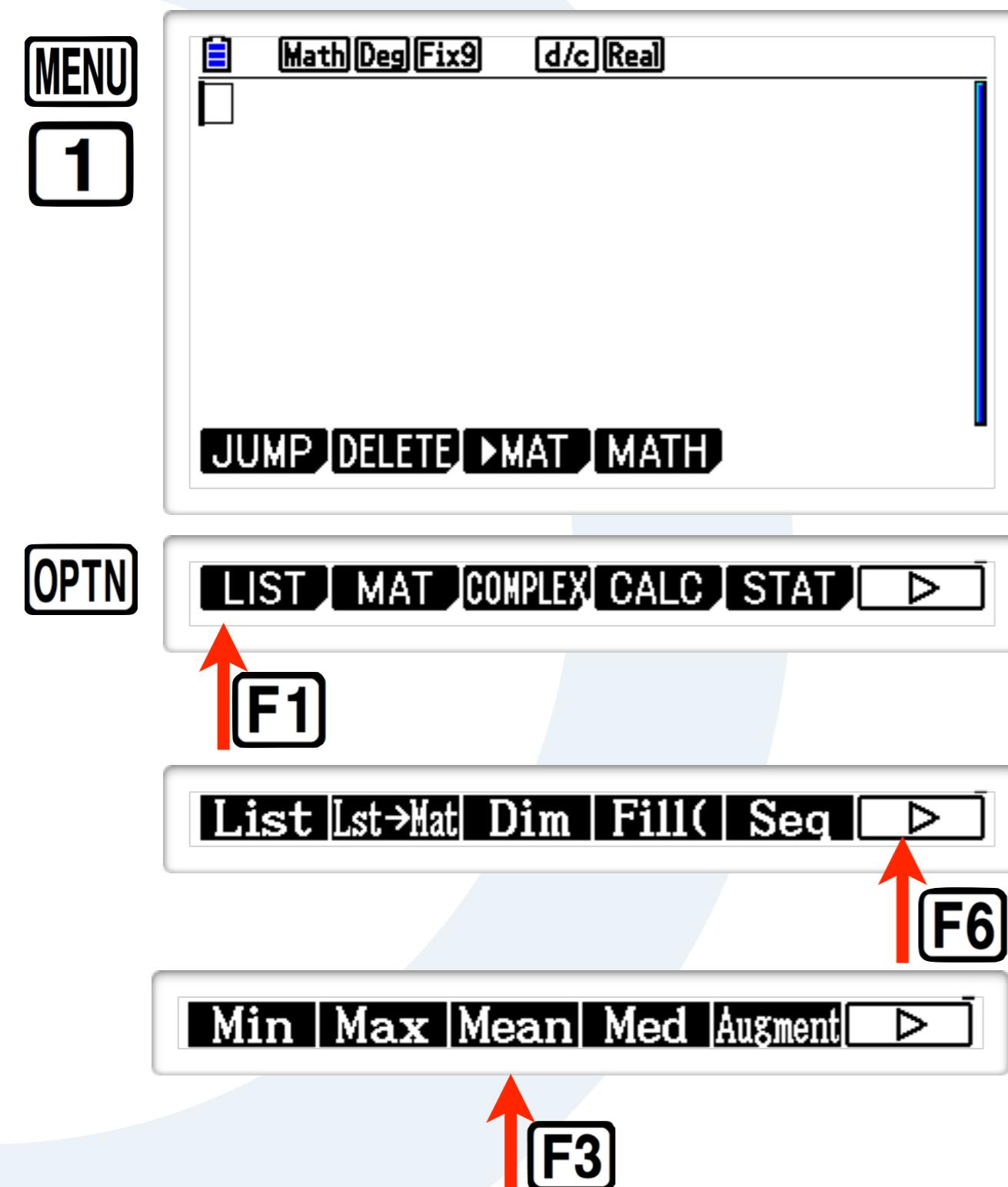
**F1**

EXE

	List 1	List 2	List 3	List 4
SUB	AGE	ARGENT		
1	12	4	-1	
2	12	3	-1	
3	15	11	2	
4	14	10	1	

- 1 . 0000000000

List Lst→Mat Dim Fill( Seq >

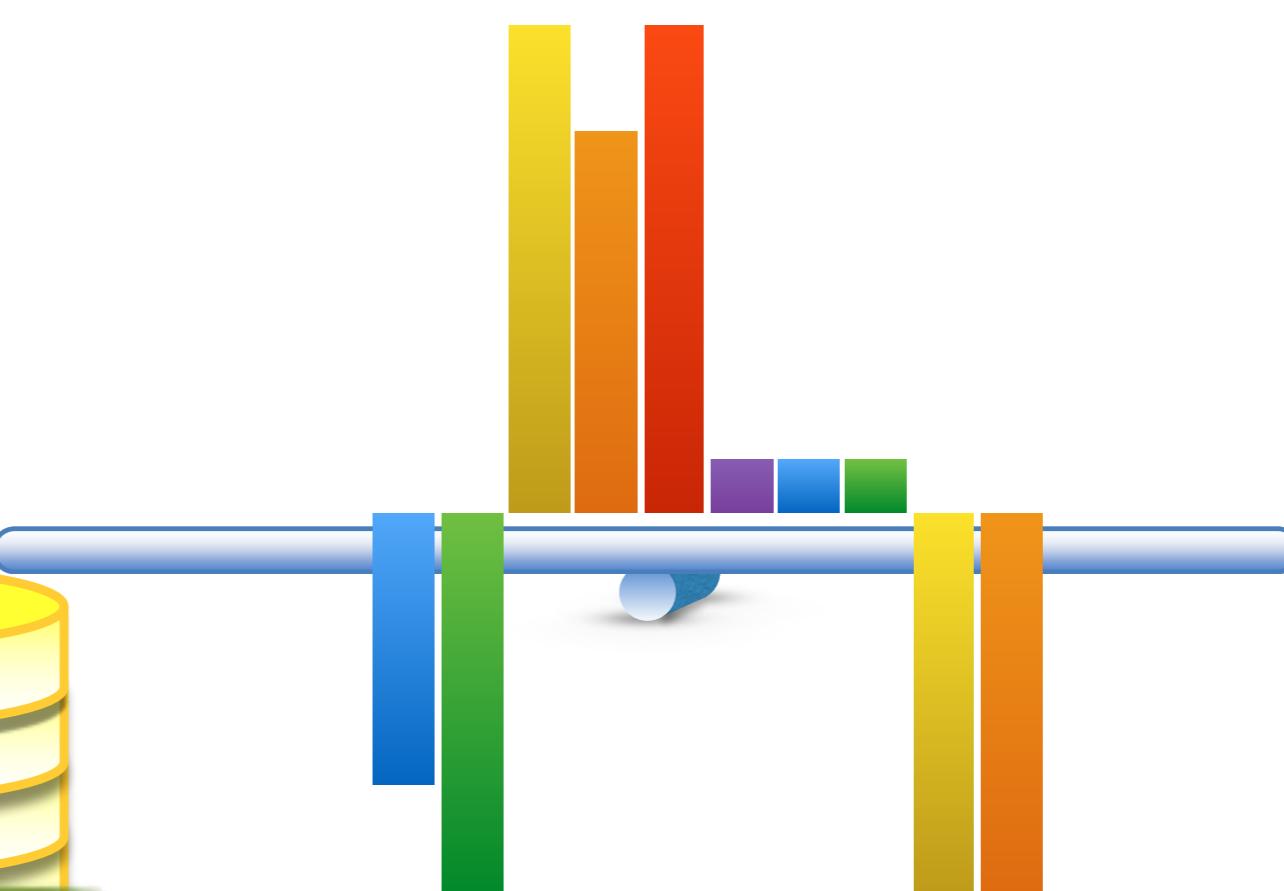


$i$	$y_i$	$y_i - \bar{y}$
1	4	-2,5
2	3	-3,5
3	11	4,5
4	10	3,5
5	11	4,5
6	7	0,5
7	6	0,5
8	7	0,5
9	3	-3,5
10	3	-3,5



« La moyenne des déviations est nulle »

$$\frac{\sum_{i=1}^n (y_i - \bar{y})}{n} = 0$$



	List 1	List 2	List 3	List 4
SUB	AGE	ARGENT	DEV X	
1	12	4	-1	
2	12	3	-1	
3	15	11	2	
4	14	10	1	

List 2-3.5|

EXE

EXIT

	List 1	List 2	List 3	List 4
SUB	AGE	ARGENT	DEV X	DEV Y
1	12	4	-1	-2.5
2	12	3	-1	-3.5
3	15	11	2	4.5
4	14	10	1	3.5

-2.500000000

GRAPH CALC TEST INTR DIST >

 $y_i$  $y_i - \bar{y}$ 

MENU

1

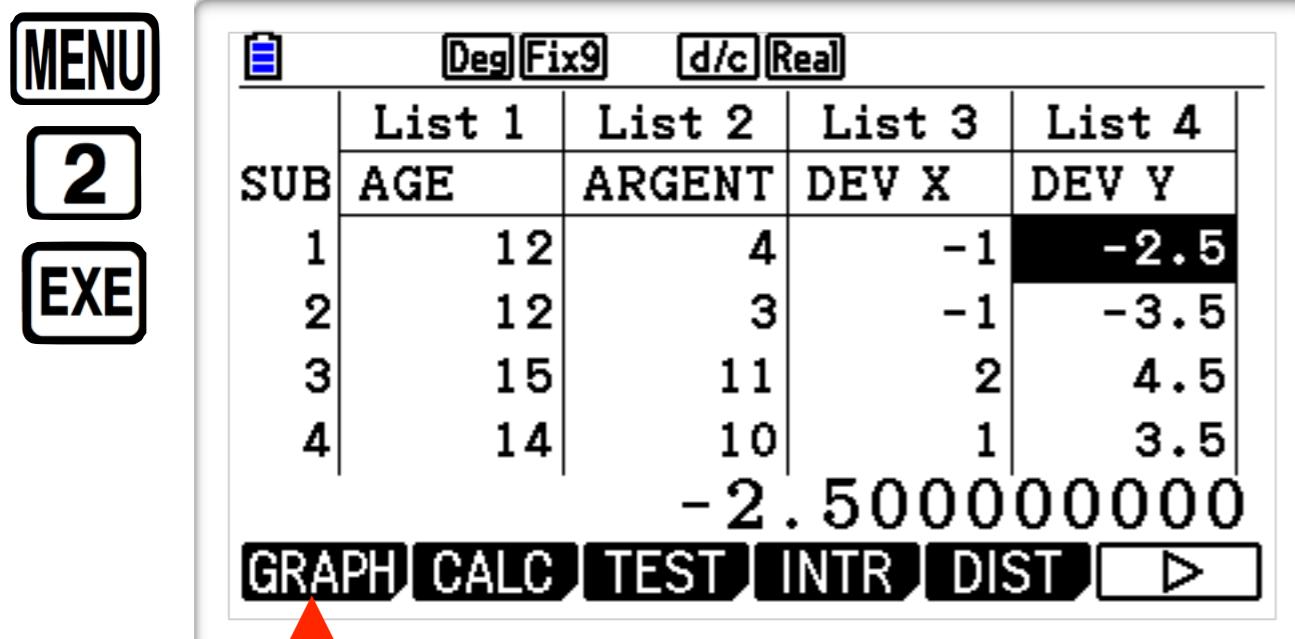
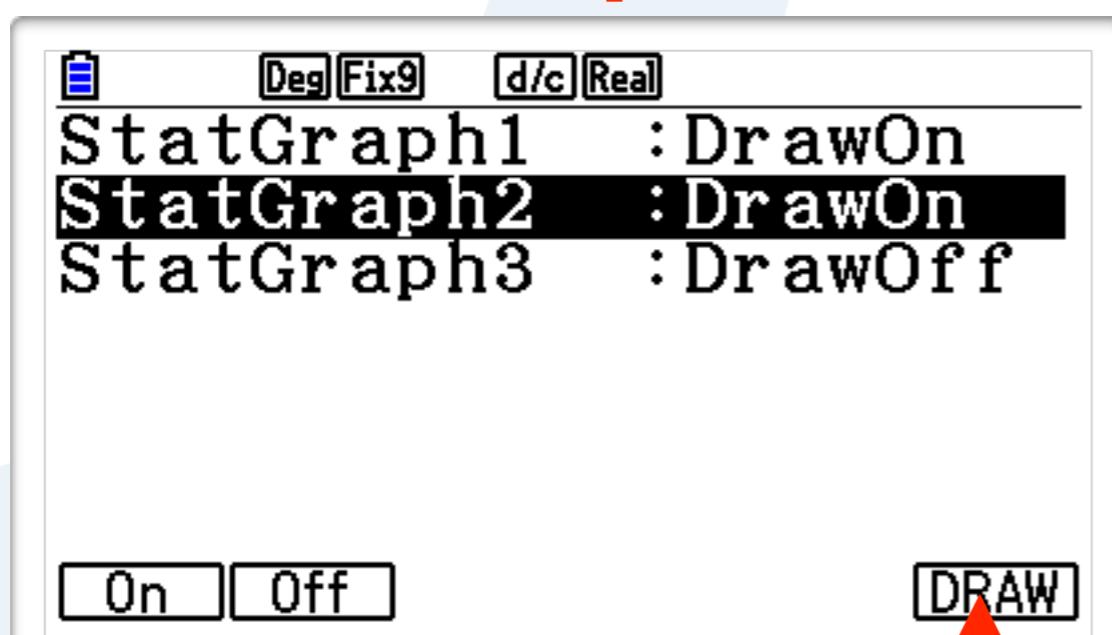
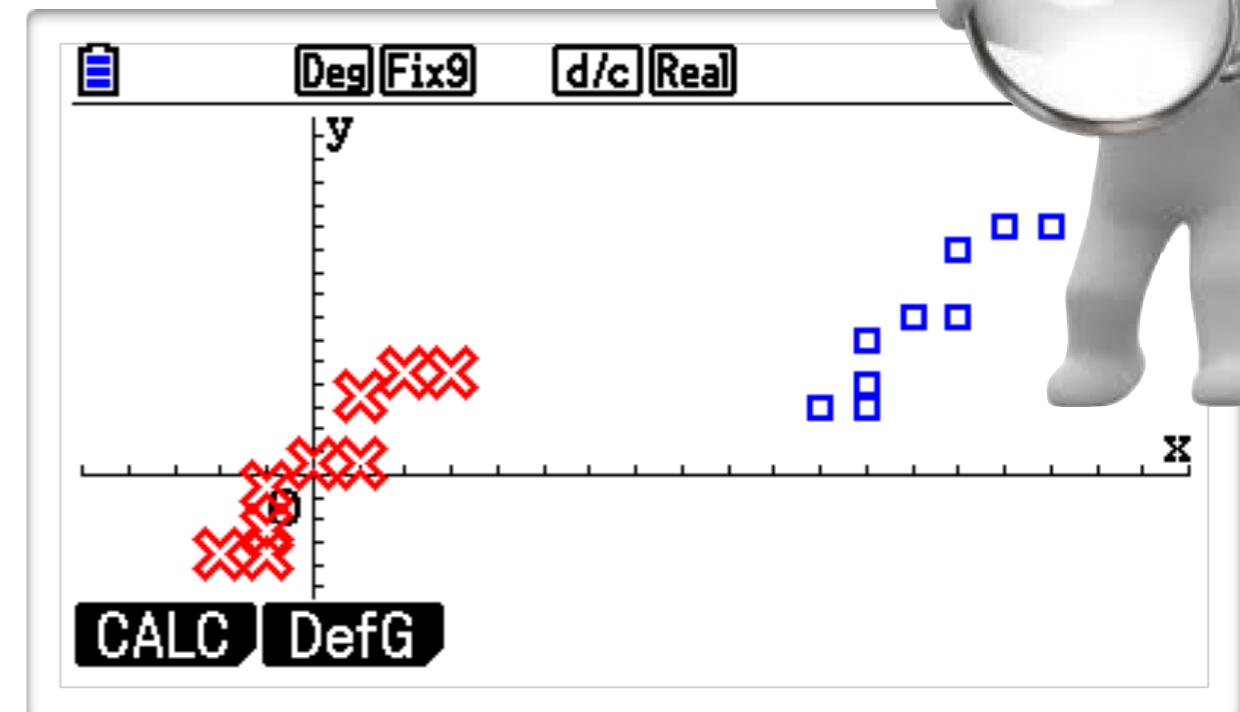
	Math	Deg	Fix9	d/c	Real
Mean (List 3)	0.000000000				
	List	Lst→Mat	Dim	Fill(	Seq >
	F6				

	Min	Max	Mean	Med	Augment
	F3				
	Sum	Prod	Cuml	%	ΔList >
	F6				
	List	Lst→Mat	Dim	Fill(	Seq >
	F1	4	)	EXE	

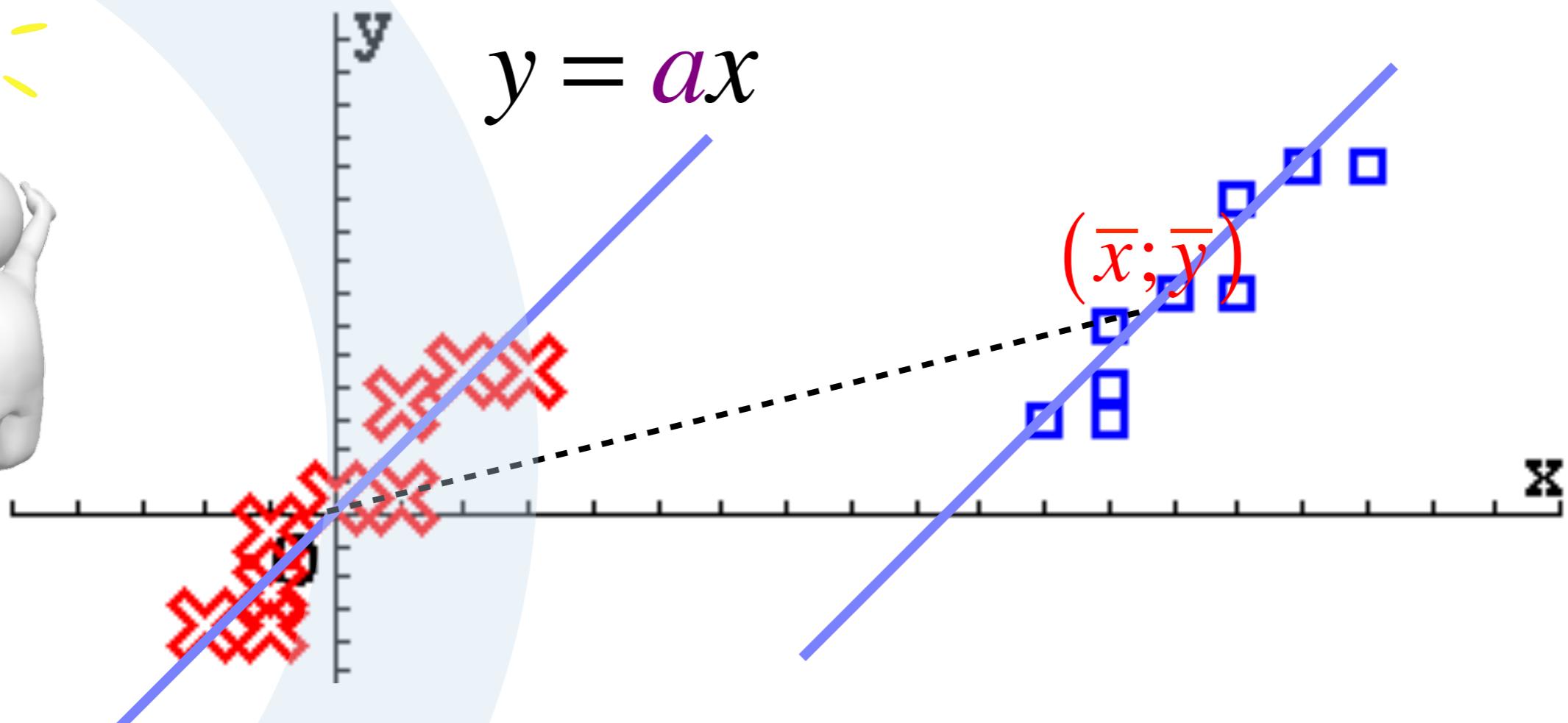
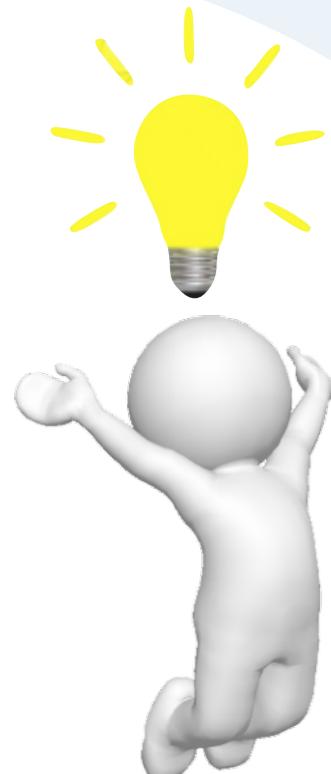


La moyenne des déviations, ou déviation moyenne, est toujours nulle.

La déviation moyenne ne nous donne aucune information sur l'orientation du nuage.

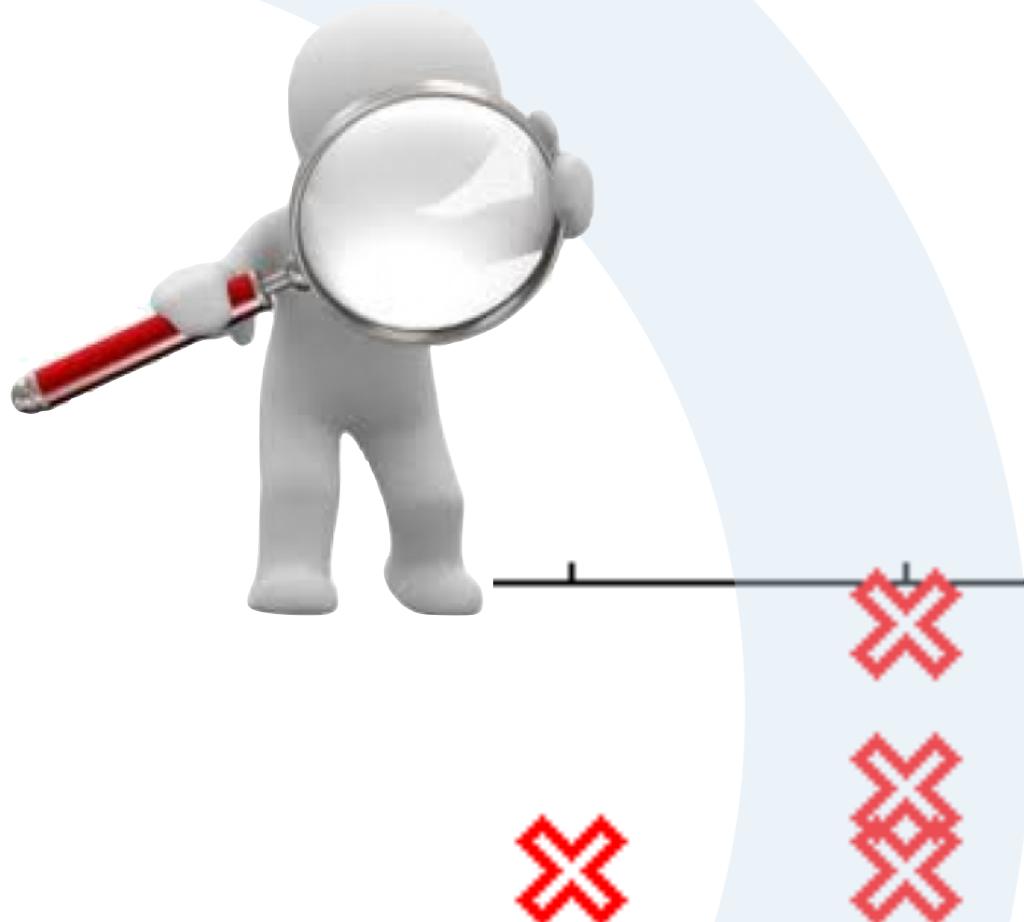
**F1****GRAPH1** **GRAPH2** **GRAPH3** **SELECT****SET****F4****F6**

Le nuage des déviations a la même orientation que le nuage de points.



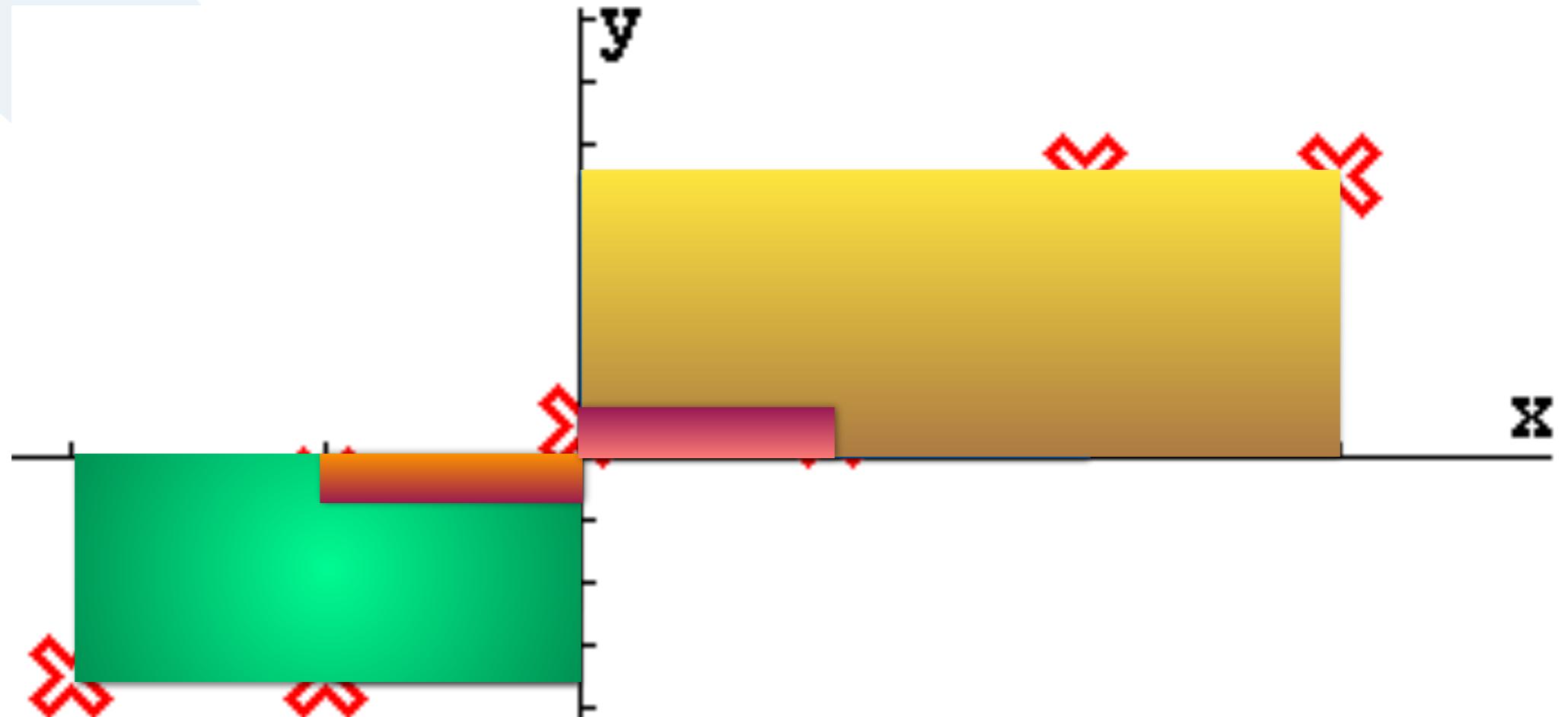
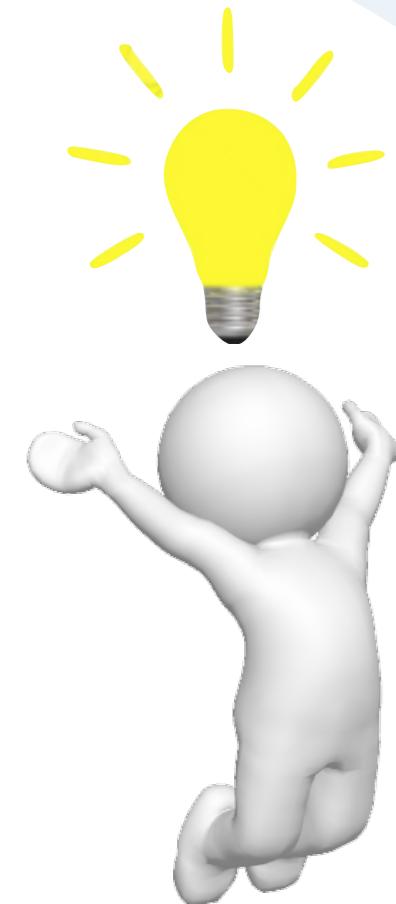
Si on détermine la pente du nuage des déviations, on a l'équation de la droite de régression

$$y = \bar{y} + a(x - \bar{x})$$

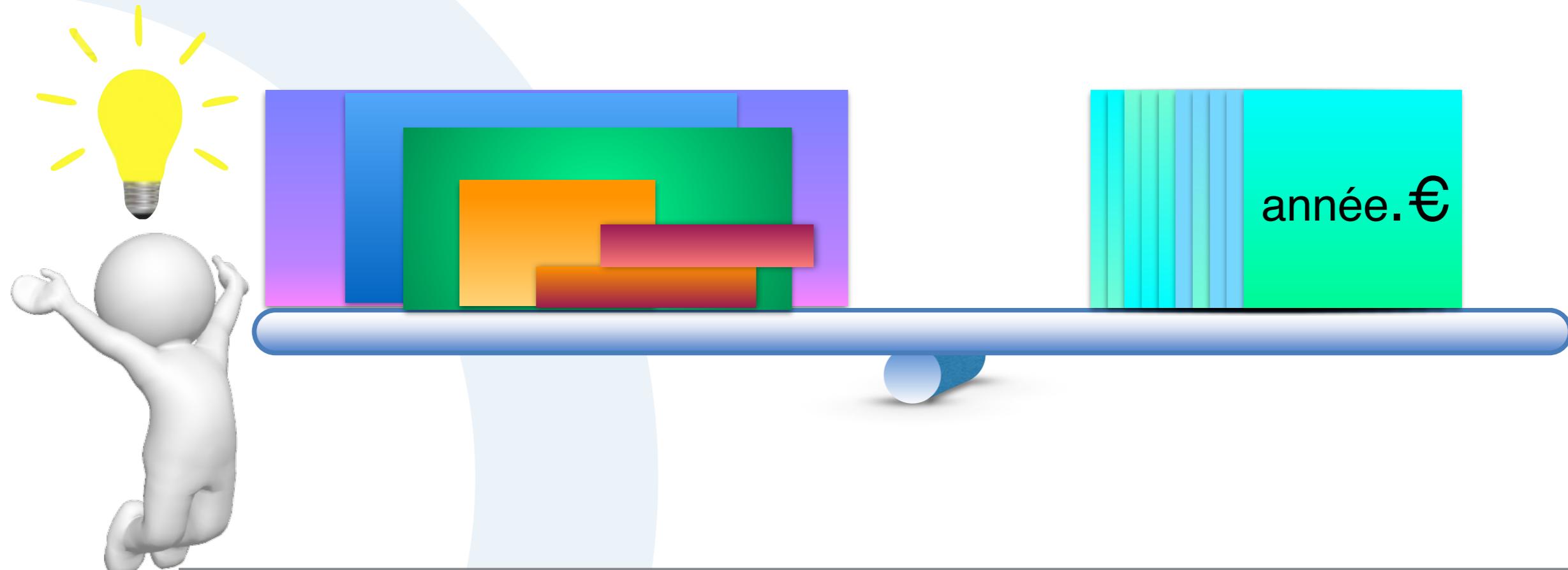


Les points du nuage des déviations se répartissent de part et d'autre de l'origine.

Les points du nuage des déviations sont dans les premier et troisième quadrants.



Les déviations sont positives ou négatives en même temps.  
Ici, le produit des déviations est positif.  
Un produit négatif correspondrait à des déviations de sens opposé.



$$s_{xy} := \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})$$

Une covariance positive correspond à des variations des deux variables dans le même sens.

Inversement une covariance négative correspond à des variations en sens opposé.

	Deg	Fix9	d/c	Real	
SUB	List 1	List 2	List 3	List 4	
1	AGE	ARGENT	DEV X	DEV Y	
2	12	4	-1	-2.5	
3	12	3	-1	-3.5	
4	15	11	2	4.5	
	14	10	1	3.5	
	- 2.500000000				
	GRAPH	CALC	TEST	INTR	DIST



Les déviations des abscisses et des ordonnées sont de même signe.

	Deg	Fix9	d/c	Real	
SUB	List 3	List 4	List 5	List 6	
1	DEV X	DEV Y	DX.DY		
2	-1	-2.5	2.5		
3	-1	-3.5	3.5		
4	2	4.5	9		
	1	3.5	3.5		
	Min	Max	Mean	Med	Augment

D	Math	Deg	Fix9	d/c	Real	
	Mean	(List 3)			0.000000000	
	Mean	(List 4)			0.000000000	
	Mean	(List 5)			4.700000000	
	List	Lst→Mat	Dim	Fill(	Seq	

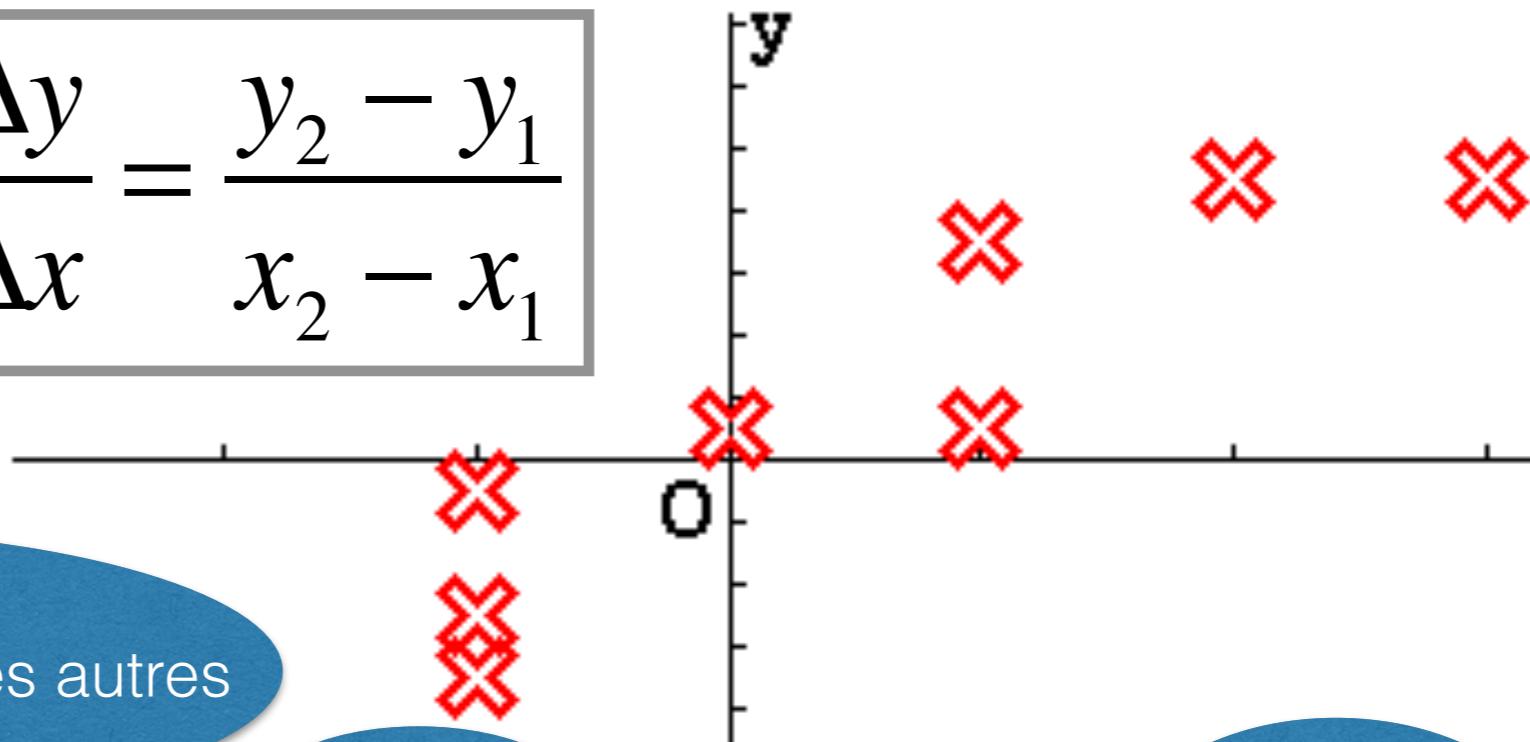


$$\frac{y_n - y_1}{x_n - x_1}$$

contrexemple  
avec des points dans les autres  
quadrants

$$\frac{1}{n} \sum_{i=1}^n \frac{y_i - \bar{y}}{x_i - \bar{x}}$$

$$a = \frac{\Delta y}{\Delta x} = \frac{y_2 - y_1}{x_2 - x_1}$$

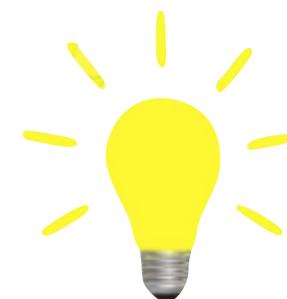


0/0

toujours &gt;0

$$\frac{\frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})}{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})}$$

$$\frac{\frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^2}{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2}$$



$$\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})$$
 traduit l'orientation du nuage

$$\frac{\frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})}{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})}$$

$$\sum_{i=1}^n (x_i - \bar{x})^2 > 0$$

$$\frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(x_i - \bar{x})}$$

$$a := \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} = \frac{s_{xy}}{s_x^2}$$

Plus un élève est éloigné en âge de l'âge moyen, plus il influence la pente de la droite ; plus il est proche de l'âge moyen, moins il l'influence.

$$a = \frac{\frac{1}{n} \sum_{i=1}^n \frac{(x_i - \bar{x})}{s_x^2} (y_i - \bar{y})}{1}$$

La pente représente par année supplémentaire, la moyenne des déviations des ordonnées « pondérées » par la déviation de leurs abscisses correspondantes.

# CONSTRUCTION DE LA DROITE

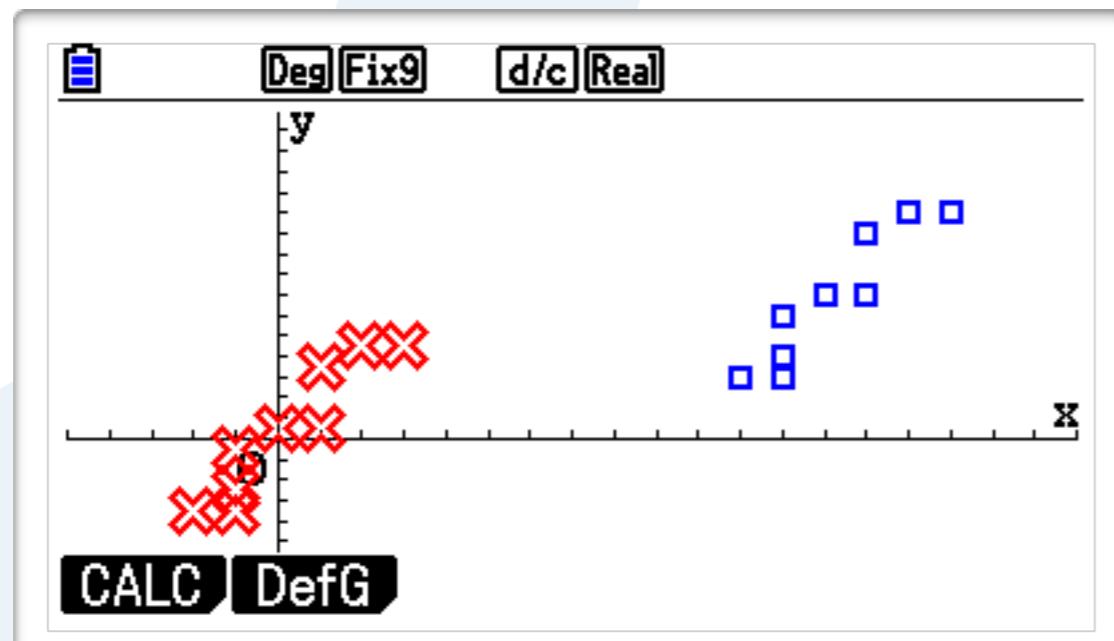
**MENU**

**2**

**Deg Fix9 d/c Real**

	List 1	List 2	List 3	List 4
SUB	AGE	ARGENT	DEV X	DEV Y
1	12	4	-1	-2.5
2	12	3	-1	-3.5
3	15	11	2	4.5
4	14	10	1	3.5
12.00000000				
<b>GRAPH</b>	<b>CALC</b>	<b>TEST</b>	<b>INTR</b>	<b>DIST</b>

**GRAPH1** **GRAPH2** **GRAPH3** **SELECT** **SET**



# Equation de la droite

$$a = 4,7 / 1,61^2$$

$$= 1,81$$

**Deg Fix9 d/c Real**

**Graph Func** :  $Y =$

$Y_1 = 1.81x$  [—]

$Y_2 = 6.5 + Y_1(x-13)$  [—]

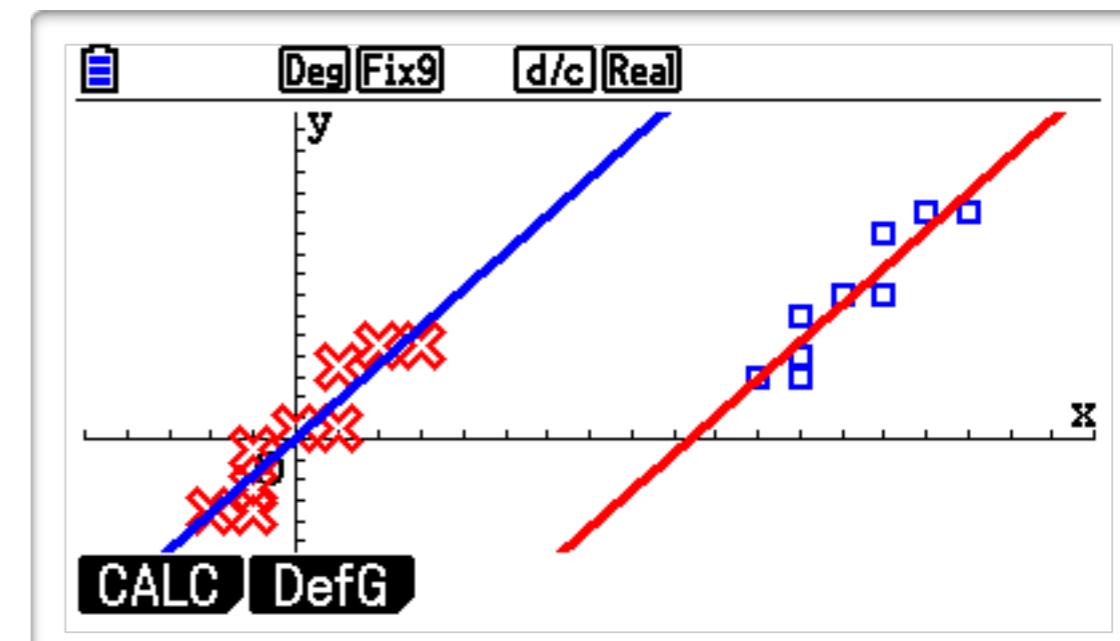
**Y3:** [—]

**Y4:** [—]

**Y5:** [—]

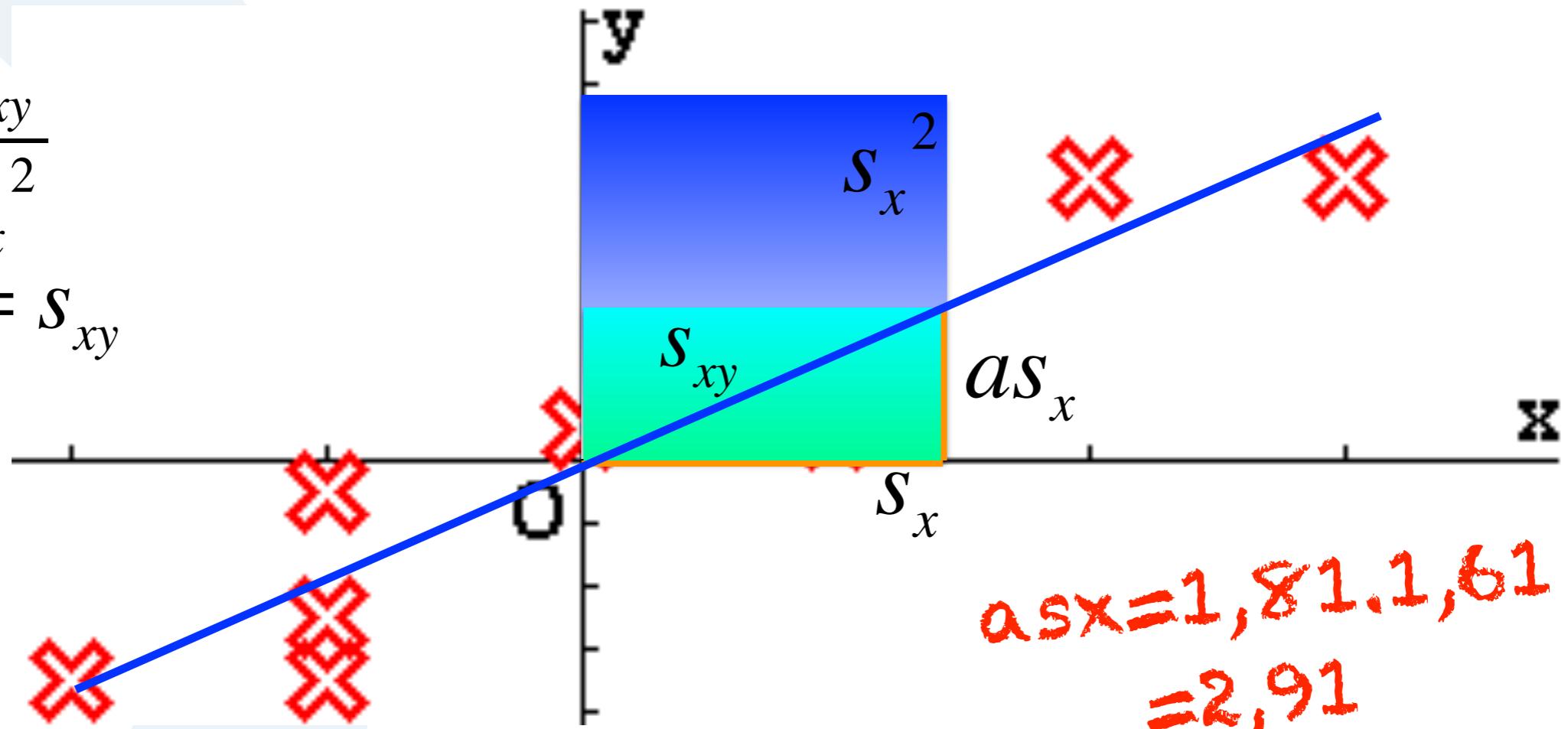
**Y6:** [—]

**SELECT** **DELETE** **Y** **STYLE** **DRAW**



$$a = \frac{s_{xy}}{s_x^2}$$

$$as_x^2 = s_{xy}$$



Un élève âgé de 14,5 ans reçoit théoriquement presque 3 euros en plus que l'élève moyen.

# QUALITÉ DE L'AJUSTEMENT



# QUALITÉ DE L'AJUSTEMENT

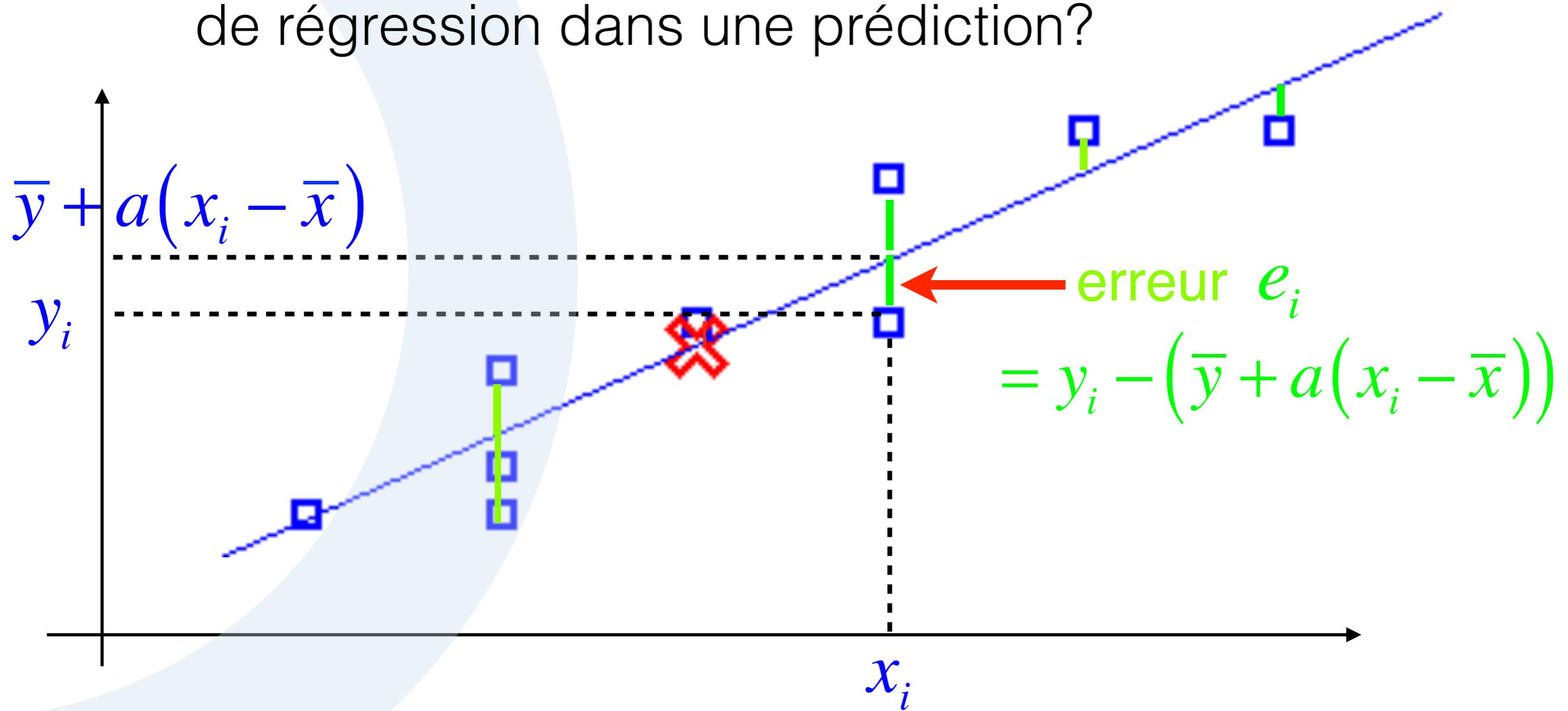
- Méthode des moindres carrés
- Coefficient de corrélation

$$\text{Minimiser} \frac{1}{n} \sum_{i=1}^n e_i^2$$

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}}$$



Quelle est l'erreur commise en utilisant la droite de régression dans une prédiction?



Résidu = la différence entre l'ordonnée observée et l'ordonnée théorique

$$\begin{aligned}
 e_i &= y_i - (\bar{y} + a(x_i - \bar{x})) \\
 &= (y_i - \bar{y}) - a(x_i - \bar{x})
 \end{aligned}$$

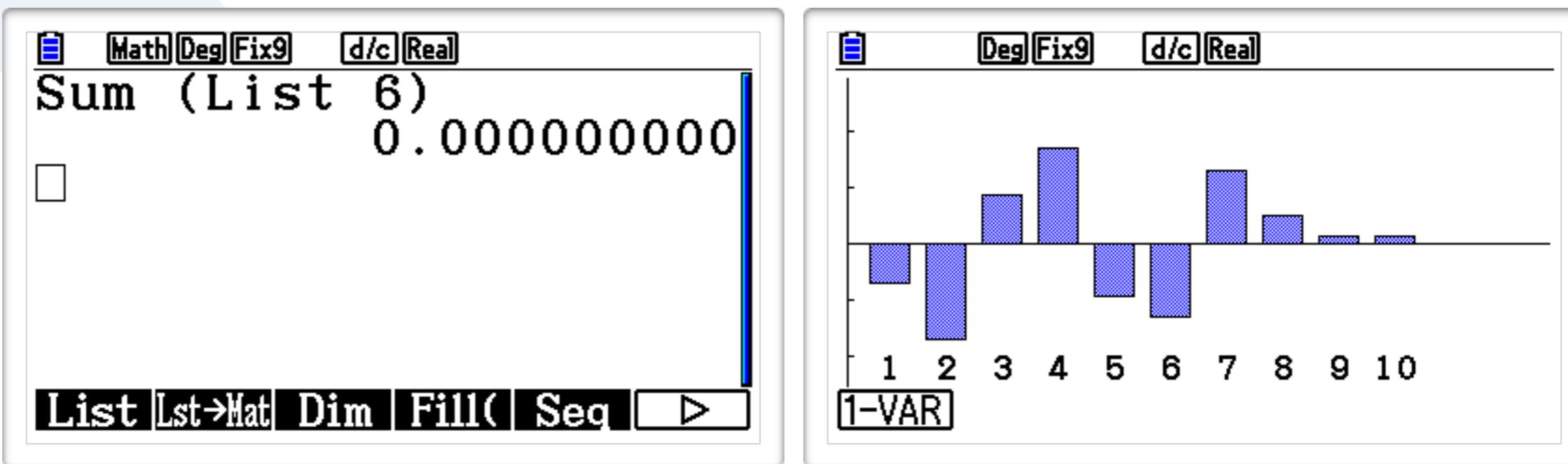
$$(y_i - \bar{y}) - 1.81(x_i - \bar{x})$$

		Deg	Fix9	d/c	Real	
SUB	DEV X	List 3	List 4	List 5	List 6	
		DEV Y	DX.DY			
1	-1	-2.5	2.5			
2	-1	-3.5	3.5			
3	2	4.5	9			
4	1	3.5	3.5			

Min | Max | Mean | Med | Augment | ►

		Deg	Fix9	d/c	Real	
SUB	DEV X	List 3	List 4	List 5	List 6	
		DEV Y	DX.DY			
1	-1	-2.5	2.5		-0.69	
2	-1	-3.5	3.5		-1.69	
3	2	4.5	9		0.88	
4	1	3.5	3.5		1.69	
						-0.690000000

List | Lst→Mat | Dim | Fill( | Seq | ►



Les résidus sont tantôt positifs, tantôt négatifs.

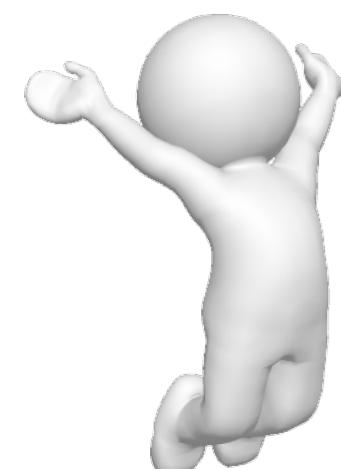
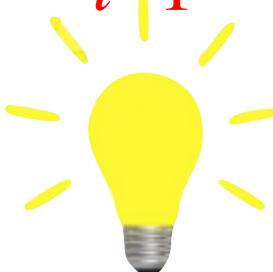
La moyenne des résidus est toujours nulle.

Elle n'apporte donc aucune information pour mesurer

l'ajustement aux données.

$$\sum_{i=1}^n e_i = \underbrace{\sum_{i=1}^n (y_i - \bar{y})}_{=0} - a \underbrace{\sum_{i=1}^n (x_i - \bar{x})}_{=0} = 0$$

$$\sum_{i=1}^n e_i^2$$



$$e_i = (y_i - \bar{y}) - a(x_i - \bar{x})$$

$$\sum_{i=1}^n e_i^2 = \sum_{i=1}^n (y_i - \bar{y})^2 + a^2 \sum_{i=1}^n (x_i - \bar{x})^2 - 2a \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})$$

$$\boxed{\frac{1}{n} \sum_{i=1}^n e_i^2 = s_y^2 + a^2 s_x^2 - 2as_{xy}}$$

	Deg	Fix9	d/c	Real	List 4	List 5	List 6	List 7
SUB	DEV Y				DX.DY	DY-ADX		
1	-2.5				2.5	-0.69		
2	-3.5				3.5	-1.69		
3	4.5				9	0.88		
4	3.5				3.5	1.69		
	List 6 <sup>2</sup>							

 $e_i$  
 $e_i^2$  

	Deg	Fix9	d/c	Real	List 4	List 5	List 6	List 7
SUB	DEV Y				DX.DY	DY-ADX	RES <sup>2</sup>	
1	-2.5				2.5	-0.69	0.4761	
2	-3.5				3.5	-1.69	2.8561	
3	4.5				9	0.88	0.7744	
4	3.5				3.5	1.69	2.8561	
							0.476100000	
					List	Lst→Mat	Dim	Fill(
							Seq	▶

	Math	Deg	Fix9	d/c	Real
	Sum	(List 6)			0.000000000
	Mean	(List 7)			1.153860000
					□
					List Lst→Mat Dim Fill( Seq ▶

$$\frac{1}{n} \sum_{i=1}^n e_i^2 = s_y^2 + a^2 s_x^2 - 2as_{xy}$$

$$= s_x^2 a^2 - 2s_{xy}a + s_y^2$$

↑ ↑

$$\frac{-b}{2a} = \frac{+2s_{xy}}{2s_x^2} = \frac{s_{xy}}{s_x^2}$$

$$\frac{-D}{4a} = -\frac{(2s_{xy})^2 - 4s_x^2 s_y^2}{4s_x^2}$$

$$= -\frac{s_{xy}^2 - s_x^2 s_y^2}{s_x^2}$$

La moyenne quadratique minimale est atteinte quand  $s_x^2$

$$a = \frac{s_{xy}}{s_x^2}$$

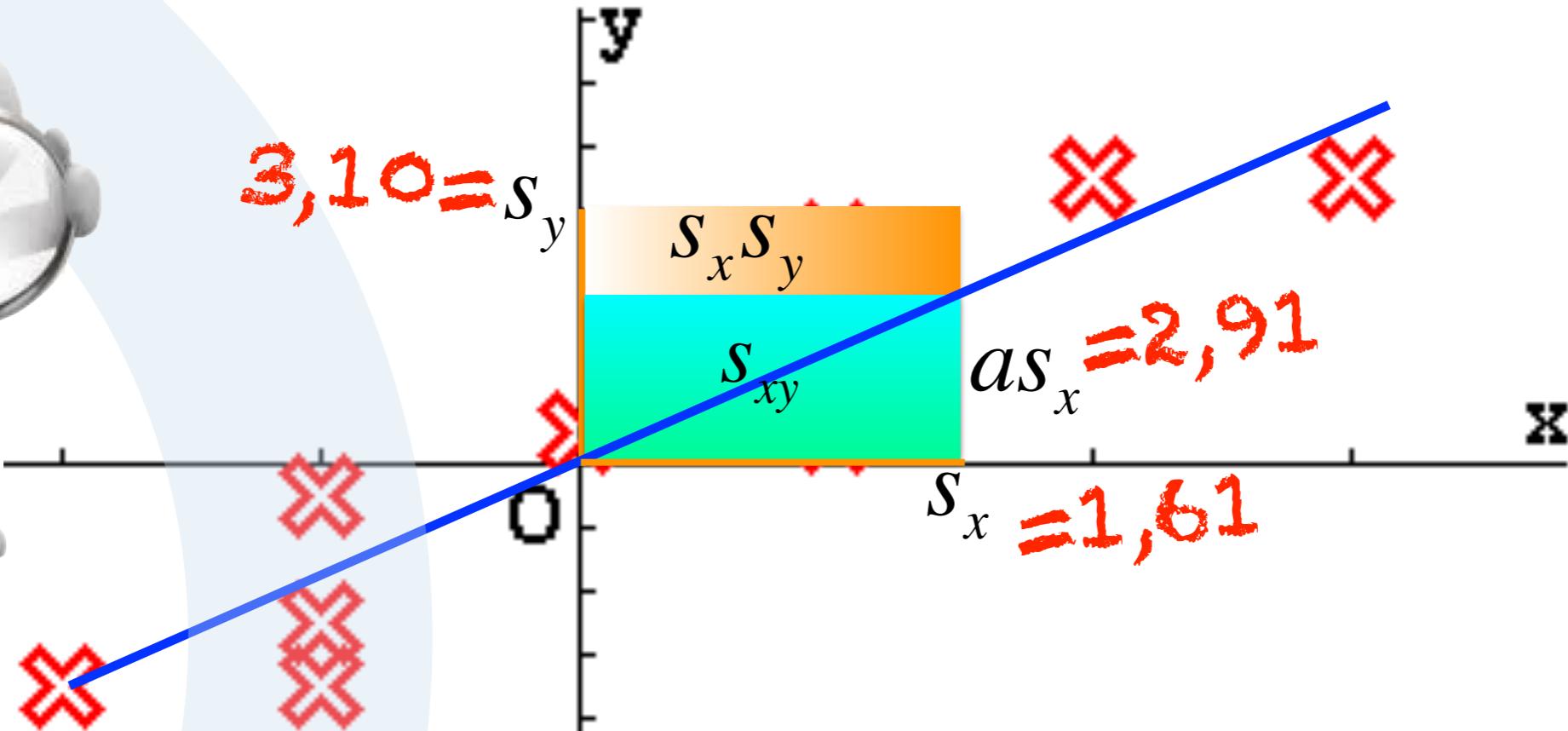


Ce minimum, erreur quadratique moyenne, vaut

$$EQM = s_y^2 - \frac{s_{xy}^2}{s_x^2}$$

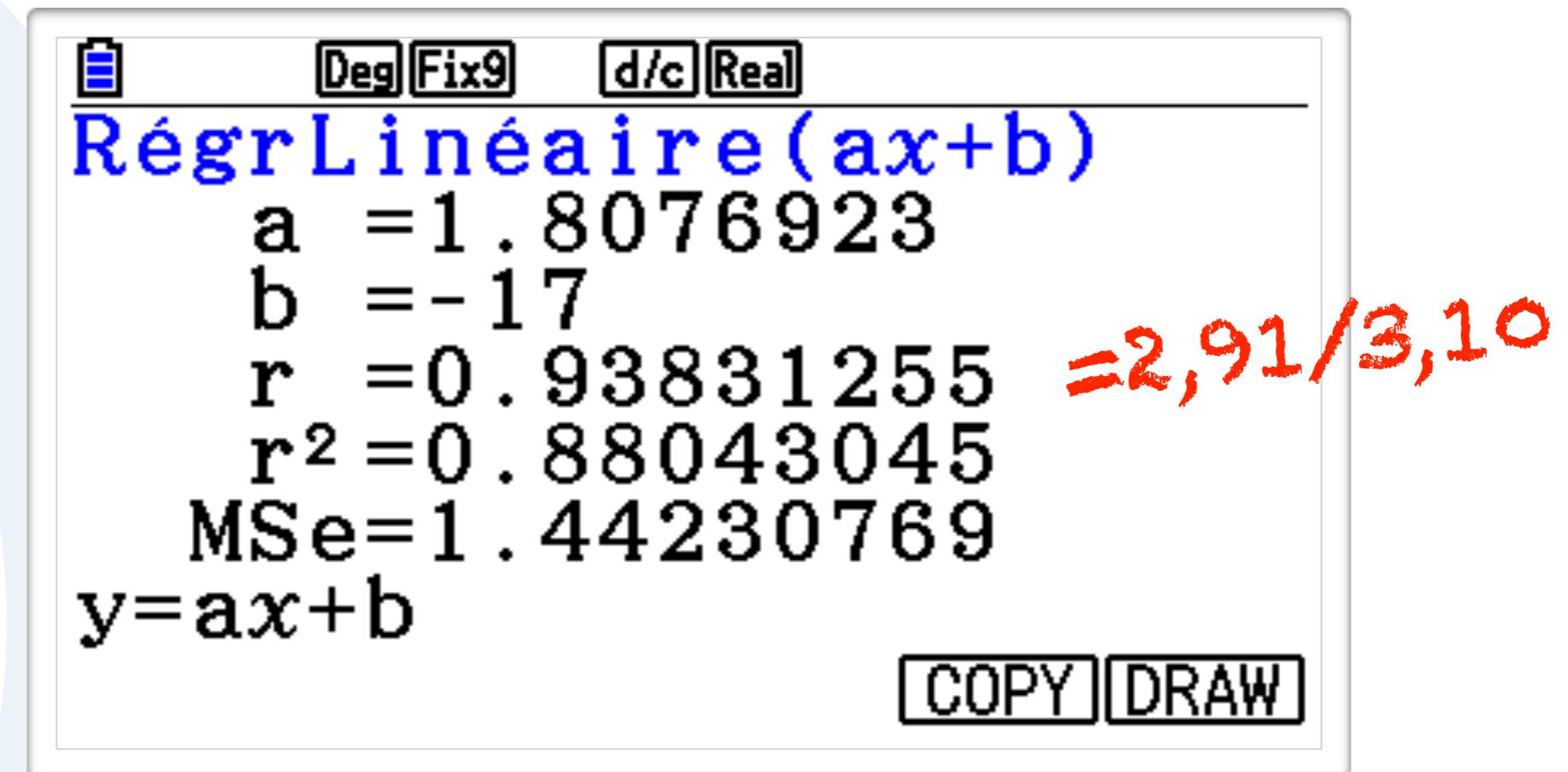
$$\begin{aligned}
 EQM &= s_y^2 - \frac{s_{xy}^2}{s_x^2} \\
 \frac{EQM}{s_y^2} &= 1 - \left( \frac{s_{xy}}{s_x s_y} \right)^2 \\
 &= 1 - r^2
 \end{aligned}
 \quad r := \frac{s_{xy}}{s_x s_y} = \frac{as_x}{s_y}$$

qualité d'ajustement



$$r = \frac{s_{xy}}{s_x s_y} = \frac{a s_x}{s_y}$$

Le coefficient de corrélation représente le rapport entre l'écart type attendu et l'écart type observé.



Plus le nuage est concentré,  
plus l'erreur quadratique est faible,  
meilleure est la qualité de l'ajustement,  
plus le coefficient de corrélation est proche de 1.

# Statistiques à 2 variables

## ... et calculatrice graphique



# Merci !

## Congrès SBPMef 2015

Equipe **CASIO** *Education*

Françoise Delpérée – [fdelperee@hotmail.com](mailto:fdelperee@hotmail.com)  
Virginie Loward – [virginie.loward@skynet.be](mailto:virginie.loward@skynet.be)

*Ce document, en tout ou en partie, peut être reproduit uniquement **à des fins pédagogiques** et ce, pour autant que la source soit citée.*

*Merci,  
l'équipe **CASIO** Education*