Module : Statistiques

Chapitre : L'élaboration des séries statistiques

Paragraphe I : Détermination du champ de l'étude et des outils d'observation :

1-1- Détermination du champ de l'étude

On veut connaître l'âge moyen des marocains vivant sur le territoire national.

Exploitation :

Deux méthodes sont possibles pour le faire :

- Relever l'âge de tous les marocains habitant sur le territoire national et diviser par le nombre de marocains. On procède dans ce cas à un **recensement**.
- Relever au hasard l'âge de quelques milliers de marocains à travers le territoire national et diviser par le nombre. On procède dans ce cas à **un sondage**.
- La première solution **(recensement)**, qui paraît plus sûre, manque parfois de précisions. En effet, pendant qu'on relève les âges, la population change (il y a des naissances et des décès), si bien qu'en pratique on ne peut pas obtenir l'âge moyen exact mais seulement une bonne approximation. En outre, il faut signaler que c'est une opération coûteuse.
- La deuxième solution qui consiste à prendre un **échantillon** (une partie des marocains) dans la population (c'est-à-dire dans l'ensemble de tous les marocains) est moins coûteuse et assez précise dans la mesure où :
- Les âges prélevés sont pris au hasard dans l'ensemble de la population ;
- Le nombre d'âges relevés est important.

1-2- Les outils d'observation :

1-2-1- L'observation des faits peut se faite :

- de manière continue :

. les notes des élèves d'une classe dans une matière déterminée ;

. les absences dans un lycée ;

. les pièces défectueuses fabriquées par une machine par période.

- de manière périodique :
. le recensement de la population, ex : au Maroc il y a eu un recensement en 1971, 1980, 1994, 2004 et 2014.
. l'étude des relevés bancaires (tous les mois).

- de manière rationnelle :
- événements passagers : épidémie...

1-2-2- L'observation des faits peut porter sur :
- l'ensemble de la population : **recensement** ;
- un sous ensemble de la population : **sondage.**

1-2-3- Le sondage peut être :
- au hasard : on prélève d'une manière arbitraire des individus composant la population statistique jusqu'à atteindre la dimension de l'échantillon. Ce choix au hasard peut être source de plusieurs erreurs. Par exemple dans le cas de l'étude de la taille des élèves d'un établissement scolaire, le choix au hasard de quelques élèves peut donner un résultat erroné.
- Par choix raisonné : reprenons l'exemple de l'étude de la taille des élèves d'un établissement scolaire, on procédera par la création de quatre catégories d'élèves : les élèves ayant une petite taille, les élèves ayant une taille moyenne, ceux ayant une taille au-dessus de la moyenne et ceux ayant une grande taille. On choisira ensuite le nombre d'élèves que contient l'échantillon et les différentes catégories d'élèves représentés proportionnellement dans l'échantillon. Si le nombre total des élèves est de 700 et l'échantillon est de 100 et si les petits représentent la moitié, ils devront être au nombre de 50 dans l'échantillon.

1-2-4- Le questionnaire :
Ce sont les questions qui peuvent être :
- posées oralement à la personne concernée par l'enquête ;
- remises en main propre par l'enquêteur ou envoyées par courrier à son adresse afin qu'elles soient remplies par ce dernier.

Paragraphe II : Dépouillement et saisie des observations :
Le but de tout travail préliminaire en statistique est de pouvoir établir une distribution permettant de déterminer comment les données d'un échantillon se répartissent dans l'échelle des valeurs possibles.
Les données sont recueillies à partir des questionnaires établis préalablement en fonction d'objectifs arrêtés au début de l'étude.

La présentation des données sous forme de distribution statistique dépend du caractère étudié.

2-1- Cas d'un caractère qualitatif :

Exemple :

Une enquête menée auprès d'un échantillon de 40 élèves choisis au hasard dans un établissement secondaire qualifiant de 700 élèves, sur leur opinion quant à l'ouverture d'une buvette à l'entrée de l'établissement a permis de recueillir les réponses suivantes :

- très défavorable :
- défavorable :
- indécis :
- indifférent :
- favorable :
- très favorable.

Chacune des réponses (très défavorable, défavorable…) constitue une modalité du caractère « opinions sur l'ouverture d'une buvette à l'entrée de l'établissement ».

L'image de la consignation des réponses des 40 élèves sur une feuille de papier est la suivante :

Très défavorable – très défavorable – favorable – indécis – indifférent – favorable – très défavorable – favorable – indécis – indécis – très défavorable – très défavorable – très défavorable – favorable – indifférent – indécis – très défavorable – très défavorable – favorable – indécis – indifférent – indécis – favorable - favorable – très défavorable – favorable – indécis – favorable – indifférent – très défavorable – très défavorable – favorable – défavorable – indécis – indifférent – favorable – très défavorable – favorable – indécis – très favorable.

- Exploitation :

Le caractère a six modalités : très défavorables, défavorable, indécis, indifférent, favorable, très favorable. On peut attribuer à chacune des modalités un code qui peut être soit alphabétique soit numérique soit alphanumérique avant de commencer à dénombrer les élèves correspondant à chaque modalité.

- par exemple à :

- très défavorable on peut associer soit 1 soit a sait a1
- défavorable on peut associer soit 2 soit b soit b2
- indécis on peut associer soit 3 soit c soit c3
- indifférent on peut associer soit 4 soit d soit d4
- favorable on peut associer soit 5 soit e soit e5
- très favorable on peut associer soit 6 soit f soit f6

Associons le code numérique 1-2-3-4-5-6 aux différentes modalités du caractère et essayons de constituer un tableau à deux lignes, une réservée aux modalités du caractère et l'autre au nombre d'élèves correspondants. On dénombre ensuite le nombre de réponses correspondantes à chaque modalité ce qui nous donne le tableau suivant :

Modalité du caractère	1	2	3	4	5	6
Nombre d'élèves correspondants	12	1	9	5	12	1

Le nombre d'élèves de l'échantillon (40) est appelé **fréquence absolue total** ou **effectif total**.

2-2- Cas d'un caractère quantitatif appelé également variable statistique :

2-2-1- Caractère quantitatif discret (ou discontinu) :

Exemple :

Dans un nouveau quartier de la ville de Casablanca, une enquête est menée auprès d'un échantillon de 70 familles sur le nombre d'enfants par famille ayant un âge compris entre 5 et 12 ans pour juger de l'opportunité de la construction d'une école primaire dans ce quartier. Le nombre de familles parmi lesquelles on a choisi l'échantillon des 70 familles est de 1746.

Les résultats obtenus et consignés sont les suivants

0 - 1 - 1 - 5 - 6 - 4 - 3 - 6 - 0 - 5 -
4 - 3 - 2 - 1 - 6 - 4 - 3 - 7 - 0 - 6 -
5 - 4 - 6 - 1 - 7 - 4 - 5 - 3 - 6 - 7 -
5 - 4 - 6 - 7 - 3 - 2 - 4 - 6 - 3 - 1 -
6 - 5 - 4 - 6 - 0 - 3 - 7 - 2 - 2 - 1 -
3 - 4 - 4 - 5 - 6 - 3 - 2 - 1 - 6 - 7 -
6 - 4 - 3 - 5 - 6 - 2 - 1 - 3 - 2 - 6 -

- Exploitation :

On remarque que le caractère nombre d'enfants par famille ayant un âge compris entre 5 et 12 ans prend les valeurs suivantes : 0-1-2-3-4-5-6 et 7.

On peut dire que la variable statistique « nombre d'enfants par famille en âge compris entre 5 et 12 ans » ne prend que 8 valeurs sus-indiquées.

Pour une représentation plus claire des résultats obtenus, on va dresser un tableau en 3 colonnes :

Nombre d'enfants ayant un âge compris entre 5 et 12 ans	Nombre de familles	
	Résultat du dénombrement :	Effectifs :
0	IIII	4
1	IIIIIIII	8
2	IIIIIII	7
3	IIIIIIIIII	11
4	IIIIIIIIII	11
5	IIIIIIII	8
6	IIIIIIIIIIIIII	15
7	IIIIII	6
Total :		70

A chaque valeur de la variable statistique correspond le nombre de familles appelé **fréquence absolue ou l'effectif**.

L'effectif global est égal à 70.

Remarque :

Le dépouillement dans le cas du caractère quantitatif discret est le même que dans le cas d'un caractère qualitatif.

2-2-2- Caractère quantitatif continue ou variable statistique continue :

Exemple :

Le dénombrement effectué sur un échantillon de 60 exploitations agricoles prises au hasard a permis d'obtenir les superficies suivantes en hectares.

5,40 – 6 – 8,3 – 5 – 99 – 10,15 – 15,3 – 4 – 16 – 5,5 – 83,4 – 40 – 27,5 – 8 – 80,60 – 33,5 – 27 – 27,8 – 18,7 – 17
3 – 8 – 9,9 – 27,8 – 23 – 18,7 – 46 – 43 – 76 – 8,2 – 21 – 18,9 – 13 – 15 – 6,6 – 73,5 – 14,6 – 9,2 – 14 – 12 – 22,5
33,6 – 83 – 50 – 33 – 8,6 – 45 – 18 – 19 – 27,5 – 27 – 36 – 44,5 – 40,8 – 33,9 – 78,8 – 84,6 – 33,6 – 77,6 – 7,6

Exploitation :

- La variation statistique étudiée est la **superficie des exploitations agricoles**.
- Cette variable prend toutes les valeurs s'étalant entre 3 hectares (la plus petite valeur) et 99 hectares (la plus grande valeur). La différence arithmétique entre ces deux valeurs constitue **l'intervalle de variation** de la variable statistique.
- Le dépouillement et la présentation sous forme de tableau est pratiquement impossible dans la mesure où chaque valeur n'est répétée que rarement, en d'autres termes, le tableau statistique se trouve, dans ce cas composé d'une cinquantaine de lignes correspondant aux valeurs rencontrées.

- On procède à un partage de l'intervalle statistique en classes d'amplitude égale (amplitude est la différence entre la plus grande valeur de la classe et la plus petite valeur) ou inégale.
- Pour une simplification, on procède à l'adoption de classes d'amplitudes égales de telle sorte que le nombre de classes soit compris (pour la commodité de la représentation sous forme de tableau) entre 10 et 20.

- On obtient les classes suivantes :[0 à 10 [- [10 – 20 [- [20 – 30 [- [30 – 40 [- [40 – 50 [- [50 – 60 [- [60 – 70 [- [70 – 80 [- [80 – 90 [- [90 – 100 [.

- On procède ensuite à l'élaboration d'un tableau de 10 lignes (classes) et 3 colonnes : la colonne des classes, la colonne du dénombrement schématisé et la colonne des fréquences absolues ou effectifs.

- Donc la commodité de la représentation exige d'annuler ces deux classes.
- Soit en changement l'amplitude des classes tout en gardant la même amplitude en ce qui concerne l'ensemble des classes à savoir prendre une amplitude de 20, on aura des classes suivantes d'amplitude égale.

Classes (v.v.s) :	Résultat du dénombrement statistique :	Effectifs :
[0 à 10 [	IIIIIIIIIIIII	15
[10 à 20 [	IIIIIIIIIIII	14
[20 à 30 [	IIIIIIIII	10
[30 à 40 [	IIIIIII	8
[40 à 50 [	IIII	5
[50 à 60 [		0
[60 à 70 [		0
[70 à 80 [	IIII	4
[80 à 90 [	III	3
[90 à 100 [	I	1
Total		60

Mais dans ce cas, on n'aura que 5 classes.

Classes :	Effectifs :
[0 à 20 [	29
[20 à 40 [	18
[40 à 60 [	5
[60 à 80 [	4
[80 à 100 [	4

Paragraphe III : L'exploitation des observations sous forme de tableaux et les représentations graphiques usuelles:

3-1- Séries statistiques à un seul caractère :

3-1-1- Cas d'un caractère qualitatif :

Exemple 1 :

Pour tester la qualité des lampes électriques fabriquées par une machine, on a prélevé un lot de 1000 lampes et on a testé leur durée de vie. Le classement de la durée de vie des lampes a été effectué selon les modalités suivantes auxquelles on a fait correspondre un code chiffré.

Modalités*	Code
Très faible	1
Faible	2
Moyenne	3
Passable	4
Importante	5
Très importante	6

Les résultats obtenus sont consignés dans le tableau ci-après :

Modalités	Code	Nombre de lampes n_i
Très faible	1	125
Faible	2	140
Moyenne	3	200
Passable	4	400
Importante	5	100
Très importante	6	35
Total		1000

- Exploitation :

- le nombre d'observations de chaque modalité est l'effectif correspondant à cette dernière, exemple : pour « moyenne » l'effectif est de 200 lampes
- la somme des effectifs : $n_1 + n_2 + n_3 + n_4 + \ldots + n_6 = \sum_{i=1}^{i=6} n_i = 1000$
- la fréquence relative $f_i = \dfrac{n_i}{\sum n_i}$
  - La somme des fréquences = $f_1 + f_2 + f_3 + \ldots + f_6 = \sum_{i=1}^{i=n} f_i = 1$
  - Le pourcentage = $P_i = f_i \times 100$

- Si on ajoute les effectifs un à un, en commençant par le haut, on obtient les effectifs cumulés croissants (ECC)

Consignons dans un tableau les différents paramètres indications qu'on peut calculer dans le cadre d'une série à caractère qualitatif

modalités	Code	Nombre de lampes ei	ECC	ECD	Fréquences fi	FCC	FCD	Pourcentages
Très faible	1	120	125	1000	0,125	0,125	1	12,5
Faible	2	140	265	875	0,14	0,265	0,875	14
Moyenne	3	200	465	735	0,2	0,465	0,735	20
Passable	4	400	865	535	0,4	0,865	0,535	40
Très	5	100	965	135	0,1	0,965	0,135	10
Très importante	6	35	1000	35	0,035	1	0,035	3,5
Très faible		1000						100

- Lecture des effectifs cumulés croissants (ECC) : 465 lampes ont une durée de vie inférieure ou égale à la durée « moyenne ».
- Lecture des ECD : 535 lampes ont une durée de vie au moins « passable »
- Lecture des FCC : 0,265 lampes ont une durée de vie au plus faibles et très faible.
- Lecture des FCD : 0,135 lampes ont une durée de vie au moins importante ou importante et très importante.

Graphique :
Faites le graphique en tyaux d'orgue (voir TP)

3-1-2- Cas d'un caractère quantitatif :

3-1-2-1- Variable discrète :

Exemple 1 :

La distribution statistique suivante est le résultat d'une enquête faite auprès de 4268 ménages sur le nombre de personnes par ménage

Nombre de personnes par ménage V.V.S :	Nombre de ménages
1	57
2	340
3	506
4	735
5	850
6	830
7	470
8	220
9	190
10 et plus	70
Total	4268

- 1, 2, 3, 4, …. Constituent les valeurs de la variable statistique appelée xi
ex : $x_i = 1$; $x_5 = 5$; $x_7 = 7$
- L'intervalle de variation de la variable statistique est compris entre 1 et plus de 10 (indéterminé dans sa limite supérieure).
- Le nombre d'observations de chaque valeur de la variation statistique xi est appelé fréquence absolue (f_{1a}) ou effectif (symbolisé par ni)

- La somme des effectifs = $n_1 + n_2 + n_3 + n_{10} = \sum_{i=1}^{i=10} n_i$

$$57 + 340 + 506 + \ldots + 70 = 4268$$

Remarque : $\sum_{i=1}^{n} n_i = N$ et $\sum_{i=1}^{n} f_i = 1$

Donc, la somme des fréquences est égale à 1.
- On peut calculer à partir de la fréquence les pourcentages en multipliant chaque fréquence par 100.
$p_1 = f_1 \times 100$ $p_6 = f_6 \times 100$
$= 0{,}013 \times 100 = 1{,}3\%$ $x = 0{,}19 \times 100 = 19\%$
- Si on ajoute les effectifs et un à un en partant du haut, on obtient les effectifs cumulés croissants (ecc).

- Si on ajoute les effectifs ei un à un en partant du bas, on obtient les effectifs cumulés décroissants (ecd).
- On peut également obtenir les fréquences cumulées croissantes (fcc) et les fréquences cumulées décroissantes (fcd) en ajoutant les fréquences cumulées décroissantes (fcd) en ajoutant les fréquences une à une soit en commençant par le haut (fcc) soit par le bas (fcd).

Consignons dans un tableau les différents paramètres et indicateurs que l'on peut calculer dans le cadre d'une série statistique.

Nombre de personnes par ménage (vvs) x_i	Nombre de ménages (effectifs) e_i	Effectifs cumulés croissants	Effectifs cumulés décroissants	Fréquence f_i calculée	Fréquence f_i arrangée	Fréquence cumulées croissantes	Fréquence cumulées décroissantes
1	57	57	4268	0,01335	0,01	0,01	1
2	340	397	4211	0,07966	0,08	0,09	0,99
3	506	903	3871	0,11855	0,12	0,21	0,91
4	735	1638	3365	0,17221	0,17	0,38	0,79
5	850	2488	2630	0,19915	0,20	0,58	0,62
6	830	3318	1780	0,19447	0,19	0,77	0,42
7	470	3788	950	0,11012	0,11	0,88	0,23
8	220	4008	480	0,05154	0,05	0,93	0,12
9	190	4198	260	0,04451	0,05	0,98	0,07
10	70	4268	70	0,01614	0,02	1	0,02
	4268			0,99996	1,00		

Remarque :

Nous retenons simplement deux chiffres après la virgule quant au calcul de la fréquence arrangée de telle sorte que la somme soit égale à 1.

- Lecture des effectifs cumulés croissants, décroissants et des fréquences cumulées croissantes et décroissantes.

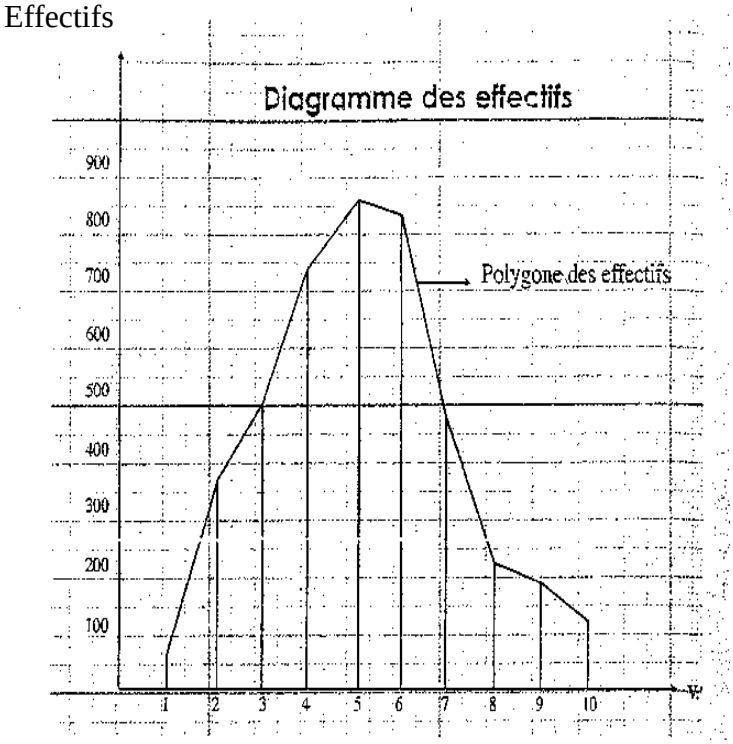
* ECC ➔ 4ème valeur : il y a 1638 ménages qui sont constitués au plus de 4 personnes, ou le nombre de personnes est inférieur ou égal à 4.
* ECD ➔ 6ème valeur : il y a 1780 ménages qui sont constitués au moins de 6 personnes, ou le nombre de personnes est supérieur ou égal à 6.
* FCC ➔ 3ème valeur : il y a 0,21 ménages qui sont constitués au plus de 3 personnes, ou le nombre de personnes est inférieur ou égal à 3.
* FCD ➔ 7ème valeur : il y a 0,23 ménages qui sont constitués au moins de 7 personnes, ou le nombre de personnes est supérieur ou égal à 7.

Représentation graphique :

1- Diagramme des effectifs :

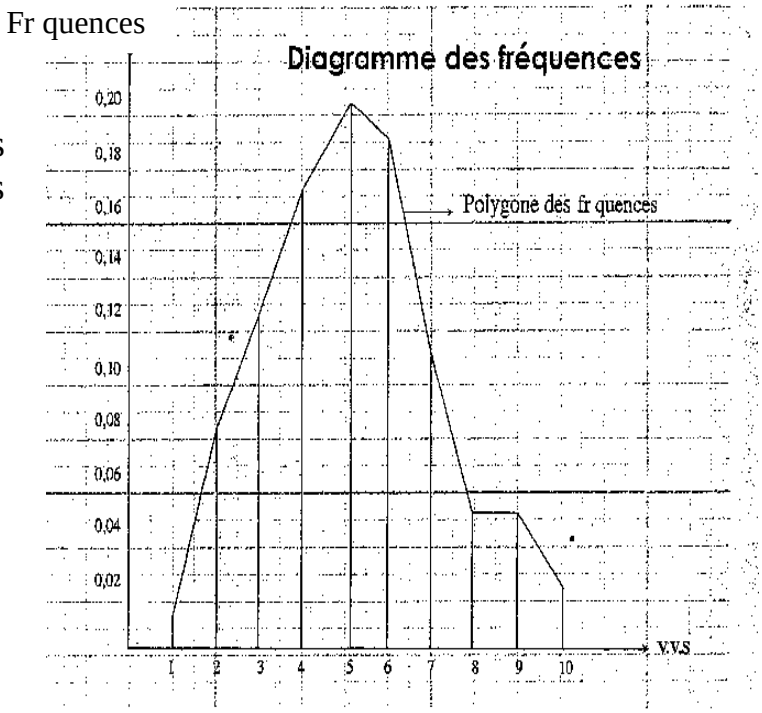
La représentation est faite à l'aide d'un repère orthonormé (axes perpendiculaire), les V.V.S sont placées sur l'axe des x et les différents effectifs sur l'axe des y. Chaque valeur de la variable statistique est représentée par un bâton proportionnel à l'effectif.

- La représentation est faite au moyen de bâtonnets proportionnels aux effectifs.
- On peut joindre les têtes des bâtonnets et on obtient ainsi un polygone des effectifs.



2- Diagramme des fréquences :

Le même principe est retenu pour le diagramme des fréquences avec la différence que l'axe des y contient les fréquences et celui des x les effectifs.



Remarque :

Le polygone des effectifs et des fréquences ont la même allure du fait que les fréquences sont proportionnelles aux effectifs.

3-1-1—Variable continue :

Exemple 1 : classes de même amplitude de variation :

On a recensé, dans une ville, 3000 ménages au cours de l'année 2005 et on a consigné l'âge des épouses. Les résultats du dépouillement nous ont permis d'établir la série statistique suivante (Tableau ci-contre) :

Age de l'épouse (année révolue) **classe** :	Nombre de ménage :
De 18 à moins 22 ans	150
De 22 à moins 26 ans	750
De 26 à moins 30 ans	600
De 30 à moins 34 ans	450
De 34 à moins 38 ans	330
De 38 à moins 42 ans	300
De 42 à moins 46 ans	210
De 46 à moins 50 ans	120
De 50 à moins 54 ans	60
De 54 à moins 58 ans	30
Total	3000

Remarque :

- La classe de 18 à moins 22 ans, par exemple, peut s'écrite à l'aide d'un intervalle fermé à gauche et ouvert à droite [18-22[.
- Le domaine de variation de la variable statistique est la différence entre la plus grande valeur de la variable et la plus petite valeur.
- Le domaine de variation de la variable statistique = 58 – 18 = 40
- Le nombre de classes est égal à 10 et l'amplitude de variation de chaque classe est 40/10=4 ou 22-18= 4

  Dans l'exemple, on remarque que les différentes classes ont la même amplitude.

Remarque :

Au lieu de 10 classes à amplitude égale, on aurait pu réduire leur nombre ou l'augmenter et opter pour des classes à amplitudes inégales. En effet, le choix du nombre des classes et leur amplitude (égale ou inégale) dépend du phénomène étudié et du choix du statisticien.

- Calcul du centre de chaque classe ou de la variable statistique X_i

$$1^{ère} \text{ classe : le centre de la classe} = \frac{\chi\text{'(borne inférieure)} + X^i \text{ (borne supérieure)}}{2} = \chi_1$$

$$c_i = \frac{18 + 22}{2} = 20$$

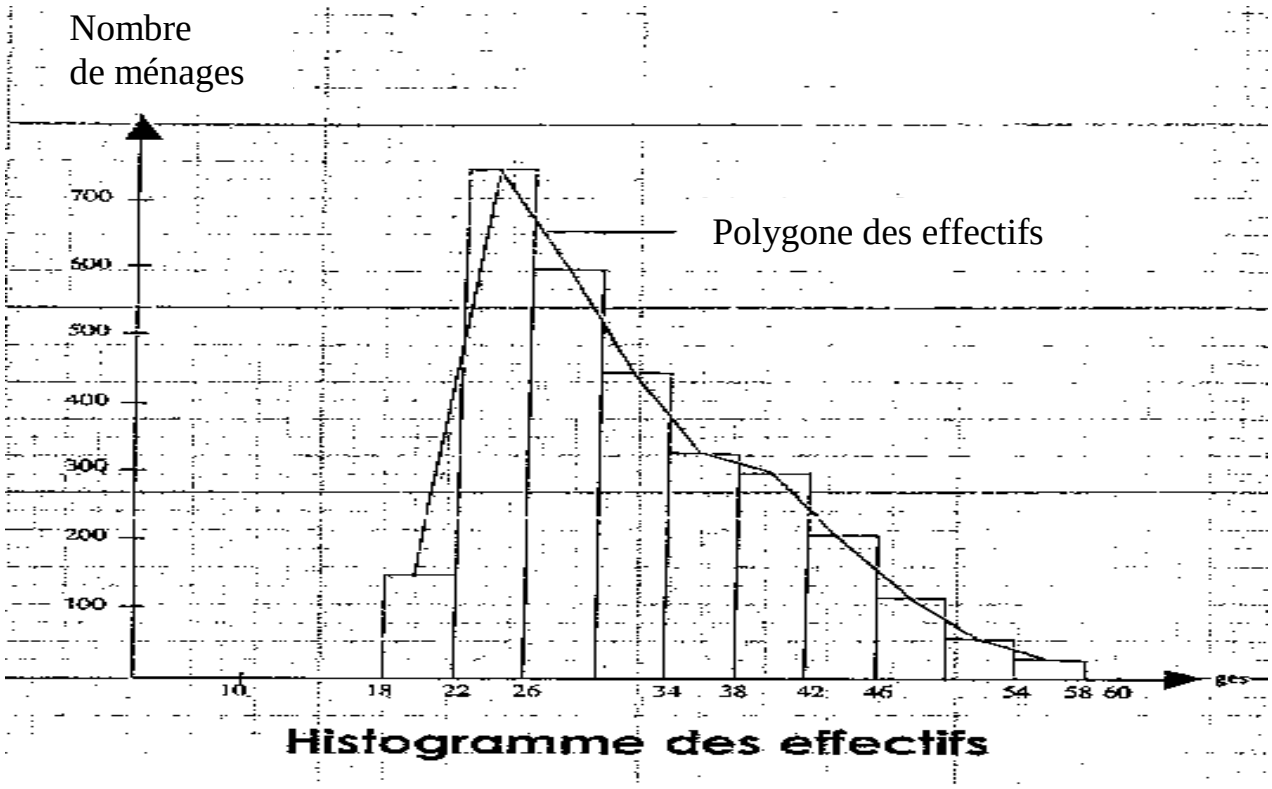
- Calcul et récapitulation dans un tableau des différentes valeurs de la variable statistique, de l'ECC (effectifs cumulés croissants), de l'ECD (fréquences cumulées décroissantes), de la fréquence, de l'FCC (fréquences cumulées croissants), de l'FCD (fréquences cumulées décroissants) et des pourcentages.

Tableau statistique :

Age de l'épouse **classe**	Nombre de ménages **ni**	Centre des classes **ci**	ECC	ECD	Fréquence fi	FCC	FCD	Pourcen-tages
De 18 à moins 22 ans	150	20	150	3000	0,05	0,05	1	5
De 22 à moins 26 ans	750	34	900	2850	0,25	0,30	0,95	25
De 26 à moins 30 ans	600	28	1500	2100	0,20	0,50	0,70	20
De 30 à moins 34 ans	450	32	1950	1500	0,15	0,65	0,50	15
De 34 à moins 38 ans	330	36	2280	1050	0,11	0,76	0,35	11
De 38 à moins 42 ans	300	40	2580	720	0,10	0,86	0,24	10
De 42 à moins 46 ans	210	44	2790	420	0,07	0,93	0,14	7
De 46 à moins 50 ans	120	48	2910	210	0,04	0,97	0,07	4
De 50 à moins 54 ans	60	52	2970	90	0,02	0,99	0,03	2
De 54 à moins 58 ans	30	56	3000	30	0,01	1	0,01	1
	3000							100%

Représentation graphique :

A partir du moment où l'amplitude des classes est la même, la représentation graphique ne pose aucun problème. Il suffit de placer les VVS sur l'axe des abscisses et les effectifs correspondants sur l'axe des ordonnées.



Remarques :

* On obtient un histogramme des effectifs (un ensemble de rectangle côté à côté).
* La hauteur de chaque rectangle est proportionnelle à l'effectif représenté.
* Si on fait le centre des classes on obtient le polygone des effectifs.

Exemple 2 : Classe d'amplitudes différentes :

Une enquête menée auprès de l'ensemble des salariés d'une entreprise sur leur salaire horaire en dirhams nous a permis d'établir la série statistique suivante :

N° d'ordre	Salaires horaires classés	Membre de salariés e_i
1	[8,50- 10 [	40
2	[10 – 11,50 [	36
3	[11,5 - 16 [	32
4	[16 - 22 [	30
5	[22 - 28 [	28
6	[28 -34 [	26
7	[34 - 40 [	22
8	[40 - 46 [	20
9	[46 – 58 et plus	12
Total		246

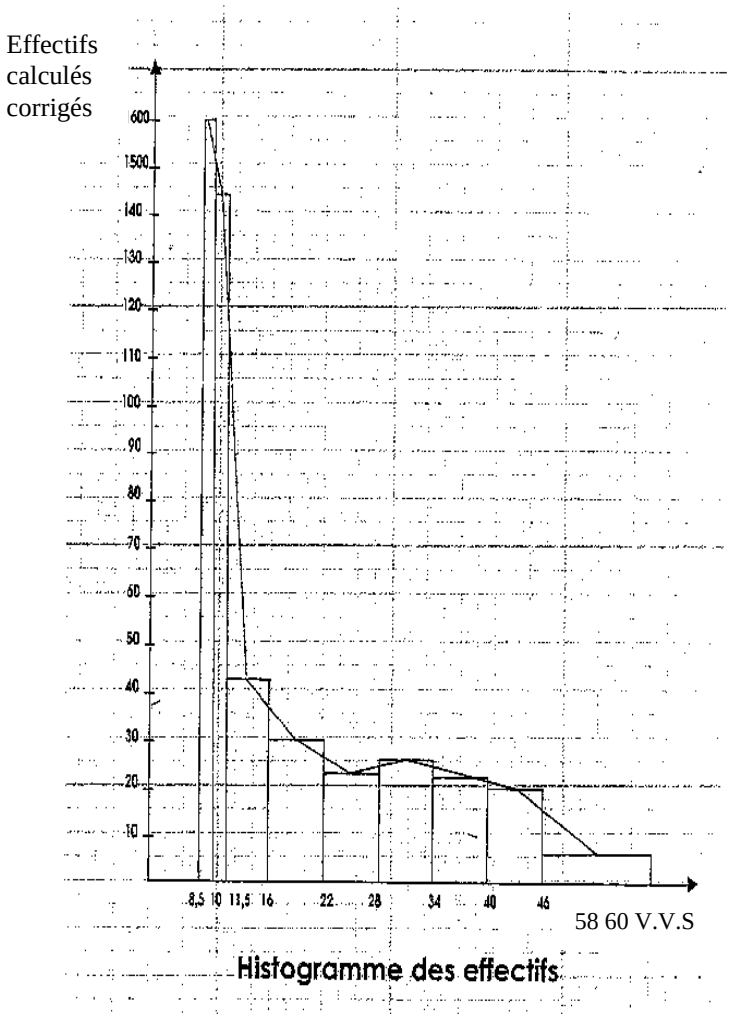
- **Tableau récapitulatif**

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Classes	Effectifs n_i	ECC	ECD	Fréquence fi	FCC	FCD	Pourcentages P_i	Effectifs corrigés
[8,50- 10 [	40	40	246	0,162	0,162	1	16,2	160
[10 – 11,50 [	36	76	206	0,146	0,308	0,838	14,6	144
[11,50 - 16 [	32	108	170	0,130.	0,438	0,692	13	42,66
[16 - 22 [	30	138	138	0,122+	0,560	0,562	12,2	30
[22 - 28 [	28	166	108	0,114+	0,674	0,440	11,4	28
[28 -34 [	26	192	80	0,106+	0,780	0,326	10,6	26
[34 - 40 [	22	214	54	0,089	0,869	0,220	8,9	22
[40 - 46 [	20	234	32	0,082	0,951	0,131	8,2	20
[46 – 58 [	12	246	12	0,049+	1	0,049	4,9	6
	246						100%	

- Représentation graphique de la série statistique :

* Prenons un repère cartésien orthogonal, plaçons les points correspondants aux limites de chaque classe sur l'axe des abscisses (points déterminants des segments de longueurs différentes).
* Sur l'axe des y on place les effectifs corrigés.

Conclusion :



3-1-3- Les séries chronologiques

Exemple 1 :

Le tableau suivant retrace l'évolution de la population marocaine en millions depuis 1960.

Années	1960	1971	1982	1994	2004
Population	11,6	15,3	20,4	26,7	29,9

Source Bank Al-Maghrib : 1960 – 71 – 82 – 94 – 04

- La série est appelée chronologique parce que l'une des variables est constituée par le temps
- plusieurs calculs peuvent être effectués sur cette série statistique :

à titre d'exemple : * l'augmentation de la population en valeur absolue

Exemple de 1960 à 2004 ➔ 29,9 – 11,6 = 18,3 millions

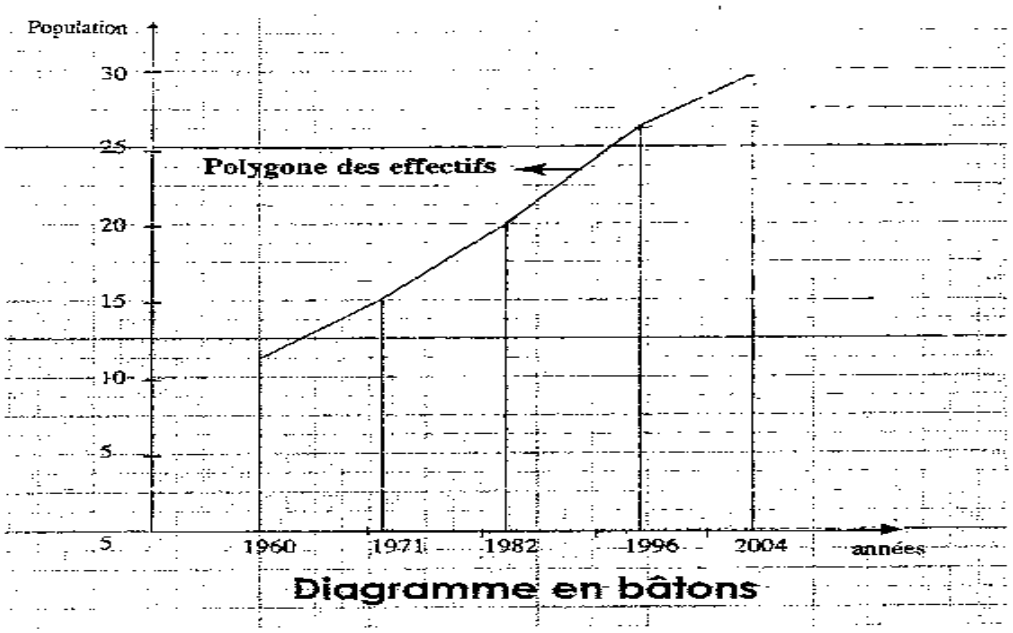
* l'augmentation de la population en valeur relative

Exemple : de 1994 à 2004 ➔ $\frac{29,9 - 26,7}{26,7} = 0,119$

Représentation graphique :

L'axe des contiendra les valeurs du temps et l'axe des y le nombre d'habitants.

Mais, il est à constater que le phénomène « connaissance de la population marocaine » s'est fait à une date précise qu'on peut considérer comme étant la fin de l'année. De ce fait, la représentation graphique sera comme dans le cas d'une variable discrète à l'aide de bâtonnets prenant leur départ du point temps considéré. Et en reliant l'ensemble des points on obtient le polygone d'évaluation.



3-2- Les séries statistiques à deux caractères :

Exemple :

Une enquête faite auprès d'un échantillon de 60 élèves d'un lycée qualifiant sur leur poids et leur taille a donné les résultats suivants.

(45-15) ; (60-162) ; (10-165) ; (80-190) ; 67-170) ; (70-175) ; (68-185) ; (57-162) ; (89-15) ; (95-180) ; (62-168) ; (72-178) ; (82-160) ; (92-180) ; (45-155) ; (57-155) ; (67-160) ; (77-182) ; (87-187) ; (95-192) ; (45-155) ; (55-158) ; (65-162) ;(75-180) ; (85-170) ; (51-159) ; (61-145) ; (71-160) ; (81-190) ; (91-184) ; (56-160) ; (66-155) ; (76-178) ; (86-190) ; (94-184) ; (92-190) ; (88-170) ; (86-192) ; (80-187) ; (50-158) ; (52-158) ; (7-158) ; (81-167) ; (92-187) ; (68-172) ; (45-155) ; (60-170) ; (71-188) ; (80-185) ; (76-178) ; (70-167) ; (75-172) ; (86-190) ; (47-160) ; (66-171) ; (90-195) ; (86-192) ; (67-180) ; (55-160) ; (49-155)

Remarque 1 :

Les résultats sont obtenus à partir d'un questionnaire et son présentés sous forme d'un couple de valeurs (xi,yi). Le 1er est le poids exprimé en Kg le deuxième est la taille exprimée en centimètre.

Remarque 2 :

Le classement est effectué en fonction de deux variations. On obtient un tableau statistique à double entrée avec soit :

-deux caractères qualitatifs

-deux caractères quantitatifs (c'est le cas de l'exemple étudié)

* tous les deux discrets
* tous les deux continus
* un discret et un continu

- un caractère qualitatif et un caractère quantitatif

Présentation des données obtenues sous forme de tableau statistique à double entrée :

- Déterminons tout d'abord l'intervalle de variation de chacune des variables :

* le poids varie entre 45 et 95 Kg l'intervalle est égal à 95 – 45 = 50.

Nous allons partager cet intervalle par 5 et obtenons 5 classes de 10 Kg d'amplitude chacune.

* la taille varie entre 195 et 155 centimètres, l'intervalle de variation est égal à 195 – 155 = 40 centimètres, partageons cet intervalle en 8 classes dont l'amplitude est égale à 5 colonnes.

Remarque :

Le choix du nombre de classes et des amplitudes est généralement arbitraire.

- Dressons un tableau à double entrée ou figurera les deux caractères, un en ligne et l'autre en colonne.

Poids / Tailles	[45-55[	[55-65[	[65-75[	[75-85[	[85-95[	Total
[155-160[						
[160-165[						
[165-170[						
[170-175[						
[175-180[						
[180-185[						
[185-190[						
[190-195[						
Total						

- Commençons par marquer tous les couples de nombre, dont le poids est compris entre 45 et 55, vérifions la taille en excluant tous ceux qui ont une taille non comprise entre 155 et 160 et dénombrons le reste qui doit être porté sur le tableau à la case correspondante, on obtient la répartition suivante :

Poids / Tailles	[45-55[	[55-65[	[65-75[	[75-85[	[85-95[	Total
[155-160[	8	3	2	-	-	13
[160-165[	1	4	3	1	-	9
[165-170[	-	1	2	1	-	4
[170-175[	-	1	3	1	2	7
[175-180[	-	-	2	2	-	4
[180-185[	-	-	1	2	3	6
[185-190[	-	-	2	2	2	6
[190-195[	-	-		1	10	11
Total	9	9	15	10	17	60

Lecture des résultats

-Quel que soit leur poids, 13 élèves ont une taille comprise entre 155 et 160 centimètres.

-Quel que soit leur taille 9 élèves ont un poids compris entre 45 et 55 Kg.

Remarque 1 :

La série à double entrée peut donner lieu à deux séries statistiques simples dont l'une étudie la taille et l'autre le poids.

Tailles	Effectifs
Reprendre les classes	13 9 4 7 4 6 6 11
	60

Poids	Effectifs
Reprendre les classes	9 9 15 10 17
	60

3-3- Séries statistiques à plus de deux caractères :

Exemple :

Population active âgée de 15 ans et plus occupée et en chômage en milliers de personnes						
Population	Milieu urbain		Milieu rural		Total	
	2003	2004	2003	2004	2003	2004
Population occupée	4343	4533	5141	5289	9484	9822
Population en chômage	1041	1021	182	172	1223	1193
Total	5384	5554	5323	5461	10707	11015

D'après rapport Bank AL Maghreb

Remarque :

Etant donné l'existence de 3 caractères (Milieu, années, population), on a été amené à dresser un tableau à double entrée avec les différentes modalités des caractères.

Cette série peut donner à 6 séries statistiques :

- la population occupée et en chômage en milieu urbain en 2003
- la population occupée et en chômage en milieu urbain en 2004
- la population occupée et en chômage en milieu urbain en 2003
- la population occupée et en chômage en milieu rural en 2004
- la population occupée et en chômage total en 2003
- la population occupée et en chômage total en 2004

Paragraphe IV : Les représentations graphiques :

4-1-Représentation graphique dans le cas d'un caractère qualitatif (voir paragraphe III)

4-2-Représentation graphique dans le cas d'un caractère quantitatif (voir paragraphe III)

Remarque :

Les séries chronologiques qui retracent l'évolution d'un phénomène dans le temps ou de deux ou de plusieurs peuvent être représentées à l'aide d'un graphique cartésien où le temps figure sur l'axe des abscisses et les valeurs prises par le caractère sur l'axe des y (on peut présenter 2 ou plusieurs phénomènes) à condition que l'unité de compte soit la même.

Exemple 1 : représenter sur le même graphique la masse monétaire et le produit (P.I.B) en dirhams pendant un certain nombre d'années pour constater leur évolution.

4-3- Les graphiques non cartésiens

4-3-1- Représentation à l'aide de graphiques circulaires :

Les représentations à l'aide de graphiques circulaires peuvent être adoptées dans le cas d'une variable, statique continue et discontinue ou dans le cas d'un caractère qualitatif avec un nombre de modalités assez réduit (ne dépassent pas en général 10).

Les représentations circulaires peuvent jouer deux rôles :

- mettre en relief les composantes d'un phénomène ;
- comparer l'évolution d'un phénomène dans le temps ou ses composantes.

4-3-1-1- Représentation des composantes d'un phénomène

Exemple :

Soit le tableau suivant retraçant les grandes des recettes ordinaires du Trésor en 2004 en millions de dirhams.

Recettes ordinaires	2004
Recettes fiscales	91.219
Recettes non fiscales	14.939
Recettes de certains comptes spéciaux	3.228
Total	109386

Rapport Bank Al-Maghreb

Remarque :

La représentation se fait à l'aide de secteurs (portions d'un cercle) proportionnels aux valeurs du caractère, la surface des secteurs est elle aussi proportionnelle au caractère, ce qui nous donne dans ce cas 109386 fi 360°. Il suffit par la suite de calculer la mesure en degrés de chaque valeur de la modalité du caractère.

Pour 91219 correspondant à x1 donc

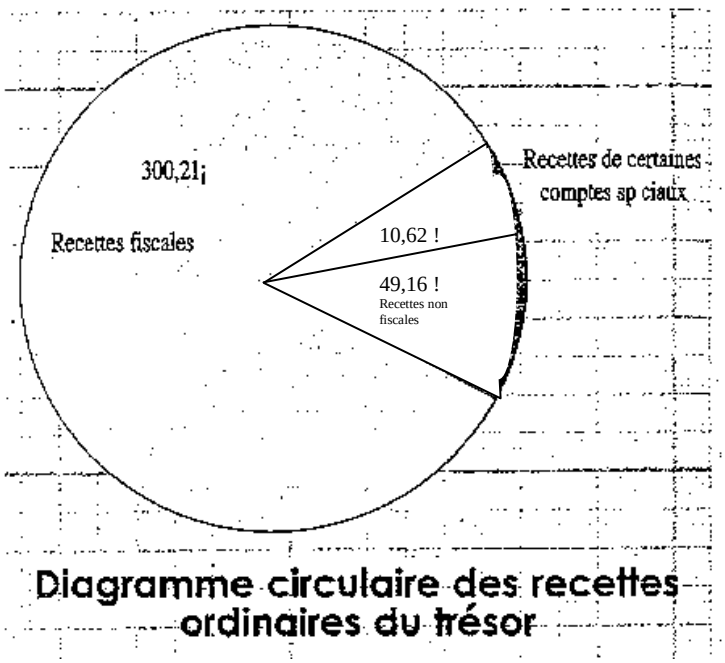
$$\chi_1 = \frac{360 \times 91219}{109386} = 300{,}21°$$

$$\chi_2 = \frac{306 \times 14939}{109386} = 49{,}16°$$

$$\chi_3 = \frac{360 \times 3228}{109386} = 10{,}62°$$

Donc les angles au centre représente chacun 300,21° , 49,16° et 10,62°.

A l'aide d'un rapporteur on procède à la représentation graphique.



Remarque :

Il serait intéressant de faire figurer sur les secteurs les pourcentages et non pas les valeurs absolues.

4-3-1-2- Comparer l'évolution d'un phénomène et ses composantes dans le temps :

Le principe consiste à tenir compte de l'évolution du phénomène lui-même en faisant en sorte que les rayons des certes soient proportionnels à la valeur globale du phénomène.

Exemple

Soit le tableau suivant, retraçant l'évolution des recettes fiscales en millions de dirhams entre 2003 et 2004 en millions de dirhams.

Recettes fiscales	Valeurs	
	2003	2004
Impôts directs	33363	36468
Droits de douane	10670	11412
Impôts indirects	41890	43901
Enregistrement et timbres	5296	5505
Total	91219	97286

Rapport Bank Al-Maghrib

- Le premier travail consiste à calculer le rayon de chaque cercle d'une manière proportionnelle à la valeur du caractère.

Si on suppose que le rayon du cercle 1 (2003)= 3 cm
Le rayon du cercle 2 (2004) sera égal à 91219 ➔ 3 cm
97286 ➔ x

$$\chi = \frac{97286 \times 3}{91219} = 3{,}20 \text{ cm}$$

- Le deuxième travail réside dans le calcul des différentes les valeurs des angles au centre pour 2003/2004, on obtient :

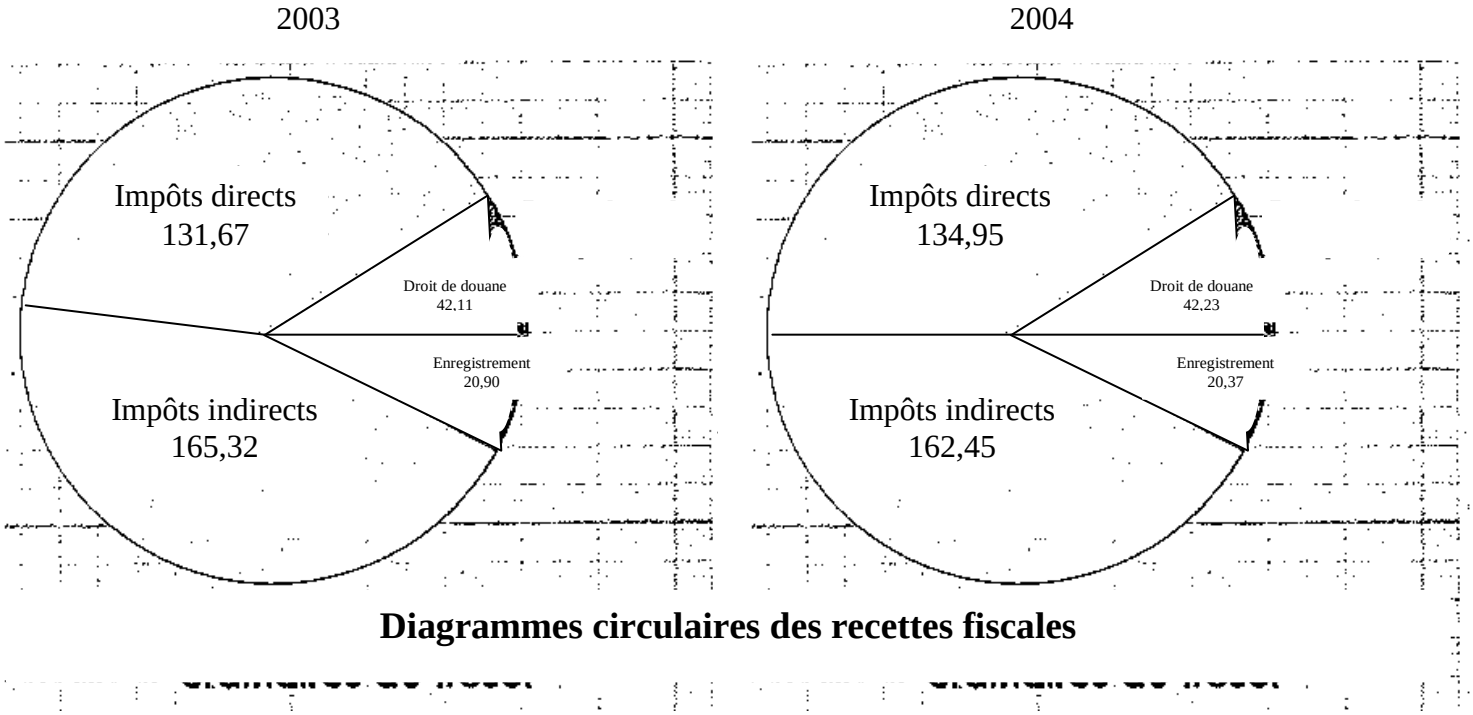
Années / Recettes fiscales	Mesure des angles au centre en degrés	
	2003	2004
Impôts directs	131,67	134,95
Droits de douane	42,11	42,23
Impôts indirects	165,32	162,45
Enregistrement et timbres	20,9	20,37
Total	360°	360°

Source : Rapport Bank Al-Maghrib 2004

Exemple de calcul :

Impôts directs en degrés $\frac{33363 \times 360}{91219} = 131{,}66\ \%$

Il suffit de reporter les mesures des angles au centre, respectivement pour 2003 et 2004.



Diagrammes circulaires des recettes fiscales

4-3-2- Représentation à l'aide des graphiques semi-circulaires

La représentation à l'aide d'un graphique semi circulaire obéit au même principe que pour la représentation d'un graphique circulaire, la différence réside dans le fait que l'ensemble des angles au centre mesure 180° au lieu de 360°. De ce fait, l'angle correspondant à l'effectif total serait égal à 180° et celui correspondant à l'effectif d'une modalité serait égal à $180 \times \frac{i}{\sum n_i}$

Exemple :

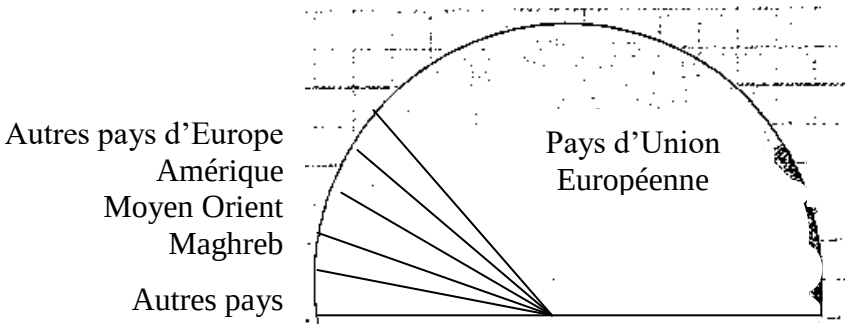
Le tableau suivant retrace les entrées des touristes de séjour au Maroc en 2004 suivant l'origine

Pays d'origine	Nombre de touristes
Pays d'Union Européenne	2236992
Autres pays d'Europe	112564
Amériques	127974
Moyen Orient	74873
Maghreb	81969
Autres pays	112975
Total	2747347

Annuaire statistique du Maroc 2004

Après calcul des différents angles au centre, on obtient le tableau suivant

Pays d'origine	Nombre de touristes	Mesure des angles au centre en degré
Pays d'Union Européenne	2236992	146,56
Autres pays d'Europe	112564	7,38
Amériques	127974	8,38
Moyen Orient	74873	4,91
Maghreb	81969	5,37
Autres pays	112975	7,40
Total		180°



Graphique semi-circulaire

4-3-3- Représentation graphique à l'aide d'un diagramme triangulaire* :

Le diagramme triangulaire permet la représentation d'une valeur composée de 3 modalités.

Exemple 1 :

On peut représenter la décomposition du PIB en 3 composantes :

- le PIB du secteur primaire
- le PIB du secteur secondaire
- le PIB du secteur tertiaire

Exemple 2 :

La décomposition prix de revient en :

- coût de matières premières
- coût de la main d'œuvre
- part des frais généraux

Exemple :

Soit le tableau ci-après constatant les ressources ordinaires (recettes fiscales, recettes non fiscales, recettes de certains comptes spéciaux) du Trésor au cours de l'année 2004 en millions de dirhams.

Remarque :

Avant d'entamer la représentation graphique, il est indispensable de transformer les valeurs des modalités du caractère en pourcentages.

Ressources	Janvier – Décembre 2004	Pourcentages
Recettes fiscales	97286	83,43
Recettes non fiscales	15761	13,52
Recettes de certains comptes spéciaux	3557	3,05
Total	116604	100

Annuaires statistique du Maroc 2004

1 er travail à faire : calcul des pourcentages (voir tableau)

2ème travail à faire : traçons un triangle équilatéral dont la hauteur : 5 cm mesure ce qui revient à choisir des côtés égaux à :

$$\text{Si } h = \frac{a\sqrt{3}}{2} \Rightarrow a = \frac{2h}{\sqrt{3}} \Rightarrow a = \frac{10}{1{,}732} = 5{,}77 \text{ cm}$$

- On peut également tracer un triangle équilatéral d'une manière géométrique.

3ème Travail à faire : graduons les 3 hauteurs du triangle de 0 à 100

4ème Travail à faire : prenons sur la hauteur AN le point des recettes fiscales et traçons une parallèle à BC passant par le point porté sur la hauteur AN
Faisons la même chose pour la hauteur BL et CM

5ème Travail à faire : du point d'insertion I menant les hauteurs aux trois côtés du triangle.

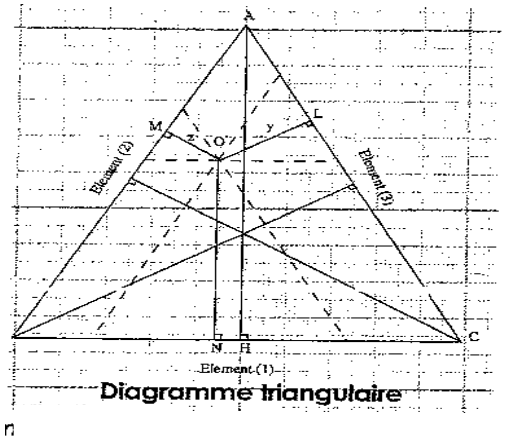


Diagramme triangulaire

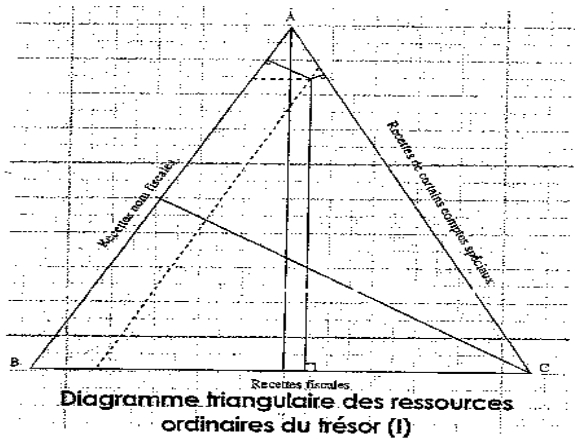


Diagramme triangulaire des ressources ordinaires du trésor (I)

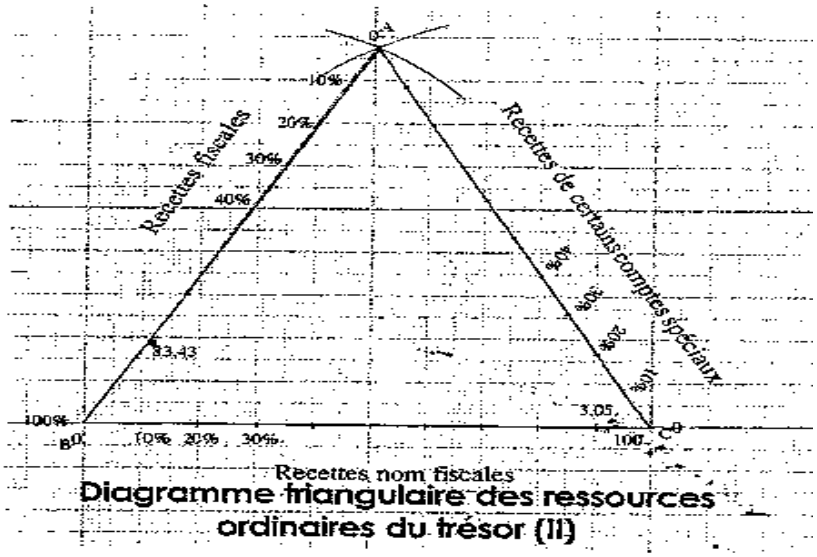


Diagramme triangulaire des ressources ordinaires du trésor (II)

Remarque 1 :

Il serait possible de procéder d'une autre manière en portant les valeurs de la modalité du caractère non pas sur la hauteur mais sur les côtés. De ce fait on choisit des côtés de 100 unités mesures.

On peut tracer la base du triangle BC = 100 et à l'aide d'un compas on détermine les 2 autres côtés en reportant la mesure de la base sur l'ouverture du compas et en traçant des portions de cercle de chaque côté de la base.

Sur le côté AB, on porte la part des recettes fiscales.

Sur le côté BC, on porte la part des recettes non fiscales.

Sur le côté AC, on porte la part des recettes de certains comptes spéciaux.

NB: les côtés doivent être gradués dans le même sens de 0 à 100.

4-3-4- Représentation graphique à l'aide de diagrammes à barres

Les barres sont des rectangles dont les hauteurs sont proportionnelles aux valeurs à représenter et les bases sont choisies arbitrairement mais doivent être de mesure égale.

Les barres permettent une représentation assez variée, on peut :

- Utiliser une ligne de base verticale ou horizontale. Les rectangles peuvent être juxtaposés ou séparés. Ils peuvent représenter le phénomène et ses composantes également.
- Représenter des valeurs positives et négatives en même temps.
- Mettre en évidence l'évolution de plusieurs phénomènes à la fois.

Exemple 1 :

Le tableau suivant représente le taux de scolarisation dans les différentes régions du Maroc des élèves de 15 à 17 ans en pourcentage.

Régions	Oued Ed-dehab Lagouira	Laayoune Boujdour Sakia El Hamra	Guelmim Es-mara	Souss Massa-Daraa	Gharb CHrarda Beni hssein	Chaouia Ourdigha	Marrakech-Tansift-Al Haouz	L'oriental
Taux	76,02	63,80	60,07	42,54	33,81	37,69	30,20	40,10
Régions	Grand Casa	Rabat-Salé Zammour-Zaer	Doukkala Abda	Tadla Azilal	Maknèse Tafilalt	Fès-Boulmane	Taza AlHouceima Taounate	Tanger-Tetouan
Taux	68,86	66,83	31,36	36,68	49,39	51,39	27,58	33,70

Bulletin d'éducation et de formation n 3(janvier-février 2005)
M.E.N.E.S.F.C.R.S

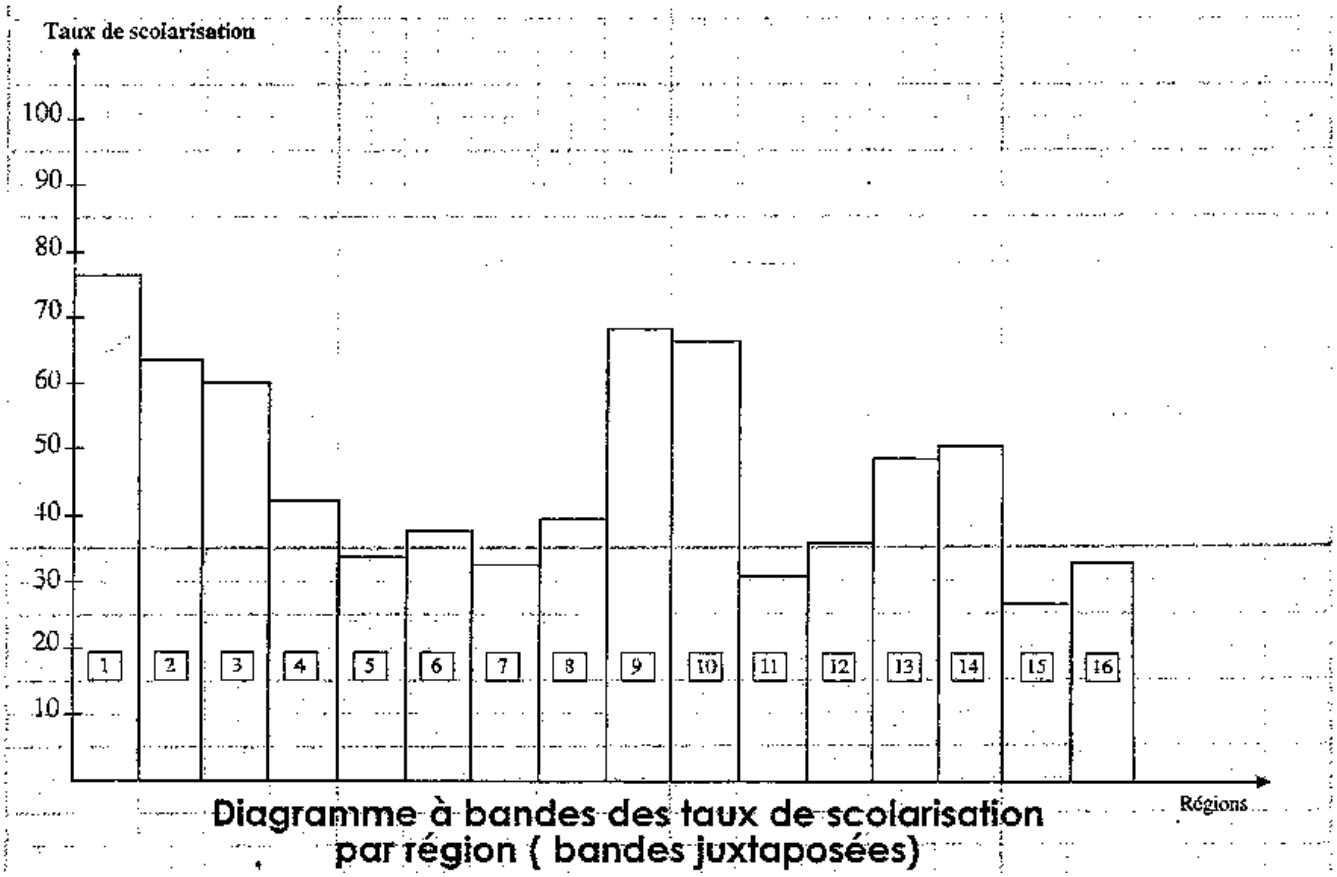


Diagramme à bandes des taux de scolarisation par région (bandes juxtaposées)

Remarque :

La représentation est faite sur la base de lignes horizontales avec des barres juxtaposées ayant une longueur proportionnelle aux mesures du phénomène étudié.

Les barres peuvent être également séparées.

La ligne de base peut être verticale et à ce moment c'est l'axe des x qui sera gradué de 0 à 100.

Exemple 2 :

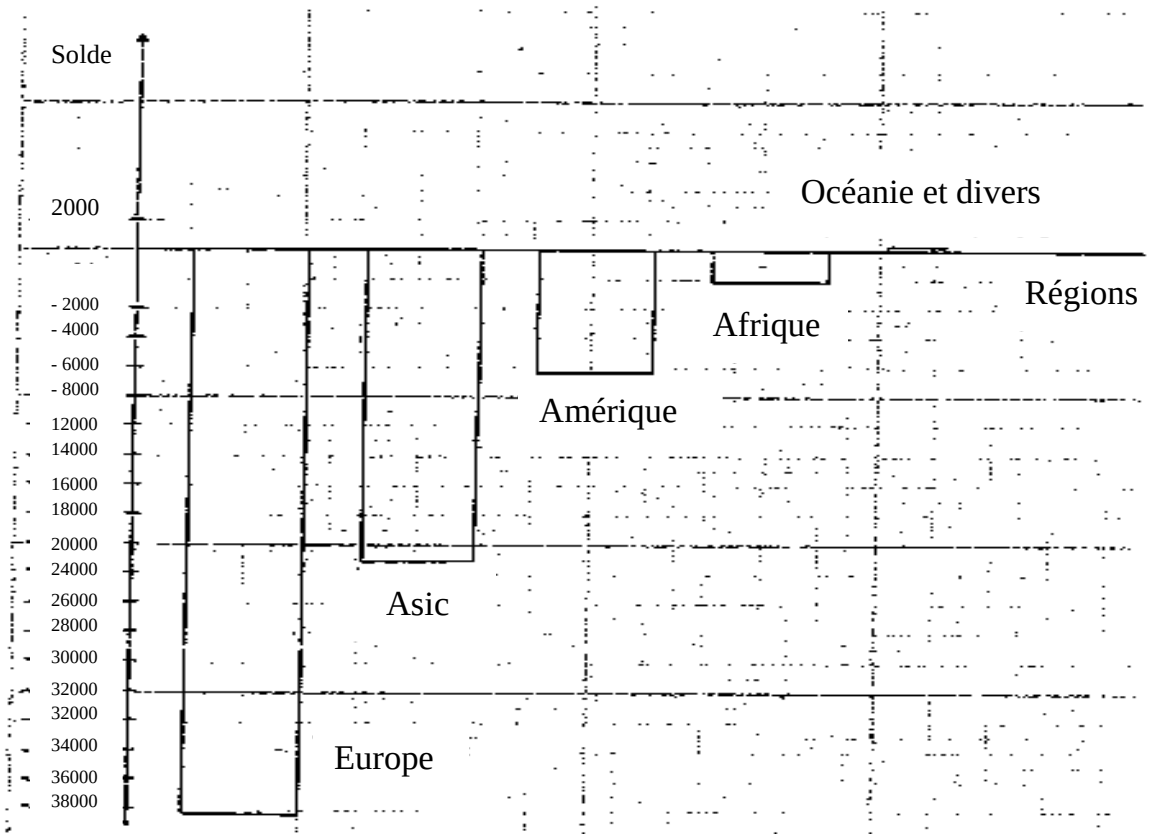
Le tableau suivant représente le solde de la balance commerciale en millions de dirhams.

Régions	Solde en 2004
Europe	-38431
Asie	-21124
Amérique	-8494
Afrique	-2061
Océanie et divers	+178
Total	-69 932

Rapport Bank Al-Maghrib 2004

Remarque :

La représentation est faite en valeurs absolues avec des barres séparées.



Exemple 3 :

Le tableau suivant indique la composition des prévisions de dépenses du budget général de l'Etat en millions de dirhams en 2002, 2003 et 2004.

Prévisions des dépenses	2002	2003	2004
Dépenses de fonctionnement	75442	78250	81047
Dépenses de la dette	46708	41759	41626
Dépenses d'équipement	19925	19547	191868
Total	142075	13956	141868

Bank Al-Maghrib 2004

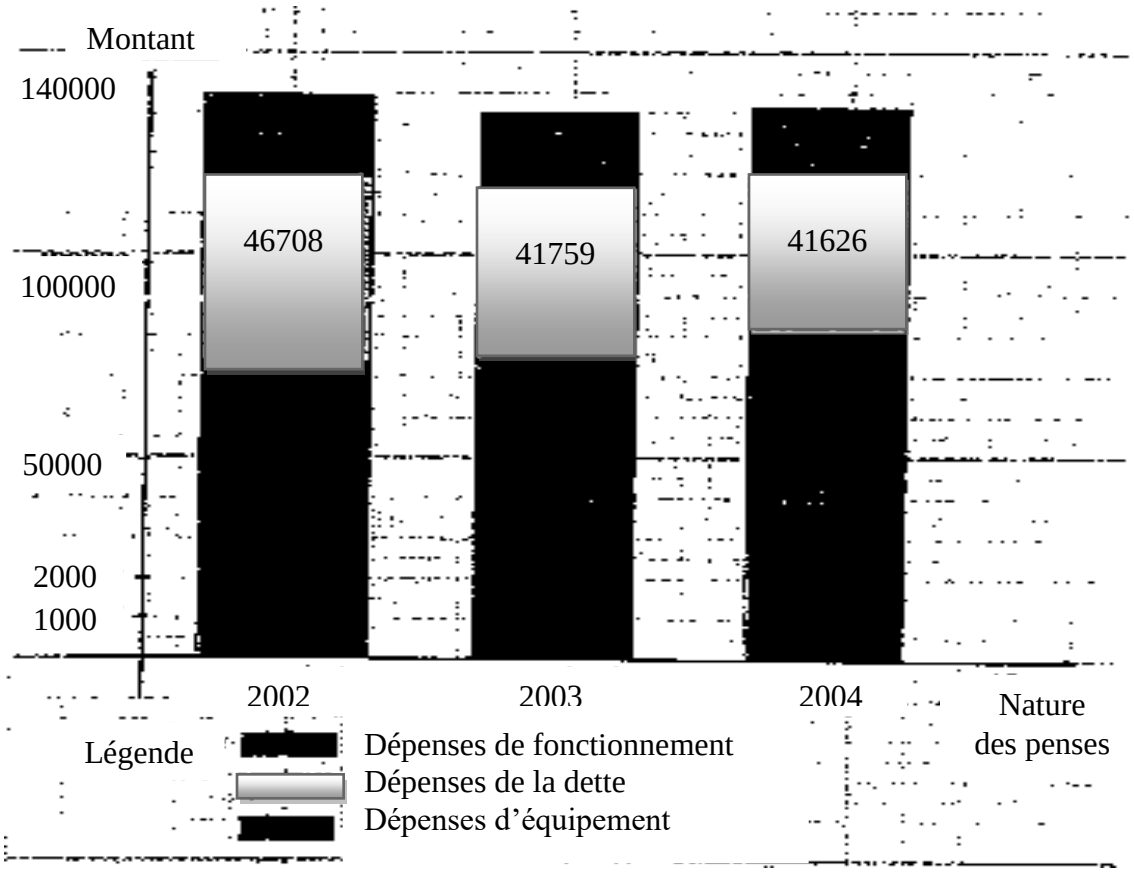
Remarque :

La représentation peut être faite en usant des valeurs absolues figurant sur le tableau ou moyen des pourcentages.

- Représentation des valeurs absolues :

Prévisions des dépenses	2002		2003		2004	
	Valeurs	Valeurs cumulées	Valeurs	Valeurs cumulées	Valeurs	Valeurs cumulées
Dépenses de fonctionnement	75442	75442	78250	78250	81047	81047
Dépenses de la dette	46708	122150	41759	120009	41626	122673
Dépenses d'équipement	19925	142075	19547	189556	191868	141868*

- Un diagramme représentera chaque année et les hauteurs des différentes composantes seront égales aux valeurs absolues.



- On trace ensuite les diagrammes représentatifs chacun d'une année et leur hauteur et respectivement 100 ; 98,22 et 99,85
- Après calcul des pourcentages représentatifs de chaque composante et après graduation des hauteurs des rectangles de 1 à 100, on place les pourcentages cumulés.

Tableau de calcul des pourcentages et des pourcentages cumulés :

Prévisions des dépenses	2002		2003		2004	
	Pourcen-tages	% cumulé	Pourcen-tages	% cumulé	Pourcen-tages	% cumulé
Dépenses de fonctionnement	53,10	53,10	56,07	56,07	57,13	57,13
Dépenses de la dette	32,88	85,98	29,93	86	29,34	86,47
Dépenses d'équipement	14,02	100	14	100	13,53	100

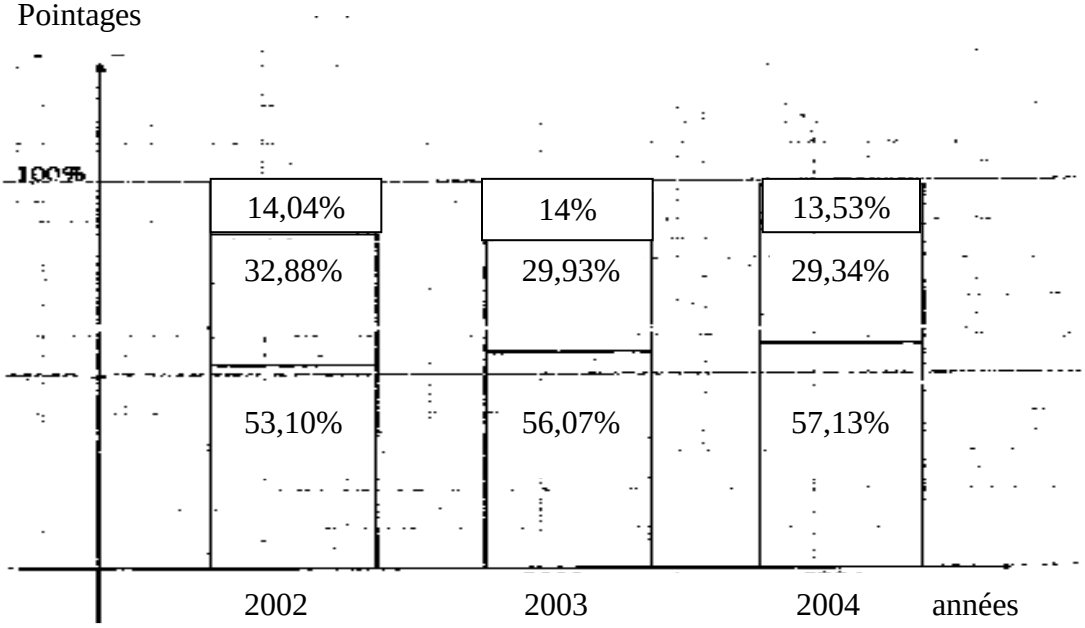


Diagramme en barres des prévisions de dépenses du budget en 2002-2003 et 2004

4-3-5- Représentation graphique à l'aide d'un diagramme carré

Le principe est simple, mais subordonné à l'existence de deux masses égales composées chacune de deux éléments ; exemple du bilan avec à l'actif les informations et l'actif circulant et au passif les capitaux permanents et les dettes.

Exemple :

Le bilan d'une entreprise se présente comme suit (tableau ci-contre) :

Actif	Montant	%	Passif	Montant	%
Actif immobilisé	130000	23,21	Capitaux prermanents	220000	39,28
Actif circulant	430000	76,79	Dettes	340000	60,72
	5600000	100		560000	100

Traduisons le bilan en pourcentage, on obtient :

On le représente grâce à un carré dont on gradue la base du haut et la base du bas de 0 à 100.

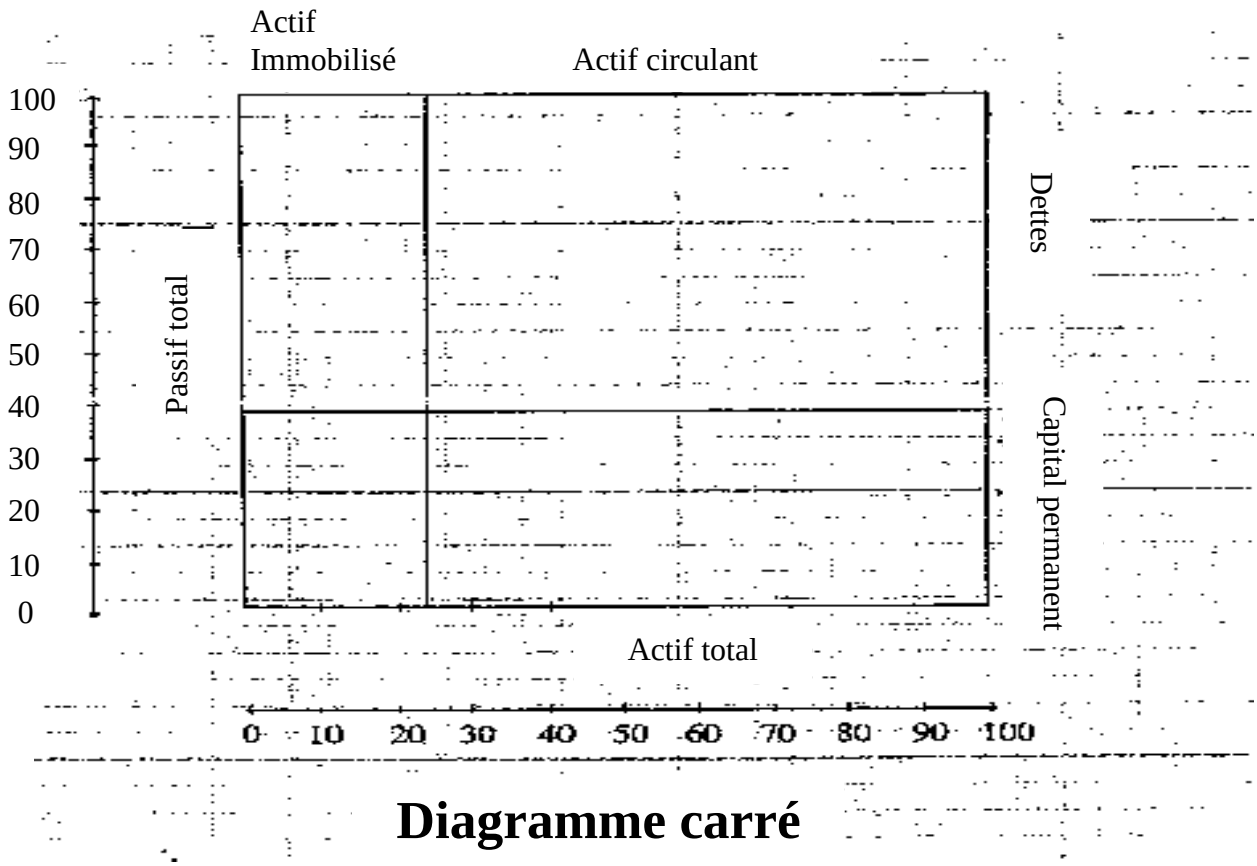


Diagramme carré

4-3-6 Représentation graphique à l'aide d'un diagramme polaire

Ce graphique est généralement utilisé dans le cadre de représentation de séries chronologiques avec différentes composantes : année et mois, année et trimestre, mois et jour etc…

Il s'agit de diviser un angle plein en autant d'angles égaux qu'il y a des périodes.

Exemple :

Le tableau suivant représente l'évolution de la monnaie fiduciaire mois par mois en 2003 et 2004 en millions de dirhams.

Mois Années	Jan.	Fév.	Mars	Avril	Mai	Juin	Juillet	Août	Sept.	Oct.	Nov.	Déc.
2003	70172	70885	70264	70392	70246	70937	74970	76671	74587	74525	74998	74890
2004	79620	76578	75429	76141	75091	75829	80114	81448	79420	79470	78959	79439

Rapport Bank Al-Maghrib 2004

Remarque :

Il est possible de représenter les valeurs absolues ou de transformer les valeurs absolues en indices. (cours de la 2ème année du Bac).

Nous procédons ici à une représentation moyennant les valeurs en partageant l'angle plein en 12 c'est-à-dire en angles égaux de 30° chacun.

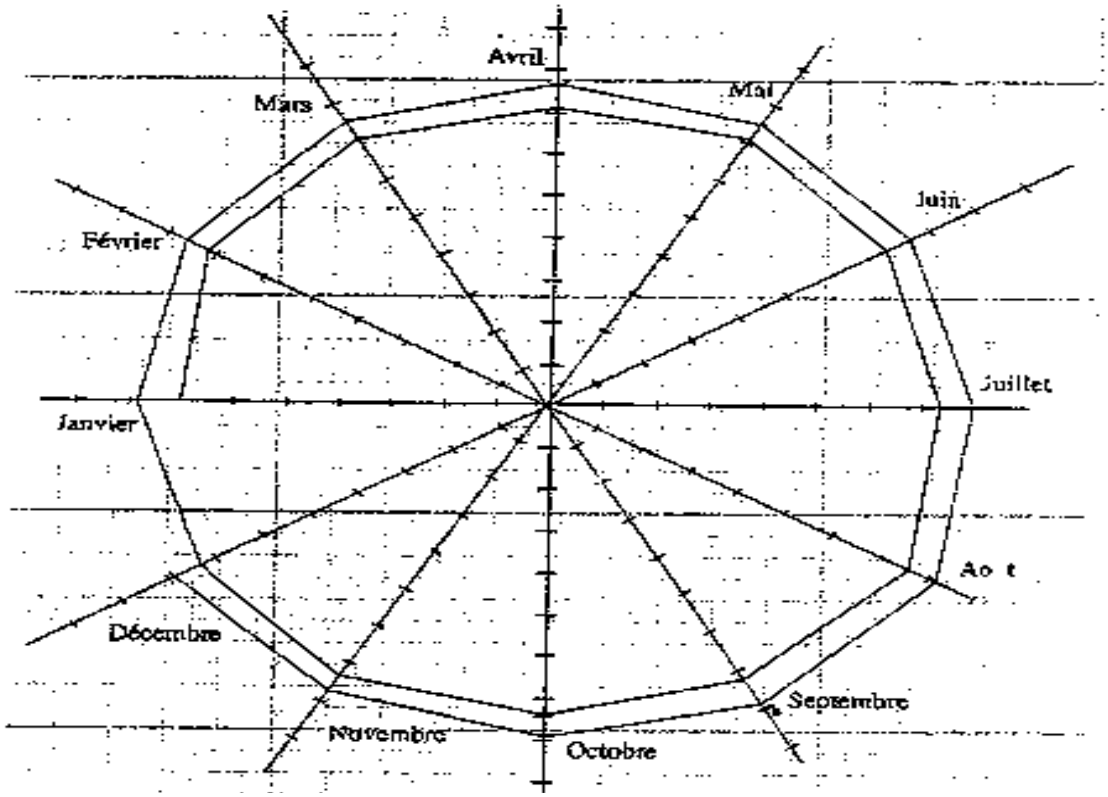


Diagramme a coordonnées polaires