محاضرات تقنيات المعالجة الإحصائية لطلبة السنة الأولى ماستر
تخصص: النشاط البدني الرياضي المدرسي

المحور الأول: مفاهيم أساسية Concepts De Base

أولا - تعريف علم الإحصاء

الإحصاء هو العلم الذي يبحث في طريقة جمع البيانات عن الظواهر التي تحيط بنا سواء كانت علمية أو اقتصادية أو اجتماعية، وفي كيفية تسجيل هذه الحقائق والبيانات في صورة دقيقة، ثم وصفها بصورة سهلة تبين علاقات واتجاهات هذه الظواهر، وأخيرا يبحث في دراسة هذه العلاقات والاتجاهات ووضعها في صورة يسهل معها فهم الظواهر المراد دراستها.

كما يعرف علم الإحصاء بأنه علم اتخاذ القرارات في جميع نواحي الحياة، وذلك من خلال جمع ودراسة وتحليل البيانات المتوفرة واستخلاص النتائج عن الظواهر المدروسة.

مما سبق يمكن تصنيف الإحصاء كعلم إلى قسمين رئيسيين هما:

1 – الإحصاء الوصفي Statistique descriptive: وهو ذلك الفرع من الإحصاء الذي يتناول طرق تنظيم وتلخيص وعرض البيانات في صورة مبسطة.

2 – الإحصاء الاستدلالي Statistique inferentiel: وهو ذلك الفرع من الإحصاء الذي يهتم بطرق الوصول إلى نتائج معينة أو توقعات ما عن المجتمع من خلال دراسة عينة من ذلك المجتمع.

فإذا كانت لدينا كمية كبيرة من البيانات العددية، فإن الإحصائي سيحاول أن يرتبها في صورة تجعل من السهل قراءتها وفهمها، وقد يتضمن هذا:

- تبويب البيانات وتقديمها في شكل جداول تكرارية، أو في شكل منحنيات بيانية ليسهل فهم معناها فورا.
- حساب بعض المقاييس أو المؤشرات الإحصائية مثل النسب أو المتوسطات.

وتدخل العمليات السابقة في نطاق الإحصاء الوصفي، أما الإحصاء الاستدلالي فهو يختص بـ:

- إجراء التنبؤات والتقديرات والاستنتاجات عن مجموعة من المتغيرات أكبر من تلك التي تمت ملاحظتها فعلا.

ثانيا – التعريف بالمصطلحات الإحصائية الضرورية:

1 – المجتمع Population: هو جميع العناصر المشتركة في الصفة التي تهم الباحث في دراسته، فقد يكون المجتمع مثلا عدد سكان مدينة، أو طلبة جامعة ما،... إلخ.

2 – العينة ECHANTILLON: هي جزء من المجتمع تحت الدراسة مثل مجموعة من سكان مدينة، أو مجموعة من طلبة جامعة ما،...الخ.

3 – الظاهرة Phénomène: هي صفة لعناصر تختلف من عنصر لآخر في الشكل أو النوع أو الكمية، ويطلق على الصفة تحت الدراسة متغير variable مثل: طول شخص ما، عدد الأخطاء الإملائية في بحث ما، ... الخ.

4 – المتغير variable: هو الصفة تحت الدراسة كما أشرنا أعلاه أو هو الشيء الذي يمكن أن يأخذ قيما مختلفة في ظروف مختلفة (زمنية، مكانية، سياسية، اقتصادية ... الخ) فمثلا سعر المكملات الغذائية يختلف من يوم لآخر ويختلف في نفس اليوم من محل لآخر.

ثالثا- أنواع المتغيرات:

تنقسم المتغيرات إلى نوعين:

1 – متغيرات نوعية (كيفية) Variable qualitative: وهي عبارة عن صفات أو أنواع معينة ليست عددية، وتنقسم بدورها إلى:

أ – بيانات نوعية خاضعة للترتيب : مثل المستوى التعليمي، الرتب العسكرية تقديرات النجاح ، المستوى الاقتصادي ... الخ.

ب – بيانات نوعية غير خاضعة للترتيب : مثل الجنسية، الجنس، أنواع السيارات، أنواع الأمراض ...إلخ.

2 – متغيرات كمية (عددية) Variable quantitatives: وهي البيانات التي يعبر عنها في صورة عددية وتنقسم إلى:

أ – متغير منفصل متقطع Variable discret: وهو المتغير الذي يأخذ أعداد صحيحة، فمثلا إذا كان (X) متغير يمثل عدد أفراد الأسرة، فإنه يمكن أن يأخذ القيم 2، 3، 4، 5 ... ولا يمكن أن يأخذ (X)القيم 1.5، 3.25، 5.17.

ب – متغير متصل (مستمر) Variable Continue: وهو المتغير الذي يمكن أن يأخذ أي قيمة بين قيمتين معنيتين، وكأمثلة عن المتغيرات المتصلة: الطول، الوزن، الزمن، السرعة ...الخ، فإذا كان (X) هو متغير الطول فمثلا فإن (X) يمكن أن تأخذ القيم 15متر، 11.3 متر، 14.75 متر، أي أن المتغير (X) يمكن أن يأخذ أي قيمة في فترة زمنية معينة.

رابعا - مصادر البيانات:

مما لا شك فيه أن في الدراسات الإحصائية تعد البيانات المادة الأساسية الرئيسية، وعليها تتوقف دقة الوصف والتحليل وسلامة الاستنتاج ومنطقيته، فإذا كانت هذه البيانات والمعلومات دقيقة وشاملة وواقعية، كان الوصف والاستنتاج والقرار الذي نحصل عليه سليما وصحيحا، وعليه فالاهتمام التام والحرص الدقيق في الحصول على بيانات سليمة وواقعية حول الظواهر تحت الدراسة يعد العمود الفقري والحجر الأساسي في علم الإحصاء، وهناك عدة مصادر للحصول على البيانات تختلف باختلاف موضوع الدراسة والغرض منها، من أهم هذه المصادر ما يلي:

- النشرات والدوريات والسجلات.
- التجارب.
- الاستبيان.
- التعدادات العامة.

II. العرض الجدولي للبيانات:

تلخص الجداول التكرارية البيانات الكمية الكثيرة في وضعها على صورة جدول منظم يوضح كيفية توزيع القيم التي حصلنا عليها من الظاهرة المدروسة، حيث يدل العمود (السطر) الأول على قيم الظاهرة، ويدل العمود (السطر) الثاني على التكرار المقابل لهذه القيم.

مثال1:

إذا كانت لدينا درجات 30 طالب في أحد الاختبارات كما يلي:

4، 8، 7، 6، 4، 11، 12، 4، 6، 8، 4، 6، 9، 6، 7، 6، 8، 4، 11، 9، 8، 11 ، 9، 4، 8، 9، 11، 12، 4، 5.

لخص هذه البيانات في صورة جدول يوضح معالمها الأساسية.

الدرجة	4	5	6	7	8	9	11	12	المجموع
التكرار	7	1	5	2	5	4	4	2	30

إن وضع البيانات بهذه الصورة أصبح أكثر وضوحا لمعرفة عدة معلومات كانت غير واضحة في الصورة الخام، فمثلا من السهل الآن معرفة أكبر وأصغر درجة عليها هؤلاء الطلبة كما يمكن معرفة الناجحين أو عدد المقصيين ببساطة.

وفي حالة ما إذا كنت الدرجات كثيرة ومنتشرة داخل مجال واسع (مثلا التنقيط كان من 100) فإنه لو وضعت في الشكل السابق الذي يبين درجات الطلبة مفردة مع تكرار كل درجة سوف نجد أن الجدول بهذه الطريقة أبسط ويوضح المعالم الرئيسية ، أي وضع البيانات في صورة مختصرة ومنظمة تعطي فكرة عامة عن هذه البيانات.

وقد وجدت طريقة أكثر اختصارا من السابقة يمكن بواسطتها وضع البيانات في جدول يبين ويوضح الخصائص العامة لهذه البيانات، يسمى هذا الجدول بجدول التوزيع التكراري ولتكوين مثل هذا الجدول نقوم بالتالي:

أولا: نحدد المجال (المدى) الذي تنتشر فيه البيانات، وهو الفرق بين أكبر قيمة للبيانات وأصغر قيمة لها، أي أن:

المدى = أكبر قيمة – أصغر قيمة.

ثانيا: نقسم المدى إلى فئات متساوية الطول بحيث يكون عددها مناسبا (ما بين 5 و 25 فئة) وهناك عدة طرق لحساب عدد الفئات نذكر منها:

1 – معادلة ستيرجس Sturages :التي تنص على أن: **عدد الفئات = 1 + 3.322 لغ عدد البيانات.**

2 – معادلة يول Yule:التي تنص على: **عدد الفئات = 2.5** $\sqrt[4]{\text{عدد البيانات}}$

ثالثا:نحسب طول الفئة وهو يساوي المدى مقسوما على عدد الفئات

$$\text{طول الفئة} = \frac{\text{المدى}}{\text{عدد الفئات}}$$

مما سبق نستخلص الطريقة المرنة في تحديد عدد الفئات وأطوالها والتي لا تعتمد على المعدلات الرياضية بل أن هذه الطريقة مرنة بطبيعتها وهي:

المدى≤ طول الفئة × عدد الفئات

ملاحظات:

أ – عند تفريغ البيانات فإنه يجب أن تنتمي كل مفردة إلى فئة واحدة فقط.

ب – عند كتابة الفئات فإنه:

- يذكر الحد الأدنى والأعلى لكل فئة إذا كان المتغير متقطع.
- يذكر الحد الأدنى ويحدد الحد الأدنى الأعلى ضمنيا أو العكس إذا كان المتغير متصل.

- يفضل استخدام الفئات المتساوية الطول، إلا أنه في بعض الحالات يمكن أن يستخدم الفئات غير المتساوية، من هذه الحالات ما يلي:

- إذا كان الغرض من الدراسة هو الاهتمام ببعض الفئات والتركيز عليها وإهمال باقي الفئات، فيمكن عندها دمج الفئات التي لا تهم الباحث في فئة واحدة.
- إذا كان التكرار لبعض الفئات صغيرا جدا مقارنة بباقي الفئات، يمكن دمج هذه الفئات معا.

مركز الفئة:

مركز الفئة (منتصف الفئة) تحسب كالتالي: $$\text{مركز الفئة} = \frac{\text{الحد الأدنى للفئة + الحد الأعلى للفئة}}{2}$$

القواعد الواجب إتباعها عند تشغيل الجداول الإحصائية هي:

1) عنوان واضح في أعلى الجدول يعطي فكرة عن البيانات التي يحتويها هذا الجدول.
2) ذكر مصدر البيانات في أسفل الجدول.
3) ذكر وحدة القياس المستعملة.
4) ذكر عنوان كل عمود (سطر).
5) وضع رقما للجدول.

*العرض الجدولي في حالة البيانات النوعية:

يتكون الجدول من عمود ين (سطرين) يحتوي العمود (السطر) الأول على رموز كتابية للخاصية المدروسة (صفات) أما الثاني فيحتوي على تكرارات كل رمز كتابي.

مثال2: أخذت عينة من طلبة جامعة ما تتكون من 100 طالب، حيث كانت الخاصية المدروسة هي الانتماء إلى كلية من كليات الجامعة فكانت النتائج كما هي مبينة في الجدول أدناه

جدول توزيع مجموعة من طلبة جامعة ما على كليات الجامعة

الكلية	الحقوق	الاقتصاد	الآداب	العلوم	رياضة	المجموع
عدد الطلبة	15	30	25	10	20	100

المصدر: مصلحة الانخراط المركزي للجامعة

*التكرارات المتجمعة:

في بعض الحالات نرغب في معرفة التكرارات أو البيانات التي تزيد قيمتها عن قيمة معينة أو تقل عن قيمة معينة، فمثلا عندما نرغب في معرفة عدد الناجحين (أي الطلبة المتحصلين على درجة تساوي أو تزيد 10) فإن هذه المعلومات غير واضحة في جدول التوزيع التكراري فنكون لهذا الغرض ما يسمى بالجدول التكراري المتجمع الصاعد أو النازل، وكتعريف فإن:

- التكرار المتجمع الصاعد(ك م ص) لأي فئة هو تكرار هذه الفئة مضافا إليه مجموع تكرارات الفئات السابقة.
- التكرار المتجمع النازل(ك م ن) لأي فئة هو عبارة عن مجموع التكرارات مطروحا منه تكرارات الفئات السابقة.

مثال 3:

الجدول الآتي يبين توزيع دخل عينة من عمال مؤسسة صناعة الكوابل الكهربائية حسب دخولهم بآلاف الدينارات. المطلوب حساب التكرار المتجمع الصاعد والتكرار المتجمع النازل لهذه البيانات.

فئة الدخل	عدد العمال
15 – 10	4
20 – 15	6
25 – 20	12
30 – 25	8
35 – 30	6
40 – 35	4
المجموع	40

الحل:

الفئة	التكرار fi	ك م ص ↗ fi+	ك م ن -fi ↘
15 – 10	4	4	40
20 – 15	6	10	36
25 – 20	12	22	30
30 – 25	8	30	18
35 – 30	6	36	10
40 – 35	4	40	4
المجموع	40		

ملاحظات:

- التكرار المتجمع الصاعد للفئة الثالثة [20 – 25[= 22 ويعني أن عدد العمال الذي يقل دخلهم عن 25 ألف دينار يساوي 22 عامل أي 4 + 6 + 12.

- التكرار المتجمع النازل للفئة الرابعة [25 – 30 [= 18 ويعني أن عدد العمال الذي يساوي دخلهم أو يزيد عن 25 ألف دينار هو 18 عامل أي 40 – 4 – 6 – 12.
- نستعمل نفس الطريقة في حساب التكرار المتجمع الصاعد والتكرار المتجمع النازل في حالة الفئات غير متساوية.

التكرار النسبي: هو وضع تكرار كل فئة كنسبة من التكرار الكلي، وهذه الطريقة لها عدة استخدامات وفوائد حيث توضح نسبة تكرار كل فئة إلى التكرار الكلي.

الجدول المزدوج:

يستعمل الجدول المزدوج عند دراسة خاصيتين في نفس الوقت في مجتمع ما ، وتوضح البيانات الإحصائية في هذا الجدول بالشكل التالي:

- نخصص الأسطر لبيانات الخاصية الأولى ونخصص الأعمدة لبيانات الخاصية الثانية.
- نرمز لقيم الخاصية الأولى بالرمز Xi حيث i تتراوح من 1 إلى n ونرمز لقيم الثانية بالرمز yi حيث j تتراوح من 1 إلى n.

مثال: عند دراسة مستوى معيشة العائلات يمكن أن نتطرق إلى خاصيتين هما: مهنة رب الأسرة والإنفاق الاستهلاكي، فإذا رمزنا للمتغير الأول مهنة رب الأسرة بالرمز Xi والمتغير الثاني الإنفاق الاستهلاكي بالرمز yj يمكن أن نشكل الجدول المزدوج التالي:

fi	y4	y3	y2	y1	المتغير y / المتغير X
f1	N14	N13	N12	N11	X1
f2	N24	N23	N22	N21	X2
f3	N34	N33	N32	N31	X3
f4	N44	N43	N42	N41	X4
f5	N55	N53	N52	N51	X5
$\sum f_i = \sum n_j$	n4	n3	n2	n1	Nj

مثال 4:

سحبت عينة عشوائية من مجتمع ما، تتكون من 183 أسرة قصد دراسة خاصيتين هما مهنة رب الأسرة والتركيبة الأسرية من حيث عدد الأطفال، فكانت النتائج كالتالي:

المجموع	6	5	4	3	2	1	0	عدد الأطفال / المهنة
32	0	1	3	4	6	8	10	إطار المتوسط
28	0	0	1	2	3	7	15	مهنة حرة
57	7	9	15	10	8	5	3	عامل متخصص
66	10	17	13	11	9	4	2	عامل بسيط
183	17	27	32	27	26	24	30	المجموع

ويمكن أن نقرأ من الجدول السابق ما يلي:

- من بين 183 أسرة المدروسة هناك 28 أسرة يشتغل رب الأسرة بالمهن الحرة.
- من بين 183 أسرة المدروسة هناك 32 أسرة لها أربع أطفال.
- أن هناك 8 عمال متخصصين لهم طفلين و 10 لهم 3 أطفال.
- أن هناك 17 أسرة فقط لها 6 أطفال منها 7 أسر يشتغل رب الأسرة كعامل متخصص و 10 أسر يشتغل رب الأسرة كعامل بسيط.

III - العرض البياني للبيانات:

بالإضافة إلى العرض الجدولي للبيانات هناك طرق أخرى تستخدم لتوضيح وتلخيص البيانات ، وهي طرق العرض البياني ، وهذه الطرق أسهل وأبسط من سابقتها، فإذا كان العرض الجدولي لا يفهمه إلا المختصين فإن طرق العرض البياني يمكن أن تجلب غير المتخصصين، كما تمكن من القيام بتحليل سريع للظاهرة المدروسة، وتستخدم أنواع مختلفة للعرض البياني حسب نوعية المتغير المدروس.

1 – المتغير الكمي المتقطع: وهنا يمكن استخدام الأعمدة البسيطة، العرض البياني للتكرارات المتجمعة الصاعدة والنازلة.

2 – المتغير الكمي المتصل: هنا يمكن استخدام المدرج التكراري، المضلع التكراري، المنحنى البياني للتكرارات المتجمعة الصاعدة والنازلة.

3 – المتغير النوعي: وهنا يمكن استخدام الدوائر، الأعمدة المجزأة، الأعمدة المستطيلة.

I - العرض البياني في حالة متغير كمي متقطع:

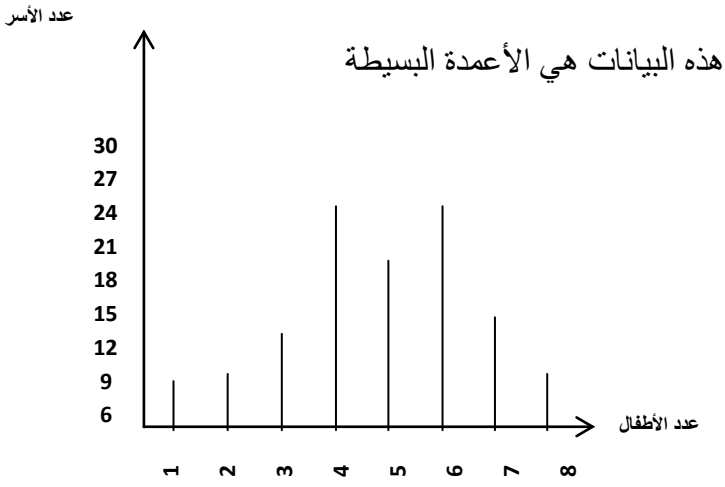
1 – الأعمدة البسيطة: هي عبارة عن أعمدة بسيطة تتناسب أطوالها مع التكرار المقابل لقيمة العينة للمتغير المدروس (Diagrammes en batons)

مثال 5: يبين الجدول التالي عدد الأطفال في العائلة لعينة تتكون من 125 أسرة، المطلوب عرض هذه البيانات بطريقة العرض المناسبة البسيطة.

عدد الأطفال في كل أسرة	عدد الأسر
0	6
1	9
2	10
3	14
4	26
5	20
6	25
7	15
8	10
المجموع	125

الحل:

أفضل طريقة لعرض هذه البيانات هي الأعمدة البسيطة



نلاحظ بسهولة أن العمود الذي يقابل القيمة 4 هو أطول الأعمدة وتكراره يساوي 26 ويعني ذلك أن أغلب العائلات لها 4 أطفال.

2 – العرض البياني للتكرار المتجمع الصاعد :

هو عبارة عن قطع مستقيمة متصاعدة حسب تصاعد التكرارات المتجمعة الصاعدة المقابلة لكل قيمة من قيم المتغير الإحصائي المدروس ، ولرسم القطع المستقيمة المقابلة لقيم المتغير المدروس في المثال السابق نقوم أولا بحساب التكرار المتجمع الصاعد.

ثانيا: العرض البياني في حالة متغير كمي متصل (مستمر):

إن العروض البيانية للمتغير الإحصائي الكمي المستمر هي أكثر العروض البيانية استعمالا ومن أهمها:

1 – المدرج التكراري Histogramme:

وهو عبارة عن مستطيلات متلاصقة طول كل منها يتناسب مع التكرار المقابل، وقاعدة كل منها تساوي طول الفئة المقابلة، ويمكن أن نميز بين حالتين عند رسم المدرج التكراري:

ملاحظة: نقوم بتعديل التكرارات في حالة الفئات غير المتساوية في الحالتين التاليتين:

- عند رسم المدرج التكراري.
- عند تحديد الفئة المنوالية وحساب المنوال.

2. المضلع التكراري Polygone de fréquence:

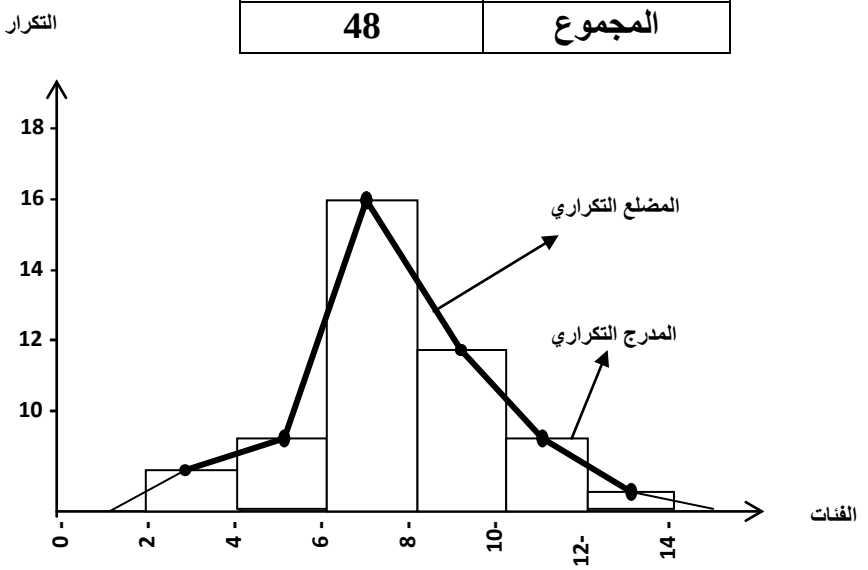
ه ي مجموعة من قطع مستقيمة متصلة ومنكسرة تتحدد بنقاط أحداثياتها مركز الفئة والتكرارات المقابلة.

مثال 6:

ليكن التوزيع التكراري الآتي، أرسم المدرج التكراري والمضلع التكراري.

الفئة	التكرار
4 – 2	4
6 – 4	8
8 – 6	16
10 – 8	12
12 – 10	6
14 – 12	2
المجموع	48

الحل:



ملاحظة: الخط المنكسر يمثل المضلع التكراري، حيث أن المساحة التي تقع تحت المضلع التكراري تساوي المساحة التي تقع تحت المدرج التكراري، وحتى نحافظ على المساحة التي تقع أسفل هذا المضلع،

نفترض أن لهذا التوزيع فئات إحداهما في بدايته والأخرى في نهايته ، تكرار كل منهما يساوي صفر، بحيث ننطلق في رسم المضلع من مركز الفئة الافتراضية الأولى (الفئة ما قبل الأولى)، وننتهي عند مركز الفئة الافتراضية الأخيرة.

3) منحنى التكرارات المتجمعة الصاعدة والنازلة:

يرسم منحنى التكرار المتجمع الصاعد عن طريق إيصال مجموعة النقاط ذات الإحداثيات التالية: الحدود العليا للفئات والتكرار المتجمع الصاعد المقابل لها، ويرسم منحنى التكرار المتجمع النازل بإيصال مجموعة النقاط التي إحداثياتها: الحدود الدنيا للفئات والتكرار المتجمع النازل مقابل لها.

يبين كل من منحنى التكرار المتجمع الصاعد ومنحنى التكرار المتجمع النازل شدة أو ضعف تطور الظاهرة المدروسة عن مستوى معين من مجال الدراسة. إن فاصلة نقطة تقاطع المنحنيين تسمى بالوسيط.

مثال7: حالة المثال رقم 3

III - العرض البياني في حالة متغير نوعي:

1) العرض الدائري: Diagramme circulaire:

ويتمثل في دائرة مقسمة إلى عدة أجزاء كل جزء يقابل زاوية مركزية تتناسب مع التكرارت المقابلة لكل خاصية من الخصائص المدروسة، ولتحقيق ذلك نضيف عمودا إلى جدول المعطيات يحتوي على الزوايا المركزية المقابلة لكل تكرار.

ملاحظة: تحسب الزوايا المركزية بالطريقة التالية: $\frac{\text{تكرار الخاصية}}{\text{مجموع التكرارات}} \times 360°$

مثال 8:

يبين الجدول التالي عدد طلبة كلية الحقوق والعلوم الاقتصادية سنة 2004 مقسمين على أقسام الكلية المختلفة.

المطلوب: عرض البيانات باستخدام القطع الدائرية.

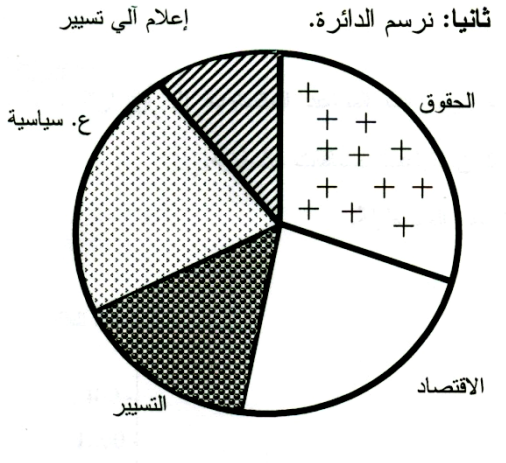
القسم	الحقوق	الاقتصاد	التسيير	علوم سياسية	إعلام آلي	المجموع
عدد الطلبة	1200	1000	800	600	400	4000

الحل:

أولا: نحسب الزاوية المركزية.

القسم	عدد الطلبة	الزاوية المركزية
الحقوق	1200	°108
الاقتصاد	1000	°90
التسيير	800	°72
ع.سياسية	600	°54
إ. آلي	400	°36

ثانيا: نرسم الدائرة.



2) العمود المجزأ Diagrammes en Barres:

وهو عبارة عن مستطيل مقسم إلى عدة أجزاء كل جزء يقابل تكرار معين للخاصية المدروسة.

مثال 15: أعرض بيانات المثال السابق باستخدام العمود المجزأ.

3) الأعمدة المستطيلة:

وهو عبارة عن مجموعة من الأعمدة المتجاو رة ذات القواعد المتساوية إلا أن ارتفاعها تتناسب مع تكرار كل خاصية، كما أن هذه الأعمدة تكون متباعدة بمسافات متساوية.

مثال 9: أعرض بيانات المثال السابق باستخدام الأعمدة المستطيلة