المستوى: طلبة السنة الثانية ماستر تخصص تسويق خدمات أستاذ المقياس: د. لمياء مكرسي

المحاضرة الثالثة عشرة: المعالجة الإحصائية للبيانات

أولا: مفهوم المعالجة الإحصائية للبيانات

تبدأ مرحلة معاجلة البيانات واستخلاص النتائج، بعد الانتهاء مباشرة من عملية جمع البيانات والمعلومات، بالوسائل والأساليب المختلفة المشار إليها في المحاضرات السابقة. ويكمن الهدف الأساسي لمعالجة البيانات واستخلاص النتائج، في التأكد من دقتها وموثوقيتها واستكمالها، باستخدام الطرق والأدوات الملائمة لطبيعة البحث، والإمكانات المتاحة للباحث ليتاح له الوصول إلى النتائج العلمية للمشكلة المدروسة.

بعد العودة من البحث الميداني، يجد الباحث نفسه في وسط بحر من المعلومات، فماذا يفعل؟ الخطوة الأولى تتمثل في غربلة المعلومات التي يحصل عليها، وذلك بواسطة إعطاء الأولوية للمصادر الأصلية، والتدقيق في المعلومات الركيكة في أكثر من مصدر والتركيز على المصادر والمراجع الأكثر حداثة: سواء في إحصاءاتها وأرقامها أو توثيقها، واستبعاد المعلومات التي لا تتعلق مباشرة بموضوع البحث تلافيا للوقت والجهد. ويمكن استخدام برامج معينة للتحليل العلمي.

بعد تجمع البيانات والمعلومات وتصنيفها وتبويبها وتمثيلها بيانيا ينتقل الباحث إلى وصفها، عن طريق إبراز الخصائص الأساسية لها، والتي يمكن التعبير عنها بمقاييس محددة. والخصائص الأساسية لأي مجموعة من البيانات والمعلومات تقاس بمقاييس معينة، وسيتم تناول البعض منها فيما يلي:

ثانيا: مقاييس النزعة المركزية

مقاييس النزعة المركزية هي مقاييس عددية تستخدم لقياس موضع تمركز البيانات أو تجمعها إذ إن بيانات أي ظاهرة تنزع في الغالب إلى التركز والتجمع حول قيم معينة، هذه القيم هي ما يسمى بمقاييس النزعة المركزية. وللنزعة المركزية مقاييس عديدة أهمها:

1. **الوسط الحسابي**: ويسمى أيضا بالمتوسط وهو أبسط أنواع المتوسطات وأكثرها استعمالا.

أ. **في حالة البيانات غير المبوبة**: يأخذ الصيغة التالية

$$\bar{x} = \frac{\Sigma x_i}{n}$$

مثال: إذا كانت أعمار خمسة موظفين في شركة ما: 22، 25، 30، 33، 40.

فإن المتوسط الحسابي لأعمارهم هو: **الوسط الحسابي= 150/ 5= 30**

ب. في حالة البيانات المبوبة: فإن صيغة الوسط الحسابي كما يلي:

$$\frac{\Sigma f_i x_i}{\Sigma f_i} = \bar{x}$$

مثال1: أوجد الوسيط الحسابي للجدول التكراري التالي لعينة عددها 70 من العاملين في الجامعة:

الفئات	التكرار
6–2	5
11–7	7
16–12	12
21–17	18
26–22	21
31–27	6
36–32	1
المجموع	70

الحل:

الفئات	التكرار f_i	مراكز الفئات x_i	$f_i x_i$
6–2	5	4	20
11–7	7	9	63
16–12	12	14	168
21–17	18	19	342
26–22	21	24	504
31–27	6	29	174
36–32	1	34	34
المجموع	70		**1305**

الوسط الحسابي= (70/1305)= 18.64

2. **الوسيط:** هي القيمة التي تقع في منتصف مجموعة من البيانات المرتبة تصاعديا أو تنازليا. ونميز بين حالتين لحسابه:

أ. **حالة البيانات غير المبوبة:** ويتم اتباع الخطوات التالية:

1.ترتب قيم المتغير تصاعديا؛ 1.2 إذا كان حجم العينة n فردي فالوسيط هو قيمة المتغير في الترتيب (n+1)/2

أي $Me = x_{\left(\frac{n+1}{2}\right)}$ أي المشاهدة التي تقع في المنتصف.

2.2 إذا كان حجم العينة n زوجي فالوسيط هو متوسط قيم المتغير في الرتبة n/2 وعليه:

$$Me = \frac{x_{(n/2)} + x_{(n/2)+1}}{2}$$

أي هو متوسط المشاهدتين التين تقعان في المنتصف.

مثال: عينة من القيم 4 6 8 10 قيمة الوسيط في هذه الحالة هي 7.

ب. **حالة البيانات المبوبة:** يمكن حساب الوسيط حسابيا وإيجاده بيانيا.

-**الوسيط حسابيا:** وفقا للصيغة التالية:

$$\mathbf{M_d} = \mathbf{L} + \frac{\frac{\Sigma f_i}{2} - f_i}{f_2 - f_1}.H$$

حيث أن:

L: الحد الأدنى لفئة الوسيط. f1: التكرار المتجمع السابق لقيمة موقع الوسيط $\frac{\Sigma fi}{2}$: قيمة موقع الوسيط

f2: التكرار المتجمع اللاحق لقيمة موقع الوسيط H: طول الفئة.

وباتباع الخطوات التالية:

-تحويل الجدول التكراري إلى جدول تكراري متجمع صاعد أو نازل.

-إيجاد ترتيب الوسيط بموجب المعادلة: (n/2)

-استخدام ترتيب الوسيط لإيجاد الفئة التي يقع فيها الوسيط، في جدول التكرار المتجمع الصاعد (تسمى الفئة الوسطية).

مثال 2: من المثال السابق (مثال 1) أحسب الوسيط.

الفئات	التكرار f_i	مراكز الفئات x_i	المتجمع الصاعد
6–2	5	4	5
11–7	7	9	12
16–12	12	14	24
21–17	18	19	42

26–22	21	24	63
31–27	6	29	69
36–32	1	34	70
المجموع	70	-	**-**

موقع الوسيط: (2/70)= 35

وبما أنه لا توجد القيمة 35 بين قيم التكرار المتجمع الصاعد، فإن موقع الوسيط بين القيمتين 24 و42، وبذلك فإن الوسيطة هي 17–21 باعتبارها الفئة اللاحقة لموقع الوسيط، وبتعويض القيم في صيغة الوسيط نحصل على قيمة الوسيط وهي:

الوسيط= 17+((24-42)/24-35)).5= 3.06+17 **=20.06**

–الوسيط بيانيا: يمكن إيجاد الوسيط بيانيا والذي يمثل على المحور الأفقي مكان إسقاط تقاطع المنحنين الصاعد والنازل معا في رسم واحد. أو من المنحنى المتجمع الصاعد أو المنحنى المتجمع النازل كل على حده حيث مستوى منتصف التكرارات (n/2)، يتم رسم خط أفقي حتى يمس المنحنى المتجمع ثم يتم إسقاط عمود على المحور الأفقي ليتم الحصول على الوسيط مباشرة.

3.المنوال: وهو القيمة التي تحدث أكثر من غيرها، أو هو القيمة الأكثر شيوعا في البيانات.

أ. في حالة البيانات غير المبوبة: هو القيمة الأكثر تكرارا بين مجموعة القيم، لذلك فإن قيمته قد لا تكون الوحيدة بل قد يكون هناك أكثر من قيمة منوالية واحدة لها ذات العدد من التكرارات.

مثال: أحسب المنوال من البيانات التالية: 5، 8، 7، 5، 4، 1.

يوجد لهذه البيانات منوال واحد وهو القيمة 5 لأنها تكررت مرتين أكثر من غيرها.

ب. في حالة البيانات المبوبة: في هذه الحالة يجب تحديد الفئة المنوالية التي هي الفئة التي يقابلها أكبر تكرار، ومن ثم تطبيق الصيغة التالية:

$$M_o = L + \frac{d_1}{d_1 + d_2}.H$$

حيث أن:

L: الحد الأدنى للفئة المنوالية. d1: تكرار الفئة المنوالية–تكرار الفئة السابقة d2: تكرار الفئة المنوالية–تكرار الفئة اللاحقة H: طول الفئة.

مثال: انطلاقا من المثال1 المطلوب إيجاد المنوال.

الحل:

الفئة المنوالية هي: 22–26 كونها تقابل أكبر تكرار، فيكون لدينا:

المنوال= $22+((15+3)/3)).5=$ 22.83