الاخطاء في البيانات السكانية وأساليب تقويمها

الإحصاءات السكانية سواء كان مصدرها التعداد العام للسكان أو التسجيل الحيوي معرضة لأخطاء يتوقف حجمها على الأسلوب المستخدم في عملية جمع البيانات أوفي عملية التسجيل .كما تتوقف أهمية هذه الأخطاء بعد تحديد حجمها على الحالة التي تستخدم فيها البيانات، حيث تتطلب بعض الحالات توافر بيانات سكانية على مستوى عالٍ من الدقة لا يتطلب الأمر توافره في مجالات أخرى. تعد عمليه تقويم البيانات الخطوات المهمه والاوليه التي يقوم بها الباحث للحصول على بيانات خاليه من الاخطاء قدر الامكان والتي قد تكون حاصله عن طريق الصدفه او قد تكون اخطاء متعمده اثناء مرحله جمع البيانات او اي مرحله من مراحل اعداد البيانات المختلفة. ولما كانت بيانات التركيب العمري والنوعي أهم البيانات التي تعتمد في حساب المعدلات او في مجالات التخطيط الاجتماعي او الاقتصادي وغير ذلك من المجالات.فلا بد ان تكون على مستوى عال من الدقه لتكون نقطه بدايه صحيحه لاعداد التحليلات الديموغرافية.

I. الاخطاء التي تتعرض لها بيانات التركيب العمري والنوعي:

تتعرض بيانات التركيب العمري والنوعي لعدد من الاخطاء التي قد تحصل لعده اسباب مختلفة ويمكن تقسيم هذه الاخطاء الى

1- أخطاء الشمول coverage errors: وترجع الى:

1-1 قصور العد undercount: بالنسبة لمجموعة من الأسر أو الأفراد داخل هذه الأسر أو عدم تسجيل بعض الواقعات الحيوية.

1-2 تكرار العد over count: أي الحصر أو التسجيل أكثر من مرة للأفراد أو الوقائع الحيوية.

1-3 عيوب في المراحل الادارية: حيث تتعرض بعض السجلات للفقد أو سوء الاستخدام أثناء تداولها في مراحل تجهيز البيانات أو عند نقلها من وحدات محلية إلى وحدات أكبر.

2- أخطاء المحتوى content errors: يمكن تقسيمها الى مايلي:

2-1 أخطاء الإجابة response errors :ويرجع هذا النوع من الأخطاء إلى عدم إعطاء الإجابة المناسبة على الأسئلة الواردة في استمارة التعداد أو قائمة التسجيل .ويعتبر الإبلاغ الخاطئ عن العمر من الأمثلة التقليدية على هذا النوع من الخطأ.

2-2 أخطاء العدادين enumerators errors:وترجع إلى فشل العداد وعدم تمكنه من توجيه الأسئلة بدقة أو إلى تهاون القائم بالتسجيل في عملية قيد الواقعات الحيوية.

2-3 أخطاء تجهيز البيانات data processing errors:وتنشأ هذه الأخطاء أثناء التجهيز الفني للبيانات وخاصة أثناء عملية وضع الدليل وتبويب البيانات وإعدادها للنشر.

II. تقويم بيانات العمر والنوع:

بما أن التعداد السكاني هو المصدر الرئيس الاول للبيانات السكانية، وهو الذي يوفر البيانات الكاملة عن التركيب العمري والنوعي للسكان، لذلك لابد من أختبار دقة تلك البيانات قبل الاعتماد عليها ويتم ذلك باستخدام عدة مقاييس أهمها:

أولا: مؤشر ويبل (Whipple's Index)

وهو مؤشر يعتمد لقياس درجة التراكم العمري من خلال التعرف على مدى التفضيل للأعمار التي آحادها (0 أو 5) للأعمار المحصورة بين (23 سنة) و (62 سنة). تتراوح قيمة هذا المؤشر بين القيمة 100 والتي تدل على عدم وجود تفضيل للأعمار المنتهية بالرقمين صفر أو خمسة (0 أو 5)، والقيمة 500 التي تدل على أن جميع الأعمار التي ذكرت في التعداد انتهت بأحد الرقمين صفر أو خمسة. ويتم حساب المؤشر باستخدام الصيغة التالية:

$$WI = \frac{(\sum P_{25} + P_{30} + P_{35} + \cdots + P_{60})}{\frac{1}{5}\sum_{X=23}^{62} P_x} \mathbf{x100}$$

لاختبار تفضيل الأعمار التي آحادها "0" يتم حساب المؤشر باستخدام الصيغة التالية:

$$WI = \frac{(\sum P_{30} + P_{40} + P_{50} + P_{60})}{\frac{1}{10}\sum_{X=23}^{62} P_x} \mathbf{x100}$$

ولاختبار تفضيل الأعمار التي آحادها "5" يتم حساب المؤشر باستخدام الصيغة التالية:

$$WI = \frac{(\sum P_{25} + P_{35} + P_{45} + P_{55})}{\frac{1}{10}\sum_{X=23}^{62} P_x} \mathbf{x100}$$

حيث أن:

WI: مؤشر ويبل.

P_x: عدد الأشخاص الذين يبلغون عن أعمارهم (X).

وتوصي الأمم المتحدة بمعيار للكشف عن مدى الدقة في الإدلاء ببيانات العمر باستخدام مؤشر ويبل كما هو موضح في الجدول الموالي:

جدول: توصية الأمم المتحدة للتراكم العمري كما حددها مؤشر ويبل

مؤشر ويبل	نوعية البيانات
أقل من 105	عالية الدقة
109-105	دقيقة
124-110	تقريبا دقيقة
175-125	غير دقيقة
أكبر من 175	غير دقيقة على الإطلاق

على الرغم من أهمية هذا المؤشر في قياس مدى التراكم العمري للأعمار، إلا أنه يعاني من بعض نقاط الضعف، حيث أنه يقتصر على قياس درجة تفضيل الأعمار التي تنتهي بصفر أو خمسة فقط، وذلك في المدى العمري من 23 إلى 62 سنة دون الأخذ بعين الاعتبار باقي سنوات العمر، ولا يأخذ في الاعتبار الطبيعة المتناقصة للتوزيع العمري بسبب التناقص بالوفاة، كما أنه يطبق فقط على بيانات سنة واحدة.

مثال: باستخدام بيانات التوزيع العمري للدولة ع الموضحة في الجدول رقم 09، حدد مدى التفضيل العمري لكل من:

- الآحاد "0" و"5"؛

- الآحاد "0" فقط؛
- الأحاد"5" فقط.

جدول رقم09: التوزيع العمري للدولة ع

العمر	السكان	العمر	السكان
23	30870	**25**	67208
24	55780	**30**	91610
29–26	120550	**35**	87456
34–31	105605	**40**	71610
39–36	115790	**45**	44510
44–41	80420	**50**	35420
49–46	75610	**55**	23450
54–51	55620	**60**	41577
59–56	58700		
61	5000		
62	6789		

الحل:

- **للأعمار التي آحادها "0" و"5":**

$$\boldsymbol{WI} = \frac{(462841)}{\frac{\mathbf{1}}{\mathbf{5}}(1173575)} \text{x} 100 = 197.19\%$$

- **للأعمار التي آحادها "0":**

$$\boldsymbol{WI} = \frac{(240217)}{\frac{\mathbf{1}}{\mathbf{10}}(1173575)} \text{x} 100 = 204.69\%$$

- **للأعمار التي آحادها "5":**

$$\boldsymbol{WI} = \frac{(222624)}{\frac{\mathbf{1}}{\mathbf{10}}(1173575)} \text{x} 100 = 189.7\%$$

2-مقياس مايرز Myer's Blended Index:

تم تطوير مؤشر مايرز لتحديد أوجه التفضيل لجميع الأعمار التي آحادها من الرقم صفر إلى تسعة. ويتراوح المدى النظري لدليل مايرز من الصفر إلى 90، وعندما تكون قيمة المؤشر مساوية للصفر لا يكون هناك أخطاء تفضيل رقمي (تفضيل لأي رقم)، وكلما بعد عن قيمة الصفر كلما دل ذلك على وجود أخطاء تفضيل رقمي عند تسجيل بيانات العمر.

ويمكن توضيح طريقة حساب مقياس مايرز من خلال اتباع الخطوات التالية:

أ-ايجاد عدد السكان المبتدئ بكل من أحاد العمر من (0 الى 9) وعلى مدى الاعمار من (10 فأكثر) فلأعمار المبتدئة ب 0 وتكون (10،20...) والاعمار المبتدئة ب 1 تكون (11،21....) وهكذا لباقي الاحاد وحتى الوصول الى الأعمار المبتدئة ب 9 فتكون (19،29...)

ب-تكرار نفس الخطوة السابقة ولكن على مدى الاعمار من (20فأكثر)

ج -يضرب عدد السكان الذي تم الحصول عليه في الخطوة أ أعلاه في المعامل (r+1) حيث r تمثل آحاد العمر فالأعمار التي أحادها 0 تضرب في المعامل 1 والأعمار التي احادها 1 تضرب في المعامل 2.... وهكذا.

د-يضرب عدد السكان الذي تم الحصول عليه من الخطوة ب أعلاه في متمم المعامل المستخدم في الخطوة ج أعلاه بحيث يصبح مجموع العاملين يساوي (10) فمثلا الأعمار التي أحادها 0 تضرب في المعامل 10-1=9 وهكذا لبقية الأعمار.

هـ-بجمع كل رقمين متقابلين تم الحصول عليهما من الخطوتين ج، د أعلاه يتم الحصول على المجموع المدمج أو المختلط للسكان لكل رقم من الأرقام العشرة من 0 الى 9 ثم ايجاد المجموع الكلي عموديا وحساب النسب المئوية عموديا.

وأخيرا يحسب انحراف كل من هذه النسب عن (% 10)ومجموع هذه الانحرافات بغض النظر عن اشاراتها يمثل مقياس مايرز.

مثال:

r	مجموع الاعمار 10-89	المعامل1	الناتج1	مجموع الاعمار 20-89	المعامل2 المتمم	الناتج2	المجموع المختلط	التوزيع %	الانحراف عن 10%
0	365	1	365	335	9	3015	3380	38.61	28.61
1	21	2	42	14	8	112	154	1.76	8.24-
2	91	3	273	72	7	504	777	8.87	1.13-
3	29	4	116	17	6	102	218	2.49	7.51-
4	33	5	165	23	5	115	280	3.20	6.80-
5	225	6	1350	199	4	796	2146	24.51	14.51
6	24	7	168	16	3	48	216	2.47	7.53-
7	48	8	384	37	2	74	458	5.23	4.77-
8	83	9	747	49	1	49	796	9.09	0.91-
9	33	10	330	22	0	0	330	3.77	6.23-
المجموع	952			784			8755	100	86.24

نلاحظ من الجدول أعلاه أن قيمة مؤشر مايرز للذكور هي 86.24 وهي بعيدة جدا عن الصفر مما يدل على وجود تفضيل لبعض الأعمار، وهذا يدل على أن الأعمار لم تذكر بدقة، حيث أظهرت النتائج وجود تفضيل للأعمار المنتهية بالأرقام 0 و5 ب 38.6 و24.51 على التوالي، في حين مثلت الأعمار المنتهية بالرقم 1 أقل نسبة تفضيل ب 1.76.

3-مقياس سكرتارية الامم المتحدة Ages-Sex Accuracy Index:

في عام 1952 اقترحت الامم المتحدة مقياسا للتعرف على درجه دقه بيانات العمر والنوع يعرف بالرقم القياسي لسكرتارية الامم المتحدة ويعتمد هذا المقياس في حسابه على ايجاد نسبة النوع والعمر حسب الفئات العمرية الخماسية ولكلا الجنسين حساب هذا المقياس كما

وتعتمد هذه الطريقة بصورة أساسية على حساب ثلاثة مؤشرات تتعلق بنسب النوع والعمر للسكان الذكور والإناث في فئات عمرية خمسية

يمتاز هذا المقياس بكونه ممكن التطبيق في حالة عدم توفر بيانات التوزيع العمري حسب آحاد العمر.

❖ **نسبة العمر Age Ratio:**

تعرف نسبه العمر بانها عدد الأشخاص في فئة معينة من فئات الأعمار لكل مائة من متوسط عدد الأشخاص في الفئتين السابقة واللاحقة، وهو من المؤشرات التي يمكن استخدامها للكشف عن وجود عيوب في بيانات العمر للسكان، ويأخذ الصيغة التالية:

$$_5AR_x = \frac{_5P_x}{½\,(_5P_{x-n} + {}_5P_{x+n})} \times 100$$

حيث أن:

$_5P_x$: يمثل عدد السكان من العمر X الى العمر (X+4)

$_5P_{X-5}$: يمثل عدد السكان في الفئة العمرية السابقة للفئة $_5P_x$

$_5P_{X+5}$: يمثل عدد السكان في الفئة العمرية اللاحقة للفئة $_5P_x$

- حساب الانحرافات المطلقة لنسب العمر عن العدد (100) ثم استخراج متوسط هذه الانحرافات (الفروق) للحصول على مقياس دقة العمر (مقياس نسبة العمر).

❖ **نسبة النوع Sex Ratio:**

تعرف نسبه النوع بانها عدد الذكور لكل مائة من الإناث، ويأخذ الصيغة التالية:

$$SR = (_5P^M_x / {}_5P^F_x) \times 100$$

حيث أن:

$_5P^M_x$: يمثل عدد الذكور في فئة عمرية معينة.

$_5P^F_x$: يمثل عدد الاناث في فئة عمرية معينة.

تحليل مؤشر نسبة النوع

قيمة نسبة النوع	التحليل
100	نفس العدد من الذكور والإناث في فئة عمرية معينة
أكبر من 100	عدد الذكور أكثر من عدد الإناث في فئة عمرية معينة
أقل من 100	عدد الذكور أقل من عدد الإناث في فئة عمرية معينة

- حساب المجموع المطلق للفروق المتتالية لنسب النوع عند كل فئة عمرية والفئة العمرية السابقة لها ثم إيجاد متوسط لهذه الفروق المطلقة.

❖ يحسب الرقم القياسي لسكرتارية الأمم المتحدة من خلال استخدم الصيغة الآتية:

AGE–SEX ACCURACY INDEX (UNAI) =3 (S) +M+F

حيث أن:

S: يمثل متوسط فروق نسب النوع.

M: يمثل متوسط الانحرافات في النسب العمر للذكور (اي مقياس دقه العمر للذكور).

F: يمثل متوسط الانحرافات في نسب العمر للإناث (أي مقياس دقه العمر للإناث).

وتعتبر البيانات جيدة إذا لم تتعدى قيمة هذا المؤشر 20، ومتوسطة الجودة إذا كانت محصورة بين 20 و40، أما إذا تعدى 40 فهي سيئة وغير موثوق فيها.

مثال تطبيقي:

3–مؤشر سكرتارية الأمم المتحدة:

فئات العمر	عدد الذكور	نسبة العمر	الانحراف عن 100	عدد الاناث	نسة العمر	الانحراف عن 100	نسبة النوع	الفروق الاولى
4-0	422			260			162.31	
9-5	114	45.60	54.40-	99	57.89	42.11-	115.15	-47.16
14-10	78	76.47	23.53-	82	45.05	54.95-	95.12	-20.03
19-15	90	88.24	11.76-	265	113.73	13.73	33.96	-61.16
24-20	126	113.51	13.51	384	131.51	31.51	32.81	-1.15
29-25	132	112.34	12.34	319	99.07	0.93-	41.38	8.57
34-30	109	97.32	2.68-	260	120.93	20.93	41.92	0.54
39-35	92	98.40	1.60-	111	59.68	40.32-	82.88	40.96
44-40	78	113.04	13.04	112	155.56	55.56	69.64	-13.24
49-45	46	65.71	34.29-	33	38.37	61.63-	139.39	69.75
54-50	62	156.96	56.96	60	210.53	110.53	103.33	-36.06
59-55	33	58.93	41.07-	24	52.75	47.25-	137.50	34.17
64-60	50	111.11	11.11	31	86.11	13.89	161.29	23.79
65+	57			48			118.75	-42.54
المجموع			276.29			493.34		399.12
المتوسط			23.02			41.11		30.70
مؤشر الامم المتحدة			**156.23**					

من الجدول نلاحظ أن انحرافات نسب العمر عن القيمة 100 بلغ مجموعها 276.29 للذكور و493.34 للإناث وهي تقريبا ضعف نسبة الرجال مما يدل على وجود أخطاء في بيانات العمر، وأن هذه الأخطاء كانت عند الإناث أكبر مما هي عليه عند الذكور، حيث بلغت قيمة مقياس دقة العمر للذكور 23.02 و41.11 للإناث، مما يدل على أن الاناث أقل دقة في الادلاء بأعمارهن من الذكور.

ونلاحظ أن قيمة المؤشر هي 156.23 وهي قيمة كبيرة جدا وأكبر من 40 وبالتالي ووفقا لمقياس الامم المتحدة فالبيانات عالية الخطأ ولا يمكن الاعتماد عليها.