

المنهجية الإحصائية

أكدت نتائج تحليل الارتباط أيضاً صلاحية اختيار المتغيرات وتصنيفها، حيث أوجدت مصفوفة الارتباط للمتغيرات المطبّعة التي تأكد التقرير من أنها تسير في اتجاه المؤشر المركب نفسه؛ كما تأكد من ضرورة إدراج المتغيرات، التي لها معاملات ارتباط عالية (أكبر من 0.9)، مع بعض المتغيرات الأخرى.

البيانات المستخدمة وجودتها

بلغ عدد المتغيرات الداخلة في بناء المؤشرات القطاعية الستة لمؤشر المعرفة العربي الحالي 432 متغير حصل عليها من مصادر مختلفة عدة، أهمها أجهزة الأمم المتحدة، والبنك الدولي، واليونسكو، والاتحاد العالمي للاتصالات، والاتحاد الأوروبي، ومنظمة التعاون والتنمية الاقتصادية، وغيرها³.

راجع فريق العمل البيانات أكثر من مرة للتأكد من عدم وجود أخطاء إدخال، وتعامل بالتالي مع البيانات على أنها خالية من الأخطاء. وقد أعيد حساب بعض المتغيرات التي ترتبط بعوامل أخرى، مثل عدد السكان أو الناتج المحلي الإجمالي، بعد حذف أثر الحجم. ولم تُقدّر القيم المفقودة تحرياً للشّافية والبساطة، ولإمكانية إعادة الحصول على النتائج؛ كما كفاً استخدام صيغة الوسط الحسابي في حساب المؤشر تقدير القيم المفقودة للمتغير بقيمة وسطه الحسابي. ومثلما هو معهود في حالات مماثلة، تُحتسب المؤشرات المركبة بتجاهل القيم المفقودة، واستخدام المتاح فحسب من بيانات كل دولة⁴.

جودة البيانات

ينبغي للبيانات الداخلة في بناء المؤشرات القطاعية أن تتصف بخصائص إحصائية، أهمها غياب القيم الشاذة، والالتواء أو التفرطح الحاد؛ لأن عدم توفر هذه الخصائص يؤدي إلى تحيز قيم المؤشر. هنا، كان لا بد من التحقق من هذه الخصائص قبل حساب المؤشر المركب وعمل المعالجات المناسبة - في حال عدم تحققها - لتجنب عدم دقة (تحيز) قيم المؤشر المركب.

يمثل بناء مؤشر المعرفة العربي للعام 2016 عملية مستمرة ومتطورة يُبنى فيها على الإصدار السابق عام 2015، حيث احتفظ بالهيكل العام للمؤشر؛ المكوّن من ستة مؤشرات قطاعية مركبة هي: التعليم ما قبل الجامعي، والتعليم التقني والتدريب المهني، والتعليم العالي، والاقتصاد، وتكنولوجيا المعلومات والاتصالات، والبحث والتطوير والابتكار. وبهدف رفع كفاءة المؤشر في قياس المعرفة، خضعت منهجية بناء المؤشر لأوجه تطوير عديدة شملت تحديث البيانات، وإدراج متغيرات جديدة وحذف بعض المتغيرات القديمة، كما يتضح خلال فصول التقرير السابقة. وشمل التطوير أيضاً التوسع في التحليلات الإحصائية للتحقق من اتساق المؤشر، واختيار المتغيرات، وتحديد الأوزان، واكتشاف القيم الشاذة والالتواء والتفرطح؛ بقصد التأكد من ملائمة البيانات لحساب المؤشر على نحو أدق. وفي ما يلي، أوجه التطوير في خطوات بناء هذه المؤشرات، مفصلة إلى حد ما.

اختيار المتغيرات

بوجه عام، احتفظت قائمة المتغيرات (المؤشرات الفردية) الداخلة في بناء الإصدار الثاني من مؤشرات المعرفة الستة 2016 بالمتغيرات المختارة في الإصدار الأول عام 2015. وقد أضيفت و/أو حُذفت بعض المتغيرات المستخدمة في بناء المؤشرات الستة، واختلف حجم الإضافة والحذف من مؤشر إلى آخر؛ حيث بلغ حجم التغيرات أقصاه في مؤشر التعليم التقني والتدريب المهني ومؤشر البحث والتطوير والابتكار.

تحقق فريق التقرير من مدى الاتساق في اختيار المتغيرات وهيكل تصنيفها على محاور رئيسية وفرعية مختلفة باستخدام نتائج تحليل المركبات الرئيسية وقيم معاملات ألفا كرونباخ؛ حيث تعدت نسبة التباين المفسر في أغلب الحالات 50 في المئة¹ وتعدت قيم معامل ألفا كرونباخ 0.70 في الأغلب الأعم من الحالات².

عدد المتغيرات التي فيها قيم شاذة أو التواء وتفرطح

المؤشر	عدد المتغيرات التي فيها قيم شاذة/ التواء وتفرطح
الاقتصاد	10
التعليم العالي	15
تكنولوجيا المعلومات والاتصالات	8
التعليم ما قبل الجامعي	12
البحث والتطوير والابتكار	17
التعليم التقني والتدريب المهني	4

تزيد على 2؛ كما أن كون قيمة التفرطح المطلقة أكثر من 3.5 يدل على أن المتغير يتصف بدرجة حادة من التفرطح. وتحتاج المتغيرات ذات الالتواء و/أو التفرطح الحاد إلى معالجة إحصائية قبل استخدامها في حساب المؤشرات المركبة؛ ويُعد تحويل اللوغاريتم من أشهر التحويلات المستخدمة في هذا الشأن.

بتطبيق قواعد اكتشاف القيم الشاذة وكذلك الالتواء و/أو التفرطح الحاد في بيانات المتغيرات الداخلة في بناء المؤشرات القطاعية لمؤشر المعرفة العربي لعام 2016، وُجد أن هناك 66 متغير فيها قيم شاذة أو التواء وتفرطح حاد. ويعرض الجدول 9 توزيع هذه المتغيرات بحسب مؤشرات المعرفة.

اتضح أيضًا أن عدد القيم الشاذة لم يزد على اثنين لأي متغير، وأن معالجة القيم الشاذة ساهمت بشكل فعال في حل مشكلة الالتواء والتفرطح الحاد في كل الحالات؛ دون الحاجة إلى إجراء تحويلات.

التطبيع

استُخدمت طريقة إعادة القياس (أو طريقة الأكبر - الأصغر) في تطبيع المتغيرات حيث يُشير الأكبر والأصغر إلى أكبر قيم المتغير المتاحة وأصغرها على التوالي؛ فيما تقع قيم المتغيرات في الفترة [1,100]، حيث تُشير القيم الأعلى إلى نتائج أفضل. وتعتمد صيغة التطبيع على ما إذا كان المتغير جيّدًا، أي يرتبط بعلاقة طردية مع المؤشر العام؛ أو سيئًا، أي

وفي ما يلي استعراض كيفية اكتشاف القيم الشاذة ومعالجتها؛ ثم مناقشة كيفية التحقق من وجود الالتواء والتفرطح الحاد، ومعالجته.

القيم الشاذة

تُعد قيمة المتغير شاذة إذا وقعت خارج سياق البيانات، وهو فترة حديها الأدنى والأعلى اللذين يُحددان بناءً على مقاييس تركيز البيانات (الربيع الأول والربيع الثالث) وتشتتها (المدى الربيعي)؛ كما يلي:

الحُد الأدنى = الربيع الأول - 1.5 * المدى الربيعي
الحُد الأعلى = الربيع الثالث + 1.5 * المدى الربيعي

وتُعد جميع القيم الواقعة داخل سياق البيانات، أي بين الحدّين الأدنى والأعلى، غير شاذة؛ بينما تُعد القيم شاذة إذا وقعت خارج السياق، أي أكبر من حدّه الأعلى أو أقل من حدّه الأدنى.

إذا كانت هناك قيمة شاذة كبرى، تُعالج بإحلال ثاني أكبر قيمة محلها؛ وإذا كانت هناك قيمة شاذة صغرى، تُعالج بإحلال ثاني أصغر قيمة محلها. وتكرّر هذه الخطوة في حال وجود أكثر من قيمة شاذة.

الالتواء والتفرطح

تُشير الأدبيات الدولية إلى أن المتغير يتصف بدرجة حادة من الالتواء إذا كانت قيمة الالتواء المطلقة

يرتبط بعلاقة عكسية مع المؤشر العام. ويمكن تطبيق المتغيرات الجيدة من خلال المعادلة التالية:

$$1 + \left(\frac{\text{قيمة المتغير الخام للدولة - أصغر قيمة خام للمتغير بين الدول}}{\text{أكبر قيمة خام للمتغير بين الدول - أصغر قيمة خام للمتغير بين الدول}} \right) \times 99 = \text{قيمة المتغير المطبوعة للدولة}$$

وفي حالة المتغيرات السيئة، أي المعكوسة، تُعدّل هذه المعادلة لتعكس القيم، بحيث تكون القيم الأعلى في الاتجاه الإيجابي؛ كما يلي:

$$1 + \left(\frac{\text{أكبر قيمة خام للمتغير بين الدول - قيمة المتغير الخام للدولة}}{\text{أكبر قيمة خام للمتغير بين الدول - أصغر قيمة خام للمتغير بين الدول}} \right) \times 99 = \text{قيمة المتغير المطبوعة للدولة}$$

وشملت هذه البيانات 22 دولة عربية. وأُتبعَت بسلسلة من التجميعات المتتالية، بدءاً من المتغيرات (المستوى الأكثر تفصيلاً)، وانتهاءً ببلوغ المؤشر العام.

نظراً إلى تعذر الحصول على بيانات كل المحاور الرئيسية لكل دولة، ورغبةً في الحفاظ على مستوى دقة مناسب، لم يُحتسب المؤشر إلا إذا توفرت بيانات اثنين على الأقل من محاوره الرئيسية. من ناحية أخرى، استُبعد عرض نتائج المحاور الفرعية التي اقترَح حسابها في حال عدم توفّر بيانات متغيرات ثلاث دول على الأقل.

استُخدم أسلوب التجميع الحسابي في حساب قيم كل المؤشرات القطاعية المركبة لمؤشر المعرفة؛ ورُكّب المؤشر المركب (CI) بتجميع المحاور الرئيسية الحسابية (SI_j)؛ بتطبيق المعادلة التالية:

$$CI = \sum_{j=1}^n w_j \times SI_j$$

تحديد الأوزان

عموماً، اعتمد الإصدار الحالي للعام 2016 لمؤشرات المعرفة على الأوزان المستخدمة في بناء الإصدار السابق عام 2015 التي تنوّعت ما بين أساليب الأوزان المتساوية، وتخصيص الميزانية، والتحليل العملي. وقد خضعت بعض هذه الأوزان لعدد من التعديلات بسبب التغير في الهيكل العام للمؤشرات، الناجم عن إضافة و/ أو حذف بعض المتغيرات.

قُدّرت الأوزان إحصائياً أيضاً باستخدام التحليل العملي لكل متغير فردي لتكوين المؤشر القطاعي بطريقتين. الأولى، استخدام قيم تحميل عامل واحد على المتغيرات الفردية المقترحة لقياس المؤشر القطاعي محلّ الدراسة والتكوين؛ والثانية استخدام قيم تحميل ثلاثة معاملات - وليس معاملاً واحداً - على المتغيرات الفردية المقترحة؛ بهدف اقتراح بدائل متعدّدة للأوزان، كأداة مساعدة أمام الباحث في تحديد الأوزان للمتغيرات المختلفة.

حساب المؤشر

استُخدمت أحدث البيانات المتاحة في حساب الإصدار الثاني لمؤشر المعرفة العربي 2016 وأفضلها؛

- 1 .Hair et al., 2015
- 2 .Tavakol and Dennick, 2011
- 3 للتعرف على مصادر بيانات مؤشرات المعرفة العربي، أنظر الملحق.
- 4 .Cornell University et al., 2015