

معهد البحوث الاجتماعية والاقتصادية المسحية

ورشة عمل تقييم السياسات والبرامج

الدوحة، قطر

19-22 يناير 2015

تمهيد : جلسة 3

- مراجعة تجارب التحكم العشوائية (RCTs)
- مخاوف تصميم التقييم :
 - انحياز الإنتقاء
 - انحياز حذف المتغير
 - علاقات زائفة
 - تفسيرات بديلة
- صدق الداخلي /الخارجي
- مقدمة عن التصاميم شبه التجريبية :
 - مقارنات قبلية مع مجموعة التحكم
 - السلسلة الزمنية المتقطعة
 - المطابقة باستخدام نقاط النزوع

الخلاصة: تجارب التحكم العشوائية

- تصميم بحثي قوي يعتمد على التعيين العشوائي في مجموعتي المعالجة والتحكم لإنشاء واقع مغاير مقنع .
- أعضاء مجموعتي المعالجة والتحكم لهم احتمالات متساوية في تلقي العلاج .
- تنشأ مجموعات "متساوية في التوقع"
- يمكن إرجاع الاختلاف في النتائج إلى تأثير العلاج .
- تسمح باستدلال سببي قوي

ما هي المشاكل التي يمكن أن يحلها التعيين العشوائي؟

- العشوائية تصحح خطأ المواصفات، مثل انحياز الإنتقاء. يشمل هذا :
 - انحياز حذف المتغير
 - العلاقات الزائفة
 - التفسيرات البديلة

انحياز الإنتقاء

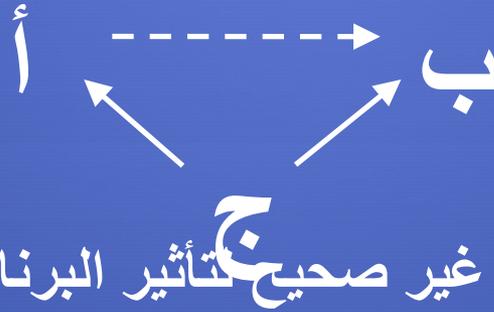
- قد يختلف الأفراد الذين يختارون المشاركة في البرنامج في طرق هامة عن الأفراد الذين لا يشاركون .
- العودة إلى مثال برنامج محو الأمية المالية : هل البرنامج /المعالجة تسبب في جعل الناس تدخر أكثر؟
- بدون التعيين العشوائي، الناس الذين هم أكثر احتمالية للادخار قد يكونوا أكثر احتمالية في المشاركة في البرنامج، لذلك يبدو وكأن البرنامج يقودهم إلى الادخار، بينما في الحقيقة أنهم كانوا أكثر احتمالية للقيام بهذا في أي حالة من الأحوال .
- هنا نرتبك في تأثير الميل الأولي للأفراد للادخار مع مشاركتهم في البرنامج .
- مع التعيين العشوائي، الناس الذين هم أكثر احتمالية للادخار هم متساوين في التوقع أن يتم تمثيلهم في مجموعتي المعالجة والتحكم، لذلك يمكن تقديم التأثير الحقيقي للبرنامج .

أنواع انحياز الإنتقاء

- انحياز حذف المتغير
- إرجاع تأثير الناتج إلى البرنامج أو السياسة عندما تكون في الحقيقة مسببة من قبل بعض العوامل الأخرى التي لا تُأخذ في الاعتبار في التصميم أو النموذج .
- العلاقات الزائفة
- تحدث العلاقات الزائفة إذا حددنا علاقة بين أ & ب عندما يوجد في الواقع متغير ثالث، ج، يؤثر على كل منهما .
- التفسيرات البديلة
- أي سبب يجعل الأفراد في مجموعة المعالجة يختلفون عن مجموعة التحكم، بالتالي لا يمكن إرجاع التأثيرات وحدها إلى البرنامج أو التدخل .

مثال

- إذا اعتقدنا أن تدخل محو الأمية المالية (أ) (يؤثر على سلوك الادخار) ب)، فإننا نهتم بالعلاقة السببية بين أ & ب
- مع ذلك، قد يكون عمر (ج) (يؤثر على كل من المشاركة في التدخل و سلوك الادخار . ربما الأفراد الأكبر سنًا أكثر احتمالية للادخار والمشاركة في التدخل .



- في هذه الحالة، قد نقوم بتقدير غير صحيح لتأثير البرنامج .

تصحيح انحياز الإنتقاء

- بالعشوائية، يمكن الأخذ في الاعتبار تفسيرات بديلة للعلاقة بين أ – ب لأن مجموعتي المعالجة والتحكم متساويتين في توقع تلقي العلاج .
- بدون العشوائية، يمكن تصحيح انحياز الإنتقاء عن طريق تضمين ضوابط إحصائية في النماذج (على سبيل المثال، التحكم في المتغيرات المربكة مثل العمر، والجنس، وسلوك الادخار المسبق، وما إلى ذلك).

مقارنة الصدق الداخلي والخارجي

● الصدق الداخلي

- الحد الذي يمكن للباحث تحديده بدون شك أن أ □ ب
- تجارب التحكم العشوائية تميل إلى إعطاء الأولوية إلى الصدق الداخلي.

● الصدق الخارجي

- الحد الذي عنده العلاقة السببية تحمل تغييرات في الناس، والسياقات، والمعالجات والنتائج .
- يشار إليها أيضاً على أنها القدرة على التعميم .
- ليست "جيدة" أو "سيئة" - "تعتمد على احتياجات البرنامج والجهات المعنية به، بما في ذلك المُقيِّم .
- التوتر بين الصدق الداخلي و الخارجي هو دائماً شيء يجب أن يتعامل معه الباحث /المُقيِّم .

تهديدات الصدق الداخلي: قائمة التدقيق

- الأسبقية الزمنية الغامضة
 - هل أ تسبق ب؟
- الاختيار
 - هل مجموعتي المعالجة والتحكم مختلفتين بطريقة منتظمة فيما عدا تلقي التدخل؟
- التاريخ
 - هل كان هناك شيء آخر حدث في نفس الوقت مع تنفيذ البرنامج قد أثر على النتائج؟
- النضوج
 - قد يكون الناتج قد حدث بدون البرنامج نتيجة التغييرات الطبيعية؟
- التراجع
 - هل المشاركين حالات شديدة أو متطرفين؟
- الإثهاك
 - هل تم إسقاط المشاركين من الدراسة بطريقة منهجية؟
- الاختبار
 - هل تم استخدام نفس الاختبار للقياسات القبليّة والبعدية؟
- الأدوات
 - هل القياس المستخدم (على سبيل المثال، المسح، بروتوكول المقابلات، الاختبار (مناسب) .
- تأثيرات مضافة وتفاعلية للتهديدات على الصدق الداخلي
 - هل هناك شيء آخر يحدث يعمل على تعقيد العلاقة بين السبب والتأثير؟

تصاميم البحث شبه التجريبي

- التقييم العشوائي للمعالجة نادر للغاية .
- غالبًا يطلب الباحث تقييم البرنامج بعد تنفيذه بالفعل، لذلك لا يمكن التحكم في التعيين في المعالجة .
- في هذه الحالة، قد يكون على المُقيم استخدام البيانات الموجودة لمحاولة تقريب الاستدلال النسبي .
- كيف يمكن للمُقيم إنشاء واقع مغاير في غياب التعيين العشوائي؟
- استخدام التصاميم شبه التجريبية يمكن أن يوفر إمكانية لعزل تأثيرات المعالجة .

التصاميم شبه التجريبية القوية

● التصاميم شبه التجريبية المتعددة

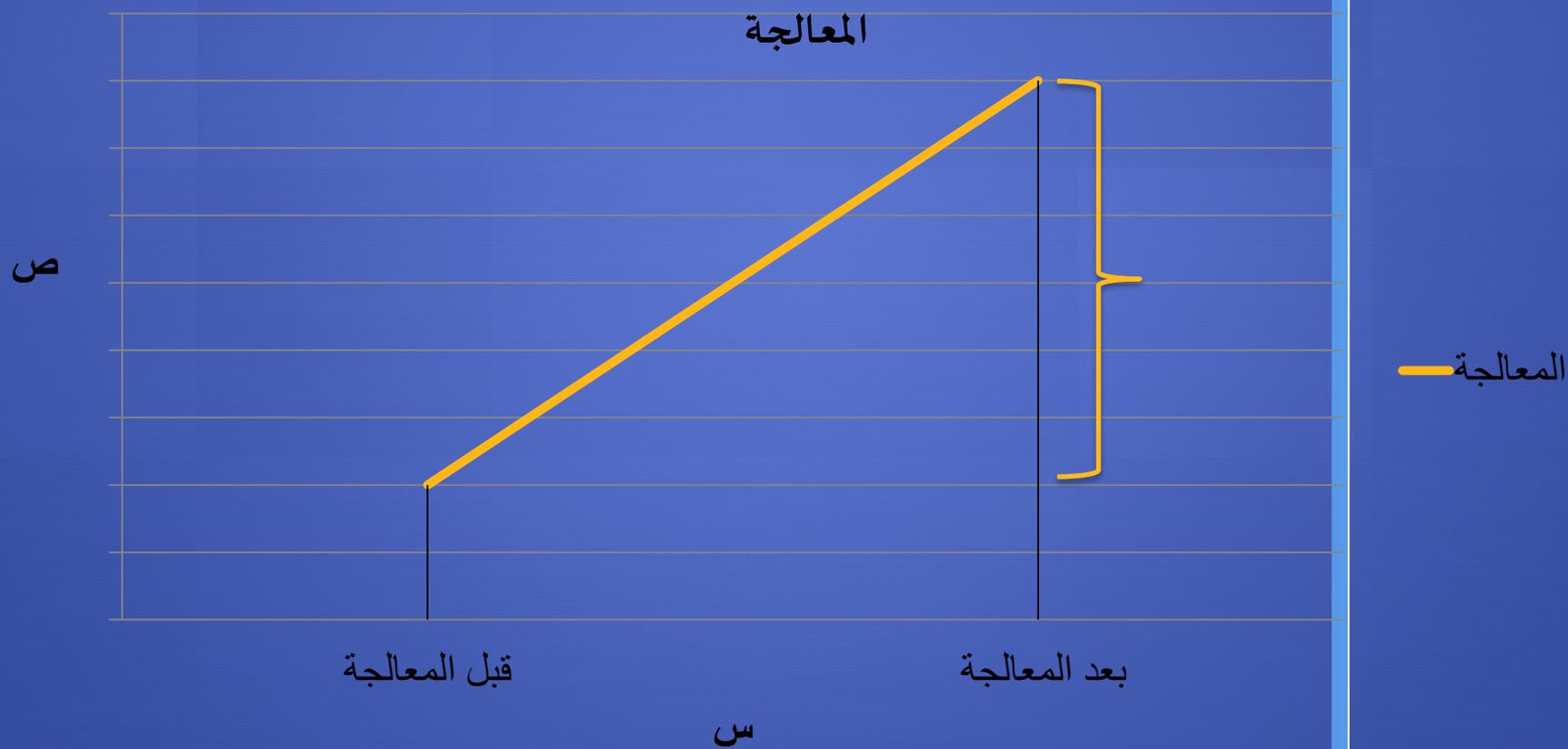
● تتغير اعتمادًا على كيفية إنشاء مجموعة التحكم أو المقارنة، وما تم عمله من مقارنات .

● سنركز على ثلاثة تصاميم شبه تجريبية مشتركة لا تحتاج من المقيم إدارة المعالجة: مقارنة ما قبل وما بعد مع مجموعة التحكم، والسلسلة الزمنية المتقطعة، والمطابقة .

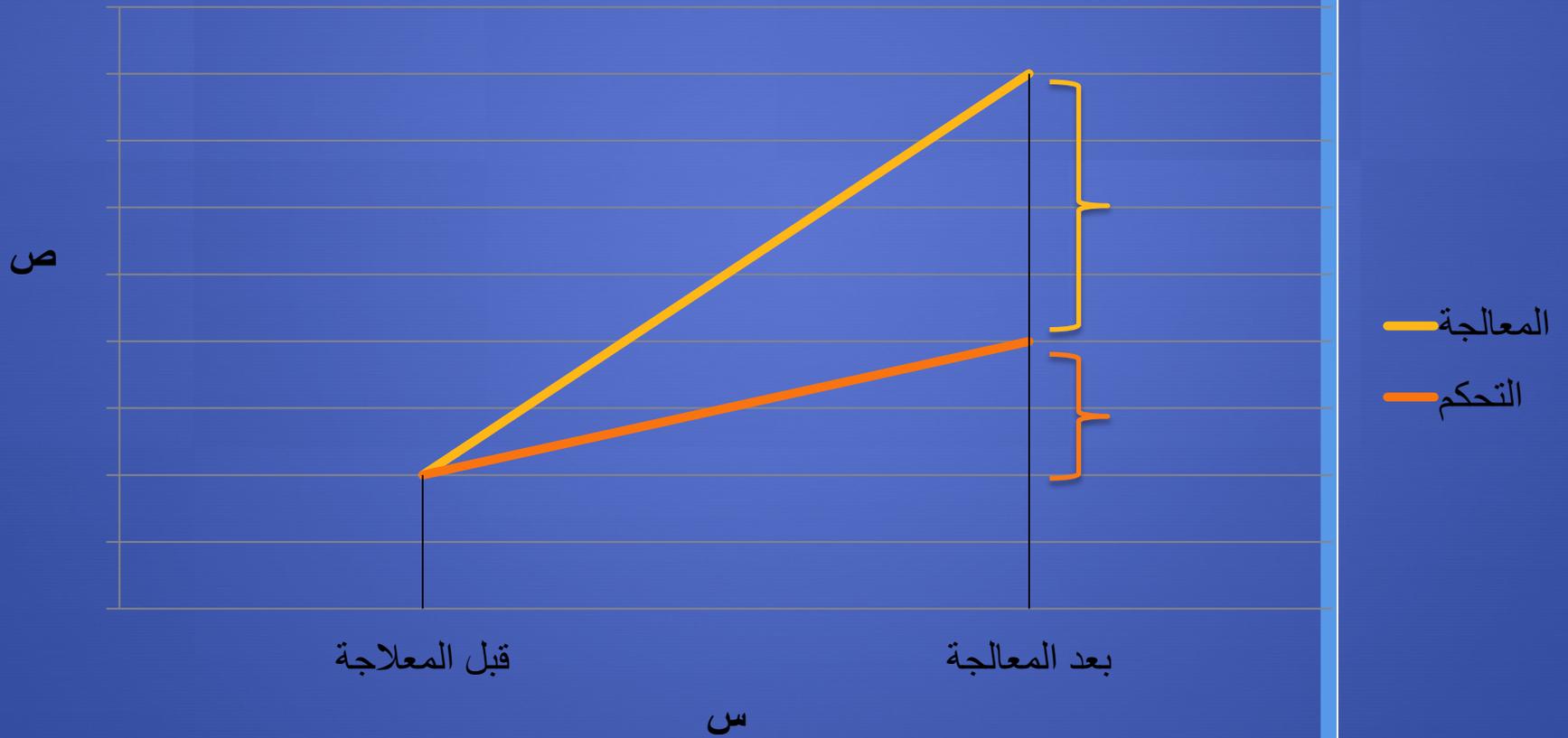
مقارنة ما قبل وما بعد مع مجموعة التحكم

- تصميم من مجموعتين، كل منها له نقطتين من البيانات .
- التأثير المقدر هو الفرق بين الاختلافات .
- المميزات التي تفوق المقارنة ما قبل وما بعد بدون مجموعة التحكم
 - قد تلتقط المتغيرات التي تم حذفها، والتفسيرات البديلة التي تؤثر على المجموعتين .
- القيود
 - قياسات محدودة للنتائج .
 - صعوبات في إيجاد مجموعة تحكم مناسبة .
 - تهديدات للصدق: النضج، الموسمية، التاريخ .

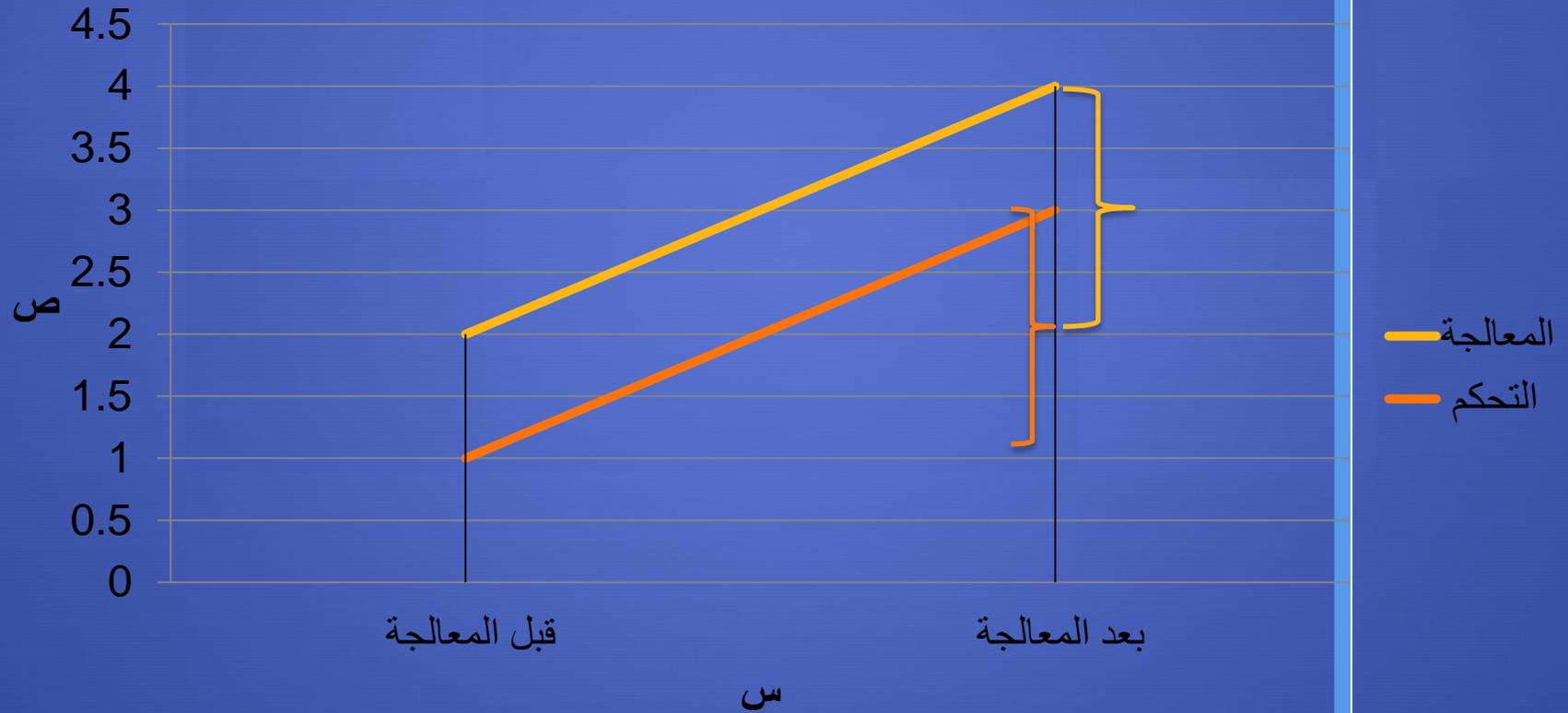
مقارنة ما قبل وما بعد بدون مجموعة التحكم



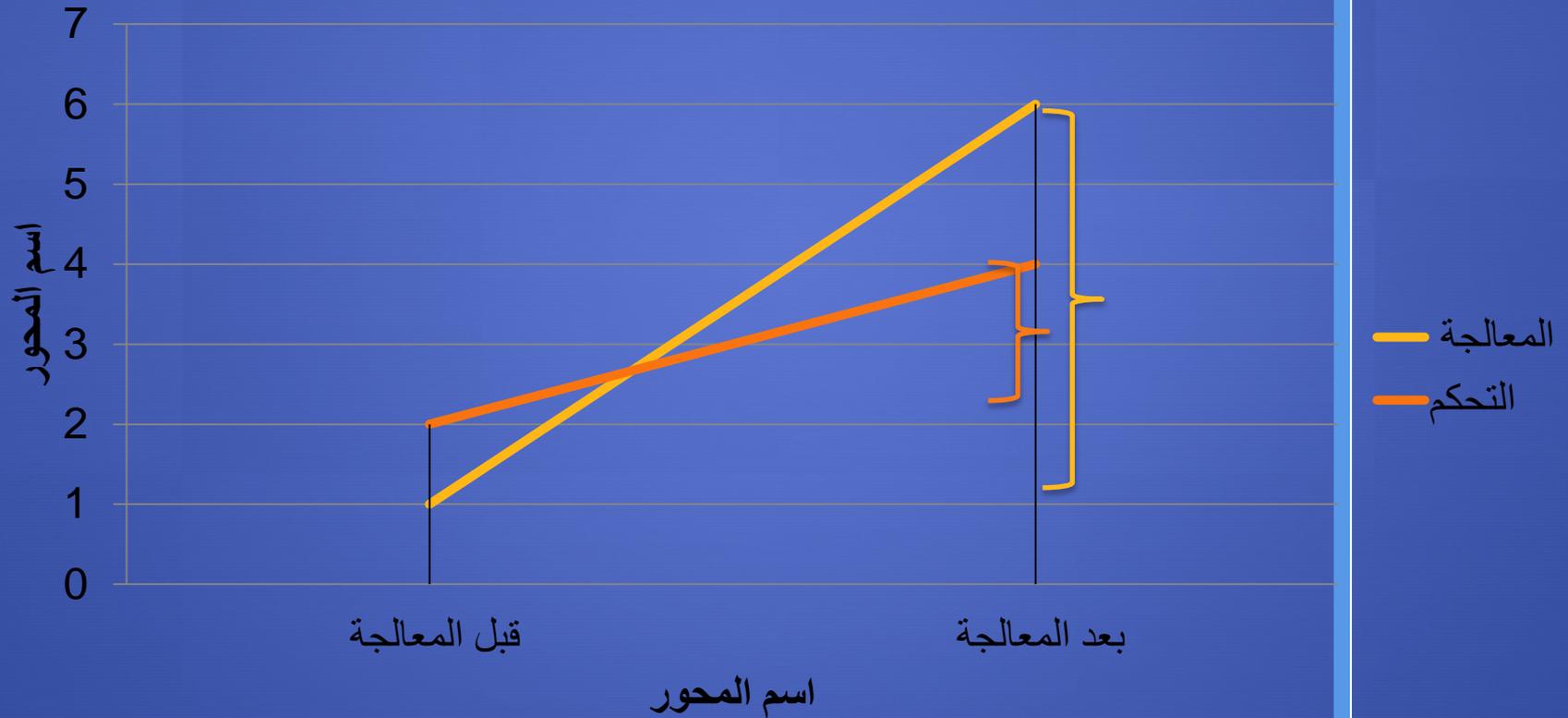
مقارنة ما قبل وما بعد مع مجموعة التحكم



مقارنة ما قبل وما بعد مع مجموعة التحكم



مقارنة ما قبل وما بعد مع مجموعة التحكم



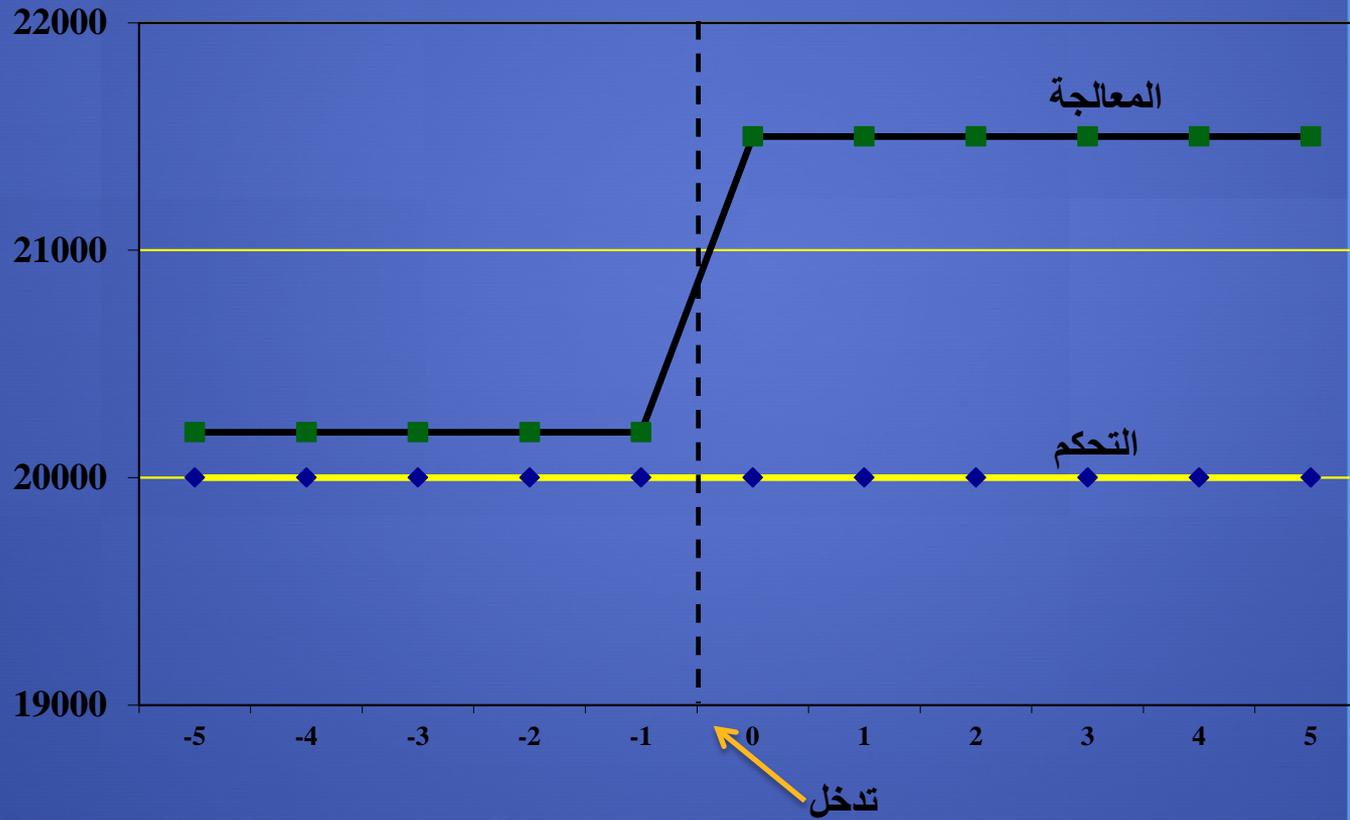
مقارنة ما قبل وما بعد مع مجموعة التحكم

- كيف يمكن اختيار مجموعة تحكم؟
 - تذكر يمكن أن يكون لدينا قيود مثل مدخل الجهة المعنية، توافر بيانات، وما إلى ذلك .
 - الإستراتيجية الجيدة هي التفكير في تجربتك المثالية ثم تحاول التقريب لأفضل ما لديك من قدرات مع اخذ قيودك بعين الاعتبار
- الإستراتيجية العامة :اختيار مجموعة تحكم تستند على قيم المتغير المتعلقة بالنتائج .
 - اختيار مخصص لمجموعة يجب أن تبدو وكأنها معالجة (على سبيل المثال، مجموعات مهاجرين أخرى في الدول المجاورة .)
 - الضبط الإحصائي لعوامل مختلفة (ع (ضمن الانحدار .
 - مطابقة رسمية تستند على خصائص ما قبل التدخل للمعالجات والتحكمات (على سبيل المثال، أرباح ما قبل التدخل)

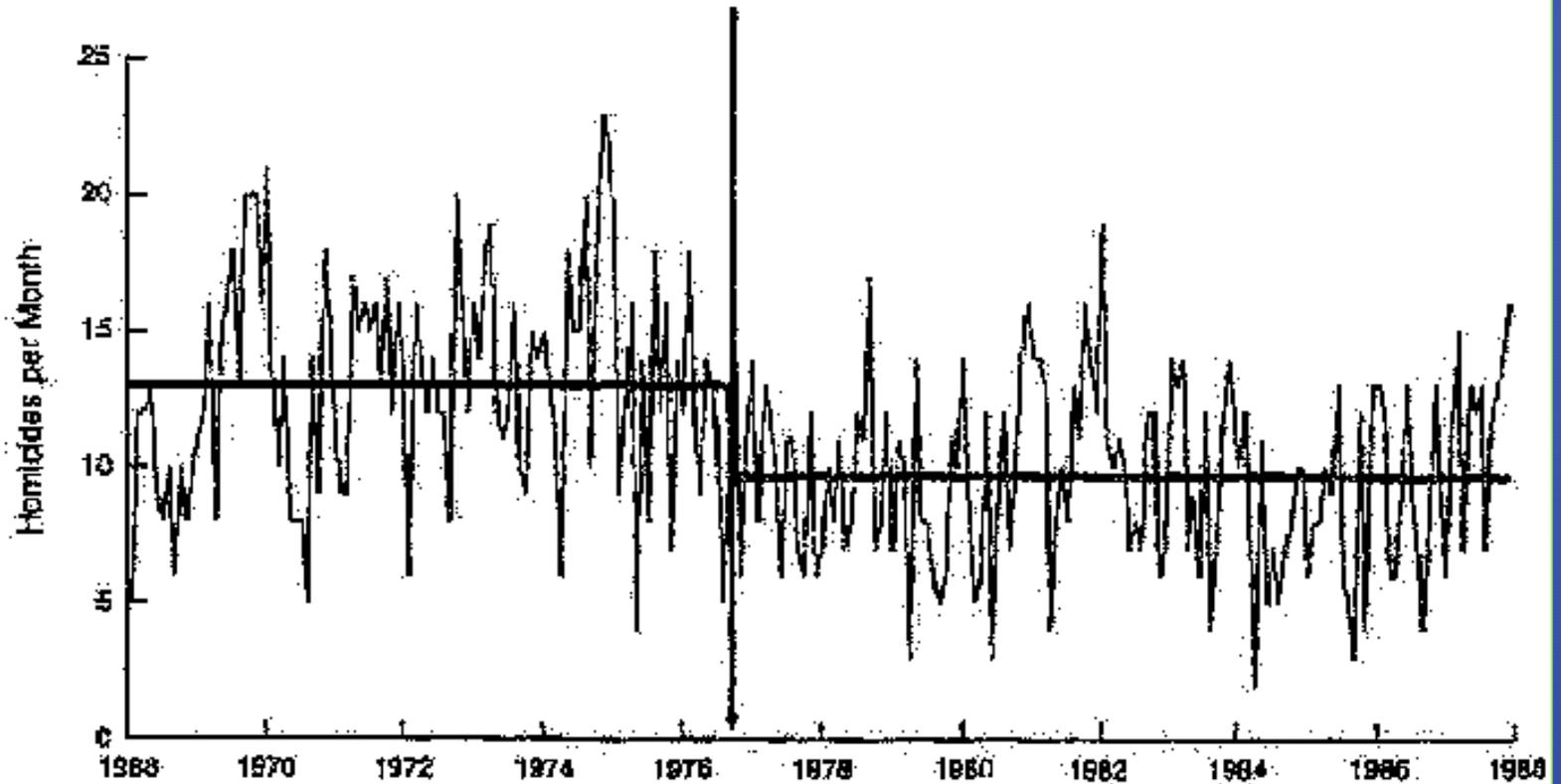
السلسلة الزمنية المتقطعة

- السلسلة الزمنية: العديد من الملاحظات التي تتم على نفس النتائج تبعاً مع مرور الوقت .
- المتقطعة: معالجة تنشأ نقطة قطع مفترضة للسلسلة الزمنية .
- مجموعة التحكم: في بعض الأحيان -تساعد في التهديدات على المصادقية .
- تمديد ما قبل /ما بعد التصميم: أكثر من ملاحظة قبل وبعد المعالجة
- الواقع المغاير
- في حالة عدم وجود مجموعة تحكم، الاتجاه نحو استقرار ما قبل المعالجة
- في حالة وجود مجموعة تحكم، قارن التغييرات ما قبل وما بعد في مجموعة المعالجة مع التغييرات ما قبل وما بعد في مجموعة التحكم .

السلسلة الزمنية المقطوعة

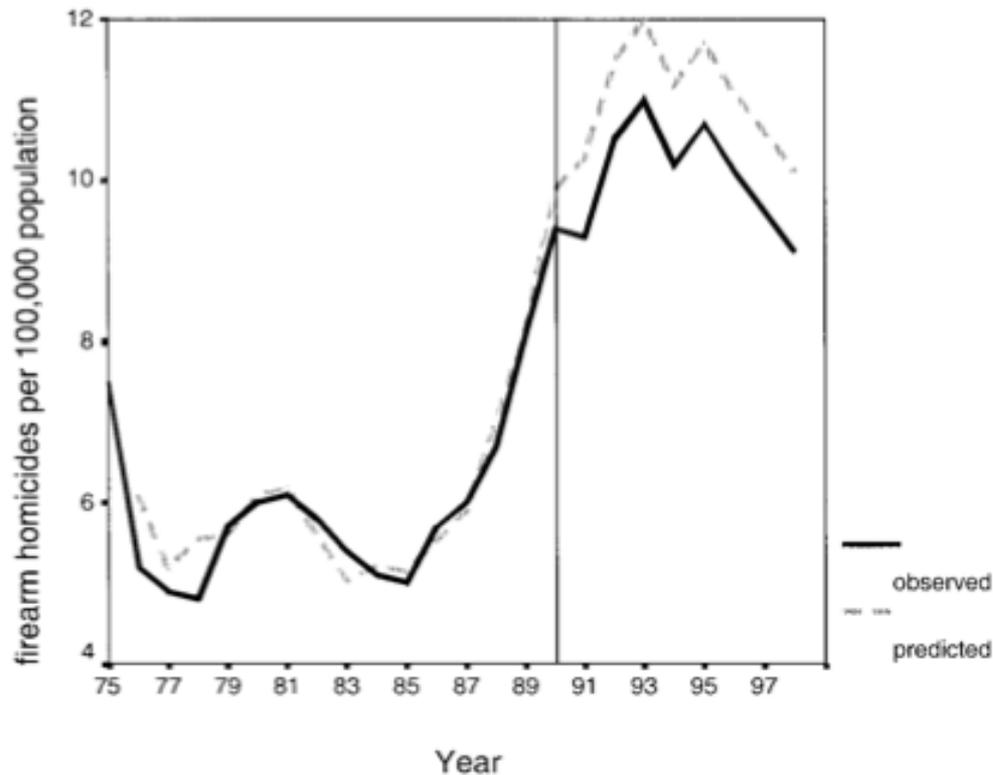


السلسلة الزمنية المقطوعة



(في الشهر) عدد جرائم القتل بالأسلحة النارية في واشنطن دي سي

TS: مجموعة مقارنة محتملة لـ جوار ولاية ميريلاند



السلسلة الزمنية المتقطعة

● أكثر فاعلية عندما :

● يكون التدخل سريع وموثق جيداً .

● إذا انتشر التدخل على عدة أشهر أو سنوات، قد لا نعرف متى نتوقع أن نرى تأثيرات .

● فرصة أكبر للتاريخ أو متغير محذوف .

● النظرية تعني تأثير فوري

● إذا لم يتم التوقع بالتأثير للتعبير عن نفسه لعدة سنوات، يكون هناك فرصة أكبر للتاريخ أو المتغيرات المحذوفة أن تنشأ انحياز .

● مثال -تأثيرات القيود على إعلانات السجائر على سرطان الرئة .

● ملاحظات متعددة ما قبل وما بعد المعالجة .

● التهديدات على صدق البحث :مثل ما قبل وما بعد مع التحكم

المطابقة

- تهدف المطابقة إلى إنشاء أزواج من أفراد العينة المعالجة وغير المعالجة التي لها احتمالية مماثلة لتلقي العلاج، على الرغم من واحدة فقط بالفعل تقوم بهذا .
- تكون مفيدة إذا تم بالفعل التنفيذ للتدخل، ولكن قد لديك الكثير من البيانات عن التعداد العام للسكان، بما في ذلك القياسات الأولية للنتائج .

المطابقة

- نقاط النزوع
- نهج واحد للمطابقة
- نقاط النزوع هي الاحتمالية المتوقعة أن يتلقى الشخص المعالجة استنادًا على جميع المتغيرات المتاحة ما قبل التدخل .
- تعطيك نقاط النزوع متغير واحد يلخص كل شخص مع البعد الذي تهتم به أكثر -بما يعني احتمالية أن يكون معالجًا .
- العملية تتم على مرحلتين:
- ما هي العوامل التي تتوقع احتمالية المعالجة؟
- ما هي احتمالية حصول الفرد على المعالجة؟
- تطبق كل حالة معالجة مع حالة تحكم التي لها نقاط نزوع مشابهة .

المطابقة

● نقاط القوة

- انحياز الإنتقاء: يقارن التغيير في النتائج لأفراد العينة المعالجة مع التغيير في نتائج أفراد التحكم التي لها احتمالية معالجة مماثلة (ولكن غير متطابقة)
- تقارن مباشرة أعضاء مماثلين آخرين لمجموعتي المعالجة والتحكم (باستخدام متوسط تأثير المعالجة "ATE" أو إحصاءات مماثلة).

● نقاط الضعف

- نوعية مجموعة التحكم تعتمد بشكل كبير على كيفية إنشاء نقاط النزوع . هل جميع العوامل ذات الصلة مدرجة في النموذج؟
- يقرر الباحث ما هو المدرج في النموذج، والذي قد يؤدي إلى مخالفة المواصفات .
- ملاحظات محدودة ما قبل وما بعد المعالجة .

تمرين المجموعة

- تخيل مرة أخرى تجارب التحكم العشوائية لتدريب محو الأمية المالية في قطر كتصميم ITS أو تصميم ما بعد وما قبل المطابقة. ما نوع البيانات التي تحتاجها؟ ما هي الانحيازات التي لا يمكن أن يحدث منها التصميم؟ هل هناك من مزايا؟
- استخدم قائمة فحص "التهديدات على الصدق الداخلي للبحث" للنقد وتحسين التصميم.

معهد البحوث الاجتماعية والاقتصادية المسحية

ورشة عمل تقييم السياسات والبرامج

الدوحة، قطر

19-22 يناير 2015

تمهيد :جلسة 4

- القياس
- أنواع ومصادر البيانات
- التعريف الإجرائي
- خطأ القياس
- ثبات وصدق البحث
- نظرية القياس
- خطأ عشوائي
- الانحياز

خطوات التعريف الإجرائي: بحث عن التحويلات المالية

نحن مهتمون بدراسة التحويلات المالية من منظور السياسة:

من الذي يرسلها، وبأي مبلغ، ولأي غرض؟

كيف يمكننا زيادة سلوك الادخار بين العمال الأجانب وأسرهم؟

ما هي التدفقات المالية الإجمالية بين الدول وكيف تتغير مع مرور الزمن؟

خطوات التعريف الإجرائي: بحث عن التحويلات المالية

ما هو التحويل؟

إرسال الأموال إلى مكان خارجي

كيف يمكننا أن نعرفه عند مشاهدته؟

إيصال الإرسال من المؤسسة المالية عند أي من طرفي المعاملة

تقارير ذاتية عن عمليات الإرسال

التدفقات المالية الإجمالية بين الدول

ماذا عن الإرسال غير الرسمي للسلع والمنتجات؟

خطوات التعريف الإجرائي: بحث عن التحويلات المالية

ما هو التحويل؟

إرسال المال إلى مكان خارجي

ما هي الوحدة التي نلاحظ أو نقيس بها التحويلات؟

إيصال الإرسال من المؤسسة المالية عند أي من طرفي المعاملة : **معاملة**

تقارير ذاتية عن عمليات الإرسال : **فرد**

التدفقات المالية الإجمالية بين الدول : **بيانات على المستوى الوطني لفترة زمنية**
(الكمية في الشهر أو السنة)

خطوات التعريف الإجرائي: بحث عن التحويلات المالية

الفرق بين الوصف و التحليل أو التفسير

ما هو متوسط حجم التحويلات المالية في قطر؟

فرضية: برنامج تعليم الادخار سيزيد من متوسط حجم التحويلات المالية .

عملية التعريف الإجرائي

إقرار وحدات القياس ووحدات التحليل، بمعنى، تحديد كيفية قياس، أو ملاحظة، أو تشكيل المتغيرات

يجب قياس جميع المتغيرات لنفس وحدات التحليل، وخاصة عند تقييم فرضية .

إقرار أي تصميم بحث سيستخدم في جمع البيانات

ماذا يعني هذا في مفهوم البيانات؟ مصفوفة بيانات افتراضية

المتغير	الجنس	العمر	المعاملة	مبلغ التحويلات (دولار)
المستجيب	1 ذكر	25	1	880
المستجيب	2 ذكر	37	0	400
المستجيب	3 ذكر	30	0	285
المستجيب	4 ذكر	28	1	750
المستجيب	4 ذكر	40	0	1000

ماذا يعني هذا في مفهوم البيانات؟ مصفوفة بيانات افتراضية

بلد	مبلغ مرسل (مليون \$) إلى			
	باكستان	سريلانكا	مصر	التوظيف الهند
قطر	1,700	17,700	931.0	613.0
الكويت	485.9	2,900	2,200	823.4
عمان	283.9	231.3	132	2,600
البحرين	759.6	187.8	—	160.9

ما الناقص هنا؟

المخطط السببي للفرضية

نتائج على المدى الطويل

- زيادة المدخرات بين المهاجرين نتيجة زيادة المعرفة بالخيارات المالية

نتائج على المدى القصير

- يقرأ المستهدفين الصحف
- تعطى نشرات للمستهدفين
- يحضر المستهدفين برنامج محو الأمية المالية

التدخل / المدخل

- إعلانات في الصحف
- نشرات إعلامية
- برنامج محو الأمية المالية

مشكلة السياسة العامة

- سلوك الادخار بين المهاجرين منخفض

ما الأمور الأساسية التي يجب أن تُأخذ في عين الاعتبار عند التصميم؟

- هل يمكن إجراء قياسات أولية عن التحويلات المالية قبل بدء برنامج تعليم الادخار؟
- هل يمكن إنشاء دراسة البائل أو ملف بيانات لدراسات على المدى الطويل مع قياسات متكررة عبر الزمن؟
- هل يمكن إنشاء مجموعة أو مجموعات تحكم مناسبة؟
- كم عدد وحدات التحليل المختلفة التي يمكن استخدامها؟

:استطراد طرق جمع البيانات

مصادر الخطأ في القياس :

الأخطاء العشوائية بسبب تغيرات الفرص والمتوسط يؤول إلى الصفر

بشكل عام تساهم في عدم الدقة

الأخطاء النظامية ليست بسبب الفرصة ولها اتجاه أو "انحياز"

يمكن أن تثير مخاوف من ناحية ثبات وصدق البحث

التفكير في القياس

يمكننا التفكير لأي قياس عن طريق الملاحظة التي تتكون من نتيجة حقيقية، بالإضافة إلى بعض الخطأ

الملاحظة/القيمة المشاهدة = القيمة الحقيقية + خطأ

بما أن الخطأ يمكن أن يكون إما عشوائياً أو نظامياً :

$$\text{القيمة المشاهدة} = \text{القيمة الحقيقية} + \left[\text{خطأ عشوائي} + \text{خطأ نظامي} \right]$$

“تعليم طرق ادخار الأموال” مفهوم

فرضية: تعليم طرق ادخار الأموال تعمل على زيادة سلوك الادخار

كيف يمكن قياسها؟

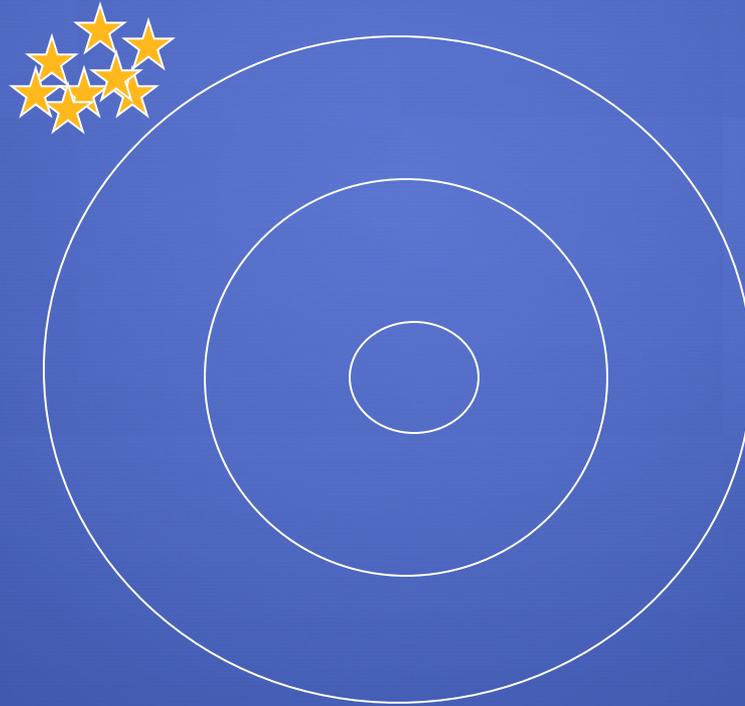
يجب الاتفاق على الوحدات. ما هي “وحدة” التعليم و “وحدة” سلوك الادخار؟

الثبات والصدق يُشيران إلى أخطاء القياس المُحتملة

الثبات يُشير إلى مدى اتساق أو دقة القياس

الصدق يُشير إلى ما إذا كنا نقيس ما نُفكر فيه (المفهوم)

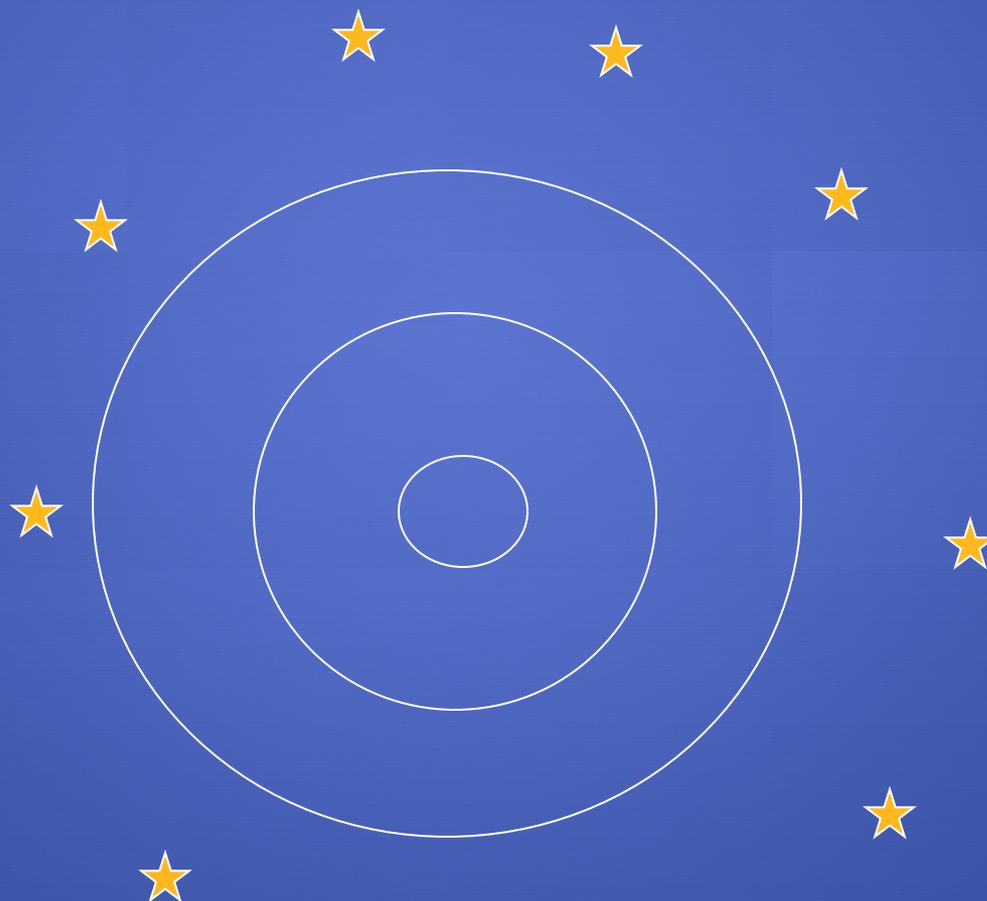
موثوق به، وغير صالح



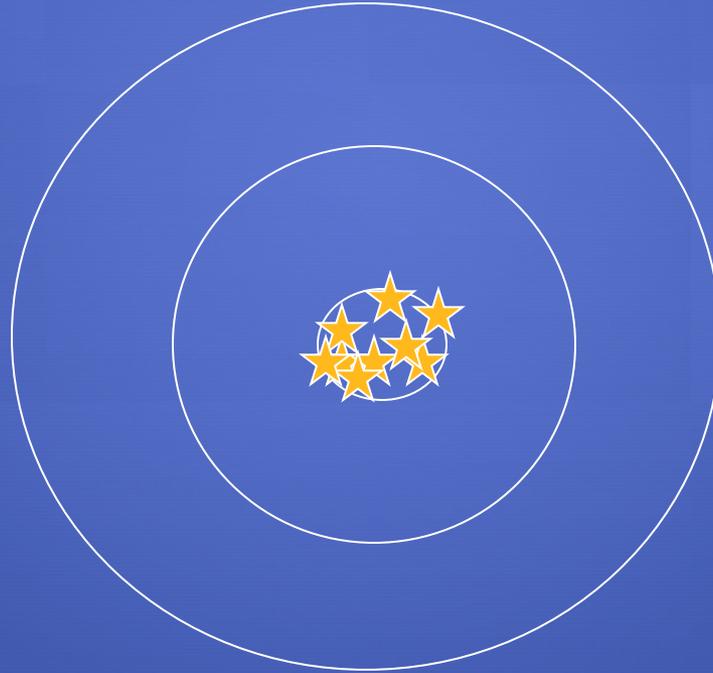
صالح، وغير موثوق به



غير صالح، و غير موثوق به



صالح وموثوق به



الثبت تهتم بالدقة والاتساق في القياس

يمكن أن تشمل على علاقة بين القياسات المتكررة لنفس المفهوم،
في شكل فرضية :

نظرية القياس

فرضية القياس:

القياس المتكرر لنفس المفهوم باستخدام نفس التفعيل يجب أن يعود لنفس القيمة، في المتوسط.

ماذا يعني هذا في مصطلحات البيانات؟

مصفوفة البيانات الفرضية

V5	V4	X _{t2}	X _{t1}	ID
7	4	نعم	نعم	Resp1
1	3	نعم	نعم	Resp2
2	1	لا	لا	Resp4
4	3	لا	نعم	Resp5

الافتراض القياسي للأخطاء العشوائية

الفرق الوحيد بين اثنين من القياسات لنفس المفهوم يجب أن يكون خطأ عشوائي

الأخطاء العشوائية :

1. يكون الاثنان موجب و سالب

2. المجموع يساوي صفر

(الانحياز) الافتراضات القياسية للأخطاء النظامية

1. تميل أن تكون إما موجب أو سالب

2. مجموعها لا يساوي صفر

طُرح عليك سؤال في أحد المسوح، وبعد أسبوع طُرح عليك نفس السؤال. إذا كان سؤال المسح وإجاباته صالحة وموثوق بها، فإن إجاباتك يجب أن تكون هي نفسها.

الارتباط

بما أن الاعتمادية تشمل على علاقة، يمكن وصفها بقياس إحصائي
للارتباط مثل معامل الارتباط .

الثبات ومعاملات الارتباط

معامل الارتباط هو قياس للتغير المشترك القياسي. ويتراوح بين 1- و

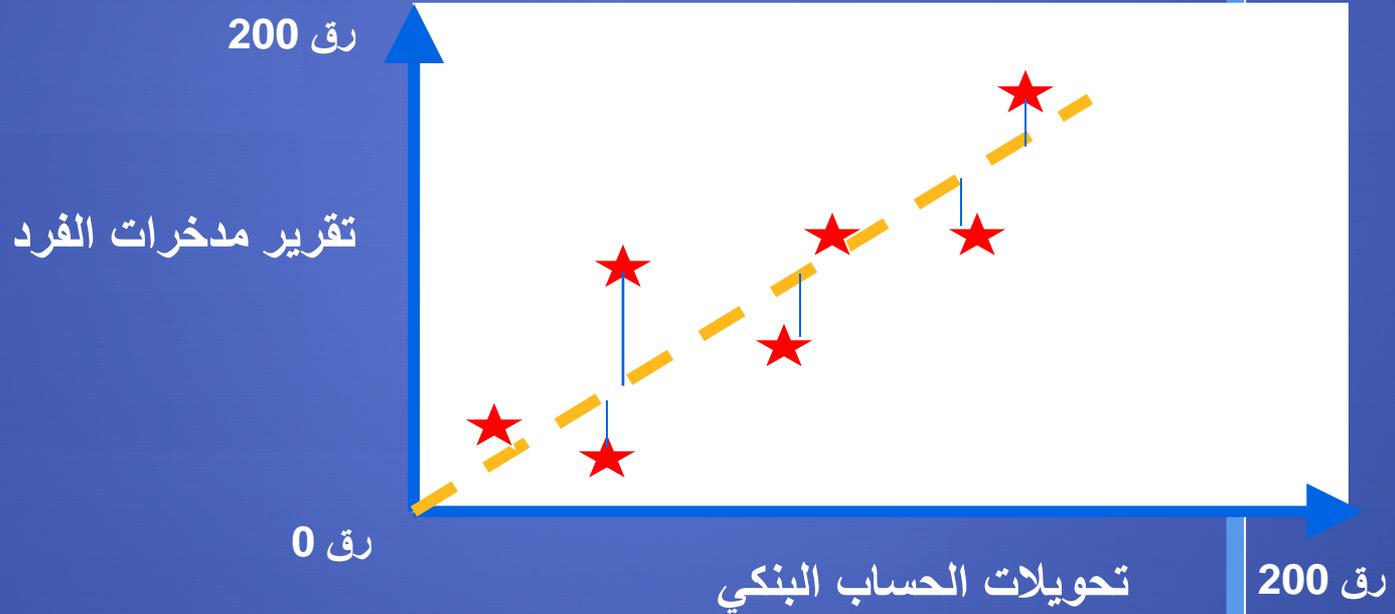
1

- القيم القريبة من 1 توحى بوجود ارتباط موجب قوي .
- القيم القريبة من -1 توحى بوجود ارتباط سالب قوي
- القيم القريبة من الصفر توحى بوجود ارتباط غير خطي
- من المفترض أن يكون الارتباط علاقة خطية، لذلك يقل إذا كانت العلاقة الحقيقية ليست خطية .

إستراتيجية القياس

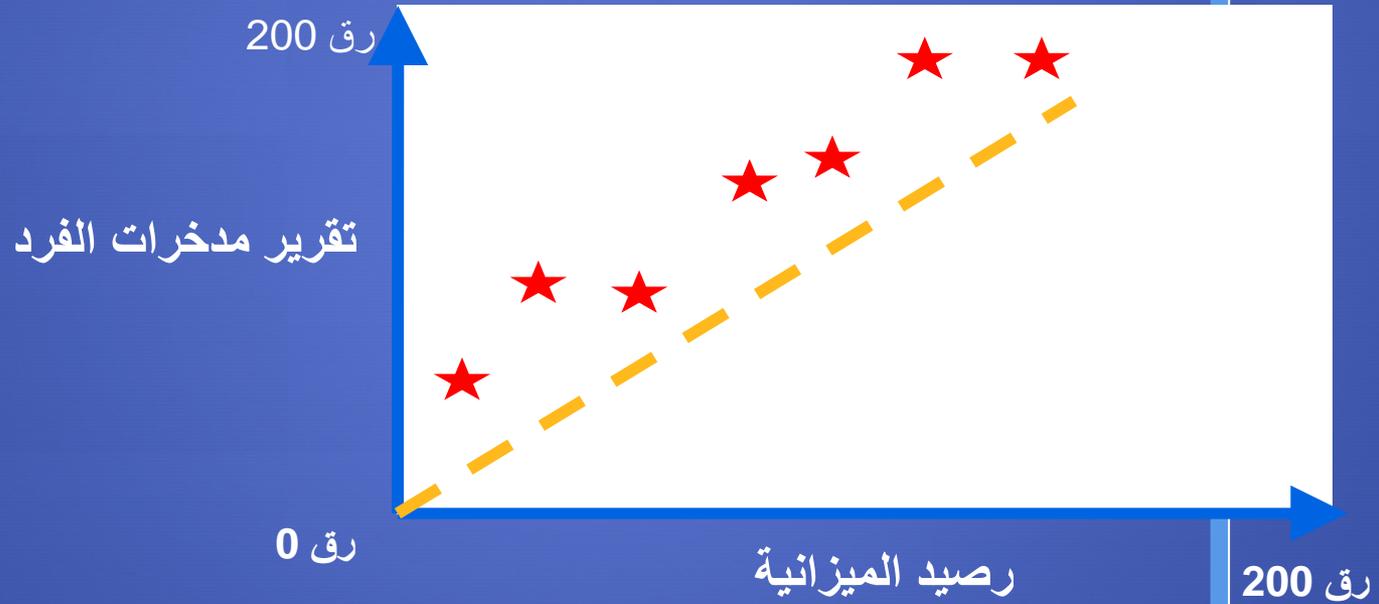
تسأل عامل مهاجر عن مقدار المبلغ من راتبه الشهري الذي يرسله إلى بلده وأنت تتنظر في السجلات الإدارية لرصيد حسابه الحالي .

القياس المتكرر مع خطأ عشوائي



ما هو الارتباط الذي يلخص ما هو احتمال أن تكون هذه العلاقة؟

(انحياز) القياس المتكرر مع خطأ عشوائي



ما هي التفسيرات المحتملة لهذه المشاهدة؟

تمرين المجموعة

- ناقش في مجموعات الطرق البديلة لتفعيل مفهوم "برنامج محو الأمية المالية" والنتائج المتضمنة لوحدات التحليل المناسبة .
- ما هي النتائج الضمنية لهذا الاختيار لقياس متوسط التحويلات المالية؟