عقد التوأمة

JO 21 ENI ST22 01

تعزيز قدرات دائرة الإحصاءات العامة في الأردن من حيث تجميع وتحليل وإعداد التقارير الخاصة بالبيانات الإحصائية بما يتماشى مع أفضل الممارسات الدولية والأوروبية

تقرير المهمة

عن

المعيار 1
إعداد خارطة طريق لتطوير نظام بيانات إدارية متكامل في الأردن
من خلال تنفيذ مشاريع تجريبية باستخدام سجلات إحصاءات الأعمال (SBR) والإحصاءات السكانية

النشاط 1.3.7:
ربط البيانات الإدارية وبيانات المسوح

تم تنفيذ المهمة بواسطة:
دكتور جان-فيليب كولب
السيد رينيه كريمر

عمان، الأردن
06-09 يناير 2025

تعزيز قدرات دائرة الإحصاءات العامة الأردنية

أسماء المؤلفين، العناوين، البريد الإلكتروني

دكتور *جان-فيليب كولب*
عالم بيانات
المكتب الفيدرالي للإحصاء في ألمانيا
11 Gustav-Stresemann-Ring
Wiesbaden 65819
ألمانيا
البريد الإلكتروني: Jan-Philipp.Kolb@destatis.de

السيد رينيه كريمر
عالم بيانات
المكتب الفيدرالي للإحصاء في ألمانيا
11 Gustav-Stresemann-Ring
Wiesbaden 65819
ألمانيا
البريد الإلكتروني: Rene.Kremer@destatis.de

جدول المحتويات

1. الملاحظات العامة ..... **Fejl! Bogmærke er ikke defineret.**
2. التقييم والنتائج ..... **Fejl! Bogmærke er ikke defineret.**
3. الاستنتاجات والتوصيات ..... **Fejl! Bogmærke er ikke defineret.**
الملحق 1. الإطار المرجعي ..... **Fejl! Bogmærke er ikke defineret.**
قائمة الاختصارات ..... **Fejl! Bogmærke er ikke defineret.**
0. الهدف والنتائج الإلزامية للمعيار ..... **Fejl! Bogmærke er ikke defineret.**
الهدف ..... **Fejl! Bogmærke er ikke defineret.**
النتائج الإلزامية ومؤشرات الإنجاز لكل عنصر فرعي ..... **Fejl! Bogmærke er ikke defineret.**
1. هدف النشاط ..... **Fejl! Bogmærke er ikke defineret.**
2. المخرجات المتوقعة للنشاط ..... **Fejl! Bogmærke er ikke defineret.**
3. المشاركون ..... **Fejl! Bogmærke er ikke defineret.**
خبراء الفترة القصيرة ..... **Fejl! Bogmærke er ikke defineret.**
خبراء دائرة الإحصاءات العامة ..... **Fejl! Bogmærke er ikke defineret.**
فريق مشروع التوأمة ..... **Fejl! Bogmærke er ikke defineret.**
4. الموارد ..... **Fejl! Bogmærke er ikke defineret.**
5. الأجندة الإجمالية ..... **Fejl! Bogmærke er ikke defineret.**
6. المعلومات الأساسية ..... **Fejl! Bogmærke er ikke defineret.**
الملحق 2. برنامج المهمة ..... **Fejl! Bogmærke er ikke defineret.**
الملحق 3. الأشخاص الذين تم الالتقاء بهم ..... **Fejl! Bogmærke er ikke defineret.**

قائمة الاختصارات

- BC – Beneficiary Country
الدولة المستفيدة
- DoS – Department of Statistics
دائرة الإحصاءات العامة
- MS – Member State
الدولة العضو
- PL – Project Leader
رئيس المشروع
- RTA – Resident Twinning Advisor
مستشار التوأمة المقيم
- STE – Short-term Expert
خبير الفترة القصيرة

1. الملاحظات العامة

تم إعداد تقرير هذه المهمة ضمن مشروع التوأمة المعنون "تعزيز قدرات دائرة الإحصاءات العامة في الأردن من حيث تجميع وتحليل وإعداد التقارير الخاصة بالبيانات الإحصائية بما يتماشى مع أفضل الممارسات الدولية والأوروبية". تتعلق هذه المهمة بالنتائج الإلزامية (MR) والمؤشرات التالية:

النتيجة الإلزامية 1.3: القيام بمشروع تجريبي حول كيفية استخدام السجلات الإدارية لتعزيز إحصاءات السكان وصياغة استمارة التعداد السكاني 2025

- المؤشر 1.3.أ: إعداد وتقييم قائمة (مخزون) لمصادر البيانات ووضع خطة عمل لإدماجها في إحصائيات دائرة الإحصاءات العامة
- المؤشر 1.3.ب: وضع منهجية لدمج البيانات الإدارية
- المؤشر 1.3.ج: إعداد الوثائق حول المعايير الإحصائية والتصنيفات والمعرفات والمخططات وما إلى ذلك.
- المؤشر 1.3.د: مراجعة لكيفية مساعدة البيانات الإدارية في صياغة استمارات التعداد لعام 2025

كان الهدف من هذا النشاط هو اتخاذ الخطوة الأولى في ربط بيانات التعداد ومصادر البيانات الإدارية.

خلال المهمة، تم تناول المواضيع التالية على المستوى النظري والمستوى العملي مع الاستناد إلى البيانات المحلية الأردنية.

- التحديات والحلول العملية المحتملة لاستخدام البيانات الإدارية في تعداد السكان والمساكن المركب في الأردن؛
- ربط البيانات بمعرفات مشتركة؛
- ربط البيانات بدون معرفات مشتركة؛
  - الطريقة الحتمية
  - مطابقات الاحتمالات للبيانات بدون معرفات مشتركة
- كيفية التعامل مع الاختلافات في المفاهيم والتعريفات والتوقيت
- تقييم الجودة

يود الخبراء أن يعبروا عن شكرهم الجزيل لجميع المسؤولين والأفراد الذين التقوا بهم على الدعم الكريم والمعلومات القيمة التي تلقوها خلال جلسات العمل التي سهلت عملهم بشكل كبير. وجهات النظر والملاحظات المذكورة في هذا التقرير هي آراء المستشارين ولا تتوافق بالضرورة مع آراء المكتب الإحصائي الفيدرالي الألماني.

2. التقييم والنتائج

يعمل الخبراء حاليًا على مشروع استكشافي (اختبار طريقة) للتحضير للجولة القادمة من التعداد 2031 في ألمانيا. تُستخدم طريقة ربط السجلات الاحتمالية المعتمدة على SPLINK لربط الإدخالات من سجل السكان بالسجلات الإدارية. SPLINK هو حزمة بايثون لربط السجلات الاحتمالية (حل الكيانات) التي تتيح لك إزالة التكرار وربط السجلات من مجموعات البيانات بدون معرفات فريدة. الهدف من هذه النشاط هو نقل تجارب الخبراء من اختبار الطريقة إلى حالة الاستخدام الأردنية.

في حالة الاستخدام الأردني، يجب ربط المصادر التالية ببيانات التعداد في العقبة:

- سجل مدني
- سجل التعليم
- سجل التأمين الصحي

الرقم الوطني يستخدم لربط مصادر البيانات. طرق ربط السجلات تستخدم فقط إذا لم يكن الرقم الوطني متاحًا في تعداد العقبة. هذا يتعلق بحوالي 30% من الإدخالات في تعداد العقبة.

بعض المتغيرات أساسية في هذه الحالة لاستخدامها لربط مجموعات البيانات بشكل فعال. أولاً، يجب تضمين ما يصل إلى ثلاثة أجزاء من اسم الشخص (الاسم الأول، اسم الأب واسم الجد)، جنبًا إلى جنب مع اسم العائلة، لأخذ الاختلافات في التهجئة أو النقل أو الترتيب في الاعتبار.

تاريخ الميلاد يعمل كمعرف رئيسي، مما يساعد على تمييز الأفراد ذوي الأسماء المتشابهة. علاوة على ذلك، يلعب الجنس دورًا مهمًا من خلال تضييق نطاق المطابقات المحتملة وزيادة دقة عملية الربط. أخيرًا، توفر الجنسية طبقة إضافية من التحقق، وهي مفيدة بشكل خاص في مجموعات البيانات متعددة الثقافات حيث قد تتداخل الأسماء عبر المناطق.

يجب تحويل السلاسل الفارغة بشكل منهجي إلى قيم null للحفاظ على التناسق. تضمنت البرمجيات الميدانية المستخدمة في تعداد العقبة قواعد التحقق من الصحة، مما يضمن عدم وجود عناصر نائبة أو تواريخ ميلاد غير صالحة مسجلة. ومع ذلك، فإن يوم الميلاد يكون مفقود أحيانًا في البيانات.

الأرقام التعريفية غير متاحة للأجانب، حيث يتم تسجيل هذه المعلومات حصريًا في بيانات إدارة الإقامة والحدود. تتحول الحروف اللاتينية إلى العربية لضمان التناسق اللغوي. تُستخدم جداول البحث أيضًا لتوحيد والتحقق من الجنسيات. يعمل البرنامج الميداني على تبسيط إدخال البيانات باستخدام drop-down menu لاختيار الخيارات المحددة مسبقًا.

يوصي الخبراء باستخدام جداول البحث كأداة قوية لتوحيد البيانات، خاصة عند التعامل مع المتغيرات الفئوية مثل الجنسية أو المدخلات غير الموحدة. تقوم هذه الجداول بربط القيم في البيانات بمجموعة محددة مسبقًا من القيم القياسية، مما يضمن التناسق والاتساق عبر مجموعة البيانات.

إنه خيار لتقييد البيانات على كلا الجانبين لرب الأسرة وإجراء المطابقة فقط على رب الأسرة. الميزة في هذا النهج هي أنه يؤدي إلى مجموعة بيانات أصغر، مما يحسن الأداء. يمكن بعد ذلك استخدام حجم الأسرة كمتغير رئيسي إضافي. قد تكون المشكلة أن هذه المعلومات متاحة فقط للأفراد الأردنيين. ومع ذلك، نتوقع أن تظهر التحديات الرئيسية مع الأفراد الذين ليسوا أردنيين.

تاريخ الميلاد الكامل ليس مثالياً كمتغير blocking لأن اليوم قد يكون مفقوداً. بدلاً من ذلك، يمكن استخدام سنة الميلاد والشهر. بالإضافة إلى ذلك، يمكن استخدام الاسم الأول بشكل فعال لل blocking. من المستحسن استبعاد الاسمين الثاني والثالث من هذه العملية. لا يمكن الاعتماد على أجزاء من الأسماء العربية، حيث إن الحروف في الكتابة العربية متصلة، ويمكن أن يتغير شكل الحرف الأول اعتمادًا على الحرف التالي.

عند مقارنة الأسماء العربية، يمكن استخدام مقاييس المسافة المختلفة اعتمادًا على طبيعة البيانات والدقة المطلوبة. تقدم الأسماء العربية تحديات فريدة، مثل التشكيل، وتهجئات مختلفة، والبادئات مثل "ال-". تقوم مسافة التحرير (Levenshtein Distance) بقياس الحد الأدنى من التعديلات (الإدراجات، الحذف، الاستبدالات) اللازمة لتحويل سلسلة نصية إلى أخرى، وهي مثالية للاختلافات الطفيفة في التهجئة. Jaro-Winkler Distance هي نوع من مقياس Jaro distance الذي يعطي الأولوية ل prefix matches وهو مفيد للأسماء ذات prefixes أو الأحرف الأولى الشائعة. عند اختيار مقياس المسافة، فإن Levenshtein أو Jaro-Winkler هما الخياران المثاليان للاختلافات البسيطة في التهجئة.

في الربط السجلي الاحتمالي، تعتبر العتبات لتصنيف أزواج السجلات إلى "عدم تطابق" و"تطابق محتمل" و"تطابق" حاسمة للتحكم في دقة وكفاءة عملية المطابقة.

يجب تقييم نتائج عملية الربط بناءً على عدة عوامل رئيسية، لا سيما معدل المطابقة وعدد المطابقات المتعددة. تشير نسبة المطابقة إلى نسبة السجلات التي ترتبط بنجاح عبر مجموعات البيانات، وتعتبر هذه المقياس مؤشراً رئيسياً على

مدى جودة أداء عملية الربط في تحديد السجلات المقابلة. معدل المطابقة الأعلى عادة ما يشير إلى أن العملية فعالة في العثور على المطابقات الدقيقة، ولكن من المهم تقييم جودة هذه المطابقات بعناية أيضًا. عدد المطابقات المتعددة، من ناحية أخرى، يعكس مدى تطابق سجل واحد من مجموعة بيانات واحدة مع سجلات متعددة في مجموعة بيانات أخرى. يمكن أن تكون هذه علامة على الغموض أو القضايا المحتملة في منطق الربط، حيث قد لا يكون لسجل معين تطابق واضح واحد إلى واحد. يمكن أن يساعد تحليل عدد هذه المطابقات المتعددة في تحديد المجالات في العملية التي قد تحتاج إلى تحسين أو مزيد من التحقيق. يجب النظر بعناية في كل من معدل المطابقة وعدد المطابقات المتعددة معًا لفهم الفعالية العامة لعملية الربط وإجراء التعديلات المستنيرة عند الضرورة.

من المفيد أيضًا فحص معدل المطابقة للمجموعات الفرعية داخل مجموعة البيانات لتحديد ما إذا كانت عملية ربط السجلات بحاجة إلى تحسين في مجالات معينة. من خلال تقييم معدلات المطابقة لمجموعات فرعية مختلفة، يمكن للمرء تحديد التباينات أو المشكلات المحتملة التي قد لا تكون واضحة عند النظر في مجموعة البيانات الكلية. قد تكون لبعض المجموعات الفرعية معدلات تطابق أقل بسبب عوامل مختلفة مثل جودة البيانات، ووجود معلومات مفقودة أو غير مكتملة، أو التعقيدات الكامنة في خصائص المجموعة الفرعية نفسها. على سبيل المثال، قد تظهر المجموعات الفرعية التي تحتوي على بيانات أقل معيارية أو ذات تباين أعلى في بعض السمات الرئيسية معدلات مطابقة أقل، مما يشير إلى أنه يمكن تحسين خوارزمية الربط للتعامل مع هذه الحالات بشكل أكثر فعالية. علاوة على ذلك، فإن مراجعة المجموعات الفرعية بشكل فردي يسمح بتقييم أكثر دقة، مما يبرز المجالات التي قد تفشل فيها عملية الربط في تحقيق تطابقات دقيقة، وبالتالي يوفر رؤى قيمة للتحسينات المستهدفة.

3. الاستنتاجات والتوصيات

تعتبر تجربة عملية في العمل مع البيانات أمراً حاسماً. كان إلقاء نظرة على البيانات وكتابة بعض أكواد بايثون معًا نشاطًا مفيدًا جدًا خلال هذه المهمة. تم تقديم توصيات حول كيفية إجراء ربط سجلات أساسي وما هو المهم عند كتابة الكود. كما أنها موثقة في الوظائف وفي ملف README للنموذج الأولي. لقد كانت التطبيقات على البيانات الاصطناعية مفيدة للغاية. يمكن تطبيق النموذج الأولي على الفور على حالة استخدام خاصة.

من المنطقي أولاً إجراء مطابقة دقيقة باستخدام المعرف وتقليل مجموعة البيانات عن طريق إزالة الإدخالات التي يمكن مطابقتها بدقة. في الخطوة التالية، يجب إجراء ربط سجلات احتمالي على الإدخالات المتبقية.

قدم خبراء التوأمة إلى الدائرة نموذجًا أوليًا يمكن تطبيقه على البيانات الحقيقية. يمكن بعد ذلك تطوير هذه السلسلة بشكل أكبر. من المهم اختبار إعدادات المعلمات المختلفة ومقارنة النتائج.

استنتج أنه يجب أخذ النقاط التالية بعين الاعتبار لإنشاء رابط سجلات فعال للتعداد الأردني:

- التوحيد والتطبيع
- اختيار متغيرات blocking (مثل سنة الميلاد كمتغير blocking)
- اختيار المتغيرات الرئيسية (استخدم أكبر قدر ممكن من المعلومات)
- اختيار مقاييس المسافة مثل: Exact, Jaro-Winkler أو Levenshtein
- اختيار عدد مستويات المقارنة لقياس المسافة
- عتبة المطابقة/عدم المطابقة

يوصي الخبراء بالاحتفاظ بتاريخ الميلاد كمتغير رئيسية. هذا مهم بشكل خاص للتفريق بين الأفراد الذين يحملون أسماء شائعة جدًا. على عكس حالة الاستخدام الألمانية، على سبيل المثال، فإن العنوان ومكان الميلاد غير متاحين للتفريق بين الأشخاص الذين يحملون أسماء عائلية شائعة.

قضية تغيير الأسماء بعد الزواج لا تنطبق على الحالة الأردنية. بدلاً من ذلك، يجب إعطاء مزيد من الاعتبار لكيفية التعامل مع العشائر.

قام الخبراء بتقديم خط إنتاج لربط السجلات يمكن تعديله واستخدامه لمعالجة البيانات الحقيقية الأردنية. يجب أخذ النقاط التالية في الاعتبار عند العمل مع هذا الخط.

عند العمل مع ملفات الإدخال الجديدة، مثل البيانات الحقيقية، من الضروري تكييف العملية مع الخصائص المحددة لهذه المجموعات الجديدة من البيانات. على سبيل المثال، عند العمل مع بيانات حقيقية، من المهم التأكد من أن أسماء الأعمدة لا تحتوي على أي مسافات. هذا لأن الدوال مثل "parse_blocking" قد تفشل في التنفيذ بشكل صحيح إذا كانت هناك مسافات في أسماء الأعمدة. يمكن أن تتسبب المسافات البيضاء في حدوث أخطاء في التحليل والمعالجة، لذلك يُوصى بمراجعة أسماء الأعمدة بعناية قبل تطبيق خط إنتاج ربط السجلات.

إذا وجدت أن بعض أسماء الأعمدة تحتوي على مسافات بيضاء، يجب تصحيحها قبل المتابعة مع خط الإنتاج. يتضمن ذلك إزالة أو استبدال أحرف المسافة لضمان سلاسة وظيفة عملية الربط. إجراء هذه التعديلات مبكرًا سيساعد في تجنب أي انقطاعات في سير العمل وضمان تشغيل خط الإنتاج بدون أخطاء.

بالإضافة إلى ذلك، عند العمل مع real_data، يمكن استخدام دالة محددة تسمى 'load_source_and_delivery_data_in_duckdb_realdata'. تؤدي هذه الوظيفة دورًا مهمًا من خلال اختيار جميع السمات ذات الصلة من مجموعة البيانات وإضافة عنصرين حاسمين: source_dataset و serial unique_id. تساعد مجموعة البيانات المصدر في تتبع أصل كل سجل، وهو أمر مهم للحفاظ على سلامة البيانات طوال العملية. المعرف الفريد التسلسلي ضروري لـ Splink، حيث يحدد كل سجل في مجموعة البيانات بشكل فريد، مما يضمن أن المطابقات والروابط يتم تعيينها بشكل صحيح خلال عملية الربط.

من المهم التأكد من أن البيانات مُنسقة بشكل صحيح ومعالجة، ثم يمكنك تطبيق خط إنتاج ربط السجلات بفعالية على مجموعات البيانات الجديدة، مثل البيانات الحقيقية، دون مواجهة مشكلات ناتجة عن أخطاء التنسيق أو المعرّفات المفقودة.

عند النظر في قواعد blocking لتوقع ربط السجلات، من المهم تقييم مدى فعالية هذه القواعد في الممارسة العملية. طريقة واحدة لقياس ما إذا كان لديك قاعدة blocking جيدة هي تقييم أدائها (على سبيل المثال، مقارنة أوقات التشغيل مع قواعد blocking المختلفة). بدلاً من الاعتماد على قاعدة blocking واحدة، غالبًا ما يكون من الجيد استخدام قواعد متعددة، خاصةً للتعامل مع الحالات التي تحتوي على قيم مفقودة. إذا كان السجل يفتقر إلى معلومات في سمة واحدة، فإن وجود عدة قواعد تعتمد على سمات مختلفة يمكن أن يساعد في ضمان تحديد المطابقات الصحيحة. تزيد هذه الطريقة من قوة عملية المطابقة وتقلل من خطر الفشل في ربط السجلات المهمة.

يمكن أن يكون من المفيد دمج سمات متعددة (مثل 2-3 سمات) في قاعدة blocking واحدة. على سبيل المثال، يمكن أن تجمع قاعدة blocking بين الصوت العربي للاسم الأول، وسنة الميلاد، والصوت العربي للاسم العائلي. يساعد هذا المزيج في التقاط جوانب مختلفة من البيانات ويمكن أن يحسن دقة عملية blocking.

كقاعدة عامة، من المهم تحقيق التوازن عند إنشاء قواعد blocking الخاصة بك. إذا كانت القواعد فضفاضة جدًا، فقد تفشل وظيفة الربط بسبب وجود عدد زائد من المقارنات غير الضرورية، مما يؤدي إلى عدم الكفاءة. من ناحية أخرى، إذا كانت القواعد صارمة جدًا، فقد تفوت الروابط الصالحة، حيث قد يتم استبعاد بعض المطابقات المحتملة من الاعتبار. تأكد من أن الكتل متساوية تقريبًا في الحجم أو على الأقل ليست غير متجانسة. تتضمن Splink وظيفة مدمجة تسمح بتقييم حجم الكتل. إذا تلفت kernel أثناء التنفيذ، فقد يتجاوز عبء العمل السعة القصوى المتاحة من ذاكرة الوصول العشوائي. قاعدة blocking فضفاضة للغاية، يمكن أن تصبح blocking الناتجة كبيرة بشكل مفرط، مما يؤدي إلى مشاكل في الذاكرة. في مثل هذه الحالات، من المحتمل أن تكون فكرة جيدة تحسين قاعدة blocking من خلال جعلها أكثر صرامة أو تعديل السمات المعنية لتقليل حجم blocking.

في ملف readme للنموذج الأولي، وصف الخبراء تعديلات مختلفة ممكنة للتعاون مع الأداء. إذا استمرت المشكلة، فقد يكون من الضروري زيادة سعة ذاكرة الوصول العشوائي للنظام. في مثل هذه الحالات، يمكن أن يؤدي تحديث ذاكرة الكمبيوتر إلى حل المشكلة وتمكين عبء العمل من العمل بسلاسة دون تعطل kernel.

قدم الخبراء أيضًا إمكانية المراجعات المكتبية، والتي تُستخدم في اختبار الطريقة الألمانية. ومع ذلك، لا يُوصى بذلك في المثال الأردني، حيث إن هذا الإجراء يتطلب الكثير من العمل. بدلاً من ذلك، يجب استخدام معدلات المطابقة، وعدد المطابقات المتعددة، والأدوات الرسومية مثل waterfall charts المقدمة للتحقق مما إذا كانت النتائج معقولة.

الملحق 1. الإطار المرجعي

الإطارات المرجعية

مشروع التوأمة مع الاتحاد الأوروبي
JO 21 ENI ST22 01

المعيار 1:
خارطة طريق لتطوير نظام بيانات إدارية متكامل في الأردن
من خلال تنفيذ مشاريع تجريبية حول سجلات الأعمال الإحصائية (SBR) وإحصاءات السكان

النشاط 1.3.7:
ربط البيانات الإدارية وبيانات المسوح
التواريخ: 06 - 09 يناير 2025

جدول المحتويات

قائمة الاختصارات 2
0. الهدف والنتائج الإلزامية للمعيار 3
الهدف 3
النتائج الإلزامية ومؤشرات الإنجاز لكل عنصر فرعي 4
1. هدف النشاط 4
2. المخرجات المتوقعة من النشاط 5
3. المشاركون 5
خبراء الفترة القصيرة من الدول الأعضاء الأوروبية 5
خبراء دائرة الإحصاءات العامة 6
فريق التوأمة 6
4. الموارد 6
5. الأجندة الإجمالية 7
6. المعلومات الأساسية 8

- **قائمة الاختصارات**

BC	**Beneficiary Country** **الدولة المستفيدة**
DoS	**Department of Statistics** **دائرة الإحصاءات العامة**
ESS	**European Statistical System** **النظام الإحصائي الأوروبي**
MS	**Member State** **الدولة العضو**
RTA	**Resident Twinning Advisor** **مستشار التوأمة المقيم**
STE	**Short Term Expert** **خبير الفترة القصيرة**
ToR	**Term of References** **الإطارات المرجعية**

0. الهدف والنتائج الإلزامية للمعيار

الهدف

إعداد خارطة طريق لتطوير نظام بيانات إدارية متكامل للأردن، وإجراء مشاريع تجريبية لإنشاء سجل إحصائي للأعمال وتعزيز إحصاءات السكان.

وبما أن تطوير نظام بيانات إدارية متكامل بالكامل يعد مشروعًا طويل الأمد، سوف يركز مشروع التوأمة بشكل رئيسي على مشاريع تجريبية محددة حيث يمكن استخدام السجلات الإدارية لمعالجة التحديات الرئيسية التي تواجهها دائرة الإحصاءات العامة حاليًا. وستشكل هذه المشاريع التجريبية الخطوات الأولى في طرح خارطة طريق للنظام الإحصائي الأردني من خلال توفير نموذج لتوسيع استخدام البيانات الإدارية عبر النظام الإحصائي الأوسع بمرور الوقت. وعلى وجه التحديد، سوف تركز المشاريع التجريبية لمشروع التوأمة على تطوير سجل إحصائي للأعمال وتحسين جودة إحصاءات السكان.

بالإضافة إلى تحسين تقديرات السكان، يمكن للبيانات الإدارية أيضًا أن تساهم في تحسين نطاق استراتيجية التنمية السكانية لعام 2025. استمارة التعداد العام للسكان والمساكن، وبالتالي تحرير الموارد في دائرة الإحصاءات العامة.

سيبحث هذا العنصر الفرعي في كيفية قدرة السجلات الإدارية على توفير بيانات مصدرية جديدة لمراقبة تدفقات السكان وحركاتهم عبر المحافظات والبلديات بشكل أفضل. سيعمل مشروع تجريبي على تقييم كيفية دمج البيانات الإدارية (على سبيل المثال، من دائرة الأحوال المدنية وجوازات السفر) مع بيانات دائرة الإحصاءات العامة مثل التعداد السكاني لتعزيز إحصاءات السكان. وقد يرغب مشروع التوأمة في استكشاف مصادر بيانات أخرى غير البيانات الإدارية - على سبيل المثال، يجمع بنك تنمية المدن والقرى البيانات على مستوى المناطق الصغيرة حول تحركات السكان.

وقد تساعد البيانات الإدارية المتعلقة بخصائص السكان أيضًا في استبدال البيانات التي يتم جمعها حاليًا في التعدادات السكانية. سيقوم هذا العنصر الفرعي بتقييم كيفية مساعدة البيانات الإدارية في صياغة الاستمارة الخاصة بتعداد عام 2025، مع التركيز بشكل خاص على إمكانية تحرير الموارد في دائرة الإحصاءات العامة.

تم إطلاق رؤية التحديث الاقتصادي الأردني 2030 مؤخراً وتم تحديد "الأردن الذكي" كواحد من محركات النمو الثمانية لتنفيذ رؤية التحديث الاقتصادي. يتضمن برنامج "سائق الأردن الذكي" سبع قطاعات حيث تعد البيانات أحدها. ويشير هذا إلى المصلحة الوطنية في ضمان مصادر بيانات ثابتة وموثوقة، وأنظمة إحصائية قوية تساهم في صنع السياسات في الوقت المناسب وبشكل مدروس. ومن المتوقع أن يكون أحد الإجراءات التي سيتم اتخاذها هو تحويل دائرة الإحصاءات العامة الأردنية إلى مركز إحصائي وطني تفاعلي.

ينقسم المعيار الأول إلى خمسة عناصر فرعية، ولكل منها نتائج إلزامية (MR) واثنين إلى أربعة مؤشرات للإنجازات المرتبطة بالعنصر الفرعي.

- ***النتائج الإلزامية ومؤشرات الإنجاز لكل عنصر فرعي***

الجدول 1: النتائج ومؤشرات الإنجاز الإلزامية لكل عنصر فرعي ضمن المعيار 1: نظام بيانات إدارية متكامل للأردن

النتائج الإلزامية من وثيقة التوأمة	المؤشرات
النتيجة الإلزامية 1.1: إعداد قائمة (مخزون) للبيانات الإدارية الخاصة بالشركات والأسر وخارطة طريق إرشادية لإدراجها في نظام متكامل	**المؤشر 1.1.أ: إعداد قائمة لمتغيرات البيانات الإدارية والبيانات الوصفية الداعمة التفصيلية** **المؤشر 1.1.ب: إعداد خارطة طريق أولية لإدراج البيانات في النظام المتكامل**
النتيجة الإلزامية 1.2: مشروع تجريبي لتطوير استراتيجية لدمج مصادر البيانات الإدارية لأغراض إنشاء سجل إحصاءات أعمال	**المؤشر 1.2.أ: تحديد مصادر البيانات الإدارية وتقييمها ووضع خطة لدمجها مع معلومات تعداد المنشآت (CoE) في سجل إحصاءات الأعمال** **المؤشر 1.2.ب: إعداد الوثائق المتعلقة بهياكل قواعد البيانات والامتثال للمعايير الإحصائية والتصنيفات (على سبيل المثال ISIC، Rev 4) وما إلى ذلك واستخدام المعرفات المشتركة وما إلى ذلك.** **المؤشر 1.2.ج: دراسة كيف يمكن لسجل إحصاءات الأعمال أن يفيد المجالات الإحصائية الأخرى في دائرة الإحصاءات العامة**
النتيجة الإلزامية 1.3: تنفيذ مشروع تجريبي حول كيفية استخدام السجلات الإدارية لتعزيز إحصاءات السكان وصياغة استمارة التعداد السكاني لعام 2025	المؤشر 1.3.أ: إعداد جرد مصادر البيانات وتقييمها ووضع خطة عمل لدمجها في إحصاءات دائرة الإحصاءات العامة **المؤشر 1.3.ب: تطوير منهجية لدمج البيانات الإدارية** المؤشر 1.3.ج: إعداد الوثائق بشأن المعايير الإحصائية والتصنيفات والمعرفات والمخططات وما إلى ذلك. **المؤشر 1.3.د: مراجعة لكيفية مساعدة البيانات الإدارية في تطوير استمارات التعداد السكاني 2025**
النتيجة الإلزامية 1.4: تطوير استراتيجية لضمان تدفق البيانات بين دائرة الإحصاءات العامة والمؤسسات النظيرة على أساس مستمر للمشاريع التجريبية المذكورة أعلاه	المؤشر 1.4.أ: مراجعة البنية التحتية التقنية لنقل البيانات وإعداد خطة العمل على أساس 1.1 و1.2 أعلاه **المؤشر 1.4.ب: توقيع مذكرات تفاهم بين دائرة الإحصاءات العامة والمؤسسات الشريكة** المؤشر 1.4.ج: اتفاقية بشأن المعايير الإحصائية والتصنيفات والمعرفات وما إلى ذلك بين دائرة الإحصاءات العامة والمؤسسات الشريكة **المؤشر 1.4.د: مراجعة تدفقات البيانات داخل دائرة الإحصاءات العامة**
النتيجة الإلزامية 1.5: تنفيذ برامج التدريب وتطوير المواد التدريبية داخل دائرة الإحصاءات العامة ومع المؤسسات الشريكة بشأن استخدام السجلات الإدارية لأغراض إحصائية، استنادًا إلى المشاريع التجريبية المذكورة أعلاه	**المؤشر 1.5.أ: تطوير توثيق مفصل للمعايير الإحصائية والتصنيفات والمعرفات وما إلى ذلك.** المؤشر 1.5.ب: برامج تدريبية وورش عمل شاملة مقدمة لموظفي دائرة الإحصاءات العامة والمؤسسات الشريكة المؤشر 1.5.ج: تعزيز الدور القيادي لدائرة الإحصاءات العامة في ضمان تطبيق المعايير الإحصائية المناسبة في النظام الإحصائي الأردني.

- **1. هدف النشاط**

الهدف من هذا النشاط هو اتخاذ الخطوة الأولى في ربط بيانات التعداد ومصادر البيانات الإدارية: سيتم مناقشة المواضيع التالية وعرض أمثلة عملية للمنهجية واختبارها على البيانات الأردنية

- التحديات والحلول العملية الممكنة لاستخدام البيانات الإدارية لتعداد السكان والمساكن الأردني المركب
- **ربط البيانات بالمعرفات المشتركة؛**
- **ربط البيانات بدون معرفات مشتركة؛**
  - **الطريقة الحتمية**
  - **مطابقات الاحتمالات للبيانات التي لا تحتوي على معرفات مشتركة؛**
- **كيفية التعامل مع الاختلافات في المفاهيم والتعريفات والتوقيت**
- **تقييم الجودة**

- **2. المخرجات المتوقعة من النشاط**
- **تقرير النشاط**
- مناقشة الفهم المشترك للتحديات والحلول العملية المحتملة لاستخدام البيانات الإدارية لتعداد السكان والمساكن الأردني المركب
- **الخبرات العملية في مطابقة بيانات التعداد والبيانات الإدارية التي تم الحصول عليها؛**

- **3. المشاركون**

- ***خبراء الفترة القصيرة من الدول الأعضاء الأوروبية***

- **الدكتور يان - فيليب كولب، عالم بيانات، المكتب الإحصائي الاتحادي في ألمانيا (Destatis). يمتلك الدكتور كولب خبرات متخصصة في مجال ربط السجلات وكذلك تطوير المفاهيم ونمذجة العمليات لتحديد الأرقام السكانية على أساس التسجيل، فضلا عن إجراءات ضمان الجودة وتصور نقل البيانات. يحمل الدكتور كولب درجة الدكتوراه في طرق توليد مجموعات المحاكاة الاصطناعية. الشائع الجيد في تكنولوجيا morderns suc مثل R وPython وPyspark وSQL وNifi وما إلى ذلك. البريد الإلكتروني: Jan-Philipp.Kolb@destatis.de**

- **السيد رينيه كريمر، عالم بيانات، المكتب الإحصائي الاتحادي في ألمانيا (Destatis). السيد كريمر، ترأس اللجان الوطنية والدولية المعنية بمواضيع الربط المسجل، وتيسير التعاون عبر مختلف المؤسسات الحكومية والبحثية. يتمتع السيد كريمر بمعرفة متخصصة في بناء وتحسين خط ربط قياسي لتعداد**

السجل، وتحسين دقة البيانات وإدماجها في الإحصاءات السكانية على الصعيد الوطني. المعرفة المتقدمة في techniSpark وPython وNiFi (شهادة Cloudera) وKafka وSQL وGit. البريد الإلكتروني: Rene.Kremer@destatis.de

- ***خبراء دائرة الإحصاءات العامة***

مديرية إدارة البيانات

قسم البيانات الإدارية

- **السيد جعفر عبابنة مدير مديرية إدارة البيانات**
- **السيد صفوت ردايدة، رئيس قسم البيانات الإدارية**
- **السيد محمد العمري، إحصائي، قسم البيانات الإدارية**
- **السيدة لينا الجزازي، إحصائي، قسم البيانات الإدارية**
- **السيد عبد الواحد الحرايزة، قسم سجل الأعمال**
- السيدة صفاء أو عيطة، قسم سجل الأعمال
- الآنسة أنوار الخصبة، قسم سجل الأعمال

مديرية المسوح الأسرية والسكانية

- **السيدة منار الجخ، مديرية المسوح الأسرية والسكانية**
- **السيدة نور نعيمات، مديرية المسوح الأسرية والسكانية**
- **السيدة نهى طويقات، مديرية المسوح الأسرية والسكانية**

مديرية المنهجيات وقسم البيانات

قسم ضمان الجودة

- **السيدة رقية الصنابرة، قسم الجودة**

مديرية التحول الالكتروني وتكنولوجيا المعلومات

- **السيدة أحلام الروسان، مدير مديرية تكنولوجيا المعلومات والتحول الإلكتروني**
- **السيد محمد الشطناوي، خبير (Python)**
- **السيد حسام أو الشكر، مديرية IT**
- **السيد أيمن الهلول، خبير (Python)**

- ***فريق التوأمة***

- **المهندس تامر الروسان، رئيس قسم إحصاءات النباتات في الأردن (نظير مستشار التوأمة المقيم)**
- **دكتور. شارلوت نيلسن (مستشار التوأمة المقيم)**
- **الآنسة زينة عميرة (المساعد اللغوي لمستشار التوأمة المقيم)**
- **الآنسة ذكرى الطورة (مساعد مستشار التوأمة المقيم)**

- **4. الموارد**

سيتم توفير الترجمة الشفوية والترجمة الكتابية طوال النشاط. سيتم توفير الترجمة كترجمة تتابعية.

سيكون مكان الاجتماع في قاعة الاجتماعات بدائرة الإحصاءات العامة ستكون المواد المكتبية متاحة.

- **5. الأجندة الإجمالية**

الاثنين 06 كانون الثاني 2025

(جميع المشاركين المذكورين في الإطار المرجعي)

- الجهة المستفيدة: الترحيب - عرض الحالة الراهنة لربط البيانات - الإنجازات والتحديات الراهنة وعرضها - جميع المشاركين (1.5 ساعة)
- **الدولة العضو: مقدمة لربط البيانات بدون معرفات مشتركة (ساعتين) - جميع المشاركين؛**
  - **الطريقة الحتمية**
  - **مطابقات الاحتمالات للبيانات بدون معرفات مشتركة**
- الجهة المستفيدة والدولة العضو: التحضير للعمل العملي في الأيام القادمة (1.5 ساعة) - جميع المشاركين؛
  - **استعراض مشترك للبيانات والبيانات الوصفية لاستخدامها في الأيام المقبلة**
  - **المراجعة والتحقق من نشر جميع أدوات التحليل في الخادم**

الثلاثاء 07 كانون الثاني 2025 والأربعاء 08 كانون الثاني 2025

(المشاركون - موظفو دائرة الإحصاءات العامة الذين لديهم خبرات عملية في ربط البيانات و/أو Python)

- الجهة المستفيدة والدولة العضو: العمل العملي على ربط البيانات بدون معرف مشترك. سيبدأ العمل فيما يلي:
  - **عينة من 25,000 فرد تعدها دائرة الإحصاءات العامة مسبقًا**
  - **البيانات التي تربط بين العمل المنجز بالفعل في دائرة الإحصاءات العامة في Python (ستقدم الدائرة وصفاً منهجياً لخبراء الفترة القصيرة قبل المهمة)**
  - **"Library "Splink التي تستند إلى كود ب Python/Spark**
    - ***دليل المستخدم:***

https://moj-analytical-services.github.io/splink/topic_guides/topic_guides_index.html

- **الشرح:**

https://moj-analytical-services.github.io/splink/demos/tutorials/00_Tutorial_Introduction.html

الخميس 09 كانون الثاني 2025

- الجهة المستفيدة والدولة العضو: صياغة منهجية ومبادئ توجيهية عملية لربط البيانات استناداً إلى الدروس المستفادة في الأيام السابقة - المشاركون من الأيام السابقة (2.5 ساعة)

- ***إذا سمح الوقت - الجهة المستفيدة والدولة العضو*: تحديد المفاهيم ذات الاختلافات في الأردن وتبادل الأفكار حول كيفية التعامل مع هذا (جميع المشاركين من الإطار المرجعي)**
- الجهة المستفيدة والدولة العضو: الاستنتاجات والتوصيات - جميع المشاركين من الإطارات المرجعية وكذلك الإدارة العليا (نصف ساعة)

الشروط المسبقة

- **العمل على عينة من الأفراد 25,000 من جميع مصادر البيانات**
- **البيانات المخزنة على حاسوب خدمة والوصول عن بعد المقدمة إلى المشاركين في دائرة الإحصاءات العامة**
- **البيانات الوصفية لجميع مصادر البيانات المقدمة**
- **تحميل Python (Python 3.8 وPython module Splink)**
- **استيفاء متطلبات وحدة المعالجة المركزية وذاكرة الوصول العشوائي (CPU/RAM)**

- **6. المعلومات الأساسية**

التعداد العام للسكان والمساكن في الأردن لعام 2025

اعتمادًا على النضج، يمكن اتباع طرق مختلفة لإجراء تعداد مركب.

بفضل الجهود المشتركة التي بذلها خبراء المشروع وخبراء دائرة الإحصاءات العامة، تم إجراء تحليل لمختلف الأساليب المتبعة للانتقال من التعداد السكاني والمساكن التقليدي إلى التعداد الكامل القائم على السجلات من حيث المخاطر والتكاليف والجودة والاستثمارات الزمنية. علاوة على ذلك، تم تصنيف أشكال التعداد السكاني المركب المختلفة حسب المستوى الذي وصلت إليه حتى الآن في الأردن. ينبغي فهم الدرجات العددية على أنها اتجاهات وليست رتبًا مطلقة. وكانت النتيجة أنه في عام 2025 سوف تستخدم دائرة الإحصاءات العامة الجولة التالية من التعداد للتحقق من جودة البيانات الإدارية، ولكنها ستستخدم أيضًا في بعض الحالات البيانات الإدارية للتعبئة المسبقة (الشكل 1).

نوع التعداد	مدى استخدام السجلات	استخدام السجلات	المزايا	المخاطر [1 ;0]	التكاليف [1 ;0]	الجودة [1 ;0]	الوقت المستغرق [1 ;0]	المستوى الذي تم الوصول إليه الان
التعداد التقليدي		لا يتم استخدام السجلات		0	1	0,9	0	
التعداد المركب		للتحقق من الجودة/ التحسين		0	1	1	0,1	
		التعبئة المسبقة		0	0,94	1	0,1	
		لتقدير بعض المتغيرات بدلاً من السؤال		0,2	0,9	1	0,3	
		لسؤال عينة فقط من السكان		0,4	0,25	0,8	1	
التعداد القائم على السجلات		فقط استخدام السجلات		يعتمد على جودة 1-0 السجلات والوقت المستغرق للتحضير	0,2	0,8-1	0,8	

الشكل 1: إمكانية الانتقال من التعداد التقليدي إلى التعداد الكامل المبني على السجل وتقدير المخاطر والتكاليف والجودة والاستثمارات الزمنية.

يمكن العثور على إرشادات مفصلة وأمثلة من بلدان مختلفة في إرشادات لجنة الأمم المتحدة الاقتصادية لأوروبا بشأن استخدام السجلات والبيانات الإدارية لتعدادات السكان والمساكن:
https://unece.org/DAM/stats/publications/2018/ECECESSTAT20184.pdf

ربط البيانات

لقد أصبح ربط بيانات المسوحات وبيانات التعداد بالبيانات الإدارية أمرا مهما بشكل متزايد بالنسبة للمؤسسات الإحصائية الوطنية التي تتجه من المسوحات والتعدادات التقليدية إلى استخدام المصادر الإدارية لإنتاج الإحصاءات الرسمية.

ومع ذلك، من أجل الجمع بين البيانات التي تم جمعها ميدانيا مع البيانات الإدارية، فإن الربط بين مصادر البيانات المختلفة مع أو بدون مفاتيح فريدة أمر ضروري. إذا كان هناك معرف فريد في جميع السجلات أو معظمها، يصبح الارتباط سهلاً نسبيًا ويكون مستوى الارتباط الناجح مرتفعًا عادةً (على الرغم من أنه لا يزال يتعين قياس جودة المتغيرات في السجل الإحصائي).

في الأردن، يتم إصدار رقم تعريفي فريد من نوعه بواسطة xxxx لجميع المواطنين الأردنيين والمقيمين الدائمين والمقيمين المؤقتين (العاملين) - جميع الأشخاص. ومع ذلك، في الأردن، لم يتم جمع هذه المعرفات الفريدة حتى الآن في المسوحات والإحصاءات، مما يجعل ربط هذه البيانات بالبيانات الإدارية أمرًا صعبًا. خلال التعداد السكاني الذي أجري في محافظة العقبة في عام 2024، تم توضيح رغبة/قدرة السكان على تزويد الباحثين برقمهم الوطني. وأظهرت النتيجة أن حوالي 30% فقط من السكان كانوا على استعداد لتزويد الباحث برقم الوطني الخاص بهم.
عندما تكون المعرفات الفريدة غائبة ولا يمكن إنشاؤها، فيجب قياس جودة الارتباط وتقييم التأثير على المخرجات الناتجة.

يمكن وصف الطريقة المستخدمة حاليًا في مرحلتين:

- ***الطريقة الحتمية*. تعتمد هذه الطريقة على مطابقة المفاتيح لربط السجلات عبر المصادر الإدارية. يتم إنشاء مفاتيح المطابقة من خلال الجمع بين متغيرات التعريف الرئيسية (أو أجزاء منها) مثل الاسم والجنس وتاريخ الميلاد والرمز البريدي. يتم إنتاج نفس مجموعة مفاتيح المطابقة لكل مجموعة بيانات. إذا كانت مفاتيح المطابقة هي نفسها في كل مصدر، فسيتم إنشاء رابط.**

- ***الطريقة الاحتمالية*. يقوم هذا النهج بتحديد الروابط بين السجلات في مجموعتي بيانات من خلال مقارنة وقياس التشابه النسبي للسجلات (على سبيل المثال، إعطاء درجة التشابه). الفرق الرئيسي عن مرحلة المطابقة الحتمية هو أن الاحتمالات**

الاختلافات في المفاهيم والتعريفات

غالبًا ما تعتمد السجلات وغيرها من مصادر البيانات الإدارية مفاهيم وتعريفات مختلفة للمتغيرات المتعلقة بالسكان عن تلك التي تنطبق عمومًا على التعدادات التقليدية. ينبغي للمؤسسات الإحصائية الوطنية أن تدرك أن مثل هذه الاختلافات قد توجد وأن تقرر ما إذا كانت هذه الاختلافات مقبولة عند الانتقال من التعداد التقليدي إلى التعداد المركب أو التعداد القائم على السجلات. وما قد يعتبر في بلد ما فرقاً مقبولاً عند تقييم التوازن بين استمرارية وتماسك الإحصاءات الناتجة وانخفاض تكاليف العمل الميداني، قد يعتبر غير مقبول للمستخدمين في أماكن أخرى. وينبغي للمؤسسات الإحصائية الوطنية أن تزن الأمور جيداً قبل أن تقرر ما إذا كانت على استعداد لدفع هذا الثمن عند الانتقال إلى تعداد قائم على السجلات أو تعداد مركب دون تعداد ميداني كامل للمتغيرات المختارة. في بعض الأحيان، من الممكن تقريب التعريفات والمفاهيم الأصلية بشكل دقيق إلى حد ما من خلال اشتقاقات من مصادر مختلفة أو عن طريق تحرير المعلومات من مصادر التعداد المكتسبة حديثًا. ومع ذلك، ليس هذا هو الحال دائما، وينبغي للمعهد الوطني للإحصاء بعد ذلك أن يوازن بين قبول الاختلافات وتكاليف الاستمرار في تعداد الحقول الكاملة للمتغيرات المختارة. وفي الأردن، لوحظ حتى الآن أن العقبة الرئيسية تتعلق بتعريف الأسر.

التوقيت المناسب

إن المشكلة الخاصة التي تواجهها المؤسسات الإحصائية الوطنية عند الانتقال إلى التعداد المركب أو المعتمد على السجلات هي أن المصادر المختلفة للبيانات الإدارية غالباً ما يكون لها تواريخ مرجعية مختلفة. في بعض الأحيان، يوفر المصدر خيار التمييز بوضوح بين تواريخ المرجع وتواريخ الأحداث، ولكن هذه الممارسة الجيدة لا تنطبق دائمًا.

قياس الجودة/ضمان الجودة

وبغض النظر عن منهجية جمع البيانات، فإن تقييم جودة مخرجات بيانات التعداد كان دائما مهمة ضرورية ومهمة. هناك عدة طرق وأساليب مختلفة لتقييم جودة الإحصاءات، بما في ذلك جودة مخرجات التعداد السكاني. إن تقييم جودة التعداد الذي يستخدم منهجية جديدة أمر مهم بشكل خاص، لأنه يوفر معلومات ذات صلة بموثوقية نتائج التعداد الجديد، وكيف قد تختلف الجودة عن نتائج التعدادات السابقة.

الملحق 2. برنامج المهمة

الإثنين 06 يناير 2025
(جميع المشاركين من الإطار المرجعي)

- **الدولة المستفيدة:** الترحيب وتقديم الحالة الحالية لربط البيانات - الإنجازات والتحديات الحالية - جميع المشاركين (1.5 ساعة)
- **الدولة العضو:** مقدمة في ربط البيانات بدون معرفات مشتركة (2 ساعة) - جميع المشاركين
  - الطريقة الحتمية
  - مطابقات الاحتمالات للبيانات بدون معرفات مشتركة
- **الدولة المستفيدة والدولة العضو:** التحضير للعمل العملي في الأيام القادمة (1.5 ساعة) - جميع المشاركين
  - مراجعة مشتركة للبيانات والبيانات الوصفية التي ستستخدم في الأيام القادمة
  - التحقق من أن جميع أدوات التحليل تم نشرها على الخادم

تعزيز قدرات دائرة الإحصاءات العامة الأردنية

الثلاثاء 07 يناير 2025 والأربعاء 08 يناير 2025
(المشاركون - موظفو الدائرة الذين لديهم بالفعل خبرات عملية في ربط البيانات و/أو بايثون)

- **الدولة المستفيدة والدولة العضو: عمل** عملي على ربط البيانات بدون معرف مشترك. سيبدأ العمل فيما يلي:
  - عينة من 25.000 فرد تم إعدادها مسبقًا بواسطة الدائرة
  - ربط البيانات الذي تم بالفعل في الدائرة في بايثون (ستقدم الدائرة وصفًا منهجيًا لخبراء الفترة القصيرة قبل المهمة
  - مكتبة "Splink" التي تعتمد على كود Python/Spark
  - دليل المستخدم (التوثيق):
https://moj-analytical-services.github.io/splink/topic_guides/topic_guides_index.html
  - *التعليمات:*
https://moj-analytical-Services.github.io/splink/demos/tutorials/00_Tutorial_Introduction.html

الخميس 09 يناير 2025

- **الدولة المستفيدة والدولة العضو: صياغة** منهجية وإرشادات عملية لربط البيانات بناءً على الدروس المستفادة من الأيام السابقة - المشاركون من الأيام السابقة (2.5 ساعة)
- إذا سمح الوقت - الدولة المستفيدة والدولة العضو: تحديد المفاهيم مع الاختلافات في الأردن والتفكير في كيفية التعامل مع ذلك (جميع المشاركين من الإطار المرجعي)
- **الدولة المستفيدة والدولة العضو: الاستنتاجات** والتوصيات - جميع المشاركين من الإطار المرجعي وكذلك الإدارة العليا (½ ساعة)

الاختصارات:
MS = دولة عضو في الاتحاد الأوروبي (الدنمارك، ألمانيا، إيطاليا، ليتوانيا، فنلندا)
DoS = دائرة الإحصاءات العامة، الأردن

الملحق 3. الأشخاص الذين تم الالتقاء بهم

مديرية إدارة البيانات
- السيد جعفر عبابنة
- السيد محمد العمري
- السيدة لينا الجزازي
- السيد عبد الواحد الحرايزة
- السيدة صفاء أبو عيطة

مديرية المسوح السكانية والأسرية
- السيدة منار الجخ
- السيدة نور نعيمات

مديرية التحول الإلكتروني وتكنولوجيا المعلومات
- السيد محمد الشطناوي (خبير في بايثون)

تعزيز قدرات دائرة الإحصاءات العامة الأردنية

- السيد حسام أبو شكر
- السيد أيمن الهلول - (خبير في بايثون)

مركز البيانات الوطني

- السيد محمد الحيارى
- السيد أيمن العثامنة
- السيد زيد أبو رشيد
- السيد قصي حمدان
- السيد صهيب عنانبة