برامج التحليل الاحصائيR، SPSS، Python

يدخل الإحصاء في كثير من الاختصاصات العلمية كأداة من أهم أدوات البحث العلمي وتحليل البيانات بهدف استخلاص النتائج وإيجاد العلاقات الرياضية بين المتحولات. ولا بد أن أي طالب علم أو باحث قد سمع بما يسمى لغات البرمجة الاحصائية التي تستعمل لتطبيق القوانين الاحصائية على البيانات ومن أشهر هذه اللغاتSPSS, R و Python.

تعتبر اللغة SPSS الأكثر شهرة بين الباحثين لأن أغلب الجامعات تشتري رخص استخدام وتتيحها للباحثين مجاناً. المهتمون بالبرمجيات المجانية والبرمجيات مفتوحة المصدر يفضلون استخدام لغة R والتي تلاقي انتشاراً واسعاً هذه الأيام حتى أن البرمجيات التجارية تتيح تصدير البيانات والاتصال مع R و Python.

- SPSS هو اختصار لـ Statistical Package for the Social Sciences المطور للعلوم الاجتماعية وهو أول لغة برمجة إحصائية للحواسب الشخصية وقد طور في جامعة Stanford عام 1968 ثم استقل في شركة خاصة باسم SPSS Inc. وبعدها اشترته شركة IBM منذ 2009
- جامعة Auckland طورت لغة البرجة R تحت رخصة المشاع GNU للتركيز على النمذجة الاحصائية في العام 2000.
- اللغة الوحيدة التي لم تطور في أي جامعة هي لغة Python فقد طورها شخص ألماني وقد حرص على تبسيطها كلغة برمجة وهي ليست لغة احصائية لكنها تقبل إضافة الكثير من الـ Modules والتي تعتبر الوحدات الإحصائية من أهمها ما جعل لغة بايثون تصنف ضمن مجال لغات البرمجة الاحصائية.

بالنسبة للسعر تعتبر كل من SPSS ذات سعر مرتفع نسبياً بينما R و Python فمجانيتين. لكن SPSS أكثر انتشاراً لأنها أبسط في التعامل وواجهاتها الرسومية GUI أبسط وتغني عن كتابة أي كود برمجي وبالتالي يمكن تعلمها بسهولة أكبر بالمقارنة مع R و Python التي تحتاج لمهارات برمجية. ولعل تعلم لغة R هو الأصعب للمبتدئين لكن التقدم فيها يصبح أسهل بعد فهم المبادئ. ومن الجدير بالذكر أنه بالرغم من صحة أن Python تعتبر لغة بسيطة لكن الخصائص الاحصائية المضافة ليست بهذه السهولة.

Python هي بيئة تطويرية متعددة الأغراض والمهام ما يعطيها القوة. كما أنها مصممة لدعم التطبيقات ومن المزايا الاحصائية لها أنها تدعم التعلم الآلي .

تملك R مكتبات (reticulate, rPython) لدعم البرامج المكتوبة في Python كما وتدعم Python تشغيل البرامج المكتوبة بلغة R من خلال مكتبة (rpy2) مم يجعل البيئتين متكاملتين وقابلتين للاستخدام معاً

SPSS تم تطويرهما في الأساس في بيئة أكاديمية من أجل تطبيقات تركز على وصف البيانات. حالياً يملك أدوات كـ SPSS Modeler لكن يحتاج المرء لدفع سعر إضافي للحصول على ترخيص استخدام هذه الميزات. وهنا تبرز أفضلية للبرامج مفتوحة المصدر في أنها دائما قيد التطوير والتحديث والإضافة من قبل مجتمع كبير ومتعدد المتطلبات. مما يجعل المشاركات أكبر ويعطي لـ Rميزات قوية في المجالين الوصفي والتنبئي .
بالنسبة لـ Pythonفهو يدعم التطبيقات التجارية ويركز بشكل أكبر على الإحصاء التنبيئي وتطبيقات البيانات الكبيرة big data.