المحاضرة 2

الدرس الثاني: برامج وأدوات معالجة النصوص بايثون وتحليل النصوص (الجزء الأول)

في هذا الدرس، سننتقل من دور المستخدم العادي لبرامج معالجة النصوص إلى دور المطور الذي يطوع التكنولوجيا لخدمة البحث النفسي والتربوي. إن الهدف ليس تحويلكم إلى مبرمجين، بل تزويدكم بـ مشرط رقمي يساعدكم في تشريح النصوص وفهم الظواهر التربوية الكامنة وراء الكلمات.[1]

1. ما هي البرمجة؟ (لغة التفاهم مع الآلة)

لتبسيط المفهوم، البرمجة هي عملية إعطاء تعليمات محددة ومرتبة للحاسوب ليقوم بتنفيذ مهمة معينة. الحاسوب في أصله لا يفهم لغتنا البشرية، بل يفهم فقط الشيفرة الثنائية (Binary Code) المكونة من (0 و 1)، وهي تمثل حالات التيار الكهربائي (تشغيل أو إيقاف).

لغة البرمجة (مثل بايثون) هي المترجم الذي يأخذ أفكارنا المكتوبة بلغة قريبة من الإنجليزية ويحولها إلى لغة يفهمها الحاسوب. بالنسبة لكم كباحثين، البرمجة هي الطريقة التي تخبرون بها الحاسوب: اقرأ هذه المقابلات النفسية، واستخرج لي الكلمات التي تدل على القلق أو التوتر.[3]

2. لماذا بايثون لطلبة علم النفس والعلوم التربوية؟

تعتبر بايثون الخيار الأول عالمياً في مجال البحث العلمي والذكاء الاصطناعي لعدة أسباب تجعلها مثالية لتخصصكم:

البساطة والقابلية للقراءة: لغة بايثون تُكتب بكلمات إنجليزية واضحة، وقواعدها بسيطة جداً مقارنة باللغات الأخرى، مما يجعلها تبدو كأنها إنجليزية مكسرة.

مجانية ومفتوحة المصدر: يمكن لأي طالب تحميلها واستخدامها دون أي تكاليف مالية.

المجتمع والدعم: هناك ملايين الباحثين الذين يشاركون أدواتهم وبرامجهم الجاهزة في بايثون، مما يعني أنك ستجد دائماً حلاً لأي مشكلة تواجهك.

تعدد الاستخدامات: تُستخدم في تحليل البيانات الإحصائية، ومعالجة اللغات الطبيعية (NLP)، وحتى في بناء الاختبارات النفسية المحوسبة.[1]

3. تجهيز بيئة العمل (أين نكتب الأكواد؟)

لا نحتاج إلى معدات معقدة للبدء. يمكننا استخدام طريقتين:

Google Colab (الخيار الأسهل): وهو مختبر رقمي توفره جوجل مجاناً، يسمح لك بكتابة وتشغيل أكواد بايثون عبر المتصفح دون الحاجة لتثبيت أي برنامج على جهازك، وهو مثالي للمبتدئين.

تثبيت محلي (VS Code): لمن يفضل العمل دون إنترنت، يمكن تحميل لغة بايثون من الموقع الرسمي (python.org) واستخدام محرر نصوص متخصص مثل Visual Studio Code.

4. اللبنات الأساسية في بايثون (المتغيرات وأنواع البيانات)

في البحث النفسي، نتعامل مع أنواع مختلفة من المعلومات. في بايثون، نسمي هذه المعلومات بيانات، ونخزنها في متغيرات.

أ. المتغيرات (Variables):

تخيل المتغير كأنه صندوق تضع فيه معلومة وتلصق عليه ملصقاً (اسماً) لتعود إليه لاحقاً.

مثال: student_name = أحمد (هنا وضعنا اسم الطالب في صندوق سميناه student_name).

ب. أنواع البيانات (Data Types):

يحتاج الحاسوب لمعرفة نوع المعلومة التي يتعامل معها ليقوم بالعمليات الصحيحة عليها:

نوع البيانات	الاسم في بايثون	مثال	الاستخدام في علم النفس
النصوص	String (str)	يعاني الطالب من تشتت	كتابة الملاحظات السلوكية والتقارير
الأعداد الصحيحة	Integer (int)	85	درجات الاختبارات التحصيلية
الأعداد العشرية	Float	3.5	معدلات الذكاء أو الانحراف المعياري
القيم المنطقية	Boolean (bool)	True أو False	هل الطالب موهوب؟ (نعم/لا)

5. العمليات الأساسية على النصوص (String Operations)

بما أن مادتنا هي تحليل النصوص، فإن النصوص (Strings) هي أهم نوع بيانات بالنسبة لنا. بايثون توفر أدوات قوية للتعامل معها:

الطباعة (:print) هي الأمر الذي يظهر لنا النتائج على الشاشة.

تحويل حالة الأحرف: في اللغات الأجنبية، نستخدم lower() لتحويل كل النص إلى أحرف صغيرة لتوحيد البيانات (مثلاً جعل Anxiety و anxiety كلمة واحدة عند العد).

التقسيم (:split) تقطيع الجملة إلى كلمات منفردة ليسهل تحليلها.

6. تطبيق عملي مبسط: تحليل استجابة نفسية

تخيل أن لديك استجابة من طالب في مقابلة إكلينيكية، وتريد التأكد من وجود كلمات مفتاحية معينة فيها:

Python

```
# تخزين نص المقابلة في متغير
interview_text = أشعر بالقلق الشديد عندما أدخل قاعة الامتحان

# التأكد من وجود كلمة قلق في النص
if قلق in interview_text:
    print(تم رصد مؤشر نفسي: قلق)

# حساب طول النص (عدد الأحرف)
print(len(interview_text))
```

هذا الكود البسيط هو حجر الزاوية في تحليل المحتوى الآلي.

7. أهمية تحليل النصوص في البحث التربوي

إن الانتقال إلى تحليل النصوص المحوسب يحل معضلة البيانات الضخمة.[4] قديماً، كان الباحث يقرأ المقابلات يدوياً ويضع علامات بالأقلام الملونة (Coding)، وهي عملية مضنية ومعرضة للخطأ البشري. أما باستخدام أدوات بايثون، فيمكننا:

تحليل المشاعر :(Sentiment Analysis) تحديد ما إذا كان نص الطالب يعكس مشاعر إيجابية أو سلبية أو محايدة.

استخراج الكلمات المفتاحية: معرفة أكثر المواضيع تكراراً في شكاوى الطلاب.

التصنيف الآلي: تصنيف التقارير التربوية إلى فئات (مثلاً: صعوبات تعلم، اضطرابات سلوكية) بناءً على الكلمات المستخدمة فيها.

الخلاصة والتوصيات

لقد تعلمنا اليوم أن البرمجة هي لغة تواصل، وأن بايثون هي الأداة الأنسب لتخصصنا بسبب بساطتها وقوتها. إن فهمنا للمتغيرات وأنواع البيانات هو الأساس الذي سنبني عليه في الدرس القادم مهارات أكثر تعمقاً في تحليل تكرار الكلمات وكيفية قراءة الملفات النصية الكبيرة.

نصيحة الأستاذ: لا تقلقوا من الأكواد في البداية؛ تعاملوا معها كأنها وصفة طبخ (Recipe) تتبعون خطواتها بدقة لتصلوا إلى نتيجة شهية (بيانات دقيقة).

■ في المحاضرة القادمة، سنقوم معاً بكتابة أول برنامج حقيقي يقوم بقراءة ملف نصي كامل واستخراج إحصائيات حول المفردات المستخدمة فيه.