المحور الثاني: تحليل البيانات باستخدام بايثون Python:

- **مقدمة في مكتبة Pandas:**

هي مكتبة برمجية مكتوبة للغة بايثون وتعد من أشـــهر المكتبات في مجتمع علوم البيانات حيث توفر هياكل بيانات قوية ومعبرة ومرنة تجعل معالجة البيانات وتحليلها أمراً ســـهلاً.[1] في لغة برمجة بايثون يمكنك اســـتخدام Python مع Pandas في مجموعة واســـعة من المجالات المتنوعة فيما يتضـــمن ذلك المجالات الأكاديمية والتجارية وأيضـا التخصـصـات المالية والاقتصـادية والإحصـائية والتحليلات وما إلى ذلك في هذا البرنامج التعليمي.[2]

1. تثبيت مكتبة pandas للبايثون:

لكي تستطيع استخدام مكتبة **Pandas** يجب عليك تثبيتها على جهازك اولا لكي تتمكن من استعمالها في python وذلك يكون عن طريق كتابة الامر التالي ســـواء في cmd او tirmenal الخاص بال vs code.

```
python -m pip install pandas –users
```

2. مميزات مكتبة pandas:

- **التعامل مع الجداول والبيانات المركبة**: إنشـــاء Data Framesلتحليل البيانات بطريقة مشـــابهة لبرامج مثل Excel.
- **تنظيف البيانات:** التعامل مع القيم الناقصة، تصحيح الأخطاء، وتحويل البيانات.
- **التصفية والفرز:** استخراج البيانات المطلوبة بسرعة باستخدام وظائف تصفية متقدمة.
- **التحليل الإحصـــائي:** حســـاب المعدلات، الانحراف المعياري، والمجموعات باســـتخدام وظائف مُدمجة.

مثال تطبيقي باستخدام:Pandas

Python
Copy code

```
Import pandas as pd

# إنشاء Data Frame بسيط
data = {'المنتج': ['منتج A', 'منتج B', 'منتج C'], 'المبيعات': [200, 340, 150]}
df = pd. DataFrame(data)
# تحليل بسيط
print(df.describe()) # الإحصائيات الأساسية
```

- **كيفية استخدامها كأداة رئيسية لمعالجة وتحليل البيانات الجدولية (CSV. Excel):**

---

[1] تحضير البيانات باستخدام مكتبة Pandas بلغة بايثون.
https://bayandata.sa/1195/%D8%AA%D8%AD%D8%B6%D9%8A%D8%B1-%D8%A7%D9%84%D8%A8%D9%8A%D8%A7%D9%86%D8%A7%D8%AA (18-04-2026)

[2] geecoders
https://geecoders.com/pandas-pandas-library-with-examples-in-html/(18-04-2026)

1. تشغيل Pandas

```
import pandas as pd
```

أ/ مصفوفة أحادية البعد تستقبل كل البيانات:

```
s = pd.Series([3, -5, 7, 4],  index=['a',  'b',  'c',  'd'])
```

النتيجة:

```
i                                3
```

أ/ مصفوفة ثنائية الأبعاد تستقبل كل البيانات:

الرقم	الدولة	العاصمة	السكان
1	بلجيكا	بروكسل	11190846
2	الهند	نيودلهي	1303171035
3	البرازيل	برازيليا	207847528

تحليل والنتيجة:

```
data = {'Country': ['Belgium',  'India',  'Brazil'],

'Capital': ['Brussels',  'New Delhi',  'Brasilia'],

'Population': [11190846, 1303171035, 207847528]}

df = pd.DataFrame(data,columns=['Country',  'Capital',  'Population'])
```

أولا: القراءة ضمن ملفات CSV:

ملف CSV هو نوع بسيط من الملفات النصية العادية التي تستخدم بنية محددة لترتيب البيانات الجدولية. يتم تحديد التنسيق القياسي لملف CSV من خلال بيانات الصفوف والأعمدة حيث ينهي سطر جديد كل صف لبدء الصف التالي، ويتم فصل كل عمود بفاصلة داخل الصف.

يعد CSV تنسيقًا شائعًا لتبادل البيانات لأنه مضغوط وبسيط وعامة. تسمح العديد من الخدمات عبر الإنترنت لمستخدميها بتصدير البيانات الجدولية من موقع الويب إلى ملف CSV ، سيتم فتح ملفات CSVفي برنامجExcel ، وتحتوي جميع قواعد البيانات تقريبًا على أداة للسماح بالاستيراد من ملفات .CSV.

تسمى البيانات الموجودة في شكل جداول أيضًا بـــ CSV (قيم مفصولة بفواصل) - حرفيًا "قيم مفصولة بفواصل"، هذا تنسيق نصي مخصص لعرض البيانات الجدولية، كل سطر من الملف هو سطر واحد من الجدول، يتم فصل قيم الأعمدة الفردية برمز فاصل - فاصلة (،) أو فاصلة منقوطة (؛) أو رمز آخر. يمكن قراءة ملف CSV ومعالجته بسهولةPython .

❖ **وظائف وحدةCSV :**

يمكنك العثور على الوظائف التالية في وثائق وحدةCSV :

- **csv. field_size_limit:** إرجاع الحد الأقصى لحجم الحقل.
- **csv.get_dialect:** تحصل على اللهجة المرتبطة بالاسم.
- **csv.list_dialects:** إظهار جميع اللهجات المسجلة.
- **csv.reader**: قراءة البيانات من ملف.CSV .
- **csv.register_dialect**: ربط اللهجة بالاسم.
- **csv.writer**: كتابة البيانات إلى ملف CSV .
- csv.unregister_dialectحذف اللهجة المرتبطة باسم سجل اللهجة
- csv.QUOTE_ALL: اقتبس كل شيء، بغض النظر عن نوعه.
- csv.QUOTE_MINIMAL: حقول الاقتباس بأحرف خاصة.
- csv.QUOTE_NONNUMERIC: اقتبس جميع الحقول التي لا تحتوي على قيم رقمية.
- csv.QUOTE_NONE: لا تقتبس أي شيء في الإخراج.

❖ **كيفية قراءة ملف CSV فيPython**

1) **الخطوة 1:** لقراءة البيانات من ملفاتCSV ، يجب عليك اســـتخدام وظيفة القارئ لإنشـــاء كائن قارئ، تم تطوير وظيفة القارئ لأخذ كل صـــف من الملف وإنشـــاء قائمة بجميع الأعمدة. ثم عليك اختيار العمود الذي تريد بيانات المتغير له.

مثال تطبيقي:

```
import necessary modules#
                    import csv
:with open('X:\data.csv','rt')as f
            data = csv.reader(f)
                :for row in data
                print(row)
```

2) **الخطوة 2:** عند تنفيذ البرنامج أعلاه ستكون النتيجة:

```
['Programming language; Designed by; Appeared; Extension']
['Python; Guido van Rossum; 1991; .py']
['Java; James Gosling; 1995; .java']
['C++; Bjarne Stroustrup;1983;.cpp']
```

❖ **قراءة ملف CSV باستخدامPandas**

Pandas هي مكتبة مفتوحة المصـــدر تســـمح لك باســـتيراد ملف CSV فيها Python وتنفيذ معالجة البيانات. توفر الباندا طريقة سهلة لإنشاء البيانات ومعالجتها وحذفها.

يجب عليك:

1. تثبيت مكتبة الباندا باستخدام الأمر pip install pandas في Windows .
2. تقوم بتنفيذ هذا الأمر في موجه الأوامر أثناء وجودك في Linux في المحطة الطرفية.
3. تعد قراءة ملف CSV في Pandas DataFrame أمرًا سريعًا.

```
import necessary modules#
import pandas
result = pandas.read_csv('X:\data.csv')
print(result)
```

النتيجة:

```
Programming language, Designed by, Appeared, Extension
Python, Guido van Rossum, 1991, .py    0
Java, James Gosling, 1995, .java    1
C++, Bjarne Stroustrup,1983,.cpp    2
```

❖ **كتابة ملف CSV باستخدام Pandas:**

إن الكتابة إلى ملف CSV باستخدام Pandas أمر سهل مثل القراءة:

1. يجب عليك إنشاء إطار بيانات بناءً على ما يلي Python الكتابة إلى رمز.CSV

```
from pandas import DataFrame
,C = {'Programming language': ['Python','Java', 'C++']
,Designed by': ['Guido van Rossum', 'James Gosling', 'Bjarne Stroustrup']'
,Appeared': ['1991', '1995', '1985']'
,Extension': ['.py', '.java', '.cpp']'
{
df = DataFrame(C, columns= ['Programming language', 'Designed by', 'Appeared',
'Extension'])
export_csv = df.to_csv (r'X:\pandaresult.csv', index = None, header=True) # here you
have to write path, where result file will be stored
print (df)
```

النتيجة:

```
Programming language, Designed by, Appeared, Extension
Python, Guido van Rossum, 1991, .py    0
Java, James Gosling, 1995, .java    1
C++, Bjarne Stroustrup,1983,.cpp    2
```

ثانيا: : القراءة ضمن ملفات Excel:

Excel هو تطبيق جداول بيانات من Microsoft لأنظمة Windows و macOS وAndroid وiOS. يتم استخدام جدول البيانات لتخزين وتحليل البيانات الجدولية المنظمة في الصفوف والأعمدة.

يعد Excel أحد التنسيقات الأكثر استخدامًا لتبادل البيانات. يرجع شعبيته إلى حقيقة أن نظام التشغيل Windows شائع الاستخدام، يحظى برنامج Excel أيضًا بشعبية لدى الأشخاص غير التقنيين. بعض الأسباب لذلك هي:

- ❖ سهل الاستخدام.
- ❖ يتمتع بوظائف قوية تجعل التصور والتحليل أسهل.
- ❖ تحتوي معظم قواعد البيانات وأدوات معالجة البيانات على واجهات للعمل مع.Excel

إن أسهل طريقة لقراءة ملف Excel باستخدام Pandas هي ببساطة تمرير مسار الملف إلى المعامل io =لذا يمكنك ببساطة تمرير السلسلة إلى الملف، يقبل المعامل كل المسارات سواء كان الملف على جهاز محلي أو مسار HTTP أو مسار FTP أو أكثر.

```
import pandas as pd

df = pd.read_excel('https://github.com/datagy/mediumdata/raw/master/Sales.xlsx')
print(df.head())

# Returns:
#         Date Customer  Sales
# 0 2022-04-01        A    191
# 1 2022-04-02        B    727
# 2 2022-04-03        A    782
# 3 2022-04-04        B    561
# 4 2022-04-05        A    969
```

- **<u>التحليل الوصفي للبيانات: استخدام وظائف Pandas لإجراء تحليلات إحصائية وصفية بسيطة (مثل: حساب المتوسط، الوسيط، الانحراف المعياري) على البيانات السياسية:</u>**

تسمح مكتبة Pandas بإجراء العديد من التحليلات الدقيقة على البيانات السياسية منها الانتخابات سواء الرئاسية أو التشريعية، ويمكن حسابها بعمليات إحصائية.

<u>مثال تطبيقي:</u>

دراسة نسبة المشاركة الانتخابية باستخدام المتوسط الحسابي والانحراف المعياري:

1. **نقوم بإدخال البيانات في برنامج Python**

```
Python
Import pandas as pd
data={"region ":  ["A "  ,    "B "   ,  "C "    ,     "D "    ] ,
"participation_rate" :  [35,  80  , 40, 66  ,  60] }
df = pd. Datafrrame(data)
```

2. **نقوم بحساب المتوسط الحسابي:**

```
Python
Std_value =
df[ " participation_rate"].mean( )
Print(   " متوسط المشاركة,  mean_value) "
```

النتيجة: إذا كانت نسبة المشاركة تقارب 52 % فإنها مشاركة نسبية.

3. **حساب الانحراف المعياري :**

```
Python
Std_value =
```

```
df[ " participation_rate"].std( )
Print(  " الانحراف المعياري,  std_value) "
```

النتيجة: إذا كانت نتيجة الانحراف المعياري تقارب 15 فإن نسبة المشاركة الانتخابية متفاوتة بين المناطق أو الولايات عبر ربوع الوطن.

4. **تحليل النتائج التي تم التوصل إليها.**

- **تصور البيانات: استخدام مكتبة Matplotlibn:**

تعتبر هذه المكتبة أحد أشهر مكتبات البايثون في ال Visualization ، وتمكنك من بناء الرسومات البيانية والتوضيحية بشكل سهل ومرن، وترتبط هذه المكتبة بعدة مكتبات وأدوات أخرى منها مكتبة **Pandas**وأداة **Jupyter** لتمكن المستخدمين من استدعاء أوامر الرسومات بشكل مباشر على إطار البيانات وبطريقة تفاعلية.

كما أن المرونة الموجودة في هذه المكتبة تمكنهم من ترتيب الأشكال والرسومات بطريقة سهلة ومريحة، وتستطيع أيضا أن تصدر المخرجات على شكل ملفات بصيغ متعددة منها Pdf, Jpg, SVG, PNG, BMP, GIFوغيرها.

❖ **تثبيت مكتبةMatplotlib :**

1) توجه أولاً إلى موقع Matplotlib.org
2) تنزيل الإصدار المتوافق مع إصدار بايثون لديك.
3) التأكد من تثبيت مكتبات pyparsing و dateutilو sixوnumpy، وربما بعض المكتبات الأخرى المذكورة في الفيديو، يمكنك الحصول على جميع هذه المكتبات أيضًا إذا كنت تستخدم نظام ويندوز من خلال زيارة الرابط التالي : http://www.lfd.uci.edu/~gohlke/pythonlibs/#matplotlib
4) بعد تثبيت مكتبةMatplotlib ، تأكد من فتح نافذة طرفية أو برنامج نصي، واكتب ما يلي:

```
matplotlib استيراد مكتبة
```

5) أكد من عدم وجود أخطاء في عملية الاستيراد. إذا وُجدت أخطاء، فاقرأ تفاصيلها. في أغلب الأحيان، يكون السبب إما عدم تطابق إصدار النظام (64 بت مقابل 32 بت)، أو أنك تفتقد حزمةً مثل dateutilأو pyparsing.

بمجرد أن تتمكن من استيراد مكتبة matplotlib بنجاح، ستكون جاهزًا للمتابعة.

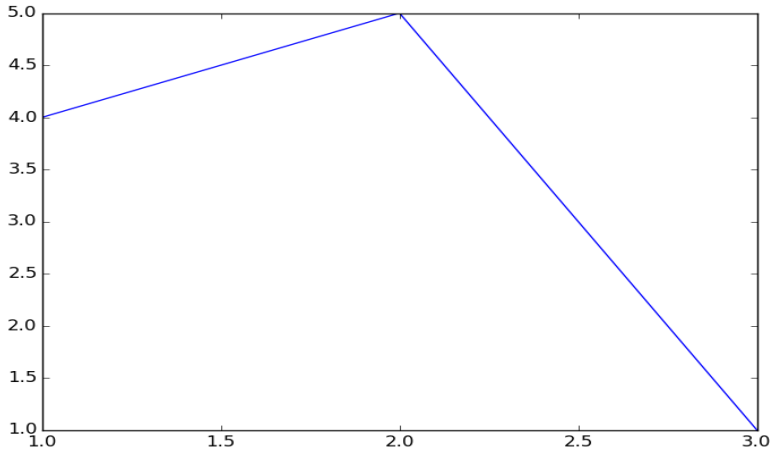
هناك بعض التعليمات البرمجية الأساسية لإنشاء أحد أبسط الرسوم البيانية التي يمكننا إنشاؤها، ولن يستغرق الأمر منا سوى 3 أسطر.

```
#استيراد pyplot
from matplotlib import pyplot as plt

#التخطيط على لوحتنا
plt.plot([1,2,3],[4,5,1])

عرض ما خططنا له
plt.show()
```

6) الرسم البياني الناتج:


5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
1.0
1.5
2.0
2.5
3.0