

الوحدة الثانية: معالجة البيانات وتحليلها Data Processing and Analysis

الدرس الثالث: جمع البيانات وتحليلها باستخدام برمجية بايثون (Python Data Collection & Analyses Methods)

القسم الأول: أساسيات التعامل مع البيانات في بايثون

ما هو الأمر الذي يُستخدم بشكل شائع لتنصيب مكتبة تحليل بيانات خارجية في بايثون، مثل Pandas أو Matplotlib ؟

أ. `install package_name`

ب. `get package_name`

ج. `pip install package_name`

د. `require package_name`

الإجابة الصحيحة: ج. `pip install package_name`

لتحقيق الاستفادة القصوى من مكتبات تحليل البيانات في بايثون، يجب عليك أولاً:

أ. إعادة تشغيل الكمبيوتر.

ب. تثبيت جميع المكتبات المتاحة.

ج. استيراد المكتبات المطلوبة إلى بيئة العمل الخاصة بك (مثل Jupyter Notebook أو ملف.py).

د. تغيير إعدادات نظام التشغيل.

الإجابة الصحيحة: ج. استيراد المكتبات المطلوبة إلى بيئة العمل الخاصة بك (مثل Jupyter Notebook أو ملف.py).

أي من الأوامر التالية هو الأكثر شيوعاً لاستيراد مكتبة Pandas في بايثون، مع استخدام اختصار متعارف عليه؟

أ. `import pandas`

ب. `import pandas as pd`

ج. `load pandas`

د. `pandas.import()`

الإجابة الصحيحة: ب. `import pandas as pd`

تُعد ملفات CSV (Comma Separated Values) صيغة شائعة لتخزين البيانات الجدولية. ما هي دالة Pandas التي تُستخدم لقراءة ملف CSV وتحويله إلى DataFrame؟

أ. `pd.read_excel()`

ب. `pd.load_csv()`

ج. `pd.from_csv()`

د. `pd.read_csv()`

الإجابة الصحيحة: د. `pd.read_csv()`

إذا كان لديك ملف بيانات بتنسيق Excel (.xlsx)، أي دالة من مكتبة Pandas يمكنك استخدامها لاستدعاء هذا الملف؟

أ. `pd.read_text()`

ب. `pd.read_json()`

ج. `pd.read_excel()`

د) `pd.read_sql()` .

الإجابة الصحيحة: ج) `pd.read_excel()` .

بعد قراءة ملف بيانات إلى **DataFrame** باسم `my_data` ، ما هي الدالة التي ستستخدمها لعرض أول 5 صفوف من هذا **DataFrame** لمراجعة سريعة للبيانات؟

أ) `my_data.tail()` .

ب) `my_data.info()` .

ج) `my_data.describe()` .

د) `my_data.head()` .

الإجابة الصحيحة: د) `my_data.head()` .

للحصول على نظرة عامة سريعة على عدد الصفوف والأعمدة وأنواع البيانات وعدد القيم غير المفقودة لكل عمود في **DataFrame** ، ما هي الدالة الأنسب؟

أ) `df.head()` .

ب) `df.shape` .

ج) `df.info()` .

د) `df.columns` .

الإجابة الصحيحة: ج) `df.info()` .

ما هي الدالة في **Pandas** التي تُستخدم لحساب الإحصائيات الوصفية الأساسية (مثل المتوسط، الانحراف المعياري، القيم الدنيا والقصوى) للأعمدة الرقمية في **DataFrame** ؟

أ) `df.summary()` .

ب. df.statistics()

ج. df.analyze()

د. df.describe()

الإجابة الصحيحة: د. df.describe()

إذا أردت عرض جميع الأعمدة في DataFrame باسم df ، ولكن دون رؤية كل الصفوف، ما هي الخاصية أو الدالة التي تمنحك أسماء الأعمدة فقط؟

أ. df.shape

ب. df.index

ج. df.columns

د. df.values

الإجابة الصحيحة: ج. df.columns

أي من المكتبات التالية في بايثون تُستخدم بشكل أساسي لتوليد الرسوم البيانية وتصور البيانات؟

أ. NumPy

ب. Pandas

ج. Scikit-learn

د. Matplotlib

الإجابة الصحيحة: د. Matplotlib

Prepared by: موقع المناهج الدراسية الأردن

الوحدة الثانية: معالجة البيانات وتحليلها Data Processing and Analysis

الدرس الثالث: جمع البيانات وتحليلها باستخدام برمجية بايثون (Python Data Collection & Analyses Methods)

القسم الثاني: تحليل وعرض البيانات الوصفية باستخدام بايثون

إذا أردت معرفة عدد القيم الفريدة في عمود معين (مثل عمود 'المدينة') في DataFrame باسم df ، فما هي الدالة الأنسب في Pandas ؟

أ] df['المدينة'].count()

ب] df['المدينة'].sum()

ج] df['المدينة'].unique()

د] df['المدينة'].value_counts()

الإجابة الصحيحة: د] df['المدينة'].value_counts()

ما هي الدالة التي تُستخدم في Pandas لحساب متوسط قيم عمود رقمي معين في DataFrame ؟

أ] df['العمود'].mode()

ب] df['العمود'].median()

ج] df['العمود'].mean()

د] df['العمود'].max()

الإجابة الصحيحة: ج] df['العمود'].mean()

عند الرغبة في تحديد القيمة الأكثر تكرارًا (المنوال) في عمود من DataFrame في Pandas، أي دالة تستخدم؟

أ. df[''].average()

ب. df[''].median()

ج. df[''].mode()

د. df[''].sum()

الإجابة الصحيحة: ج. df[''].mode()

أي مكتبة بايثون تُستخدم عادة لإنشاء رسوم بيانية تفاعلية وغنية بالمعلومات مثل المخططات الشريطية والخطية والمبعثرة؟

أ. NumPy .

ب. SciPy .

ج. Matplotlib .بالإضافة إلى. Seaborn)

د. Requests .

الإجابة الصحيحة: ج. Matplotlib .بالإضافة إلى. Seaborn)

لكتابة سطر برمجي يقوم بإنشاء رسم بياني خطي (Line Plot) لعمود بيانات اسمه "المبيعات" مقابل "الشهر" باستخدام Matplotlib، ما هو الأمر الشائع لبدء الرسم البياني؟

أ. plt.bar().الشهر, المبيعات)

ب. plt.plot().الشهر, المبيعات)

ج. plt.scatter().الشهر, المبيعات)

د. plt.pie().المبيعات)

الإجابة الصحيحة: ب) `plt.plot()`. الشهر، المبيعات).

بعد إنشاء الرسم البياني باستخدام **Matplotlib** ، ما هي الدالة التي تُستخدم لعرض الرسم البياني فعلياً في نافذة منفصلة أو ضمن مخرجات **Jupyter Notebook** ؟

أ) `plt.save()` .

ب) `plt.open()` .

ج) `plt.show()` .

د) `plt.display()` .

الإجابة الصحيحة: ج) `plt.show()` .

أي من المفاهيم التالية يُشير إلى ترتيب البيانات في **DataFrame** بناءً على قيم عمود واحد أو أكثر (تصاعدياً أو تنازلياً)؟

أ. التجميع. (Grouping)

ب. التصفية. (Filtering)

ج. الفرز. (Sorting)

د. الدمج. (Merging)

الإجابة الصحيحة: ج. الفرز. (Sorting)

عند التعامل مع القيم المفقودة (Missing Values) في **DataFrame** في بايثون، ما هي الدالة الشائعة في **Pandas** التي تُستخدم لإزالة الصفوف التي تحتوي على قيم مفقودة؟

أ) `df.fill_na()` .

ب. `df.drop_values()` .

ج. `df.dropna()` .

د. `df.remove_na()` .

الإجابة الصحيحة: ج. `df.dropna()` .

ما هي الدالة التي تُستخدم في Pandas لملء القيم المفقودة (NaN) في DataFrame بقيمة محددة (مثل المتوسط أو الصفر)؟

أ. `df.replace_na()` .

ب. `df.fillna()` .

ج. `df.add_values()` .

د. `df.update_na()` .

الإجابة الصحيحة: ب. `df.fillna()` .

إذا أردت حساب الانحراف المعياري (Standard Deviation) لعمود رقمي في DataFrame باستخدام Pandas ، فما هي الدالة المناسبة؟

أ. `df[''].mean()` العمود .

ب. `df[''].variance()` العمود .

ج. `df[''].std()` العمود .

د. `df[''].sum()` العمود .

الإجابة الصحيحة: ج. `df[''].std()` العمود .

Prepared by: موقع المناهج الدراسية الأردن

الوحدة الثانية: معالجة البيانات وتحليلها Data Processing and Analysis

الدرس الثالث: جمع البيانات وتحليلها باستخدام برمجية بايثون (Python Data Collection & Analyses Methods)

القسم الثالث: معالجة البيانات وتحضيرها في بايثون

في مكتبة Pandas ، ما هي الخاصية التي تُستخدم لمعرفة أبعاد DataFrame عدد الصفوف وعدد الأعمدة؟

أ. df.size .

ب. df.ndim .

ج. df.shape .

د. df.axes .

الإجابة الصحيحة: ج. df.shape .

ما هي الدالة التي تُستخدم للتحقق من وجود قيم مفقودة (NaN) في DataFrame ، وترجع DataFrame من القيم المنطقية (True/False) ؟

أ. df.hasna() .

ب. df.missing() .

ج. df.isnull() .

د. df.checkna() .

الإجابة الصحيحة: ج. df.isnull() .

إذا أردت معرفة عدد القيم غير المفقودة لكل عمود في DataFrame ، أي دالة من Pandas يمكن أن تمنحك هذه المعلومات كجزء من ملخص؟

أ) df.head()

ب) df.describe()

ج) df.info()

د) df.value_counts()

الإجابة الصحيحة: ج) df.info()

ما هي الدالة في Pandas التي تُستخدم لإزالة الأعمدة التي تحتوي على قيم مفقودة؟

أ) df.drop_rows(axis=1)

ب) df.drop(columns='missing_values')

ج) df.dropna(axis=1)

د) df.remove_cols()

الإجابة الصحيحة: ج) df.dropna(axis=1)

لتغيير نوع البيانات لعمود معين في DataFrame على سبيل المثال، من نص إلى رقم أو من رقم إلى تاريخ، ما هي الطريقة الشائعة في Pandas ؟

أ) df['العمود'].change_type()

ب) df['العمود'].astype(new_type)

ج) df['العمود'].convert_to()

د) df['العمود'].set_type()

الإجابة الصحيحة: ب. `df['العمود'].astype(new_type)`

إذا أردت تصفية `DataFrame` لعرض الصفوف التي تحقق شرطاً معيناً (مثال: 'العمر' أكبر من 30)، فما هو الأسلوب الصحيح في `Pandas` ؟

أ. `df[df['العمر'] < 30]`

ب. `df.filter(العمر < 30)`

ج. `df.where(العمر < 30)`

د. `df.select(العمر < 30)`

الإجابة الصحيحة: أ. `df[df['العمر'] < 30]`

ما هي الدالة التي تُستخدم لحفظ `DataFrame` من بايثون إلى ملف `CSV` جديد؟

أ. `df.to_excel()`

ب. `df.save_csv()`

ج. `df.write_csv()`

د. `df.to_csv()`

الإجابة الصحيحة: د. `df.to_csv()`

عند تحليل البيانات الوصفية، ما هو المقياس الذي يمثل القيمة الوسطى في مجموعة بيانات مرتبة (يوجد 50% من القيم أقل منها و 50% أعلى منها)؟

أ. المتوسط. (Mean)

ب. المنوال. (Mode)

ج. الوسيط.(Median)

د. الانحراف المعياري.(Standard Deviation)

الإجابة الصحيحة: ج. الوسيط.(Median)

ما هو مقياس التشتت الذي يقيس مدى انتشار القيم حول المتوسط؟

أ. المنوال.(Mode)

ب. المدى.(Range)

ج. الانحراف المعياري.(Standard Deviation)

د. الوسيط.(Median)

الإجابة الصحيحة: ج. الانحراف المعياري.(Standard Deviation)

تُستخدم دالة `value_counts()` في Pandas بشكل أساسي لتحديد:

أ. متوسط القيم في عمود.

ب. تكرار كل قيمة فريدة في عمود.

ج. أقصى قيمة في عمود.

د. أصغر قيمة في عمود.

الإجابة الصحيحة: ب. تكرار كل قيمة فريدة في عمود.

Prepared by: موقع المناهج الدراسية الأردن