

Dataanalys med Python: Grundläggande Verktyg och Tekniker

Redogörelse

- **Årskurs:** Gymnasiet
- **Ämne:** Digital teknik – Dataanalys med Python: grunder och verktyg
- **Tema:** Introduktion till datahantering och analys med Python

Ordkollen

Här listas tio ämnesord på läxans tema som är bra att känna till betydelsen av.

1. **Pandas:** Ett populärt Python-bibliotek för datahantering och analys.
2. **DataFrame:** En tvådimensionell tabellstruktur i Pandas för att lagra data.
3. **Series:** En endimensionell array i Pandas som kan innehålla data av olika typer.
4. **Importera:** Att hämta data från externa källor till ditt Python-program.
5. **Exportera:** Att spara data från ditt Python-program till externa filer.
6. **Filtrering:** Processen att välja specifika data baserat på vissa kriterier.
7. **Aggregation:** Sammanställning av data, t.ex. beräkning av medelvärden eller summor.
8. **Visualisering:** Skapandet av grafer och diagram för att presentera data visuellt.
9. **Jupyter Notebook:** En webbaserad applikation för att skapa och dela dokument som innehåller kod, visualiseringar och text.
10. **API:** Ett gränssnitt som tillåter olika mjukvarukomponenter att kommunicera med varandra.

Instuderingsfrågor

1. Vad är Pandas och varför används det inom dataanalys?
2. Beskriv vad en DataFrame är och ge ett exempel på hur den kan användas.
3. Hur skiljer sig en Series från en DataFrame i Pandas?
4. Förklara processen för att importera en CSV-fil till en Pandas DataFrame.
5. Vad innebär det att filtrera data i en DataFrame? Ge ett exempel.
6. Hur kan du exportera en DataFrame till en Excel-fil?
7. Vad är aggregation och vilka typer av aggregationsfunktioner finns tillgängliga i Pandas?
8. Beskriv hur du kan visualisera data med hjälp av Matplotlib i Python.
9. Vad är en Jupyter Notebook och vilka fördelar erbjuder den för dataanalys?
10. Förklara vad ett API är och ge ett exempel på hur du kan använda ett

API i en dataanalysprojekt.

Övning

Nedan listas uppgifter och fyra svarsalternativ. Du ska ringa in det alternativ som är korrekt. Observera att av de fyra alternativen är endast ett korrekt.

Uppgift	A	B	C	D
1. Vad används Pandas främst för i Python?	Bildbearbetning	Datahantering och analys	Webbdesign	Spelutveckling
2. Hur skapar du en DataFrame från en lista i Pandas?	<code>pd.create_dataframe()</code>	<code>pd.DataFrame()</code>	<code>pd.make_df()</code>	<code>pd.new_dataframe()</code>
3. Vilken metod använder du för att läsa en CSV-fil?	<code>pd.read_csv()</code>	<code>pd.load_csv()</code>	<code>pd.import_csv()</code>	<code>pd.get_csv()</code>
4. Hur kan du visa de första fem raderna i en DataFrame?	<code>df.tail()</code>	<code>df.head()</code>	<code>df.first()</code>	<code>df.show()</code>
5. Vilken funktion används för att beräkna medelvärde i en kolumn?	<code>df.average()</code>	<code>df.mean()</code>	<code>df.mid()</code>	<code>df.sum()</code>
6. Hur filtrerar du rader där värdet i kolumnen 'Ålder' är över 30?	<code>df.filter(df.Age > 30)</code>	<code>df[df['Ålder'] > 30]</code>	<code>df.select(Age > 30)</code>	<code>df.choose(Age > 30)</code>
7. Vilket bibliotek används ofta tillsammans med Pandas för visualisering?	Matplotlib	TensorFlow	Django	Flask
8. Hur exporterar du en DataFrame till en Excel-fil?	<code>df.to_excel('file.xlsx')</code>	<code>df.save_excel('file.xlsx')</code>	<code>df.write_excel('file')</code>	<code>df.export('file.xlsx')</code>
9. Vad gör funktionen <code>df.describe()</code> i Pandas?	Ger en beskrivning av DataFrame	Beräknar statistik för numeriska kolumner	Visar DataFrame som text	Rensar DataFrame från nullvärden
10. Vilken Jupyter-kommandorad används för att installera paket?	<code>%install</code>	<code>%pip install</code>	<code>%pkg install</code>	<code>%get install</code>

Skrivuppgifter

Här presenteras tre olika skrivuppgifter som är utformade på tre olika svårighetsnivåer: enkel, medel och svår.

Skrivuppgift 1: Introduktion till Pandas

Beskriv vad Pandas-biblioteket är och ge exempel på två vanliga användningsområden inom dataanalys. Förklara varför Pandas är ett värdefullt verktyg för dataanalytiker.

Svarslängd: ca. 250 ord (En halv sida)

Skrivuppgift 2: Dataimport och -export

Förklara stegen för att importera en CSV-fil till en Pandas DataFrame och sedan exportera en modifierad DataFrame till en Excel-fil. Inkludera kodexempel och beskriv vad varje steg gör.

Svarslängd: ca. 400 ord (En sida)

Skrivuppgift 3: Dataanalys och Visualisering

Välj ett dataset du är intresserad av och beskriv hur du skulle använda Python och Pandas för att analysera detta dataset. Inkludera vilka analyser du skulle utföra, vilka kolumner du skulle fokusera på, och hur du skulle visualisera resultaten. Diskutera även eventuella utmaningar du kan stöta på och hur du skulle hantera dem.

Svarslängd: ca. 500 ord (Ett och en halv sida)

Tags: [Läxa](#)