

Inlämningsuppgift kursen Grundläggande Programmering i Python, feb. 2024.

Syfte och mål

Syftet med inlämningsuppgiften är att du ska visa, inte minst för dig själv, att du kan skapa ett Python-program utifrån en definierad problembeskrivning där flera av de färdigheter som du fått i kursen används. Målet med inlämningsuppgiften är att du efter att ha slutförd uppgifterna har fått en fördjupad kunskap i grundläggande programmering.

Förutsättningar

All administrativ information som datum, tider och instruktioner för kamratgranskningen och komplettering samt betyg & rättning, osv finns på uppgiftssidan i Canvas.

Uppgiften går ut på att skapa ett antal funktioner som sätts samman till ett komplett program med hjälp av ett enkelt menysystem och ger oss ett sammanhållet program. Uppgiften utförs i form av 6 st. deluppgifter, där den första uppgiften blir att skapa menysystemet som sedan används för att köra funktionerna från övriga deluppgifter. Under arbetet kommer menysystemet uppdateras.

Vid inlämningstillfälle 1 bedöms hela uppgiften med bedömningen Godkänd (G) Komplettering (K) eller Underkänd (U). Bedömningen görs för deluppgift var för sig med "OK" eller "ej OK". För "OK" krävs att den är *acceptabelt* korrekt utförd. För G på inlämningsuppgiften i helhet måste alla deluppgifter vara bedömda "OK". K ges om en eller flera deluppgifter behöver kompletteras. U ges om en eller flera deluppgifter saknas eller om det fortfarande inte är korrekt efter komplettering.

En grundregel i lösandet av denna inlämningsuppgift är att ni inte får använda färdiga funktioner i moduler som går att importera, för att ev. försöka förenkla några beräkningar. Undantag ges för följande moduler som är nödvändiga för att lösa inläsning av csv-filer och plottning av diagram:

- modulen CSV (se kapitel 10 i Canvas) för att hantera csv-filer
- modulen *matplotlib.pyplot* (se kapitel 11 i Canvas) för plottning (behöver installeras)

Vi ställer inga krav på att "inmatningsfelkontroller" måste finnas, t.ex. på deluppgift 1 där användaren skulle kunna mata in ogiltiga menyval.

Några allmänna riktlinjer som skall följas:

- koden ska vara genomtänkt, lättläst och välkommenterad.
skriv varför kodrader finns, och inte vad de tekniskt gör. Se i kursmaterialet för diskussion och exempel på detta (kapitel 2 i Canvas *Introduktion och ett första programexempel*)
- variabler ska ha meningsfulla namn
- du får inte lämna in kod som genererar felmeddelande eller varningar
- du får inte lämna kvar, ej aktuell, bortkommenterad kod (gör koden svårare överskåda)

Programkoden skriver du i VS Code - använd svarsfilen "*Svarsfil_inl_lp3_24_A237TG.py*" som finns i Canvas. Denna .py-fil, med dina lösningar i, blir det enda du laddar upp vid inlämning i Canvas (dvs tillhörande .csv-filer behövs inte skickas in). Ändra inte namn på svarsfilen vid inlämning.

För att få en hanterbar och smidig rättning måste ovanstående regler och beskrivning av respektive uppgift följas noggrant. Avsteg innebär omedelbar retur.

Tänk på att innehållet i csv-filen kan importeras till Excel. Där kan ni sedan analysera data och kontrollera att lösningen på era uppgifter är korrekta. Gör gärna denna kontroll innan ni lämnar in inlämningsuppgiften så att ni slipper få retur på grund av felaktiga resultat i er lösning.

Uppgiftsbeskrivning

Vart tredje år genomförs PISA-undersökningen som visar skolelevs kunskapsutveckling i matematik. Denna undersökning brukar få ett stort medialt utrymme eftersom svenska skolelevs kunskapsutveckling över tid jämförs med motsvarande skolelevs kunskapsutveckling i andra länder. Den bifogade csv-filen *pisadata.csv* innehåller resultaten från dessa undersökningar från 2003 – 2018. Det finns flera länder som deltar i undersökningen än de som finns i filen, men vi har valt att enbart ta med de länder som rapporterat in resultat vid samtliga undersökningstillfällen under åren 2003 till 2018. Inlämningsuppgiften går ut på att analysera den data som finns i filen genom att lösa uppgifterna med hjälp av Python.

Nedanstående tabell visar några länders resultat i filen *pisadata.csv*. Resultaten är uppdelade per undersökningsår för män (M), för kvinnor (F) samt medelvärdet av resultatet för kvinnor och män (medel). Studera tabellen noga så att du är helt säker på hur den är uppbyggd. Det är helt nödvändigt för att ni ska kunna lösa uppgifterna.

	2018	2018	2015	2015	2012	2012	2009	2009	2006	2006	2003	2003	2018	2015	2012	2009	2006	2003
	M	F	M	F	M	F	M	F	M	F	M	F	medel	medel	medel	medel	medel	medel
Australia	494	488	497	491	510	498	519	509	527	513	527	522	491	494	504	514	520	524
Austria	505	492	510	483	517	494	506	486	517	494	509	502	499	497	506	496	505	506
Belgium	514	502	514	500	520	509	526	504	524	517	533	525	508	507	515	515	520	529
Brazil	388	379	385	370	397	380	394	379	380	361	365	348	384	377	389	386	370	356
Canada	514	510	520	511	523	513	533	521	534	520	541	530	512	516	518	527	527	532
Czech Republic	501	498	496	489	505	493	495	490	514	504	524	509	499	492	499	493	510	516
Denmark	511	507	516	506	507	493	511	495	518	508	523	506	509	511	500	503	513	514
Finland	504	510	507	515	517	520	542	539	554	543	548	541	507	511	519	541	548	544
France	499	492	496	490	499	491	505	489	499	492	515	507	495	493	495	497	496	511
Germany	503	496	514	498	520	507	520	505	513	494	508	499	500	506	514	513	504	503
Greece	452	451	454	454	457	449	473	459	462	457	455	436	451	454	453	466	459	445
Hong Kong (China)	548	554	549	547	568	553	561	547	555	540	552	548	551	548	561	555	547	550
Hungary	486	477	481	473	482	473	496	484	496	486	494	486	481	477	477	490	491	490
Iceland	490	500	487	489	490	496	508	505	503	508	508	523	495	488	493	507	506	515
Indonesia	374	383	385	387	377	373	371	372	399	382	362	358	379	386	375	371	391	360
Ireland	503	497	512	495	509	494	491	483	507	496	510	495	500	504	501	487	501	503
Italy	494	479	500	480	494	476	490	475	470	453	475	457	487	490	485	483	462	466
Japan	532	522	539	525	545	527	534	524	533	513	539	530	527	532	536	529	523	534
Korea	528	524	521	528	562	544	548	544	552	543	552	528	526	524	554	546	547	542
Latvia	500	493	481	483	489	493	483	481	489	484	485	482	496	482	491	482	486	483
Luxembourg	487	480	491	480	502	477	499	479	498	482	502	485	483	486	490	489	490	493
Macao (China)	560	556	540	548	540	537	531	520	530	520	538	517	558	544	538	525	525	527
Mexico	415	403	412	404	420	406	425	412	410	401	391	380	409	408	413	419	406	385
Netherlands	520	519	513	511	528	518	534	517	537	524	540	535	519	512	523	526	531	538
New Zealand	499	490	499	491	507	492	523	515	527	517	531	516	494	495	500	519	522	523

Uppgifter.

Nedan följer uppgifterna. Observera att ni måste skriva dina funktioner i en sådan ordning så att de är deklarerade innan de anropas. Alla deluppgifter kommer skrivas i form av funktioner så att du inte behöver (och inte ska) kopiera och klistra in tidigare skriven kod i ert huvudprogram. Ni skall skapa funktioner som ska återanvändas i senare deluppgifter, detta för att hålla nere mängden kod och göra det hela mera överskådligt. Det ska inte förekomma någon kod utanför funktionerna och huvudprogrammet.

När ni är färdig med deluppgiften bör ni lägga lite tid på att gå igenom ert program och förvissa er om att de ger korrekta resultat, samt dubbelkolla att koden inte bryter mot någon regel eller riktlinje som finns under *Förutsättningar*.

Innan inlämning bör du även spara din kod, stänga ner ditt program och sedan starta om VS Code och öppna programmet igen för att försäkra dig om att du verkligen kör rätt version av programmet.

När du är färdig med alla deluppgifter, d.v.s. hela inlämningsuppgiften, kommer den slutliga menyn få utseendet som nedan. Detta ger dig också en uppfattning om vad som kommer göras vid de kommande deluppgifterna.

Detta ska du göra:

1. Ladda ner och spara en er inlämningsfil med namnet "Svarsfil_inl_lp3_24_A237TG.py"
2. Placera all modulimport överst i filen.
3. Placera funktionerna från de olika deluppgifterna under resp. kommentarsrubrik i filen.
Placera dem i ordning så att funktioner som anropar andra ligger efter dessa. Ex:

- # Deluppgift 1
Funktioner från deluppgift 1 i ordning.
- # Deluppgift 2
Funktioner från deluppgift 2 i ordning.
- # Deluppgift 3
Funktioner från deluppgift 3 i ordning.
- # Deluppgift 4
Funktioner från deluppgift 4 i ordning.
- # Deluppgift 5
Funktioner från deluppgift 5 i ordning.
- # Huvudprogram med Meny från deluppgift 0. Använd nu dessa menyrubriker:

1. Läs in csv-filen.
2. Bästa resp. sämsta resultat år 2018.
3. Matematikkunskaper i nordnorden år 2003 - 2018.
4. Kontinuerligt förbättrat resp. försämrat år 2003 - 2018.
5. Kvinnor presterar bättre än män under åren 2003-2018.
6. Avsluta programmet.

Välj ett menyalternativ (1-6):

4. Se till så att koden från samtliga deluppgifter fungerar enligt anvisningarna och finns samlade på ett strukturerat sätt.
5. Kommentera varje funktionshuvud med en beskrivning av vad funktionen gör och vilka argument och returvärden funktionen har.
6. Kommentera övrig kod utifrån exemplet längst ner på sidan "Introduktion och ett första programexempel"
7. Testa programmet så att det inte ger några felmeddelanden annat än vid felaktiga inmatningar.

Uppgift 0:

Skapa skalet till huvudprogrammet under kommentaren # Huvudprogram med Meny i svarsfilen. Det innehåller nedanstående meny där användaren anger ett av menyalternativen (1 – 6). När menyalternativet är utfört ska användaren på nytt kunna ange ett menyalternativ osv. Först när användaren anger menyalternativ 6 avslutas programmet. Eftersom det inte finns några funktioner ännu ska menyalternativen bara skriva ut vilket alternativ som valts, förutom menyalternativ 6 som avslutar programmet och meddelar användaren om det.

Program för att läsa in och analysera data från PISA-undersökningen

1. Läs in csv-filen.
2. Bästa resp. sämsta resultat år 2018.
3. Matematikkunskaper i Norden år 2003 – 2018.
4. Kontinuerligt förbättrat resp. försämrat år 2003 – 2018.
5. Kvinnor presterar bättre än män under åren 2003–2018.
6. Avsluta programmet.

Välj ett menyalternativ (1 – 6):

Beskrivning av menyalternativen:

1. Frågar användaren efter namnet på datafilen. Därefter läses csv-filen in och listan ***pisadata*** skapas enligt uppgift 1. Vi förutsätter att csv-filen ligger i arbetskatalogen för VC-code.
2. Skapar tabeller över de tio länder med bästa resp. sämsta resultat 2018 enligt uppgift 2.
3. Skapar en tabell och ritat ett diagram över matematikkunskaperna i Norden enligt uppgift 3.
4. Skapar tabeller över de länder där resultaten kontinuerligt förbättrats resp. försämrats enligt uppgift 4.
5. Skapar en tabell på när kvinnor presterat bättre än män, enligt uppgift 5.
6. Programmet avslutas och en bekräftelse t.ex. "Programmet är avslutat" skrivs på skärmen.

Uppgift 1:

- a) Skriv en egendefinierad funktion `read_file` som tar ett filnamn (t.ex. data.csv) som argument och returnerar en lista. Funktionen skall öppna filen och därefter läsa in filens innehåll till en lista. Till sist skall funktionen returnera denna lista.

Observera: csv-filen använder semikolon (;) som avgränsningstecken. När du använder funktionen `open()` för att påbörja inläsningen av filerna, skicka med `encoding = 'UTF-8'` som argument för att säkerställa att inläsningen lyckas, och att teckentolkningen på bl.a. åäö blir korrekt (om inte så varierar det lite hur det beter sig, beroende på operativsystem).

- b) Uppdatera **Menyalternativ 1** (se menyn under Uppgift 0) så att programmet frågar användaren efter namnet på datafilen. Anges inget filnamn skall defaultvärdet användas. Ett bra sätt att ange defaultvärdet är för t.ex. **data.csv** är att skriva: `filename = input('Ange filnamn eller tryck bara Enter för data.csv') or 'data.csv'`
- c) Kontrollera att funktionen returnerar förväntad information genom att anropa funktionen ***read_file*** och spara den returnerade listan i en lista med namnet ***pisadata***. Därefter skall de fem första raderna i ***pisadata*** skrivas ut. Numeriska värdena kan vara antingen heltal eller text.

```
[['2018', '2018', '2015', '2015', '2012', '2012', '2009', '2009', '2006', '2006', '2003', '2003', '2018',  
'2015', '2012', '2009', '2006', '2003'], ['M', 'F', 'M', 'F', 'M', 'F', 'M', 'F', 'M', 'F', 'M', 'F', 'M', 'F', 'medel',  
'medel', 'medel', 'medel', 'medel'], ['Australia', 494, 488, 497, 491, 510, 498, 519, 509, 527,  
513, 527, 522, 491, 494, 504, 514, 520, 524], ['Austria', 505, 492, 510, 483, 517, 494, 506, 486, 517,  
494, 509, 502, 499, 497, 506, 496, 505, 506], ['Belgium', 514, 502, 514, 500, 520, 509, 526, 504, 524,  
517, 533, 525, 508, 507, 515, 515, 520, 529]]
```

Uppgift 2:

I denna uppgift ska ni ta reda på hur det gick för länderna år 2018.

- a) Skapa en funktion med valfritt namn som sorterar innehållet i en lista utifrån vald kolumn. Funktionen har två parametrar, listan som ska sorteras och indexet för den kolumn som bestämmer sorteringsordningen. Funktionen returnerar den sorterade listan efter att den skapat och skrivit ut två tabeller. En tabell som visar resultatet för de tio länder med bäst resultat och en tabell som visar resultatet för de tio länder med sämst resultat. Tabellerna ska ha följande utseende (observera att värdena i tabellen är exempelvärden och inte korrekta värden):

De tio länder som hade bäst resultat år 2018

Land	Resultat
Japan	507
Switzerland	503
Turkey	501
...	

De tio länder som hade sämst resultat år 2018

Land	Resultat
Russia	434
United Kingdom	456
Mexico	467
...	

Tips: Använd metoden `sorted()` för att sortera listan ***pisadata***. Spara gärna resultatet i en ny lista. Är ni osäkra på hur `sorted()` används läs t.ex. på <https://docs.python.org/3/library/index.html>. Det är en del i denna uppgift att inhämta information från externa källor för att lösa uppgiften.

Tips 2. Använd gärna f-string för att få till en snygg tabell här och i kommande uppgifter.

- b) Uppdatera **Menyalternativ 2** (se menyn under Uppgift 0) så att huvudprogrammet anropar funktionen med ***pisadata*** och indexet för kolumnen med medelvärden år 2018 som argument. Spara gärna den returnerade listan i variabeln ***pisadata_sorterad***. Huvudprogrammet behöver först hitta indexet för kolumnen med medelvärden år 2018.

Uppgift 3:

I denna deluppgift skall ni jämföra hur matematikkunskaperna i de nordiska länderna (Sverige, Norge, Danmark, Finland och Island) har utvecklats under åren 2003 – 2018 i jämförelse med medelvärdet per undersökningstillfälle för alla länder som finns i den tidigare inlästa listan ***pisadata***.

- a) Skapa funktionen ***kolumnmedel*** som beräknar medelvärdet på en kolumn i en lista. Funktionen har två parametrar, listan och indexet för kolumn för medelvärdesberäkningen. Funktionen returnerar medelvärdet. **OBS!** Ni måste själva skriva programkoden för medelvärdesberäkningen. Pythons inbyggda funktion `mean()` får alltså inte användas.
- b) Skapa funktionen ***arsmedel*** som har en lista som parameter och returnerar listan ***armedel***. Funktionen skapar listan ***armedel*** med medelvärdet/undersökningsår, 2003 till 2018 i samma ordning som kolumnerna i originallistan, för alla länder som finns i den. Använd kolumnerna märkta "medel" för dessa beräkningar. Funktionen anropar i sin tur funktionen ***kolumnmedel*** för varje år som skall beräknas.

- c) Skapa funktionen **nordTabell** med 2 el. 3 listor som parametrar. Funktionen saknar returvärde, Funktionen anropas från huvudprogrammet med listorna **pisadata** och **armedel** och skapar en tabell som jämför resultaten för de nordiska länderna med de i uppgift 3a beräknade medelvärdena. Använd kolumnen "medel" för de nordiska länderna. Du ska söka rätt på de nordiska länderna med hjälp av deras namn inte radnummer. Dessa kan hårdkodas i funktionen eller läggas till som en tredje parameter (gäller även 3d). Tabellen ska ha nedanstående utseende (observera att värdena i tabellen är exempelvärden och inte korrekta värden):

Kunskapsutveckling i matematik enligt PISA-undersökningen 2003 - 2018.

Länder:						
År	Sweden	Norway	Denmark	Finland	Iceland	Medelvärde alla länder
2018	507	504	506	511	509	479
2015	501	509	504	509	510	477
.						
.						
.						
2003	504	507	501	509	505	475

- d) Skapa funktionen **nordGraf** med 2 el. 3 listor som parametrar. Funktionen saknar returvärde, Funktionen skapa ett diagram som visar samma information som tabellen i uppgift 3c. Diagrammet ska ha en kurva per nordiskt land och undersökningsår samt en kurva för medelvärdena per undersökningsår. Diagrammet skall ha följande principiella utseende med rubriker, förklaringstexter m.m. (kurvorna visas inte här):



- e) Uppdatera **Menyalternativ 3** (se menyn under Uppgift 0) så att huvudprogrammet anropar funktionen **arsmedel** med **pisadata** som argument. Spara returvärdet i variabeln **medelvAr**

Anropa sen funktionen **nordTabell** med **pisadata**, **medelvAr** och ev. **lander** som argument.
Anropa sen funktionen **nordGraf** med **pisadata**, **medelvAr** och ev. **lander** som argument.

Uppgift 4:

Denna deluppgift går ut på att presentera trender i data och du skall skapa två tabeller.

Den första tabellen ska presentera de länder som kontinuerligt har förbättrat sina resultat mellan år 2003 till 2018. Använd kolumnerna märkta "medel" i listan **pisadata**. Tabellen ska ha följande utseende (observera att värdena i tabellen är exempelvärden och inte korrekta värden):

Länder som hela tiden har förbättrat sina resultat mellan 2003 - 2018

År och resultat:						
Land	2018	2015	2012	2009	2006	2003
Greece	490	487	479	474	471	469
Italy	487	481	477	472	469	465
.						
.						
.						

Den andra tabellen ska presentera de länder som kontinuerligt har försämrat sina resultat under samma tidsperiod. Tabellen för de länder där resultatet har försämrats ska ha samma utseende.

- Skapa en funktion **bättreSamre** med två parametrar en lista och en boolesk variabel. Funktionen har inget returvärde. Om den booleska variabeln är True skall tabellen med förbättrade resultat skapas annars den med försämrade.
- Uppdatera **Menyalternativ 4** (se menyn under Uppgift 0) så att huvudprogrammet: Anropar funktionen **bättreSamre** med **pisadata** och **True** som argument.

Anropar funktionen **bättreSamre** med **pisadata** och **False** som argument.

Uppgift 5:

Hur har kvinnorna klarat sig i förhållande till männen sett för alla länder, inte bara de nordiska?

- Skapa en funktion **kvinna_man** som har en lista som parameter och saknar returvärde. Funktionen ska finna länder och år för de mätningar där kvinnorna har presterat bättre resultat än männen. Tabellen ska ha nedanstående utseende (observera att värdena i tabellen är exempelvärden och inte korrekta värden):

År och länder när kvinnorna presterar bättre än männen under åren 2003-2018.			
År	Land	Kvinnor	Män
2018	Norway	507	503
	Germany	503	501
2015	Finland	514	512
2012			
.			
2003	France	509	506
	Japan	517	514

Tips: Skapa först en ny lista med alla de fall som uppfyller villkoret. Skriv sedan ut den nya tabellen. För att få till en utskrift som bara visar årtal första gången använd en if-sats och testa mot året.

- Uppdatera **Menyalternativ 5** (se menyn under Uppgift 0) så att huvudprogrammet: Anropar funktionen **kvinna_man** med **pisadata** argument.