

Inlämningsuppgift i kursen Grundläggande Programmering i Python, lp2-2022

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 uppgiften är att du, efter slutförd uppgift, har fått ytterligare ökad förståelse för programmeringens grunder och att du fått insikt om hur man bryter ned en programmeringsuppgift i mindre delar.

Förutsättningar

All administrativ information om datum för inlämning, hur kamratgranskningen går till, betyg & rättning, osv finner du på Uppgifts-sidan i Canvas.

Uppgiften utförs i form av 6 st. deluppgifter, där den sista uppgiften blir att sätta ihop programmen från de olika uppgifterna och skapa ett sammanhållet program via ett enkelt menysystem.

Varje deluppgift bedöms var för sig med "OK" eller "ej OK". För "OK" krävs att den är *acceptabelt* korrekt utförd, och för att inlämningsuppgiften i helhet ska bli godkänd vid lärarinlämningen i slutet (se flödesdiagrammet på Uppgifts-sidan i Canvas) måste alla deluppgifter vara bedömda "OK".

En grundregel i lösandet av denna inlämningsuppgift är att du inte får använda färdiga funktioner i moduler som går att importera, för att ev. försöka förenkla några beräkningar. Men med undantag för följande moduler som är helt nödvändiga:

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

Vi ställer inga krav på att "inmatningsfelkontroller" måste finnas, t.ex. på deluppgift 6 där användaren skulle kunna mata in ogiltiga menyval.

Några allmänna riktlinjer som skall följas:

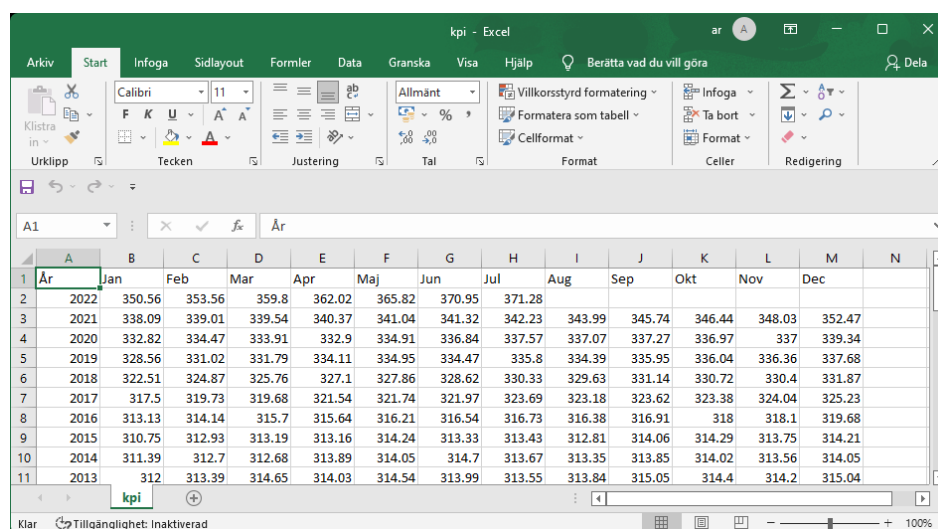
- koden ska vara genomtänkt, lättläst och välkommenterad
 - o skriv varför kodrader finns, och inte vad de tekniskt gör. Se i kursmaterialet för diskussion och exempel på detta (kapitel 2 i Canvas *Introduktion och ett första programexempel*)
- variabler ska ha meningsfulla namn
- du får inte lämna in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

Programkoden skriver du i Jupyter Notebook - använd den tillhandahållna förbereda svarsfilen "*Svarsfil_inluppg_1122_A237TG.ipynb*". Denna Jupyter Notebook-fil med dina lösningar 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. Vill du kommentera dina lösningar på deluppgifter på ett allmänt sätt (ex. "*jag tänkte så här ...*", "*jag valde att lösa problemet på detta sättet pga ...*" osv), så använd de förskapade *Markdown*-cellerna i svarsfilen till detta.

Uppgiftsbeskrivning

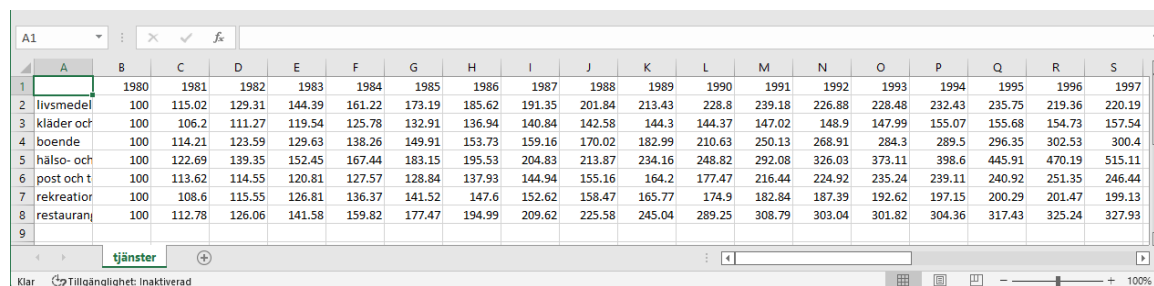
I denna uppgift ska du undersöka den allmänna prisutvecklingen för den privata konsumtionen genom att analysera KPI (Key Performance Indicators) under åren 1980-2022. Vi ska också mera i detalj analysera prisutvecklingen för ett antal varor och tjänster under ungefär samma tidsperiod. Vi börjar med att studera konsumentprisindex (KPI) som är mätetal vilka anger den privata konsumtionens kostnadsutveckling under en given tidsperiod. Statistiska Centralbyrån SCB är den myndighet som ansvarar för att uppdatera KPI och på deras hemsida (<https://www.scb.se>) kan du läsa mera om KPI. Den KPI-data som ska användas i denna inlämningsuppgift finns sparad i filen *kpi.csv*. Vi kommer även att analysera prisutvecklingen för några utvalda varor och tjänster baserat på den information som finns sparad i filerna *tjänster.csv* och *livsmedel.csv*. Nedanstående tre figurer visar en delmängd av innehållet i respektive fil.

kpi.csv:



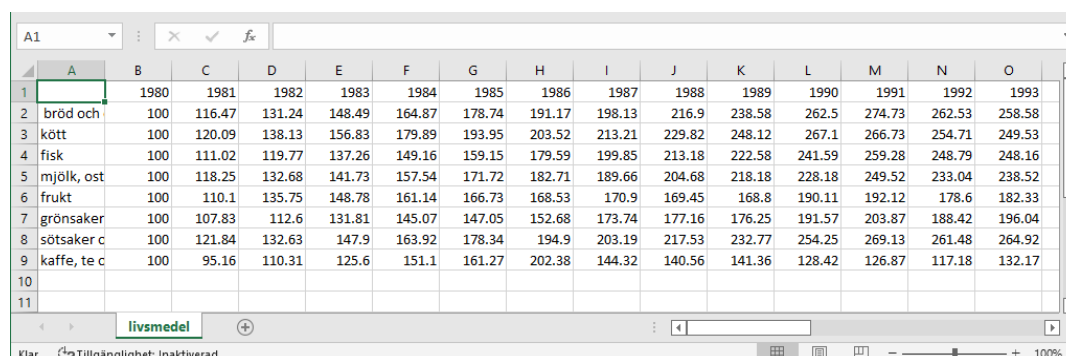
År	Jan	Feb	Mar	Apr	Maj	Jun	Jul	Aug	Sep	Okt	Nov	Dec
2022	350.56	353.56	359.8	362.02	365.82	370.95	371.28					
2021	338.09	339.01	339.54	340.37	341.04	341.32	342.23	343.99	345.74	346.44	348.03	352.47
2020	332.82	334.47	333.91	332.9	334.91	336.84	337.57	337.07	337.27	336.97	337	339.34
2019	328.56	331.02	331.79	334.11	334.95	334.47	335.8	334.39	335.95	336.04	336.36	337.68
2018	322.51	324.87	325.76	327.1	327.86	328.62	330.33	329.63	331.14	330.72	330.4	331.87
2017	317.5	319.73	319.68	321.54	321.74	321.97	323.69	323.18	323.62	323.38	324.04	325.23
2016	313.13	314.14	315.7	315.64	316.21	316.54	316.73	316.38	316.91	318	318.1	319.68
2015	310.75	312.93	313.19	313.16	314.24	313.33	313.43	312.81	314.06	314.29	313.75	314.21
2014	311.39	312.7	312.68	313.89	314.05	314.7	313.67	313.35	313.85	314.02	313.56	314.05
2013	312	313.39	314.65	314.03	314.54	313.99	313.55	313.84	315.05	314.4	314.2	315.04

tjänster.csv:



	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997
livsmedel	100	115.02	129.31	144.39	161.22	173.19	185.62	191.35	201.84	213.43	228.8	239.18	226.88	228.48	232.43	235.75	219.36	220.19
kläder och skor	100	106.2	111.27	119.54	125.78	132.91	136.94	140.84	142.58	144.3	144.37	147.02	148.9	147.99	155.07	155.68	154.73	157.54
boende	100	114.21	123.59	129.63	138.26	149.91	153.73	159.16	170.02	182.99	210.63	250.13	268.91	284.3	289.5	296.35	302.53	300.4
hälso- och sjukvård	100	122.69	139.35	152.45	167.44	183.15	195.53	204.83	213.87	234.16	248.82	292.08	326.03	373.11	398.6	445.91	470.19	515.11
post och telekommunikation	100	113.62	114.55	120.81	127.57	128.84	137.93	144.94	155.16	164.2	177.47	216.44	224.92	235.24	239.11	240.92	251.35	246.44
rekreation	100	108.6	115.55	126.81	136.37	141.52	147.6	152.62	158.47	165.77	174.9	182.84	187.39	192.62	197.15	200.29	201.47	199.13
restauranger	100	112.78	126.06	141.58	159.82	177.47	194.99	209.62	225.58	245.04	289.25	308.79	303.04	301.82	304.36	317.43	325.24	327.93

livsmedel.csv:



	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993
bröd och bakvaror	100	116.47	131.24	148.49	164.87	178.74	191.17	198.13	216.9	238.58	262.5	274.73	262.53	258.58
kött	100	120.09	138.13	156.83	179.89	193.95	203.52	213.21	229.82	248.12	267.1	266.73	254.71	249.53
fisk	100	111.02	119.77	137.26	149.16	159.15	179.59	199.85	213.18	222.58	241.59	259.28	248.79	248.16
mjölk, ost	100	118.25	132.68	141.73	157.54	171.72	182.71	189.66	204.68	218.18	228.18	249.52	233.04	238.52
frukt	100	110.1	135.75	148.78	161.14	166.73	168.53	170.9	169.45	168.8	190.11	192.12	178.6	182.33
grönsaker	100	107.83	112.6	131.81	145.07	147.05	152.68	173.74	177.16	176.25	191.57	203.87	188.42	196.04
sötsaker	100	121.84	132.63	147.9	163.92	178.34	194.9	203.19	217.53	232.77	254.25	269.13	261.48	264.92
kaffe, te och alkohol	100	95.16	110.31	125.6	151.1	161.27	202.38	144.32	140.56	141.36	128.42	126.87	117.18	132.17



Uppgifter

Nedan följer uppgifterna. Observera att i Jupyter Notebook så kan skrivna funktioner från en tidigare kodcell (i detta fallet en deluppgift) anropas i en senare kodcell (så länge den tidigare kodcellen har exekverats minst en gång). Det innebär att du inte behöver (och inte ska) kopiera & klistra in tidigare skrivna funktioner som ska återanvändas i senare deluppgifter, för att hålla nere mängden kod och göra det hela mera överskådligt.

När du är färdig bör du lägga lite tid på att gå igenom dina program och förvissa dig om att dessa ger korrekta resultat, samt dubbelkolla att koden inte bryter mot någon regel eller riktlinje som fanns under *Förutsättningar*.

Innan inlämning bör du även i Jupyter Notebook rensa alla variabler i minnet och testköra programmen en sista gång för säkerhets skull. Detta för att Jupyter Notebook kan "komma ihåg" gamla variabler från kod som är borttagen, vilket i vissa fall kan göra att ett program som fungerar vid provkörningen före rensning inte fungerar som tänkt nästa gång någon provkör det. Du rensar variablerna genom att välja alternativet "Restart & Clear Output" under menyalternativet "Kernel" (alternativt att starta om datorn och Jupyter Notebook, vilket garanterat ger samma effekt).

Uppgift 1:

- a) Skriv en egendefinierad funktion med namnet *read_file* som har ett angivet filnamn (t.ex. *data.csv*) som inargument. Funktionen ska kopiera innehållet i det angivna filnamnet till en lista. Funktionen ska sedan returnera denna lista.

Obs! csv-filerna 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) Kontrollera att funktionen returnerar kopior av de filer som ska användas i inlämningsuppgiften genom att anropa funktionen *read_file* för varje csv-fil. Spara de returnerade listorna i listor med namnen ***kpiData***, ***tjansteData*** och ***livsmedelData***. Skriv därefter ut några rader i respektive listor och jämför innehållet i motsvarande csv-filer:

kpiData

```
['År', 'Jan', 'Feb', 'Mar', 'Apr', 'Maj', 'Jun', 'Jul', 'Aug', 'Sep',  
'Okt', 'Nov', 'Dec']  
['2022', '350.56', '353.56', '359.8', '362.02', '365.82', '370.95',  
'371.28', '', '', '', '', '', '']
```

tjansteData

```
['', '1980', '1981', '1982', '1983', '1984', '1985', '1986', '1987',  
'1988', '1989', '1990', '1991', '1992', '1993', '1994', '1995', '1996',  
'1997', '1998', '1999', '2000', '2001', '2002', '2003', '2004', '2005',  
'2006', '2007', '2008', '2009', '2010', '2011', '2012', '2013', '2014',  
'2015', '2016', '2017', '2018', '2019', '2020', '2021']  
['livsmedel och alkoholfria drycker', '100.00', '115.02', '129.31',  
'144.39', '161.22', '173.19', '185.62', '191.35', '201.84', '213.43',  
'228.80', '239.18', '226.88', '228.48', '232.43', '235.75', '219.36',  
'220.19', '222.69', '225.97', '225.98', '232.47', '240.02', '240.79',  
'239.74', '238.09', '239.91', '244.77', '261.69', '269.25', '273.15',  
'276.65', '280.76', '286.96', '288.02', '294.73', '297.84', '304.22',  
'311.45', '320.19', '327.02', '328.36']
```

och livsmedelData

```
['', '1980', '1981', '1982', '1983', '1984', '1985', '1986', '1987',  
'1988', '1989', '1990', '1991', '1992', '1993', '1994', '1995', '1996',  
'1997', '1998', '1999', '2000', '2001', '2002', '2003', '2004', '2005',  
'2006', '2007', '2008', '2009', '2010', '2011', '2012', '2013', '2014',  
'2015', '2016', '2017', '2018', '2019', '2020', '2021']  
['bröd och övriga spannmålsprodukter', '100.00', '116.47', '131.24',  
'148.49', '164.87', '178.74', '191.17', '198.13', '216.90', '238.58',  
'262.50', '274.73', '262.53', '258.58', '251.93', '245.47', '229.23',  
'228.76', '230.67', '236.64', '240.13', '243.04', '249.16', '249.30',  
'249.54', '247.38', '245.60', '254.06', '279.80', '286.05', '288.29',  
'297.25', '300.04', '300.28', '300.44', '303.63', '304.96', '305.06',  
'308.29', '321.92', '325.85', '329.38']
```

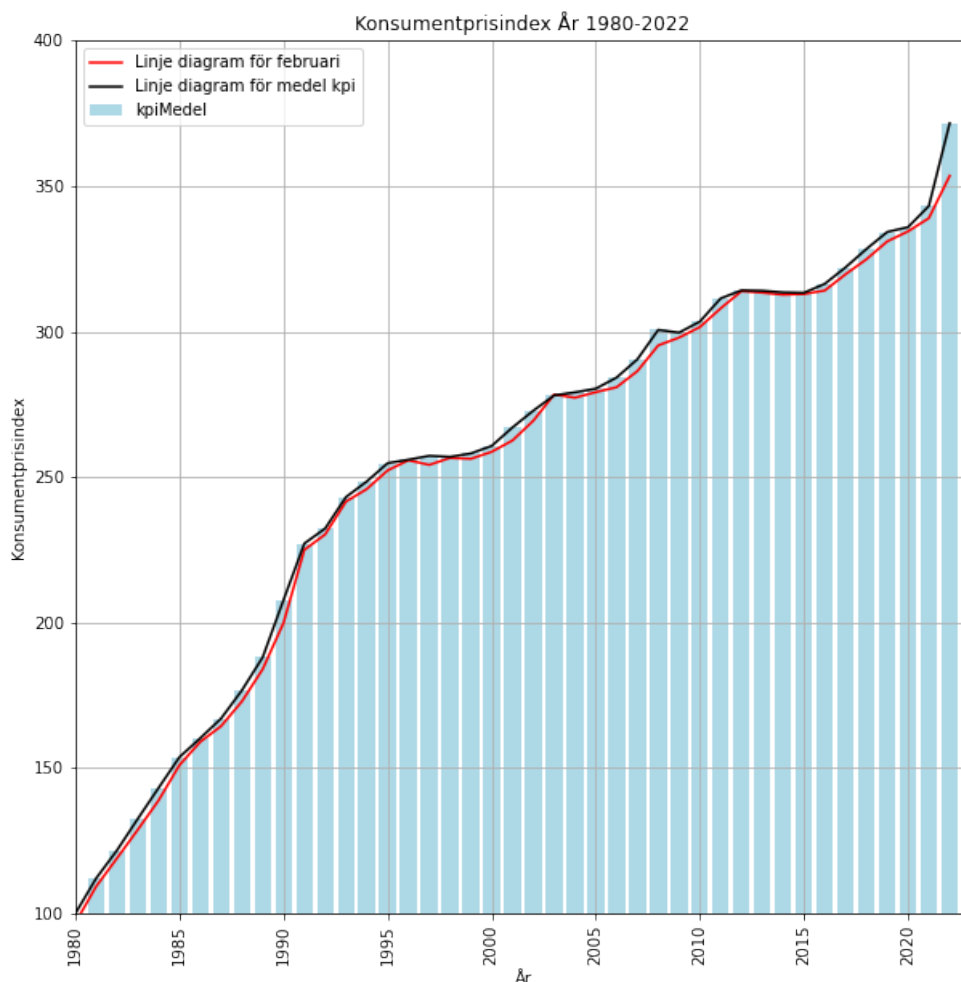
Uppgift 2:

I denna uppgift ska du använda listan **kpiData** för att skapa grafer och ett stapeldiagram enligt nedanstående figur. OBS! I denna uppgift måste du själv skriva programkoden för summering och medelvärdesberäkningen.

Programmet ska först fråga efter en månad (1 – 12). Därefter ska programmet beräkna medelvärdet av KPI för varje år under åren 1980 – 2022 (år 2022 baseras medelvärdet på månaderna jan – juli). Resultatet ska sedan ritas ut som både ett linjediagram (svart linje) och som ett stapeldiagram. Tillsammans med ett linjediagram (röd linje) för KPI för den angivna månaden under tidsperioden 1980 – 2022.

Exempelkörning: *(Kursiv text anger inmatning vid programkörning).*

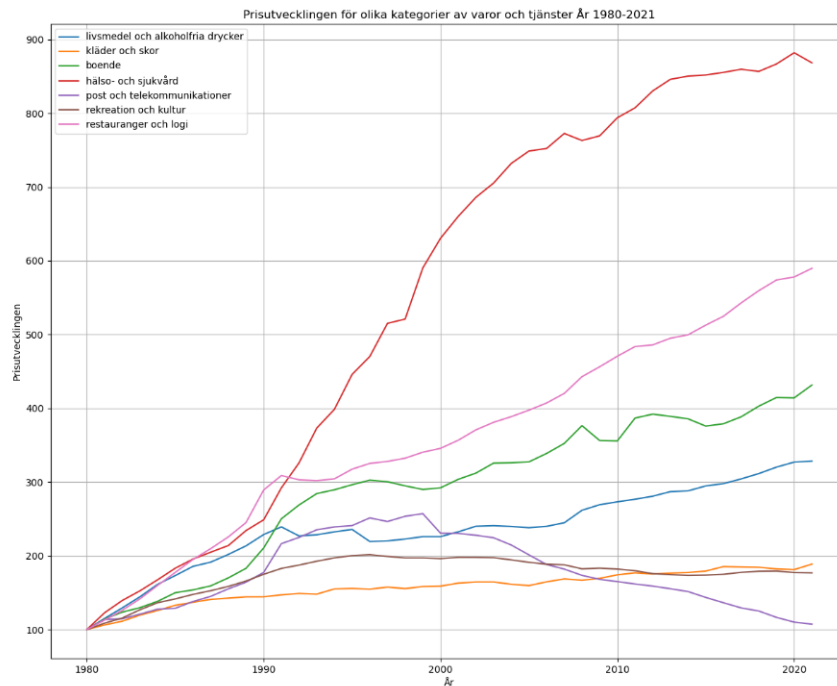
Ange vilken månad som ska presenteras: 2



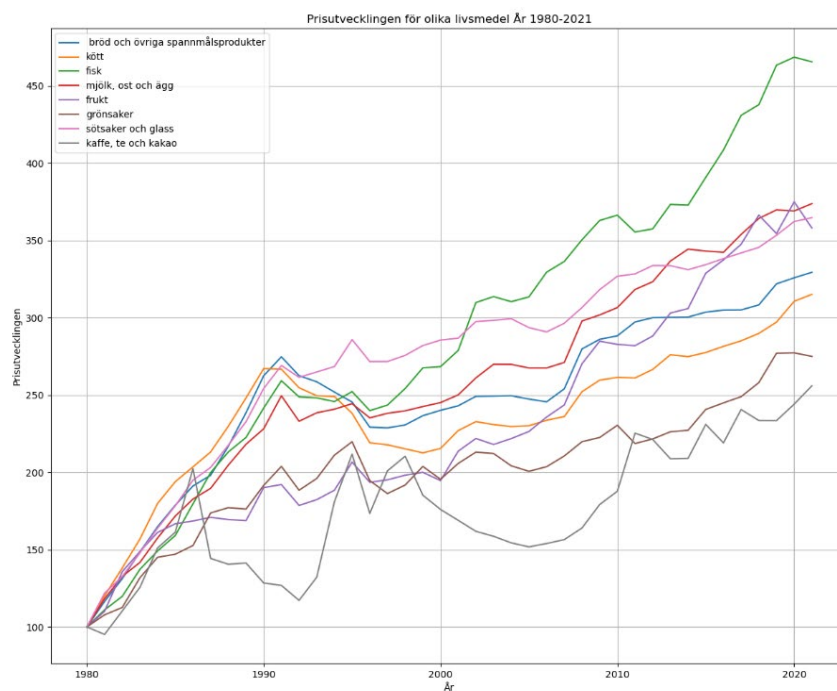
Uppgift 3:

I denna uppgift skall du plotta prisutvecklingen under åren 1980 – 2021 som baseras på innehållet i listorna **tjansteData** eller **livsmedelData**. Skapa därför en egendefinierad funktion med namnet **plotta_data** som har en lista som inargument. Funktionen ska skapa en av nedanstående diagram beroende på om inargumentet är listan **tjansteData** eller listan **livsmedelData**.

tjansteData:



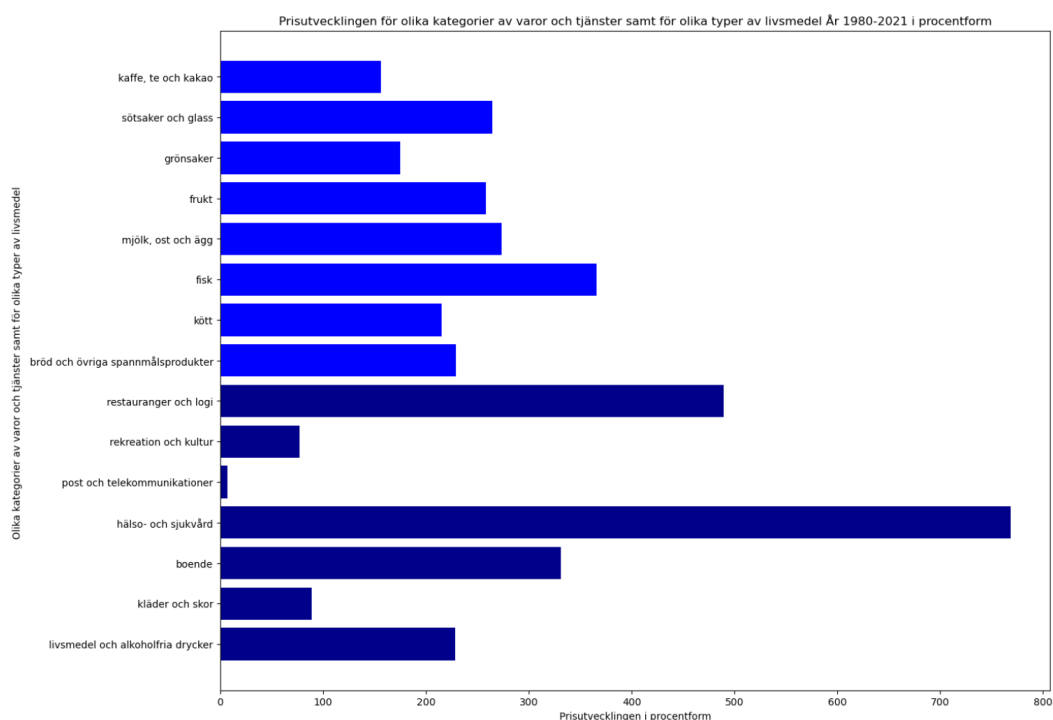
livsmedelData:



Uppgift 4:

I denna uppgift skall du skapa en egendefinerad funktion som beräknar prisutvecklingen i procentform för de olika kategorier av varor och tjänster samt för olika livsmedel år 1980 – 2021 och likt nedan skriver ut och ritat resultatet i en tabell och ett stapeldiagram. Tabellen och diagrammet ska baseras på de olika kategorier i de två listorna (*tjansteData* och *livsmedelData*).

Kategorier av varor och tjänster	Prisutvecklingen i procentform
livsmedel och alkoholfria drycker	228.36
kläder och skor	88.68
boende	331.43
hälso- och sjukvård	768.47
post och telekommunikationer	7.29
rekreation och kultur	76.84
restauranger och logi	489.68
Kategorier olika typer av livsmedel	Prisutvecklingen i procentform
bröd och övriga spannmålsprodukter	229.38
kött	215.13
fisk	365.55
mjölk, ost och ägg	273.80
frukt	258.05
grönsaker	175.00
sötsaker och glass	264.72
kaffe, te och kakao	155.98



Uppgift 5:

I deluppgift a ska du träna på att analysera data och skapa en tabell med hjälp av 'f-strängar' som innehåller resultatet av din analys (i Canvas under kapitel 4 'Strängar' finns exempel på hur du skapar en tabell med 'f-strängar'). I deluppgift b ska du presentera resultatet från deluppgift b i ett punktdiagram. OBS! I denna uppgift måste du själv skriva programkoden för summering och medelvärdesberäkningen.

- Skriv ett program som går igenom listan **kpiData** under åren 2000-2022 och beräknar och presenterar den månad varje år som har haft den största KPI-förändringen från en månad till nästakommande.

Exempel: KPI-förändringen för november definierar vi som $(KPI_{nov} - KPI_{okt})/KPI_{okt}$.

Både KPI-förändringen (uttryckt i procent) och månaden ska finnas med i tabellen. Tabellen ska även innehålla en kolumn med årsmedelvärdet av KPI. Tabellen avslutas med information som anger året och månaden för den största KPI-förändringen som uppmättes under mätperioden. Tabellen ska ha följande utseende **(OBS! värdena på denna uppgift är bara exempel och inte de korrekta värdena):**

ANALYS AV KPI UNDER ÅREN 2000 - 2022			

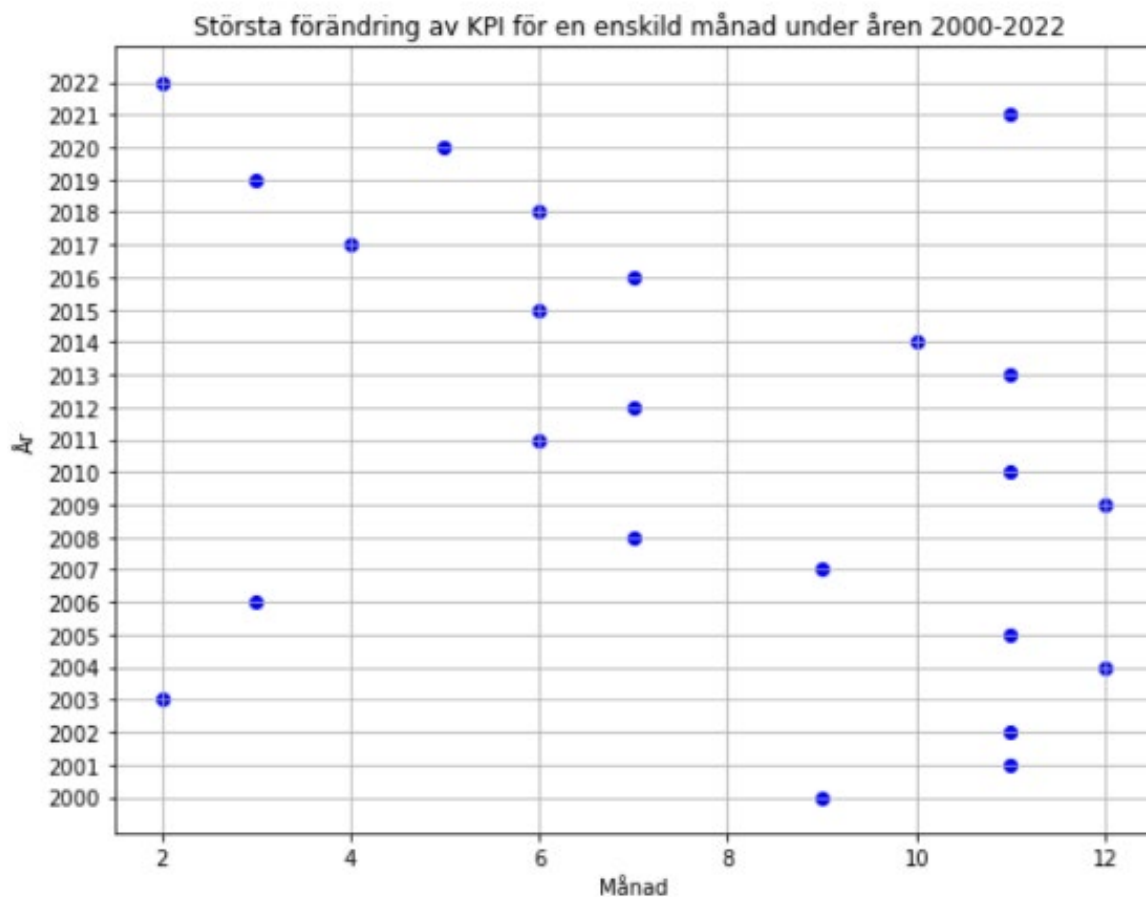
Största förändring			

År	%	månad	Årsmedelvärde

2000	0.71	Nov	262.23
2001	0.83	Mar	268.56
...
...
...
2022	-0.57	Feb	304.10

Största förändring:	1.94	Dec 2008	
=====			

- b. Att hitta mönster (trender) i datamängder är ofta av stort intresse. Mönster upptäcker vi människor bättre genom att studera bildinformation och inte information i tabellform. Låt oss därför utreda om det finns någon trend som visar när på året de största förändringarna i KPI brukar inträffa genom att skapa ett enkelt punktdiagram enligt nedan. Vi sätter en 'o' (fylld cirkel) för den månad i respektive år där den största förändringen skett (dvs en grafisk presentation av informationen som finns i tabellen i deluppgift a). Punktdiagrammet ska ha följande utseende **(OBS! värdena på denna uppgift är bara exempel och inte de korrekta värdena)**:



Uppgift 6:

Skapa ett program som 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.

Program för att läsa in och analysera resultatet i uppgift 1 – 5

1. Läser in csv-filerna
2. Konsumentprisindex under åren 1980 – 2022
3. Prisutvecklingen för de olika kategorierna i filerna "Varor och tjänster" samt "Livsmedel" under åren 1980 – 2021
4. Prisutvecklingen i procentform för de olika kategorierna i filerna "Varor och tjänster" samt "Livsmedel" under åren 1980 – 2021
5. Förändringar i KPI under åren 2000 – 2022
6. Avsluta programmet

Välj menyalternativ (1-6):

Funktionsbeskrivning av menyalternativen:

1. Programmet frågar användaren efter namnen på de tre aktuella datafilerna. Anges inget namn utan man bara trycker på Enter så används standardnamnen. Vi förutsätter att csv-filerna ligger i samma mapp som Jupyter-filen.

Tips: använd input-varianten i stil med detta:

```
input("namn för kpi.csv: ") or "kpi.csv"
```

...för att lätt uppnå denna funktionalitet

Sedan läses de tre csv-filerna in och programmet skapar listorna **kpiData**, **tjansteData** och **livsmedelData** enligt uppgift 1. Skriv ut några rader från respektive listor för att bekräfta för användaren att filerna lästes in korrekt.

2. Ritar KPI under åren 1980 – 2022 enligt uppgift 2.
3. Prisutvecklingen för de olika kategorierna i filerna "Varor och tjänster" samt "Livsmedel" under åren 1980 – 2021 enligt uppgift 3.
4. Prisutvecklingen i procentform för de olika kategorierna i filerna "Varor och tjänster" samt "Livsmedel" under åren 1980 – 2021 enligt uppgift 4.
5. Tabell och punktdiagram som skapas i uppgift 5 skrivs ut.
6. Programmet avslutas och "Programmet är avslutat" eller en liknande bekräftelse skrivs ut på skärmen.