

Data cleaning and filtering

Baggrund

På Industriens Uddannelsers hjemmeside <https://iu.dk/>, kan man hente et regnark med alle fag og målepinde for *Datatekniker med speciale i programmerings* uddannelsen.

Gå til <https://hentdata.stil.dk/uddannelser>

Vælg (1205) Data- og kommunikationsuddannelsen + alle kolonner + "Eksporter inkl målepinde"

Uddannelsesordning

Uddannelse

(1205) Data- og kommunikationsuddannelsen

Drafttrædelsesdato

01-08-2023

+ Ændringer i forhold til tidligere

+ Forslag til skole- og oplæringsperioder

+ Varighed af valgfri uddannelsesspecifikke specialfag

+ Varighed af valgfag

+ Allokering for voksne

+ Link til censurvejledning/skuenestervejledning

+ Fag med videreuddannelsespædsidv

+ Yderligere oplysninger

Uddannelsesordning for (1205) Data- og kommunikationsuddannelsen:

Uddannelsesordning for data- og kommunikationsuddannelsen

Hovedforløb:

1205 Data- og kommunikationsuddannelsen

Bekræftigelses:

Bekræftigelse om data- og kommunikationsuddannelsen (01-08-2023)

Drafttrædelsesdato:

01-08-2023

Fag på uddannelsen

Speciale

Datatekniker med speciale i itteknologi

Eksporter

Datatekniker med speciale i programmering

Kolonner

Resultatform

Fagkategori

Tilknytning

Standard

Slutdato

Bilag til kompetencemål

Se vejledning

Elevertype

Ungdom og EUV3

EUV2

EUV1

EUV & EUV3

EUV & EUV2

EUV & EUV1

(eller hent regnearket i Moodle)

Desværre er formatet ikke helt lige til. Denne opgave går ud på at indlæse regnearket i en Python data struktur, og derefter filtre data, så vi kan finde præcist det som vi søger.

Uddannelsesspecifikke fag									
NUMMER	TITEL	MÅLPINDE	NIVEAU	RESULTATFORM	FAGKATEGORI	TILKNYTNING	STARTDATO	SLUTDATO	BIDRAGET TIL KOMPETENCEN
1595	Netværksd 1. Lærlingen kan...	4	-/STA/7TRIN	Uddannelsesspecifikke fa Valgri, valgfrit niveau			01-08-2023		11, 23
15938	Game-desi 1. Lærlingen kan b2		-/STA/7TRIN	Uddannelsesspecifikke fa Valgri			01-08-2023		24, 33
16475	GUI-progra 1. Lærlingen kan f3		DEL-/7TRIN	Uddannelsesspecifikke fa Bundet, valgfrit niveau			01-08-2023		24, 28
16475	GUI-progra 1. Lærlingen kan f5		-/STA/7TRIN	Uddannelsesspecifikke fa Bundet, valgfrit niveau			01-08-2023		24, 28
16478	App progra 1. Lærlingen kan f5		-/STA/7TRIN	Uddannelsesspecifikke fa Bundet, valgfrit niveau			01-08-2023		28
16867	Netværksd 1. Lærlingen kan r3		-/STA/7TRIN	Uddannelsesspecifikke fa Valgri			01-08-2023		9, 17, 22
16495	Programud 1. Lærlingen kan f-		PE/-/-	Oplæringsmål	Bundet		01-08-2023		26
10842	Kemi, eux 1. Eleven eller lær C		DEL-/7TRIN	Grundfag	Valgri		01-08-2023		
10842	Kemi, eux 1. Eleven eller lær C		DEL-/7TRIN	Grundfag	Valgri		01-08-2023		
10842	Kemi, eux 1. Eleven eller lær C		-/STA/7TRIN	Grundfag	Valgri		01-08-2023		
6678	Informatik 1. Informatik B		MDT/EKS/7TRIN	hhx	Bundet		01-08-2023		
6678	Informatik 1. Informatik B		MDT/GSI/7TRIN	hhx	Bundet		01-08-2023		
6678	Informatik 1. Informatik B		MDT/STA/7TRIN	hhx	Bundet		01-08-2023		
6277	Projektsstyr 1. Lærlingen kan f4		-/STA/7TRIN	Uddannelsesspecifikke fa Bundet, valgfrit niveau			01-08-2023		51
6279	Afsluttende 1. Lærlingen kan f5		-/STA/7TRIN	Uddannelsesspecifikke fa Bundet			01-08-2023		1-10, 23-34
9446	LINQ og En 1. Lærlingen kan r3		-/STA/7TRIN	Uddannelsesspecifikke fa Valgri, valgfrit niveau			01-08-2023		10
9446	LINQ og En 1. Lærlingen kan f4		-/STA/7TRIN	Uddannelsesspecifikke fa Valgri, valgfrit niveau			01-08-2023		10
12250	Komp-mål 1. Lærlingen kan f-		PE/-/-	Kompetencemål	Bundet		01-08-2023		
12885	Sikkerhed - 1. Lærlingen kan i-		PE/-/-	Oplæringsmål	Bundet		01-08-2023		32, 33
12889	Programud 1. Lærlingen kan f-		PE/-/-	Oplæringsmål	Bundet		01-08-2023		23, 24, 26-28, 30, 31

Kolonne nr. tre: Målpinde indeholder en liste af fagets målpinde i et sjovt format med linie-skift.

Vi vil gerne have den som en liste vi kan slå op i.

Uddannelsesspecifikke fag									
NUMMER	TITEL	MÅLPINDE	NIVEAU	RESULTATFORM	FAGKATEGORI	TILKNYTNING	STARTDATO	SLUTDATO	BIDRAGET TIL KOMPETENCEN
1595	Netværksd 1. Lærlingen kan...	4	-/STA/7TRIN	Uddannelsesspecifikke fa Valgri, valgfrit niveau			01-08-2023		11, 23
15938	Game-desi 1. Lærlingen kan b2		-/STA/7TRIN	Uddannelsesspecifikke fa Valgri			01-08-2023		24, 33
16475	GUI-progra 1. Lærlingen kan f3		DEL-/7TRIN	Uddannelsesspecifikke fa Bundet, valgfrit niveau			01-08-2023		24, 28
16475	GUI-progra 1. Lærlingen kan f5		-/STA/7TRIN	Uddannelsesspecifikke fa Bundet, valgfrit niveau			01-08-2023		24, 28
16478	App progra 1. Lærlingen kan f5		-/STA/7TRIN	Uddannelsesspecifikke fa Bundet, valgfrit niveau			01-08-2023		28
16867	Netværksd 1. Lærlingen kan r3		-/STA/7TRIN	Uddannelsesspecifikke fa Valgri			01-08-2023		9, 17, 22
16495	Programud 1. Lærlingen kan f-		PE/-/-	Oplæringsmål	Bundet		01-08-2023		26
10842	Kemi, eux 1. Eleven eller lær C		DEL-/7TRIN	Grundfag	Valgri		01-08-2023		
10842	Kemi, eux 1. Eleven eller lær C		DEL-/7TRIN	Grundfag	Valgri		01-08-2023		
10842	Kemi, eux 1. Eleven eller lær C		-/STA/7TRIN	Grundfag	Valgri		01-08-2023		
6678	Informatik 1. Informatik B		MDT/EKS/7TRIN	hhx	Bundet		01-08-2023		
6678	Informatik 1. Informatik B		MDT/GSI/7TRIN	hhx	Bundet		01-08-2023		
6678	Informatik 1. Informatik B		MDT/STA/7TRIN	hhx	Bundet		01-08-2023		
6277	Projektsstyr 1. Lærlingen kan f4		-/STA/7TRIN	Uddannelsesspecifikke fa Bundet, valgfrit niveau			01-08-2023		51
6279	Afsluttende 1. Lærlingen kan f5		-/STA/7TRIN	Uddannelsesspecifikke fa Bundet			01-08-2023		1-10, 23-34
9446	LINQ og En 1. Lærlingen kan r3		-/STA/7TRIN	Uddannelsesspecifikke fa Valgri, valgfrit niveau			01-08-2023		10
9446	LINQ og En 1. Lærlingen kan f4		-/STA/7TRIN	Uddannelsesspecifikke fa Valgri, valgfrit niveau			01-08-2023		10
12250	Komp-mål 1. Lærlingen kan f-		PE/-/-	Kompetencemål	Bundet		01-08-2023		
12885	Sikkerhed - 1. Lærlingen kan i-		PE/-/-	Oplæringsmål	Bundet		01-08-2023		32, 33
12889	Programud 1. Lærlingen kan f-		PE/-/-	Oplæringsmål	Bundet		01-08-2023		23, 24, 26-28, 30, 31

Udfordringer

- Regnarket indholder 4 faneblade: ét til fælles fag, ét til programmering, ét til infrastruktur og ét til IT-supporter
- Arket indeholder (af én eller anden grund) fag med kategorier som Grundfag og htx. Disse vil vi gerne filtrere væk.
- Nogle fag findes flere gange. Bla. én gang for hvert niveau (2-4). Vi ønsker blot at kigge på højeste niveau, som vi finde for hvert fag
- Nogle fag findes endda flere gang med samme niveau, men med forskellige Resultatformer. Vi ønsker at finde dag med resultatformen "-/STA/7TRIN"

- Målepindene står i ét felt, formateret med linie skift.

Se næste side

Opgaven

- Indlæs regnearket i en Pandas datastruktur
- Læs en Python funktion som tager et fag-nummer som input og returnerer en (Python) liste

```
def hent_maalepinde(fagnr : int) -> list[str]

    # your magic code here

m1 = hent_maalepinde(17348)
sw_test = hent_maalepinde(16484)

for m in m1:
    print(m)

for m in sw_test:
    print(m)
```

1. Lærningen kan redegøre for, hvordan en maskine kan lære.
2. Lærningen kan anvende maskinlæring til at løse praktiske opgaver.
3. Lærningen kan samle og forberede data til maskinlæring.
4. Lærningen kan redegøre for forskellige maskinlæringsmetoder.
5. Lærningen kan træne og teste sin model.
6. Lærningen kan anvende klassificering og regression til at løse praktiske opgaver.
7. Lærningen kan anvende et maskinlærings-framework til at løse en praktisk opgave.

1. Lærningen kan, ud fra en case-opgave, omhandlende krav til en nyudviklet applikation, selvstændigt planlægge, gennemføre og dokumentere en softwaretest med fokus på funktionalitet, performance og sikkerhed, og kan herigennem demonstrere viden, færdigheder og kompetencer ud over de i faget beskrevne mål og målniveauer.
2. Lærningen har en generel viden om metoder og løsninger til at fejl- og funktionsteste en udviklet softwareløsning, samt om løsninger, der kan forbedre sikkerheden på en web-enabled applikation.
3. Lærningen kan tilrettelægge en testplan, som indeholder test til funktionelle krav.
4. Lærningen kan anvende forskellige relevante testmetodikker.
5. Lærningen kan vurdere, hvornår en software-release er tilstrækkeligt testet.
6. Lærningen kan gennemføre en scriptet test på en app- eller webplatform.
7. Lærningen kan integrere automatiserede tests i software udviklingscyklussen.
8. Lærningen kan gennemføre en tryktest, hvor performance og krav til hardware medtages.
9. Lærningen kan aflæse performancetabeller og identificere flaskehalse i et kompliceret system.
10. Lærningen kan redegøre for UX-tests (User Experience), som evaluerer kundeoplevelsen.
11. Lærningen har en generel viden om data-kryptering, og kan anvende krypteringsformer som f.eks. SSL i forbindelse med udvikling af webbaserede applikationer.
12. Lærningen kan indhente viden omkring softwaresikkerhedsstandarder og kendte sikkerhedshuller i eksisterende software.
13. Lærningen har en generel viden om authentication og kan udvikle programmer, der anvender authentication.
14. Lærningen har en generel viden om Hashing-kodning, og kan anvende kodningsformer som f.eks. MD5 i forbindelse med udvikling af webbaserede applikationer.
15. Lærningen har en generel viden om forskellige former for hacking angreb, som f.eks. SQL injection og XSS (Cross-Site Scripting), og kan anvende denne viden til at sikre egne udviklede webbaserede applikationer.
16. Lærningen kan anvende forskellige website security tools til at teste sikkerheden på udviklet software.

Hints

Python pakken Pandas kan bruges til at indlæs CSV filer og Excel filer.

```
excel_file = 'Fagtabel_Excel_2023-09-28T13 31 53.165576.xlsx'

df = pd.read_excel(excel_file, sheet_name=1)
```

Brug Panda tutorials til at lære at organisere og filtrere data, så vi kan lave den ønskede funktion!