



UNIVERSITETET I BERGEN

MER CSV

INF100

VÅR 2024

Torstein Strømme

DATASTRUKTURER FOR CSV

- En CSV-fil: ren tekst som representerer tabell-data

```
personnummer,medisin,startdato  
01010111111,vitaminbjørner,2022-08-22  
22020222222,fluortabletter,2022-01-11  
01010111111,tran,2022-06-01
```



	A	B	C
1	personnummer	medisin	startdato
2	1010111111	vitaminbjørner	22/08/2022
3	2202022222	fluortabletter	11/01/2022
4	1010111111	tran	01/06/2022

DATASTRUKTURER FOR CSV

- En CSV-fil: ren tekst som representerer tabell-data

```
personnummer,medisin,startdato
01010111111,vitaminbjørner,2022-08-22
22020222222,fluortabletter,2022-01-11
01010111111,tran,2022-06-01
```

- I Python: tabell-data representert som streng

```
table_string = (  
    'personnummer,medisin,startdato\n'  
    '01010111111,vitaminbjørner,2022-08-22\n'  
    '22020222222,fluortabletter,2022-01-11\n'  
    '01010111111,tran,2022-06-01\n'  
)
```

DATASTRUKTURER FOR CSV

- En CSV-fil: ren tekst som representerer tabell-data

```
personnummer,medisin,startdato  
01010111111,vitaminbjørner,2022-08-22  
22020222222,fluortabletter,2022-01-11  
01010111111,tran,2022-06-01
```

- I Python: tabell-data representert som liste av strenger

```
table_string_list = [  
    'personnummer,medisin,startdato',  
    '01010111111,vitaminbjørner,2022-08-22',  
    '22020222222,fluortabletter,2022-01-11',  
    '01010111111,tran,2022-06-01',  
]
```

DATASTRUKTURER FOR CSV

- En CSV-fil: ren tekst som representerer tabell-data

```
personnummer,medisin,startdato
01010111111,vitaminbjørner,2022-08-22
22020222222,fluortabletter,2022-01-11
01010111111,tran,2022-06-01
```

- I Python: tabell-data representert som 2D-liste

```
table_2d_list = [
    ['personnummer', 'medisin', 'startdato'],
    ['01010111111', 'vitaminbjørner', '2022-08-22'],
    ['22020222222', 'fluortabletter', '2022-01-11'],
    ['01010111111', 'tran', '2022-06-01'],
]
```

DATASTRUKTURER FOR CSV

- En CSV-fil: ren tekst som representerer tabell-data

```
personnummer,medisin,startdato  
01010111111,vitaminbjørner,2022-08-22  
22020222222,fluortabletter,2022-01-11  
01010111111,tran,2022-06-01
```

- I Python: tabell-data representert som liste av oppslagsverk

```
table_dict_list = [  
    {'personnummer': '01010111111', 'medisin': 'vitaminbjørner', 'startdato': '2022-08-22'},  
    {'personnummer': '22020222222', 'medisin': 'fluortabletter', 'startdato': '2022-01-11'},  
    {'personnummer': '01010111111', 'medisin': 'tran', 'startdato': '2022-06-01'}  
]
```

DATASTRUKTURER FOR CSV

- Liste av lister: hvilke medisiner er registrert på person 01010111111?

```
table_2d_list = [  
    ['personnummer', 'medisin', 'startdato'],  
    ['01010111111', 'vitaminbjørner', '2022-08-22'],  
    ['22020222222', 'fluortabletter', '2022-01-11'],  
    ['01010111111', 'tran', '2022-06-01'],  
]
```

```
for row in table_2d_list:  
    if row[0] == '01010111111':  
        print(row[1])
```

magiske tall (fy!)

DATASTRUKTURER FOR CSV

- Liste av oppslagsverk: hvilke medisiner er registrert på person 01010111111?

```
table_dict_list = [  
    {'personnummer': '01010111111', 'medisin': 'vitaminbjørner', 'startdato': '2022-08-22'},  
    {'personnummer': '22020222222', 'medisin': 'fluortabletter', 'startdato': '2022-01-11'},  
    {'personnummer': '01010111111', 'medisin': 'tran', 'startdato': '2022-06-01'}  
]
```

```
for row in table_dict_list:  
    if row['personnummer'] == '01010111111':  
        print(row['medisin'])
```



FRA CSV TIL LISTE AV OPPSLAGSVERK

- Fra CSV til streng

```
path = 'org.csv'  
with open(path, 'r', encoding='utf-8') as file:  
    file_content_as_string = file.read()
```

FRA CSV TIL LISTE AV OPPSLAGSVERK

- Fra CSV til streng
- **Fra streng til liste av linjer**
 - Øverste linje er header
 - Øvrige linjer er data

```
lines = file_content_as_string.splitlines()  
header_line = lines[0]  
data_lines = lines[1:]
```

FRA CSV TIL LISTE AV OPPSLAGSVERK

- Fra CSV til streng
- Fra streng til liste av linjer
 - Øverste linje er header
 - Øvrige linjer er data
- **Fra linje til liste av «celler»**

```
header_cells = header_line.split(',')
```

```
for line in data_lines:  
    cells = line.split(',')
```

FRA CSV TIL LISTE AV OPPSLAGSVERK

- Fra CSV til streng
- Fra streng til liste av linjer
 - Øverste linje er header
 - Øvrige linjer er data
- Fra linje til liste av «celler»
- **Fra liste av celler til oppslagsverk**

```
# header_cells = ['personnummer', 'medisin', ...  
# cells = ['1010111111', 'tran', ...  
# number_of_cols = 3
```

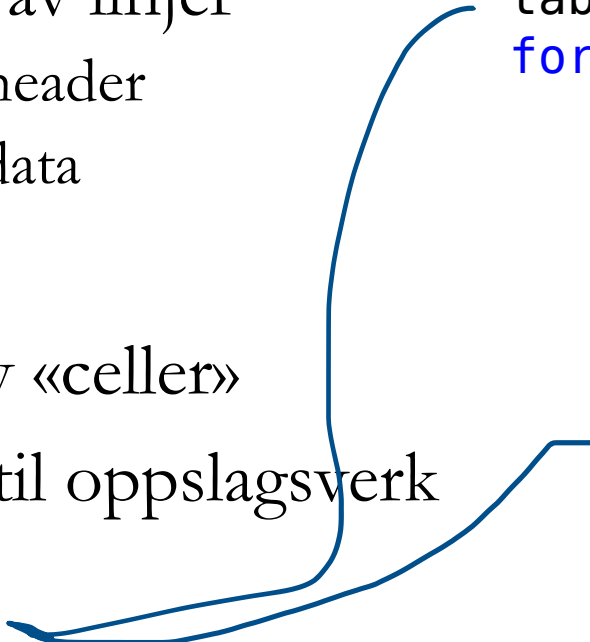
```
row_dict = {}  
for i in range(number_of_cols):  
    header = header_cells[i]  
    value = cells[i]  
    row_dict[header] = value
```

```
# row_dict = {  
#     'personnummer': '1010111111',  
#     'medisin': 'tran',  
#     ...  
# }
```

FRA CSV TIL LISTE AV OPPSLAGSVERK

- Fra CSV til streng
- Fra streng til liste av linjer
 - Øverste linje er header
 - Øvrige linjer er data
- Fra linje til liste av «celler»
- Fra liste av celler til oppslagsverk
- **Legg til i liste**

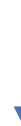
```
header_cells = header_line.split(',')
number_of_cols = len(header_cells)
table_dict = []
for line in data_lines:
    cells = line.split(',')
    row_dict = {}
    for i in range(number_of_cols):
        header = header_cells[i]
        value = cells[i]
        row_dict[header] = value
    table_dict.append(row_dict)
```



CASE: REFORMATTERING AV CSV

- Du har en CSV-fil med data, men det er på feil format

```
personnummer,medisin,startdato,sluttdato
01010111111,vitaminbjørner,2022-08-22,2022-11-28
22020222222,fluortabletter,2022-01-11,2022-03-03
01010111111,tran,2022-06-01,2022-06-30
```



- Produser en CSV-fil hvor data er på ønsket format

```
personnummer,medisin,dato,endring
01010111111,vitaminbjørner,2022-08-22,1
01010111111,vitaminbjørner,2022-11-28,0
22020222222,fluortabletter,2022-01-11,1
22020222222,fluortabletter,2022-03-03,0
01010111111,tran,2022-06-01,1
01010111111,tran,2022-06-30,0
```



UNIVERSITETET I BERGEN

PANDEMISIMULERING

INF100

VÅR 2024

Torstein Strømme



UNIVERSITETET I BERGEN

ÅPEN LAB

INF100

VÅR 2024

Torstein Strømme

ÅPEN LAB

- Gjør (nesten) akkurat hva du vil
- Bestått tilsvarende bestått på lab 7 og lab 8
- Krav:
 - Må benytte *uib_inf100_graphics.event_app* som rammeverk
 - Må være interaktivt (reagerer på mus eller tastatur)
 - Må ha en viss kompleksitet. Spør hvis du er usikker.