

```
In [1]: import pandas as pd
import seaborn as sns
from matplotlib import pyplot as plt
```

Globale variabelen: zelf instellen

```
In [2]: Woo_type= '2i'
Bestuursorgaan= 'gm0268' # 'gm1963' #
Jaar='2023'

DOI= '.'.join(['nl', Bestuursorgaan, Woo_type, Jaar])
URL = 'https://doi.wooverheid.nl/?doi='+DOI+ '&csv=true'

print(URL)
```

https://doi.wooverheid.nl/?doi=nl.gm0268.2i.2023&csv=true

Haal data op en maak spreadsheet/dataframe

- Onderstaand spreadsheet toont dezelfde informatie als op <https://doi.wooverheid.nl/?doi=nl.gm0268.2i.2023>, maar dan als Excel bestand.
- De analyses en plaatjes die hieronder staan kunnen natuurlijk ook in Excel gemaakt worden.
- Wij doen het zo omdat we dat net wat makkelijker vinden, en nu is het herhaalbaar voor elk bestuursorgaan met die notebook.

```
In [3]: df= pd.read_csv(URL, index_col='dc_identifier')
parse_dates={'foi_publishedDate', 'dc_date_year', 'foi_requestDate',
              'foi_decisionDate', 'foi_retrievedDate'}
dates_in_csv= list(set(df.columns) & parse_dates)
df= pd.read_csv(URL, index_col='dc_identifier', parse_dates= dates_in_csv )

print(df.shape)
df.head(2)
```

(54, 21)

```
Out[3]:
```

	dc_title	dc_type	dc_description	dc_publisher	dc_creator	tooiwl_topic	foi_r
	dc_identifier						

nl.gm0268.2i.2022.5	Algoritmes	2i	Woo verzoek gedaan op 14 december 2022 over al...	gm0268	gm0268		NaN
---------------------	------------	----	--	--------	--------	--	-----

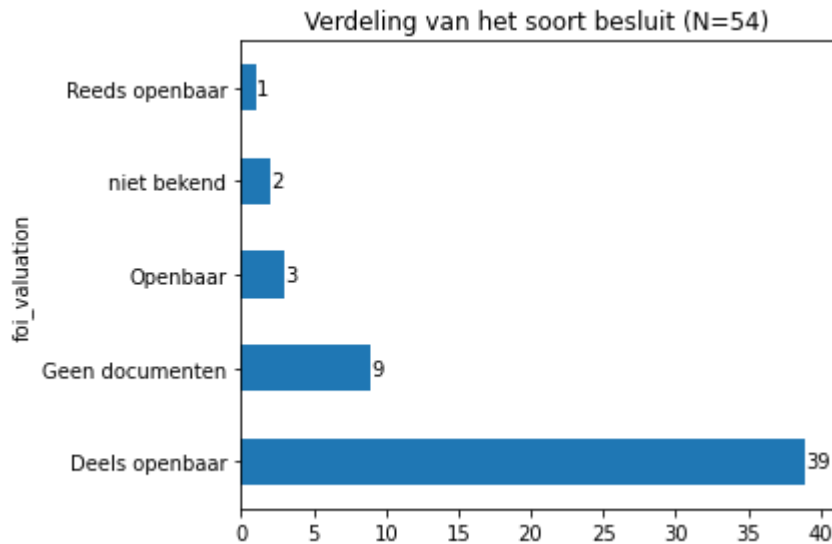
nl.gm0268.2i.2022.6	brief minister BZK inzake besluit op verzoek t...	2i	Woo verzoek gedaan op 13 december 2022 over de...	gm0268	gm0268		recht
---------------------	---	----	--	--------	--------	--	-------

2 rows x 21 columns

De overzichten

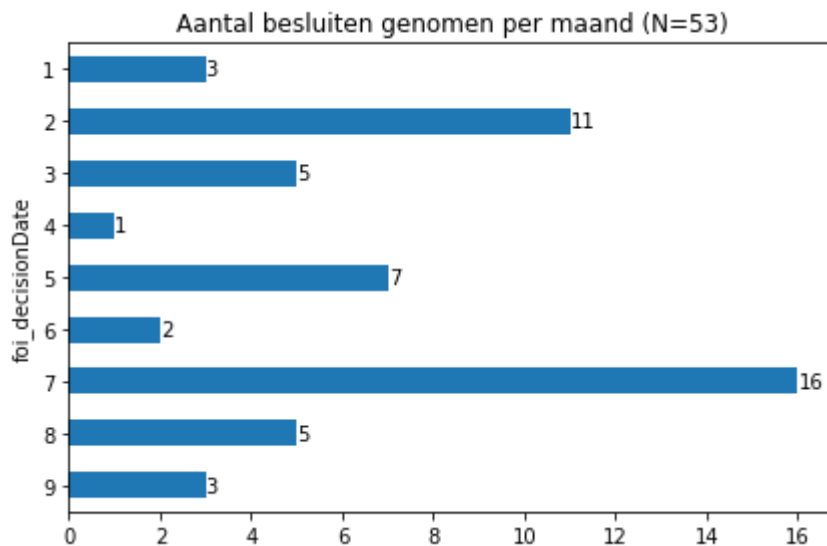
Verdeling van het soort besluit

```
In [4]: t='Verdeling van het soort besluit (N=%s)' % df.shape[0]
ax=df.foi_valuation.fillna('niet bekend').value_counts().plot(kind='barh', title=t);
ax.bar_label(ax.containers[0]);
fig= ax.get_figure()
fig.tight_layout()
fig.savefig('soort.png')
```



Aantal besluiten genomen per maand

```
In [5]: t='Aantal besluiten genomen per maand (N=%s)' % df[df.foi_decisionDate.notnull()].shape[0]
ax=df[df.foi_decisionDate.notnull()].foi_decisionDate.dt.month.value_counts().sort_index()
ax.bar_label(ax.containers[0]);
fig= ax.get_figure()
fig.tight_layout()
fig.savefig('permaand.png')
```



Verdeling van de besluiten over de Tooi themas

```
In [6]: t='Verdeling van de besluiten over de Tooi themas (N=%s)' % df.shape[0]
ax=df.tooiwl_topic.fillna('niet bekend').value_counts().plot(kind='barh', title=t);
ax.bar_label(ax.containers[0]);
fig= ax.get_figure()
fig.tight_layout()
fig.savefig('tooi.png')
```



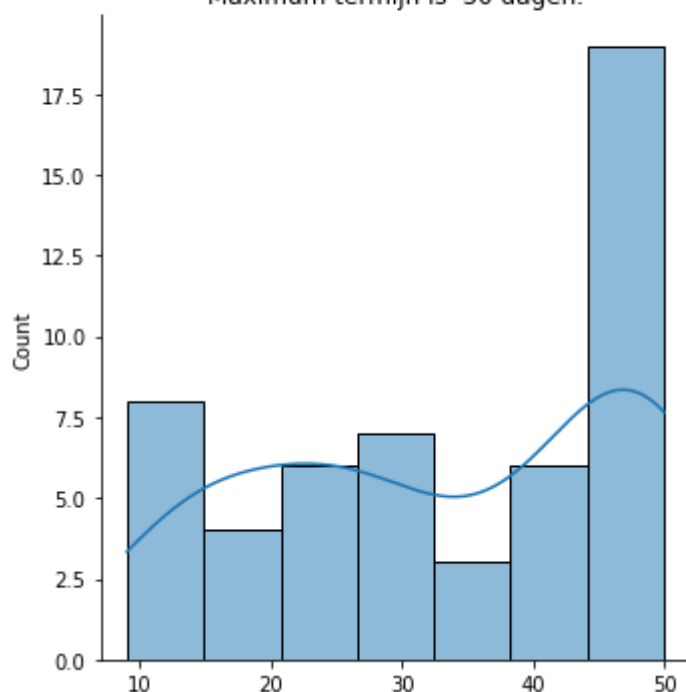
Beslistermijnen

```
In [7]: beslistijden=(df.foi_decisionDate - df.foi_requestDate).dt.days
t='Histogram van de doorlooptijden van de Woo-besluiten (N=%s). \n Gemiddelde beslist
)

ax= sns.displot(beslistijden,kde=True).set(title=t);
ax.savefig('doorlooptijden.png')
beslistijden.describe().round(0).astype(int)
```

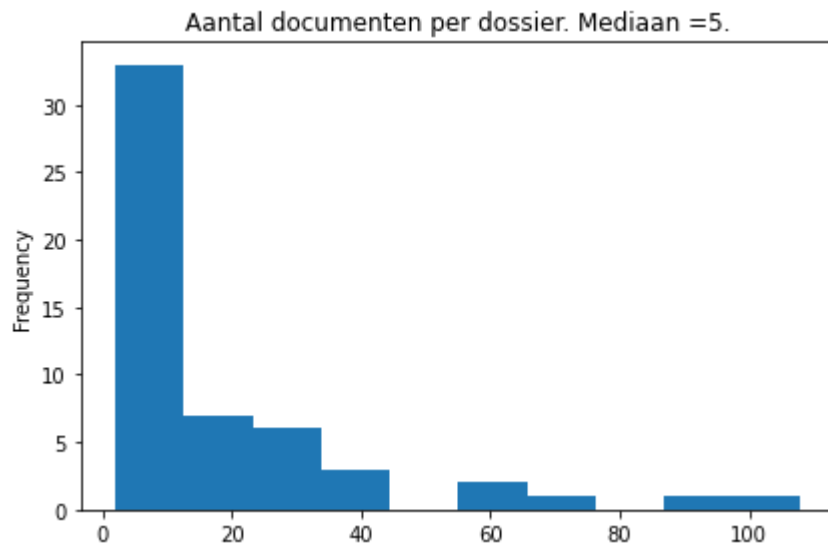
```
Out[7]: count    53
mean      34
std       14
min        9
25%       22
50%       35
75%       49
max       50
dtype: int64
```

Histogram van de doorlooptijden van de Woo-besluiten (N=53).
 Gemiddelde beslistermijn is 33 dagen.
 Maximum termijn is 50 dagen.



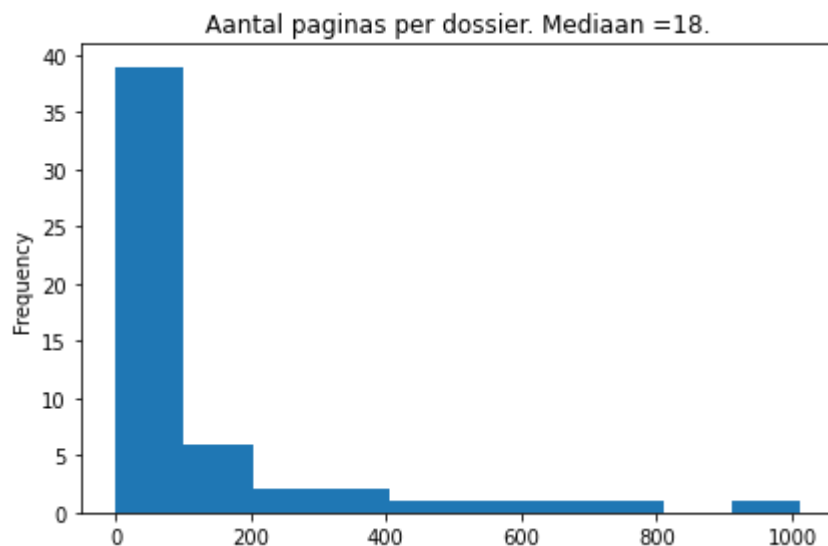
Aantal documenten per dossier

```
In [8]: t='Aantal documenten per dossier. Mediaan =%s.' % int(df.foi_nrDocuments.median())
ax=df.foi_nrDocuments.plot.hist(title=t);
fig= ax.get_figure()
fig.tight_layout()
fig.savefig('nrdocs.png')
```



Aantal paginas per dossier

```
In [9]: t='Aantal paginas per dossier. Mediaan =%s.' % int(df.foi_nrPagesInDossier.median())
ax=df.foi_nrPagesInDossier.plot.hist(title=t);
fig= ax.get_figure()
fig.tight_layout()
fig.savefig('nrpages.png')
```



Ingevulde metadata per dossier in procenten

```
In [10]: round(100 *df.notnull().sum()/df.shape[0]).astype(int).sort_index()
```

```
Out[10]: dc_creator      100
          dc_date_year    100
          dc_description   100
          dc_publisher     100
          dc_publisher_name 100
          dc_title         100
          dc_type          100
          dc_type_description 100
          foi_active       100
          foi_approved     100
          foi_decisionDate   98
          foi_decisionText   98
          foi_isAdjourned    70
          foi_nrDocuments   100
          foi_nrPagesInDossier 100
          foi_publishedDate  100
          foi_requestDate    98
          foi_requestText    100
          foi_retrievedDate  100
          foi_valuation      96
          tooiwl_topic       96
          dtype: int64
```

```
In [ ]:
```