

# notebook

April 13, 2022

## 1 Lista projektów realizowanych z Funduszy Europejskich w Polsce w latach 2014-2020 - 31 stycznia 2022 r.

Autor analizy: Maciej Jur

Data udostępnienia danych: 2 lutego 2022, 14:01

Dane na dzień: 2 lutego 2022

Źródło: <https://dane.gov.pl/pl/dataset/1176,lista-beneficjentow-projektow-funduszy-europejskich-resource/36335/table>

Licencja wykorzystanego zbioru danych: [CC0 1.0](#)

### 1.1 Narzędzia

Do przeprowadzenia analizy wykorzystano środowisko Jupyter, język Python 3.9.7 oraz pakiet Pandas.

### 1.2 Wykaz kolumn

| Index | Label  | Data type |
|-------|--|-----------|
| 0     | Tytuł projektu/<br>Project name                        | String    |
| 1     | Skrócony opis [do 2000<br>znaków] / Project<br>summary | String    |
| 2     | Numer<br>umowy/decyzji/<br>Contract number             | String    |
| 3     | Nazwa beneficjenta/<br>Beneficiary name                | String    |
| 4     | Fundusz/ Fund  | String    |
| 5     | Program/ Programme                                     | String    |
| 6     | Priorytet/ Priority axis                               | String    |
| 7     | Działanie/ Measure                                     | String    |
| 8     | Poddziałanie/<br>Submeasure                            | String    |

| Index | Label  | Data type                                     |
|-------|--|---|
| 9     | Wartość projektu<br>(w zł, dla projektów<br>EWT w euro)/<br>Total project value<br>(PLN, for ETC<br>projects EUR)                      | Float   |
| 10    | Wydatki<br>kwalifikowalne (w zł,<br>dla projektów EWT w<br>euro)/ Total eligible<br>expenditure (PLN, for<br>ETC projects EUR)         | Float   |
| 11    | Wartość unijnego<br>dofinansowania (w zł,<br>dla projektów EWT w<br>euro)/ Amount of EU<br>co-financing (PLN, for<br>ETC projects EUR) | Float   |
| 12    | Poziom unijnego<br>dofinansowania w<br>procentach (w %)/<br>Union co-financing rate<br>(%)   | Float   |
| 13    | Forma finansowania/<br>Form of finance   | String  |
| 14    | Miejsce realizacji pro-<br>jektu/ Project location   | String  |
| 15    | Typ obszaru, na<br>którym realizowany<br>jest projekt/ Territory<br>type   | String  |
| 16    | Data rozpoczęcia<br>realizacji projektu/<br>Project start date   | Date  |
| 17    | Data zakończenia<br>realizacji projektu/<br>Project end date   | Date  |
| 18    | Projekt konkursowy<br>czy pozakonkursowy/<br>Project implemented<br>under competitive or<br>non-competitive<br>procedure               | Konkursowy / Nadzwyczajny /<br>Pozakonkursowy |

| Index | Label  | Data type                |
|-------|--|--------------------------|
| 19    | Dziedzina działalności gospodarczej, której dotyczy projekt/ Area of economic activity                                       | String                   |
| 20    | Obszar wsparcia projektu/ Area of project intervention   | String                   |
| 21    | Cel projektu/ Project thematic objective   | String                   |
| 22    | Cel uzupełniający dla projektów EFS/ ESF secondary theme   | String                   |
| 23    | Projekt realizowany w ramach terytorialnych mechanizmów wdrażania/ Project implemented under territorial delivery mechanisms | ZIT / RKLS / Nie dotyczy |
| 24    | Finansowanie zakończone/ Funding completed   | Tak/Nie                  |

### 1.3 Ładowanie danych

Usuwane są kolumny: 1. Skrócony opis [do 2000 znaków] / Project summary 2. Numer umowy/decyzji/ Contract number

```
[ ]: import pandas as pd
import matplotlib as mpl
import matplotlib.pyplot as plt

mpl.rcParams['figure.dpi'] = 150

df = pd.read_csv("data.csv", sep=";").set_axis([
    "Name",
    "Summary",
    "Contract",
    "Beneficiary",
    "Fund",
    "Program",
    "Priority",
    "Measure",
    "Submeasure",
```

```

    "Total project value",
    "Total eligible expenditure",
    "Amount of EU co-financing",
    "Union co-financing rate (%)",
    "Form of finance",
    "Project location",
    "Territory type",
    "Project start date",
    "Project end date",
    "Procedure type",
    "Area of economic activity",
    "Area of project intervention",
    "Project thematic objective",
    "ESF secondary theme",
    "Territorial delivery mechanisms",
    "Funding completed"
], axis=1, inplace=False)
df.drop(df.columns[[1, 2]], axis=1, inplace=True)

```

### 1.3.1 Czyszczenie typów dla kolumn z liczbami oraz Tak/Nie

```

[ ]: fix_format = lambda n: float(n.replace(',', '.')) if type(n) == str else
    ↪float(n)
df["Total project value"] = df["Total project value"].map(fix_format)
df["Total eligible expenditure"] = df["Total eligible expenditure"].
    ↪map(fix_format)
df["Amount of EU co-financing"] = df["Amount of EU co-financing"].
    ↪map(fix_format)
df["Union co-financing rate (%)"] = df["Union co-financing rate (%)"].
    ↪map(fix_format)

```

```

[ ]: map_boolean = lambda x: False if x.lower() == "nie" else True
df["Funding completed"] = df["Funding completed"].map(map_boolean)

```

### 1.3.2 Usunięcie testowej danej

```

[ ]: dropped = df.iloc[[41743]]
_ = df.drop(41743, axis=0, inplace=True)
dropped

```

```

[ ]:
      Name Beneficiary  Fund \
41743  Test           TEST  EFRR

                                           Program \
41743  Regionalny Program Operacyjny Województwa Łódz...

                                           Priority \

```

41743 2. Innowacyjna i konkurencyjna gospodarka

|       | Measure                               | Submeasure \           |
|-------|---------------------------------------|------------------------|
| 41743 | 2.3. Zwiększenie konkurencyjności MŚP | 2.3.1. Innowacje w MŚP |

|       | Total project value | Total eligible expenditure \ |
|-------|---------------------|------------------------------|
| 41743 | 1.0                 | 0.0                          |

|       | Amount of EU co-financing ... \ |
|-------|---------------------------------|
| 41743 | 0.0 ...                         |

|       | Territory type                                    | Project start date \ |
|-------|---|----------------------|
| 41743 | 01 Duże obszary miejskie (o ludności >50 000 i... | 2020-11-01           |

|       | Project end date | Procedure type | Area of economic activity \ |
|-------|------------------|----------------|-----------------------------|
| 41743 | 2021-01-31       | Nadzwyczajny   | 19 Edukacja                 |

|       | Area of project intervention \                    |
|-------|---|
| 41743 | 001 Ogólne inwestycje produkcyjne w małych i ś... |

|       | Project thematic objective \                      |
|-------|---|
| 41743 | 03 Wzmacnianie konkurencyjności małych i średn... |

|       | ESF secondary theme \                     |
|-------|---|
| 41743 | Projekt nie jest realizowany w ramach EFS |

|       | Territorial delivery mechanisms | Funding completed |
|-------|---------------------------------|-------------------|
| 41743 | Nie dotyczy                     | False             |

[1 rows x 23 columns]

```
[ ]: df["Name"].count()
```

```
[ ]: 97701
```

```
[ ]: df.head(1)
```

|   | Name  | Beneficiary     | Fund \ |
|---|---|-----------------|--------|
| 0 | Restoration of common culture heritage as a ba... | GMINA PUCHACZÓW | EFRR   |

|   | Program \   |
|---|---|
| 0 | The ENI Cross-border Cooperation Programme Pol... |

|   | Priority \  |
|---|---|
| 0 | 1. Promotion of local culture and preservation... |

|  | Measure | Submeasure \ |
|--|---------|--------------|
|--|---------|--------------|

```

0 1.1. Promotion of local culture and history Brak poddziałania

Total project value Total eligible expenditure Amount of EU co-financing \
0 2655000.0 2655000.0 2375000.0

... Territory type Project start date \
0 ... 03 Obszary wiejskie (o małej gęstości zaludnie... 2019-07-01

Project end date Procedure type \
0 2022-06-30 Konkursowy

Area of economic activity \
0 15 Turystyka oraz działalność związana z zakwa...

Area of project intervention \
0 094 Ochrona, rozwój i promowanie dóbr publiczn...

Project thematic objective \
0 02 Zwiększanie dostępności, stopnia wykorzysta...

ESF secondary theme Territorial delivery mechanisms \
0 Projekt nie jest realizowany w ramach EFS Nie dotyczy

Funding completed
0 False

[1 rows x 23 columns]

```

#### 1.4 Liczba unikalnych wartości dla poszczególnych kolumn

```

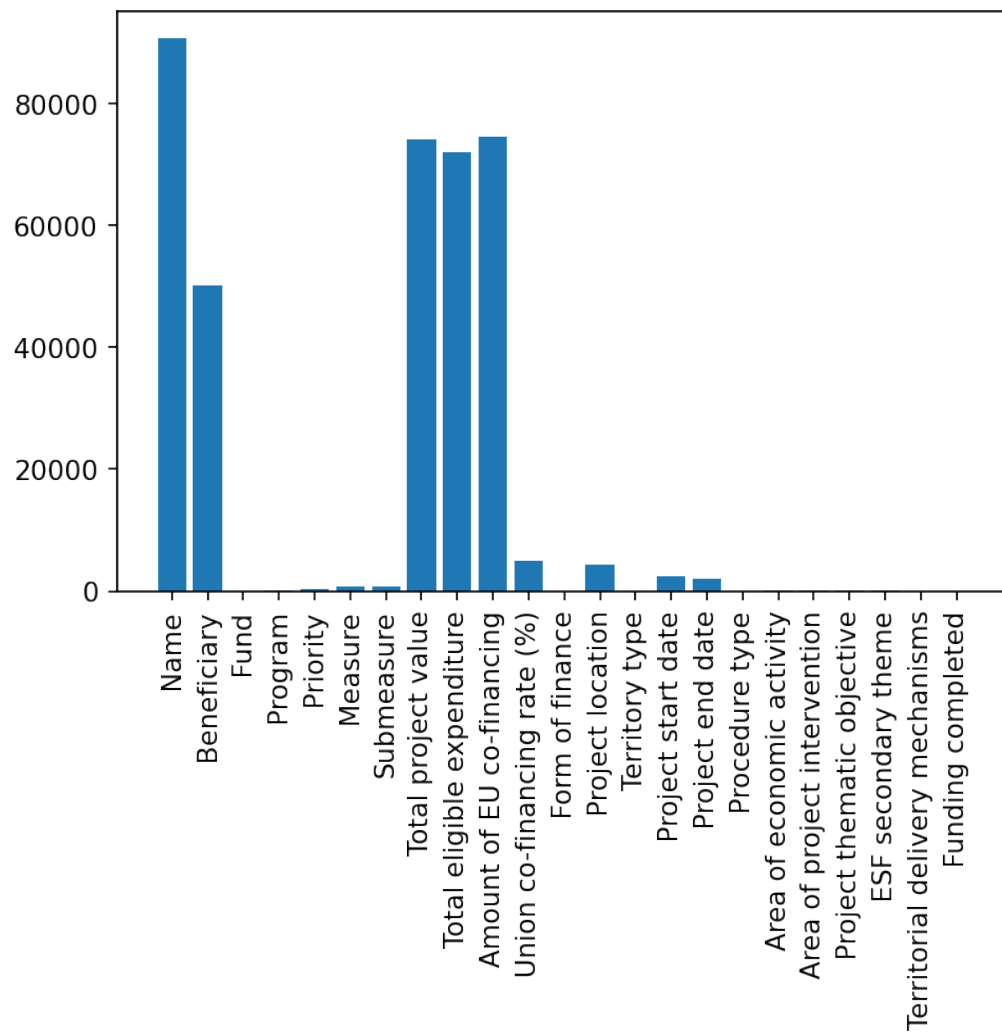
[ ]: all_unique = df.unique()
print(all_unique)
plt.xticks(rotation='vertical')
plt.bar(all_unique.index, all_unique, )

```

|                             |       |
|-----------------------------|-------|
| Name                        | 90747 |
| Beneficiary                 | 50074 |
| Fund                        | 3     |
| Program                     | 27    |
| Priority                    | 222   |
| Measure                     | 835   |
| Submeasure                  | 739   |
| Total project value         | 74220 |
| Total eligible expenditure  | 72015 |
| Amount of EU co-financing   | 74638 |
| Union co-financing rate (%) | 4944  |
| Form of finance             | 5     |
| Project location            | 4269  |

|                                 |      |
|---------------------------------|------|
| Territory type                  | 7    |
| Project start date              | 2374 |
| Project end date                | 2022 |
| Procedure type                  | 3    |
| Area of economic activity       | 24   |
| Area of project intervention    | 110  |
| Project thematic objective      | 20   |
| ESF secondary theme             | 9    |
| Territorial delivery mechanisms | 3    |
| Funding completed               | 2    |
| dtype: int64                    |      |

[ ]: <BarContainer object of 23 artists>



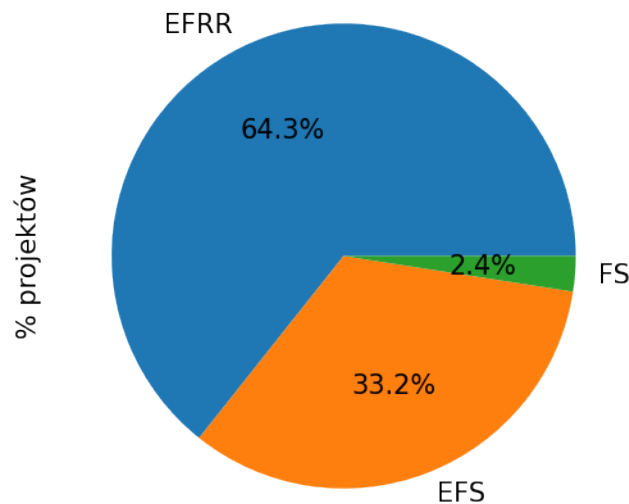
## 1.5 Fundusze

Analiza wydatków w ramach funduszy: - EFRR - Europejski Fundusz Rozwoju Regionalnego - EFS - Europejski Fundusz Społeczny - FS - Fundusz Spójności

```
[ ]: group_by_fund = df.groupby("Fund")
      print(group_by_fund_count := group_by_fund["Name"].count())
      group_by_fund_count.plot.pie(ylabel="% projektów", autopct="%1.1f%%")
```

```
Fund
EFRR    62826
EFS     32483
FS       2392
Name: Name, dtype: int64
```

```
[ ]: <AxesSubplot:ylabel='% projektów'>
```



Zdecydowana większość (około 64.3%) środków stanowiących dofinansowanie do projektów realizowanych w Polsce przekazywana była w ramach EFRR, natomiast 33.2% przekazywane było w ramach EFS, podczas gdy 2.4% przekazane zostało w ramach FS.

```
[ ]: round(group_by_fund["Total project value"].describe(), 2)
```

```
[ ]:
      count      mean      std      min      25%      50% \
Fund
EFRR  62824.0  4755071.32  3.068986e+07  2541.0  108578.57  667499.75
EFS   32483.0  2178508.62  6.429627e+06  1000.0  485960.54  996698.40
FS     2392.0  76819530.24  3.172936e+08  4800.0  2782750.00  8668188.49
```



|      | 75%         | max          |
|------|-------------|--------------|
| Fund |             |              |
| EFRR | 3020941.50  | 2.779093e+09 |
| EFS  | 1962028.40  | 3.350751e+08 |
| FS   | 30618978.19 | 5.176690e+09 |

Na podstawie tych danych zauważyć można, że w ramach Funduszu Społecznego na ogół realizowano projekty bardziej kosztownych niż w pozostałych funduszach, chociaż jest to fundusz w ramach którego zrealizowano ich najmniej. Dla FS mediana wynosiła 8 mln złotych, co stanowi około 10 razy więcej niż mediany dla EFRR oraz dla EFS.

W ramach funduszy EFRR oraz FS realizowano też bardziej zróżnicowane projekty, tj. takie które wymagają różnych nakładów kosztowych, co objawia się najmniejszym odchyleniem standardowym dla funduszu EFS.

```
[ ]: round(group_by_fund["Union co-financing rate (%)"].describe(), 2)
```

```
[ ]:
      count    mean    std    min    25%    50%    75%    max
Fund
EFRR  62826.0   79.60  18.33    0.00   68.46   85.0  100.0  100.0
EFS   32483.0   84.88   2.46   38.98   84.28   85.0   85.0  100.0
FS     2392.0   78.71  13.51   15.55   85.00   85.0   85.0  100.0
```

Projekty realizowane w ramach funduszu EFS na ogół są dofinansowywane na poziomie około 85% co objawia się wyjątkowo niskim odchyleniem standardowym, oraz średnią i medianą wynoszącą około 85%.

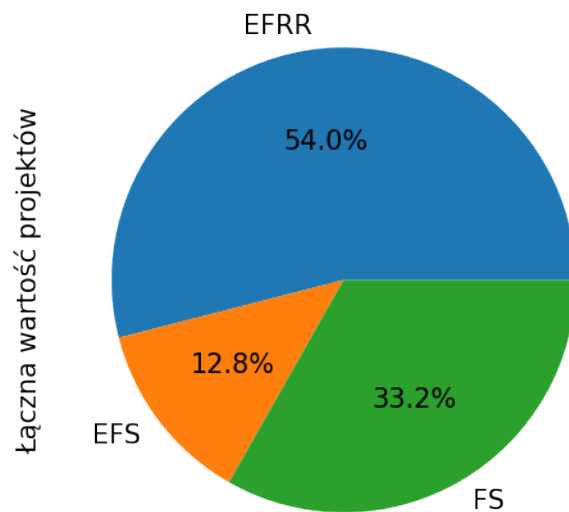
Procent dofinansowania dla projektów w funduszach EFRR oraz FS jest znacznie bardziej zróżnicowany, a w przypadku funduszu EFRR zauważyć można, że dla 25 centyla dofinansowanie jest na poziomie około 68%, a dla centyla 75% jest na poziomie 100%. Wywnioskować z tego można, że w ramach funduszu EFRR występuje spora liczba projektów, które są dofinansowane w mniej niż 70%, lub w więcej niż 85% całkowitej wartości projektu.

## 1.6 Procentowa ilość pieniędzy wydana w ramach funduszy

```
[ ]: print(group_by_fund_sum := group_by_fund["Total project value"].sum())
      group_by_fund_sum.plot.pie(ylabel="Łączna wartość projektów", autopct="%1.1f%%")
```

```
Fund
EFRR  2.987326e+11
EFS   7.076450e+10
FS    1.837523e+11
Name: Total project value, dtype: float64
```

```
[ ]: <AxesSubplot:ylabel='Łączna wartość projektów'>
```

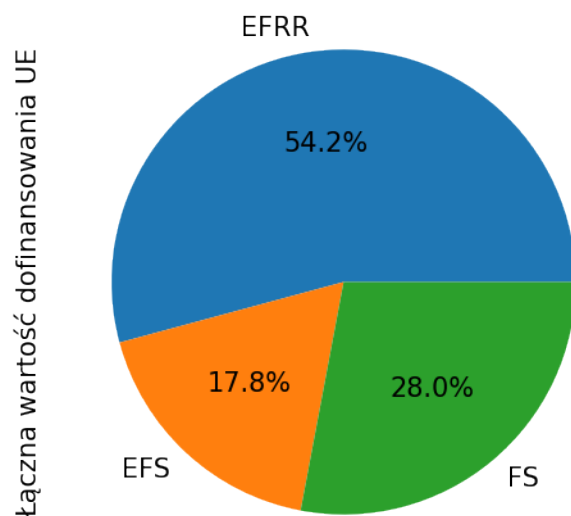


Pomimo tego, że fundusz FS obejmuje tylko około 2.4% wszystkich prowadzonych projektów, to wartość tych projektów obejmuje 33.2% całkowitej wartości wszystkich projektów. Na podstawie tego można wywnioskować, że na ogół projekty z funduszu FS mają większą wartość niż te z funduszy EFS oraz EFRR.

```
[ ]: print(group_by_fund_sum := group_by_fund["Amount of EU co-financing"].sum())
      group_by_fund_sum.plot.pie(ylabel="Łączna wartość dofinansowania UE",
      ↪autopct="%1.1f%%")
```

```
Fund
EFRR    1.829437e+11
EFS      6.021946e+10
FS       9.441420e+10
Name: Amount of EU co-financing, dtype: float64
```

```
[ ]: <AxesSubplot:ylabel='Łączna wartość dofinansowania UE'>
```



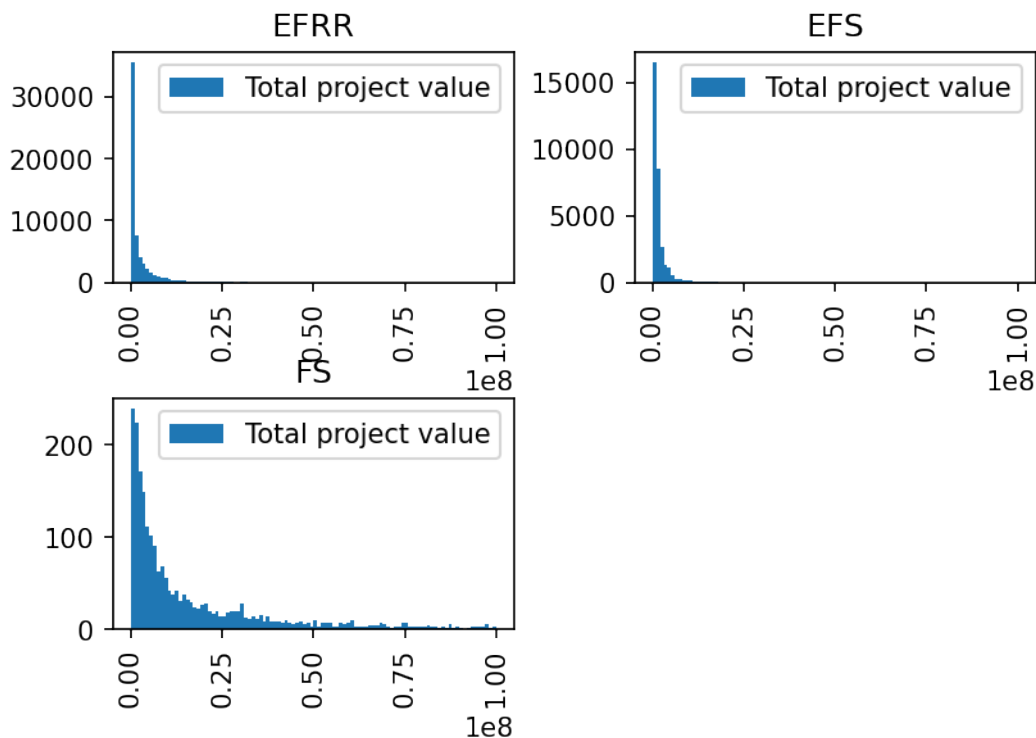
Większa wartość projektów FS oznacza, że Unia Europejska musi wydawać więcej pieniędzy na ich częściowe dofinansowanie, co jest również odzwierciedlone w procentym podziale wydatków UE na poszczególne fundusze.

### 1.6.1 Asymetria i kurtoza

#### Histogramy

```
[ ]: df["Total project value"].hist(by=df["Fund"], bins=100, range=[0, 1e8],
    ↪grid=True, legend=True)
```

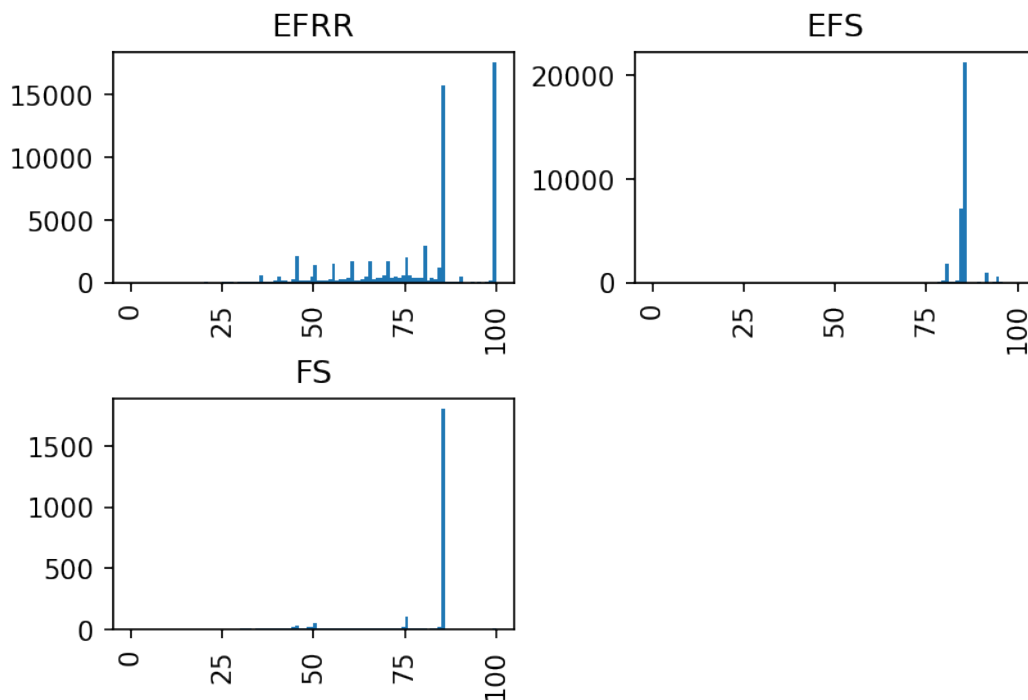
```
[ ]: array([[<AxesSubplot:title={'center': 'EFRR'}>,
    <AxesSubplot:title={'center': 'EFS'}>],
    [<AxesSubplot:title={'center': 'FS'}>, <AxesSubplot:>]],
    dtype=object)
```



Fundusze EFRR oraz EFS obejmują projekty o mniejszej wartości, natomiast fundusz FS obejmuje projekty o większej wartości, przy czym na fundusz FS przypada mniejsza liczba projektów.

```
[ ]: df["Union co-financing rate (%)"].hist(by=df["Fund"], bins=100, range=[0, 100],  
      grid=True, legend=False)
```

```
[ ]: array([[<AxesSubplot:title={'center': 'EFRR'}>,  
            <AxesSubplot:title={'center': 'EFS'}>],  
          [<AxesSubplot:title={'center': 'FS'}>, <AxesSubplot:>]],  
          dtype=object)
```



Projekty dofinansowywane w ramach funduszu EFRR mają najbardziej różnorodny poziom dofinansowania, co można zauważyć na histogramie. Zdarzają się projekty, które są dofinansowane na poziomie 100%, ale też na poziomach 85%, lub mniej. W przypadku funduszu EFS większość dofinansowań oscyluje na poziomie 80%-85%. W przypadku funduszu FS zdecydowana większość projektów jest dofinansowana na poziomie 85%.

### Asymetria

```
[ ]: group_by_fund.skew(numeric_only=True)
```

```
[ ]:      Total project value  Total eligible expenditure  \
```

| Fund | Total project value | Total eligible expenditure |
|------|---------------------|----------------------------|
| EFRR | 42.426347           | 37.244156                  |
| EFS  | 22.183675           | 22.183675                  |
| FS   | 8.634722            | 9.042194                   |

```
      Amount of EU co-financing  Union co-financing rate (%)  \
```

| Fund | Amount of EU co-financing | Union co-financing rate (%) |
|------|---------------------------|-----------------------------|
| EFRR | 45.447921                 | -0.740372                   |
| EFS  | 23.199221                 | 0.174834                    |
| FS   | 9.326420                  | -2.028145                   |

```
      Funding completed
```

| Fund | Funding completed |
|------|-------------------|
| EFRR | -0.735270         |

|     |           |
|-----|-----------|
| EFS | -0.510070 |
| FS  | 0.417054  |

Współczynnik skośności dla każdego funduszu w przypadku wartości projektu jest większa od 0, co oznacza że w każdym funduszu dominują projekty o mniejszych wartościach, tych bardziej wartościowych projektów jest zdecydowanie mniej licznościowo.

### Kurtoza

```
[ ]: group_by_fund.apply(lambda x: pd.DataFrame.kurt(x, numeric_only=True))
```

```
[ ]:      Total project value  Total eligible expenditure \
```

Fund

|      |             |             |
|------|-------------|-------------|
| EFRR | 2756.129627 | 2304.673147 |
|------|-------------|-------------|

|     |            |            |
|-----|------------|------------|
| EFS | 794.442196 | 794.442196 |
|-----|------------|------------|

|    |           |            |
|----|-----------|------------|
| FS | 92.190158 | 106.346624 |
|----|-----------|------------|

```
      Amount of EU co-financing  Union co-financing rate (%) \
```

Fund

|      |             |           |
|------|-------------|-----------|
| EFRR | 3431.291855 | -0.235357 |
|------|-------------|-----------|

|     |            |           |
|-----|------------|-----------|
| EFS | 864.137608 | 23.012291 |
|-----|------------|-----------|

|    |            |          |
|----|------------|----------|
| FS | 112.663084 | 2.944495 |
|----|------------|----------|

```
      Funding completed
```

Fund

|      |           |
|------|-----------|
| EFRR | -1.459425 |
|------|-----------|

|     |           |
|-----|-----------|
| EFS | -1.739936 |
|-----|-----------|

|    |           |
|----|-----------|
| FS | -1.827595 |
|----|-----------|

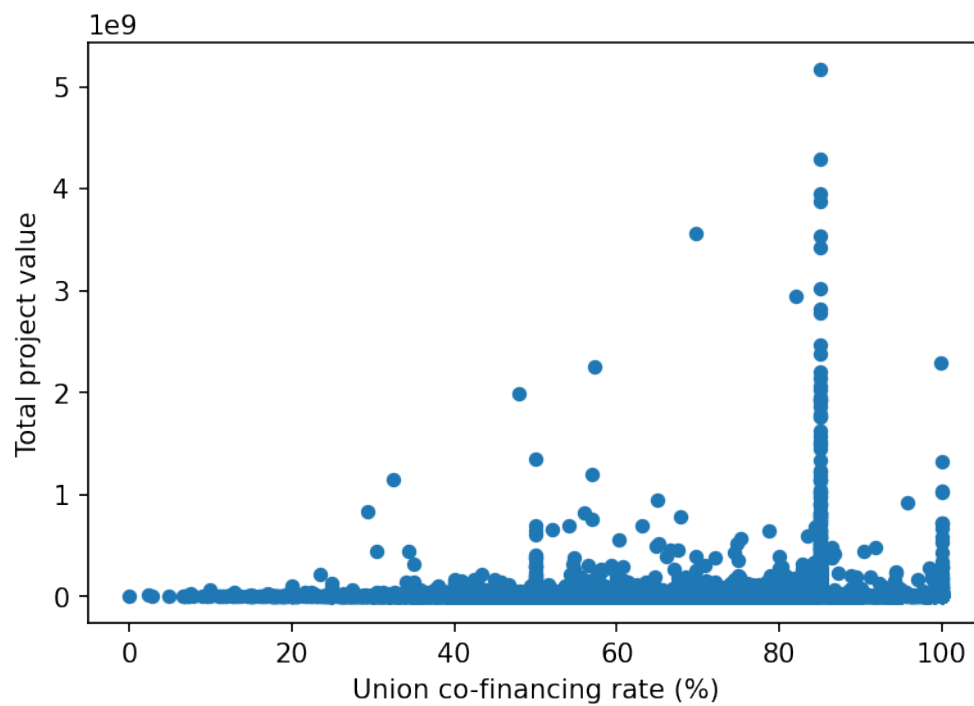
Występuje spora kurttoza dla wartości projektów co oznacza, że zdecydowana większość projektów ma bardzo zbliżoną wartość. Biorąc pod uwagę poprzednie wykresy można dojść do wniosku, że istnieje ogromna ilość projektów o bardzo niskiej wartości i są to zazwyczaj projekty regionalne prowadzone przez samorządy gmin.

## 1.7 Korelacja wartości projektu oraz % dofinansowania

```
[ ]: print(df["Union co-financing rate (%)"].corr(df["Total project value"]))
df.plot.scatter(x="Union co-financing rate (%)", y="Total project value")
```

```
-0.03129948290662574
```

```
[ ]: <AxesSubplot:xlabel='Union co-financing rate (%)', ylabel='Total project value'>
```



Pomiędzy wartością projektu oraz punktami procentowymi dofinansowania nie występuje korelacja, ale na podstawie wykresu zauważyć można, że przeważnie projekty o wyższych całkowitych wartościach mają więcej punktów procentowych dofinansowania UE.