

```
# Import library
import numpy as np
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
import matplotlib.patches as mpatches

import warnings
warnings.filterwarnings("ignore")
```

source of the dataset=

- ▼ ["https://www.kaggle.com/code/harits/salary-analysis-data-science-jobs/data"](https://www.kaggle.com/code/harits/salary-analysis-data-science-jobs/data)

```
from google.colab import drive
drive.mount('/content/drive')
```

Mounted at /content/drive

```
# Load the dataset
ds_job_salaries = pd.read_csv('/content/drive/MyDrive/ds_salaries.csv')
ds_job_salaries.head()
```

	work_year	employment_type	job_title	salary	salary_currency	salary_in_usd
0	2020	FT	Data Scientist	70000	EUR	77000
1	2020	FT	Machine Learning Scientist	260000	USD	260000
2	2020	FT	Big Data Engineer	85000	GBP	106000
3	2020	FT	Product Data Analyst	20000	USD	20000
4	2022	FT	Data Analyst	155000	USD	155000

```
ds_job_salaries.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 110 entries, 0 to 109
Data columns (total 6 columns):
 #   Column              Non-Null Count  Dtype
---  -
 0   work_year           110 non-null    int64
 1   employment_type     110 non-null    object
 2   job_title           110 non-null    object
 3   salary              110 non-null    int64
 4   salary_currency     110 non-null    object
 5   salary_in_usd       110 non-null    int64
```

```
dtypes: int64(3), object(3)
memory usage: 5.3+ KB
```

```
ds_job_salaries['salary'].describe()
```

```
count      1.100000e+02
mean       1.784182e+05
std        5.622533e+05
min        2.000000e+04
25%        9.370000e+04
50%        1.276875e+05
75%        1.591500e+05
max        6.000000e+06
Name: salary, dtype: float64
```

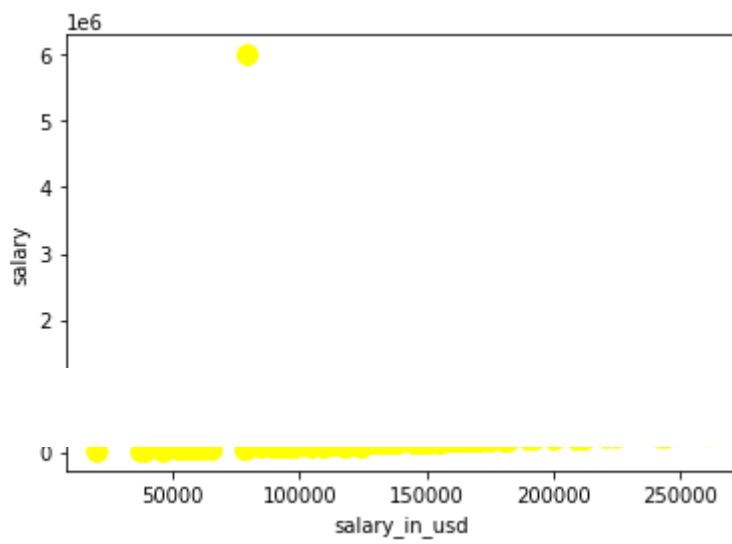
```
ds_job_salaries['employment_type'].describe()
```

```
count      110
unique      1
top         FT
freq       110
Name: employment_type, dtype: object
```

```
ds_job_salaries.sort_values("salary")
```

	work_year	employment_type	job_title	salary	salary_currency	salary_in_usd
3	2020	FT	Product Data Analyst	20000	USD	20000
100	2022	FT	Machine Learning Engineer	28500	GBP	35625
26	2022	FT	Data Scientist	30000	GBP	37500
30	2022	FT	Data Engineer	35000	GBP	43750
28	2022	FT	Data Engineer	40000	GBP	50000
...	...	...	...	...	...	...
56	2022	FT	Head of Data Science	224000	USD	224000
23	2022	FT	Data Engineer	242000	USD	242000
51	2022	FT	Data Engineer	243900	USD	243900
1	2020	FT	Machine Learning Scientist	260000	USD	260000

```
ds_job_salaries.plot.scatter(x = 'salary_in_usd', y = 'salary', s = 100, c = 'yellow');
```



[Colab paid products](#) - [Cancel contracts here](#)

✓ 0s completed at 3:55 PM

