Welcome to 30 ML | Day 8

Import Library | pandas

```
In [1]: import pandas as pd
```

Import Datasets: Titanic

```
In [2]: df = pd.read_csv('train.csv')
```

```
In [4]: df.head()
```

Out[4]:

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Cabin	Emb
0	1	0	3	Braund, Mr. Owen Harris	male	22.0	1	0	A/5 21171	7.2500	NaN	
1	2	1	1	Cumings, Mrs. John Bradley (Florence Briggs Th...	female	38.0	1	0	PC 17599	71.2833	C85	
2	3	1	3	Heikkinen, Miss. Laina	female	26.0	0	0	STON/O2. 3101282	7.9250	NaN	
3	4	1	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0	1	0	113803	53.1000	C123	
4	5	0	3	Allen, Mr. William Henry	male	35.0	0	0	373450	8.0500	NaN	

Pandas Profiling

In [7]: ```
!pip install ydata-profiling
```

```
Requirement already satisfied: ydata-profiling in c:\users\asus\anaconda3\lib\site-p
ackages (4.6.3)
Requirement already satisfied: scipy<1.12,>=1.4.1 in c:\users\asus\anaconda3\lib\sit
e-packages (from ydata-profiling) (1.11.1)
Requirement already satisfied: pandas!=1.4.0,<3,>1.1 in c:\users\asus\anaconda3\lib
\site-packages (from ydata-profiling) (2.0.3)
Requirement already satisfied: matplotlib<3.9,>=3.2 in c:\users\asus\anaconda3\lib\s
ite-packages (from ydata-profiling) (3.7.2)
Requirement already satisfied: pydantic>=2 in c:\users\asus\anaconda3\lib\site-packa
ges (from ydata-profiling) (2.5.2)
Requirement already satisfied: PyYAML<6.1,>=5.0.0 in c:\users\asus\anaconda3\lib\sit
e-packages (from ydata-profiling) (6.0)
Requirement already satisfied: jinja2<3.2,>=2.11.1 in c:\users\asus\anaconda3\lib\si
te-packages (from ydata-profiling) (3.1.2)
Requirement already satisfied: visions[type_image_path]==0.7.5 in c:\users\asus\anac
onda3\lib\site-packages (from ydata-profiling) (0.7.5)
Requirement already satisfied: numpy<1.26,>=1.16.0 in c:\users\asus\anaconda3\lib\si
te-packages (from ydata-profiling) (1.24.3)
Requirement already satisfied: htmlmin==0.1.12 in c:\users\asus\anaconda3\lib\site-p
ackages (from ydata-profiling) (0.1.12)
Requirement already satisfied: phik<0.13,>=0.11.1 in c:\users\asus\anaconda3\lib\sit
e-packages (from ydata-profiling) (0.12.3)
Requirement already satisfied: requests<3,>=2.24.0 in c:\users\asus\anaconda3\lib\si
te-packages (from ydata-profiling) (2.31.0)
Requirement already satisfied: tqdm<5,>=4.48.2 in c:\users\asus\anaconda3\lib\site-p
ackages (from ydata-profiling) (4.65.0)
Requirement already satisfied: seaborn<0.13,>=0.10.1 in c:\users\asus\anaconda3\lib
\site-packages (from ydata-profiling) (0.12.2)
Requirement already satisfied: multimethod<2,>=1.4 in c:\users\asus\anaconda3\lib\si
te-packages (from ydata-profiling) (1.10)
Requirement already satisfied: statsmodels<1,>=0.13.2 in c:\users\asus\anaconda3\lib
\site-packages (from ydata-profiling) (0.14.0)
Requirement already satisfied: typeguard<5,>=4.1.2 in c:\users\asus\anaconda3\lib\si
te-packages (from ydata-profiling) (4.1.5)
Requirement already satisfied: imagehash==4.3.1 in c:\users\asus\anaconda3\lib\site-
packages (from ydata-profiling) (4.3.1)
Requirement already satisfied: wordcloud>=1.9.1 in c:\users\asus\anaconda3\lib\site-
packages (from ydata-profiling) (1.9.3)
Requirement already satisfied: dacite>=1.8 in c:\users\asus\anaconda3\lib\site-packa
ges (from ydata-profiling) (1.8.1)
Requirement already satisfied: numba<0.59.0,>=0.56.0 in c:\users\asus\anaconda3\lib
\site-packages (from ydata-profiling) (0.57.1)
Requirement already satisfied: PyWavelets in c:\users\asus\anaconda3\lib\site-packag
es (from imagehash==4.3.1->ydata-profiling) (1.4.1)
Requirement already satisfied: pillow in c:\users\asus\anaconda3\lib\site-packages
(from imagehash==4.3.1->ydata-profiling) (9.4.0)
Requirement already satisfied: attrs>=19.3.0 in c:\users\asus\anaconda3\lib\site-pac
kages (from visions[type_image_path]==0.7.5->ydata-profiling) (22.1.0)
Requirement already satisfied: networkx>=2.4 in c:\users\asus\anaconda3\lib\site-pac
kages (from visions[type_image_path]==0.7.5->ydata-profiling) (3.1)
Requirement already satisfied: tangled-up-in-unicode>=0.0.4 in c:\users\asus\anacond
a3\lib\site-packages (from visions[type_image_path]==0.7.5->ydata-profiling) (0.2.0)
Requirement already satisfied: MarkupSafe>=2.0 in c:\users\asus\anaconda3\lib\site-p
ackages (from jinja2<3.2,>=2.11.1->ydata-profiling) (2.1.1)
Requirement already satisfied: contourpy>=1.0.1 in c:\users\asus\anaconda3\lib\site-
packages (from matplotlib<3.9,>=3.2->ydata-profiling) (1.0.5)
Requirement already satisfied: cycler>=0.10 in c:\users\asus\anaconda3\lib\site-pack
ages (from matplotlib<3.9,>=3.2->ydata-profiling) (0.11.0)
Requirement already satisfied: fonttools>=4.22.0 in c:\users\asus\anaconda3\lib\site
-packages (from matplotlib<3.9,>=3.2->ydata-profiling) (4.25.0)
Requirement already satisfied: kiwisolver>=1.0.1 in c:\users\asus\anaconda3\lib\site
-packages (from matplotlib<3.9,>=3.2->ydata-profiling) (1.4.4)
Requirement already satisfied: packaging>=20.0 in c:\users\asus\anaconda3\lib\site-p
ackages (from matplotlib<3.9,>=3.2->ydata-profiling) (23.1)
Requirement already satisfied: pyparsing<3.1,>=2.3.1 in c:\users\asus\anaconda3\lib
```

Import Profile Report

In [8]:
```python
from ydata_profiling import ProfileReport
prof = ProfileReport(df)
prof.to_file(output_file='output.html')
```

Summarize dataset: 0%| | 0/5 [00:00<?, ?it/s]

Generate report structure: 0%| | 0/1 [00:00<?, ?it/s]

Render HTML: 0%| | 0/1 [00:00<?, ?it/s]

Export report to file: 0%| | 0/1 [00:00<?, ?it/s]

Import Datasets: Hotel

In [9]:
```python
hotel = pd.read_csv('hotel.csv')
```

```
In [11]: hotel
```

Out[11]:

	hotel	is_canceled	lead_time	arrival_date_year	arrival_date_month	arrival_date_week_number	a
0	Resort Hotel	0	342	2015	July	27	
1	Resort Hotel	0	737	2015	July	27	
2	Resort Hotel	0	7	2015	July	27	
3	Resort Hotel	0	13	2015	July	27	
4	Resort Hotel	0	14	2015	July	27	
...	...	...	...	...	...	...	
119385	City Hotel	0	23	2017	August	35	
119386	City Hotel	0	102	2017	August	35	
119387	City Hotel	0	34	2017	August	35	
119388	City Hotel	0	109	2017	August	35	
119389	City Hotel	0	205	2017	August	35	

119390 rows × 32 columns

```
In [12]: pip install ydata-profiling
```

```
Requirement already satisfied: ydata-profiling in c:\users\asus\anaconda3\lib\site-p
ackages (4.6.3)Note: you may need to restart the kernel to use updated packages.

Requirement already satisfied: scipy<1.12,>=1.4.1 in c:\users\asus\anaconda3\lib\sit
e-packages (from ydata-profiling) (1.11.1)
Requirement already satisfied: pandas!=1.4.0,<3,>1.1 in c:\users\asus\anaconda3\lib
\site-packages (from ydata-profiling) (2.0.3)
Requirement already satisfied: matplotlib<3.9,>=3.2 in c:\users\asus\anaconda3\lib\s
ite-packages (from ydata-profiling) (3.7.2)
Requirement already satisfied: pydantic>=2 in c:\users\asus\anaconda3\lib\site-packa
ges (from ydata-profiling) (2.5.2)
Requirement already satisfied: PyYAML<6.1,>=5.0.0 in c:\users\asus\anaconda3\lib\sit
e-packages (from ydata-profiling) (6.0)
Requirement already satisfied: jinja2<3.2,>=2.11.1 in c:\users\asus\anaconda3\lib\si
te-packages (from ydata-profiling) (3.1.2)
Requirement already satisfied: visions[type_image_path]==0.7.5 in c:\users\asus\anac
onda3\lib\site-packages (from ydata-profiling) (0.7.5)
Requirement already satisfied: numpy<1.26,>=1.16.0 in c:\users\asus\anaconda3\lib\si
te-packages (from ydata-profiling) (1.24.3)
Requirement already satisfied: htmlmin==0.1.12 in c:\users\asus\anaconda3\lib\site-p
ackages (from ydata-profiling) (0.1.12)
Requirement already satisfied: phik<0.13,>=0.11.1 in c:\users\asus\anaconda3\lib\sit
e-packages (from ydata-profiling) (0.12.3)
Requirement already satisfied: requests<3,>=2.24.0 in c:\users\asus\anaconda3\lib\si
te-packages (from ydata-profiling) (2.31.0)
Requirement already satisfied: tqdm<5,>=4.48.2 in c:\users\asus\anaconda3\lib\site-p
ackages (from ydata-profiling) (4.65.0)
Requirement already satisfied: seaborn<0.13,>=0.10.1 in c:\users\asus\anaconda3\lib
\site-packages (from ydata-profiling) (0.12.2)
Requirement already satisfied: multimethod<2,>=1.4 in c:\users\asus\anaconda3\lib\si
te-packages (from ydata-profiling) (1.10)
Requirement already satisfied: statsmodels<1,>=0.13.2 in c:\users\asus\anaconda3\lib
\site-packages (from ydata-profiling) (0.14.0)
Requirement already satisfied: typeguard<5,>=4.1.2 in c:\users\asus\anaconda3\lib\si
te-packages (from ydata-profiling) (4.1.5)
Requirement already satisfied: imagehash==4.3.1 in c:\users\asus\anaconda3\lib\site-
packages (from ydata-profiling) (4.3.1)
Requirement already satisfied: wordcloud>=1.9.1 in c:\users\asus\anaconda3\lib\site-
packages (from ydata-profiling) (1.9.3)
Requirement already satisfied: dacite>=1.8 in c:\users\asus\anaconda3\lib\site-packa
ges (from ydata-profiling) (1.8.1)
Requirement already satisfied: numba<0.59.0,>=0.56.0 in c:\users\asus\anaconda3\lib
\site-packages (from ydata-profiling) (0.57.1)
Requirement already satisfied: PyWavelets in c:\users\asus\anaconda3\lib\site-packag
es (from imagehash==4.3.1->ydata-profiling) (1.4.1)
Requirement already satisfied: pillow in c:\users\asus\anaconda3\lib\site-packages
(from imagehash==4.3.1->ydata-profiling) (9.4.0)
Requirement already satisfied: attrs>=19.3.0 in c:\users\asus\anaconda3\lib\site-pac
kages (from visions[type_image_path]==0.7.5->ydata-profiling) (22.1.0)
Requirement already satisfied: networkx>=2.4 in c:\users\asus\anaconda3\lib\site-pac
kages (from visions[type_image_path]==0.7.5->ydata-profiling) (3.1)
Requirement already satisfied: tangled-up-in-unicode>=0.0.4 in c:\users\asus\anacond
a3\lib\site-packages (from visions[type_image_path]==0.7.5->ydata-profiling) (0.2.0)
Requirement already satisfied: MarkupSafe>=2.0 in c:\users\asus\anaconda3\lib\site-p
ackages (from jinja2<3.2,>=2.11.1->ydata-profiling) (2.1.1)
Requirement already satisfied: contourpy>=1.0.1 in c:\users\asus\anaconda3\lib\site-
packages (from matplotlib<3.9,>=3.2->ydata-profiling) (1.0.5)
Requirement already satisfied: cycler>=0.10 in c:\users\asus\anaconda3\lib\site-pack
ages (from matplotlib<3.9,>=3.2->ydata-profiling) (0.11.0)
Requirement already satisfied: fonttools>=4.22.0 in c:\users\asus\anaconda3\lib\site
-packages (from matplotlib<3.9,>=3.2->ydata-profiling) (4.25.0)
Requirement already satisfied: kiwisolver>=1.0.1 in c:\users\asus\anaconda3\lib\site
-packages (from matplotlib<3.9,>=3.2->ydata-profiling) (1.4.4)
Requirement already satisfied: packaging>=20.0 in c:\users\asus\anaconda3\lib\site-p
ackages (from matplotlib<3.9,>=3.2->ydata-profiling) (23.1)
```

```
Requirement already satisfied: pyparsing<3.1,>=2.3.1 in c:\users\asus\anaconda3\lib
\site-packages (from matplotlib<3.9,>=3.2->ydata-profiling) (3.0.9)
Requirement already satisfied: python-dateutil>=2.7 in c:\users\asus\anaconda3\lib\s
ite-packages (from matplotlib<3.9,>=3.2->ydata-profiling) (2.8.2)
Requirement already satisfied: llvmlite<0.41,>=0.40.0dev0 in c:\users\asus\anaconda3
\lib\site-packages (from numba<0.59.0,>=0.56.0->ydata-profiling) (0.40.0)
Requirement already satisfied: pytz>=2020.1 in c:\users\asus\anaconda3\lib\site-pack
ages (from pandas!=1.4.0,<3,>1.1->ydata-profiling) (2023.3.post1)
Requirement already satisfied: tzdata>=2022.1 in c:\users\asus\anaconda3\lib\site-pa
ckages (from pandas!=1.4.0,<3,>1.1->ydata-profiling) (2023.3)
Requirement already satisfied: joblib>=0.14.1 in c:\users\asus\anaconda3\lib\site-pa
ckages (from phik<0.13,>=0.11.1->ydata-profiling) (1.1.1)
Requirement already satisfied: annotated-types>=0.4.0 in c:\users\asus\anaconda3\lib
\site-packages (from pydantic>=2->ydata-profiling) (0.6.0)
Requirement already satisfied: pydantic-core==2.14.5 in c:\users\asus\anaconda3\lib
\site-packages (from pydantic>=2->ydata-profiling) (2.14.5)
Requirement already satisfied: typing-extensions>=4.6.1 in c:\users\asus\anaconda3\l
ib\site-packages (from pydantic>=2->ydata-profiling) (4.7.1)
Requirement already satisfied: charset-normalizer<4,>=2 in c:\users\asus\anaconda3\l
ib\site-packages (from requests<3,>=2.24.0->ydata-profiling) (2.0.4)
Requirement already satisfied: idna<4,>=2.5 in c:\users\asus\anaconda3\lib\site-pack
ages (from requests<3,>=2.24.0->ydata-profiling) (3.4)
Requirement already satisfied: urllib3<3,>=1.21.1 in c:\users\asus\anaconda3\lib\sit
e-packages (from requests<3,>=2.24.0->ydata-profiling) (1.26.16)
Requirement already satisfied: certifi>=2017.4.17 in c:\users\asus\anaconda3\lib\sit
e-packages (from requests<3,>=2.24.0->ydata-profiling) (2023.7.22)
Requirement already satisfied: patsy>=0.5.2 in c:\users\asus\anaconda3\lib\site-pack
ages (from statsmodels<1,>=0.13.2->ydata-profiling) (0.5.3)
Requirement already satisfied: colorama in c:\users\asus\anaconda3\lib\site-packages
(from tqdm<5,>=4.48.2->ydata-profiling) (0.4.6)
Requirement already satisfied: six in c:\users\asus\anaconda3\lib\site-packages (fro
m patsy>=0.5.2->statsmodels<1,>=0.13.2->ydata-profiling) (1.16.0)
```

In [13]:
```python
from ydata_profiling import ProfileReport
prof=ProfileReport(hotel)
prof.to_file(output_file='output.html')
```

```
Summarize dataset:   0%|          | 0/5 [00:00<?, ?it/s]

Generate report structure:   0%|          | 0/1 [00:00<?, ?it/s]

Render HTML:   0%|          | 0/1 [00:00<?, ?it/s]

Export report to file:   0%|          | 0/1 [00:00<?, ?it/s]
```

In [ ]: