

```
!pip install pandas-profiling
```

```

Looking in indexes: https://pypi.org/simple, https://us-python.pkg.dev/colab-wheels/public/simple/
Requirement already satisfied: pandas-profiling in /usr/local/lib/python3.9/dist-packages (3.2.0)
Requirement already satisfied: PyYAML>=5.0.0 in /usr/local/lib/python3.9/dist-packages (from pandas-profiling) (6.0)
Requirement already satisfied: missingno>=0.4.2 in /usr/local/lib/python3.9/dist-packages (from pandas-profiling) (0.5.2)
Requirement already satisfied: pandas!=1.0.0,!1.0.1,!1.0.2,!1.1.0,>=0.25.3 in /usr/local/lib/python3.9/dist-packages (from pandas-pr
Requirement already satisfied: Jinja2>=2.11.1 in /usr/local/lib/python3.9/dist-packages (from pandas-profiling) (3.1.2)
Requirement already satisfied: tangled-up-in-unicode==0.2.0 in /usr/local/lib/python3.9/dist-packages (from pandas-profiling) (0.2.0)
Requirement already satisfied: MarkupSafe>=2.1.1 in /usr/local/lib/python3.9/dist-packages (from pandas-profiling) (2.1.2)
Requirement already satisfied: requests>=2.24.0 in /usr/local/lib/python3.9/dist-packages (from pandas-profiling) (2.27.1)
Requirement already satisfied: Seaborn>=0.10.1 in /usr/local/lib/python3.9/dist-packages (from pandas-profiling) (0.12.2)
Requirement already satisfied: Pydantic>=1.8.1 in /usr/local/lib/python3.9/dist-packages (from pandas-profiling) (1.10.7)
Requirement already satisfied: Visions[type_image_path]==0.7.4 in /usr/local/lib/python3.9/dist-packages (from pandas-profiling) (0.7.4)
Requirement already satisfied: Scipy>=1.4.1 in /usr/local/lib/python3.9/dist-packages (from pandas-profiling) (1.10.1)
Requirement already satisfied: HTMLmin>=0.1.12 in /usr/local/lib/python3.9/dist-packages (from pandas-profiling) (0.1.12)
Requirement already satisfied: Tqdm>=4.48.2 in /usr/local/lib/python3.9/dist-packages (from pandas-profiling) (4.65.0)
Requirement already satisfied: Phik>=0.11.1 in /usr/local/lib/python3.9/dist-packages (from pandas-profiling) (0.12.3)
Requirement already satisfied: Joblib>=1.1.0 in /usr/local/lib/python3.9/dist-packages (from pandas-profiling) (1.1.1)
Requirement already satisfied: Numpy>=1.16.0 in /usr/local/lib/python3.9/dist-packages (from pandas-profiling) (1.22.4)
Requirement already satisfied: Multimethod>=1.4 in /usr/local/lib/python3.9/dist-packages (from pandas-profiling) (1.9.1)
Requirement already satisfied: Matplotlib>=3.2.0 in /usr/local/lib/python3.9/dist-packages (from pandas-profiling) (3.7.1)
Requirement already satisfied: Attrs>=19.3.0 in /usr/local/lib/python3.9/dist-packages (from visions[type_image_path]==0.7.4->pandas-pr
Requirement already satisfied: NetworkX>=2.4 in /usr/local/lib/python3.9/dist-packages (from visions[type_image_path]==0.7.4->pandas-pr
Requirement already satisfied: Pillow in /usr/local/lib/python3.9/dist-packages (from visions[type_image_path]==0.7.4->pandas-profiling)
Requirement already satisfied: Imagehash in /usr/local/lib/python3.9/dist-packages (from visions[type_image_path]==0.7.4->pandas-profiling)
Requirement already satisfied: Python-dateutil>=2.7 in /usr/local/lib/python3.9/dist-packages (from matplotlib>=3.2.0->pandas-profiling)
Requirement already satisfied: Pyparsing>=2.3.1 in /usr/local/lib/python3.9/dist-packages (from matplotlib>=3.2.0->pandas-profiling) (3
Requirement already satisfied: Contourpy>=1.0.1 in /usr/local/lib/python3.9/dist-packages (from matplotlib>=3.2.0->pandas-profiling) (1
Requirement already satisfied: Importlib-resources>=3.2.0 in /usr/local/lib/python3.9/dist-packages (from matplotlib>=3.2.0->pandas-pro
Requirement already satisfied: Packaging>=20.0 in /usr/local/lib/python3.9/dist-packages (from matplotlib>=3.2.0->pandas-profiling) (23
Requirement already satisfied: Cycler>=0.10 in /usr/local/lib/python3.9/dist-packages (from matplotlib>=3.2.0->pandas-profiling) (0.11.
Requirement already satisfied: Kiwisolver>=1.0.1 in /usr/local/lib/python3.9/dist-packages (from matplotlib>=3.2.0->pandas-profiling) (
Requirement already satisfied: Fonttools>=4.22.0 in /usr/local/lib/python3.9/dist-packages (from matplotlib>=3.2.0->pandas-profiling) (
Requirement already satisfied: Pytz>=2020.1 in /usr/local/lib/python3.9/dist-packages (from pandas!=1.0.0,!1.0.1,!1.0.2,!1.1.0,>=0.2
Requirement already satisfied: Typing-extensions>=4.2.0 in /usr/local/lib/python3.9/dist-packages (from pydantic>=1.8.1->pandas-profili
Requirement already satisfied: Charsheet-normalizer>=2.0.0 in /usr/local/lib/python3.9/dist-packages (from requests>=2.24.0->pandas-profi
Requirement already satisfied: IDNA<4,>=2.5 in /usr/local/lib/python3.9/dist-packages (from requests>=2.24.0->pandas-profiling) (3.4)
Requirement already satisfied: Urllib3<1.27,>=1.21.1 in /usr/local/lib/python3.9/dist-packages (from requests>=2.24.0->pandas-profiling)
Requirement already satisfied: Certifi>=2017.4.17 in /usr/local/lib/python3.9/dist-packages (from requests>=2.24.0->pandas-profiling) (
Requirement already satisfied: Zipp>=3.1.0 in /usr/local/lib/python3.9/dist-packages (from importlib-resources>=3.2.0->matplotlib>=3.2.
Requirement already satisfied: Six>=1.5 in /usr/local/lib/python3.9/dist-packages (from python-dateutil>=2.7->matplotlib>=3.2.0->pandas
Requirement already satisfied: PyWavelets in /usr/local/lib/python3.9/dist-packages (from imagehash->visions[type_image_path]==0.7.4->p

```

```
from pandas_profiling import ProfileReport
```

```

import pandas as pd
df=pd.read_csv('/content/sample_data/california_housing_train.csv')
df.head()

```

	longitude	latitude	housing_median_age	total_rooms	total_bedrooms	population	ho
0	-114.31	34.19	15.0	5612.0	1283.0	1015.0	
1	-114.47	34.40	19.0	7650.0	1901.0	1129.0	
2	-114.56	33.69	17.0	720.0	174.0	333.0	
3	-114.57	33.64	14.0	1501.0	337.0	515.0	
4	-114.57	33.57	20.0	1454.0	326.0	624.0	

```

res=ProfileReport(df)
res

```

Summarize dataset:
100%

Generate report structure: 100%

Render HTML: 100%

103/103 [00:24<00:00, 4.10it/s,
Completed]

1/1 [00:07<00:00, 7.86s/it]

1/1 [00:04<00:00, 4.95s/it]

Overview

Dataset statistics

Number of variables	9
Number of observations	17000
Missing cells	0
Missing cells (%)	0.0%
Duplicate rows	0
Duplicate rows (%)	0.0%
Total size in memory	1.2 MiB
Average record size in memory	72.0 B

Variable types

res.to_file('result.html');

Export report to file: 100%

1/1 [00:00<00:00, 14.34it/s]

longitude is highly correlated with latitude	High correlation
latitude is highly correlated with longitude	High correlation
total_rooms is highly correlated with total_bedrooms and 2 other fields (total_bedrooms, population, households)	High correlation
total bedrooms is highly correlated with total_rooms and 2 other fields	High correlation

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

✓ 0s completed at 10:33 PM

