Data analytics workflow

Data analysis example workbook

Kunal Khurana

2024-03-16

Table of contents


```
#!pip install calmap
!pip install ydata-profiling
```

```
Collecting ydata-profiling
  Downloading ydata_profiling-4.6.5-py2.py3-none-any.whl.metadata (20 kB)
Collecting scipy<1.12,>=1.4.1 (from ydata-profiling)
  Using cached scipy-1.11.4-cp311-cp311-win_amd64.whl.metadata (60 kB)
Requirement already satisfied: pandas!=1.4.0,<3,>1.1 in c:\users\khurana_kunal\appdata\local\
Requirement already satisfied: matplotlib<3.9,>=3.2 in c:\users\khurana_kunal\appdata\roaming
Collecting pydantic>=2 (from ydata-profiling)
  Downloading pydantic-2.6.4-py3-none-any.whl.metadata (85 kB)
     ---------------------------------------- 0.0/85.1 kB ? eta -:--:--
     ---- ----------------------------------- 10.2/85.1 kB ? eta -:--:--
     ------------------------------------ - 81.9/85.1 kB 1.1 MB/s eta 0:00:01
     ---------------------------------------- 85.1/85.1 kB 947.4 kB/s eta 0:00:00
Requirement already satisfied: PyYAML<6.1,>=5.0.0 in c:\users\khurana_kunal\appdata\local\pro
Requirement already satisfied: jinja2<3.2,>=2.11.1 in c:\users\khurana_kunal\appdata\local\pr
Collecting visions==0.7.5 (from visions[type_image_path]==0.7.5->ydata-profiling)
  Downloading visions-0.7.5-py3-none-any.whl.metadata (6.3 kB)
```

```
Collecting numpy<1.26,>=1.16.0 (from ydata-profiling)
  Downloading numpy-1.25.2-cp311-cp311-win_amd64.whl.metadata (5.7 kB)
Collecting htmlmin==0.1.12 (from ydata-profiling)
  Downloading htmlmin-0.1.12.tar.gz (19 kB)
  Installing build dependencies: started
  Installing build dependencies: finished with status 'done'
  Getting requirements to build wheel: started
  Getting requirements to build wheel: finished with status 'done'
  Installing backend dependencies: started
  Installing backend dependencies: finished with status 'done'
  Preparing metadata (pyproject.toml): started
  Preparing metadata (pyproject.toml): finished with status 'done'
Collecting phik<0.13,>=0.11.1 (from ydata-profiling)
  Downloading phik-0.12.4-cp311-cp311-win_amd64.whl.metadata (5.6 kB)
Requirement already satisfied: requests<3,>=2.24.0 in c:\users\khurana_kunal\appdata\local\pr
Requirement already satisfied: tqdm<5,>=4.48.2 in c:\users\khurana_kunal\appdata\local\progra
Collecting seaborn<0.13,>=0.10.1 (from ydata-profiling)
  Downloading seaborn-0.12.2-py3-none-any.whl.metadata (5.4 kB)
Collecting multimethod<2,>=1.4 (from ydata-profiling)
  Downloading multimethod-1.11.2-py3-none-any.whl.metadata (9.1 kB)
Collecting statsmodels<1,>=0.13.2 (from ydata-profiling)
  Downloading statsmodels-0.14.1-cp311-cp311-win_amd64.whl.metadata (9.8 kB)
Collecting typeguard<5,>=4.1.2 (from ydata-profiling)
  Downloading typeguard-4.1.5-py3-none-any.whl.metadata (3.7 kB)
Collecting imagehash==4.3.1 (from ydata-profiling)
  Downloading ImageHash-4.3.1-py2.py3-none-any.whl.metadata (8.0 kB)
Collecting wordcloud>=1.9.1 (from ydata-profiling)
  Downloading wordcloud-1.9.3-cp311-cp311-win_amd64.whl.metadata (3.5 kB)
Collecting dacite>=1.8 (from ydata-profiling)
  Downloading dacite-1.8.1-py3-none-any.whl.metadata (15 kB)
Collecting numba<0.59.0,>=0.56.0 (from ydata-profiling)
  Using cached numba-0.58.1-cp311-cp311-win_amd64.whl.metadata (2.8 kB)
Collecting PyWavelets (from imagehash==4.3.1->ydata-profiling)
  Downloading pywavelets-1.5.0-cp311-cp311-win_amd64.whl.metadata (9.0 kB)
Requirement already satisfied: pillow in c:\users\khurana_kunal\appdata\local\programs\pythor
Requirement already satisfied: attrs>=19.3.0 in c:\users\khurana_kunal\appdata\local\programs
Collecting networkx>=2.4 (from visions==0.7.5->visions[type_image_path]==0.7.5->ydata-profili
  Downloading networkx-3.2.1-py3-none-any.whl.metadata (5.2 kB)
Collecting tangled-up-in-unicode>=0.0.4 (from visions==0.7.5->visions[type_image_path]==0.7.5
  Downloading tangled_up_in_unicode-0.2.0-py3-none-any.whl.metadata (4.8 kB)
Requirement already satisfied: MarkupSafe>=2.0 in c:\users\khurana_kunal\appdata\local\progra
Requirement already satisfied: contourpy>=1.0.1 in c:\users\khurana_kunal\appdata\roaming\pyt
Requirement already satisfied: cycler>=0.10 in c:\users\khurana_kunal\appdata\roaming\python\
```

```
Requirement already satisfied: fonttools>=4.22.0 in c:\users\khurana_kunal\appdata\roaming\py
Requirement already satisfied: kiwisolver>=1.0.1 in c:\users\khurana_kunal\appdata\roaming\py
Requirement already satisfied: packaging>=20.0 in c:\users\khurana_kunal\appdata\local\progra
Requirement already satisfied: pyparsing>=2.3.1 in c:\users\khurana_kunal\appdata\roaming\pyt
Requirement already satisfied: python-dateutil>=2.7 in c:\users\khurana_kunal\appdata\local\p
Collecting llvmlite<0.42,>=0.41.0dev0 (from numba<0.59.0,>=0.56.0->ydata-profiling)
  Using cached llvmlite-0.41.1-cp311-cp311-win_amd64.whl.metadata (4.9 kB)
Requirement already satisfied: pytz>=2020.1 in c:\users\khurana_kunal\appdata\local\programs\
Requirement already satisfied: tzdata>=2022.7 in c:\users\khurana_kunal\appdata\local\progra
Collecting joblib>=0.14.1 (from phik<0.13,>=0.11.1->ydata-profiling)
  Using cached joblib-1.3.2-py3-none-any.whl.metadata (5.4 kB)
Collecting annotated-types>=0.4.0 (from pydantic>=2->ydata-profiling)
  Using cached annotated_types-0.6.0-py3-none-any.whl.metadata (12 kB)
Collecting pydantic-core==2.16.3 (from pydantic>=2->ydata-profiling)
  Downloading pydantic_core-2.16.3-cp311-none-win_amd64.whl.metadata (6.6 kB)
Requirement already satisfied: typing-extensions>=4.6.1 in c:\users\khurana_kunal\appdata\loc
Requirement already satisfied: charset-normalizer<4,>=2 in c:\users\khurana_kunal\appdata\loc
Requirement already satisfied: idna<4,>=2.5 in c:\users\khurana_kunal\appdata\local\programs\
Requirement already satisfied: urllib3<3,>=1.21.1 in c:\users\khurana_kunal\appdata\local\pro
Requirement already satisfied: certifi>=2017.4.17 in c:\users\khurana_kunal\appdata\local\pro
Collecting patsy>=0.5.4 (from statsmodels<1,>=0.13.2->ydata-profiling)
  Using cached patsy-0.5.6-py2.py3-none-any.whl.metadata (3.5 kB)
Requirement already satisfied: colorama in c:\users\khurana_kunal\appdata\local\programs\pyth
Requirement already satisfied: six in c:\users\khurana_kunal\appdata\local\programs\python\py
Downloading ydata_profiling-4.6.5-py2.py3-none-any.whl (357 kB)
   ---------------------------------------- 0.0/357.9 kB ? eta -:--:--
   ---------------- ----------------------- 143.4/357.9 kB 2.8 MB/s eta 0:00:01
   -------------------------------- ------- 286.7/357.9 kB 2.9 MB/s eta 0:00:01
   ---------------------------------------- 357.9/357.9 kB 2.5 MB/s eta 0:00:00
Downloading ImageHash-4.3.1-py2.py3-none-any.whl (296 kB)
   ---------------------------------------- 0.0/296.5 kB ? eta -:--:--
   ---------------------- ----------------- 163.8/296.5 kB 4.8 MB/s eta 0:00:01
   ------------------------------------ - 286.7/296.5 kB 3.5 MB/s eta 0:00:01
   ---------------------------------------- 296.5/296.5 kB 3.0 MB/s eta 0:00:00
Downloading visions-0.7.5-py3-none-any.whl (102 kB)
   ---------------------------------------- 0.0/102.7 kB ? eta -:--:--
   ---------------------------------------- 102.7/102.7 kB 2.9 MB/s eta 0:00:00
Downloading dacite-1.8.1-py3-none-any.whl (14 kB)
Downloading multimethod-1.11.2-py3-none-any.whl (10 kB)
Using cached numba-0.58.1-cp311-cp311-win_amd64.whl (2.6 MB)
Downloading numpy-1.25.2-cp311-cp311-win_amd64.whl (15.5 MB)
   ---------------------------------------- 0.0/15.5 MB ? eta -:--:--
   ---------------------------------------- 0.1/15.5 MB 2.9 MB/s eta 0:00:06
```

```
 ------------------------------------ 0.3/15.5 MB 3.0 MB/s eta 0:00:06
- ------------------------------------ 0.4/15.5 MB 3.4 MB/s eta 0:00:05
- ------------------------------------ 0.6/15.5 MB 3.0 MB/s eta 0:00:05
- ------------------------------------ 0.7/15.5 MB 3.3 MB/s eta 0:00:05
-- ------------------------------------ 0.9/15.5 MB 3.1 MB/s eta 0:00:05
-- ------------------------------------ 1.0/15.5 MB 3.1 MB/s eta 0:00:05
-- ------------------------------------ 1.1/15.5 MB 3.2 MB/s eta 0:00:05
--- ------------------------------------ 1.3/15.5 MB 3.0 MB/s eta 0:00:05
--- ------------------------------------ 1.4/15.5 MB 3.0 MB/s eta 0:00:05
---- ------------------------------------ 1.6/15.5 MB 3.0 MB/s eta 0:00:05
---- ------------------------------------ 1.7/15.5 MB 3.1 MB/s eta 0:00:05
---- ------------------------------------ 1.9/15.5 MB 3.1 MB/s eta 0:00:05
----- ------------------------------------ 2.0/15.5 MB 3.1 MB/s eta 0:00:05
----- ------------------------------------ 2.2/15.5 MB 3.1 MB/s eta 0:00:05
----- ------------------------------------ 2.3/15.5 MB 3.1 MB/s eta 0:00:05
------ ------------------------------------ 2.5/15.5 MB 3.1 MB/s eta 0:00:05
------ ------------------------------------ 2.6/15.5 MB 3.1 MB/s eta 0:00:05
------- ------------------------------------ 2.8/15.5 MB 3.1 MB/s eta 0:00:05
------- ------------------------------------ 2.9/15.5 MB 3.1 MB/s eta 0:00:05
------- ------------------------------------ 3.1/15.5 MB 3.1 MB/s eta 0:00:05
-------- ------------------------------------ 3.2/15.5 MB 3.1 MB/s eta 0:00:04
-------- ------------------------------------ 3.3/15.5 MB 3.1 MB/s eta 0:00:04
-------- ------------------------------------ 3.5/15.5 MB 3.1 MB/s eta 0:00:04
--------- ------------------------------------ 3.7/15.5 MB 3.1 MB/s eta 0:00:04
--------- ------------------------------------ 3.8/15.5 MB 3.1 MB/s eta 0:00:04
---------- ------------------------------------ 3.9/15.5 MB 3.1 MB/s eta 0:00:04
---------- ------------------------------------ 4.1/15.5 MB 3.1 MB/s eta 0:00:04
---------- ------------------------------------ 4.2/15.5 MB 3.1 MB/s eta 0:00:04
----------- ------------------------------------ 4.4/15.5 MB 3.1 MB/s eta 0:00:04
----------- ------------------------------------ 4.5/15.5 MB 3.1 MB/s eta 0:00:04
------------ ------------------------------------ 4.7/15.5 MB 3.1 MB/s eta 0:00:04
------------ ------------------------------------ 4.8/15.5 MB 3.1 MB/s eta 0:00:04
------------ ------------------------------------ 5.0/15.5 MB 3.1 MB/s eta 0:00:04
------------- ------------------------------------ 5.1/15.5 MB 3.1 MB/s eta 0:00:04
------------- ------------------------------------ 5.3/15.5 MB 3.1 MB/s eta 0:00:04
------------- ------------------------------------ 5.4/15.5 MB 3.1 MB/s eta 0:00:04
-------------- ------------------------------------ 5.6/15.5 MB 3.1 MB/s eta 0:00:04
-------------- ------------------------------------ 5.7/15.5 MB 3.1 MB/s eta 0:00:04
--------------- ------------------------------------ 5.9/15.5 MB 3.1 MB/s eta 0:00:04
--------------- ------------------------------------ 6.0/15.5 MB 3.1 MB/s eta 0:00:04
--------------- ------------------------------------ 6.2/15.5 MB 3.1 MB/s eta 0:00:04
---------------- ------------------------------------ 6.3/15.5 MB 3.1 MB/s eta 0:00:03
---------------- ------------------------------------ 6.5/15.5 MB 3.1 MB/s eta 0:00:03
```

```
---------------- ---------------------- 6.6/15.5 MB 3.1 MB/s eta 0:00:03
---------------- ---------------------- 6.8/15.5 MB 3.1 MB/s eta 0:00:03
---------------- ---------------------- 6.9/15.5 MB 3.1 MB/s eta 0:00:03
----------------- --------------------- 7.1/15.5 MB 3.1 MB/s eta 0:00:03
----------------- --------------------- 7.2/15.5 MB 3.1 MB/s eta 0:00:03
----------------- --------------------- 7.4/15.5 MB 3.1 MB/s eta 0:00:03
------------------ -------------------- 7.5/15.5 MB 3.1 MB/s eta 0:00:03
------------------ -------------------- 7.7/15.5 MB 3.1 MB/s eta 0:00:03
------------------- ------------------- 7.8/15.5 MB 3.1 MB/s eta 0:00:03
------------------- ------------------- 7.9/15.5 MB 3.1 MB/s eta 0:00:03
------------------- ------------------- 8.1/15.5 MB 3.1 MB/s eta 0:00:03
-------------------- ------------------ 8.2/15.5 MB 3.1 MB/s eta 0:00:03
-------------------- ------------------ 8.4/15.5 MB 3.1 MB/s eta 0:00:03
-------------------- ------------------ 8.5/15.5 MB 3.1 MB/s eta 0:00:03
--------------------- ----------------- 8.7/15.5 MB 3.1 MB/s eta 0:00:03
--------------------- ----------------- 8.9/15.5 MB 3.1 MB/s eta 0:00:03
---------------------- ---------------- 9.0/15.5 MB 3.1 MB/s eta 0:00:03
---------------------- ---------------- 9.2/15.5 MB 3.1 MB/s eta 0:00:03
---------------------- ---------------- 9.3/15.5 MB 3.1 MB/s eta 0:00:02
----------------------- --------------- 9.5/15.5 MB 3.1 MB/s eta 0:00:02
----------------------- --------------- 9.6/15.5 MB 3.1 MB/s eta 0:00:02
------------------------ -------------- 9.7/15.5 MB 3.1 MB/s eta 0:00:02
------------------------ -------------- 9.9/15.5 MB 3.1 MB/s eta 0:00:02
------------------------- ------------- 10.1/15.5 MB 3.1 MB/s eta 0:00:02
------------------------- ------------- 10.3/15.5 MB 3.1 MB/s eta 0:00:02
-------------------------- ------------ 10.4/15.5 MB 3.1 MB/s eta 0:00:02
-------------------------- ------------ 10.5/15.5 MB 3.1 MB/s eta 0:00:02
-------------------------- ------------ 10.6/15.5 MB 3.1 MB/s eta 0:00:02
--------------------------- ----------- 10.7/15.5 MB 3.1 MB/s eta 0:00:02
--------------------------- ----------- 10.9/15.5 MB 3.1 MB/s eta 0:00:02
--------------------------- ----------- 11.0/15.5 MB 3.1 MB/s eta 0:00:02
---------------------------- ---------- 11.2/15.5 MB 3.1 MB/s eta 0:00:02
---------------------------- ---------- 11.3/15.5 MB 3.1 MB/s eta 0:00:02
---------------------------- ---------- 11.5/15.5 MB 3.1 MB/s eta 0:00:02
----------------------------- --------- 11.6/15.5 MB 3.1 MB/s eta 0:00:02
----------------------------- --------- 11.8/15.5 MB 3.1 MB/s eta 0:00:02
----------------------------- --------- 11.9/15.5 MB 3.1 MB/s eta 0:00:02
------------------------------ -------- 12.1/15.5 MB 3.1 MB/s eta 0:00:02
------------------------------ -------- 12.2/15.5 MB 3.1 MB/s eta 0:00:02
------------------------------ -------- 12.3/15.5 MB 3.1 MB/s eta 0:00:02
------------------------------- ------- 12.5/15.5 MB 3.1 MB/s eta 0:00:01
------------------------------- ------- 12.7/15.5 MB 3.1 MB/s eta 0:00:01
------------------------------- ------- 12.8/15.5 MB 3.1 MB/s eta 0:00:01
```

```
   ---------------------------------- ------ 12.9/15.5 MB 3.1 MB/s eta 0:00:01
   ---------------------------------- ------ 13.1/15.5 MB 3.1 MB/s eta 0:00:01
   ---------------------------------- ----- 13.3/15.5 MB 3.1 MB/s eta 0:00:01
   ---------------------------------- ----- 13.4/15.5 MB 3.1 MB/s eta 0:00:01
   ---------------------------------- ----- 13.6/15.5 MB 3.1 MB/s eta 0:00:01
   ---------------------------------- ---- 13.7/15.5 MB 3.1 MB/s eta 0:00:01
   ---------------------------------- ---- 13.9/15.5 MB 3.1 MB/s eta 0:00:01
   ---------------------------------- --- 14.0/15.5 MB 3.1 MB/s eta 0:00:01
   ---------------------------------- --- 14.1/15.5 MB 3.1 MB/s eta 0:00:01
   ---------------------------------- --- 14.3/15.5 MB 3.1 MB/s eta 0:00:01
   ---------------------------------- -- 14.4/15.5 MB 3.1 MB/s eta 0:00:01
   ---------------------------------- -- 14.6/15.5 MB 3.1 MB/s eta 0:00:01
   ---------------------------------- -- 14.7/15.5 MB 3.1 MB/s eta 0:00:01
   ---------------------------------- - 14.9/15.5 MB 3.1 MB/s eta 0:00:01
   ---------------------------------- - 15.0/15.5 MB 3.1 MB/s eta 0:00:01
   ------------------------------------  15.2/15.5 MB 3.1 MB/s eta 0:00:01
   ------------------------------------  15.3/15.5 MB 3.1 MB/s eta 0:00:01
   ------------------------------------  15.3/15.5 MB 3.1 MB/s eta 0:00:01
   ------------------------------------  15.5/15.5 MB 3.1 MB/s eta 0:00:01
   ------------------------------------ 15.5/15.5 MB 3.0 MB/s eta 0:00:00
Downloading phik-0.12.4-cp311-cp311-win_amd64.whl (667 kB)
   ---------------------------------------- 0.0/667.1 kB ? eta -:--:--
   ----- ---------------------------------- 92.2/667.1 kB 2.6 MB/s eta 0:00:01
   -------------- ------------------------- 245.8/667.1 kB 3.0 MB/s eta 0:00:01
   ---------------------- ----------------- 399.4/667.1 kB 3.1 MB/s eta 0:00:01
   ---------------------------------- ----- 593.9/667.1 kB 3.1 MB/s eta 0:00:01
   ---------------------------------------- 667.1/667.1 kB 2.8 MB/s eta 0:00:00
Downloading pydantic-2.6.4-py3-none-any.whl (394 kB)
   ---------------------------------------- 0.0/394.9 kB ? eta -:--:--
   ---------------- ----------------------- 163.8/394.9 kB 3.3 MB/s eta 0:00:01
   ------------------------------- -------- 307.2/394.9 kB 3.2 MB/s eta 0:00:01
   ------------------------------------  389.1/394.9 kB 3.0 MB/s eta 0:00:01
   ---------------------------------------- 394.9/394.9 kB 2.5 MB/s eta 0:00:00
Downloading pydantic_core-2.16.3-cp311-none-win_amd64.whl (1.9 MB)
   ---------------------------------------- 0.0/1.9 MB ? eta -:--:--
   ---- ----------------------------------- 0.2/1.9 MB 3.9 MB/s eta 0:00:01
   -------- ------------------------------- 0.4/1.9 MB 3.4 MB/s eta 0:00:01
   ----------- ---------------------------- 0.5/1.9 MB 3.4 MB/s eta 0:00:01
   -------------- ------------------------- 0.7/1.9 MB 3.3 MB/s eta 0:00:01
   ----------------- ---------------------- 0.8/1.9 MB 3.3 MB/s eta 0:00:01
   --------------------- ------------------ 1.0/1.9 MB 3.3 MB/s eta 0:00:01
   ----------------------- ---------------- 1.1/1.9 MB 3.2 MB/s eta 0:00:01
   ------------------------- -------------- 1.3/1.9 MB 3.2 MB/s eta 0:00:01
```

```
   ---------------------------- --------- 1.4/1.9 MB 3.2 MB/s eta 0:00:01
   ------------------------------- ------ 1.6/1.9 MB 3.2 MB/s eta 0:00:01
   ---------------------------------- --- 1.7/1.9 MB 3.2 MB/s eta 0:00:01
   ---------------------------------------  1.9/1.9 MB 3.2 MB/s eta 0:00:01
   ---------------------------------------  1.9/1.9 MB 3.2 MB/s eta 0:00:01
   ---------------------------------------- 1.9/1.9 MB 2.9 MB/s eta 0:00:00
Using cached scipy-1.11.4-cp311-cp311-win_amd64.whl (44.1 MB)
Downloading seaborn-0.12.2-py3-none-any.whl (293 kB)
   ---------------------------------------- 0.0/293.3 kB ? eta -:--:--
   ------------------ -------------------- 143.4/293.3 kB 2.8 MB/s eta 0:00:01
   ---------------------------------------  286.7/293.3 kB 3.0 MB/s eta 0:00:01
   ---------------------------------------- 293.3/293.3 kB 2.6 MB/s eta 0:00:00
Downloading statsmodels-0.14.1-cp311-cp311-win_amd64.whl (9.9 MB)
   ---------------------------------------- 0.0/9.9 MB ? eta -:--:--
   ---------------------------------------- 0.1/9.9 MB 2.2 MB/s eta 0:00:05
   - -------------------------------------- 0.3/9.9 MB 3.2 MB/s eta 0:00:04
   - -------------------------------------- 0.4/9.9 MB 2.8 MB/s eta 0:00:04
   -- ------------------------------------- 0.6/9.9 MB 3.2 MB/s eta 0:00:03
   -- ------------------------------------- 0.7/9.9 MB 3.1 MB/s eta 0:00:03
   --- ------------------------------------ 0.8/9.9 MB 3.1 MB/s eta 0:00:03
   ---- ----------------------------------- 1.0/9.9 MB 3.0 MB/s eta 0:00:03
   ---- ----------------------------------- 1.1/9.9 MB 3.0 MB/s eta 0:00:03
   ----- ---------------------------------- 1.3/9.9 MB 3.0 MB/s eta 0:00:03
   ----- ---------------------------------- 1.4/9.9 MB 3.1 MB/s eta 0:00:03
   ------ --------------------------------- 1.6/9.9 MB 3.1 MB/s eta 0:00:03
   ------- -------------------------------- 1.7/9.9 MB 3.1 MB/s eta 0:00:03
   ------- -------------------------------- 1.9/9.9 MB 3.1 MB/s eta 0:00:03
   -------- ------------------------------- 2.0/9.9 MB 3.1 MB/s eta 0:00:03
   -------- ------------------------------- 2.2/9.9 MB 3.1 MB/s eta 0:00:03
   --------- ------------------------------ 2.3/9.9 MB 3.1 MB/s eta 0:00:03
   ---------- ----------------------------- 2.5/9.9 MB 3.1 MB/s eta 0:00:03
   ---------- ----------------------------- 2.6/9.9 MB 3.1 MB/s eta 0:00:03
   ----------- ---------------------------- 2.8/9.9 MB 3.1 MB/s eta 0:00:03
   ----------- ---------------------------- 2.9/9.9 MB 3.1 MB/s eta 0:00:03
   ------------ --------------------------- 3.1/9.9 MB 3.1 MB/s eta 0:00:03
   ------------- -------------------------- 3.2/9.9 MB 3.1 MB/s eta 0:00:03
   ------------- -------------------------- 3.4/9.9 MB 3.1 MB/s eta 0:00:03
   -------------- ------------------------- 3.5/9.9 MB 3.1 MB/s eta 0:00:03
   -------------- ------------------------- 3.7/9.9 MB 3.1 MB/s eta 0:00:02
   --------------- ------------------------ 3.8/9.9 MB 3.1 MB/s eta 0:00:02
   ---------------- ----------------------- 4.0/9.9 MB 3.1 MB/s eta 0:00:02
   ---------------- ----------------------- 4.1/9.9 MB 3.1 MB/s eta 0:00:02
   ----------------- ---------------------- 4.3/9.9 MB 3.1 MB/s eta 0:00:02
```

```
   ---------------- ---------------------- 4.4/9.9 MB 3.1 MB/s eta 0:00:02
   ---------------- ---------------------- 4.6/9.9 MB 3.1 MB/s eta 0:00:02
   ----------------- --------------------- 4.7/9.9 MB 3.1 MB/s eta 0:00:02
   ----------------- --------------------- 4.9/9.9 MB 3.1 MB/s eta 0:00:02
   ------------------ -------------------- 5.0/9.9 MB 3.1 MB/s eta 0:00:02
   ------------------ -------------------- 5.2/9.9 MB 3.1 MB/s eta 0:00:02
   ------------------- ------------------- 5.3/9.9 MB 3.1 MB/s eta 0:00:02
   -------------------- ------------------ 5.4/9.9 MB 3.1 MB/s eta 0:00:02
   -------------------- ------------------ 5.6/9.9 MB 3.1 MB/s eta 0:00:02
   --------------------- ----------------- 5.8/9.9 MB 3.1 MB/s eta 0:00:02
   --------------------- ----------------- 5.9/9.9 MB 3.1 MB/s eta 0:00:02
   ---------------------- ---------------- 6.1/9.9 MB 3.1 MB/s eta 0:00:02
   ---------------------- ---------------- 6.2/9.9 MB 3.1 MB/s eta 0:00:02
   ----------------------- --------------- 6.4/9.9 MB 3.1 MB/s eta 0:00:02
   ----------------------- --------------- 6.5/9.9 MB 3.1 MB/s eta 0:00:02
   ------------------------ -------------- 6.7/9.9 MB 3.1 MB/s eta 0:00:02
   ------------------------ -------------- 6.8/9.9 MB 3.1 MB/s eta 0:00:01
   ------------------------- ------------- 7.0/9.9 MB 3.1 MB/s eta 0:00:01
   ------------------------- ------------- 7.1/9.9 MB 3.1 MB/s eta 0:00:01
   ------------------------- ------------- 7.3/9.9 MB 3.1 MB/s eta 0:00:01
   -------------------------- ------------ 7.4/9.9 MB 3.1 MB/s eta 0:00:01
   -------------------------- ------------ 7.5/9.9 MB 3.1 MB/s eta 0:00:01
   -------------------------- ------------ 7.6/9.9 MB 3.1 MB/s eta 0:00:01
   --------------------------- ----------- 7.8/9.9 MB 3.1 MB/s eta 0:00:01
   --------------------------- ----------- 8.0/9.9 MB 3.1 MB/s eta 0:00:01
   ---------------------------- ---------- 8.1/9.9 MB 3.1 MB/s eta 0:00:01
   ---------------------------- ---------- 8.3/9.9 MB 3.1 MB/s eta 0:00:01
   ----------------------------- --------- 8.5/9.9 MB 3.1 MB/s eta 0:00:01
   ----------------------------- --------- 8.7/9.9 MB 3.1 MB/s eta 0:00:01
   ----------------------------- --------- 8.8/9.9 MB 3.1 MB/s eta 0:00:01
   ------------------------------ -------- 9.0/9.9 MB 3.1 MB/s eta 0:00:01
   ------------------------------ -------- 9.1/9.9 MB 3.1 MB/s eta 0:00:01
   ------------------------------- ------- 9.3/9.9 MB 3.1 MB/s eta 0:00:01
   ------------------------------- ------- 9.4/9.9 MB 3.1 MB/s eta 0:00:01
   ------------------------------- ------- 9.6/9.9 MB 3.1 MB/s eta 0:00:01
   -------------------------------- ------  9.8/9.9 MB 3.1 MB/s eta 0:00:01
   -------------------------------- ------  9.9/9.9 MB 3.1 MB/s eta 0:00:01
   -------------------------------- ------ 9.9/9.9 MB 3.1 MB/s eta 0:00:00
Downloading typeguard-4.1.5-py3-none-any.whl (34 kB)
Downloading wordcloud-1.9.3-cp311-cp311-win_amd64.whl (300 kB)
   ---------------------------------------- 0.0/300.2 kB ? eta -:--:--
   ------------------- -------------------- 163.8/300.2 kB 4.8 MB/s eta 0:00:01
   ---------------------------------------  297.0/300.2 kB 3.7 MB/s eta 0:00:01
```

```
   ---------------------------------------- 300.2/300.2 kB 3.1 MB/s eta 0:00:00
Using cached annotated_types-0.6.0-py3-none-any.whl (12 kB)
Using cached joblib-1.3.2-py3-none-any.whl (302 kB)
Using cached llvmlite-0.41.1-cp311-cp311-win_amd64.whl (28.1 MB)
Downloading networkx-3.2.1-py3-none-any.whl (1.6 MB)
   ---------------------------------------- 0.0/1.6 MB ? eta -:--:--
   ---- ----------------------------------- 0.2/1.6 MB 5.3 MB/s eta 0:00:01
   -------- --------------------------------- 0.3/1.6 MB 3.5 MB/s eta 0:00:01
   ----------- ---------------------------- 0.5/1.6 MB 3.9 MB/s eta 0:00:01
   --------------- ------------------------ 0.6/1.6 MB 3.3 MB/s eta 0:00:01
   ------------------- -------------------- 0.8/1.6 MB 3.4 MB/s eta 0:00:01
   ------------------------ --------------- 1.0/1.6 MB 3.3 MB/s eta 0:00:01
   ---------------------------- ----------- 1.2/1.6 MB 3.2 MB/s eta 0:00:01
   ------------------------------- -------- 1.3/1.6 MB 3.3 MB/s eta 0:00:01
   ---------------------------------- ----- 1.5/1.6 MB 3.2 MB/s eta 0:00:01
   ---------------------------------------  1.6/1.6 MB 3.2 MB/s eta 0:00:01
   ---------------------------------------- 1.6/1.6 MB 3.1 MB/s eta 0:00:00
Using cached patsy-0.5.6-py2.py3-none-any.whl (233 kB)
Downloading tangled_up_in_unicode-0.2.0-py3-none-any.whl (4.7 MB)
   ---------------------------------------- 0.0/4.7 MB ? eta -:--:--
   - -------------------------------------- 0.1/4.7 MB 3.6 MB/s eta 0:00:02
   -- ------------------------------------- 0.3/4.7 MB 3.4 MB/s eta 0:00:02
   --- ------------------------------------ 0.4/4.7 MB 3.0 MB/s eta 0:00:02
   ---- ----------------------------------- 0.6/4.7 MB 3.2 MB/s eta 0:00:02
   ------ --------------------------------- 0.7/4.7 MB 3.1 MB/s eta 0:00:02
   ------- -------------------------------- 0.9/4.7 MB 3.1 MB/s eta 0:00:02
   --------- ------------------------------ 1.1/4.7 MB 3.1 MB/s eta 0:00:02
   ---------- ----------------------------- 1.2/4.7 MB 3.1 MB/s eta 0:00:02
   ----------- ---------------------------- 1.4/4.7 MB 3.1 MB/s eta 0:00:02
   ------------ --------------------------- 1.5/4.7 MB 3.1 MB/s eta 0:00:02
   -------------- ------------------------- 1.7/4.7 MB 3.1 MB/s eta 0:00:01
   --------------- ------------------------ 1.8/4.7 MB 3.1 MB/s eta 0:00:01
   ---------------- ----------------------- 2.0/4.7 MB 3.1 MB/s eta 0:00:01
   ----------------- ---------------------- 2.1/4.7 MB 3.1 MB/s eta 0:00:01
   ------------------- -------------------- 2.3/4.7 MB 3.1 MB/s eta 0:00:01
   -------------------- ------------------- 2.5/4.7 MB 3.1 MB/s eta 0:00:01
   ---------------------- ----------------- 2.6/4.7 MB 3.1 MB/s eta 0:00:01
   ----------------------- ---------------- 2.7/4.7 MB 3.1 MB/s eta 0:00:01
   ------------------------ --------------- 2.9/4.7 MB 3.1 MB/s eta 0:00:01
   ------------------------- -------------- 3.1/4.7 MB 3.1 MB/s eta 0:00:01
   --------------------------- ------------ 3.2/4.7 MB 3.1 MB/s eta 0:00:01
   ---------------------------- ----------- 3.3/4.7 MB 3.1 MB/s eta 0:00:01
   ----------------------------- ---------- 3.5/4.7 MB 3.1 MB/s eta 0:00:01
```

```
   ---------------------------- --------- 3.6/4.7 MB 3.1 MB/s eta 0:00:01
   ------------------------------ ------- 3.8/4.7 MB 3.1 MB/s eta 0:00:01
   -------------------------------- ------ 3.9/4.7 MB 3.1 MB/s eta 0:00:01
   ---------------------------------- ----- 4.1/4.7 MB 3.1 MB/s eta 0:00:01
   ----------------------------------- ---- 4.2/4.7 MB 3.1 MB/s eta 0:00:01
   ------------------------------------ --- 4.4/4.7 MB 3.1 MB/s eta 0:00:01
   ------------------------------------- -- 4.5/4.7 MB 3.1 MB/s eta 0:00:01
   -------------------------------------  4.7/4.7 MB 3.1 MB/s eta 0:00:01
   -------------------------------------  4.7/4.7 MB 3.1 MB/s eta 0:00:01
   -------------------------------------- 4.7/4.7 MB 3.0 MB/s eta 0:00:00
Downloading pywavelets-1.5.0-cp311-cp311-win_amd64.whl (4.3 MB)
   ---------------------------------------- 0.0/4.3 MB ? eta -:--:--
   - -------------------------------------- 0.1/4.3 MB 3.2 MB/s eta 0:00:02
   -- ------------------------------------- 0.3/4.3 MB 3.2 MB/s eta 0:00:02
   --- ------------------------------------ 0.4/4.3 MB 3.2 MB/s eta 0:00:02
   ----- ---------------------------------- 0.6/4.3 MB 3.2 MB/s eta 0:00:02
   ------ --------------------------------- 0.7/4.3 MB 3.2 MB/s eta 0:00:02
   ------- -------------------------------- 0.8/4.3 MB 3.1 MB/s eta 0:00:02
   --------- ------------------------------ 1.0/4.3 MB 3.1 MB/s eta 0:00:02
   ---------- ----------------------------- 1.1/4.3 MB 3.1 MB/s eta 0:00:01
   ------------ --------------------------- 1.3/4.3 MB 3.1 MB/s eta 0:00:01
   ------------- -------------------------- 1.4/4.3 MB 3.1 MB/s eta 0:00:01
   --------------- ------------------------ 1.6/4.3 MB 3.1 MB/s eta 0:00:01
   ---------------- ----------------------- 1.8/4.3 MB 3.1 MB/s eta 0:00:01
   ------------------ --------------------- 1.9/4.3 MB 3.2 MB/s eta 0:00:01
   -------------------- ------------------- 2.1/4.3 MB 3.1 MB/s eta 0:00:01
   --------------------- ------------------ 2.2/4.3 MB 3.1 MB/s eta 0:00:01
   ----------------------- ---------------- 2.4/4.3 MB 3.1 MB/s eta 0:00:01
   ------------------------ --------------- 2.5/4.3 MB 3.2 MB/s eta 0:00:01
   -------------------------- ------------- 2.7/4.3 MB 3.1 MB/s eta 0:00:01
   --------------------------- ------------ 2.8/4.3 MB 3.1 MB/s eta 0:00:01
   ---------------------------- ----------- 2.9/4.3 MB 3.1 MB/s eta 0:00:01
   ----------------------------- ---------- 3.1/4.3 MB 3.1 MB/s eta 0:00:01
   ------------------------------ --------- 3.2/4.3 MB 3.1 MB/s eta 0:00:01
   -------------------------------- ------- 3.4/4.3 MB 3.1 MB/s eta 0:00:01
   --------------------------------- ------ 3.6/4.3 MB 3.2 MB/s eta 0:00:01
   ---------------------------------- ----- 3.8/4.3 MB 3.2 MB/s eta 0:00:01
   ----------------------------------- ---- 3.9/4.3 MB 3.1 MB/s eta 0:00:01
   ------------------------------------ --- 4.0/4.3 MB 3.1 MB/s eta 0:00:01
   -------------------------------------  4.2/4.3 MB 3.2 MB/s eta 0:00:01
   -------------------------------------  4.2/4.3 MB 3.2 MB/s eta 0:00:01
   ---------------------------------------- 4.3/4.3 MB 3.0 MB/s eta 0:00:00
Building wheels for collected packages: htmlmin
```

```
  Building wheel for htmlmin (pyproject.toml): started
  Building wheel for htmlmin (pyproject.toml): finished with status 'done'
  Created wheel for htmlmin: filename=htmlmin-0.1.12-py3-none-any.whl size=27092 sha256=35968
  Stored in directory: c:\users\khurana_kunal\appdata\local\pip\cache\wheels\8d\55\1a\19cd535
Successfully built htmlmin
Installing collected packages: htmlmin, typeguard, tangled-up-in-unicode, pydantic-core, nump
  Attempting uninstall: numpy
    Found existing installation: numpy 1.26.4
    Uninstalling numpy-1.26.4:
      Successfully uninstalled numpy-1.26.4
  Attempting uninstall: llvmlite
    Found existing installation: llvmlite 0.42.0
    Uninstalling llvmlite-0.42.0:
      Successfully uninstalled llvmlite-0.42.0
  Attempting uninstall: scipy
    Found existing installation: scipy 1.12.0
    Uninstalling scipy-1.12.0:
      Successfully uninstalled scipy-1.12.0
  Attempting uninstall: numba
    Found existing installation: numba 0.59.0
    Uninstalling numba-0.59.0:
      Successfully uninstalled numba-0.59.0
  Attempting uninstall: seaborn
    Found existing installation: seaborn 0.13.2
    Uninstalling seaborn-0.13.2:
      Successfully uninstalled seaborn-0.13.2
Successfully installed PyWavelets-1.5.0 annotated-types-0.6.0 dacite-1.8.1 htmlmin-0.1.12 ima
```

1. Libraries

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
import warnings as wrn

wrn.filterwarnings('ignore', category = DeprecationWarning)
wrn.filterwarnings('ignore', category = FutureWarning)
wrn.filterwarnings('ignore', category = UserWarning)
#from pandas_profiling import ProfileReport
```

Context

1. Invoice ID: A unique identifier for each invoice or transaction.
2. Branch: The branch or location where the transaction occurred.
3. City: The city where the branch is located.
4. Customer Type: Indicates whether the customer is a regular or new customer.
5. Gender: The gender of the customer.
6. Product Line: The category or type of product purchased.
7. Unit Price: The price of a single unit of the product.
8. Quantity: The number of units of the product purchased.
9. Tax 5%: The amount of tax (5% of the total cost) applied to the transaction.
10. Total: The total cost of the transaction, including tax.
11. Date: The date when the transaction took place.
12. Time: The time of day when the transaction occurred.
13. Payment: The payment method used (e.g., credit card, cash).
14. COGS (Cost of Goods Sold): The direct costs associated with producing or purchasing the products sold.

15. Gross Margin Percentage: The profit margin percentage for the transaction.
16. Gross Income: The total profit earned from the transaction.
17. Rating: Customer satisfaction rating or feedback on the transaction.

For instance, if you were interested in predicting customer satisfaction, Rating might be a suitable label. If you were trying to predict sales or revenue, Total or Gross Income could be a potential label.

2. Initial Data Exploration

```
df = pd.read_csv("/kaggle/input/super-market-sales/supermarket_sales.csv")
```

```
df.head(10)
```

	Invoice ID	Branch	City	Customer type	Gender	Product line	Unit price	Quant
0	750-67-8428	A	Yangon	Member	Female	Health and beauty	74.69	7
1	226-31-3081	C	Naypyitaw	Normal	Female	Electronic accessories	15.28	5
2	631-41-3108	A	Yangon	Normal	Male	Home and lifestyle	46.33	7
3	123-19-1176	A	Yangon	Member	Male	Health and beauty	58.22	8
4	373-73-7910	A	Yangon	Normal	Male	Sports and travel	86.31	7
5	699-14-3026	C	Naypyitaw	Normal	Male	Electronic accessories	85.39	7
6	355-53-5943	A	Yangon	Member	Female	Electronic accessories	68.84	6
7	315-22-5665	C	Naypyitaw	Normal	Female	Home and lifestyle	73.56	10
8	665-32-9167	A	Yangon	Member	Female	Health and beauty	36.26	2
9	692-92-5582	B	Mandalay	Member	Female	Food and beverages	54.84	3

```
df.columns
```

```
Index(['Invoice ID', 'Branch', 'City', 'Customer type', 'Gender',
       'Product line', 'Unit price', 'Quantity', 'Tax 5%', 'Total', 'Date',
       'Time', 'Payment', 'cogs', 'gross margin percentage', 'gross income',
       'Rating'],
      dtype='object')
```

```
df.dtypes
```

```
Invoice ID                 object
Branch                     object
City                       object
Customer type              object
```

```
Gender                      object
Product line                object
Unit price                 float64
Quantity                     int64
Tax 5%                     float64
Total                      float64
Date                        object
Time                        object
Payment                     object
cogs                       float64
gross margin percentage    float64
gross income               float64
Rating                     float64
dtype: object
```

```
df['Date'] = pd.to_datetime(df['Date'])
```

```
df.dtypes
```

```
Invoice ID                        object
Branch                            object
City                              object
Customer type                     object
Gender                            object
Product line                      object
Unit price                       float64
Quantity                           int64
Tax 5%                           float64
Total                            float64
Date                      datetime64[ns]
Time                              object
Payment                           object
cogs                             float64
gross margin percentage          float64
gross income                     float64
Rating                           float64
dtype: object
```

```
df.set_index("Date", inplace=True)
```

```
df.describe()
```

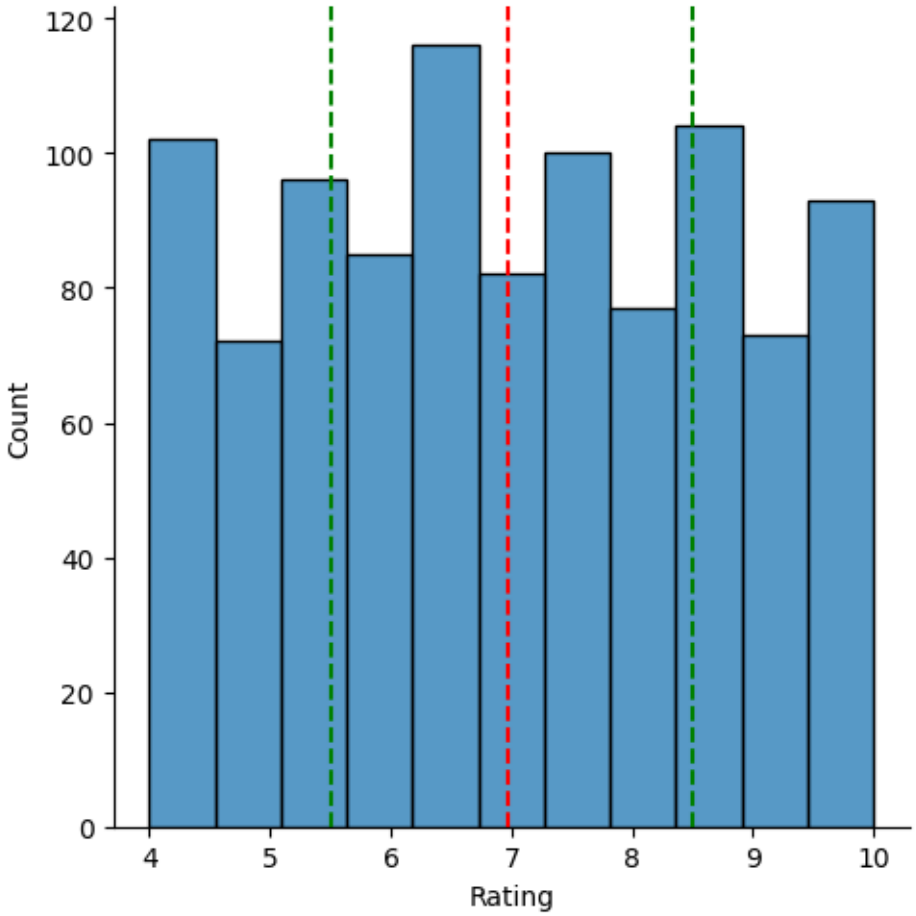
	Unit price	Quantity	Tax 5%	Total	cogs	gross margin percentage	gross
count	1000.000000	1000.000000	1000.000000	1000.000000	1000.00000	1000.000000	1000
mean	55.672130	5.510000	15.379369	322.966749	307.58738	4.761905	15.3
std	26.494628	2.923431	11.708825	245.885335	234.17651	0.000000	11.7
min	10.080000	1.000000	0.508500	10.678500	10.17000	4.761905	0.508
25%	32.875000	3.000000	5.924875	124.422375	118.49750	4.761905	5.92
50%	55.230000	5.000000	12.088000	253.848000	241.76000	4.761905	12.08
75%	77.935000	8.000000	22.445250	471.350250	448.90500	4.761905	22.4
max	99.960000	10.000000	49.650000	1042.650000	993.00000	4.761905	49.6

3. Univariate Analysis

Q1 What does the disribution of customer rating looks like? Is it skewed?

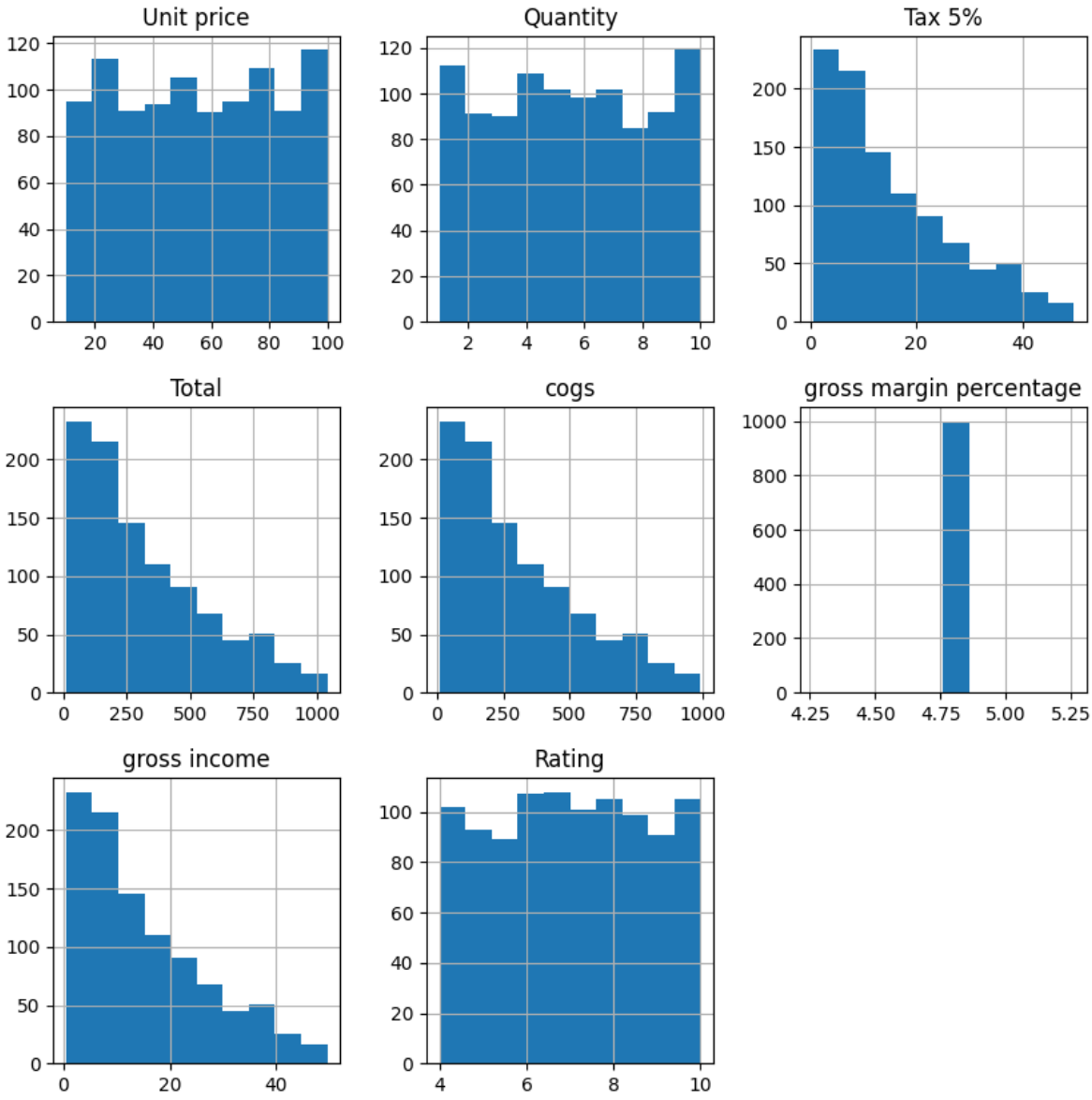
```
sns.displot(df["Rating"])
plt.axvline(x=np.mean(df["Rating"]), c='red', ls= "--")
plt.axvline(x=np.percentile(df["Rating"],25), c='green', ls= "--")
plt.axvline(x=np.percentile(df["Rating"],75), c='green', ls= "--")
```

```
<matplotlib.lines.Line2D at 0x7fa762ae94b0>
```




```
df.hist(figsize=(10,10))
```

```
array([[<Axes: title={'center': 'Unit price'}>,
        <Axes: title={'center': 'Quantity'}>,
        <Axes: title={'center': 'Tax 5%'}>],
       [<Axes: title={'center': 'Total'}>,
        <Axes: title={'center': 'cogs'}>,
        <Axes: title={'center': 'gross margin percentage'}>],
       [<Axes: title={'center': 'gross income'}>,
        <Axes: title={'center': 'Rating'}>, <Axes: >]], dtype=object)
```




```
df['Branch'].value_counts()
```

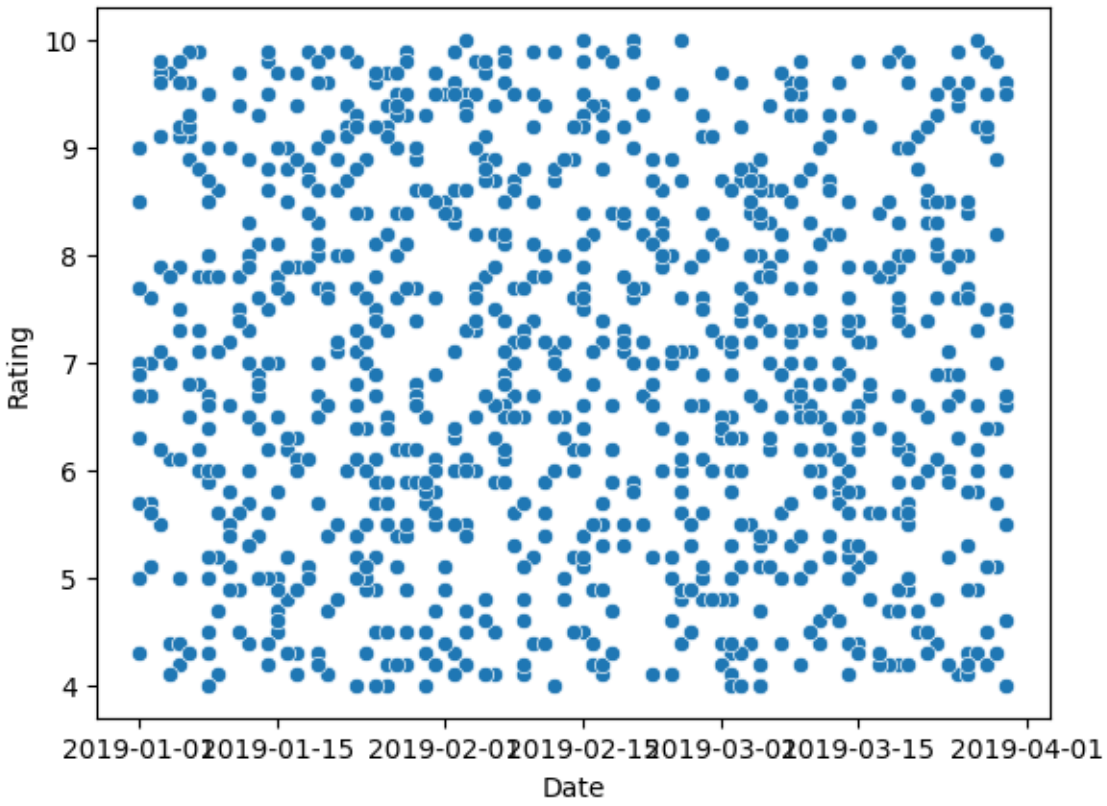
```
Branch
A    340
B    332
C    328
Name: count, dtype: int64
```

4. Bivariate analysis

```
#sns.countplot(df['Payment'])
```

```
# comparison between two columns
sns.scatterplot(df['Rating'])
```

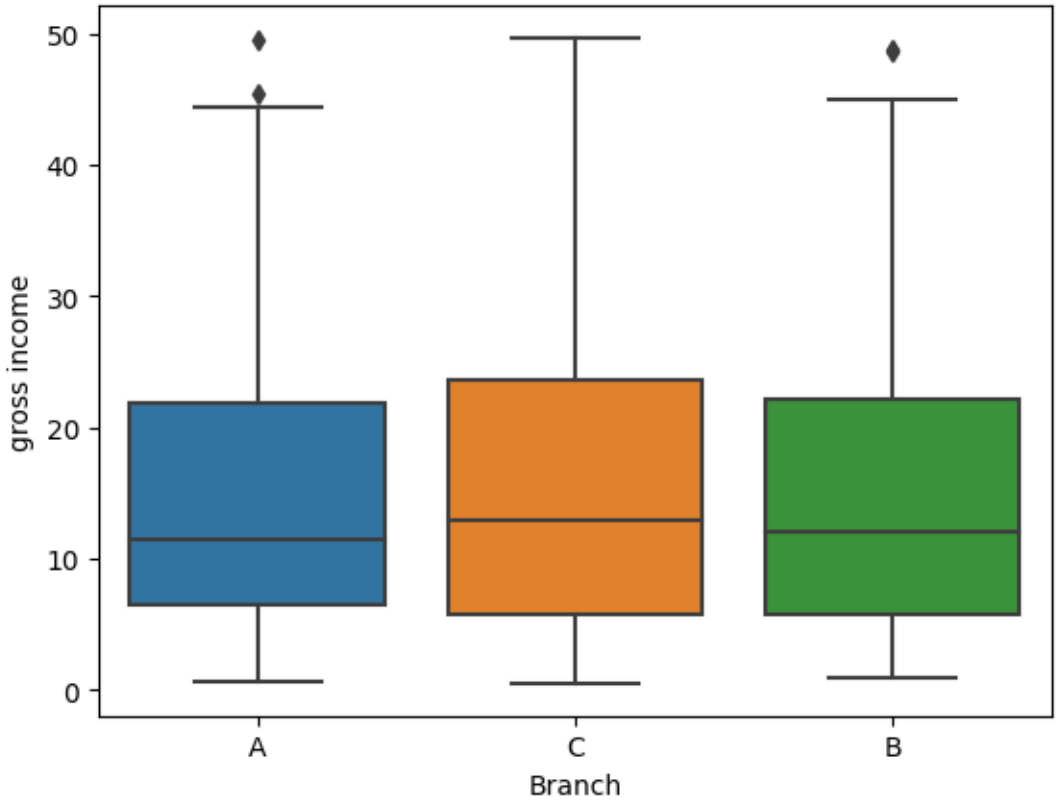
```
<Axes: xlabel='Date', ylabel='Rating'>
```



Q2: is there a noticiable time trend in gross income?

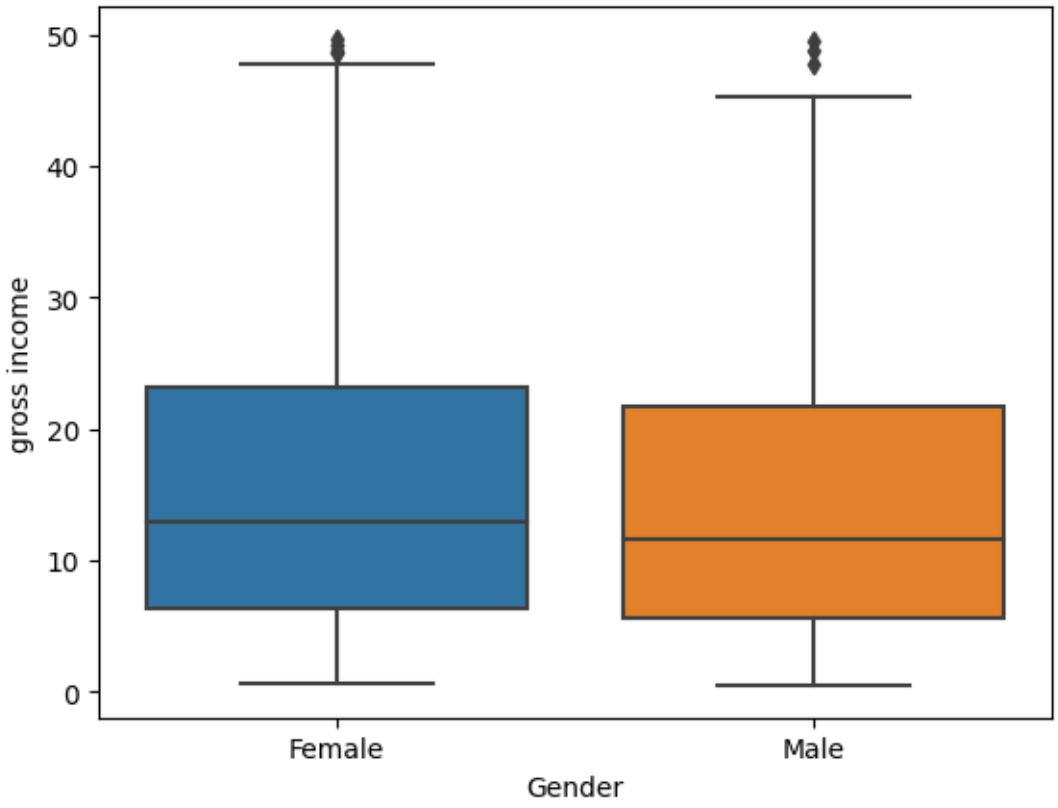
```
sns.boxplot(df, x='Branch', y='gross income')
```

```
<Axes: xlabel='Branch', ylabel='gross income'>
```




```
sns.boxplot(df, x="Gender", y="gross income")
```

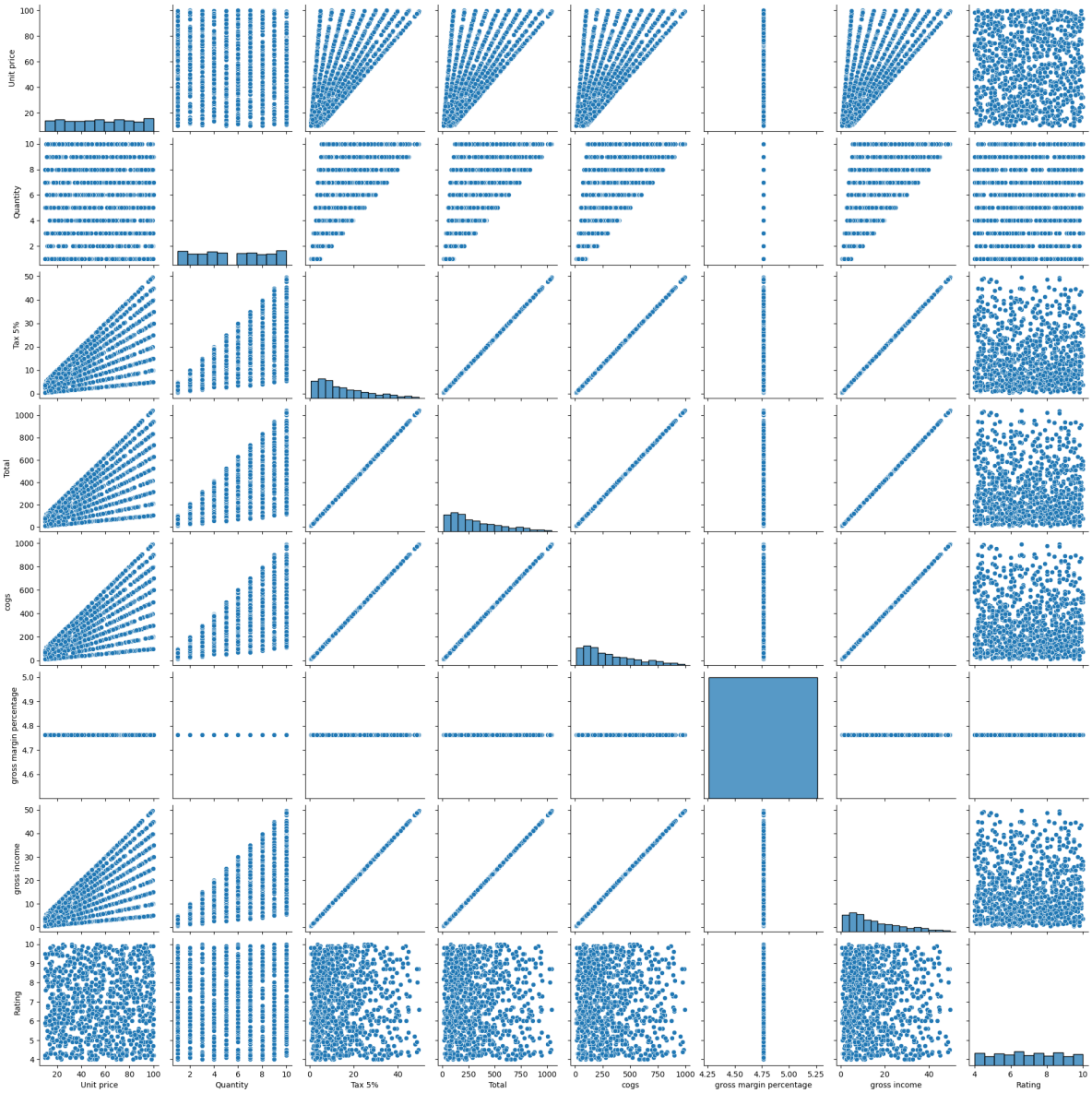
```
<Axes: xlabel='Gender', ylabel='gross income'>
```




```
df.groupby(by='gross income')
```

```
<pandas.core.groupby.generic.DataFrameGroupBy object at 0x7fa75e5eb910>
```

```
sns.pairplot(df)
```


Unit price
Quantity
Tax 5%
Total
cogs
gross margin percentage
gross income
Rating

5. Dealing with duplicate rows and missing values

```
df.duplicated()
```

```
Date
2019-01-05    False
2019-03-08    False
2019-03-03    False
2019-01-27    False
2019-02-08    False
              ...
2019-01-29    False
2019-03-02    False
2019-02-09    False
2019-02-22    False
2019-02-18    False
Length: 1000, dtype: bool
```

```
df.duplicated().sum()
```

```
0
```

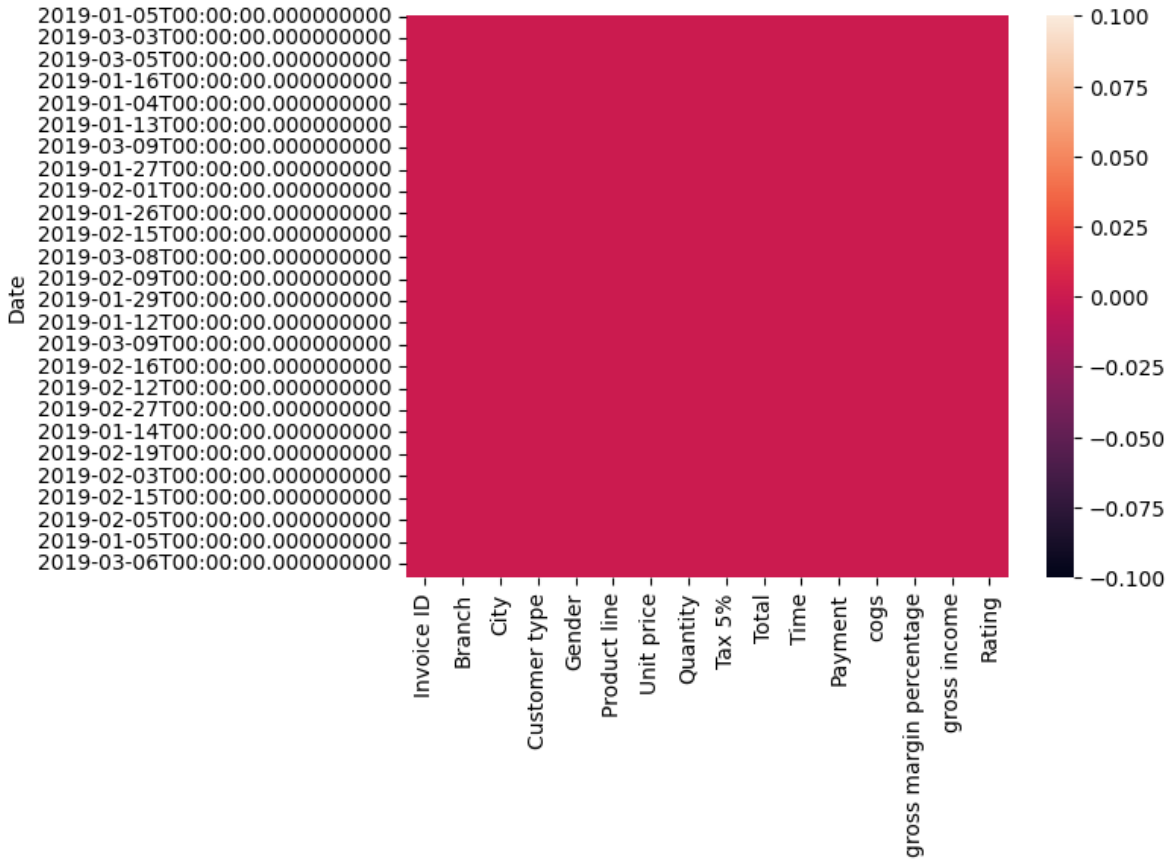
```
df.isna().sum()
```

```
Invoice ID                0
Branch                    0
City                      0
Customer type             0
Gender                    0
Product line              0
Unit price                0
```

```
Quantity                  0
Tax 5%                    0
Total                     0
Time                      0
Payment                   0
cogs                      0
gross margin percentage   0
gross income              0
Rating                    0
dtype: int64
```

```
sns.heatmap(df.isnull())
```

```
<Axes: ylabel='Date'>
```



```
df.mode()
```

	Invoice ID	Branch	City	Customer type	Gender	Product line	Unit price	Quantity
0	101-17-6199	A	Yangon	Member	Female	Fashion accessories	83.77	10.0
1	101-81-4070	NaN	NaN	NaN	NaN	NaN	NaN	NaN
2	102-06-2002	NaN	NaN	NaN	NaN	NaN	NaN	NaN
3	102-77-2261	NaN	NaN	NaN	NaN	NaN	NaN	NaN
4	105-10-6182	NaN	NaN	NaN	NaN	NaN	NaN	NaN
...	...	...	...	...	...	...	...	...
995	894-41-5205	NaN	NaN	NaN	NaN	NaN	NaN	NaN
996	895-03-6665	NaN	NaN	NaN	NaN	NaN	NaN	NaN
997	895-66-0685	NaN	NaN	NaN	NaN	NaN	NaN	NaN
998	896-34-0956	NaN	NaN	NaN	NaN	NaN	NaN	NaN
999	898-04-2717	NaN	NaN	NaN	NaN	NaN	NaN	NaN

```
df.mode().iloc[0]
```

```
Invoice ID                  101-17-6199
Branch                                A
City                             Yangon
Customer type                    Member
Gender                           Female
Product line       Fashion accessories
Unit price                       83.77
Quantity                          10.0
Tax 5%                           4.154
Total                           87.234
Time                             14:42
Payment                        Ewallet
cogs                             83.08
gross margin percentage       4.761905
gross income                     4.154
Rating                             6.0
Name: 0, dtype: object
```

6. Correlation analysis

```
np.corrcoef(df["gross income"], df['Rating'])
```

```
array([[ 1.       , -0.0364417],
       [-0.0364417,  1.       ]])
```

```
np.corrcoef(df["gross income"], df['Rating'])[1][0]
```

```
-0.03644170499701839
```

```
# rounding off
round(np.corrcoef(df['gross income'], df['Rating'])[1][0],2)
```

```
-0.04
```

7. Profiling

```
dataset = pd.read_csv("/kaggle/input/super-market-sales/supermarket_sales.csv")

from ydata_profiling import ProfileReport
profile = ProfileReport(dataset, title='Profiling Report')
profile
```

```
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]

<IPython.core.display.HTML object>
```

8. Resources

1. https://www.data-to-viz.com/
2. https://seaborn.pydata.org/examples/index.html
3. https://pypi.org/project/pandas-profiling/