

```
In [31]: #First of all we have to install all libraries needed.
!pip install numpy
```

Requirement already satisfied: numpy in /Users/reanadalasmari/anaconda3/envs/learningEnv/lib/python3.10/site-packages (1.26.4)

```
In [2]: pip install scikit-learn
```

Collecting scikit-learn

Downloading scikit_learn-1.7.2-cp310-cp310-macosx_10_9_x86_64.whl.metadata (11 kB)

Requirement already satisfied: numpy>=1.22.0 in /Users/reanadalasmari/anaconda3/envs/learningEnv/lib/python3.10/site-packages (from scikit-learn) (2.2.6)

Collecting scipy>=1.8.0 (from scikit-learn)

Downloading scipy-1.15.3-cp310-cp310-macosx_14_0_x86_64.whl.metadata (61 kB)

Collecting joblib>=1.2.0 (from scikit-learn)

Downloading joblib-1.5.3-py3-none-any.whl.metadata (5.5 kB)

Collecting threadpoolctl>=3.1.0 (from scikit-learn)

Downloading threadpoolctl-3.6.0-py3-none-any.whl.metadata (13 kB)

Downloading scikit_learn-1.7.2-cp310-cp310-macosx_10_9_x86_64.whl (9.3 MB)

9.3/9.3 MB 13.0 MB/s 0:00:00
10.6 MB/s eta 0:00:01

Downloading joblib-1.5.3-py3-none-any.whl (309 kB)

Downloading scipy-1.15.3-cp310-cp310-macosx_14_0_x86_64.whl (25.1 MB)

25.1/25.1 MB 19.2 MB/s 0:00:00
1 eta 0:00:01 [36m0:00:01

Downloading threadpoolctl-3.6.0-py3-none-any.whl (18 kB)

Installing collected packages: threadpoolctl, scipy, joblib, scikit-learn
4/4 [scikit-learn]0m 3/4 [scikit-learn]

Successfully installed joblib-1.5.3 scikit-learn-1.7.2 scipy-1.15.3 threadpoolctl-3.6.0

Note: you may need to restart the kernel to use updated packages.

```
In [3]: pip install pandas
```

Collecting pandas

Downloading pandas-2.3.3-cp310-cp310-macosx_10_9_x86_64.whl.metadata (91 kB)

Requirement already satisfied: numpy>=1.22.4 in /Users/reanadalasmari/anaconda3/envs/learningEnv/lib/python3.10/site-packages (from pandas) (2.2.6)

Requirement already satisfied: python-dateutil>=2.8.2 in /Users/reanadalasmari/anaconda3/envs/learningEnv/lib/python3.10/site-packages (from pandas) (2.9.0.post0)

Collecting pytz>=2020.1 (from pandas)

Downloading pytz-2026.1.post1-py2.py3-none-any.whl.metadata (22 kB)

Requirement already satisfied: tzdata>=2022.7 in /Users/reanadalasmari/anaconda3/envs/learningEnv/lib/python3.10/site-packages (from pandas) (2025.3)

Requirement already satisfied: six>=1.5 in /Users/reanadalasmari/anaconda3/envs/learningEnv/lib/python3.10/site-packages (from python-dateutil>=2.8.2->pandas) (1.17.0)

Downloading pandas-2.3.3-cp310-cp310-macosx_10_9_x86_64.whl (11.6 MB)

11.6/11.6 MB 13.7 MB/s 0:00:0

05.1 MB/s eta 0:00:01

Downloading pytz-2026.1.post1-py2.py3-none-any.whl (510 kB)

Installing collected packages: pytz, pandas

2/2 [pandas] — 1/2 [pandas]

Successfully installed pandas-2.3.3 pytz-2026.1.post1

Note: you may need to restart the kernel to use updated packages.

In [5]: `pip install tensorflow`

```

Collecting tensorflow
  Downloading tensorflow-2.16.2-cp310-cp310-macosx_10_15_x86_64.whl.metad
ata (4.1 kB)
Collecting absl-py>=1.0.0 (from tensorflow)
  Downloading absl_py-2.4.0-py3-none-any.whl.metadata (3.3 kB)
Collecting astunparse>=1.6.0 (from tensorflow)
  Downloading astunparse-1.6.3-py2.py3-none-any.whl.metadata (4.4 kB)
Collecting flatbuffers>=23.5.26 (from tensorflow)
  Downloading flatbuffers-25.12.19-py2.py3-none-any.whl.metadata (1.0 kB)
Collecting gast!=0.5.0,!0.5.1,!0.5.2,>=0.2.1 (from tensorflow)
  Downloading gast-0.7.0-py3-none-any.whl.metadata (1.5 kB)
Collecting google-pasta>=0.1.1 (from tensorflow)
  Downloading google_pasta-0.2.0-py3-none-any.whl.metadata (814 bytes)
Collecting h5py>=3.10.0 (from tensorflow)
  Downloading h5py-3.16.0-cp310-cp310-macosx_10_9_x86_64.whl.metadata (3.0
kB)
Collecting libclang>=13.0.0 (from tensorflow)
  Downloading libclang-18.1.1-py2.py3-none-macosx_10_9_x86_64.whl.metadata
(5.2 kB)
Collecting ml-dtypes~=0.3.1 (from tensorflow)
  Downloading ml_dtypes-0.3.2-cp310-cp310-macosx_10_9_universal2.whl.metad
ata (20 kB)
Collecting opt-einsum>=2.3.2 (from tensorflow)
  Downloading opt_einsum-3.4.0-py3-none-any.whl.metadata (6.3 kB)
Requirement already satisfied: packaging in /Users/reanadalasmari/anaconda
3/envs/learningEnv/lib/python3.10/site-packages (from tensorflow) (26.0)
Collecting protobuf!=4.21.0,!4.21.1,!4.21.2,!4.21.3,!4.21.4,!4.21.5,<
5.0.0dev,>=3.20.3 (from tensorflow)
  Downloading protobuf-4.25.8-cp37-abi3-macosx_10_9_universal2.whl.metadat
a (541 bytes)
Requirement already satisfied: requests<3,>=2.21.0 in /Users/reanadalasmari/anaconda3/envs/learningEnv/lib/python3.10/site-packages (from tensorflo
w) (2.32.5)
Requirement already satisfied: setuptools in /Users/reanadalasmari/anacond
a3/envs/learningEnv/lib/python3.10/site-packages (from tensorflow) (82.0.
0)
Requirement already satisfied: six>=1.12.0 in /Users/reanadalasmari/anacon
da3/envs/learningEnv/lib/python3.10/site-packages (from tensorflow) (1.17.
0)
Collecting termcolor>=1.1.0 (from tensorflow)
  Downloading termcolor-3.3.0-py3-none-any.whl.metadata (6.5 kB)
Requirement already satisfied: typing-extensions>=3.6.6 in /Users/reanadal
asmari/anaconda3/envs/learningEnv/lib/python3.10/site-packages (from tenso
rflow) (4.15.0)
Collecting wrapt>=1.11.0 (from tensorflow)
  Downloading wrapt-2.1.2-cp310-cp310-macosx_10_9_x86_64.whl.metadata (7.4
kB)
Collecting grpcio<2.0,>=1.24.3 (from tensorflow)
  Downloading grpcio-1.78.0-cp310-cp310-macosx_11_0_universal2.whl.metadat
a (3.8 kB)
Collecting tensorboard<2.17,>=2.16 (from tensorflow)
  Downloading tensorboard-2.16.2-py3-none-any.whl.metadata (1.6 kB)
Collecting keras>=3.0.0 (from tensorflow)
  Downloading keras-3.12.1-py3-none-any.whl.metadata (5.9 kB)
Collecting tensorflow-io-gcs-filesystem>=0.23.1 (from tensorflow)
  Downloading tensorflow_io_gcs_filesystem-0.37.1-cp310-cp310-macosx_10_14
_x86_64.whl.metadata (14 kB)
Collecting numpy<2.0.0,>=1.23.5 (from tensorflow)
  Downloading numpy-1.26.4-cp310-cp310-macosx_10_9_x86_64.whl.metadata (61
kB)

```

```

Requirement already satisfied: charset_normalizer<4,>=2 in /Users/reanadalasmari/anaconda3/envs/learningEnv/lib/python3.10/site-packages (from requests<3,>=2.21.0->tensorflow) (3.4.5)
Requirement already satisfied: idna<4,>=2.5 in /Users/reanadalasmari/anaconda3/envs/learningEnv/lib/python3.10/site-packages (from requests<3,>=2.21.0->tensorflow) (3.11)
Requirement already satisfied: urllib3<3,>=1.21.1 in /Users/reanadalasmari/anaconda3/envs/learningEnv/lib/python3.10/site-packages (from requests<3,>=2.21.0->tensorflow) (2.6.3)
Requirement already satisfied: certifi>=2017.4.17 in /Users/reanadalasmari/anaconda3/envs/learningEnv/lib/python3.10/site-packages (from requests<3,>=2.21.0->tensorflow) (2026.2.25)
Collecting markdown>=2.6.8 (from tensorboard<2.17,>=2.16->tensorflow)
  Downloading markdown-3.10.2-py3-none-any.whl.metadata (5.1 kB)
Collecting tensorboard-data-server<0.8.0,>=0.7.0 (from tensorboard<2.17,>=2.16->tensorflow)
  Downloading tensorboard_data_server-0.7.2-py3-none-macosx_10_9_x86_64.whl.metadata (1.1 kB)
Collecting werkzeug>=1.0.1 (from tensorboard<2.17,>=2.16->tensorflow)
  Downloading werkzeug-3.1.6-py3-none-any.whl.metadata (4.0 kB)
Requirement already satisfied: wheel<1.0,>=0.23.0 in /Users/reanadalasmari/anaconda3/envs/learningEnv/lib/python3.10/site-packages (from astunparse>=1.6.0->tensorflow) (0.46.3)
Collecting rich (from keras>=3.0.0->tensorflow)
  Downloading rich-14.3.3-py3-none-any.whl.metadata (18 kB)
Collecting namex (from keras>=3.0.0->tensorflow)
  Downloading namex-0.1.0-py3-none-any.whl.metadata (322 bytes)
Collecting optree (from keras>=3.0.0->tensorflow)
  Downloading optree-0.19.0-cp310-cp310-macosx_10_9_x86_64.whl.metadata (34 kB)
Requirement already satisfied: markupsafe>=2.1.1 in /Users/reanadalasmari/anaconda3/envs/learningEnv/lib/python3.10/site-packages (from werkzeug>=1.0.1->tensorboard<2.17,>=2.16->tensorflow) (3.0.3)
Collecting markdown-it-py>=2.2.0 (from rich->keras>=3.0.0->tensorflow)
  Downloading markdown_it_py-4.0.0-py3-none-any.whl.metadata (7.3 kB)
Requirement already satisfied: pygments<3.0.0,>=2.13.0 in /Users/reanadalasmari/anaconda3/envs/learningEnv/lib/python3.10/site-packages (from rich->keras>=3.0.0->tensorflow) (2.19.2)
Collecting mdurl~=0.1 (from markdown-it-py>=2.2.0->rich->keras>=3.0.0->tensorflow)
  Downloading mdurl-0.1.2-py3-none-any.whl.metadata (1.6 kB)
Downloading tensorflow-2.16.2-cp310-cp310-macosx_10_15_x86_64.whl (259.5 MB)
_____ 259.5/259.5 MB 4.5 MB/s 0:00:45 eta 0:00:01[36m0:00:02
Downloading grpcio-1.78.0-cp310-cp310-macosx_11_0_universal2.whl (11.8 MB)
_____ 11.8/11.8 MB 16.5 MB/s 0:00:00
07.7 MB/s eta 0:00:01
Downloading ml_dtypes-0.3.2-cp310-cp310-macosx_10_9_universal2.whl (389 kB)
Downloading numpy-1.26.4-cp310-cp310-macosx_10_9_x86_64.whl (20.6 MB)
_____ 20.6/20.6 MB 14.3 MB/s 0:00:01
15.7 MB/s eta 0:00:01
Downloading protobuf-4.25.8-cp37-abi3-macosx_10_9_universal2.whl (394 kB)
Downloading tensorboard-2.16.2-py3-none-any.whl (5.5 MB)
_____ 5.5/5.5 MB 14.0 MB/s 0:00:00
19.6 MB/s eta 0:00:01
Downloading tensorboard_data_server-0.7.2-py3-none-macosx_10_9_x86_64.whl (4.8 MB)
_____ 4.8/4.8 MB 20.3 MB/s 0:00:00

```

```

21.9 MB/s eta 0:00:01
Downloading absl_py-2.4.0-py3-none-any.whl (135 kB)
Downloading astunparse-1.6.3-py2.py3-none-any.whl (12 kB)
Downloading flatbuffers-25.12.19-py2.py3-none-any.whl (26 kB)
Downloading gast-0.7.0-py3-none-any.whl (22 kB)
Downloading google_pasta-0.2.0-py3-none-any.whl (57 kB)
Downloading h5py-3.16.0-cp310-cp310-macosx_10_9_x86_64.whl (3.7 MB)
_____ 3.7/3.7 MB 20.1 MB/s 0:00:00
Downloading keras-3.12.1-py3-none-any.whl (1.5 MB)
_____ 1.5/1.5 MB 27.8 MB/s 0:00:00
Downloading libclang-18.1.1-py2.py3-none-macosx_10_9_x86_64.whl (26.5 MB)
_____ 26.5/26.5 MB 19.4 MB/s 0:00:00
19.4 MB/s eta 0:00:01:01
Downloading markdown-3.10.2-py3-none-any.whl (108 kB)
Downloading opt_einsum-3.4.0-py3-none-any.whl (71 kB)
Downloading tensorflow_io_gcs_filesystem-0.37.1-cp310-cp310-macosx_10_14_x
86_64.whl (2.5 MB)
_____ 2.5/2.5 MB 13.7 MB/s 0:00:00
Downloading termcolor-3.3.0-py3-none-any.whl (7.7 kB)
Downloading werkzeug-3.1.6-py3-none-any.whl (225 kB)
Downloading wrapt-2.1.2-cp310-cp310-macosx_10_9_x86_64.whl (60 kB)
Downloading namex-0.1.0-py3-none-any.whl (5.9 kB)
Downloading optree-0.19.0-cp310-cp310-macosx_10_9_x86_64.whl (392 kB)
Downloading rich-14.3.3-py3-none-any.whl (310 kB)
Downloading markdown_it_py-4.0.0-py3-none-any.whl (87 kB)
Downloading mdurl-0.1.2-py3-none-any.whl (10.0 kB)
Installing collected packages: namex, libclang, flatbuffers, wrapt, werkze
ug, termcolor, tensorflow-io-gcs-filesystem, tensorboard-data-server, prot
obuf, optree, opt-einsum, numpy, mdurl, markdown, grpcio, google-pasta, ga
st, absl-py, tensorboard, ml-dtypes, markdown-it-py, h5py, astunparse, ric
h, keras, tensorflow
Attempting uninstall: numpy_____ 8/26 [proto
buf]bclang]
Found existing installation: numpy 2.2.6 _____
8/26 [protobuf]
Uninstalling numpy-2.2.6: _____ 8/26 [protobuf]
Successfully uninstalled numpy-2.2.6[38;2;249;38;114m-_____
_____ 11/26 [numpy]
_____ 26/26 [tensorflow]0m 25/26 [te
nsorflow]s]board]
Successfully installed absl-py-2.4.0 astunparse-1.6.3 flatbuffers-25.12.19
gast-0.7.0 google-pasta-0.2.0 grpcio-1.78.0 h5py-3.16.0 keras-3.12.1 libcl
ang-18.1.1 markdown-3.10.2 markdown-it-py-4.0.0 mdurl-0.1.2 ml-dtypes-0.3.
2 namex-0.1.0 numpy-1.26.4 opt-einsum-3.4.0 optree-0.19.0 protobuf-4.25.8
rich-14.3.3 tensorboard-2.16.2 tensorboard-data-server-0.7.2 tensorflow-2.
16.2 tensorflow-io-gcs-filesystem-0.37.1 termcolor-3.3.0 werkzeug-3.1.6 wr
apt-2.1.2
Note: you may need to restart the kernel to use updated packages.

```

In [6]: `pip install keras`

Requirement already satisfied: keras in /Users/reanadalasmari/anaconda3/envs/learningEnv/lib/python3.10/site-packages (3.12.1)

Requirement already satisfied: absl-py in /Users/reanadalasmari/anaconda3/envs/learningEnv/lib/python3.10/site-packages (from keras) (2.4.0)

Requirement already satisfied: numpy in /Users/reanadalasmari/anaconda3/envs/learningEnv/lib/python3.10/site-packages (from keras) (1.26.4)

Requirement already satisfied: rich in /Users/reanadalasmari/anaconda3/envs/learningEnv/lib/python3.10/site-packages (from keras) (14.3.3)

Requirement already satisfied: namex in /Users/reanadalasmari/anaconda3/envs/learningEnv/lib/python3.10/site-packages (from keras) (0.1.0)

Requirement already satisfied: h5py in /Users/reanadalasmari/anaconda3/envs/learningEnv/lib/python3.10/site-packages (from keras) (3.16.0)

Requirement already satisfied: optree in /Users/reanadalasmari/anaconda3/envs/learningEnv/lib/python3.10/site-packages (from keras) (0.19.0)

Requirement already satisfied: ml-dtypes in /Users/reanadalasmari/anaconda3/envs/learningEnv/lib/python3.10/site-packages (from keras) (0.3.2)

Requirement already satisfied: packaging in /Users/reanadalasmari/anaconda3/envs/learningEnv/lib/python3.10/site-packages (from keras) (26.0)

Requirement already satisfied: typing-extensions>=4.6.0 in /Users/reanadalasmari/anaconda3/envs/learningEnv/lib/python3.10/site-packages (from optree->keras) (4.15.0)

Requirement already satisfied: markdown-it-py>=2.2.0 in /Users/reanadalasmari/anaconda3/envs/learningEnv/lib/python3.10/site-packages (from rich->keras) (4.0.0)

Requirement already satisfied: pygments<3.0.0,>=2.13.0 in /Users/reanadalasmari/anaconda3/envs/learningEnv/lib/python3.10/site-packages (from rich->keras) (2.19.2)

Requirement already satisfied: mdurl<=0.1 in /Users/reanadalasmari/anaconda3/envs/learningEnv/lib/python3.10/site-packages (from markdown-it-py>=2.2.0->rich->keras) (0.1.2)

Note: you may need to restart the kernel to use updated packages.

In [7]: `pip install xgboost`

Collecting xgboost
 Downloading xgboost-3.2.0-py3-none-macosx_10_15_x86_64.whl.metadata (2.1 kB)
 Requirement already satisfied: numpy in /Users/reanadalasmari/anaconda3/envs/learningEnv/lib/python3.10/site-packages (from xgboost) (1.26.4)
 Requirement already satisfied: scipy in /Users/reanadalasmari/anaconda3/envs/learningEnv/lib/python3.10/site-packages (from xgboost) (1.15.3)
 Downloading xgboost-3.2.0-py3-none-macosx_10_15_x86_64.whl (2.5 MB)

 9 MB/s eta 0:00:01:01
 Installing collected packages: xgboost
 Successfully installed xgboost-3.2.0
 Note: you may need to restart the kernel to use updated packages.

In [8]: `pip install matplotlib`

```

Collecting matplotlib
  Downloading matplotlib-3.10.8-cp310-cp310-macosx_10_12_x86_64.whl.metadata
ta (52 kB)
Collecting contourpy>=1.0.1 (from matplotlib)
  Downloading contourpy-1.3.2-cp310-cp310-macosx_10_9_x86_64.whl.metadata
(5.5 kB)
Collecting cycler>=0.10 (from matplotlib)
  Downloading cycler-0.12.1-py3-none-any.whl.metadata (3.8 kB)
Collecting fonttools>=4.22.0 (from matplotlib)
  Downloading fonttools-4.61.1-cp310-cp310-macosx_10_9_x86_64.whl.metadata
(114 kB)
Collecting kiwisolver>=1.3.1 (from matplotlib)
  Downloading kiwisolver-1.4.9-cp310-cp310-macosx_10_9_x86_64.whl.metadata
(6.3 kB)
Requirement already satisfied: numpy>=1.23 in /Users/reanadalasmari/anacon
da3/envs/learningEnv/lib/python3.10/site-packages (from matplotlib) (1.26.
4)
Requirement already satisfied: packaging>=20.0 in /Users/reanadalasmari/an
aconda3/envs/learningEnv/lib/python3.10/site-packages (from matplotlib) (2
6.0)
Collecting pillow>=8 (from matplotlib)
  Downloading pillow-12.1.1-cp310-cp310-macosx_10_10_x86_64.whl.metadata
(8.8 kB)
Collecting pyparsing>=3 (from matplotlib)
  Downloading pyparsing-3.3.2-py3-none-any.whl.metadata (5.8 kB)
Requirement already satisfied: python-dateutil>=2.7 in /Users/reanadalasma
ri/anaconda3/envs/learningEnv/lib/python3.10/site-packages (from matplotli
b) (2.9.0.post0)
Requirement already satisfied: six>=1.5 in /Users/reanadalasmari/anaconda
3/envs/learningEnv/lib/python3.10/site-packages (from python-dateutil>=2.7
->matplotlib) (1.17.0)
Downloading matplotlib-3.10.8-cp310-cp310-macosx_10_12_x86_64.whl (8.2 MB)
_____ 8.2/8.2 MB 12.4 MB/s 0:00:00
9.3 MB/s eta 0:00:01
Downloading contourpy-1.3.2-cp310-cp310-macosx_10_9_x86_64.whl (268 kB)
Downloading cycler-0.12.1-py3-none-any.whl (8.3 kB)
Downloading fonttools-4.61.1-cp310-cp310-macosx_10_9_x86_64.whl (2.4 MB)
_____ 2.4/2.4 MB 8.8 MB/s 0:00:00m
16.3 MB/s eta 0:00:01
Downloading kiwisolver-1.4.9-cp310-cp310-macosx_10_9_x86_64.whl (66 kB)
Downloading pillow-12.1.1-cp310-cp310-macosx_10_10_x86_64.whl (5.3 MB)
_____ 5.3/5.3 MB 27.5 MB/s 0:00:00
100.4 MB/s eta 0:00:01
Downloading pyparsing-3.3.2-py3-none-any.whl (122 kB)
Installing collected packages: pyparsing, pillow, kiwisolver, fonttools, c
ycler, contourpy, matplotlib
_____ 7/7 [matplotlib] 6/7 [matplotl
ib]ow]
Successfully installed contourpy-1.3.2 cycler-0.12.1 fonttools-4.61.1 kiwi
solver-1.4.9 matplotlib-3.10.8 pillow-12.1.1 pyparsing-3.3.2
Note: you may need to restart the kernel to use updated packages.

```

In [9]: `pip install seaborn`

Collecting seaborn

Downloading seaborn-0.13.2-py3-none-any.whl.metadata (5.4 kB)

Requirement already satisfied: numpy!=1.24.0,>=1.20 in /Users/reanadalasmari/anaconda3/envs/learningEnv/lib/python3.10/site-packages (from seaborn) (1.26.4)

Requirement already satisfied: pandas>=1.2 in /Users/reanadalasmari/anaconda3/envs/learningEnv/lib/python3.10/site-packages (from seaborn) (2.3.3)

Requirement already satisfied: matplotlib!=3.6.1,>=3.4 in /Users/reanadalasmari/anaconda3/envs/learningEnv/lib/python3.10/site-packages (from seaborn) (3.10.8)

Requirement already satisfied: contourpy>=1.0.1 in /Users/reanadalasmari/anaconda3/envs/learningEnv/lib/python3.10/site-packages (from matplotlib!=3.6.1,>=3.4->seaborn) (1.3.2)

Requirement already satisfied: cycler>=0.10 in /Users/reanadalasmari/anaconda3/envs/learningEnv/lib/python3.10/site-packages (from matplotlib!=3.6.1,>=3.4->seaborn) (0.12.1)

Requirement already satisfied: fonttools>=4.22.0 in /Users/reanadalasmari/anaconda3/envs/learningEnv/lib/python3.10/site-packages (from matplotlib!=3.6.1,>=3.4->seaborn) (4.61.1)

Requirement already satisfied: kiwisolver>=1.3.1 in /Users/reanadalasmari/anaconda3/envs/learningEnv/lib/python3.10/site-packages (from matplotlib!=3.6.1,>=3.4->seaborn) (1.4.9)

Requirement already satisfied: packaging>=20.0 in /Users/reanadalasmari/anaconda3/envs/learningEnv/lib/python3.10/site-packages (from matplotlib!=3.6.1,>=3.4->seaborn) (26.0)

Requirement already satisfied: pillow>=8 in /Users/reanadalasmari/anaconda3/envs/learningEnv/lib/python3.10/site-packages (from matplotlib!=3.6.1,>=3.4->seaborn) (12.1.1)

Requirement already satisfied: pyparsing>=3 in /Users/reanadalasmari/anaconda3/envs/learningEnv/lib/python3.10/site-packages (from matplotlib!=3.6.1,>=3.4->seaborn) (3.3.2)

Requirement already satisfied: python-dateutil>=2.7 in /Users/reanadalasmari/anaconda3/envs/learningEnv/lib/python3.10/site-packages (from matplotlib!=3.6.1,>=3.4->seaborn) (2.9.0.post0)

Requirement already satisfied: pytz>=2020.1 in /Users/reanadalasmari/anaconda3/envs/learningEnv/lib/python3.10/site-packages (from pandas>=1.2->seaborn) (2026.1.post1)

Requirement already satisfied: tzdata>=2022.7 in /Users/reanadalasmari/anaconda3/envs/learningEnv/lib/python3.10/site-packages (from pandas>=1.2->seaborn) (2025.3)

Requirement already satisfied: six>=1.5 in /Users/reanadalasmari/anaconda3/envs/learningEnv/lib/python3.10/site-packages (from python-dateutil>=2.7->matplotlib!=3.6.1,>=3.4->seaborn) (1.17.0)

Downloading seaborn-0.13.2-py3-none-any.whl (294 kB)

Installing collected packages: seaborn

Successfully installed seaborn-0.13.2

Note: you may need to restart the kernel to use updated packages.

```
In [66]: #after installation all libraries have completed we will invoke the Panda
import pandas as pd
#Note that the -Stress level- feature is the target value in our data set
df = pd.read_csv('student_productivity_distraction_dataset_20000.csv')
```

```
In [84]: #Here I start exploring my data set.
df.head(10)
```

Out [84]:

	student_id	age	gender	study_hours_per_day	sleep_hours	phone_usage_hour
0	1	23	Female	4.35	3.63	3.3
1	2	20	Male	6.14	6.58	5.4
2	3	29	Female	4.98	3.26	4.8
3	4	27	Female	3.19	4.58	10.0
4	5	24	Male	7.67	6.21	3.0
5	6	29	Other	7.18	3.52	4.0
6	7	21	Female	9.06	6.36	11.4
7	8	23	Female	6.37	4.86	3.3
8	9	26	Male	4.19	4.87	9.6
9	10	19	Female	7.28	9.56	2.1

In [68]: `df.tail()`

Out [68]:

	student_id	age	gender	study_hours_per_day	sleep_hours	phone_usage_hour
19995	19996	26	Female	4.81	7.83	
19996	19997	22	Other	0.83	4.49	
19997	19998	25	Male	9.36	4.33	
19998	19999	22	Female	0.74	4.44	
19999	20000	19	Female	3.64	6.35	

In [69]: `df.describe()`

Out [69]:

	student_id	age	study_hours_per_day	sleep_hours	phone_usage_hour
count	20000.000000	20000.000000	20000.000000	20000.000000	20000.000000
mean	10000.500000	23.00745	5.254562	6.517799	6.517799
std	5773.647028	3.75489	2.742876	2.029784	2.029784
min	1.000000	17.00000	0.500000	3.000000	3.000000
25%	5000.750000	20.00000	2.900000	4.770000	4.770000
50%	10000.500000	23.00000	5.250000	6.510000	6.510000
75%	15000.250000	26.00000	7.640000	8.310000	8.310000
max	20000.000000	29.00000	10.000000	10.000000	10.000000

In [70]: `len(df)`

Out[70]: 20000

In [71]: `df.loc[0]`

Out[71]:

student_id	1
age	23
gender	Female
study_hours_per_day	4.35
sleep_hours	3.63
phone_usage_hours	3.38
social_media_hours	2.73
youtube_hours	1.83
gaming_hours	5.26
breaks_per_day	6
coffee_intake_mg	347
exercise_minutes	111
assignments_completed	2
attendance_percentage	57.21
stress_level	10
focus_score	57
final_grade	81.87
productivity_score	33.78

Name: 0, dtype: object

In [72]: *#In this step we are checking if any missing values occurs in our data so*
`df.isnull().sum()`

Out[72]:

student_id	0
age	0
gender	0
study_hours_per_day	0
sleep_hours	0
phone_usage_hours	0
social_media_hours	0
youtube_hours	0
gaming_hours	0
breaks_per_day	0
coffee_intake_mg	0
exercise_minutes	0
assignments_completed	0
attendance_percentage	0
stress_level	0
focus_score	0
final_grade	0
productivity_score	0

dtype: int64

In [73]: *#Handling the missing values*
`missingValuesPercentage = df.isnull().mean() * 100`

#Filtering the columns based on the percentage of the missing values of e
`columnsWithMissing30 = missingValuesPercentage[missingValuesPercentage >`

#Printing the missing values that is greater than 30 percent.
`print('Column with more than 30 percent missing values :')`
`if columnsWithMissing30.empty:`
 `print("No missing value detected")`
`else:`
 `print(columnsWithMissing30)`

Column with more than 30 percent missing values :
No missing value detected

```
In [74]: #Hence no missing values were detected, we will find in this step the out
import numpy as np
import pandas as pd

data = df.copy()

# Function to detect outliers using Z-score
def detect_outliers_zscore(column, threshold=3):
    mean = np.mean(column)
    std_dev = np.std(column)
    z_scores = (column - mean) / std_dev
    return np.abs(z_scores) > threshold

# We select only numeric columns to apply
numeric_cols = data.select_dtypes(include=np.number).columns

# Apply Z-score only on numeric columns
for col in numeric_cols:
    data[f'{col}_Outlier'] = detect_outliers_zscore(data[col])

# Show rows that contain any outlier
outliers_only = data[data.filter(like='_Outlier').any(axis=1)]

outliers_only
```

```
Out[74]:
```

	student_id	age	gender	study_hours_per_day	sleep_hours	phone_usage_
61	62	21	Female	9.06	9.27	
1073	1074	21	Female	9.86	8.50	
2106	2107	23	Male	9.64	7.59	
4533	4534	20	Female	0.67	3.73	
5008	5009	23	Female	9.76	9.81	
6236	6237	17	Female	0.65	3.06	
8901	8902	20	Male	0.51	3.76	
10597	10598	21	Male	1.16	3.34	
16228	16229	18	Male	1.15	3.20	
19125	19126	23	Male	9.83	9.31	

10 rows x 35 columns

```
In [75]: import matplotlib.pyplot as plt
import seaborn as sns

plt.figure(figsize=(12, 6))

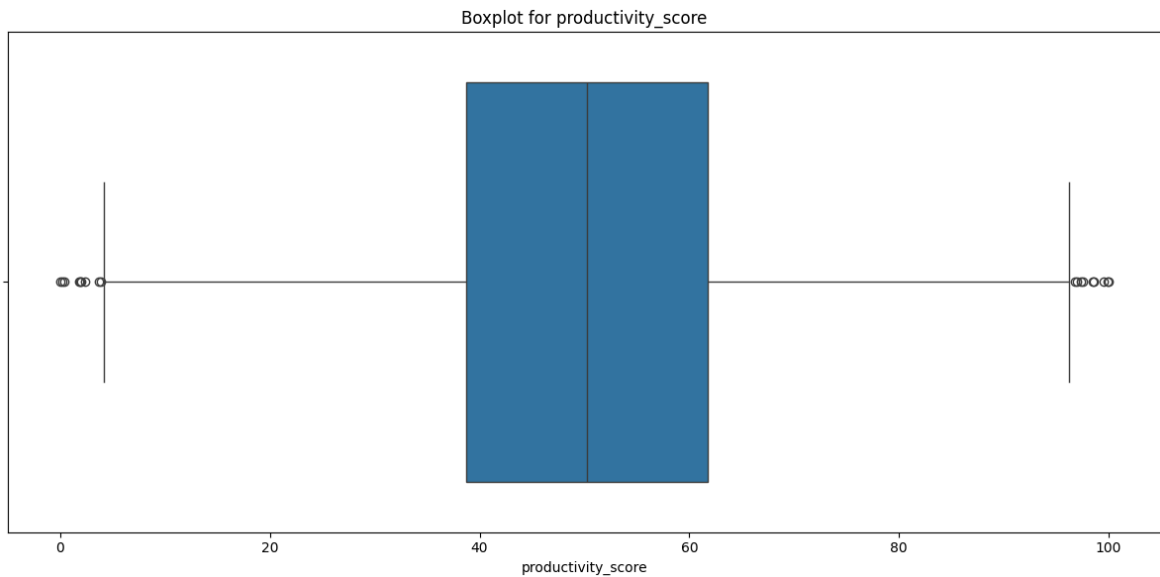
#The only column that has outliers is the productivity_score column so we
```

```

for i, col in enumerate(df.columns[17:18]): # Only plotting the feature
    sns.boxplot(x=df[col])
    plt.title(f"Boxplot for {col}")

plt.tight_layout()
plt.show()

```



In [137... *#Since the data frame has only 10 rows that contains outliers we will rep*

data = df.copy() # Create a copy to avoid modifying the original datafra

```
def replace_outliers_with_null(data, threshold=3):
```

Boolean DataFrame indicating if a value is an outlier

```
outliers = pd.DataFrame(False, index=df.index, columns=df.columns)
```

Looping through each numeric column

```
for column in data.columns:
```

```
    if data[column].dtype in ['float64', 'int64']: # Applicable on n
```

```
        mean = data[column].mean()
```

```
        std_dev = data[column].std()
```

```
        z_scores = (data[column] - mean) / std_dev
```

Mark True when abs(z_score) > threshold

```
outliers[column] = abs(z_scores) > threshold
```

Replace outliers with null values

```
data.loc[outliers[column], column] = np.nan
```

```
return data
```

#This variable will hold the data after replacing the outliers with null
cleaned_data = replace_outliers_with_null(data, threshold=3)

DataFrame that contains rows with NaN values after outlier replacement
rows_with_nan = cleaned_data[cleaned_data.isna().any(axis=1)].copy()
rows_with_nan

```
print("Original Data length:")
print(len(data))
```

```
print("\ncleaned Data length:")
print(len(cleaned_data))

print("\nData after replacing outliers with null values:")
cleaned_data.head()
```

Original Data length:
20000

cleaned Data length:
20000

Data after replacing outliers with null values:

Out [137...

	student_id	age	gender	study_hours_per_day	sleep_hours	phone_usage_hou
0	1.0	23.0	Female	4.35	3.63	3.3
1	2.0	20.0	Male	6.14	6.58	5.4
2	3.0	29.0	Female	4.98	3.26	4.8
3	4.0	27.0	Female	3.19	4.58	10.0
4	5.0	24.0	Male	7.67	6.21	3.0

In [87]: `len(cleaned_data)`

Out [87]: 20000

In [88]: `cleaned_data.head()`

Out [88]:

	student_id	age	gender	study_hours_per_day	sleep_hours	phone_usage_hou
0	1.0	23.0	Female	4.35	3.63	3.3
1	2.0	20.0	Male	6.14	6.58	5.4
2	3.0	29.0	Female	4.98	3.26	4.8
3	4.0	27.0	Female	3.19	4.58	10.0
4	5.0	24.0	Male	7.67	6.21	3.0

In [89]: `cleaned_data.isnull().sum()`

```
Out[89]: student_id      0
        age            0
        gender         0
        study_hours_per_day  0
        sleep_hours     0
        phone_usage_hours  0
        social_media_hours  0
        youtube_hours    0
        gaming_hours     0
        breaks_per_day   0
        coffee_intake_mg  0
        exercise_minutes  0
        assignments_completed  0
        attendance_percentage  0
        stress_level      0
        focus_score       0
        final_grade       0
        productivity_score  9
        dtype: int64
```

```
In [91]: #I wanted to print the rows that has a null values
rows_with_null = cleaned_data[cleaned_data.isnull().any(axis=1)]
print(len(rows_with_null))
rows_with_null
```

9

```
Out[91]:
```

	student_id	age	gender	study_hours_per_day	sleep_hours	phone_usage
61	62.0	21.0	Female	9.06	9.27	
1073	1074.0	21.0	Female	9.86	8.50	
4533	4534.0	20.0	Female	0.67	3.73	
5008	5009.0	23.0	Female	9.76	9.81	
6236	6237.0	17.0	Female	0.65	3.06	
8901	8902.0	20.0	Male	0.51	3.76	
10597	10598.0	21.0	Male	1.16	3.34	
16228	16229.0	18.0	Male	1.15	3.20	
19125	19126.0	23.0	Male	9.83	9.31	

```
In [93]: # Replace null values in numeric columns with the column mean
cleaned_data = cleaned_data.fillna(cleaned_data.mean(numeric_only=True))

print("Missing values after replacement:")
print(cleaned_data.isnull().sum())
```

Missing values after replacement:

student_id	0
age	0
gender	0
study_hours_per_day	0
sleep_hours	0
phone_usage_hours	0
social_media_hours	0
youtube_hours	0
gaming_hours	0
breaks_per_day	0
coffee_intake_mg	0
exercise_minutes	0
assignments_completed	0
attendance_percentage	0
stress_level	0
focus_score	0
final_grade	0
productivity_score	0

dtype: int64

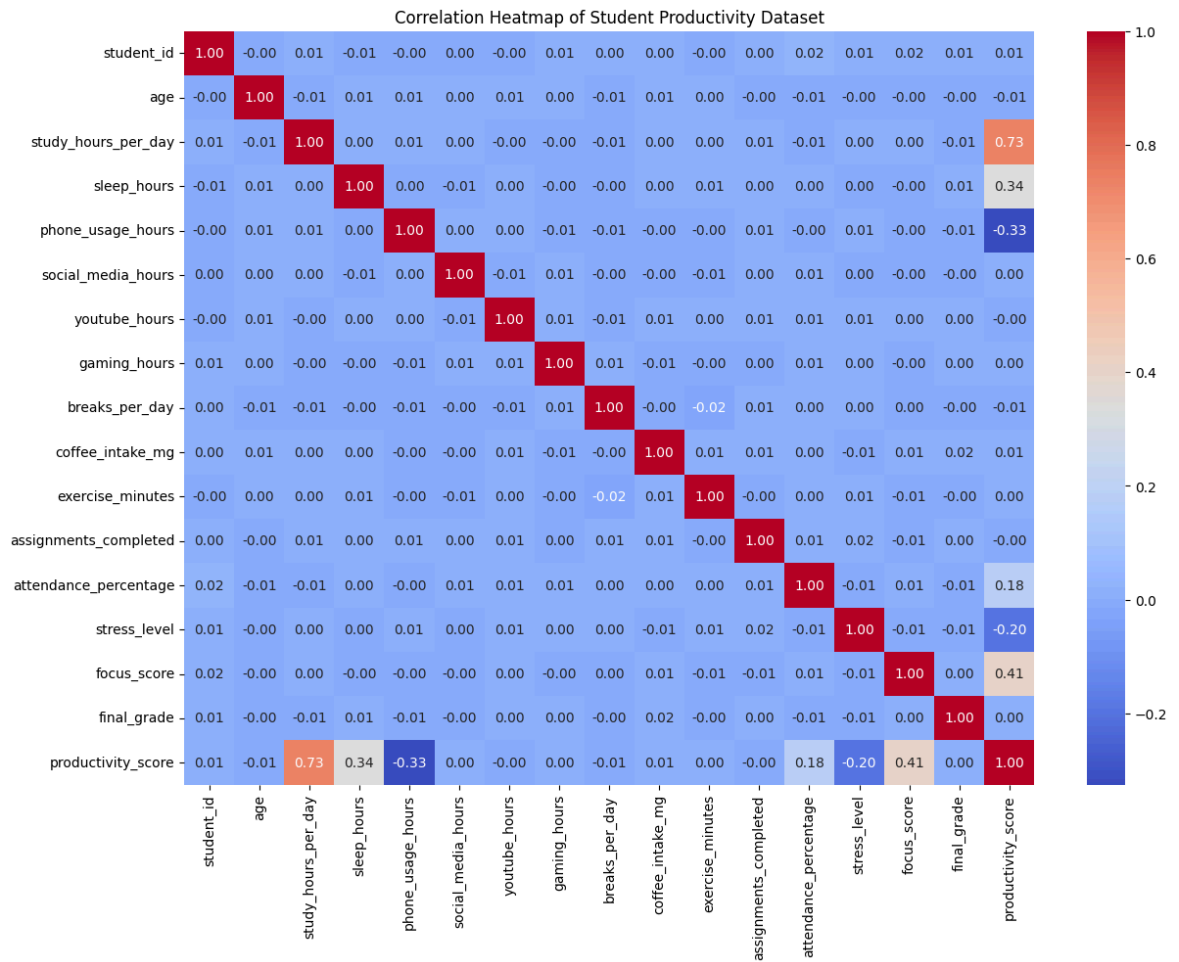
```
In [94]: import matplotlib.pyplot as plt
import seaborn as sns

# Select only numeric columns
numeric_data = cleaned_data.select_dtypes(include='number')

# Create correlation matrix
corr_matrix = numeric_data.corr()

# Plot heatmap
plt.figure(figsize=(14,10))
sns.heatmap(corr_matrix, annot=True, cmap='coolwarm', fmt=".2f")

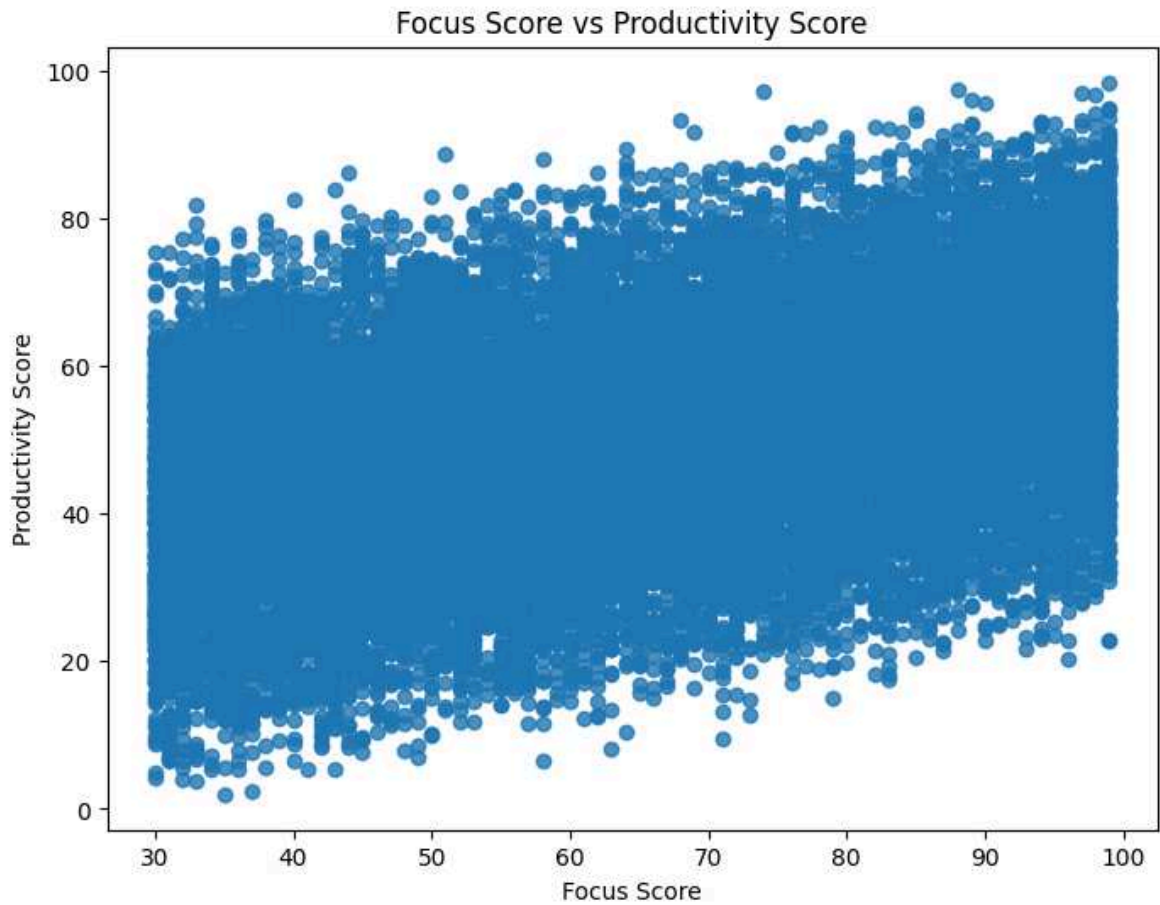
plt.title("Correlation Heatmap of Student Productivity Dataset")
plt.show()
```



```
In [119... plt.figure(figsize=(8,6))
#I used the regplot cuz we have a direct relationship between focus_score
sns.regplot(
    x='focus_score',
    y='productivity_score',
    data=cleaned_data
)

plt.title("Focus Score vs Productivity Score")
plt.xlabel("Focus Score")
plt.ylabel("Productivity Score")

plt.show()
```



```
In [104... # Count number of males and females
gender_counts = cleaned_data['gender'].value_counts()

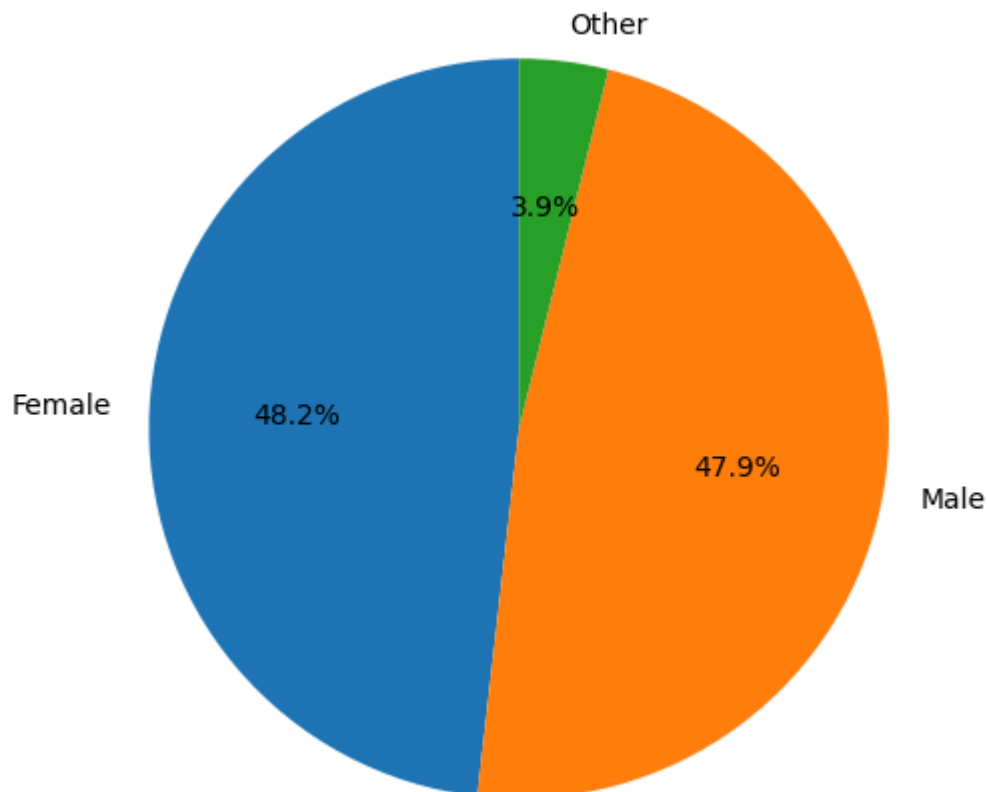
# Create pie chart
plt.figure(figsize=(6,6))

plt.pie(
    gender_counts,
    labels=gender_counts.index,
    autopct='%1.1f%%',
    startangle=90
)

plt.title("Gender Distribution of Students")

plt.show()
```

Gender Distribution of Students



```
In [103... # Calculate average stress level by gender
stress_by_gender = cleaned_data.groupby('gender')['stress_level'].mean()

plt.figure(figsize=(6,6))

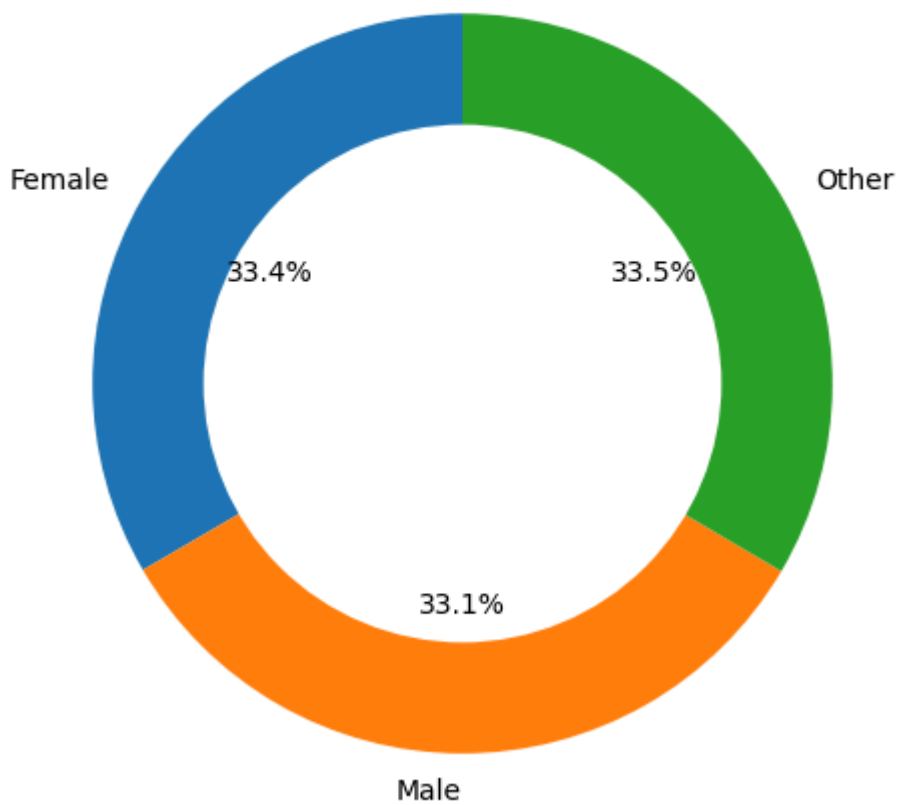
# Create pie chart
wedges, texts, autotexts = plt.pie(
    stress_by_gender,
    labels=stress_by_gender.index,
    autopct='%1.1f%%',
    startangle=90
)

# Create white circle in the center to make it a donut
centre_circle = plt.Circle((0,0),0.70,fc='white')
fig = plt.gcf()
fig.gca().add_artist(centre_circle)

plt.title("Average Stress Level by Gender")

plt.show()
```

Average Stress Level by Gender

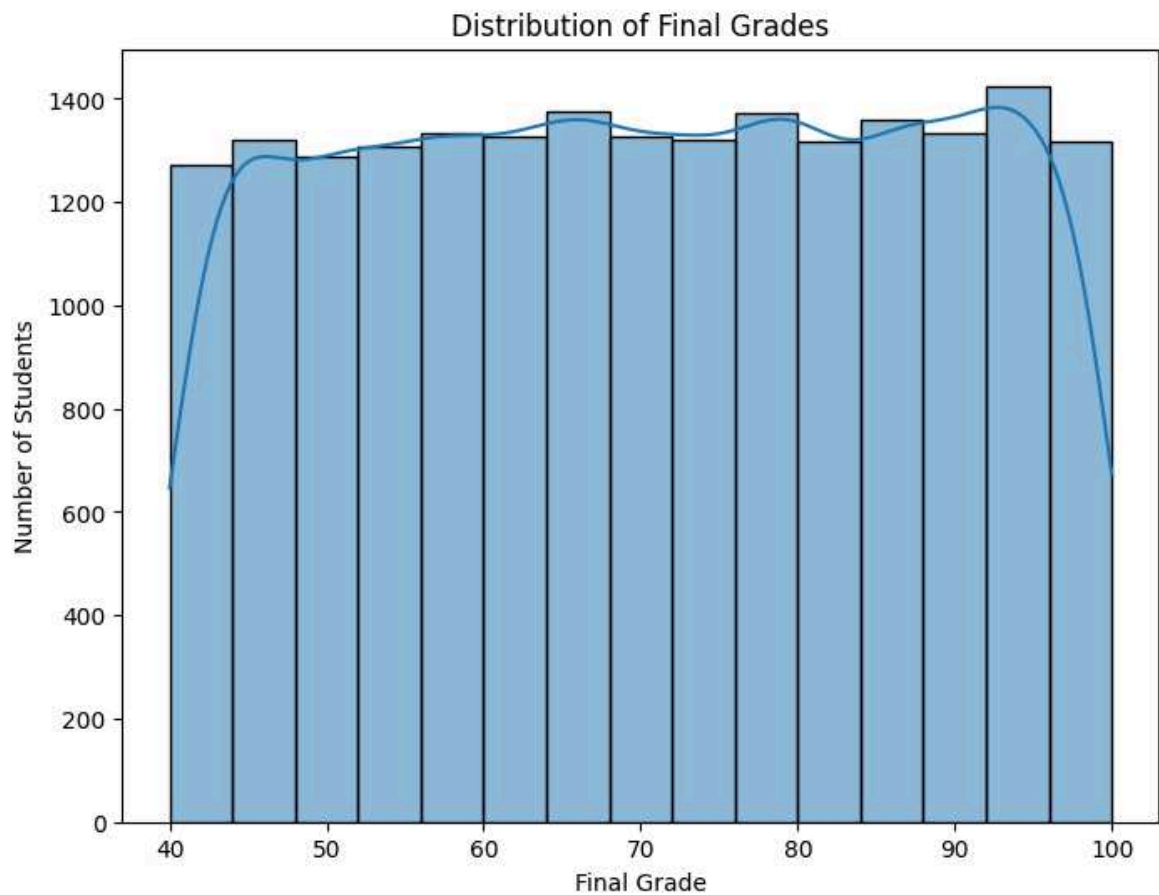


```
In [138... plt.figure(figsize=(8,6))

sns.histplot(
    cleaned_data['final_grade'],
    bins=15,
    kde=True
)

plt.title("Distribution of Final Grades")
plt.xlabel("Final Grade")
plt.ylabel("Number of Students")

plt.show()
```

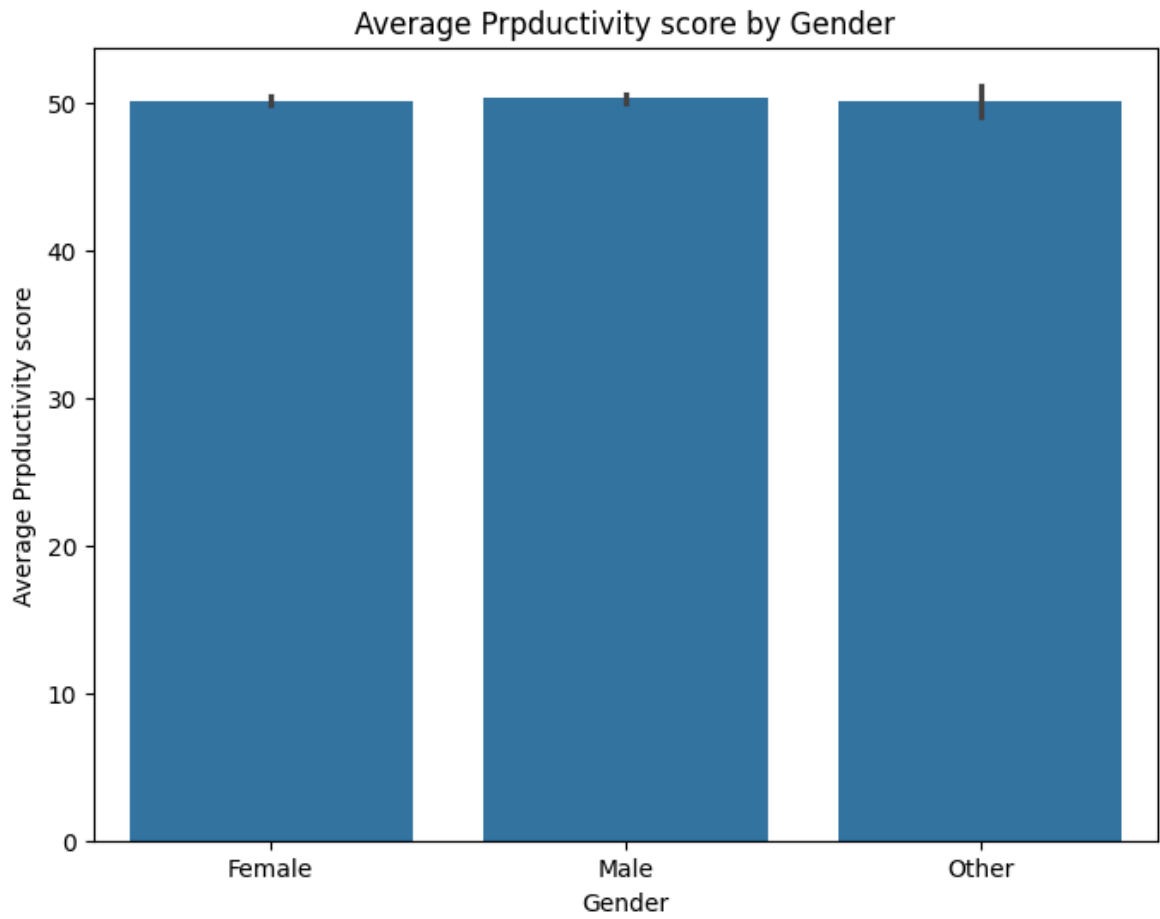


```
In [109... plt.figure(figsize=(8,6))

sns.barplot(
    x='gender',
    y='productivity_score',
    data=cleaned_data,
    estimator='mean'
)

plt.title("Average Prpductivity score by Gender")
plt.xlabel("Gender")
plt.ylabel("Average Prpductivity score")

plt.show()
```



In []: