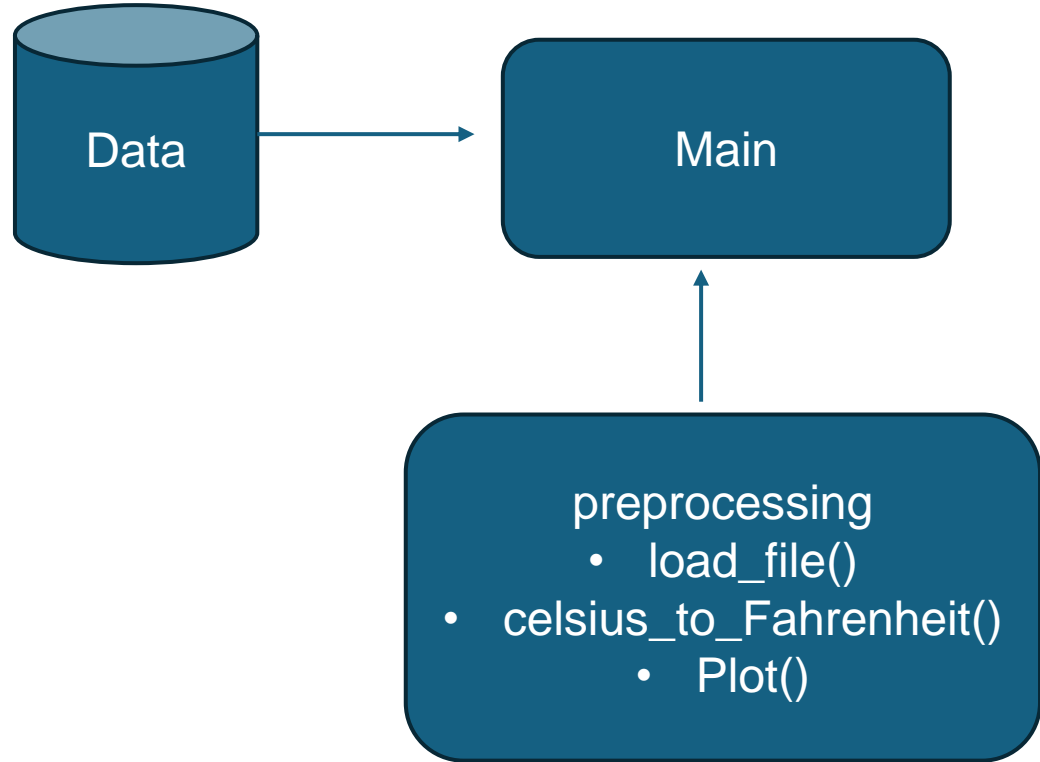# PA1
CMPS 470

Byte Me: Sanele Harmon Taeden Kitchens

TEAM MEMBERS & ROLES

Sanele Harmon: Create the GitHub repository. Preprocess the temperature columns and plot

Taeden Kitchen: Preprocess the Rainfall column and plot. Visualize the distribution of each data column.

MODULE COMMUNICATION GRAPH

DESCRIPTION OF THE PROJECT

• Goal: Predicting whether there will be rainfall on a given day by means of ML algorithms ANN, SVM, DT, and K-NN.
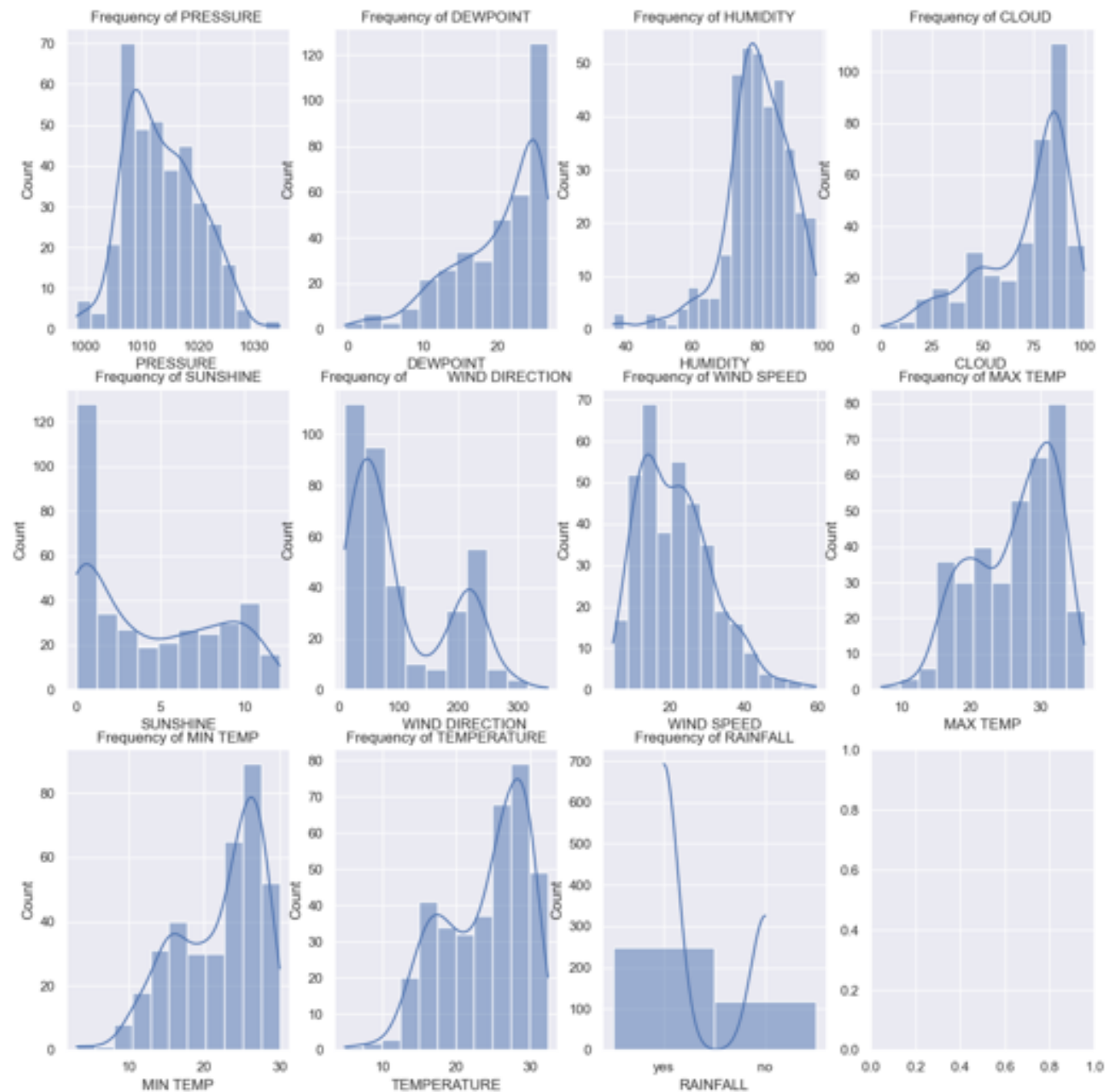
 Tasks:

- Visualize data
- Preprocess data

DESCRIPTION OF THE RAW DATA

- Data Source: Seek Geek

| ATTRIBUTE	TYPE
SAMPLE ID	INT
RAINFALL (TARGET)	STRING ('yes', 'no')
DAY	INT
PRESSURE	FLOAT
MAX TEMP	FLOAT
TEMPERATURE	FLOAT
MIN TEMP	FLOAT
DEW POINT	FLOAT
HUMIDITY	INT
CLOUD	INT
SUNSHINE	FLOAT
WIND DIRECTION	INT
WIND SPEED	FLOAT

DISTRIBUTION OF DATA

PREPROCESSING

- For preprocessing, the values for the RAINFALL column were mapped from 'yes' and 'no' to 1 and 0, respectively
- The MAX TEMP, MIN TEMP, and TEMPERATURE columns were converted from Celsius to Fahrenheit

RAW DATA: MAX_TEMP, MIN_TEMP, TEMPERATURE

SAMPLE ID	RAINFALL	DAY	PRESSURE	MAX TEMP	TEMPERATURE	MIN TEMP
1	yes	1	1025.9	19.9	18.3	16.8
2	yes	2	1022	21.7	18.9	17.2
3	yes	3	1019.7	20.3	19.3	18
4	yes	4	1018.9	22.3	20.6	19.1
5	yes	5	1015.9	21.3	20.7	20.2
6	yes	6	1018.8	24.3	20.9	19.2
7	no	7	1021.8	21.4	18.8	17
8	no	8	1020.8	21	18.4	16.5
9	no	9	1020.6	18.9	18.1	17.1
10	yes	10	1017.5	18.5	18	17.2
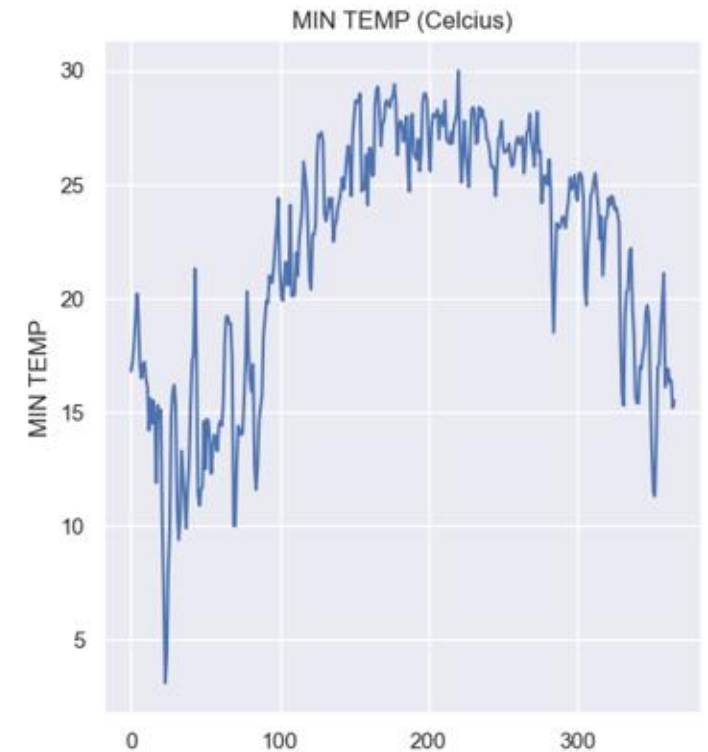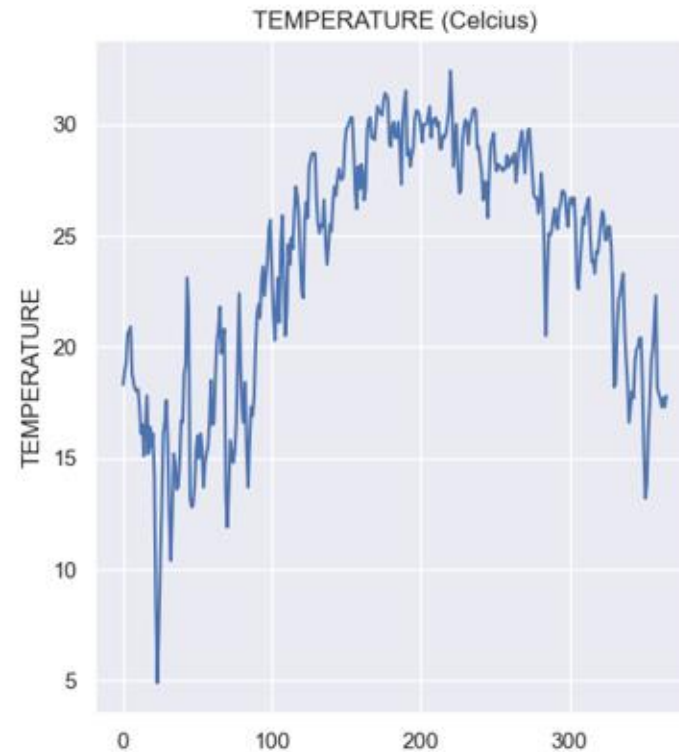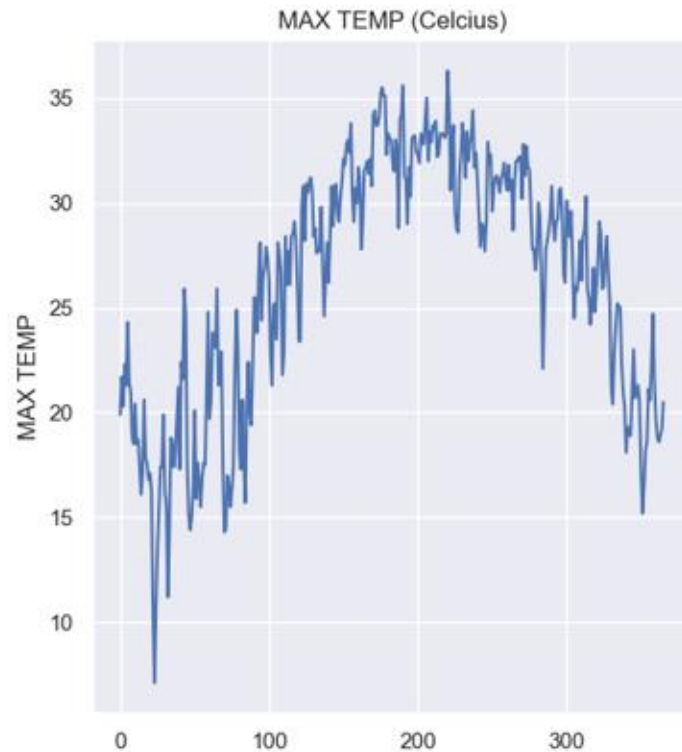
PREPROCESSED DATA: MAX_TEMP, MIN_TEMP, TEMPERATURE

SAMPLE ID	RAINFALL	DAY	PRESSURE	MAX TEMP	TEMPERATURE	MIN TEMP
1	yes	1	1025.9	67.82	64.94	62.24
2	yes	2	1022	71.06	66.02	62.96
3	yes	3	1019.7	68.54	66.74	64.4
4	yes	4	1018.9	72.14	69.08	66.38
5	yes	5	1015.9	70.34	69.26	68.36
6	yes	6	1018.8	75.74	69.62	66.56
7	no	7	1021.8	70.52	65.84	62.6
8	no	8	1020.8	69.8	65.12	61.7

MAX TEMP (Farenheit) — TEMPERATURE (Farenheit) — MIN TEMP (Farenheit)
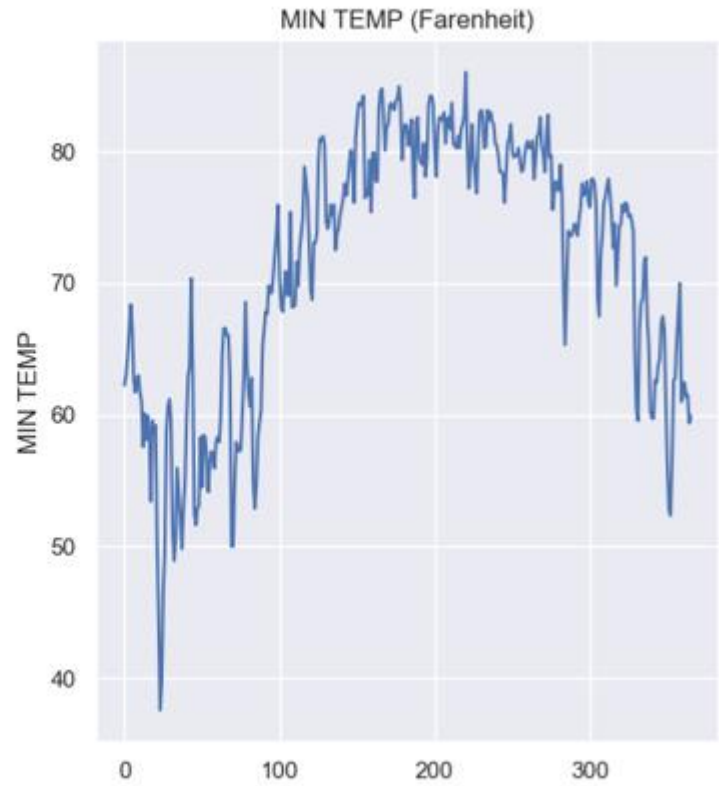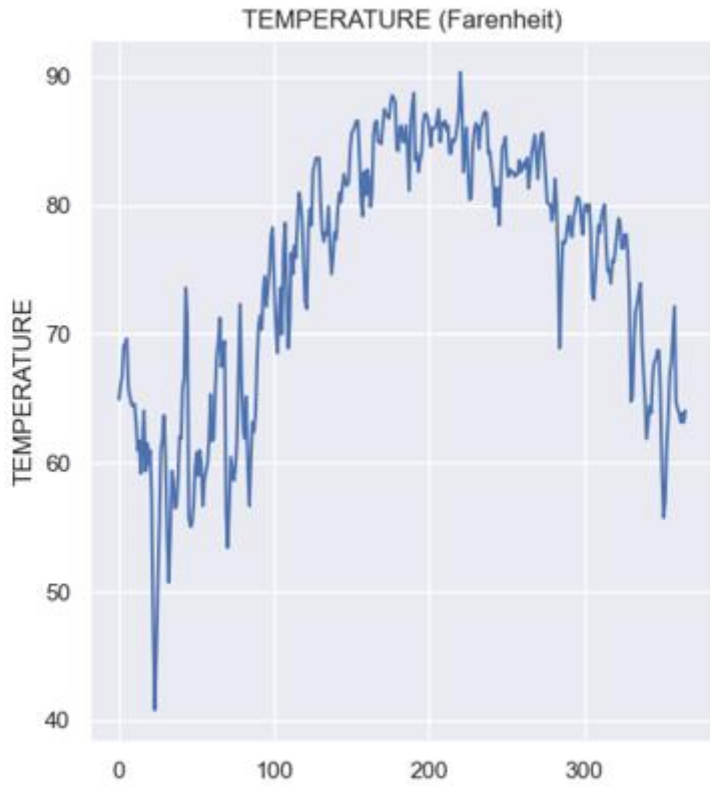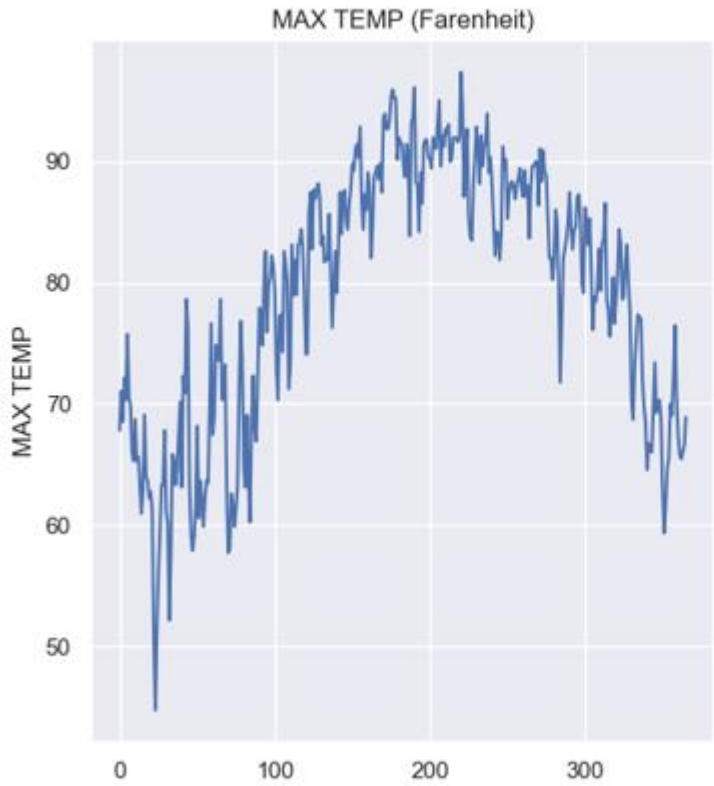
RAW DATA: Rainfall

SAMPLE ID	RAINFALL
1	yes
2	yes
3	yes
4	yes
5	yes
6	yes
7	no
8	no
9	no
10	yes
11	yes
12	no
13	no
14	yes
15	yes
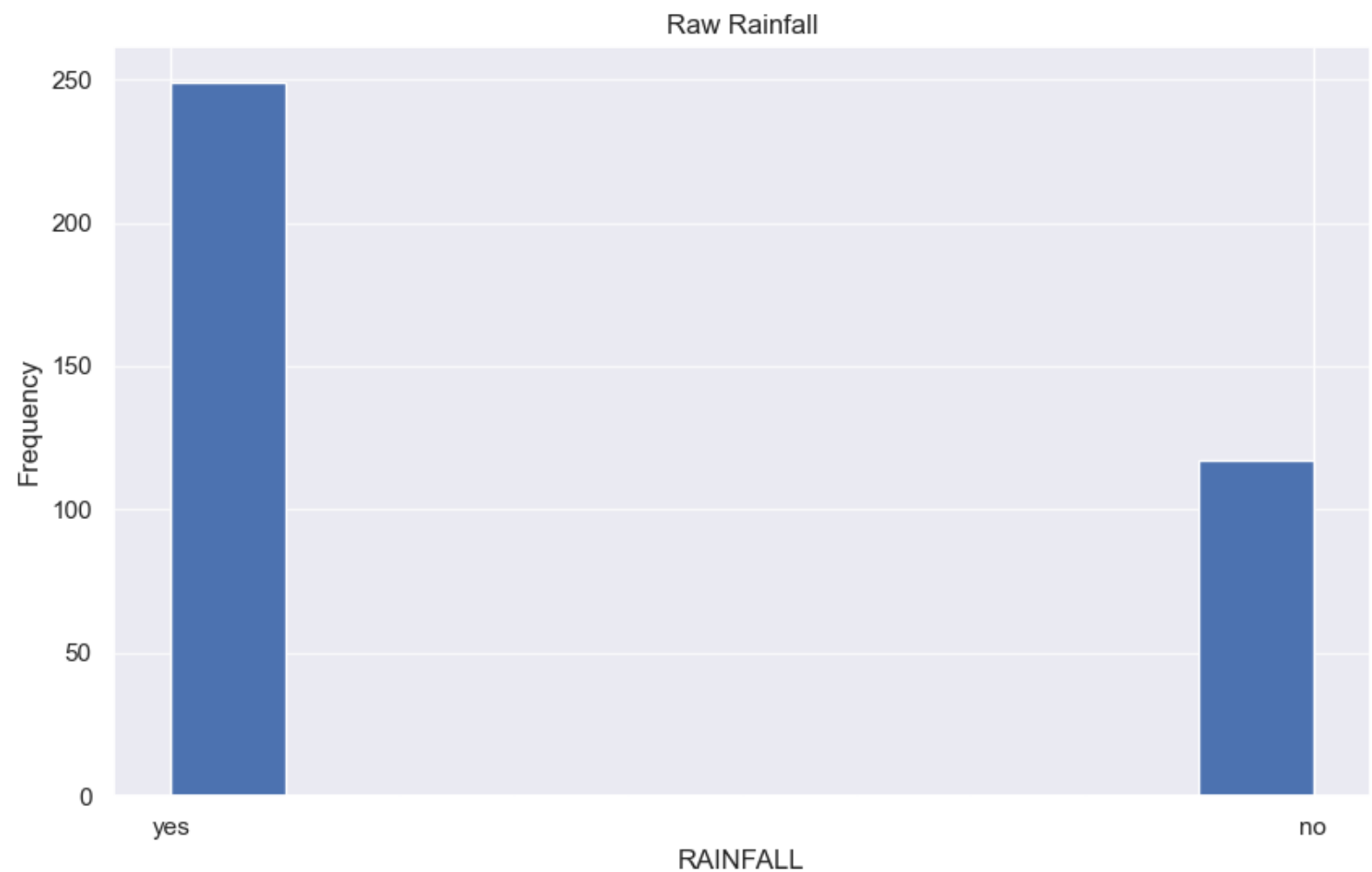
Raw Rainfall

PREPROCESSED DATA: Rainfall

SAMPLE ID	RAINFALL
1	1
2	1
3	1
4	1
5	1
6	1
7	0
8	0
9	0
10	1
11	1
12	0
13	0
14	1
15	1

Preprocessed Rainfall