In [5]:
```
import pandas as pd
import numpy as np
import seaborn as sns
import os
```

```
C:\Users\joykaaria\anaconda3\lib\site-packages\scipy\__init__.py:146: UserWarning: A
NumPy version >=1.16.5 and <1.23.0 is required for this version of SciPy (detected ve
rsion 1.26.1
  warnings.warn(f"A NumPy version >={np_minversion} and <{np_maxversion}"
```

In [6]:
```
df =pd.read_csv(r"C:\Users\joykaaria\Desktop\Temp_data.csv")
```

Displaying dataframe

In [7]:
```
df
```

Out[7]:

	Unnamed: 0	Temp	Humidity
0	0	1	1.0
1	1	NaN	NaN
2	2	3	31.0
3	3	2	22.0
4	4	3	33.0
5	5	1	11.0
6	6	2	21.0
7	7	N/a	24.0
8	8	1	12.0
9	9	NaN	32.0

IF THE Dataframe is too large use to_string to view entire data

In [29]:
```
df.to_string()
```

Out[29]:
```
'   Unnamed: 0  Temp  Humidity\n0           0   1.0       1.0\n1           1   NaN
NaN\n2           2   3.0      31.0\n3           3   2.0      22.0\n4           4   3.
0      33.0\n5           5   1.0      11.0\n6           6   2.0      21.0\n7
7   NaN      24.0\n8           8   1.0      12.0\n9           9   NaN      32.0'
```

Displays first five rows by default

In [8]:
```
df.head()
```

Out[8]:

	Unnamed: 0	Temp	Humidity
0	0	1	1.0
1	1	NaN	NaN
2	2	3	31.0
3	3	2	22.0
4	4	3	33.0

Displays last five rows by default

In [31]:
```
df.tail()
```

Out[31]:

	Unnamed: 0	Temp	Humidity
5	5	1.0	11.0
6	6	2.0	21.0
7	7	NaN	24.0
8	8	1.0	12.0
9	9	NaN	32.0

In [9]:
```
missing_value =["N/a","na","NaN","np.nan"]
df =pd.read_csv(r"C:\Users\joykaaria\Desktop\Temp_data.csv",na_values = missing_value)
```

In [10]:
```
df
```

Out[10]:

	Unnamed: 0	Temp	Humidity
0	0	1.0	1.0
1	1	NaN	NaN
2	2	3.0	31.0
3	3	2.0	22.0
4	4	3.0	33.0
5	5	1.0	11.0
6	6	2.0	21.0
7	7	NaN	24.0
8	8	1.0	12.0
9	9	NaN	32.0

identifying null values

In [11]:
```
df.isnull().sum()
```

Out[11]:
```
Unnamed: 0    0
Temp          3
Humidity      1
dtype: int64
```

In [13]:
```
df.shape
```

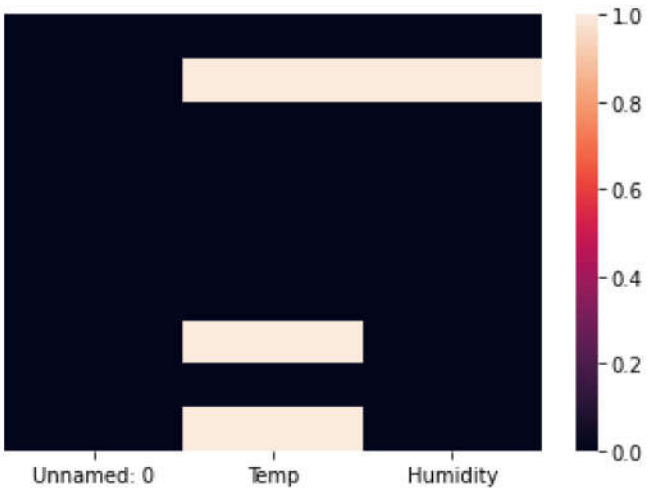
Out[13]:
```
(10, 3)
```

In [14]:
```
df.isnull().any()
```

Out[14]:
```
Unnamed: 0    False
Temp           True
Humidity       True
dtype: bool
```

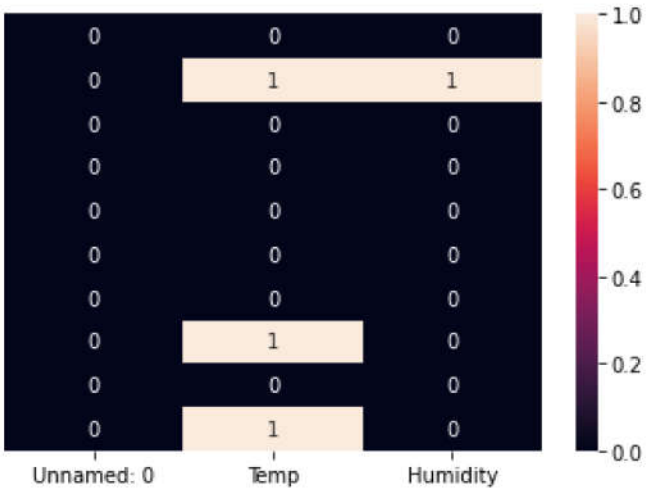
In [15]:
```
sns.heatmap(df.isnull(),yticklabels = False)
```

Out[15]:
```
<AxesSubplot:>
```



In [16]:
```
sns.heatmap(df.isnull(),yticklabels = False,annot = True)
```

Out[16]:
```
<AxesSubplot:>
```



In [ ]:
```
Discarding null values
```

In [17]:
```
df.dropna()
```

Out[17]:

	Unnamed: 0	Temp	Humidity
0	0	1.0	1.0
2	2	3.0	31.0
3	3	2.0	22.0
4	4	3.0	33.0
5	5	1.0	11.0
6	6	2.0	21.0
8	8	1.0	12.0

In [32]:
```
df.dropna(how="all")
```

Out[32]:

	Unnamed: 0	Temp	Humidity
0	0	1.0	1.0
1	1	NaN	NaN
2	2	3.0	31.0
3	3	2.0	22.0
4	4	3.0	33.0
5	5	1.0	11.0
6	6	2.0	21.0
7	7	NaN	24.0
8	8	1.0	12.0
9	9	NaN	32.0

filling null values

In [19]:
```
df.fillna(0)
```

Out[19]:

	Unnamed: 0	Temp	Humidity
0	0	1.0	1.0
1	1	0.0	0.0
2	2	3.0	31.0
3	3	2.0	22.0
4	4	3.0	33.0
5	5	1.0	11.0
6	6	2.0	21.0
7	7	0.0	24.0
8	8	1.0	12.0
9	9	0.0	32.0

In [20]:
```
df.fillna(method = "ffill")
```

Out[20]:

	Unnamed: 0	Temp	Humidity
0	0	1.0	1.0
1	1	1.0	1.0
2	2	3.0	31.0
3	3	2.0	22.0
4	4	3.0	33.0
5	5	1.0	11.0
6	6	2.0	21.0
7	7	2.0	24.0
8	8	1.0	12.0
9	9	1.0	32.0

In [21]:
```
df.fillna(method="bfill")
```

Out[21]:

	Unnamed: 0	Temp	Humidity
0	0	1.0	1.0
1	1	3.0	31.0
2	2	3.0	31.0
3	3	2.0	22.0
4	4	3.0	33.0
5	5	1.0	11.0
6	6	2.0	21.0
7	7	1.0	24.0
8	8	1.0	12.0
9	9	NaN	32.0

In [22]:
```
df.interpolate()
```

Out[22]:

	Unnamed: 0	Temp	Humidity
0	0	1.0	1.0
1	1	2.0	16.0
2	2	3.0	31.0
3	3	2.0	22.0
4	4	3.0	33.0
5	5	1.0	11.0
6	6	2.0	21.0
7	7	1.5	24.0
8	8	1.0	12.0
9	9	1.0	32.0

In [23]:
```
df_dropped = df.dropna()
```

In [24]:
```
df_dropped
```

Out[24]:

	Unnamed: 0	Temp	Humidity
0	0	1.0	1.0
2	2	3.0	31.0
3	3	2.0	22.0
4	4	3.0	33.0
5	5	1.0	11.0
6	6	2.0	21.0
8	8	1.0	12.0

In [26]:
```
df.dropna(how="all")
```

Out[26]:

	Unnamed: 0	Temp	Humidity
0	0	1.0	1.0
1	1	NaN	NaN
2	2	3.0	31.0
3	3	2.0	22.0
4	4	3.0	33.0
5	5	1.0	11.0
6	6	2.0	21.0
7	7	NaN	24.0
8	8	1.0	12.0
9	9	NaN	32.0

In [27]:
```
df.fillna({
    'Temp': 5.0
})
```

Out[27]:

	Unnamed: 0	Temp	Humidity
0	0	1.0	1.0
1	1	5.0	NaN
2	2	3.0	31.0
3	3	2.0	22.0
4	4	3.0	33.0
5	5	1.0	11.0
6	6	2.0	21.0
7	7	5.0	24.0
8	8	1.0	12.0
9	9	5.0	32.0

In [28]:
```
df.fillna({
    'Humidity':25.0
})
```

Out[28]:

	Unnamed: 0	Temp	Humidity
0	0	1.0	1.0
1	1	NaN	25.0
2	2	3.0	31.0
3	3	2.0	22.0
4	4	3.0	33.0
5	5	1.0	11.0
6	6	2.0	21.0
7	7	NaN	24.0
8	8	1.0	12.0
9	9	NaN	32.0

In [ ]: