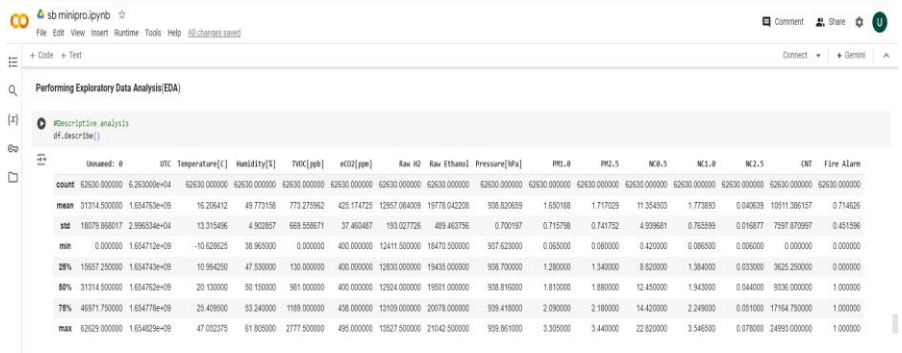
Data Collection and Preprocessing Phase

Date	11-07-2024
Team ID	740047
Project Title	SMOKE DETECTION USING IOT DATASET
Maximum Marks	6 Marks

Data Exploration and Preprocessing Report

Dataset variables will be statistically analyzed to identify patterns and outliers, with Python employed for preprocessing tasks like normalization and feature engineering. Data cleaning will address missing values and outliers, ensuring quality for subsequent analysis and modeling, and forming a strong foundation for insights and predictions.

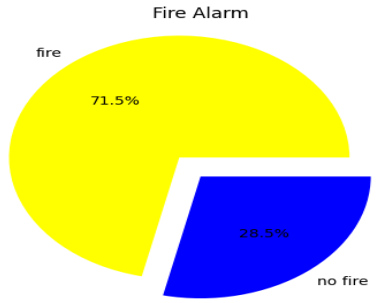
Section	**Description**
Data Overview	Dimension: 614 rows × 13 columns Descriptive statistics: 
Univariate Analysis	

sb minipro.ipynb

File Edit View Insert Runtime Tools Help Last saved at 11:42 AM

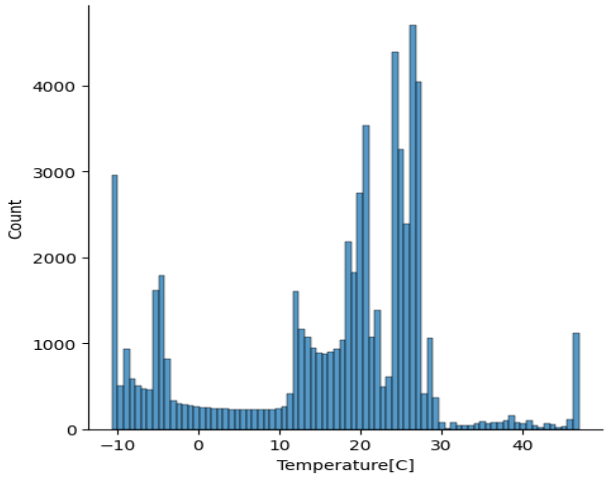
+ Code + Text

```
#Visualizing data
#Univariate analysis
plt.pie(df['Fire Alarm'].value_counts(),[0.2,0],labels=['fire','no fire'],autopct='%1.1f%%',colors=['yellow','blue'])
plt.title('Fire Alarm')
plt.show()
```



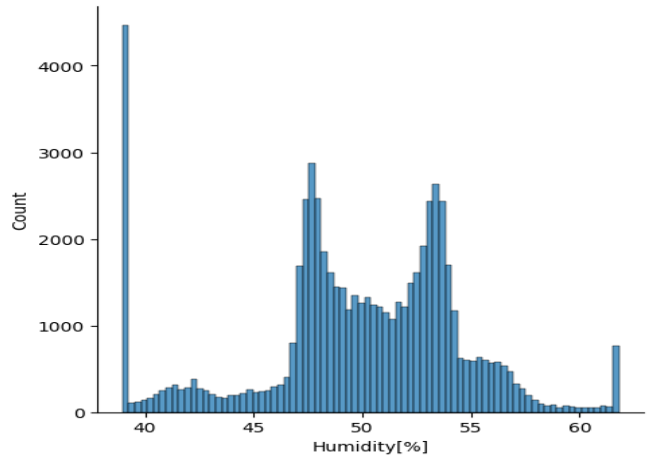

```
#visualizing temperature
sns.displot(df['Temperature[C]'])
```

```
<seaborn.axisgrid.FacetGrid at 0x7d8edd23a8c0>
```



```
#visualizing Humidity
sns.displot(df['Humidity[%]'])
```

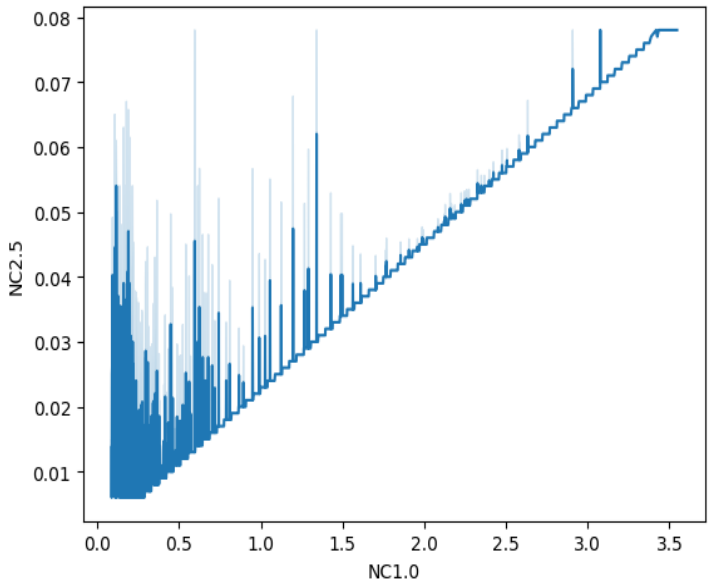
```
<seaborn.axisgrid.FacetGrid at 0x7d8edcad3cd0>
```



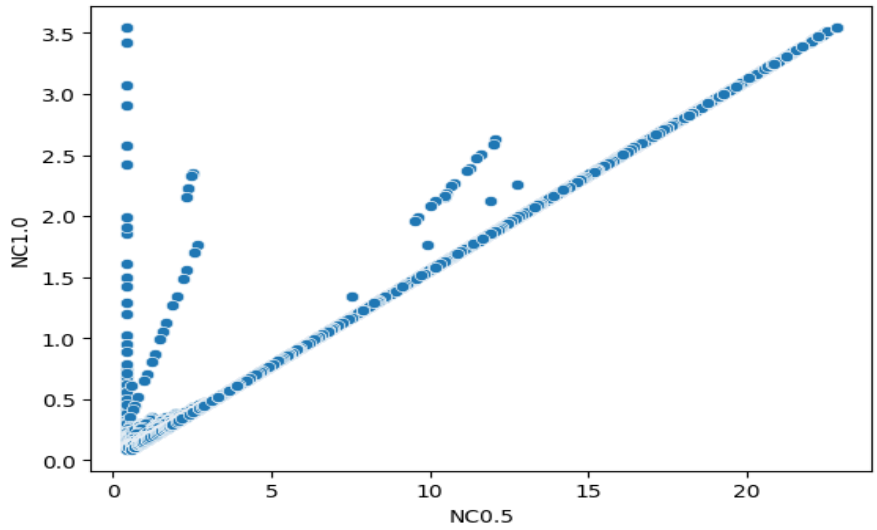
Bivariate Analysis

Bivariate Analysis

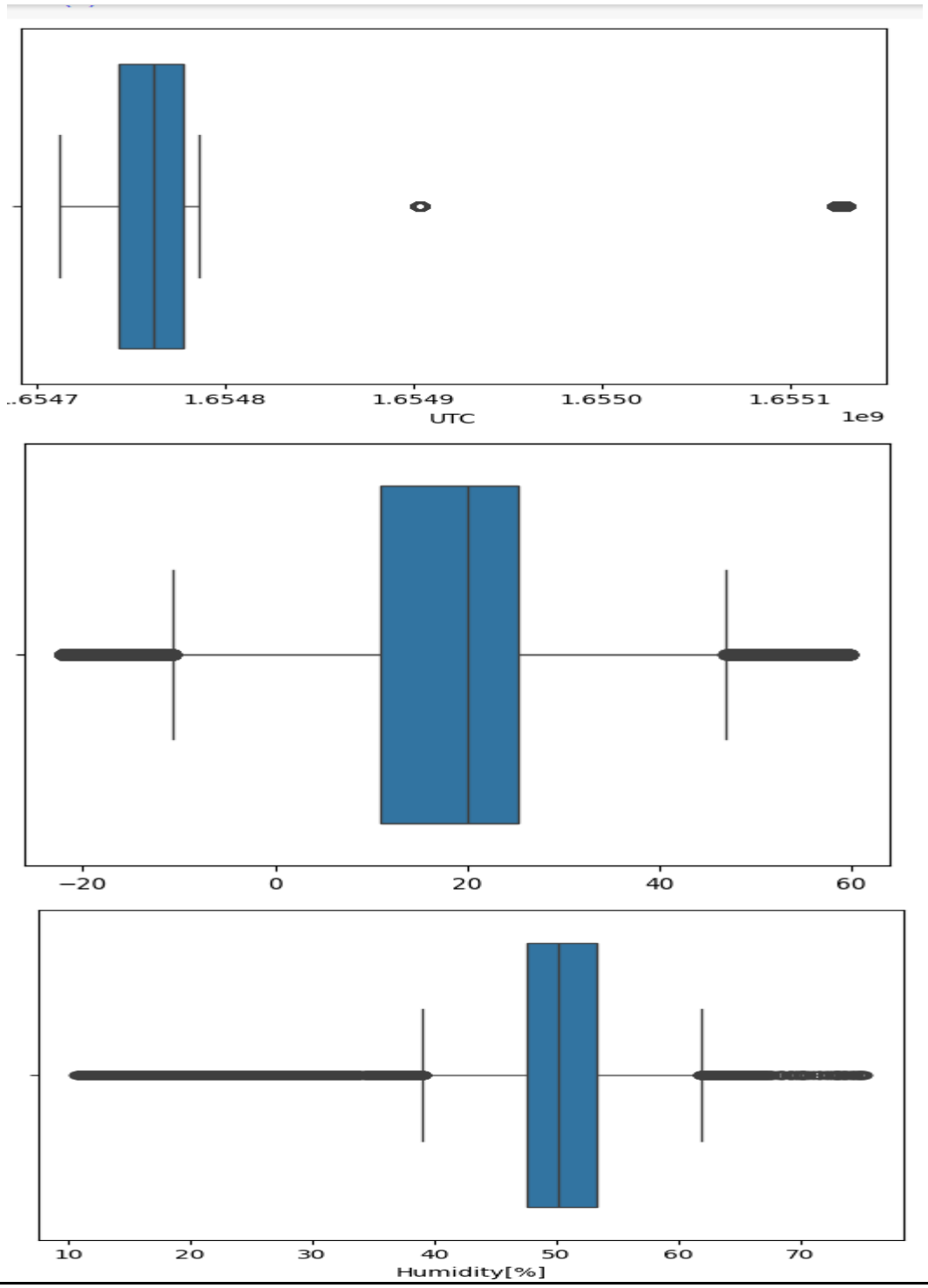
```
sns.lineplot(x='NC1.0',y='NC2.5', data=df)
plt.xlabel=('NC1.0')
plt.ylabel=('NC2.5')
plt.show()
```

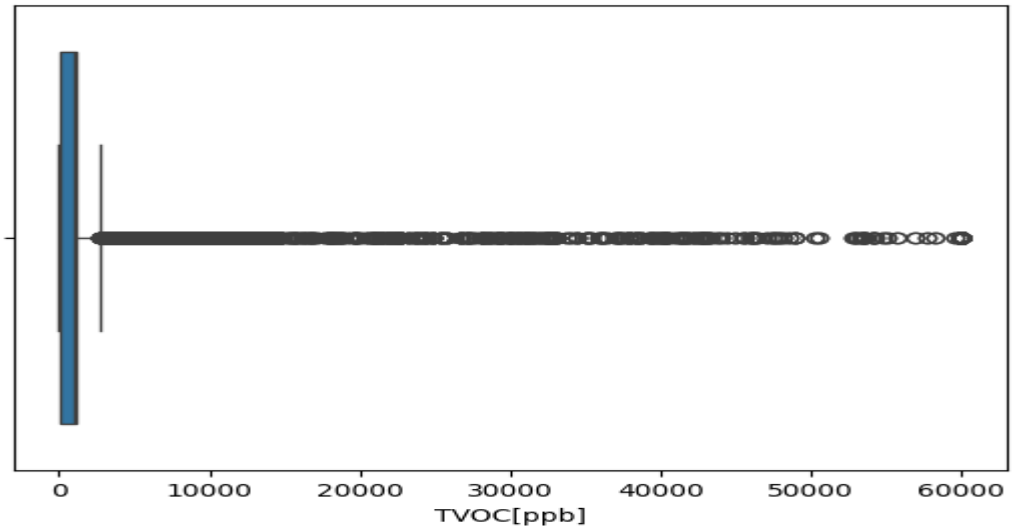


```
sns.scatterplot(x='NC0.5',y='NC1.0', data=df)
plt.xlabel=('NC0.5')
plt.ylabel=('NC1.0')
plt.show()
```



Multivariate Analysis

Outliers and Anomalies	



Data Preprocessing Code Screenshots

Loading Data

```
[ ] df=pd.read_csv('/content/smoke_detection_iot.csv')

df.head()
```

	Unnamed: 0	UTC	Temperature[C]	Humidity[%]	TVOC[ppb]	eCO2[ppm]	Raw H2	Raw Ethanol	Pressure[hPa]	PM1.0	PM2.5	NC0.5	NC1.0	NC2.5	CNT	Fire Alarm
0	0	1654733331	20.000	57.36	0	400	12306	18520	939.735	0.0	0.0	0.0	0.0	0.0	0	0
1	1	1654733332	20.015	56.67	0	400	12345	18651	939.744	0.0	0.0	0.0	0.0	0.0	1	0
2	2	1654733333	20.029	55.96	0	400	12374	18764	939.738	0.0	0.0	0.0	0.0	0.0	2	0
3	3	1654733334	20.044	55.28	0	400	12390	18849	939.736	0.0	0.0	0.0	0.0	0.0	3	0
4	4	1654733335	20.059	54.69	0	400	12403	18921	939.744	0.0	0.0	0.0	0.0	0.0	4	0

Feature Engineering	Attached the codes in final submission.
Save Processed Data	-