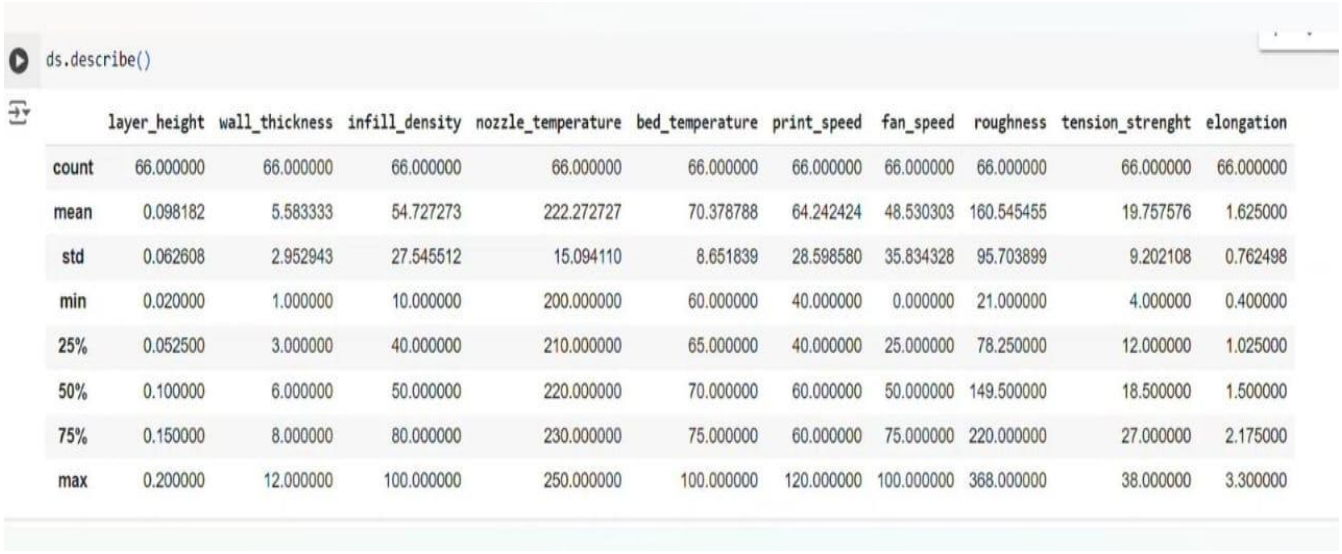
Data Collection and Preprocessing Phase

Date	7 July 2024
Team ID	739728
Project Title	3D printer material prediction using machine learning
Maximum Marks	6 Marks

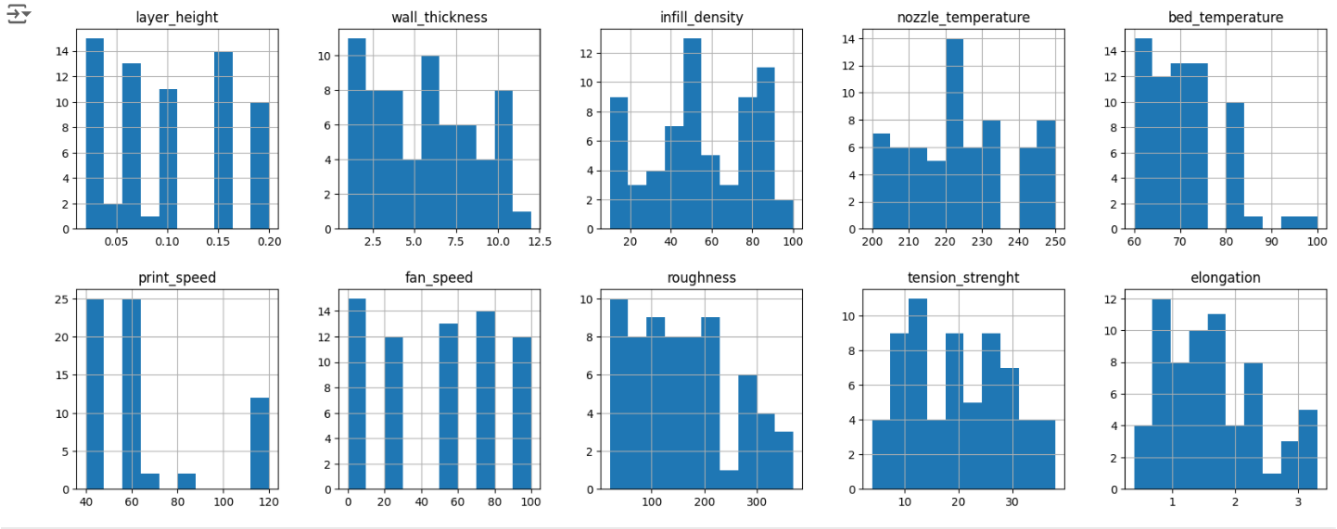
Data Exploration and Preprocessing Template

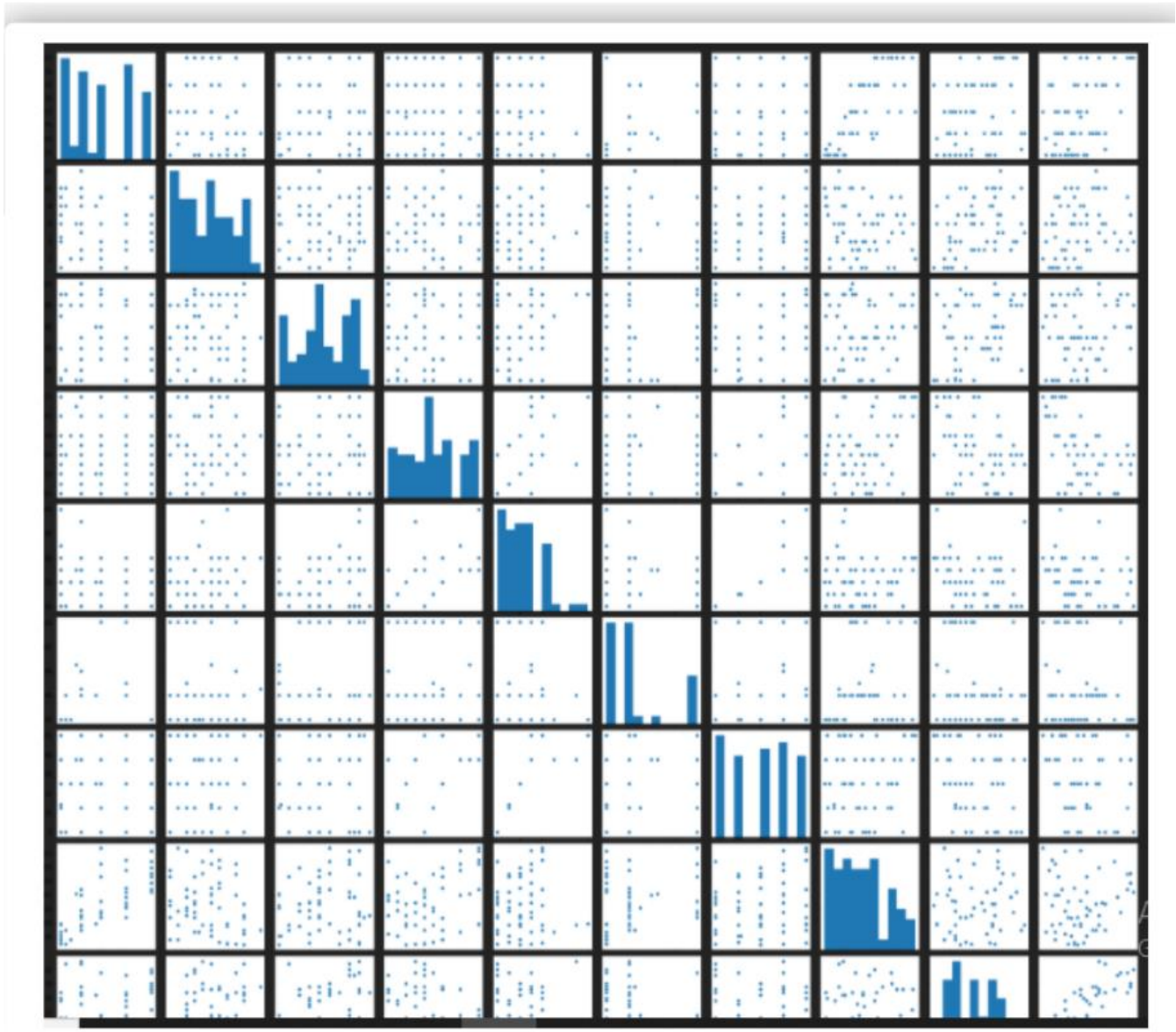
Identifies data sources, assesses quality issues like missing values and duplicates, and implements resolution plans to ensure accurate and reliable analysis.

Section	Description
Data Overview	

```
ds.describe()
```

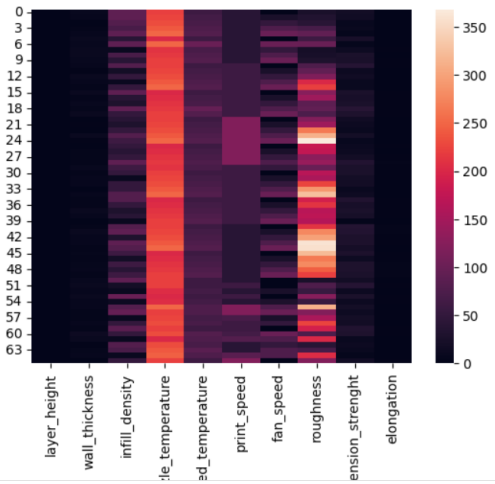
	layer_height	wall_thickness	infill_density	nozzle_temperature	bed_temperature	print_speed	fan_speed	roughness	tension_strenght	elongation
count	66.000000	66.000000	66.000000	66.000000	66.000000	66.000000	66.000000	66.000000	66.000000	66.000000
mean	0.098182	5.583333	54.727273	222.272727	70.378788	64.242424	48.530303	160.545455	19.757576	1.625000
std	0.062608	2.952943	27.545512	15.094110	8.651839	28.598580	35.834328	95.703899	9.202108	0.762498
min	0.020000	1.000000	10.000000	200.000000	60.000000	40.000000	0.000000	21.000000	4.000000	0.400000
25%	0.052500	3.000000	40.000000	210.000000	65.000000	40.000000	25.000000	78.250000	12.000000	1.025000
50%	0.100000	6.000000	50.000000	220.000000	70.000000	60.000000	50.000000	149.500000	18.500000	1.500000
75%	0.150000	8.000000	80.000000	230.000000	75.000000	60.000000	75.000000	220.000000	27.000000	2.175000
max	0.200000	12.000000	100.000000	250.000000	100.000000	120.000000	100.000000	368.000000	38.000000	3.300000

Univariate Analysis	```ds.hist(figsize=(20,20), layout=(5,5), sharex=False);``` 
Bivariate Analysis	



```
sns.heatmap(ds[['layer_height','wall_thickness','infill_density','nozzle_temperature','bed_temperature','print_speed','fan_speed','roughness','tension_strenght','elongation']])
```

```
<Axes: >
```



Data Preprocessing Code Screenshots

Loading Data

```
[3] ds=pd.read_csv(r"/content/3D printer Material Prediction (1).csv")
```

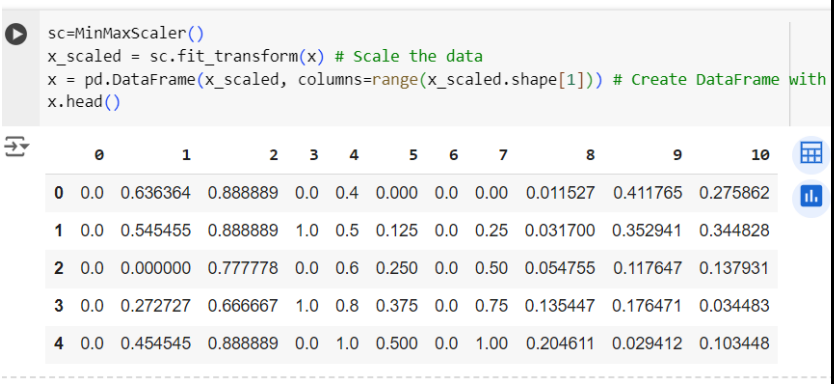
```
[4] ds.head()
```

	layer_height	wall_thickness	infill_density	infill_pattern	nozzle_temperature	bed_temperature	print_speed	material	fan_speed	roughness
0	0.02	8.0	90	grid	220	60	40	abs	0	25
1	0.02	7.0	90	honeycomb	225	65	40	abs	25	32
2	0.02	1.0	80	grid	230	70	40	abs	50	40
3	0.02	4.0	70	honeycomb	240	75	40	abs	75	68
4	0.02	6.0	90	grid	250	80	40	abs	100	92

Handling Missing Data

```
ds.isnull().any()
```

```
layer_height          False
wall_thickness        False
infill_density        False
infill_pattern        False
nozzle_temperature    False
bed_temperature       False
print_speed           False
material              False
fan_speed             False
roughness             False
tension_strenght      False
elongation            False
dtype: bool
```

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

```
sc=MinMaxScaler()
x_scaled = sc.fit_transform(x) # Scale the data
x = pd.DataFrame(x_scaled, columns=range(x_scaled.shape[1])) # Create DataFrame with
x.head()
```

	0	1	2	3	4	5	6	7	8	9	10
0	0.0	0.636364	0.888889	0.0	0.4	0.000	0.0	0.00	0.011527	0.411765	0.275862
1	0.0	0.545455	0.888889	1.0	0.5	0.125	0.0	0.25	0.031700	0.352941	0.344828
2	0.0	0.000000	0.777778	0.0	0.6	0.250	0.0	0.50	0.054755	0.117647	0.137931
3	0.0	0.272727	0.666667	1.0	0.8	0.375	0.0	0.75	0.135447	0.176471	0.034483
4	0.0	0.454545	0.888889	0.0	1.0	0.500	0.0	1.00	0.204611	0.029412	0.103448