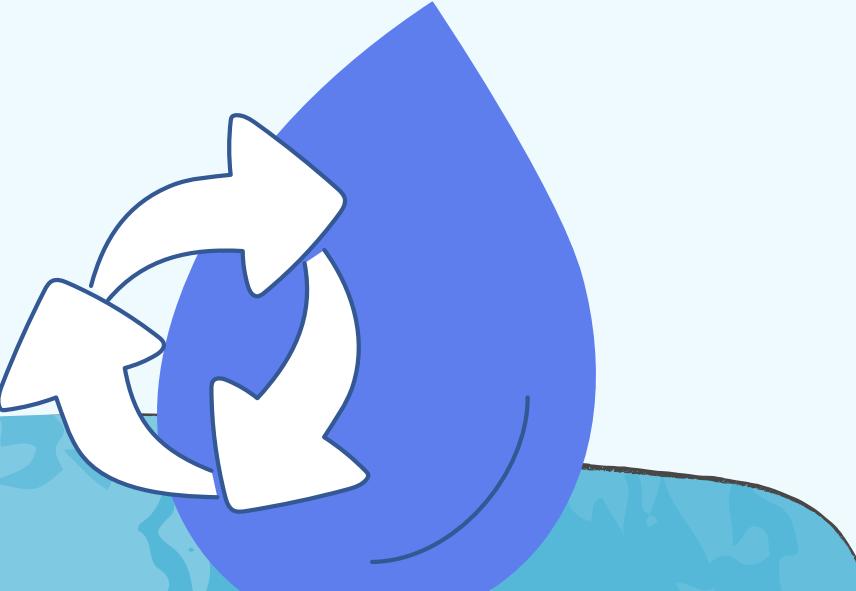


WATER QUALITY ANALYSIS

KAVYA SRI GULLIPALLI-AP21110010060



CONTENTS

Introduction

libraries
used

Abstract

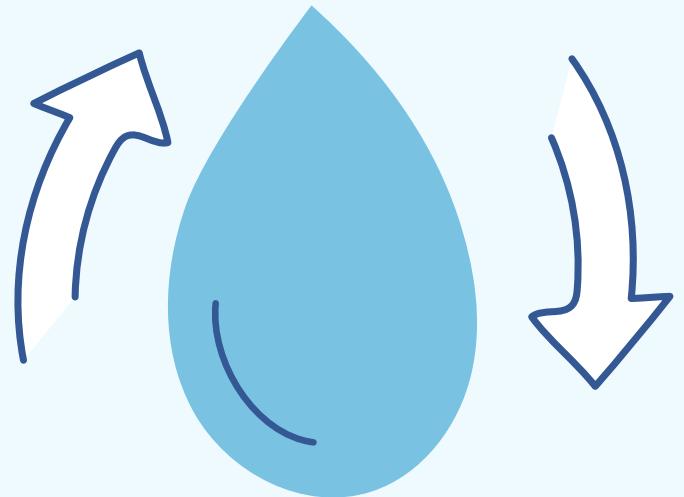
Visualisations

SYSTEM
REQUIREMENTS

Conclusion

INTRODUCTION

This data analysis project is based on verifying the samples of water, to know what type of samples are suitable for drinking.



ABSTRACT

- An extensive dataset is used in this study, which is obtained from kaggle.com
- The primary objective of this project include an in-depth analysis of various chemical forms such as ph, hardness, solids, etc...
- The project will provide a comprehensive report depicting key insights of chemical forms of water samples

SYSTEM REQUIREMENTS



LIBRARIES USED

PANDAS :

- Used for data manipulation and analysis.
- Offers powerful data structures like `DataFrames` for handling structured data.
- Provides functionalities for filtering, grouping, and transforming data efficiently

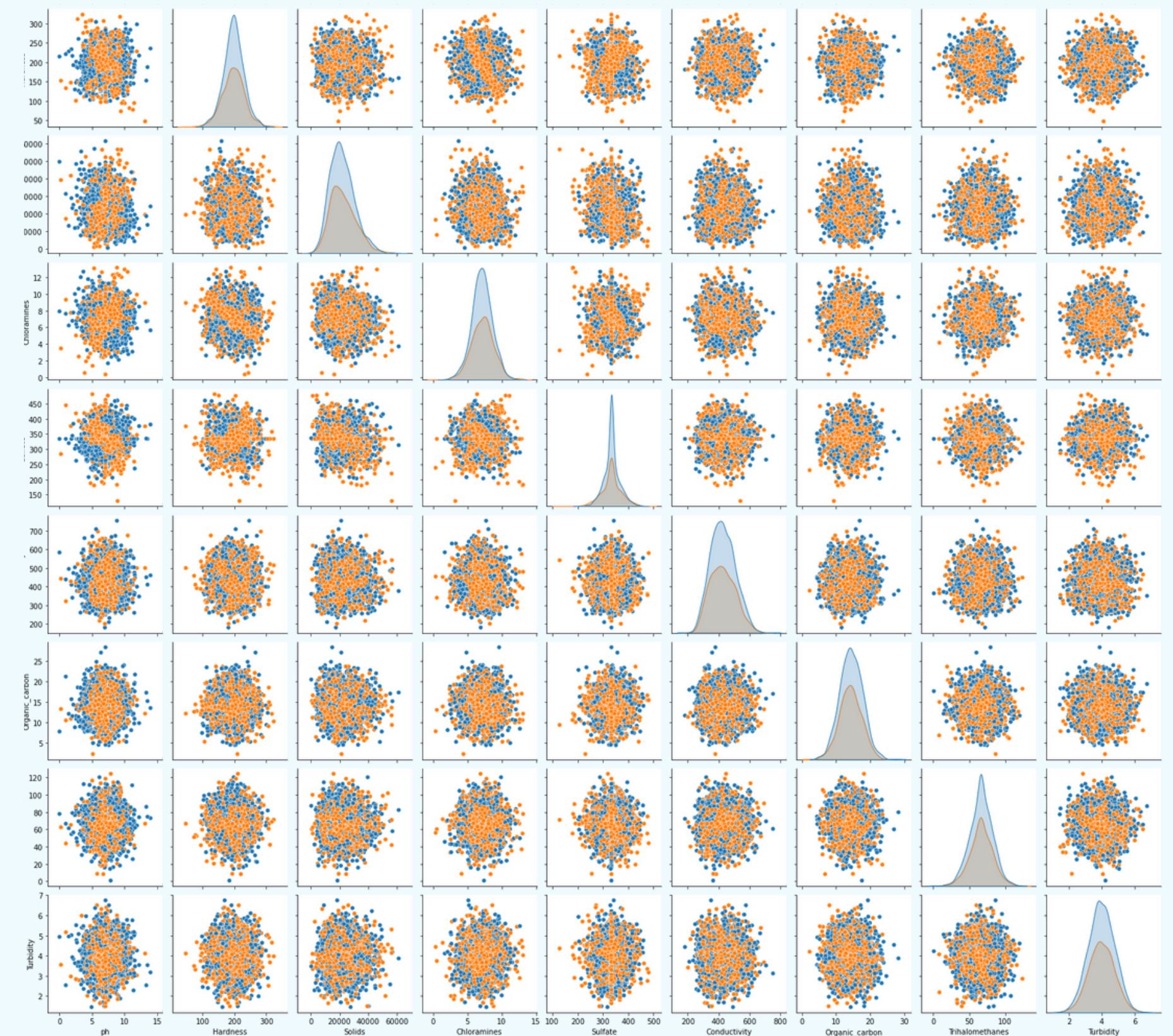
NUMPY :

- Used for numerical computations and array operations.

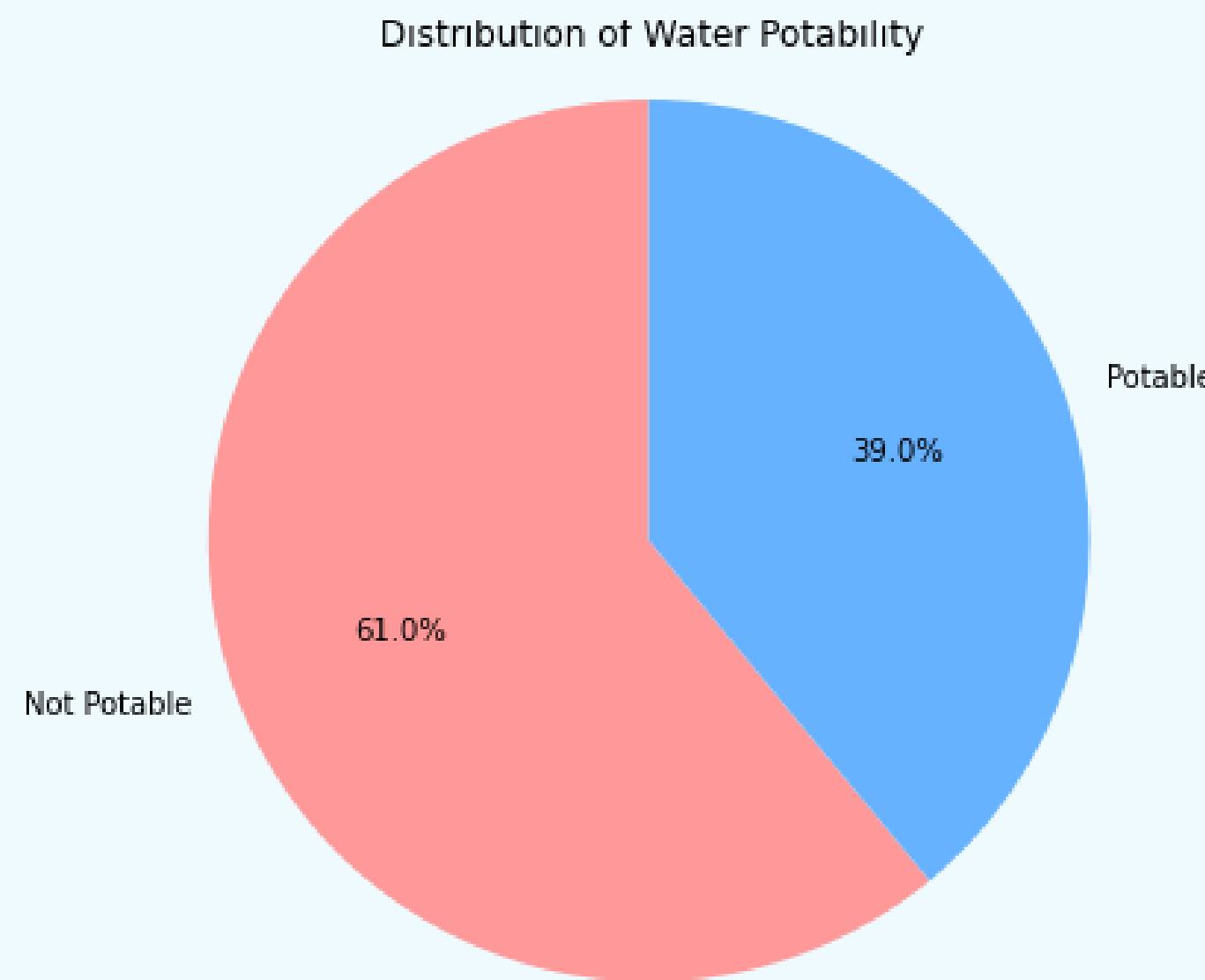
MATPLOTLIB:

- Used for creating basic static visualizations like line plots, bar charts, and histogram

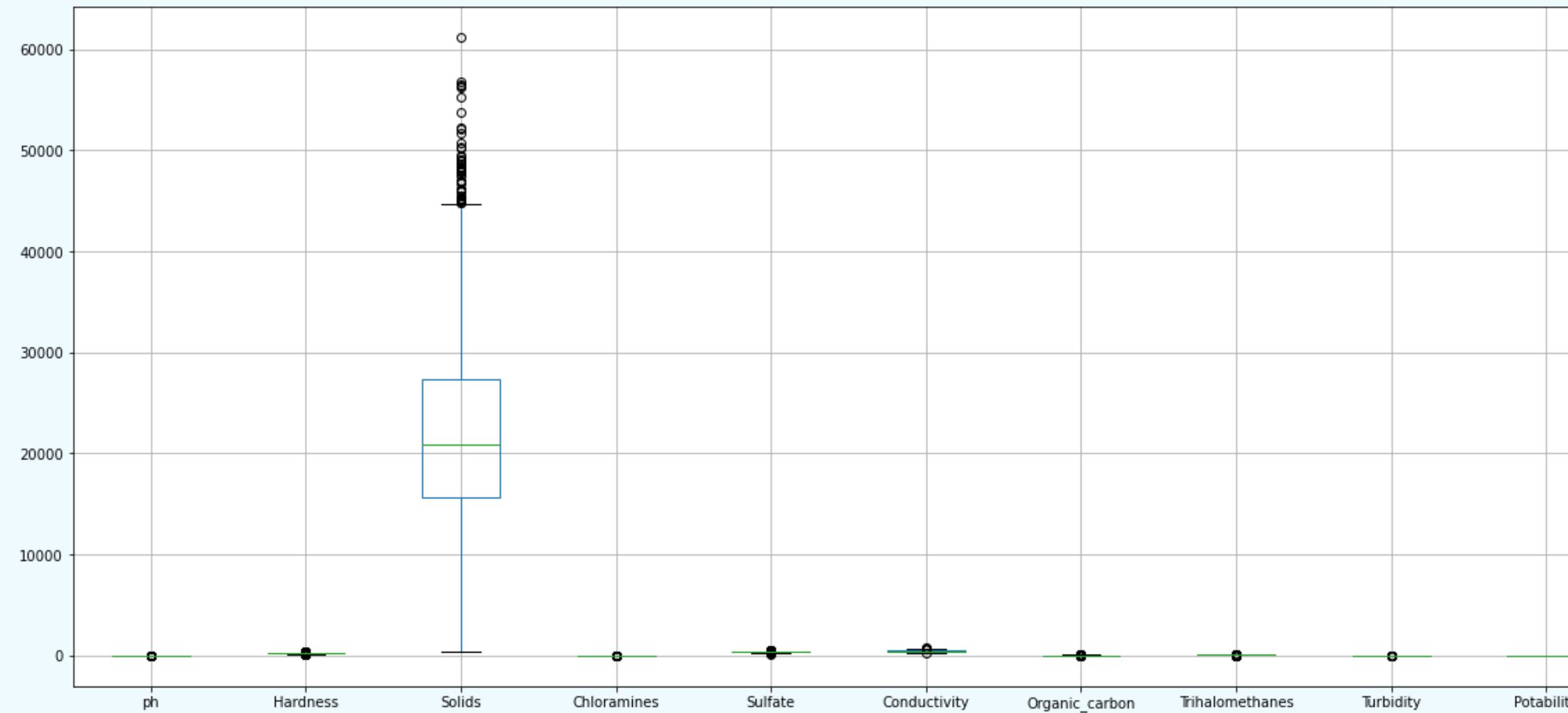
PAIR PLOT OF ALL THE CHEMICAL FORMS



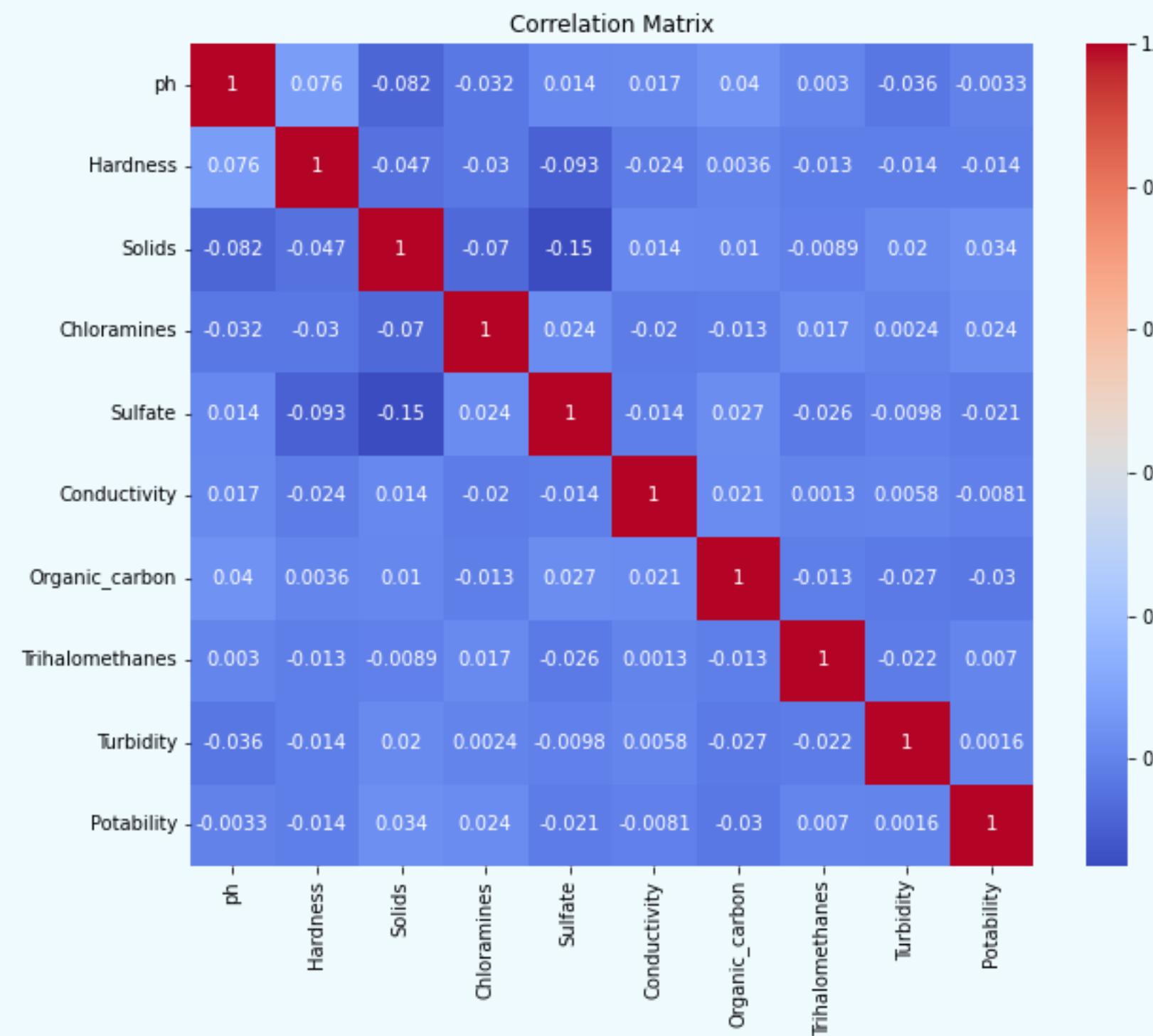
PERCENTAGE OF DISTRIBUTION OF WATER POTABILITY



BOX PLOT OF ALL CHEMICAL FORMS



CORRELATION MATRIX



WATER SAMPLES SUITABLE FOR DRINKING

Water samples predicted to be suitable for drinking:

	ph	Hardness	Solids	Chloramines	Sulfate	\
26	3.445062	207.926260	33424.768678	8.782147	384.007006	
30	7.181449	209.625601	15196.229987	5.994679	338.336431	
32	10.433291	117.791230	22326.892046	8.161505	307.707509	
44	4.758439	183.349454	21568.428779	4.731349	333.775777	
67	7.080795	103.464759	27420.167425	8.417305	333.775777	
...	
3269	11.491011	94.812545	37188.826022	9.263166	258.930600	
3272	7.808856	193.553212	17329.802160	8.061362	333.775777	
3273	9.419510	175.762646	33155.578218	7.350233	333.775777	
3274	5.126763	230.603758	11983.869376	6.303357	333.775777	
3275	7.874671	195.102299	17404.177061	7.509306	333.775777	

Conductivity Organic_carbon Trihalomethanes Turbidity Potabil

26	441.785876	13.805902	30.284597	4.184397	
30	342.111286	7.922598	71.537953	5.088860	
32	412.986834	12.890709	65.733478	5.057311	
44	403.944168	18.668229	66.912400	4.542801	
67	485.974500	11.351133	67.869964	4.620793	
...	
3269	439.893618	16.172755	41.558501	4.369264	
3272	392.449580	19.903225	66.396293	2.798243	
3273	432.044783	11.039070	69.845400	3.298875	
3274	402.883113	11.168946	77.488213	4.708658	
3275	327.459760	16.140368	78.698446	2.309149	

Predicted_Potability

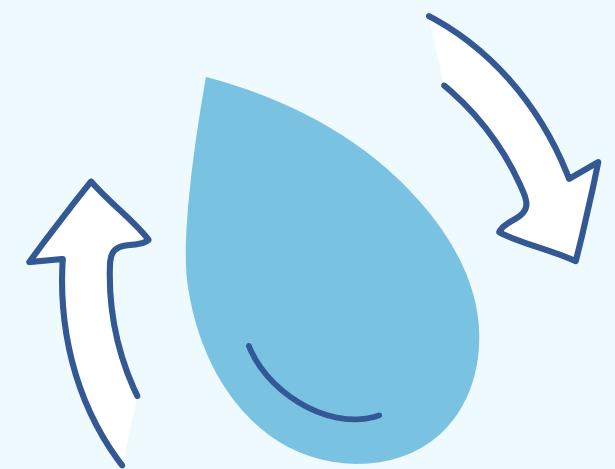
26	1
30	1
32	1
44	1
67	1
...	...
3269	1
3272	1
3273	1
3274	1
3275	1

[1185 rows x 11 columns]

CONCLUSION

As we conclude this data analysis project, our report offers a detailed information of chemical forms involved in water samples,

By Data Analysis using python, we provide valuable insights and gain deeper understanding of water samples, which provide the detailed understanding of what type of water is suitable for drinking and unsafe for drinking



**THANK
you**

