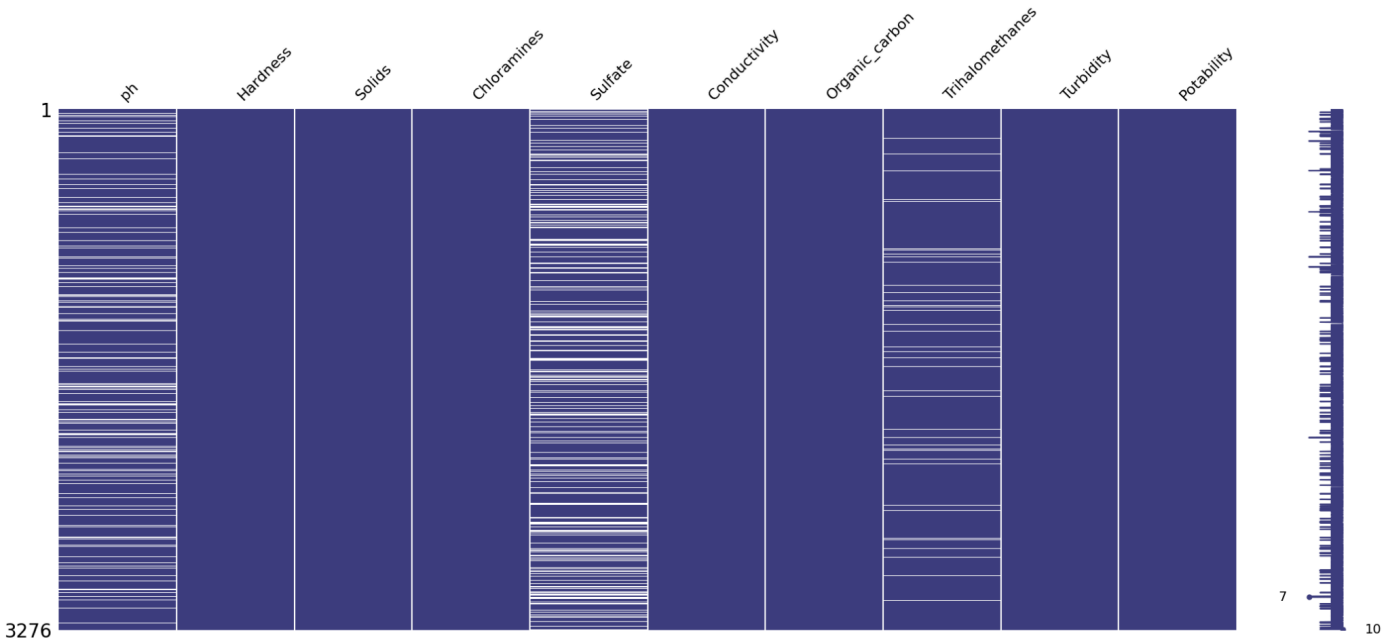


Fig. 1: Visualization of Missing Values.

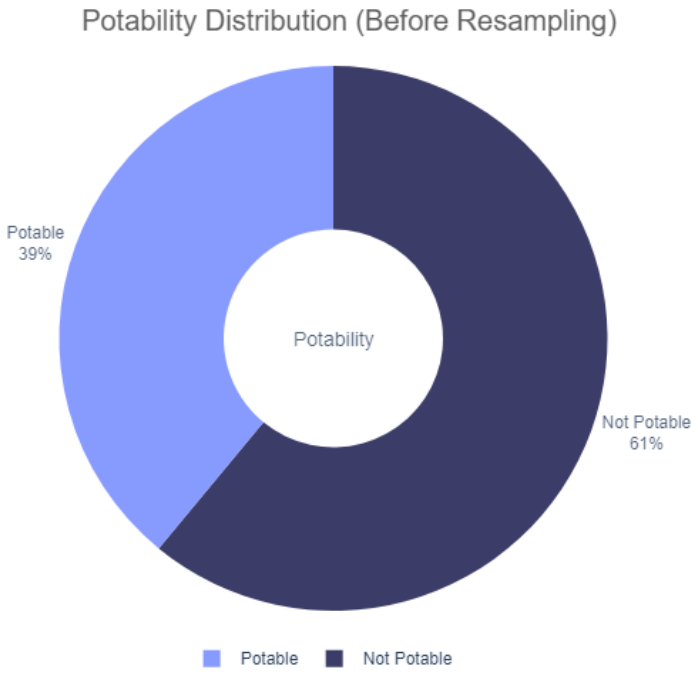


Fig. 2: Potability Distribution (Before Resampling)

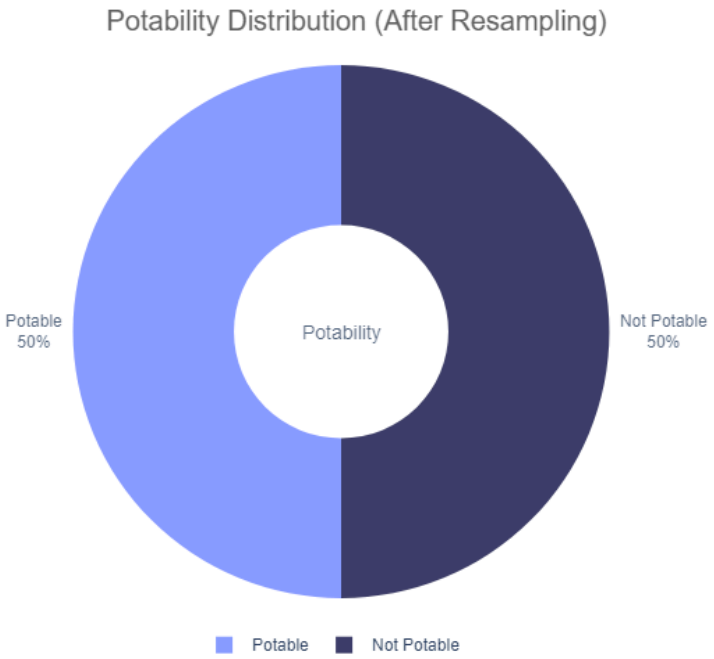


Fig.3: Potability Distribution (After Resampling)

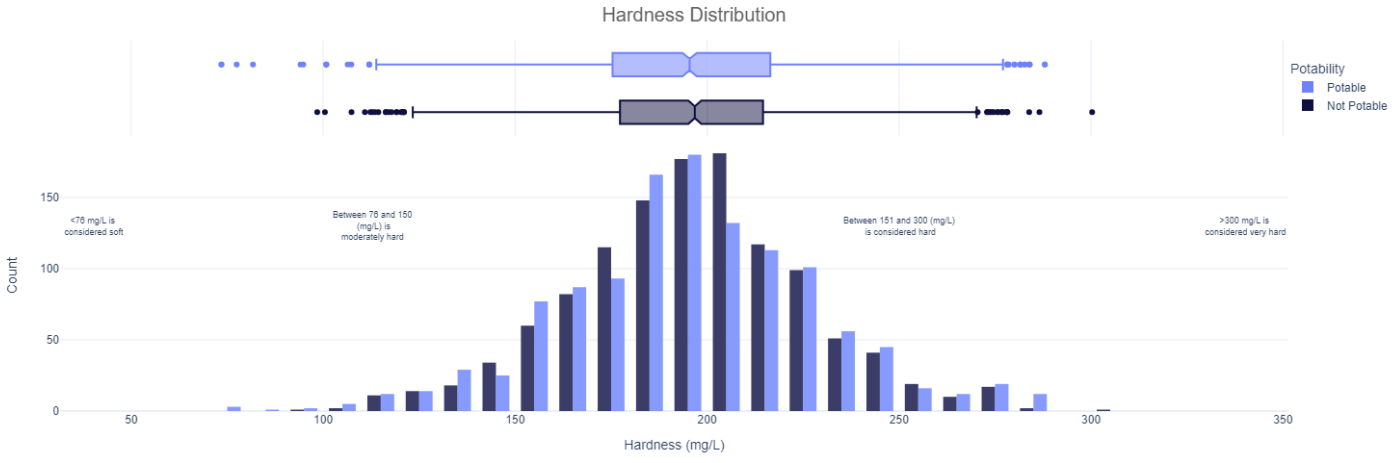


Fig. 4: Distribution of Hardness

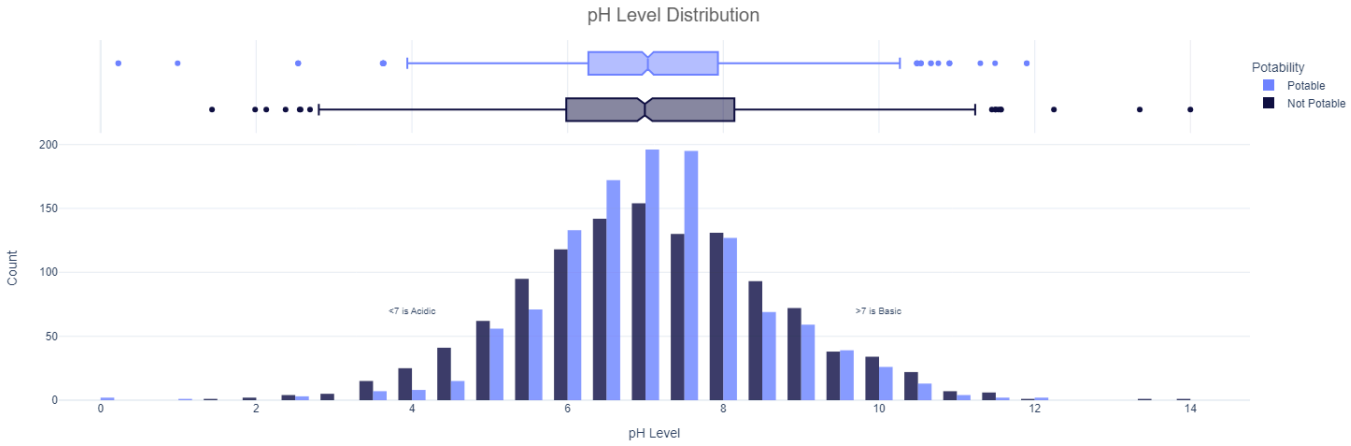


Fig. 5: Distribution of pH Level

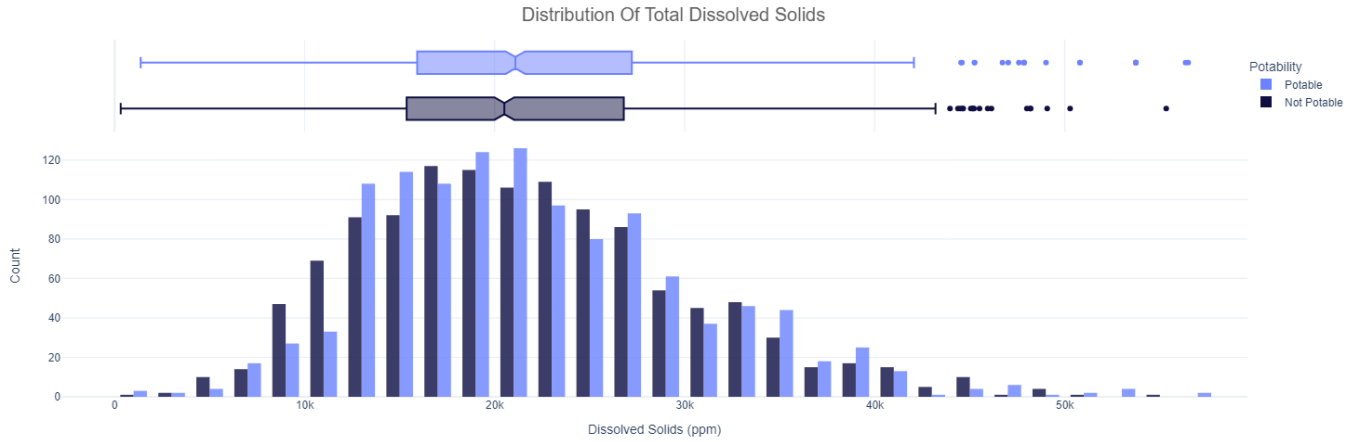


Fig. 6: Distribution of Total Dissolved Solids

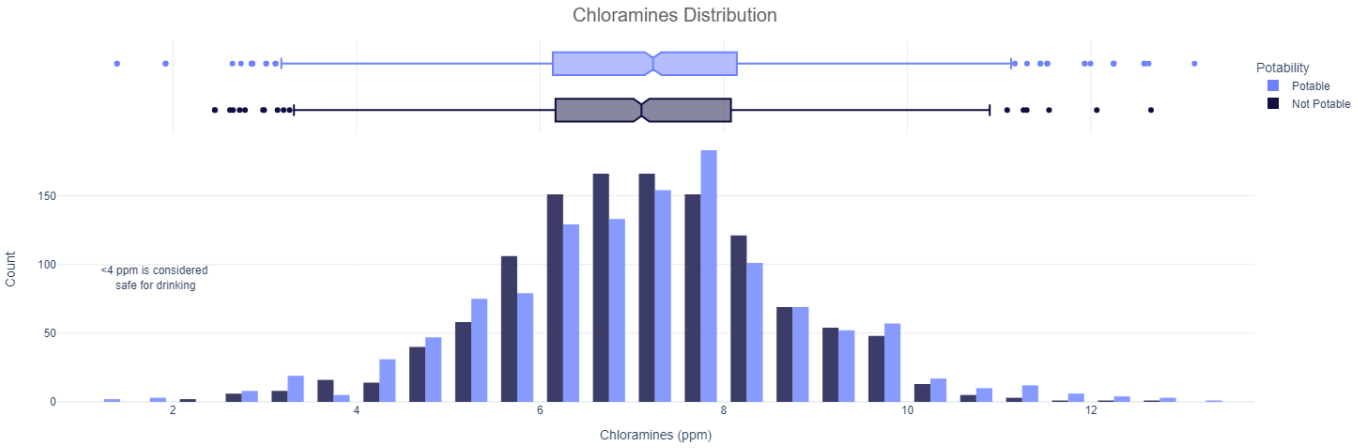


Fig. 7: Distribution of Chloramines

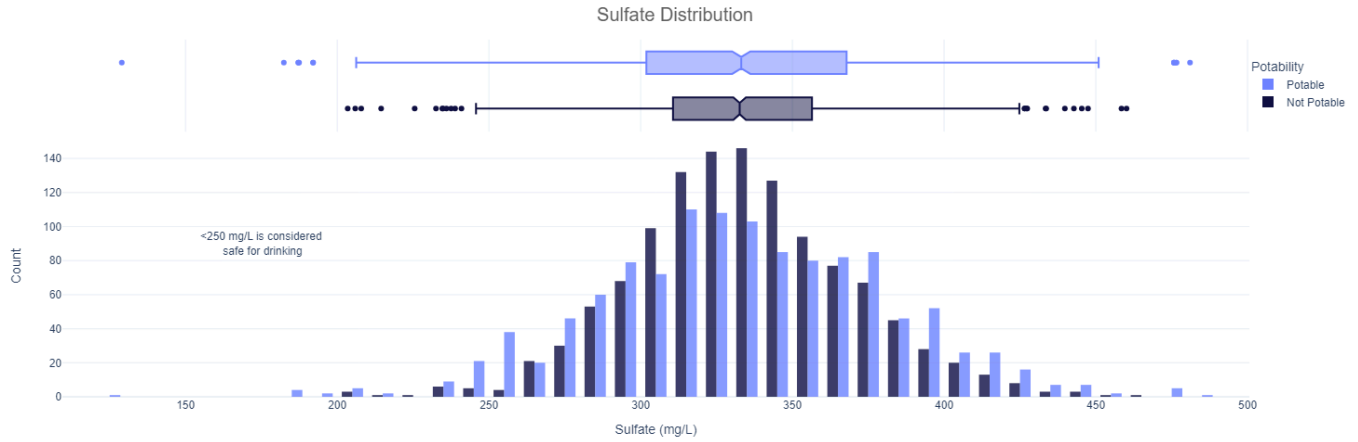


Fig. 8: Distribution of Sulfates

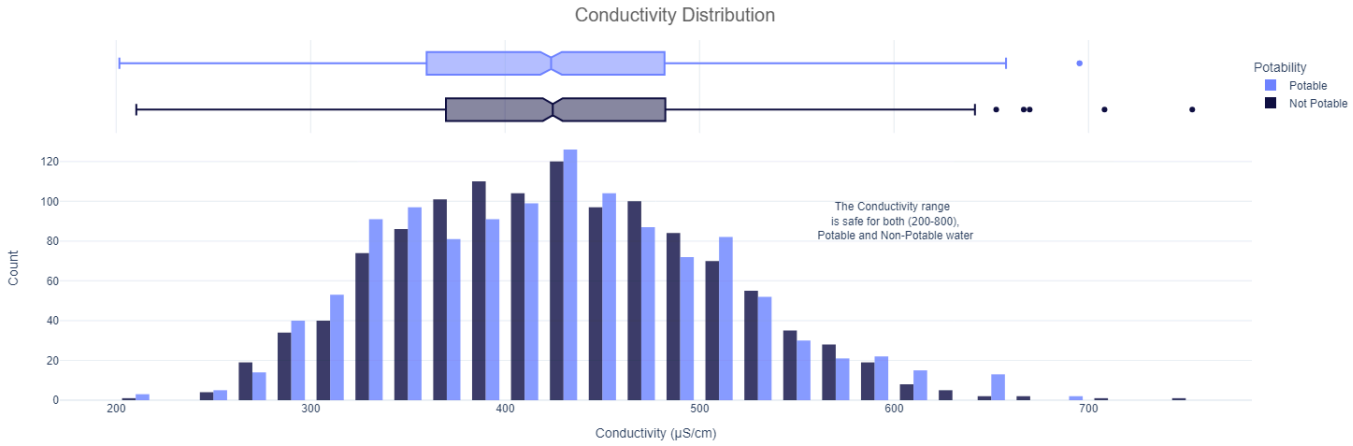


Fig. 9: Distribution of Conductivity

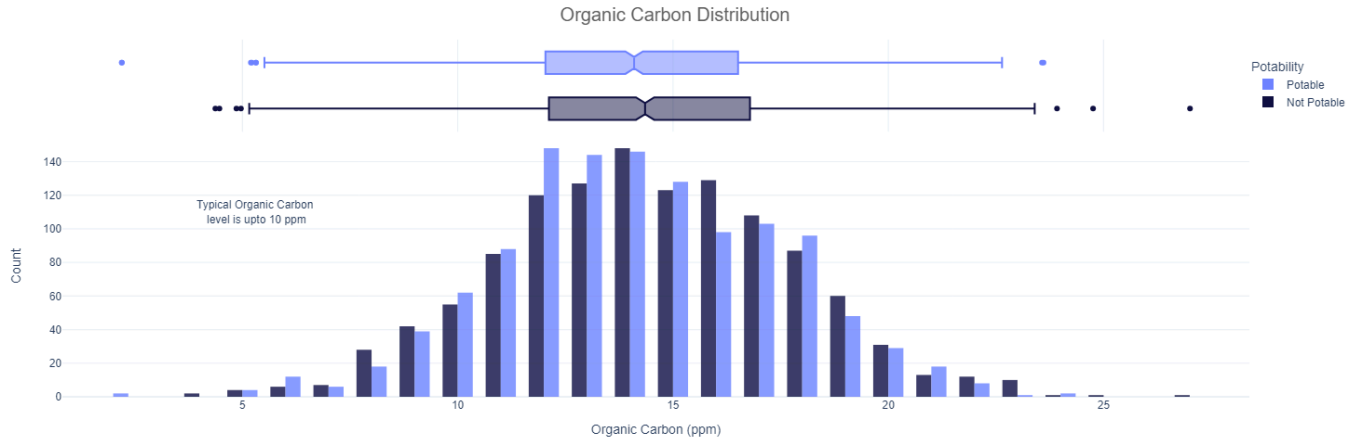


Fig. 10: Distribution of Organic Carbon

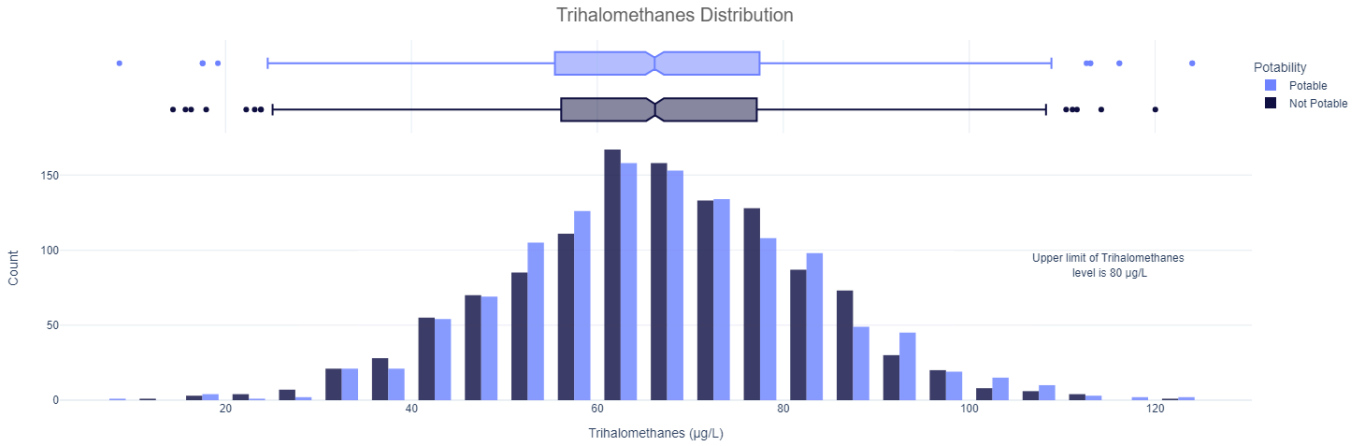


Fig. 11: Distribution of Trihalomethanes

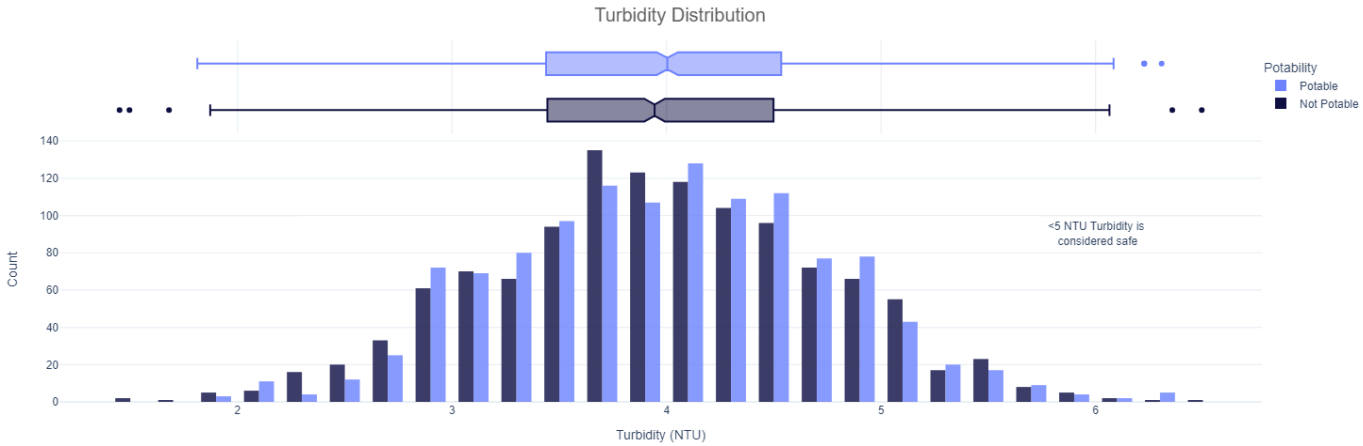


Fig. 12: Distribution of Turbidity

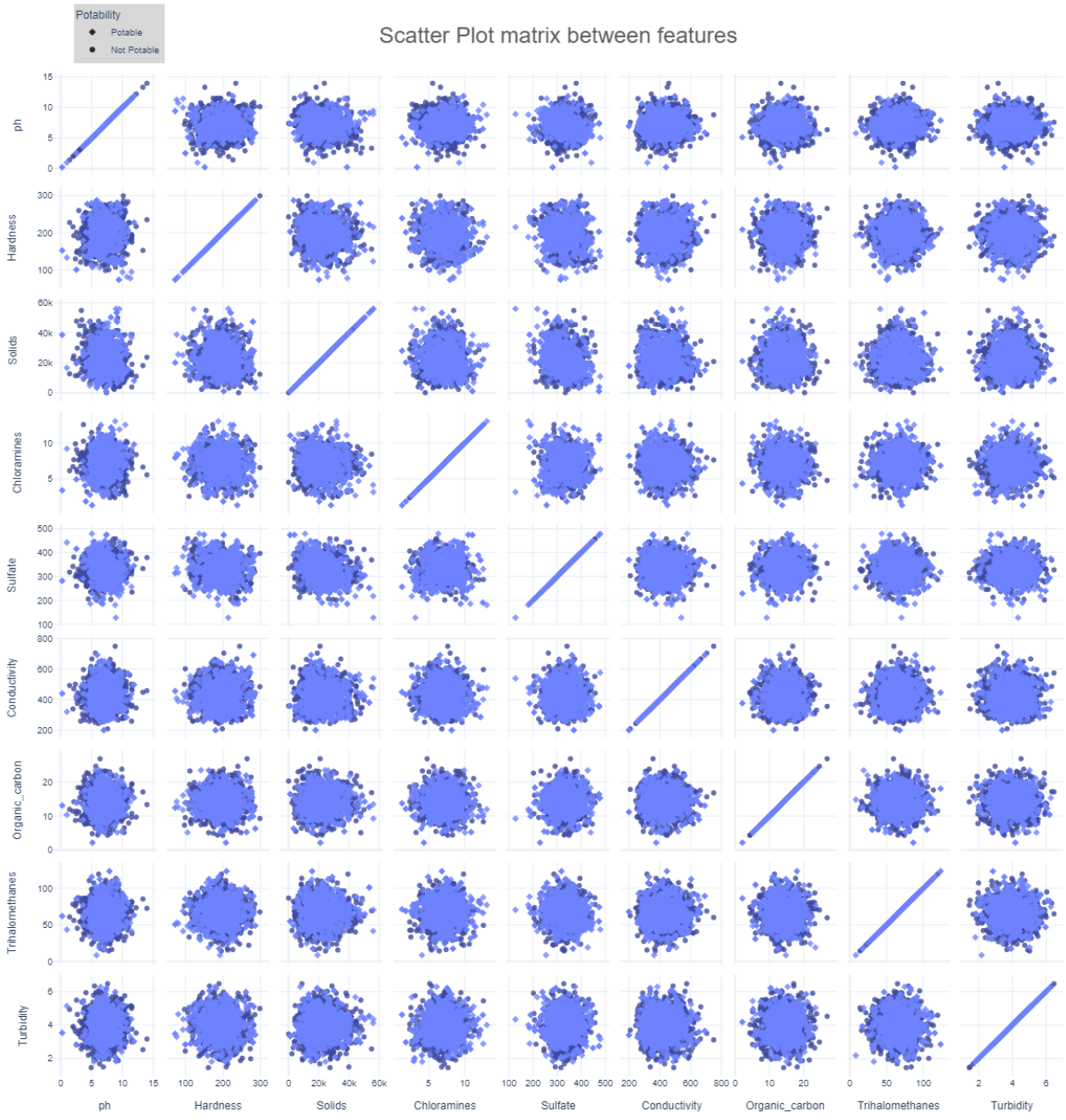


Fig. 13: Scatter Plot between Features

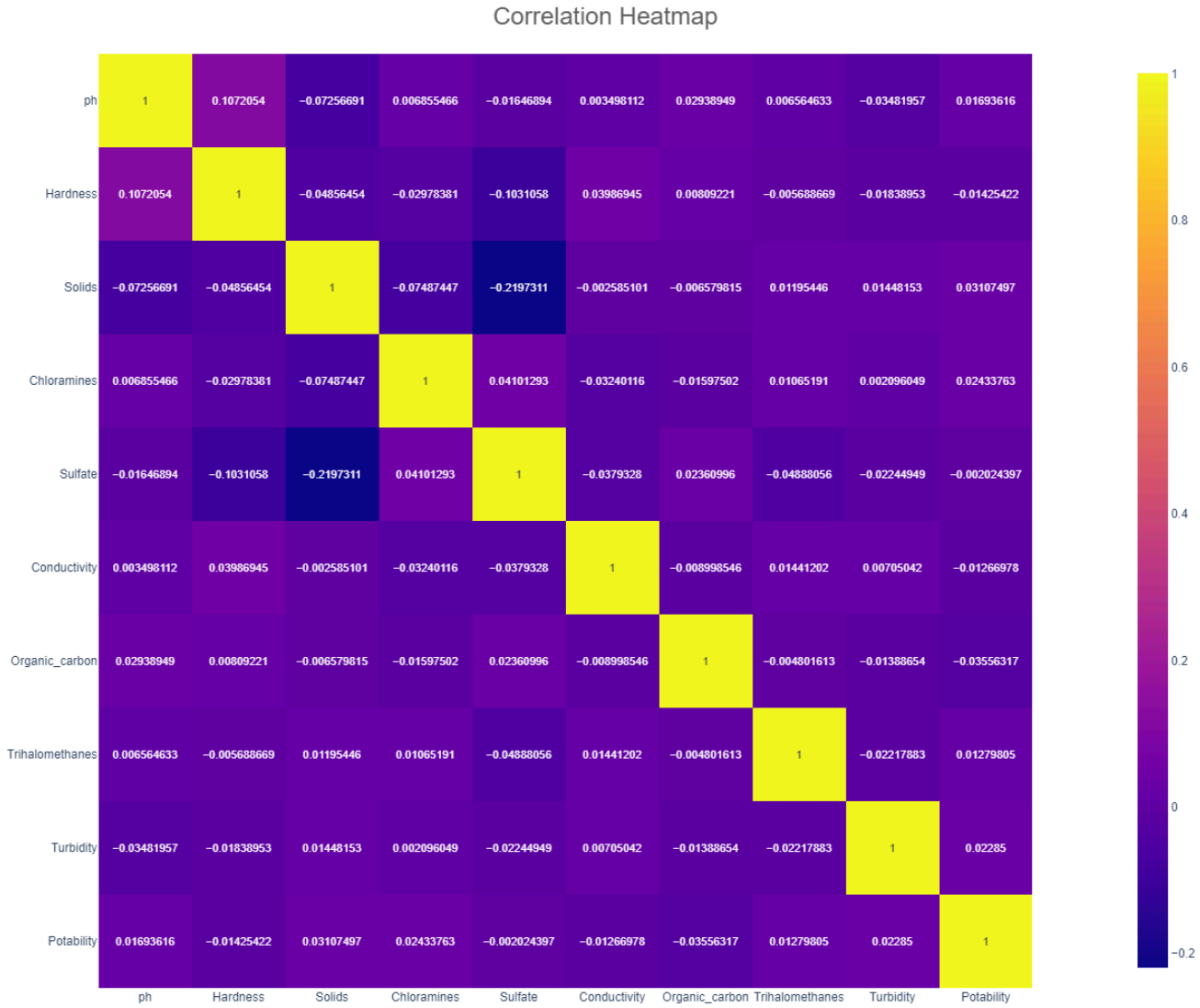
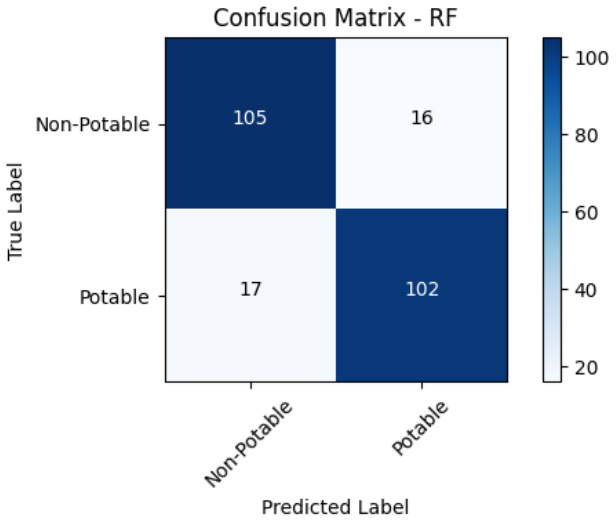
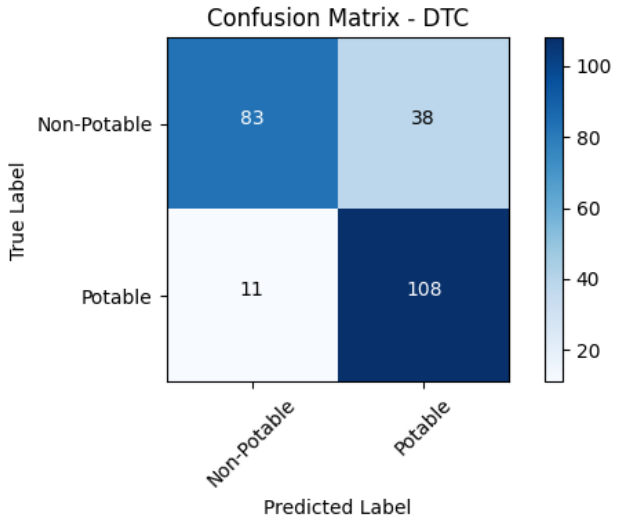


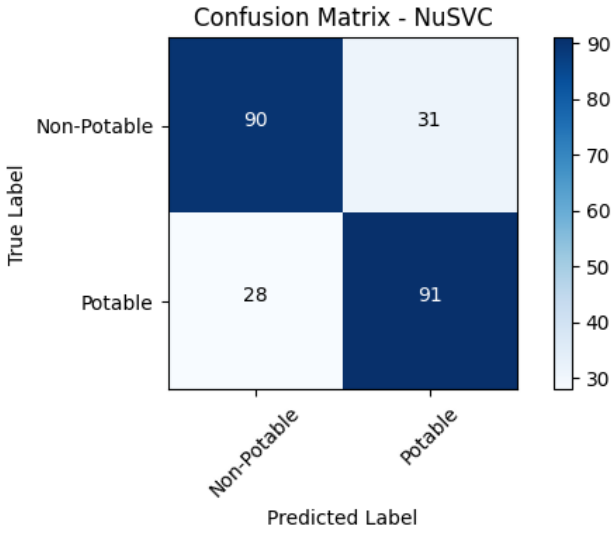
Fig. 14: Correlation Heatmap



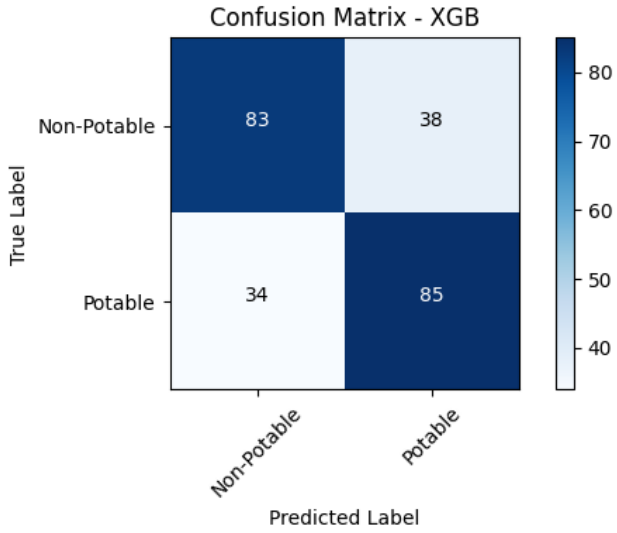
(a) Random Forest



(b) Decision Tree Classifier

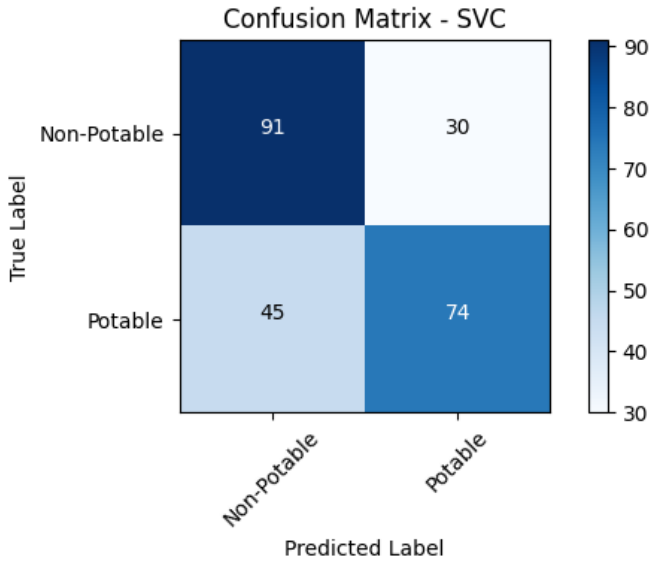


(c) Nu-Support Vector Classifier

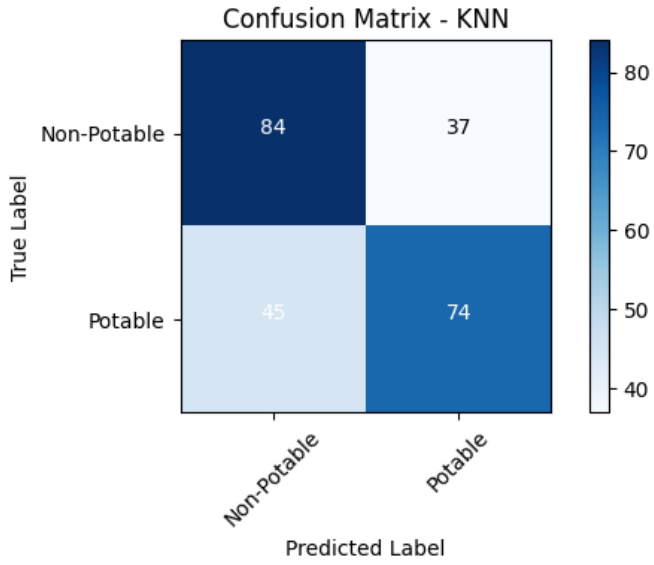


(d) Gradient Boosting Classifier

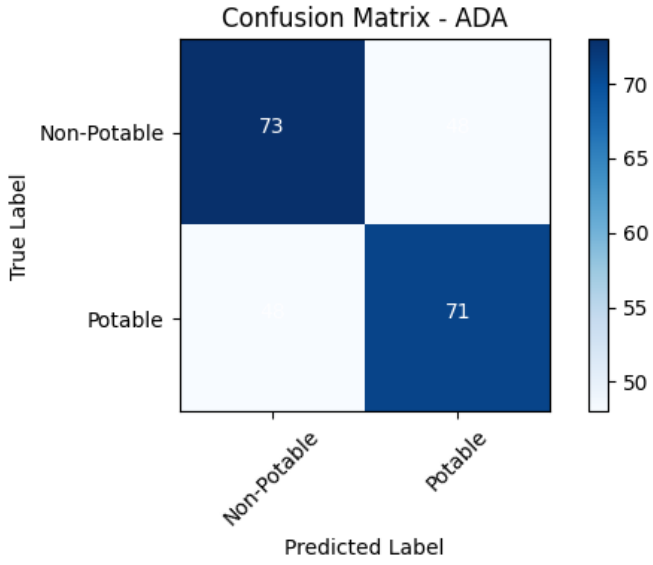
Fig. 15: Confusion Matrix



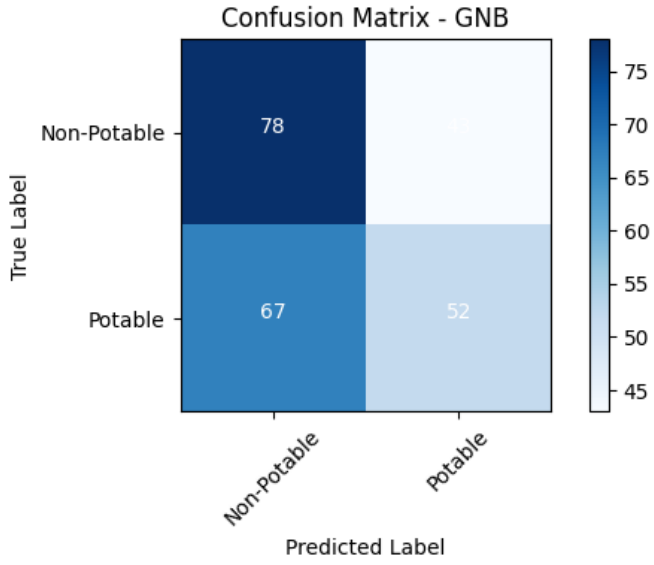
(e) Support Vector Classifier



(f) K-Nearest Neighbors



(g) AdaBoost Classifier



(h) Gaussian Naive Bayes

Fig. 15: Confusion Matrix

Machine Learning Model

Visualizations

	ph	Hardness	Solids	Chloramines	Sulfate	Conductivity	Organic_carbon	Trihalomethanes	Turbidity	Potability
0	0.2566417391	0.5402640694	0.7963480066	0.6249685203	0.5543444707	0.5285383944	0.2605027313	0.6634736839	0.7009406497	0
1	0.3678849875	0.7089020888	0.4716277576	0.1892009971	0.6834872757	0.6694579867	0.4263271806	0.3127511301	0.4108141506	0
2	0.5458176064	0.5971250538	0.4865287798	0.3605704531	0.5789944806	0.354231704	0.4146536627	0.5597459009	0.7069373546	0
3	0.4509654983	0.5689835267	0.46516572	0.3736378468	0.5495310655	0.6626187702	0.4209332626	0.408272696	0.3826474942	1
4	0.461016216	0.3647682808	0.2216780558	0.7130994111	0.6043490315	0.393952719	0.3298403444	0.6041098792	0.4601445428	0
5	0.5324300212	0.6602286442	0.25669768	0.4550682535	0.6903629387	0.5069973457	0.2701047276	0.4640086794	0.5565797932	1
6	0.345193151	0.650168286	0.3679540773	0.5619227128	0.2941471595	0.311760028	0.4045347702	0.3237918015	0.5996628884	0
7	0.4113206567	0.2904318732	0.1902356733	0.3341067124	0.5974704462	0.4438363263	0.3372906727	0.7027704949	0.3690596612	1
8	0.4014504453	0.4342128081	0.1382623691	0.5038933248	0.5713530376	0.4370654665	0.369446505	0.7565345862	0.5058887452	1
…	…	…	…	…	…	…	…	…	…	…
2392	0.5183643091	0.7085901056	0.2710852667	0.5504821277	0.6678557601	0.3473377384	0.5225029521	0.4865174526	0.3854404736	1
2393	0.5769966723	0.7528951934	0.5482485821	0.284253978	0.4440864146	0.393313147	0.5108951555	0.4148839479	0.6460981848	1
2394	0.4934073403	0.8422979227	0.2787617141	0.8150331743	0.3779913888	0.3826009274	0.1255824316	0.5843504791	0.698133328	1
2395	0.4945066463	0.5795013711	0.4323860974	0.5699736945	0.4310579951	0.2976594164	0.3469850707	0.4447591457	0.5525939651	0
2396	0.4857016854	0.6056792691	0.5152889668	0.393719261	0.3500585313	0.3457240128	0.4310209444	0.5934932791	0.6241385891	1
2397	0.5681898222	0.5533602585	0.1549668605	0.4964939466	0.5948258036	0.2298497624	0.2798647802	0.2960952861	0.4330928524	1
2398	0.4274954135	0.6728043775	0.4340318076	0.4530158366	0.4810073871	0.4134085265	0.4972880952	0.5412134015	0.4441969376	0
2399	0.2954254996	0.9023592235	0.3445313647	0.3037072296	0.8113973385	0.2328118811	0.3665616219	0.5379216767	0.5378521747	1

Table 1: Normalized Dataset