

# Final Project

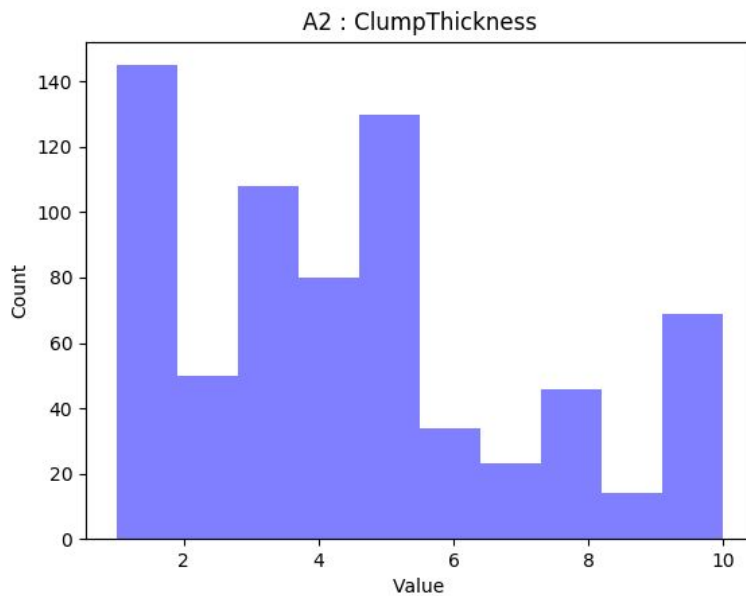
<b>Phase 1 - Statistical Analysis</b>	<b>2</b>
Summary Table	2
Histograms	2
<b>Phase 2 - k-means Clustering</b>	<b>7</b>
Program Output	7

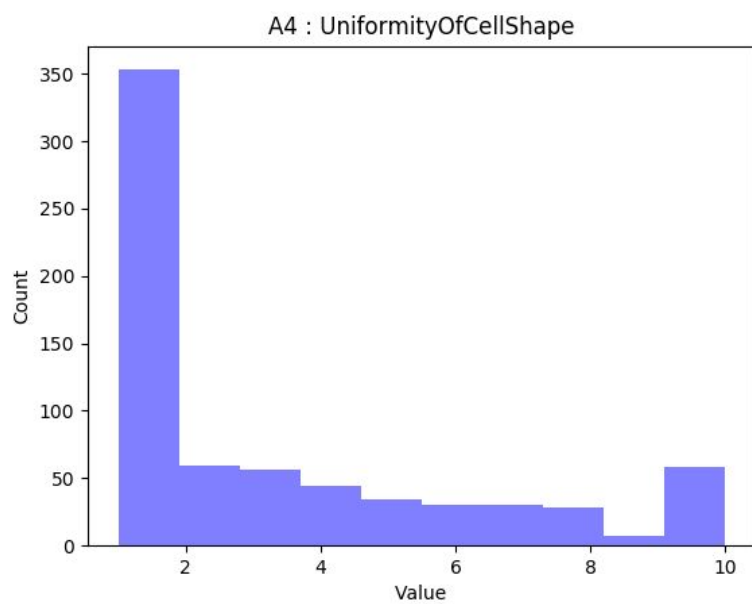
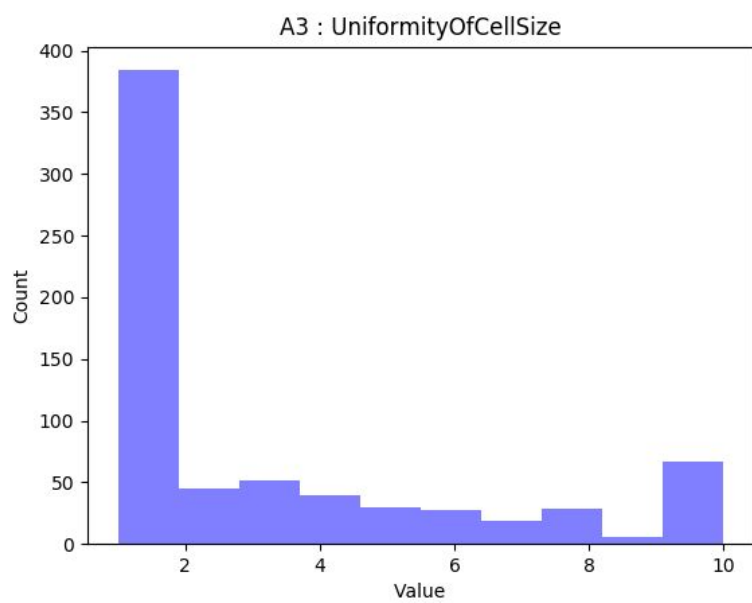
# Phase 1 - Statistical Analysis

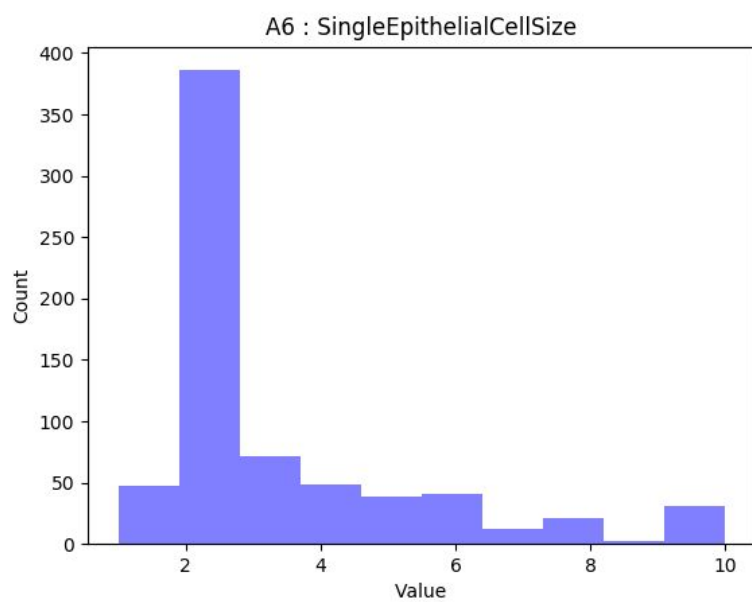
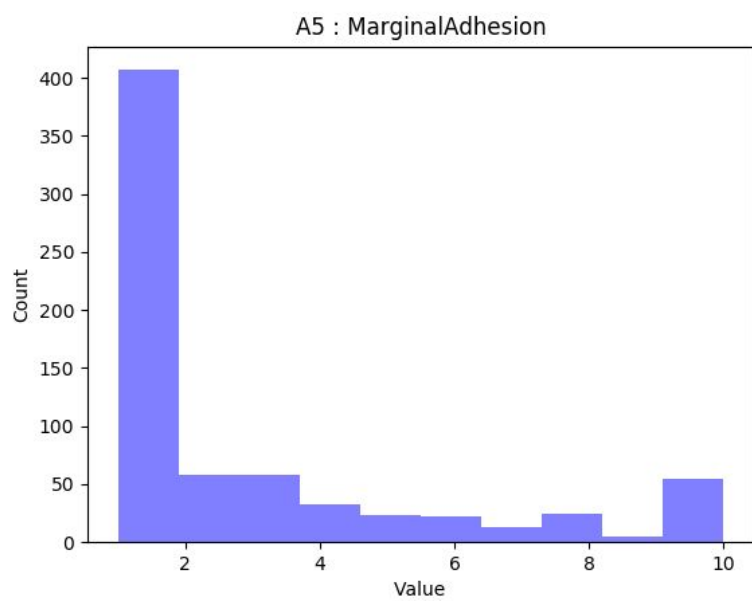
## Summary Table

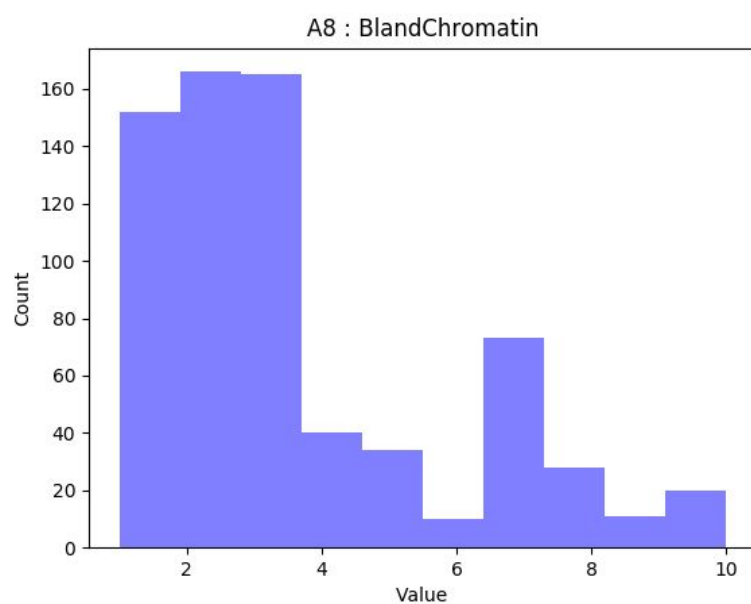
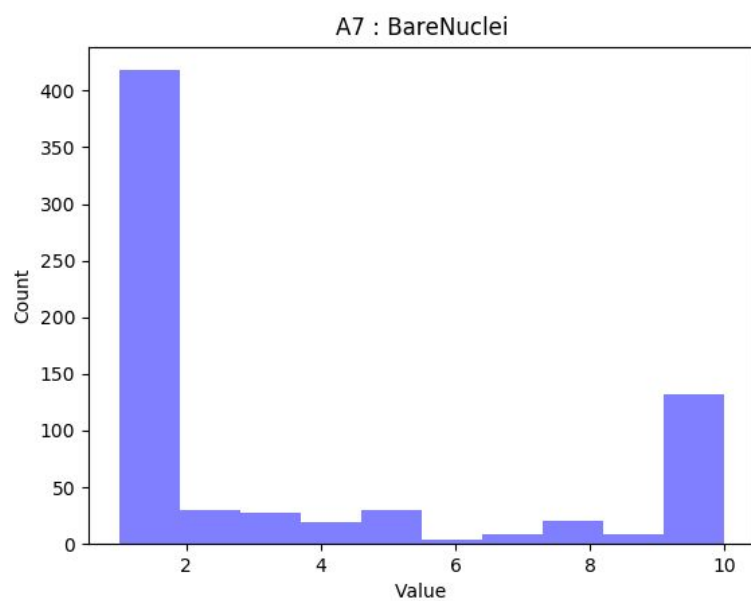
Attribute		Min	Max	Mean	Median	StdDev	Variance
A2	ClumpThickness	1.000	10.000	4.418	4.000	2.816	7.928
A3	UniformityOfCellSize	1.000	10.000	3.134	1.000	3.051	9.311
A4	UniformityOfCellShape	1.000	10.000	3.207	1.000	2.972	8.832
A5	MarginalAdhesion	1.000	10.000	2.807	1.000	2.855	8.153
A6	SingleEpithelialCellSize	1.000	10.000	3.216	2.000	2.214	4.903
A7	BareNuclei	1.000	10.000	3.486	1.000	3.622	13.118
A8	BlandChromatin	1.000	10.000	3.438	3.000	2.438	5.946
A9	NormalNucleoli	1.000	10.000	2.867	1.000	3.054	9.325
A10	Mitoses	1.000	10.000	1.589	1.000	1.715	2.941

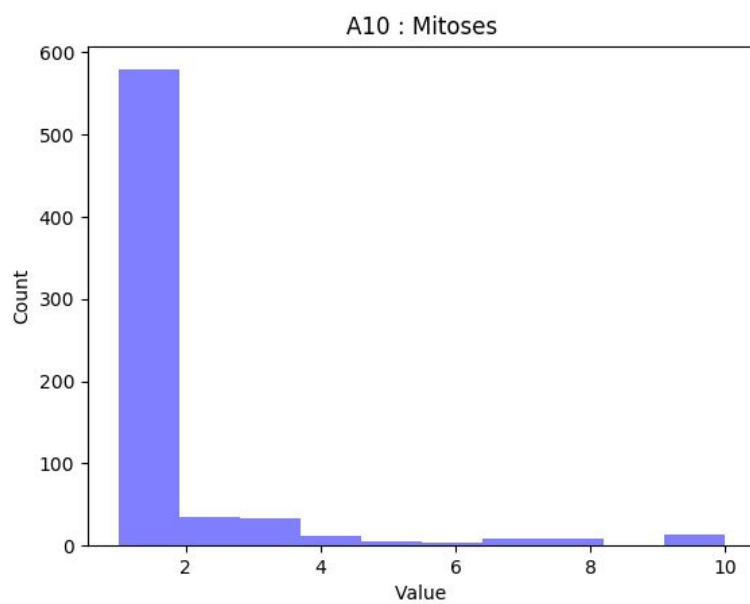
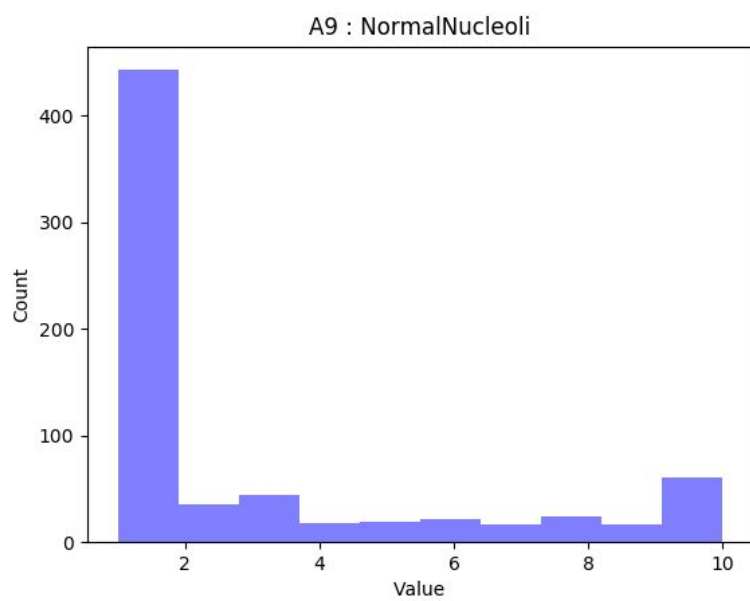
## Histograms













u4: A2=3.04301 A3=1.30108 A4=1.44301 A5=1.33763 A6=2.08817 A7=1.29677 A8=2.10323 A9=1.25161  
A10=1.10968

Cluster Assignment: u2-Benign

N	Index	Id	Class	Predicted
0	1	1002945	2	2
1	3	1016277	2	2
2	5	1017122	4	2
3	14	1044572	4	2
4	18	1050670	4	2
5	20	1054590	4	2
6	21	1054593	4	2
7	25	1065726	4	2
8	32	1072179	4	2
9	36	1080185	4	2
10	38	1084584	4	2
11	39	1091262	4	2
12	40	1096800	2	2
13	41	1099510	4	2
14	42	1100524	4	2
15	43	1102573	4	2
16	44	1103608	4	2
17	46	1105257	4	2
18	49	1106829	4	2
19	52	1110102	4	2

Cluster Assignment: u4-Malignant

N	Index	Id	Class	Predicted
0	0	1000025	2	4
1	2	1015425	2	4
2	4	1017023	2	4
3	6	1018099	2	4
4	7	1018561	2	4
5	8	1033078	2	4
6	9	1033078	2	4
7	10	1035283	2	4
8	11	1036172	2	4
9	12	1041801	4	4
10	13	1043999	2	4
11	15	1047630	4	4
12	16	1048672	2	4
13	17	1049815	2	4
14	19	1050718	2	4
15	22	1056784	2	4
16	23	1057013	4	4
17	24	1059552	2	4
18	26	1066373	2	4
19	27	1066979	2	4

real 43m27.623s  
user 43m21.896s  
sys 0m3.906s



