

# «به نام خدا»

تکلیف دوم – سوال اول – مرضیه علیدادی – 9631983  
(کد های مربوط، در دو فرمت py و ipynb. ضمیمه شده اند.)

1.

(a)

	COMPACTNESS	CIRCULARITY	DISTANCE_CIRCULARITY	RADIUS_RATIO	PR.AXIS_ASPECT_RATIO	MAX.LENGTH_ASPECT_RATIO	SCATTER_RATIO	ELON
0	95	48	83	178	72	10	162	
1	91	41	84	141	57	9	149	
2	104	50	106	209	66	10	207	
3	93	41	82	159	63	9	144	
4	85	44	70	205	103	52	149	
...	...	...	...	...	...	...	...	...
841	93	39	87	183	64	8	169	
842	89	46	84	163	66	11	159	
843	106	54	101	222	67	12	222	
844	86	36	78	146	58	7	135	
845	85	36	66	123	55	5	120	

846 rows × 19 columns

(b)

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 846 entries, 0 to 845
Data columns (total 19 columns):
#   Column                                Non-Null Count  Dtype
---  -
0   COMPACTNESS                          846 non-null    int64
1   CIRCULARITY                          846 non-null    int64
2   DISTANCE_CIRCULARITY                 846 non-null    int64
3   RADIUS_RATIO                        846 non-null    int64
4   PR.AXIS_ASPECT_RATIO                 846 non-null    int64
5   MAX.LENGTH_ASPECT_RATIO              846 non-null    int64
6   SCATTER_RATIO                       846 non-null    int64
7   ELONGATEDNESS                       846 non-null    int64
8   PR.AXIS_RECTANGULARITY               846 non-null    int64
9   MAX.LENGTH_RECTANGULARITY            846 non-null    int64
10  SCALED_VARIANCE_MAJOR                 846 non-null    int64
11  SCALED_VARIANCE_MINOR                 846 non-null    int64
12  SCALED_RADIUS_OF_GYRATION             846 non-null    int64
13  SKEWNESS_ABOUT_MAJOR                  846 non-null    int64
14  SKEWNESS_ABOUT_MINOR                  846 non-null    int64
15  KURTOSIS_ABOUT_MAJOR                  846 non-null    int64
16  KURTOSIS_ABOUT_MINOR                  846 non-null    int64
17  HOLLOWS_RATIO                        846 non-null    int64
18  Class                                846 non-null    object
dtypes: int64(18), object(1)
memory usage: 125.7+ KB
```

(c)

	COMPACTNESS	CIRCULARITY	DISTANCE_CIRCULARITY	RADIUS_RATIO	PR.AXIS_ASPECT_RATIO	MAX.LENGTH_ASPECT_RATIO	SCATTER_RATIO	ELON
5	107	57	106	172	50	6	255	
15	96	55	103	201	65	9	204	
18	104	54	100	186	61	10	216	
19	101	56	100	215	69	10	208	
24	99	53	105	219	66	11	204	
...	...	...	...	...	...	...	...	...
823	105	54	105	213	67	10	200	
824	106	55	96	196	60	12	221	
827	99	57	100	177	54	13	224	
835	111	58	105	183	51	6	265	
843	106	54	101	222	67	12	222	

195 rows × 19 columns

◀		▶
---	--	---

نتیجه ی ذخیره در فایل csv :

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1		COMPACT	CIRCULAR	DISTANCE	RADIUS_R	PR.AXIS_A	MAX.LENG	SCATTER	ELONGATE	PR.AXIS_R	MAX.LENG	SCALED_V	SCALED_V	SCALED_R
2	5	107	57	106	172	50	6	255	26	28	169	280	957	264
3	15	96	55	103	201	65	9	204	32	23	166	227	624	246
4	18	104	54	100	186	61	10	216	31	24	173	225	686	220
5	19	101	56	100	215	69	10	208	32	24	169	227	651	223
6	24	99	53	105	219	66	11	204	32	23	165	221	623	224
7	27	107	54	98	203	65	11	218	31	25	167	229	696	216
8	38	106	53	98	176	54	10	216	31	24	171	235	691	218
9	44	119	54	106	220	65	12	213	31	24	167	223	675	232
10	52	98	55	101	228	70	9	210	31	24	168	236	661	245
11	54	101	56	104	185	53	6	257	26	28	168	275	956	230
12	58	107	53	103	221	66	11	209	32	24	163	222	653	212
13	67	98	55	101	219	69	11	225	30	25	178	231	748	216
14	70	96	55	98	161	54	10	215	31	24	175	226	683	221
15	71	97	59	108	227	70	11	224	30	25	186	225	732	218
16	75	101	53	103	203	63	9	195	34	22	162	210	571	210
17	78	104	52	94	208	66	5	208	31	24	161	227	666	218
18	81	99	51	92	203	65	5	209	31	24	159	232	671	214
19	85	110	58	106	180	51	6	261	26	28	171	278	998	257
20	90	104	55	107	222	68	11	218	31	24	173	232	703	229
21	92	104	52	100	191	59	9	197	33	23	158	218	583	234
22	105	108	53	103	202	64	10	220	30	25	168	224	711	214
23	117	109	53	109	221	69	12	221	31	25	169	226	712	212
24	134	102	54	100	163	53	10	213	31	24	173	219	669	201
25	145	103	54	107	218	64	12	222	30	25	174	221	728	199
26	155	110	56	103	223	64	5	250	26	27	169	280	928	239

(d)

میانگین: 168.94

کمترین: 104

بیشترین: 333

انحراف معیار: 33.45

(e) بخشی از نتیجه:

	ELONGATEDNESS	ElonBin
0	42	MEDIUM
1	45	MEDIUM
2	32	MEDIUM
3	46	HIGH
4	45	MEDIUM
5	26	LOW
6	42	MEDIUM
7	48	HIGH
8	54	HIGH
9	36	MEDIUM
10	50	HIGH
11	54	HIGH
12	43	MEDIUM
13	44	MEDIUM
14	37	MEDIUM
15	32	MEDIUM
16	57	HIGH
17	36	MEDIUM
18	31	MEDIUM
19	32	MEDIUM

(f)

	Class	DISTANCE_CIRCULARITY
0	bus	76.720183
1	opel	89.089623
2	saab	88.672811
3	van	73.331658

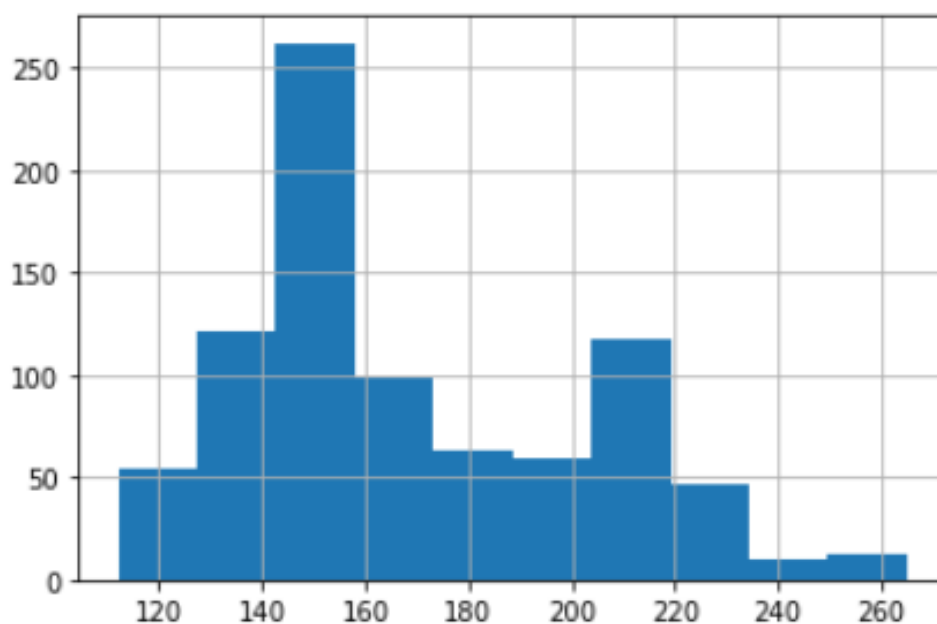
(g) با استفاده از متد pearson این کار را انجام دادم:

	COMPACTNESS	CIRCULARITY	DISTANCE_CIRCULARITY	RADIUS_RATIO	PR.AXIS_ASPECT_RATIO	MAX.LENGTH_ASPECT_RATIO
COMPACTNESS	1.000000	0.692869	0.792444	0.691659	0.093222	0.148249
CIRCULARITY	0.692869	1.000000	0.798492	0.622778	0.149692	0.247467
DISTANCE_CIRCULARITY	0.792444	0.798492	1.000000	0.771644	0.161529	0.264309
RADIUS_RATIO	0.691659	0.622778	0.771644	1.000000	0.665407	0.448048
PR.AXIS_ASPECT_RATIO	0.093222	0.149692	0.161529	0.665407	1.000000	0.648096
MAX.LENGTH_ASPECT_RATIO	0.148249	0.247467	0.264309	0.448048	0.648096	1.000000
SCATTER_RATIO	0.813003	0.860367	0.907280	0.738460	0.106064	0.166064
ELONGATEDNESS	-0.788647	-0.828755	-0.912307	-0.792557	-0.185086	-0.180086
PR.AXIS_RECTANGULARITY	0.813437	0.857925	0.895326	0.711497	0.079982	0.161497
MAX.LENGTH_RECTANGULARITY	0.676143	0.965776	0.774524	0.570154	0.129168	0.305168
SCALED_VARIANCE_MAJOR	0.764415	0.808496	0.864432	0.797588	0.274705	0.319705
SCALED_VARIANCE_MINOR	0.818632	0.852679	0.889661	0.725781	0.092182	0.144182
SCALED_RADIUS_OF_GYRATION	0.585347	0.936080	0.705818	0.539069	0.124408	0.189408
SKEWNESS_ABOUT_MAJOR	-0.252984	0.058669	-0.231660	-0.182412	0.152304	0.294304
SKEWNESS_ABOUT_MINOR	0.233693	0.148433	0.116478	0.050534	-0.056807	0.015807
KURTOSIS_ABOUT_MAJOR	0.156779	-0.015482	0.264522	0.174299	-0.034180	0.043180
KURTOSIS_ABOUT_MINOR	0.297780	-0.113047	0.147980	0.380757	0.238195	-0.027195
HOLLOWS_RATIO	0.365552	0.038677	0.335453	0.469278	0.267381	0.143381

متغیرهای CIRCULARITY و MAX.LENGTH\_RECTANGULARITY عدد مربوط به رابطه ی بینشان، خیلی نزدیک به 1 است. و باید بررسی شود که چه رابطه و وابستگی بین آن ها وجود دارد. بین متغیرهای زیر هم دو به دو این قضیه برقرار است:

CIRCULARITY و SCALED\_RADIUS\_OF\_GYRATION  
DISTANCE\_CIRCULARITY و SCATTER\_RATIO  
DISTANCE\_CIRCULARITY و ALONGATEDNESS  
DISTANCE\_CIRCULARITY و PR.AXIS\_RECTANGULARITY  
SCATTER\_RATIO و ALONGATEDNESS  
SCATTER\_RATIO و PR.AXIS\_RECTANGULARITY  
SCATTER\_RATIO و SCALED\_VARIANCE\_MAJOR  
SCATTER\_RATIO و SCALED\_VARIANCE\_MINOR  
ALONGATEDNESS و PR.AXIS\_RECTANGULARITY  
ALONGATEDNESS و SCALED\_VARIANCE\_MAJOR  
ALONGATEDNESS و SCALED\_VARIANCE\_MINOR  
PR.AXIS\_RECTANGULARITY و SCALED\_VARIANCE\_MAJOR  
PR.AXIS\_RECTANGULARITY و SCALED\_VARIANCE\_MINOR  
SCALED\_VARIANCE\_MAJOR و SCALED\_VARIANCE\_MINOR

(h)



از نمودار هیستوگرام مشخص است که تجمع داده های مربوط به ستون SCATTER\_RATIO ، در بازه 140 تا 160، به بیشترین مقدار می رسد. و کمترین تراکم، مربوط به بازه ی 240 تا 260 است.