

Summary statistics for Z-Alizadehsani dataset

n = number of samples

distinct = number of distinct samples

Info = relative information measure for continuous variables (max value of 1; lower values indicate ties, high skew)

Gmd = Gini mean difference (also known as mean absolute difference)

55 Variables 303 Observations

Age

n	missing	distinct	Info	Mean	Gmd	.05	.10
303	0	46	0.999	58.9	11.83	43	47
.25	.50	.75	.90	.95			
51	58	66	73	76			

lowest : 30 36 38 40 41, highest: 79 80 81 85 86

Weight

n	missing	distinct	Info	Mean	Gmd	.05	.10
303	0	54	0.999	73.83	13.46	55.0	60.0
.25	.50	.75	.90	.95			
65.0	74.0	81.0	89.0	93.9			

lowest : 48 50 51 53 54, highest: 102 103 105 108 120

Height

n	missing	distinct	Info	Mean	Gmd	.05	.10
303	0	44	0.998	164.7	10.65	150	152
.25	.50	.75	.90	.95			
158	165	171	176	179			

lowest : 140 144 145 147 148, highest: 184 185 186 187 188

Sex

n	missing	distinct
303	0	2

Value Female Male

Frequency 127 176

Proportion 0.419 0.581

BMI

n	missing	distinct	Info	Mean	Gmd	.05	.10
303	0	262	1	27.25	4.595	21.16	22.34
.25	.50	.75	.90	.95			
24.51	26.78	29.41	33.11	34.81			

lowest : 18.11541 18.36547 18.82679 18.93878 19.03114

highest: 37.16519 37.37024 38.28125 39.03338 40.90066

DM

n	missing	distinct
303	0	2

Value 0 1

Frequency 213 90

Proportion 0.703 0.297

HTN

n	missing	distinct
303	0	2

Value 0 1

Frequency 124 179

Proportion 0.409 0.591

Current smoker

	n	missing	distinct
	303	0	2

Value 0 1

Frequency 240 63

Proportion 0.792 0.208

Ex-smoker

	n	missing	distinct
	303	0	2

Value 0 1

Frequency 293 10

Proportion 0.967 0.033

FH

	n	missing	distinct
	303	0	2

Value 0 1

Frequency 255 48

Proportion 0.842 0.158

Obesity

	n	missing	distinct
	303	0	2

Value N Y

Frequency 92 211

Proportion 0.304 0.696

CRF

	n	missing	distinct
	303	0	2

Value N Y

Frequency 297 6

Proportion 0.98 0.02

CVA

	n	missing	distinct
	303	0	2

Value N Y

Frequency 298 5

Proportion 0.983 0.017

Airway disease

	n	missing	distinct
	303	0	2

Value N Y

Frequency 292 11

Proportion 0.964 0.036

Thyroid disease

	n	missing	distinct
	303	0	2

Value N Y

Frequency 296 7

Proportion 0.977 0.023

CHF

	n	missing	distinct
	303	0	2

Value N Y
Frequency 302 1
Proportion 0.997 0.003

DLP

 n missing distinct
303 0 2

Value N Y
Frequency 191 112
Proportion 0.63 0.37

BP

 n missing distinct Info Mean Gmd .05 .10
303 0 17 0.973 129.6 20.79 100 110

 .25 .50 .75 .90 .95
120 130 140 160 160

Value 90 100 105 110 115 118 120 125 130 132 135
Frequency 4 21 2 40 3 1 49 4 73 1 4
Proportion 0.013 0.069 0.007 0.132 0.010 0.003 0.162 0.013 0.241 0.003 0.013

Value 140 150 160 170 180 190
Frequency 53 16 19 4 7 2
Proportion 0.175 0.053 0.063 0.013 0.023 0.007

PR

 n missing distinct Info Mean Gmd .05 .10
303 0 21 0.897 75.14 9.043 64.0 70.0

 .25 .50 .75 .90 .95
70.0 70.0 80.0 85.8 90.0

lowest : 50 60 64 65 68, highest: 88 90 95 100 110

Edema

 n missing distinct
303 0 2

Value 0 1
Frequency 291 12
Proportion 0.96 0.04

Weak peripheral pulse

 n missing distinct
303 0 2

Value N Y
Frequency 298 5
Proportion 0.983 0.017

Lung rales

 n missing distinct
303 0 2

Value N Y
Frequency 292 11
Proportion 0.964 0.036

Systolic murmur

 n missing distinct
303 0 2

Value N Y
Frequency 262 41
Proportion 0.865 0.135

Diastolic murmur

 n missing distinct

303	0	2		
Value	N	Y		
Frequency	294	9		
Proportion	0.97	0.03		
Typical chest pain				
n	missing	distinct		
303	0	2		
Value	0	1		
Frequency	139	164		
Proportion	0.459	0.541		
Dyspnea				
n	missing	distinct		
303	0	2		
Value	N	Y		
Frequency	169	134		
Proportion	0.558	0.442		
Function class				
n	missing	distinct		
303	0	4		
Value	0	1	2	3
Frequency	211	1	73	18
Proportion	0.696	0.003	0.241	0.059
Atypical				
n	missing	distinct		
303	0	2		
Value	N	Y		
Frequency	210	93		
Proportion	0.693	0.307		
Nonanginal CP				
n	missing	distinct		
303	0	2		
Value	N	Y		
Frequency	287	16		
Proportion	0.947	0.053		
Low Th Ang				
n	missing	distinct		
303	0	2		
Value	N	Y		
Frequency	301	2		
Proportion	0.993	0.007		
Q wave				
n	missing	distinct		
303	0	2		
Value	0	1		
Frequency	287	16		
Proportion	0.947	0.053		
ST elevation				
n	missing	distinct		
303	0	2		
Value	0	1		
Frequency	289	14		
Proportion	0.954	0.046		

ST depression
n missing distinct
303 0 2

Value 0 1
Frequency 232 71
Proportion 0.766 0.234

T inversion
n missing distinct
303 0 2

Value 0 1
Frequency 213 90
Proportion 0.703 0.297

LVH
n missing distinct
303 0 2

Value N Y
Frequency 283 20
Proportion 0.934 0.066

Poor R progression
n missing distinct
303 0 2

Value N Y
Frequency 294 9
Proportion 0.97 0.03

BBB
n missing distinct
303 0 3

Value LBBB N RBBB
Frequency 13 282 8
Proportion 0.043 0.931 0.026

FBS
n missing distinct Info Mean Gmd .05 .10
303 0 113 0.999 119.2 49.04 77.0 80.0

.25 .50 .75 .90 .95
88.5 98.0 130.0 191.8 218.9

lowest : 62 65 69 70 71, highest: 300 320 360 377 400

Cr
n missing distinct Info Mean Gmd .05 .10
303 0 18 0.984 1.056 0.2862 0.7 0.8

.25 .50 .75 .90 .95
0.9 1.0 1.2 1.4 1.5

Value 0.50 0.60 0.70 0.80 0.90 1.00 1.10 1.15 1.20 1.30 1.40
Frequency 2 7 17 41 47 50 42 1 37 25 14
Proportion 0.007 0.023 0.056 0.135 0.155 0.165 0.139 0.003 0.122 0.083 0.046

Value 1.50 1.60 1.70 1.80 1.90 2.00 2.20
Frequency 7 5 1 1 4 1 1
Proportion 0.023 0.017 0.003 0.003 0.013 0.003 0.003

TG
n missing distinct Info Mean Gmd .05 .10
303 0 147 1 150.3 88.51 69.0 76.0

.25 .50 .75 .90 .95
90.0 122.0 177.0 250.0 308.6

lowest : 37 41 43 46 50, highest: 447 450 470 680 1050

LDL

n	missing	distinct	Info	Mean	Gmd	.05	.10
303	0	110	0.999	104.6	39.34	55.0	65.2

.25	.50	.75	.90	.95
80.0	100.0	122.0	152.8	169.8

lowest : 18 30 36 39 40, highest: 190 199 203 213 232

HDL

n	missing	distinct	Info	Mean	Gmd	.05	.10
303	0	47	0.998	40.23	11.17	26.0	28.0

.25	.50	.75	.90	.95
33.5	39.0	45.5	52.8	55.0

lowest : 15.9 18.0 22.0 23.0 24.0, highest: 66.0 67.0 82.0 83.0 111.0

BUN

n	missing	distinct	Info	Mean	Gmd	.05	.10
303	0	33	0.996	17.5	7.065	10	11

.25	.50	.75	.90	.95
13	16	20	25	31

lowest : 6 8 9 10 11, highest: 39 40 41 43 52

ESR

n	missing	distinct	Info	Mean	Gmd	.05	.10
303	0	58	0.999	19.46	16.37	3	4

.25	.50	.75	.90	.95
9	15	26	41	51

lowest : 1 2 3 4 5, highest: 68 69 76 80 90

HB

n	missing	distinct	Info	Mean	Gmd	.05	.10
303	0	66	0.999	13.15	1.809	10.14	11.00

.25	.50	.75	.90	.95
12.20	13.20	14.20	15.08	15.60

lowest : 8.9 9.0 9.4 9.9 10.0, highest: 16.6 17.0 17.2 17.5 17.6

K

n	missing	distinct	Info	Mean	Gmd	.05	.10
303	0	27	0.994	4.231	0.5009	3.50	3.70

.25	.50	.75	.90	.95
3.90	4.20	4.50	4.80	4.99

lowest : 3.0 3.1 3.2 3.3 3.4, highest: 5.2 5.3 5.4 5.6 6.6

Na

n	missing	distinct	Info	Mean	Gmd	.05	.10
303	0	25	0.99	141	4.061	135.0	137.0

.25	.50	.75	.90	.95
139.0	141.0	143.0	145.0	146.9

lowest : 128 130 131 132 133, highest: 149 150 152 153 156

WBC

n	missing	distinct	Info	Mean	Gmd	.05	.10
303	0	78	1	7562	2558	4700	5100

.25	.50	.75	.90	.95
5800	7100	8800	10400	12100

lowest : 3700 3800 4100 4200 4500, highest: 15400 16000 17000 17800 18000

Lymph

n	missing	distinct	Info	Mean	Gmd	.05	.10
303	0	50	0.998	32.4	11.17	16	19

.25	.50	.75	.90	.95
26	32	39	44	49

lowest : 7 8 9 10 11, highest: 54 55 58 59 60

Neut

n	missing	distinct	Info	Mean	Gmd	.05	.10
303	0	52	0.997	60.15	11.48	44.1	49.0

.25	.50	.75	.90	.95
52.5	60.0	67.0	72.8	78.0

lowest : 32 33 35 38 39, highest: 83 84 85 86 89

PLT

n	missing	distinct	Info	Mean	Gmd	.05	.10
303	0	135	1	221.5	60.06	159.1	170.0

.25	.50	.75	.90	.95
183.5	210.0	250.0	292.8	330.0

lowest : 25 118 129 131 145, highest: 377 380 392 410 742

EF

n	missing	distinct	Info	Mean	Gmd	.05	.10
303	0	11	0.945	47.23	9.278	30	35

.25	.50	.75	.90	.95
45	50	55	55	55

Value	15	20	25	30	35	40	45	46	50	55	60
Frequency	2	7	6	7	13	36	55	1	70	99	7
Proportion	0.007	0.023	0.020	0.023	0.043	0.119	0.182	0.003	0.231	0.327	0.023

Region RWMA

n	missing	distinct
303	0	5

Value	0	1	2	3	4
Frequency	217	26	32	14	14
Proportion	0.716	0.086	0.106	0.046	0.046

VHD

n	missing	distinct
303	0	4

Value	mild	Moderate	N	Severe
Frequency	149	27	116	11
Proportion	0.492	0.089	0.383	0.036

Cath

n	missing	distinct
303	0	2

Value	CAD	Normal
Frequency	216	87
Proportion	0.713	0.287
