No. of Printed Pages : 6 **MSTL-014(Set-II)**

M. SC. (APPLIED STATISTICS)

(MSCAST)

Term-End Practical Examination

June, 2025

MSTL-014(Set-II) : DATA ANALYSIS WITH PYTHON LAB

Time : 2 Hours *Maximum Marks : 25*

Note : *(i) Attempt any **one** question.*

(ii) Solve the question using Python software and create script file.

(iii) Mention necessary formulae, steps, hypotheses, interpretation etc.

(iv) Symbols have their usual meanings.

1. The accident related data were collected from traffic police records to examine the relationship between the number of accidents on a particular stretch of a

highway with the average speed of the vehicles and the accident time (0, for daytime and 1 for nighttime). A total of 30 subjects involved in accidents were sampled. The recorded data are shown as follows :

S. No.	No. of Accidents	Average Speed (km/h)	Accident Time
1	7	46	1
2	13	65	1
3	10	59	1
4	15	84	0
5	17	91	1
6	15	78	1
7	10	53	0
8	7	49	1
9	5	42	0
10	16	87	0
11	11	53	1
12	8	54	0
13	13	65	1
14	11	61	1

15	13	71	0
16	13	71	0
17	12	68	0
18	13	65	1
19	10	55	1
20	12	66	0
21	11	55	1
22	9	55	0
23	9	62	0
24	11	65	0
25	8	54	0
26	5	45	0
27	6	45	0
28	10	62	0
29	5	40	0
30	9	58	1

(i) Prepare a scatter plot to get an idea about the relationship among the variables. 7

(ii) Fit a linear regression model and its related analysis at 1% level of significance. 10

(iii) Does the fitted regression model satisfy the linearity and normality assumptions ? 8

2. The following data present the texture measurements for 30 food items on five measurement variables :

Oil (% oil), Density (product's density), Crispy (Crispiness measurement), Fracture (angle, in degree) and hardness :

Food Item	Oil	Density	Crispy	Fracture	Hardness
1	16.5	2955	10	23	97
2	17.7	2660	14	9	139
3	16.2	2870	12	17	143
4	16.7	2920	10	31	95
5	16.3	2957	11	26	143
6	19.1	2790	13	16	189

7	18.4	2750	13	17	114
8	17.5	2770	10	26	63
9	15.7	2955	11	23	123
10	16.4	2945	11	24	132
11	18.0	2830	12	15	121
12	17.4	2835	12	18	172
13	18.4	2860	14	11	170
14	13.9	2965	12	19	169
15	15.8	2930	9	26	65
16	16.4	2770	15	16	183
17	18.9	2650	14	20	114
18	17.3	2890	12	17	142
19	16.7	2695	13	13	111
20	19.1	2795	14	10	140
21	13.7	3000	10	27	177
22	14.7	2980	10	20	133
23	18.1	2780	13	14	150
24	17.2	2750	8	27	113
25	18.7	2825	13	20	166
26	18.1	2875	12	15	150
27	16.6	2945	10	25	100
28	17.1	2920	10	25	123
29	17.4	2845	13	19	129
30	19.4	2645	12	18	68

(i) Represent the given data on five measurement variables using a box plot. 7

(ii) Carry out the factor analysis using the number of factors equal to 2 and 3. Interpret the findings. 18

× × × × ×