SAIDS
Assignment on Module 4

(All problems to be solved by writing a python code as well exploring the same in Microsoft Excel)

For data , famous Gettysburg Address by Abraham Lincoln

Is given below :

Gettysburg Address by Abraham Lincoln

Four score and seven years ago our fathers brought forth on this continent, a new nation, conceived in Liberty, and dedicated to the proposition that all men are created equal. Now we are engaged in a great civil war, testing whether that nation, or any nation so conceived and so dedicated, can long endure. We are met on a great battlefield of that war. We have come to dedicate a portion of that field, as a final resting place for those who here gave their lives that that nation might live. It is altogether fitting and proper that we should do this. But, in a larger sense, we cannot dedicate—we cannot consecrate—we cannot hallow—this ground. The brave men, living and dead, who struggled here, have consecrated it, far above our poor power to add or detract. The world will little note, nor long remember what we say here, but it can never forget what they did here. It is for us the living, rather, to be dedicated here to the unfinished work which they who fought here have thus far so nobly advanced. It is rather for us to be here dedicated to the great task remaining before us—that from these honoured dead we take increased devotion to that cause for which they gave the last full measure of devotion—that we here highly resolve that these dead shall not have died in vain—that this nation, under God, shall have a new birth of freedom—and that government of the people, by the people, for the people, shall not perish from the earth.

Passage Data is given below :

Sr. no. of word	Letters in word
1	4
2	5
3	3
4	5
5	5
6	3
7	3
8	7
9	7
10	5
11	4
12	4
13	9
14	1
15	3
16	6
17	9
18	2
19	7
20	3
21	9
22	2
23	3
24	11
25	4
26	3
27	3
28	3
29	7
30	5
31	3
32	2
33	3
34	7
35	2
36	1
37	5

Sr. no. of word	Letters in word
44	2
45	3
46	6
47	2
48	9
49	3
50	2
51	9
52	3
53	4
54	6
55	2
56	3
57	3
58	2
59	1
60	5
61	11
62	2
63	4
64	3
65	2
66	4
67	4
68	2
69	8
70	1
71	7
72	2
73	4
74	5
75	2
76	1
77	5
78	7
79	5
80	3

Sr. no. of word	Letters in word
87	4
88	4
89	6
90	5
91	4
92	2
93	2
94	10
95	7
96	3
97	6
98	4
99	2
100	5
101	2
102	4
103	3
104	2
105	1
106	6
107	5
108	2
109	6
110	8
111	2
112	6
113	10
114	2
115	6
116	6
117	4
118	6
119	3
120	5
121	3
122	6
123	3

Sr. no. of word	Letters in word
130	2
131	3
132	5
133	3
134	4
135	5
136	2
137	3
138	2
139	7
140	3
141	5
142	4
143	6
144	4
145	3
146	4
147	8
148	4
149	2
150	3
151	4
152	3
153	2
154	3
155	5
156	6
157	4
158	4
159	3
160	4
161	2
162	2
163	3
164	2
165	3
166	6

38	5
39	3
40	7
41	7
42	4
43	6

81	5
82	3
83	4
84	4
85	5
86	5

124	4
125	3
126	9
127	4
128	4
129	11

167	6
168	2
169	2
170	9
171	4
172	2

Sr. no. of word	Letters in word
173	3
174	10
175	4
176	5
177	4
178	3
179	6
180	4
181	4
182	4
183	3
184	2
185	5
186	8
187	2
188	2
189	6
190	3
191	2
192	2
193	2
194	4
195	9
196	2
197	3
198	5
199	4
200	9
201	6
202	2
203	4
204	4
205	5
206	7
207	4
208	2
209	4

Sr. no. of word	Letters in word
216	5
217	4
218	4
219	3
220	4
221	4
222	7
223	2
224	8
225	4
226	2
227	4
228	6
229	7
230	4
231	5
232	4
233	5
234	3
235	4
236	4
237	2
238	4
239	4
240	4
241	6
242	5
243	3
244	5
245	4
246	1
247	3
248	5
249	2
250	7
251	3
252	4

Sr. no. of word	Letters in word
259	6
260	3
261	3
262	6
263	5
264	3
265	6
266	4
267	3
268	5

210	9
211	8
212	2
213	4
214	5
215	3

253	10
254	2
255	3
256	6
257	2
258	3

USE ABOVE DATA TO SOLVE THE PROBLEMS GIVEN BELOW :

Q.1 Find Mean, Mode, Median, Variance, Standard Deviation of the above population.

Q. 2 Find 10^{th}, 25^{th}, 50^{th}, 75^{th} , 90^{th} percentile for the above data.

Q. 3 Plot Bar chart & Histogram for the above population

Q. 4 Plot Scattered Plot for above population. Find correlation coefficient between col.1 & Col. 2

Q. 5 Draw box plot for above population.

Q 6 Prints word numbers whose

a. Letters are less than or equal to 4
b. Letters are less than or equal to 10

Q 7 Calculate Z score for [4,5,6,6,6,7,8,12,13,13,14,18]

Q8. Draw scattered plot & find correlation coefficient for the following data :

x	y
14.2	215
16.4	325
11.9	185
15.2	332
18.5	406
22.1	522
19.4	412
25.1	614
23.4	544
18.1	421

Q9. A clinical trial is run to compare weight loss programs and participants are randomly assigned to one of the comparison programs and are counselled on the details of the assigned program. Participants follow the assigned program for 8 weeks. The outcome of interest is weight loss, defined as the difference in weight measured at the start of the study (baseline) and weight measured at the end of the study (8 weeks), measured in pounds. (one way Anova)

Low Calorie	Low Fat	Low Carbohydrate	Control
8	2	3	2
9	4	5	2
6	3	4	-1
7	5	2	0
3	1	3	3

ANSWER:

We reject H_0 because 8.43 $\geq$ 3.24. We have statistically significant evidence at α=0.05 to show that there is a difference in mean weight loss among the four diets.

10. Solve using One-way ANOVA

Observation	A	B	C
1	8	7	6
2	10	7	8
3	6	8	10
4	7	9	6
5	9	8	4
6	0	5	5
7	0	0	7

Q 11..

Let's assume that Starbucks uses "secret shoppers" who appear to be customers to enter a store and document their experience in terms of customer service, cleanliness, and quality. The secret shoppers receive standardized training by Starbucks to ensure consistency and objectivity in their store reviews.

For its locations in the Australian cities of Sydney, Brisbane, and Melbourne, Starbucks has trained 6 secret shoppers. Each of the 6 secret shoppers will be assigned to visit the same store in each of the 3 cities. The visit sequence will be assigned randomly (hence *randomized* block design).

We would like to know if a **difference** in secret shopper ratings exists **among the cities**. Are they all about the same? Is one significantly higher than the other two? Are all three different from each other?

	col 1	col 2	col 3
Block-1	75	75	90
Block-2	70	70	70
Block-3	50	55	75
Block-4	65	60	85
Block-5	80	65	80
Block-6	65	65	65

ANSWER:

F (MSC/MSE)	5.526316
F (MSB/MSE)	3.157895

F(MSC/MSE) critical value 4.1, hence null hypothesis is rejcted