Practice (HW) 5
Due: March 07, 2019

- Asking questions to TAs and collaborating with classmates are encouraged, but copying, sharing, or distributing any material is strictly prohibited. Homework should be students' original work.
- Please submit
    1) SAS code (.SAS) with detailed comments
    2) PDF document with relevant output and interpretations
- Late homework will not be accepted.

Oral condition of cancer patients

Dataset 'Cancer.xlsx' contains a part of dataset for a study of oral condition of cancer patients conducted at the Mid-Michigan Medical Center. The oral conditions of the patients were measured and recorded at the initial stage and at the end of the second, fourth, and eighth weeks. Furthermore, age, initial weight and initial cancer stage of the patients were recorded.

Patients were divided into two groups at random: one group received a placebo and the other group received aloe juice treatment. The dataset consists of 25 patients with following variables:

Variable	Description
ID	Patient ID
TRT	Treatment group (0 = Placebo / 1 = Aloe juice)
AGE	Patient's age in years
WEIGHTIN	Patient's weight at the initial stage
STAGE	Initial cancer stage, coded 1 through 4
TOTALCW0	Oral condition at the initial stage
TOTALCW2	Oral condition at the end of week 2
TOTALCW4	Oral condition at the end of week 4
TOTALCW8	Oral condition at the end of week 8

a) Import the dataset, name it 'Cancer', and include labels and formats below in DATA step.

Variable	Label
TRT	Treatment
WEIGHTIN	Initial weight
TOTALCW0	Oral condition in week 0
TOTALCW2	Oral condition in week 2
TOTALCW4	Oral condition in week 4
TOTALCW8	Oral condition in week 8

Format	Description
TRTFMT	0 = Placebo
	1 = Aloe juice

Print the first 5 observations of dataset with the labels and formats above.

b) Transpose the dataset 'Cancer' from wide to long format.

i. PROC TRANSPOSE: Dataset 'Cancer_long1'	ii. ARRAY: Dataset 'Cancer_long2'

Obs	ID	Treatment	AGE	Initial weight	STAGE	TOTALC
1	1	Placebo	52	124	2	6
2	1	Placebo	52	124	2	6
3	1	Placebo	52	124	2	6
4	1	Placebo	52	124	2	7
5	2	Placebo	46	163.8	2	7
6	2	Placebo	46	163.8	2	16
7	2	Placebo	46	163.8	2	9
8	2	Placebo	46	163.8	2	10
9	5	Placebo	77	160	1	9
10	5	Placebo	77	160	1	6

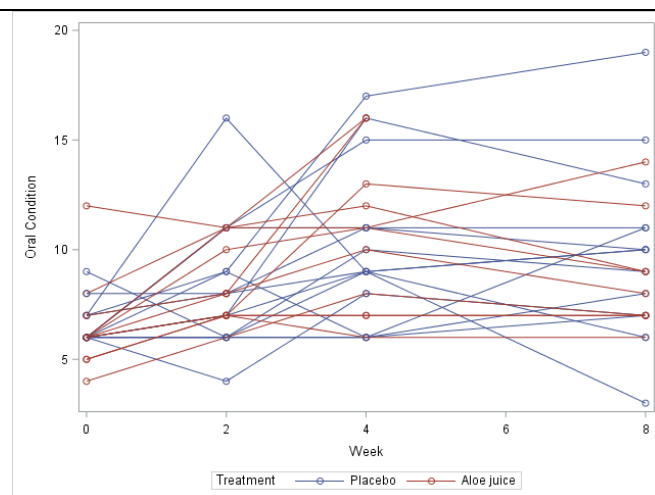| Obs	ID	week	Treatment	AGE	Initial weight	STAGE	TOTALC
1	1	0	Placebo	52	124	2	6
2	1	2	Placebo	52	124	2	6
3	1	4	Placebo	52	124	2	6
4	1	8	Placebo	52	124	2	7
5	2	0	Placebo	46	163.8	2	7
6	2	2	Placebo	46	163.8	2	16
7	2	4	Placebo	46	163.8	2	9
8	2	8	Placebo	46	163.8	2	10
9	5	0	Placebo	77	160	1	9
10	5	2	Placebo	77	160	1	6

c) Export the dataset 'Cancer_long2' into an excel file. Save it as 'Cancer2.xlsx'.

d) Use the dataset 'Cancer_long2' and RETAIN statement to calculate the grand total of TOTALC for each individual. Keep the last observation of each individual only. Namely, the first 10 observations of the new dataset are shown below:

Obs	ID	week	Treatment	AGE	Initial weight	STAGE	GrandTotal
1	1	8	Placebo	52	124	2	25
2	2	8	Placebo	46	163.8	2	42
3	5	8	Placebo	77	160	1	34
4	6	8	Placebo	60	136.5	4	52
5	9	8	Placebo	61	179.6	1	25
6	11	8	Placebo	59	175.8	2	42
7	12	8	Aloe juice	56	227.2	4	36
8	14	8	Aloe juice	42	162.6	1	25
9	15	8	Placebo	69	167.6	1	29
10	16	8	Aloe juice	44	261.4	2	42

e) Produce the following plots. Use Output Delivery System (ODS) to save them in a single RTF file.

i. SERIES	ii. PBSPLINE