Lab 11 *Some of these problems may be more challenging than others. Please feel free to work with others, attend office hours, or post on the course discussion forum if you need help. While collaboration with other students is encouraged, each student is responsible for submitting his or her own work. This assignment should be submitted in one well-commented SAS program. For any questions that require a written answer, do so in the SAS comments. Be sure to re-name the uploaded SAS scripts according to the naming convention* `LastnameFirstinitial_Lab#.sas` (*e.g.,* `PileggiS_Lab11.sas`).

Research on the harmful effects of lead poisoning began in the 1970's. Lead poisoning affects the development of the nervous system and therefore has a higher impact in children than adults. Exposure can be occupational or recreational through contaminated air, soil, water, or food. Effects of lead poisoning include learning disabilities and behavioral problems; high levels of exposure can lead to seizures, coma, and even death.

The data presented here are from one of the first quantitative research articles on the topic, published in the *The Lancet* in 1975 [1]. One-hundred and twenty-four children living near a lead-emitting smelter in El Paso, Texas were studied for two years. Various tests were used to assess neurological responses, including standing on one foot, tandem walking, alternate tapping, tapping with a stylus, visual reaction time, and auditory reaction time. Wechsler intelligence tests were also administered. The investigators compared results between children with high blood level concentration and matched 'control' children with low blood level concentration. The data set is `lead.sas7bdat`.

`Id`	subject identifier
`Area`	residence of the child in 1972: 1 = 0-1 miles from smelter 2 = 1-2.5 miles from smelter 3 = 2.5-4 miles from smelter
`Sex`	1=male, 2=female
`Iqv_inf`	information subtest in WISC and WPPSI
`Iqv_comp`	comprehension subtest in WISC and WPPSI
`Iqv_ar`	arithmetic subtest in WISC and WPPSI
`Iqv_ds`	digit span subtest (WISC) and sentence completion (WPPSI)
`Iqv_raw`	raw score/verbal IQ (this is total of previous 4 `iqv`)
`Iqp_pc`	picture completion subtest in WISC and WPPSI
`Iqp_bd`	block design subtest in WISC and WPPSI

[1] Philip J Landrigan, Robert W Baloh, William F Barthel, Randolph H Whitworth, NormanW Staehling, and Bernard F Rosenblum. Neuropsychological dysfunction in children with chronic low-level lead absorption. *The Lancet*, 305:708-712, 2012.

Iqp_oa	object assembly subtest (WISC), animal house subtest (WPPSI)
Iqp_cod	coding subtest (WISC), geometric design subtest (WPPSI)
Iqp_raw	raw score/performance IQ
HH_index	Hollingshead index of social status
Iqv	verbal IQ
Iqp	performance IQ
Iqf	full scale IQ
Iq_type	1=WISC, 2=WPPSI
Lead_type	1=blood lead <40 mg/100mL in both 1972 and 1973
	2=blood lead ≥40 mg/100mL in both 1972 and 1973
	3=blood lead ≥40 mg/100mL in 1972 and < 40 in 1973
Ld72	blood lead levels (micrograms/100mL) in 1972
Ld73	blood lead levels (micrograms/100mL) in 1973
Fst2yrs	1=child lived within 1 mile of smelter for 1^{st} two years, 2=child did not
Totyrs	total number of years spent within 4.1 miles of smelter
Pica	1=yes, 2=no
Colic	1=yes, 2=no
Clumsi	1=yes, 2=no
Irrit	1=yes, 2=no
Convul	1=yes, 2=no
X2Plat_r	# of taps for right hand in the 2-plate tapping test (# of taps in one 10 second trial)
X2Plat_l	# of taps for left hand in the 2-plate tapping test (# of taps in one 10 second trial)
Visrea_r	visual reaction time right hand (milliseconds)
Visrea_l	visual reaction time left hand (milliseconds)
Audrea_r	auditory reaction time right hand (milliseconds)
Audrea_l	auditory reaction time left hand (milliseconds)
FWT_r	finger-wrist tapping test right hand (# of taps in one 10 second trial)
FWT_l	finger-wrist tapping test left hand (# of taps in one 10 second trial)
Hyperact	Werry-Weiss-Peters scale for hyperactivitiy (as reported by parents),
	0=no activity... 4=severely hyperactive
Group	control = children whose blood-lead levels < 40 in 1972
	lead = children whose blood-lead levels ≥ 40 in 1972
age_years	age in years (4.5 = 4 years, 6 months)

1. Create a library reference called `mylib` to access the `lead` SAS data set.

2. For each of the following scenarios, implement the appropriate statistical method. In a comment in your SAS code, indicate (1) the method you chose, and (2) a brief interpretation of your results. All tests may be done at the $\alpha = 0.05$ level of significance.

   (a) Researchers would like to know if population average lead blood concentration changed between 1972 (`Ld72`) and 1973 (`Ld73`).

   (b) The CDC recommends that "safe" limits for lead blood levels to be no more than 5 micrograms per 100 mL. Did the population average in 1973 (`Ld73`) exceed

this safe limit?

(c) Describe the strength of the association between verbal IQ (`Iqv`) and performance IQ (`Iqp`).

(d) Does population average left hand finger wrist tapping ability (`FWT_l`) differ by whether or not children were in the lead or control group (`Group`)?

3. As you can see, there is a lot of data to analyze!

(a) Create macro that can perform the two-sample t-test with the following four parameters:

- `dsn` - data set name
- `quantvar` - the quantitative variable to analyze
- `catvar` - the categorical grouping variable
- `siglev` - the level of significance
- `plotoption` - this accompanies the `PLOTS` option and should take one of two values :

  `all` - all two sample t-test plots produced

  `none` - no plots produced

(b) Your SAS output should have an informative title that displays information about the macro variables.

(c) Execute your macro for the following two scenarios to test it out:

- At the 0.01 significance level, does population average full scale IQ (`Iqf`) differ by whether or not the children had pica (`Pica`)? Do not produce any plots.
- At the 0.05 significance level, does population average visual reaction time in the left hand (`Visrea_l`) differ by gender (`Sex`)? Do produce plots.

(d) Investigate the effects of lead poisoning on your own by choosing three quantitative variables of interest to you. Execute your macro on these three variables with the categorical grouping variable of `Group` to see if the lead and control groups have different characteristics. In a comment in your SAS code, note which variable(s) has the strongest evidence of an association with `Group`.

(e) Return to the three executions of your macro in question 3d. Export the output from these three macros into a single PDF file using the Output Delivery System. Upload the pdf file to the Lab 11 assignment, in addition to your SAS code.