Write programs that will ***exactly*** reproduce the output shown ***except*** for variable names. You may read the raw data files from g:\shared\bio113 or you may copy them to your p drive and read them from there. Choose your own variable names (***not*** v1, v2, etc.). Use infile options if necessary. The line (123456789 12345…) above the data is a rule or number line to get you oriented; it's ***not*** data. Display your SAS data sets with PROC PRINT (without a VAR statement). **Turn in your programs, SAS logs and SAS output.**

1. Write a program using ***list input*** to read first name, date of birth, sex, visit date, height (inches), weight (pounds), systolic blood pressure, and diastolic blood pressure from the raw data set g:\shared\bio113\hw1_1.dat. Save the data in a permanent SAS data set on your p drive with the name **visit**. Use **physical** as the *fileref* and **exam** as the *libref*.

```
123456789 123456789 123456789 123456789 1234567

Anne  5/19/1956  f  11/15/2003 63 128 119 78
Madeleine 11/22/1982  f 3/22/2004 67 141 132
Richard 7/11/1958  m 2/23/2004  . 177 120 75
Stephan  4/15/1961  m 5/2/2004 71 156 113 77
```

Obs	v1	v2	v3	v4	v5	v6	v7	v8
1	Anne	-1322	16024	f	63	128	119	78
2	Madeleine	8361	16152	f	67	141	132	.
3	Richard	-539	16124	m	.	177	120	75
4	Stephan	470	16193	m	71	156	113	77

2. Write a program that reads the raw data set g:\shared\bio113\hw1_2.dat and writes a permanent SAS data set named **birds** *with 7 observations and 4 variables* to your p drive. Notice that the raw dataset has two records for each type of bird while the SAS data set should have only one. Use **column or formatted input** to read the variables—kind of bird and total number observed on the first record and number identified as male and number identified as femail on the second. Please use **annual** for the *fileref* and **count** for the *libref*.

```
123456789 123456789 123

cardinal    174
 95 35
chickadee   382
178105
goldfinch   132
 70 62
pigeon      15244202
892978
purple finch446
326120
nuthatch    267
 89 77
woodpecker   65
 39 26
```

Obs	v1	v2	v3	v4
1	cardinal	174	95	35
2	chickadee	382	178	105
3	goldfinch	132	70	62
4	pigeon	15244202	892	978
5	purple finch	446	326	120
6	nuthatch	267	89	77

```
7     woodpecker          65      39      26
```

3. Write a program to input length and width when the object is a rectangle and smallest radius when the object is an oval. ***Choose the input style***. Save the data in a permanent SAS data set on your p drive with the name **shapes**. Use **shapes** for the *libref* as well. Use an infile statement only to access the raw data in the external ASCII file, g:\shared\bio113\hw1_3.dat.

```
123456789 123456789 12345

oval      374
rectangle 18    2
oval        57.42
oval      18.3
rectangle  24    16
```

Obs	v1	v2	v3	v4
1	oval	374.00	.	.
2	rectangle	.	18	2
3	oval	57.42	.	.
4	rectangle	.	24	16

4. Write a program to input the intraventricular hemorrhage data described on the next page. Use whatever input style or styles you like (see how little typing you can do!) but please use the variable names in the column headed 'Variable name.' The raw data are in g:\shared\bio113\ivh.dat. Save the SAS data set on your p drive with the name **ivh**. Print the *first 10 observations* of this data set and study the output to be certain that the variables look like what's described under 'Variable descriptions and codes' and 'Missing codes.' The SAS statement that will print 10 observations is: *proc print data=**dsn** (obs=10);* [where ***dsn*** is the name you gave the **permanent** SAS data set]. We will use these data in later homework exercises.

Data dictionary for g:\shared\bio113\ivh.dat

This is REAL data. It is a subset of babies and variables from a large dataset collected on infants born between 1991 and 1993 at hospitals in Boston and New York City. The babies weighed 500 to 1500 grams at birth and each was assessed for brain abnormalities with cranial ultrasound. Additional data were collected by maternal interview and review of maternal and infant medical records.

Variable name	Column location	Variable description and codes	Missing code
id	1-6	subject ID	999999
hosp	7	birth hospital : 1=St. Elsewhere, 2=Chicago Hope	9
sex	8	baby' sex: 0=female, 1=male	9
race	9	baby's race: 1=white, 2=black, 3=hispanic, 4=other, 5=mixed	9
ga	10-11	gestational age (completed weeks)	99
bw	12-15	birth weight (grams)	9999
ivh	16	intraventricular hemorrhage: 0=no, 1=yes	9
medu	17	mother's education (yrs): 1 is <12, 2 is 12-15, 3 is 16+	9
single	18	single mother: 0=no, 1=yes	9
cs	19	cesarian delivery: 0=no, 1=yes	9
pih	20	pregnancy induced hypertension: 0=no, 1=yes	9
labor	21-26	duration of labor (hrs)	9999.9
rom	27-32	duration of membrane rupture (hrs)	9999.9
acs	33	antenatal corticosteriods: 1=none, 2=partial, 3=complete	9
mage	34-35	mother's age (yrs)	99
apg1	36-37	apgar score, 1 minute	99
apg5	38-39	apgar score, 5 minutes	99
vent	40	mode of ventilation: 1=none, 2=CPAP, 3=mechanical	9
los	41-43	length of stay (days)	999
dead	44	baby died: 0=no, 1=yes	9
wt1-wt4	45-60, 4.*	weight (grams) on days 1, 2, 3, and 4	9999
map1-map4	61-68, 2.	mean blood pressure (mmHg) on days 1, 2, 3, and 4	99
pco2_1-pco2_4	69-76, 2.	lowest pCO_2 on days 1, 2, 3, and 4	99
pda1-pda4	77-80, 1.	PDA on days 1, 2, 3, and 4: 0=no, 1=yes	9
dopa1-dopa4	81-84, 1.	dopamine on days 1, 2, 3, and 4: 0=no, 1=yes	9
fluid1-fluid4	85-96, 3.	total fluids (cc) on days 1, 2, 3, and 4	999
cry1-cry4	97-108, 3.	total crystalloid (cc) on days 1, 2, 3, and 4	999
col1-col4	109-116, 2.	total colloids (cc) on days 1, 2, 3, and 4	99
ptx1-ptx4	117-120, 1.	pneumothorax on days 1, 2, 3, and 4: 0=no, 1=yes	9
t4	121-125	Total thyroxine (t4)	99.99
t4age	126-127	Postnatal age when T4 measure taken	99

* This notation indicates that the variables wt1, wt2, wt3, and wt4 are in order beginning in column 35 and run to column 50 and each variable occupies 4 columns (i.e., wt1 in cols 35-38, wt2 in cols 39-42, wt3 in cols 43-46, wt4 in cols 47-50).