1. Four payroll data files are available.

```
Payroll1.dat contains:
ID  Dept   Date    Amount
001 Math 10/10/2005 $100
002 Math 10/10/2005 $200

Payroll2.dat contains:
ID  Dept   Date    Amount
003 CS 10/10/2005 $1,000
004 CS 10/10/2005 $2,000

Payroll3.dat contains:
ID    Dept      Date    Amount
005 English 10/10/2005 $1,000
006 English 10/10/2005 $2,000

Payroll4.dat contains:
ID  Dept        Date     Amount
007 Astronamy 10/10/2005 $1,000
008 Astronamy 10/10/2005 $2,000
```

Create a macro with the number of files as the parameter, which will read in every data and run a print procedure for every data. Please include the department information in the title of each printout. For example, name the macro printall. A call "%printall(numfiles=4)" will produce the following output:

Payroll for the Math department

Obs	ID	Dept	Date	Amount
1	001	Math	October 10, 2005	100
2	002	Math	October 10, 2005	200

Payroll for the CS department

Obs	ID	Dept	Date	Amount
1	003	CS	October 10, 2005	1000
2	004	CS	October 10, 2005	2000

Payroll for the English department

Obs	ID	Dept	Date	Amount
1	005	English	October 10, 2005	1000
2	006	English	October 10, 2005	2000

Payroll for the Astronamy department

Obs	ID	Dept	Date	Amount
1	007	Astronamy	October 10, 2005	1000
2	008	Astronamy	October 10, 2005	2000

2. The exponential distribution is a popular distribution for modeling liftetimes or waiting times. The exponential with a mean of λ has a density function $f(x) = (1/\lambda)e^{-x/\lambda}$ (for $x \geq 0$) and a C.D.F. $F(x) = P(X \leq x) = 1 - e^{-x/\lambda}$.

   (a) Write a macro which has its argument the mean λ which calculates and plots (using a high resolution plot) the density and the C.D.F. These should be two separate plots. For the density, the values of x should range from 0 up to a point where $f(x) \leq .001$. For the CDF, the values of x should range from 0 up to a point where $F(X) \geq .999$. Execute the macro and obtain the plots using $\lambda = 10$. (For the exponential the CDF is obtained using cdf('exponential',x,lambda) and similarly for the pdf.)

   (b) A study was conducted which yielded estimated mean survival times for five types of cancer as below (see Cancer Survival on the web page if interested in details). Using an an exponential model, which is reasonable here, find the estimated probability of surviving a year or more, and two years or more, for each of the five types of cancer. Put the cancer names and means into arrays and make use of a do loop to do the calculations. Each line of your resulting data set should list cancer type, mean survival, prob1 and prob2 where probj is probability of surviving j years or more.

```
Type      Mean survival(in days)
Breast    1395.91
Bronchus   211.59
Colon     457.41
Ovary     884.33
Stomach   286.00
```

3. Write a macro that, from the TropicalSales data, prints the records of a selected customer in the order of a specified variable. For example, if we call this macro "select", the following call

   %select(customer=356W, sortvar=Quantity);

   prints the records for 356W sorted according to the variable quantity.

4. Write a macro that calculates the n-moving average of a variable, i.e., the following call

   %movingaverage(data=a, var=x, n=3);

   generates a SAS data set containing the 3-moving average of the variable x in dataset a. The definition of the n-moving average can be found here

   http://mathworld.wolfram.com/MovingAverage.html

   Test your macro using

```
data test;
do i=1 to 10;
output;
end;
%movingaverage(data=test, var=i, n=3)
```