```
*SAS COncate Datasets;
/*
Multiple SAS data sets can be concatenated to give a single data set using
the SET statement. The total number of observations in the concatenated
data set is the sum of the number of observations in the original data sets.
The order of observations is sequential.
*/

/*Syntax
SET data-set 1 data-set 2 data-set 3.....;
*/
DATA ITDEPT;
   INPUT empid name $ salary  ;
DATALINES;
1 Rick 623.3
3 Mike 611.5
6 Tusar 578.6
;
RUN;
DATA NON_ITDEPT;
   INPUT empid name $ salary  ;
DATALINES;
2 Dan 515.2
4 Ryan 729.1
5 Gary 843.25
7 Pranab 632.8
8 Rasmi 722.5
RUN;
DATA All_Dept;
   SET ITDEPT NON_ITDEPT;
RUN;
PROC PRINT DATA = All_Dept;
RUN;

*Different number of variables;
/*
If one of the original data set has more number of variables then another,
then the data sets still get combined but in the smaller data set
those variables appear as missing.
*/
DATA ITDEPT;
   INPUT empid name $ salary DOJ date9.  ;
DATALINES;
1 Rick 623.3 02APR2001
3 Mike 611.5 21OCT2000
6 Tusar 578.6 01MAR2009
;
RUN;
DATA NON_ITDEPT;
   INPUT empid name $ salary  ;
DATALINES;
2 Dan 515.2
4 Ryan 729.1
5 Gary 843.25
7 Pranab 632.8
8 Rasmi 722.5
RUN;
DATA All_Dept;
   SET ITDEPT NON_ITDEPT;
RUN;
PROC PRINT DATA = All_Dept;
RUN;

*Different variable name;
/*
In this scenario the data sets have same number of variables but a variable
name differs between them. In that case a normal concatenation will produce all
the variables in the result set and giving missing results for the two variables
```

```
which differ. While we may not change the variable name in the original data sets
we can apply the RENAME function in the concatenated data set we create.
*/
DATA ITDEPT;
   INPUT empid ename $ salary  ;
DATALINES;
1 Rick 623.3
3 Mike 611.5
6 Tusar 578.6
;
RUN;
DATA NON_ITDEPT;
   INPUT empid empname $ salary  ;
DATALINES;
2 Dan 515.2
4 Ryan 729.1
5 Gary 843.25
7 Pranab 632.8
8 Rasmi 722.5
RUN;
DATA All_Dept;
   SET ITDEPT(RENAME =(ename = Employee) ) NON_ITDEPT(RENAME =(empname = Employee) );
RUN;
PROC PRINT DATA = All_Dept;
RUN;

*Different Variable Lengths;

/*
If the variable lengths in the two data sets is different than the concatenated
data set will have values in which some data is truncated for the variable
with smaller length. It happens if the first data set has a smaller length.
To solve this we apply the higher length to both the data
*/
DATA ITDEPT;
   INPUT  empid 1-2 ename $ 3-7 salary 8-14  ;
DATALINES;
1 Rick  623.3
3 Mike  611.5
6 Tusar 578.6
;
RUN;
DATA NON_ITDEPT;
   INPUT  empid 1-2 ename $ 3-9  salary 10-16 ;
DATALINES;
2 Dan    515.2
4 Ryan   729.1
5 Gary   843.25
7 Pranab 632.8
8 Rasmi  722.5
RUN;
DATA All_Dept;
   LENGTH ename $ 7   ;
   SET ITDEPT  NON_ITDEPT ;
RUN;
PROC PRINT DATA = All_Dept;
RUN;
```