# Assignment 1

## Problem 1:

1. **Read AAUP data**

SAS Code:

DATA ACADEME;

INFILE 'C:\Users\mummidra\Dropbox\UC\FirstSem\Stat\_Computing\Assignment\_1\AAUP\_data.txt' LRECL=32767 delimiter=',';

length FICE $4;

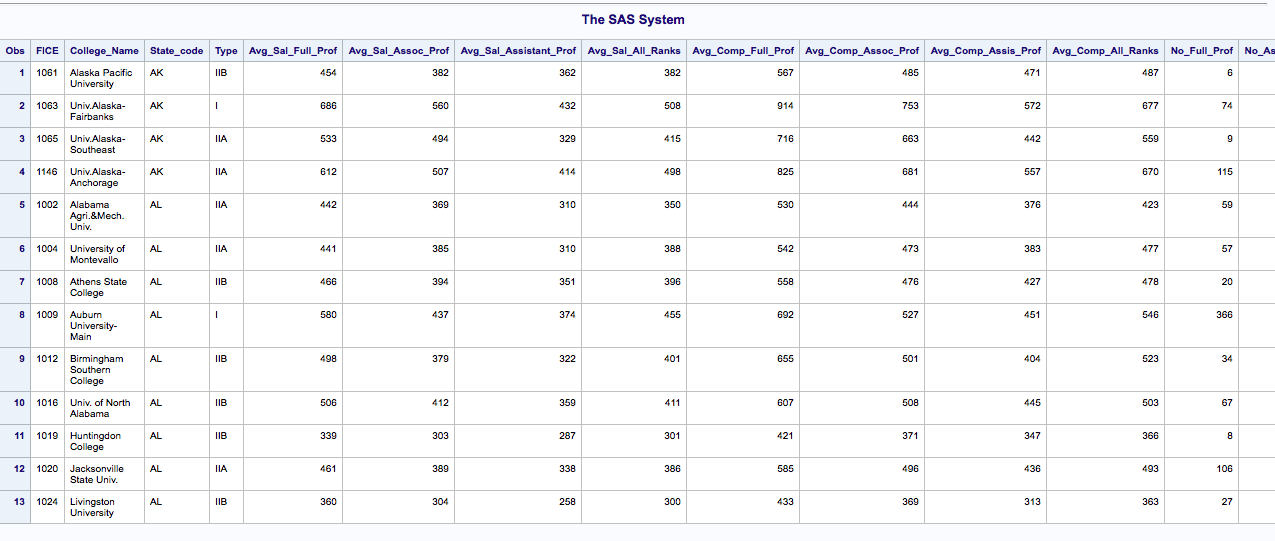
length College\_Name $40;

INPUT FICE College\_Name$ State\_code$ Type$ Avg\_Sal\_Full\_Prof Avg\_Sal\_Assoc\_Prof Avg\_Sal\_Assistant\_Prof Avg\_Sal\_All\_Ranks Avg\_Comp\_Full\_Prof Avg\_Comp\_Assoc\_Prof Avg\_Comp\_Assis\_Prof Avg\_Comp\_All\_Ranks No\_Full\_Prof No\_Assoc\_Prof No\_Assis\_Professors No\_Instructors No\_Faculty\_All\_Ranks;

RUN;

PROC PRINT DATA=ACADEME;

RUN;

SAS Output:

b) **Creating dataset with all average salaries and average compensations.**

SAS Code:

DATA ALLAVERAGES;

SET ACADEME;

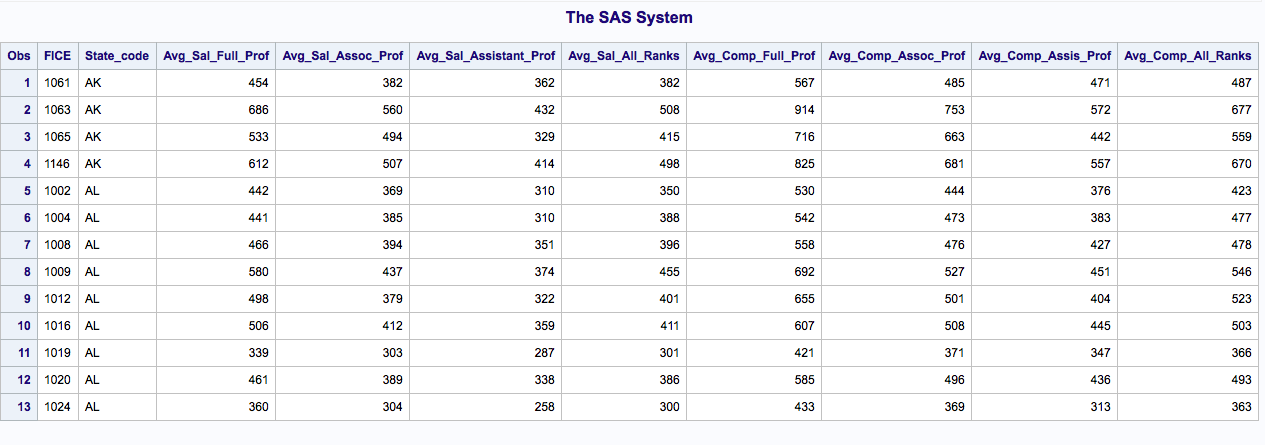
DROP College\_Name Type No\_Full\_Prof No\_Assoc\_Prof No\_Assis\_Professors No\_Instructors No\_Faculty\_All\_Ranks;

RUN;

PROC PRINT DATA=ALLAVERAGES;

RUN;

SAS Output:



c) **create two subsets whose colleges are from AK or AL separately**

SAS Code:

DATA ALDDATASET;

SET ALLAVERAGES;

IF State\_code='AL';

RUN;

DATA AKDDATASET;

SET ALLAVERAGES;

IF State\_code='AK';

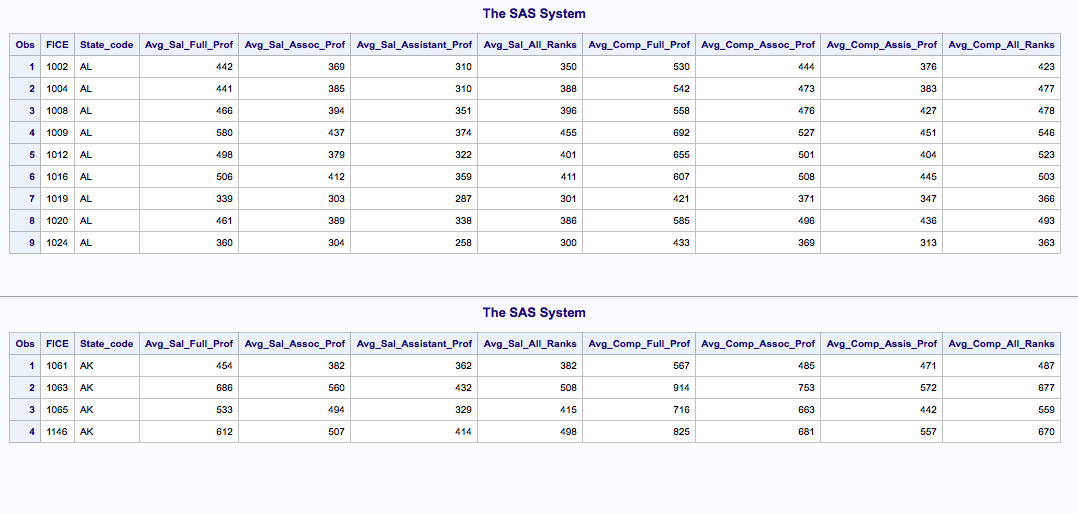
RUN;

PROC PRINT DATA=ALDDATASET;

PROC PRINT DATA=AKDDATASET;

RUN;

SAS Output:



**d) total income dataset**

SAS Code:

DATA ALTOTALDATASET;

SET ALDDATASET;

TOT\_Sal\_Full\_Prof=Avg\_Sal\_Full\_Prof+Avg\_Comp\_Full\_Prof;

TOT\_Sal\_Assoc\_Prof=Avg\_Sal\_Assoc\_Prof+Avg\_Comp\_Assoc\_Prof;

TOT\_Sal\_Assis\_Prof=Avg\_Sal\_Assistant\_Prof+Avg\_Comp\_Assis\_Prof;

TOT\_Sal\_All\_Ranks=Avg\_Sal\_All\_Ranks+Avg\_Comp\_All\_Ranks;

DROP Avg\_Sal\_Full\_Prof Avg\_Sal\_Assoc\_Prof Avg\_Sal\_Assistant\_Prof Avg\_Sal\_All\_Ranks Avg\_Comp\_Full\_Prof Avg\_Comp\_Assoc\_Prof Avg\_Comp\_Assis\_Prof Avg\_Comp\_All\_Ranks;

RUN;

DATA AKTOTALDATASET;

SET AKDDATASET;

TOT\_Sal\_Full\_Prof=Avg\_Sal\_Full\_Prof+Avg\_Comp\_Full\_Prof;

TOT\_Sal\_Assoc\_Prof=Avg\_Sal\_Assoc\_Prof+Avg\_Comp\_Assoc\_Prof;

TOT\_Sal\_Assis\_Prof=Avg\_Sal\_Assistant\_Prof+Avg\_Comp\_Assis\_Prof;

TOT\_Sal\_All\_Ranks=Avg\_Sal\_All\_Ranks+Avg\_Comp\_All\_Ranks;

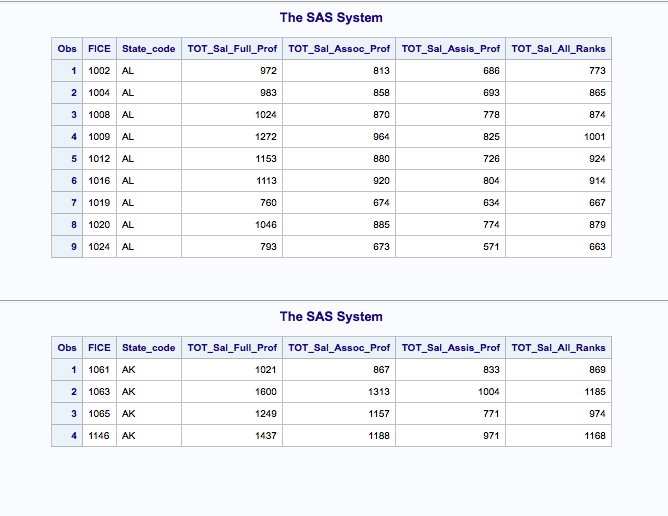
DROP Avg\_Sal\_Full\_Prof Avg\_Sal\_Assoc\_Prof Avg\_Sal\_Assistant\_Prof Avg\_Sal\_All\_Ranks Avg\_Comp\_Full\_Prof Avg\_Comp\_Assoc\_Prof Avg\_Comp\_Assis\_Prof Avg\_Comp\_All\_Ranks;

RUN;

PROC PRINT DATA=ALTOTALDATASET;

PROC PRINT DATA=AKTOTALDATASET;

RUN;



Problem 2:

1.Code after correction:

SAS Code:

DATA SCORE\_LIST;

INPUT SCHOOL$ MATH\_BEFORE MATH\_AFTER@@;

MATH\_IMPROVE= MATH\_AFTER- MATH\_BEFORE;

CARDS;

UC 80 81 UC 80 80 UC 90 87 OSU 70 80

OSU 78 80

;

RUN;

DATA LIST\_NEW;

KEEP SCHOOL MATH\_IMPROVE;

RUN;

PROC SORT DATA=LIST\_NEW OUT=LIST2;

BY MATH\_IMPROVE ;

DATA LIST2;

SET LIST2;

IF MATH\_IMPROVE>0 THEN GOOD=’YES’;

ELSE GOOD=’NO’;

RUN;

PROC PRINT DATA=LIST2;

RUN;

2 Running corrected code in SAS

SAS Code:

DATA SCORE\_LIST;

INPUT SCHOOL$ MATH\_BEFORE MATH\_AFTER@@;

MATH\_IMPROVE=(MATH\_AFTER-MATH\_BEFORE);

CARDS;

UC 80 81 UC 80 80 UC 90 87 OSU 70 80

OSU 78 80

;

RUN;

DATA LIST1\_NEW;

SET SCORE\_LIST;

KEEP SCHOOL MATH\_IMPROVE;

RUN;

PROC SORT DATA=LIST1\_NEW OUT=LIST2;

BY MATH\_IMPROVE;

DATA LIST2;

SET LIST2;

IF MATH\_IMPROVE>0 THEN GOOD='YES';

ELSE GOOD='NO';

RUN;

PROC PRINT DATA=LIST2;

RUN;

SAS Output:

