Brad Miller

Math 510

Homework 7, SAS 1

Chapter 3.2 Level 1

a.

1. How many observations are there in the orion.country data set? 7
2. How many variables are there in the orion.country data set? 6
3. What is the name of the last country in the data set? South Africa

b.

Submit a PROC CONTENTS step to generate a list of all members in the orion library. What is the name of the last member listed? US\_SUPPLIERS

Chapter 3.2 Level 2

1. Examine the general data set properties or orion.staff

proc contents data=orion.staff

/\*this line displays the descriptor portion of the orion.staff data set\*/;

run;

b. What sort information is stored for this data set? The General information section indicates that the orion.staff data set is not sorted.

Chapter 3.2 Challenge

1. What is the name of the file? Autoexec.sas
2. What is its purpose? The autoexec file contains SAS statements that are executed automatically when you invoke SAS or when you start another SAS process.
3. How is it created? The file is created with any text editor but the recommended method is to use the SAS text editor.
4. How could this be useful in a SAS session? An autoexec file can contain LIBNAME statements for SAS libraries that you access routinely in SAS sessions.

Chapter 4.1 Level 1

proc print data=orion.order\_fact noobs

/\* this line prints the orion.order\_fact data set and removes the obs column\*/;

SUM total\_retail\_price

/\*this line displays the sum of the total\_retail\_price column\*/;

where total\_retail\_price > 500

/\*this line selects only the observations where the total\_retail\_price is greater than 500\*/;

ID customer\_id

/\* this line sets the customer\_id column to be the identifying variable\*/;

var order\_id order\_type quantity total\_retail\_price

/\* this line displays the variables order\_id order\_type quantity and total\_retail\_price\*/;

run;

c. What did you notice about the Obs column? The original observation numbers are displayed. Therefore, there are gaps in the observation numbers.

Did the sum of the Total\_Retail\_Price change to reflect only the subset? Yes

d. If the Obs column is suppressed, how can you verify the number of observations in the results? Check the log.

e. When the ID statement was added, how did the output change? Customer\_ID is the leftmost colum and is displayed on each line for an observation.

f. When the VAR statement is added, what do you notice about Customer\_ID? There are two Customer\_ID columns. The first Customer\_ID column is the ID column and the second Customer\_ID column is included because Customer\_ID is in the VAR statement.

Chapter 4.1 Level 2

proc print data=orion.customer\_dim noobs

/\*this line prints the orion.customer\_dim data set and removes the obs column\*/;

where Customer\_Age between 30 and 40

/\*this line selects only the customers between the ages of 30 and 40\*/;

id Customer\_ID

/\*this line sets the Customer\_ID as the identifying column\*/;

var Customer\_Name Customer\_Age Customer\_Type

/\*this line sets the output to displaythe Customer\_Name Customer\_Age and Customer\_Type variables\*/;

run;

Chapter 4.2 Level 1

5.

proc sort data=orion.employee\_payroll out=work.sort\_salary;

/\*this line reads the observations from the orion.employee\_payroll data set and places them into the work.sort\_salary data set\*/

by Salary;

/\*this line sorts the data set by Salary\*/

run;

proc print data=work.sort\_salary; /\*this line prints the work.sort\_salary data set\*/

run;

6.

proc sort data=orion.employee\_payroll out=work.sort\_salary2

/\*this line reads the observations from the orion.employee\_payroll data set and places them into the work.salary\_sort2 data set\*/;

by Employee\_Gender descending Salary

/\*this line sorts by gender and within gender, by salary in descending order\*/;

run;

proc print data=work.sort\_salary2

/\*this line prints the sort\_salary2 data set\*/;

by Employee\_Gender

/\*this line groups the sort\_salary2 data set by Gender\*/;

run;

Chapter 4.2 Level 2

proc sort data=orion.employee\_payroll out=work.sort\_sal

/\*this line reads the observations from orion.employee\_payroll and places them into work.sort\_sal\*/;

by Employee\_Gender descending Salary

/\*this line sorts by Gender and by descending Salary within Gender\*/;

run;

proc print data=work.sort\_sal noobs

/\*this line prints the work.sort\_sal data set and removes the obs column\*/;

by Employee\_Gender

/\*this line groups by Gender\*/;

sum Salary

/\*this line adds the total and subtotals for Salary\*/;

Where Employee\_Term\_Date is missing and Salary>65000

/\*this line selects only the employees who do not have a term data and whose salary is greater than 65000\*/;

var Employee\_ID Salary Marital\_Status

/\*this line sets the data set to display only the Employee\_ID Salary and Marital\_Status columns\*/;

run;

Chapter 4.3 Level 1

9.

title1 'Austrailian Sales Employees'/\*this line creates the first title\*/;

title2 'Senior Sales Representatives'/\*this line creates the second title\*/;

footnote 'Job Title: Sales Rep. IV'/\*this line creates the footnote\*/;

proc print data=orion.sales noobs

/\*this line prints the orion.sales data set and removes the obs column\*/;

where Country='AU' and Job\_Title contains 'Rep. IV'/

\*this line selects only Austrailian Sales Rep. IV\*/;

var employee\_id first\_name last\_name gender salary

/\*this line sets the data set to display only the employee\_id first\_name last\_name gender and salary columns\*/;

run;

title;

footnote;

10.

title 'Entry-level Sales Representatives'/\*this line creates the title\*/;

footnote 'Job\_Title: Sales Rep. I'/\*this line creates the footnote\*/;

proc print data=orion.sales noobs split='\*'

/\* this line prints the orion.sakes data set, removes the obs column, and sets the split character to '\*' \*/;

where Country='US' and Job\_Title='Sales Rep. I'

/\*this line selects only observations where the Country is US and the Job title is Sales Rep. I \*/;

var Employee\_ID First\_Name Last\_Name Gender Salary

/\*this line sets the data set to display only the variables First\_Name Last\_Name Gender and Salary\*/;

label Employee\_ID="Employee\*ID"

First\_Name="First\*Name"

Last\_Name="Last\*Name"

Salary="Annual\*Salary"/\*this line creates the labels for each column\*/;

run;

title;

footnote;

Chapter 4.3 Level 2

proc sort data=orion.employee\_addresses out=work.employee\_addresses

/\*this line reads the observations from orion.employee\_addresses and places them into a new data set called work.employee\_addresses\*/;

where Country='US'

/\*this line selects only observations where Country is US\*/;

by State City Employee\_Name

/\*this line sorts the output data st by State, City, and Employee\_Name\*/;

run;

title "US Employees by State";/\* this line creates the title\*/

proc print data=work.employee\_addresses noobs split='\*';

/\*this line prints the work.employee\_addresses data set, removes the obs column, and sets the split character to '\*' \*/

var Employee\_ID Employee\_Name City Postal\_Code;

/\*This line sets the data set to only display the Employee\_ID Employee\_Name City and Postal\_Code columns \*/

label Employee\_ID='Employee\*ID'

Employee\_Name='Name'

Postal\_Code='Zip Code'/\*this line creates the labels\*/;

by State/\*this line groups the data set by State\*/;

run;

Chapter 5.1 Level 1

proc print data=orion.employee\_payroll /\*this line prints the orion.employee\_payroll data set\*/;

var employee\_ID salary birth\_Date Employee\_Hire\_date/\* this line sets the data set to display only the employee\_ID salary Birth\_Date and Employee\_Hire\_date columns\*/;

format Salary dollar11.2 Birth\_date mmddyy10. Employee\_Hire\_date date9.

/\*this line sets the format for Salary Birth\_Date and Employee\_Hire\_Date\*/;

run;

Chapter 5.1 Level 2

title1 'US Sales Employees'/\*this line creates the title1\*/;

title2 'Earning Under $26,000'/\*this line creates the title2\*/;

proc print data=orion.sales split='\*' noobs

/\*this line prints the orion.sales data set, sets the split character to \*, and removes the obs column \*/;

where Country='US' and Salary<26000

/\*this line selects only the observations where Country is US and Salary is less than 26000\*/;

var Employee\_ID First\_Name Last\_Name Job\_Title Salary Hire\_Date

/\*this line sets the data set to include only the Employee\_ID First\_Name Last\_Name Job\_Title Salary and Hire\_Date columns\*/;

label First\_Name='First\*Name'

Last\_Name='Last Name'

Job\_Title='Title'

Hire\_Date='Date\*Hired' /\*this line creates the labels\*/;

format Salary dollar10. Hire\_Date monyy7./\*this line sets the format for Salary and Hire\_Date\*/;

run;

title;

Chapter 5.2 Level 1

data Q1Birthdays/\*This line creates a new data set called Q1Birthdays\*/;

set orion.employee\_payroll /\*This line reads the observations from the orion.employee\_payroll data set\*/;

BirthMonth=month(Birth\_Date) /\*This line creates the BirthMonth column\*/;

if BirthMonth le 3 /\*this line only adds observations where the BirthMonth is less than or equal to 3 to the data set\*/ ;

run;

proc format;

value $gender /\*this line creates the gender format\*/

'F'='Female'

'M'='Male';

value mname /\*this line creates the mname format\*/

1='January'

2='February'

3='March';

run;

title 'Employees with Birthdays in Q1'/\*this line creates the title\*/;

proc print data=Q1Birthdays/\*this line prints the data set Q1Birthdays\*/;

var Employee\_ID Employee\_Gender BirthMonth /\*this line sets the data set to only display the columns Employee\_ID Employee\_Gender and BirthMonth\*/ ;

format Employee\_Gender $gender.

BirthMonth mname. /\*this line sets the format for Gender and BirthMonth\*/;

run;

title;

Chapter 5.2 Level 2

proc format;

value $gender /\*this line creates the gender format \*/

'F'='Female'

'M'='Male'

other='invalid code';

value salrange /\*this line creates the salrange format\*/

.='Missing Salary'

20000-<100000='Below $100,000'

100000-500000='$100,000 or more'

other='Invalid Salary';

run;

title1 'Salary and Gender Values'; /\*this line creates the title1\*/

title2 'for Non-Sales Employees'; /\*this line creates the title2\*/

proc print data=orion.nonsales;

/\*this line prints the orion.nonsales data set \*/

var Employee\_ID Job\_Title Salary Gender;

/\* this line sets the data set to display only the Employee\_ID Job\_Title Salary and Gender columns\*/

format Salary salrange. Gender $gender.; /\* this line sets the format for salary and gender \*/

run;

title;

Chapter 6.2 Level 2

data work.delays; /\*this line creates the output data set work.delays\*/

set orion.orders;/\*this line sets orion.orders to be the input dataset\*/

Order\_month=month(order\_date);/\*this line creates the Order\_Month variable\*/

where order\_date+4<delivery\_date

and Employee\_id=99999999;

\*this line selects observations where the delivery date are more than four days beyond the order date and the employee\_ID is 99999999\*/

if order\_month=8; /\*this line selects only observations with order\_moth=8\*/

label Order\_date='Date Ordered'

delivery\_Date='Date Delivered'

order\_month='Month Ordered'; /\*this line creates the labels\*/

format order\_date delivery\_date mmddyy10.; /\*this line sets the format for order\_date and delivery\_date\*/

keep employee\_id customer\_id Order\_date Delivery\_date

order\_month; /\*this line sets the data set to keep the employee\_id customer\_id Order\_date Delivery\_date

and order\_month variables\*/

run;

proc contents data=work.delays; /\*this line displays the descriptor portion of the work.delays data set \*/

run;

proc print data=work.delays; /\*this line prints the work.delays data set\*/

run;

Chapter 9.1 Level 2

data work.birthday; /\*this line creates a new data set called work.birthday\*/

set orion.customer; /\*this line reads the orion.cusotmer data set\*/

Bday2012=mdy(Month(Birth\_Date),day(Birth\_Date),2012);

/\*this line creates the Bday2012 variables\*/

BdayDOW2012=weekday(Bday2012); /\*this line creates the BdayDOW2012 variable\*/

Age2012=(Bday2012-Birth\_Date)/365.25; /\*this line creates the Age2012 variables\*/

keep Customer\_Name Birth\_Date Bday2012 BdayDOW2012 Age2012;

/\*this line sets the data set to display the Customer\_Name Birth\_Date Bday2012 BdayDOW2012 and Age2012 variables \*/

format Bday2012 date9. Age2012 3.; /\*this line sets the format for Bday2012 and Age 2012 \*/

run;

proc print data=work.birthday; /\*this line prints the work.birthday data set\*/

run;

Chapter 9.2 Level 2

6.

data work.season; /\*this line creates a new data set work.season\*/

set orion.customer\_dim; /\*this line reads the data from the orion.customer\_dim data set\*/

length Promo $ 6 Promo2 $ 6;

/\*this line sets the length of the Promo and Promo2 columns to 6\*/

Quarter=qtr(Customer\_BirthDate); /\*this line creates the Quarter variable \*/

if Quarter=1 then Promo='Winter';

else if Quarter=2 then Promo='Spring';

else if Quarter=3 then Promo='Summer';

else if Quarter=4 then Promo='Fall';

/\*this line assigns the Promo values based on the quarter values\*/

if customer\_age>=18 and customer\_age<=25 then Promo2='YA';

else if customer\_age>=65 then Promo2='Senior';

/\*this line assignes the Promo2 values based on the customer\_age values\*/

keep Customer\_firstname customer\_lastname customer\_birthdate customer\_age promo promo2;

/\*this line sets the data set to include the Customer\_firstname Customer\_lastname c ustomer\_birthdate customer\_age promo and promo2 variables\*/

run;

proc print data=work.season; /\*this line prints the work.season data set\*/

run;

7.

data work.ordertype; /\*this line creates a new data set work.ordertype\*/

set orion.orders; /\*this line reads the data set orion.orders\*/

Length Type $ 13 SaleAds $ 5; /\* this line sets the length for the Type and SaleAds variables\*/

DayOfWeek=Weekday(Order\_Date); /\*this line creates the DayOfWeek variable\*/

if order\_type=1 then Type='Retail Sale';

else if order\_type=2 then do;

Type='Catalog Sale';

SaleAds='Mail';

end;

else if order\_type=3 then do;

Type='Internet Sale';

SaleAds='Email';

end; /\*This line creates the Type variable based on the values of Order\_Type\*/

drop order\_type employee\_ID customer\_ID;

/\*this line removes the order\_type employee\_ID and customer\_ID variables from the data set\*/

run;

proc print data=work.ordertype; /\*this line prints the work.ordertype data set\*/

run;

Chapter 10.1 Level 2

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | |  | Code | Code | Company | Company | ContactType | ContactType | |  | Type | Length | Type | Length | Type | Length | | orion.charities | Char | 6 | Char | 40 | Char | 10 | | orion.us\_suppliers | Char | 6 | Char | 30 | Char | 1 | | orion.consultants | Char | 6 | Char | 30 | Num | 8 | |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |

1. data work.contacts; /\*this line creates a data set work.contacts\*/

set orion.charities orion.us\_suppliers;

/\*this line concatenates the orion.charities and orion.us\_suppliers data sets\*/

run;

proc contents data=work.contacts; /\*this line prints the descriptor portion of the work.contacts data set\*/

run;

1. Submit a PROC CONTENTS step to examine work.contacts. From which input data set were the variable attributes assigned? From orion.charities, the first data set in the set statement
2. data work.contacts2; /\*this line creates a data set work.contacts2\*/

set orion.us\_suppliers orion.charities; /

\*this line concantenates the orion.us\_suppliers and the orion.charities data sets\*/

run;

proc contents data=work.contacts2; /\*this line prints the descriptor portion of the work.contacts2 data set\*/

run;

1. Submit a PROC CONTENTS step to examine work.contacts2. From which input data set were the variable attributes assigned? From orion.us\_suppliers, the first data set in the set statement
2. Why did the DATA step fail? ContactType has been defined as both character and numeric.

Chapter 10.3 Level 2

proc sort data=orion.product\_list /\*this line reads the data from orion.product\_list\*/

out=work.product\_list; /\*this line creates a new data set work.product\_list\*/

by Product\_Level; /\*this line sorts the data set by Product\_Level\*/

run;

data work.listlevel; /\*this line creates a new data set work.listlevel\*/

merge orion.product\_level work.product\_list;

/\*this line merges orion.product\_list and work.product\_list\*/

by Product\_Level; /\*this line performs the merge by the Product\_Level column\*/

keep Product\_ID Product\_Name Product\_Level Product\_Level\_Name;

/\*this line sets the data set to include the Product\_ID Product\_Name Product\_Level and Product\_Level\_Name \*/

run;

proc print data=work.listlevel noobs; /\*this line prints the work.listlevel data set and removes the obs column\*/

where Product\_Level=3; /\*this line sets the data set to include only observations where the Product\_Level is 3\*/

run;

Chapter 10.4 Level 2

proc sort data=orion.customer /\*this line reads the data from the orion.customer data set\*/

out=work.customer; /\*this line creates a new data set called work.customer\*/

by Country; /\*this line sorts by the Country column\*/

run;

data work.allcustomer; /\*this line creates a new data set called work.allcustomer\*/

merge work.customer(in=Cust)

orion.lookup\_country(rename=(Start=Country

Label=Country\_Name) in=Ctry);

/\*this line merges the work.customer and orion.lookup\_country data sets and renames the Start and Label colunns\*/

by Country; /\*this line merges by the Country column\*/

keep Customer\_ID Country Customer\_Name Country\_Name;

/\*this line sets the data set to include the Customer\_ID Country Customer\_Name and Country\_Name columns\*/;

if Cust=1 and Ctry=1;

/\*this line includes only observations where both data sets contributed to the observation\*/

run;

proc print data=work.allcustomer; /\*this line prints the work.allcustomer data set\*/

run;