```
1          OPTIONS NONOTES NOSTIMER NOSOURCE NOSYNTAXCHECK;
NOTE: ODS statements in the SAS Studio environment may disable some output features.
62
63
/*************************************************************************
********/
64         /* Program Name:     pulkit.jain_HW16.sas    */
65         /* Program Location: C:\Users\Pulkit
Jain\Documents\sasuniversityedition\myfolders\assign16 */
66         /* Date Created:     12/04/2017          */
67         /* Author:  Pulkit Jain                  */
68         /* Purpose:   Assignment 16, Read from Raw Datafile    */
69
/*************************************************************************
********/
70
71         /* 1 Use a fileref to access dat file, include headers*/
72         /* Create two libname statements;   */
73         /* Assign library to locaion of hw data with access only; */
74         /* Assign another library with read and write access;     */
75
76
77         filename andro '/folders/myfolders/hw_data/andromeda.dat';
78
79         libname hw_data '/folders/myfolders/hw_data' access=readonly;
NOTE: Libref HW_DATA was successfully assigned as follows:
      Engine:        V9
      Physical Name: /folders/myfolders/hw_data
80         libname pulkit16 '/folders/myfolders/assign16';
NOTE: Libref PULKIT16 was successfully assigned as follows:
      Engine:        V9
      Physical Name: /folders/myfolders/assign16
81
82         /* Specify a fileref to designate output of pdf */
83
84         filename HW16 '/folders/myfolders/assign16/pulkit.jain_HW16_output.pdf';
85
86         /* 2 Read and create datasest with 4 variables, Level, Name, Designation
& Salary*/
87
88         /* Specify options for output pdf file */
89         ods pdf file = HW16 bookmarkgen=yes;
NOTE: Writing ODS PDF output to DISK destination "HW16", printer "PDF".
90         options dtreset;
91
92         data work.andro_data (Keep= Level Employee_Name Job_Title Salary);
93         length Level 3 Employee_Name $25 Job_Title $25;
94         infile andro truncover;
95         * read in levels and check what category they belong to;
96         input @1 row_st1 $8.
97           @10 row_st2 $8.
98           @19 row_st3 $8.
99           @28 row_st4 $8.
100          @37 row_st5 $8.
101          @46 row_st6 $8.
102          @;
103        if row_st1 = '(Level1)' then do;
104        Level = 1;
105        input @10 employee_info $50. @;
```

```
106        end;
107        else if row_st2 = '(Level2)' then do;
108        Level = 2;
109        input @19 employee_info $50. @;
110        end;
111            else if row_st3 = '(Level3)' then do;
112        Level = 3;
113        input @28 employee_info $50. @;
114        end;
115        else if row_st4 = '(Level4)' then do;
116        Level = 4;
117        input @37 employee_info $50. @;
118        end;
119          else if row_st5 = '(Level5)' then do;
120        Level = 5;
121        input @46 employee_info $50. @;
122        end;
123        else do;
124        Level = 6;
125        input @54 employee_info $50. @;
126        end;
127        input @106 Salary dollar10.0 @;
128        * parse job title & employee name from the employee info variable;
129        Job_Title = substr(employee_info, 1, find(employee_info,'(') - 1);
130        Employee_Name = substr(employee_info, find(employee_info,'(') + 1);
131        Employee_Name = compress(Employee_Name, ')');
132        run;

NOTE: The infile ANDRO is:
      Filename=/folders/myfolders/hw_data/andromeda.dat,
      Owner Name=root,Group Name=vboxsf,
      Access Permission=-rwxrwx---,
      Last Modified=04Dec2017:18:11:33,
      File Size (bytes)=49537

NOTE: 424 records were read from the infile ANDRO.
      The minimum record length was 87.
      The maximum record length was 122.
NOTE: The data set WORK.ANDRO_DATA has 424 observations and 4 variables.
NOTE: DATA statement used (Total process time):
      real time          0.01 seconds
      cpu time           0.00 seconds

132     !

133
134
135        /* 3 Use Frequency Procedure on Job_Title */
136
137        PROC FREQ data = andro_data;
138        tables Job_Title;
139        title1 "Analysis of Andromeda Employee Data for Clean Up";
140        title3 "Frequency Report of Job Title";
141        run;

NOTE: There were 424 observations read from the data set WORK.ANDRO_DATA.
NOTE: PROCEDURE FREQ used (Total process time):
      real time          0.25 seconds
      cpu time           0.24 seconds
```

```
142
143         /* 4 Use Univariate Procedure on Salary variable */
144
145         PROC univariate data = andro_data;
146         var Salary;
147         title1 "Analysis of Andromeda Employee Data for Clean Up";
148         title2 "Analysis of Salary Values";
149         run;

NOTE: PROCEDURE UNIVARIATE used (Total process time):
      real time           0.11 seconds
      cpu time            0.11 seconds


150
151         /* 5 Print irregular salaries data */
152
153         PROC PRINT data = andro_data ;
154         where  24000 > Salary  or Salary > 433800;
155         title2 "Salary Values to be Investigated";
156         RUN;

NOTE: There were 7 observations read from the data set WORK.ANDRO_DATA.
      WHERE  not (Salary>=24000 and Salary<=433800);
NOTE: PROCEDURE PRINT used (Total process time):
      real time           0.04 seconds
      cpu time            0.04 seconds


157
158         /* 6 Clean Up the job titles conditionally */
159
160         data work.andro_clean;
161             set work.andro_data;
162             if Job_Title='Accountant i'
163             then Job_Title='Accountant I';
164             else if Job_Title='Accountant ii'
165             then Job_Title='Accountant II';
166             else if Job_Title='Accountant iii'
167             then Job_Title='Accountant III';
168             else if Job_Title='Warehouse Assistant i'
169             then Job_Title='Warehouse Assistant I';
170             else if Job_Title='Warehouse Assistant ii'
171             then Job_Title='Warehouse Assistant II';
172         run;

NOTE: There were 424 observations read from the data set WORK.ANDRO_DATA.
NOTE: The data set WORK.ANDRO_CLEAN has 424 observations and 4 variables.
NOTE: DATA statement used (Total process time):
      real time           0.00 seconds
      cpu time            0.01 seconds


173
174         /* 7 Use Freq procedure to show job titles in cleaned data*/
175
176         proc freq data = work.andro_clean nlevels;
177         table Job_Title / noprint;
178         title1 "Number of Different Jobs in Cleaned Data";
```

```
 179        run;

 NOTE: There were 424 observations read from the data set WORK.ANDRO_CLEAN.
 NOTE: PROCEDURE FREQ used (Total process time):
       real time           0.02 seconds
       cpu time            0.02 seconds


 180
 181        /* 8 Print Employees with titles Chief, Director, or Temp. or Vice
President*/
 182
 183        /* Sort Data By Job level */
 184        proc sort data = work.andro_clean;
 185        by Level;
 186        RUN;

 NOTE: There were 424 observations read from the data set WORK.ANDRO_CLEAN.
 NOTE: The data set WORK.ANDRO_CLEAN has 424 observations and 4 variables.
 NOTE: PROCEDURE SORT used (Total process time):
       real time           0.00 seconds
       cpu time            0.01 seconds


 187
 188        title "List of Andromeda Employees to be Reviewed for Orion Positions";
 189
 190        /* Print required observations */
 191        proc print data=work.andro_clean;
 192          id Level;
 193          by Level;
 194          var Job_Title Employee_Name;
 195          where Job_Title like '%Chief%' or
 196          Job_Title like '%Director%' or
 197          Job_Title like '%Temp.%' or
 198          Job_Title like '%Vice President%';
 199        run;

 NOTE: There were 34 observations read from the data set WORK.ANDRO_CLEAN.
       WHERE Job_Title like '%Chief%' or Job_Title like '%Director%' or Job_Title
like '%Temp.%' or Job_Title like '%Vice
       President%';
 NOTE: PROCEDURE PRINT used (Total process time):
       real time           0.07 seconds
       cpu time            0.07 seconds


 200
 201        /* 9 Houskeeping to make sure title or footnote dont carry over */
 202
 203        title;
 204        footnote;
 205
 206        /* 10 Close PDF Output */
 207        ods pdf close;
 NOTE: ODS PDF printed 11 pages to
/folders/myfolders/assign16/pulkit.jain_HW16_output.pdf.
 208        ods listing;
 209
 210
```

```
211         OPTIONS NONOTES NOSTIMER NOSOURCE NOSYNTAXCHECK;
224
```