```
 1          OPTIONS NONOTES NOSTIMER NOSOURCE NOSYNTAXCHECK;
 NOTE: ODS statements in the SAS Studio environment may disable some
output features.
 62
 63
/*************************************************************************
/
 64         /*Program Name- HW12.sas*/
 65         /* Date Created: November 1 2017 */
 66         /* Author: Dilip Lalwani */
 67         /* Purpose: SAS and working with datasets */
 68
 69         /*1 Output filename and libname statements*/
 70         libname data "/folders/myfolders/data" access=readonly ;
 NOTE: Libref DATA was successfully assigned as follows:
       Engine:        V9
       Physical Name: /folders/myfolders/data
 71         libname hw12 "/folders/myfolders/HW12" ;
 NOTE: Libref HW12 was successfully assigned as follows:
       Engine:        V9
       Physical Name: /folders/myfolders/HW12
 72         filename output
"/folders/myfolders/HW12/dilip.k.lalwani_HW12_output.pdf";
 73         data hw12.zip_codes(drop=estpopulation lastword_county
decommissioned);
 74
 75         /*2a Reduce the length of the county variable to 31 */
 76         length county $ 31;
 77         set data.zip_codes(keep=county decommissioned
estimated_population primary_city state timezone zip
 78         rename=(estimated_population=estpopulation));
 NOTE: Data file DATA.ZIP_CODES.DATA is in a format that is native to
another host, or the file encoding does not match the session
       encoding. Cross Environment Data Access will be used, which
might require additional CPU resources and might reduce
       performance.
 79
 80         /*2 Remove decommissioned zip codes*/
 81         if decommissioned eq 1 then delete;
 82
 83         /*2 Remove observations with states equal to AE, AA or
AP*/
 84         if state in ('AE', 'AA', 'AP') then delete;
```

```
 85
 86
 87         /*2a Use manipulation functions to modify county values*/
 88         lastword_county=scan(county,-1);
 89         if UPCASE(lastword_county) in
('COUNTY','PARISH','BOROUGH') then do;
 90         county=substr(county,1,length(county)-length(scan(county,-
1)));
 91         end;
 92
 93         /*2b Convert estimated_population variable from character
to numeric*/
 94         estimated_population=input(estpopulation, 8.);
 95
 96         /*2c Replace underscores with blank space in timezone
variable*/
 97         if timezone eq "America/New_York" then
 98         do;
 99         substr(timezone,12,1)=' ';
 100        end;
 101        else if timezone eq "America/Los_Angeles" then
 102        do;
 103        substr(timezone,12,1)=' ';
 104        end;
 105        else if timezone eq "America/Puerto_Rico" then
 106        do;
 107        substr(timezone,15,1)=' ';
 108        end;
 109        else if timezone eq "America/Indiana/Tell_City" then
 110        do;
 111        substr(timezone,21,1)=' ';
 112        end;
 113        else if timezone eq "America/North_Dakota/Center" then
 114        do;
 115        substr(timezone,14,1)=' ';
 116        end;
 117
 118        label zip="Zip Code"
 119        primary_city="City"
 120        state="State"
 121        timezone="Time Zone"
 122        county="County"
 123        estimated_population="Est Population";
 124        run;
```

```
 NOTE: Character values have been converted to numeric values at the
places given by: (Line):(Column).
       81:4
 WARNING: Multiple lengths were specified for the variable county by
input data set(s). This can cause truncation of data.
 NOTE: There were 42522 observations read from the data set
DATA.ZIP_CODES.
 NOTE: The data set HW12.ZIP_CODES has 41317 observations and 6
variables.
 NOTE: DATA statement used (Total process time):
       real time           0.20 seconds
       cpu time            0.16 seconds


 125
 126        /*3a Use a sort procedure to sort the clean data set -
sorting by state and primary_city*/
 127        proc sort data=hw12.zip_codes;
 128        by state primary_city;
 129        run;

 NOTE: There were 41317 observations read from the data set
HW12.ZIP_CODES.
 NOTE: The data set HW12.ZIP_CODES has 41317 observations and 6
variables.
 NOTE: PROCEDURE SORT used (Total process time):
       real time           0.08 seconds
       cpu time            0.04 seconds


 130
 131        /*3b Remove zip and estimated_population from the
dataset*/
 132        data work.zipstats(drop=zip estimated_population
timezone);
 133        set hw12.zip_codes;
 134        retain zip_codes;
 135        length zip_codes $1700;
 136        by state primary_city;
 137
 138        /*3c Store total of estimated population values for each
city in est_city_population*/
 139        if First.primary_city then do;
 140        est_city_population = 0;
 141        zip_codes='';
 142        end;
```

```
 143        est_city_population+estimated_population;
 144
 145        /*3d List all zips into Zip Codes variable and create
summary*/
 146        zip_codes=catx(',',zip_codes, zip);
 147        if Last.primary_city;
 148        label est_city_population="Est. City Population"
 149        zip_codes="Zip Codes"
 150        primary_city="City"
 151        state="State"
 152        county="County";
 153        format est_city_population COMMA10.;
 154
 155        /*3e Output cities with estimated city population greater
than 0*/
 156        if est_city_population<=0 then delete;
 157        run;

 NOTE: There were 41317 observations read from the data set
HW12.ZIP_CODES.
 NOTE: The data set WORK.ZIPSTATS has 21404 observations and 5
variables.
 NOTE: DATA statement used (Total process time):
       real time           0.14 seconds
       cpu time            0.13 seconds


 158
 159        /*4 Open PDF Destination and output observations only for
selected cities*/
 160        ods pdf file=output bookmarkgen=no;
 NOTE: Writing ODS PDF output to DISK destination "OUTPUT", printer
"PDF".
 161        proc contents data=hw12.zip_codes;
 162        title "4.1 Descriptor Portion of Cleaned Zip Code Data
Set";
 163        run;

 NOTE: PROCEDURE CONTENTS used (Total process time):
       real time           0.12 seconds
       cpu time            0.12 seconds


 164        proc print data =hw12.zip_codes label;
 165        title "4.2 Cleaned Zip Codes from Selected Cities";
 166        where (primary_city ='Center'
```

```
 167        or primary_city='Buffalo'
 168        or primary_city='Las Vegas'
 169        or primary_city='Bristow'
 170        or primary_city='Muleshoe'
 171        or primary_city='Athens'
 172        or primary_city='Carolina'
 173        or primary_city='Auke Bay'
 174        or primary_city='Washington');
 175        var zip primary_city state timezone county
estimated_population;
 176        run;

 NOTE: There were 489 observations read from the data set
HW12.ZIP_CODES.
       WHERE primary_city in ('Athens', 'Auke Bay', 'Bristow',
'Buffalo', 'Carolina', 'Center', 'Las Vegas', 'Muleshoe',
       'Washington');
 NOTE: PROCEDURE PRINT used (Total process time):
       real time           1.08 seconds
       cpu time            1.05 seconds


 177        proc contents data=work.zipstats;
 178        title "4.3 Descriptor Portion of Summarized Zip Codes Data
Set";
 179        run;

 NOTE: PROCEDURE CONTENTS used (Total process time):
       real time           0.06 seconds
       cpu time            0.06 seconds


 180        proc print data =work.zipstats label;
 181        title "4.4 Summarized Zip Codes from Selected Cities";
 182        where (primary_city ='Center'
 183        or primary_city='Buffalo'
 184        or primary_city='Las Vegas'
 185        or primary_city='Bristow'
 186        or primary_city='Muleshoe'
 187        or primary_city='Athens'
 188        or primary_city='Carolina'
 189        or primary_city='Auke Bay'
 190        or primary_city='Washington');
 191        var primary_city state county zip_codes
est_city_population;
 192        run;
```

```
 NOTE: There were 61 observations read from the data set
WORK.ZIPSTATS.
       WHERE primary_city in ('Athens', 'Auke Bay', 'Bristow',
'Buffalo', 'Carolina', 'Center', 'Las Vegas', 'Muleshoe',
       'Washington');
 NOTE: PROCEDURE PRINT used (Total process time):
       real time           0.17 seconds
       cpu time            0.18 seconds


 193         ods pdf close;
 NOTE: ODS PDF printed 20 pages to
/folders/myfolders/HW12/dilip.k.lalwani_HW12_output.pdf.
 194
 195         OPTIONS NONOTES NOSTIMER NOSOURCE NOSYNTAXCHECK;
 208
```