

```

data cleaned_tourism;
set work.tourism;
*put suitable length for column category;
length category $50;
*ratain country_name,Tourism_type to meets output like requirements;
retain country_name;
retain Tourism_type;
*reorganizing country column;
* 1 Tourism_type valid values are Inbound tourism,Outbound tourism;
if Country="Inbound tourism" or Country="Outbound tourism" then Tourism_type=country;
* 2 category have 6 valid values in requirements;
if country="Arrivals - Thousands" then category="Arrivals";
else if country="Departures - Thousands" then category="Departures";
else if country="Passenger transport - US$ Mn" then category="Passenger transport - US$";
else if country="Tourism expenditure in other countries - US$
Mn" then category="Tourism expenditure in other countries - US$";
else if country="Tourism expenditure in the country - US$ Mn" then category="Tourism expenditure in the country - US$";
else if country="Travel - US$ Mn" then category="Travel - US$";
* make series upercase;
Series=upcase(Series);
if Series=".." then Series="";
* 3 ConversionType can be mn or thousands from column country;
ConversionType=scan(country,-1,' -');
* create numeric column and convert char col to numeric;
if ConversionType="Mn" then Y2014=input(_2014,16.)*1000000;
else if ConversionType="Thousands" then Y2014=input(_2014,16.)*1000;
format Y2014 comma12.;
* 4 create country_name from country col;
if A ne "." then Country_Name = country;
keep Country_Name Tourism_Type Category Series Y2014;
*sort cleaned_tourism before merging;
proc sort data=cleaned_tourism;
by country_name;
*sort country_info before merging;
proc sort data=country_info(rename=(country=country_name));
by country_name;
*creat a format for Continent column;
proc format;
value Continent
1 = "North America"
2 = "South America"

```

```
3 = "Europe"
4 = "Africa"
5 = "Asia"
6 = "Oceania"
7 = "Antarctica";
run;
*Merging matching columns and put non matching in another table;
data final_tourism nocountryfound(keep=Country_Name);
merge cleaned_tourism(in=ct)
country_info(in=ci);
by country_name;
if(ct=1 and ci=1) then output final_tourism;
if(ct=1 and ci=0) then output nocountryfound;
format Continent continent.;
```