



Advanced SAS Programming Project

Presented by,

PRIYANKA SAXENA



QUESTION:-1

Produce a customized table report of total population by variable region and output the report as PDF file.



QUES.1 :Output

Regionwise Population Total
Date Created:- 11JUL18
Prepared By Priyanka Saxena

Region Name	Total Population
AFR	738,083,274
AMR	886,334,010
EMR	538,001,455
EUR	1,061,732,810
SEAR	1,656,528,818
WPR	1,791,767,503

QUES1: Coding

Proc sql;

Options nodate nonumber;

Title1 'Regionwise Population Total';

Title2 "Date Created:- &sysdate";

Title3 'Prepared By Priyanka Saxena';

Footnote j=r 'Demographic Data Source:World Population Prospects - 2004 Rev.';

Create Table RegionwisePopulation as select region label='Region Name'

sum(pop)format=comma15. as T_Pop label='Total Population'

from sashelp.Demographics

Group by region;

Quit;

ods pdf file='C:\Users\Priyanka\Documents\Pri SAS Folder\Regionwise Population.pdf';

proc print data=RegionwisePopulation label noobs

style(header)={backgroundcolor=Pink color=Green just=l};

var region T_Pop;

Run;

ods pdf close;



QUES:2

Create an indicator for adult literacy rate percentage with cutoff points at 90% and 50%. Then create a table report to show total population, average population annual growth rate percentage and average population in urban areas percentage for each region and your indicator level .

OBJECT:-2 Output

Tabular report for Regional Adult Literacy Rate Level against Total Population,%Population AGR & % Population in Urban Areas

Date Created:- 11JUL18

Prepared By Priyanka

Region Vs Adult Literacy Rate		Population	% Population Annual Growth Rate	% Population in Urban Areas
		Sum of Population	Average	Average
Region Name	Adult Literacy rate Level			
AFR	Low Literacy	195,489,273	2.45%	34.77%
	Medium Literacy	511,056,571	2.16%	39.21%
	High Literacy	13,090,188	0.75%	43.05%
AMR	Medium Literacy	276,518,961	1.57%	59.33%
	High Literacy	277,834,534	1.18%	70.71%
EMR	Low Literacy	178,909,730	2.51%	30.55%
	Medium Literacy	287,823,808	2.44%	73.03%
EUR	Medium Literacy	73,594,468	1.00%	79.70%
	High Literacy	412,932,333	0.18%	58.29%
SEAR	Low Literacy	168,954,905	1.93%	20.40%
	Medium Literacy	1,376,671,781	1.29%	35.73%
	High Literacy	85,304,861	1.40%	27.73%
WPR	Medium Literacy	51,441,032	2.02%	28.66%
	High Literacy	1,499,118,146	1.25%	52.68%

Demographic Data Source:World Population Prospects - 2004 Rev.

OBJECT:-2 Coding

```
*Project Problem:-2;
Proc Format;
Value Avg_Lit_Rate
0-<0.5= 'Low Literacy'
0.5-<0.9 = 'Medium Literacy'
0.9-high = 'High Literacy'
other ='No Data';
Run;

Proc Tabulate data=sashelp.Demographics;
class region AdultLiteracypct;
Format AdultLiteracypct Avg_Lit_Rate.;
Var pop popAGR popUrban;
label pop='Population'
popAGR='% Population Annual Growth Rate'
popUrban='% Population in Urban Areas'
region='Region Name'
AdultLiteracypct='Adult Literacy rate Level';
Table region*AdultLiteracypct,
(pop*f=comma15.) (popAGR*f=percent9.2
popUrban*f=percent9.2)*mean
/ box='Region Vs Adult Literacy Rate';
keylabel
Sum='Sum of Population'
mean='Average';
Title1 'Tabular report for Regional Adult Literacy Rate Level against Total Population
,%Population AGR & % Population in Urban Areas';
Title2 "Date Created:- &sysdate";
Title3 'Prepared By Priyanka';
Footnote j=r 'Demographic Data Source:World Population Prospects - 2004 Rev.';
Run;
```



QUES :3

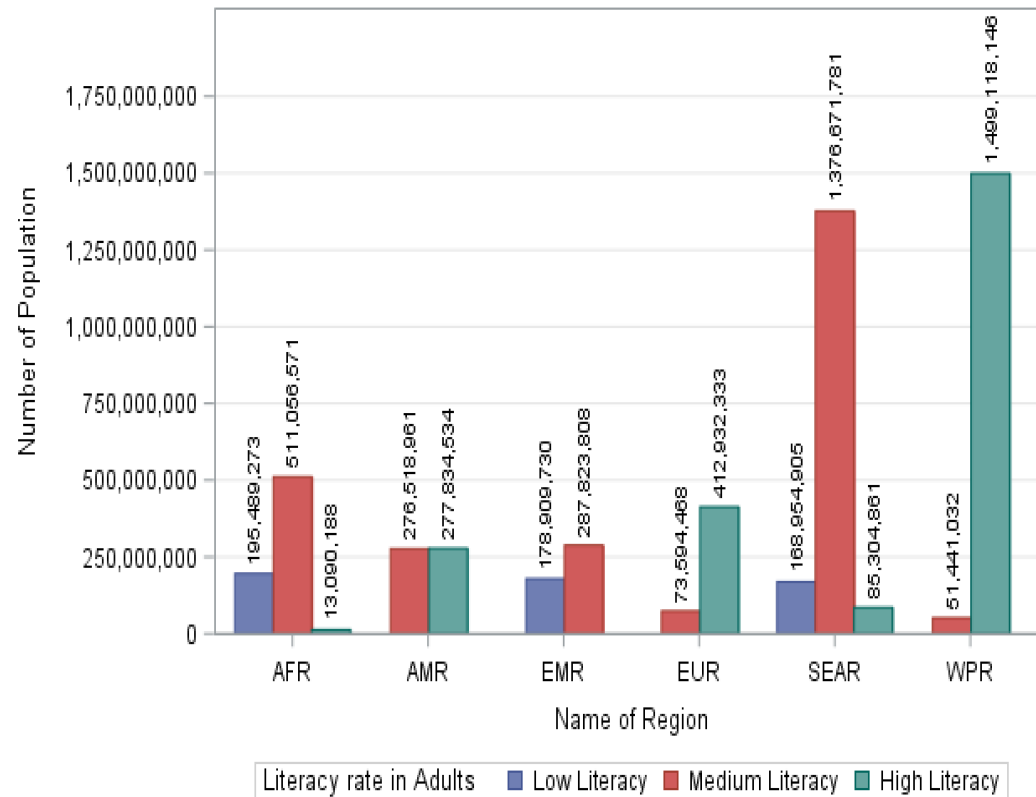
Produce a bar chart for total population by regions with groups by indicator created in object 2. Also create a histogram with normal density curve for population in urban areas in year 2005 with 10 bins.

QUES.3 :Output

Regionwise Vertical Bar Chart for Total Population with Groups by Adult Literacy rate Level (new indicator)

Date Created:- 11JUL18

Prepared By Priyanka

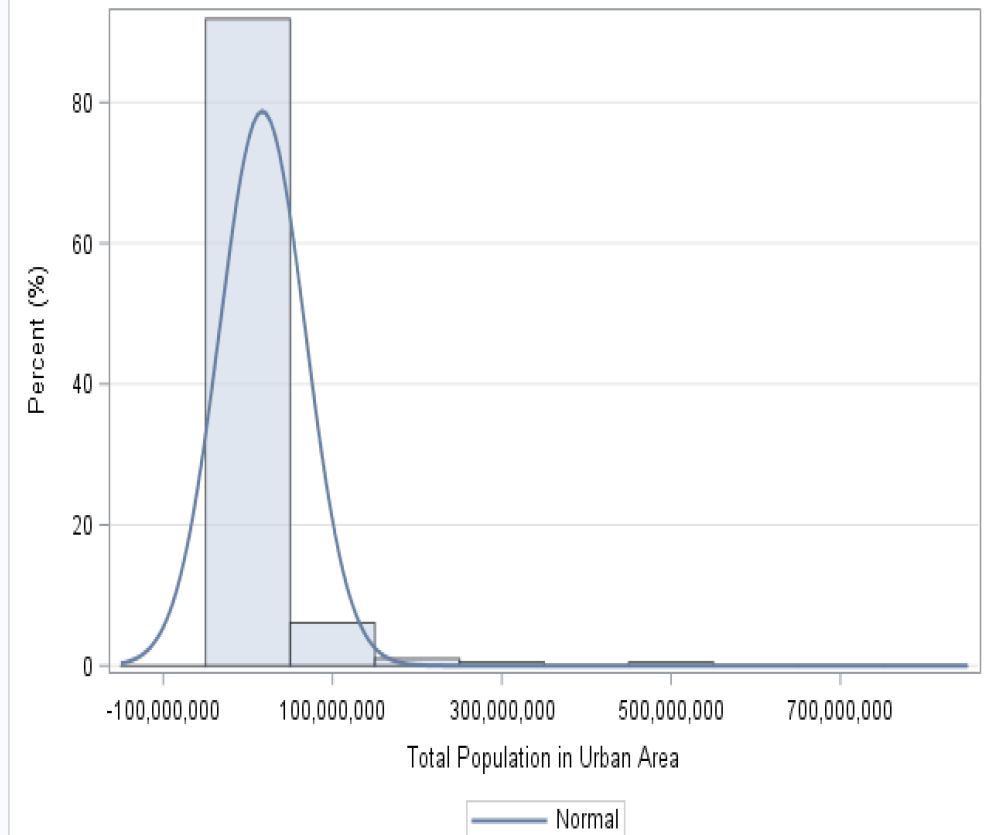


Demographic Data Source: World Population Prospects - 2004 Rev.

Histogram_Normal Density Curve_Urban Population

Date Created:- 11JUL18

Prepared By PRIYANKA



Demographic Data Source: World Population Prospects - 2004 Rev.

OBJECT:-3 Coding

```
|
❏ Proc Sgplot Data=sashelp.Demographics;
  VBAR region
  /GROUP=AdultLiteracypct
  RESPONSE= pop
  datalabel groupdisplay=cluster;
  yaxis grid
  label= 'Number of Population'
  values= (0 to 1750000000 by 250000000);
  Format AdultLiteracypct Avg_Lit_Rate.;
  Label region='Name of Region'
  AdultLiteracypct='Literacy rate in Adults'
  pop='Population' ;
  Title " Regionwise Vertical Bar Chart for Total Population
  with Groups by Adult Literacy rate Level (new indicator)";
  Title2 "Date Created:- &sysdate";
  Title3 'Prepared By Priyanka';
  Footnote j=r 'Demographic Data Source:World Population
  Prospects - 2004 Rev.';;
  Run;
```

QUES.3 : Coding

❏ PROC SQL;

```
Create Table UrbanPopulation as select pop Label='TOTAL POPULATION',  
popUrban label='%POPULATION URBAN',  
pop*popUrban as TotPopUrban label='URBAN POPULATION_TOTAL' format=comma15.  
from sashelp.Demographics;  
quit;
```

❏ Proc sgplot data=UrbanPopulation;

```
Histogram TotPopUrban/  
showbins nbins=10  
transparency=.40;  
yaxis grid label='Percent (%)';  
Density TotPopUrban;  
label TotPopUrban='Total Population in Urban Area';  
Title 'Histogram_Normal Density Curve_Urban Population';  
Title2 "Date Created:- &sysdate";  
Title3 'Prepared By PRIYANKA';  
Footnote j=R 'Demographic Data Source:World Population  
Prospects - 2004 Rev.';  
Run;
```



QUES:4

Create a new SAS data set with variables GLC country name, region, population, population in urban areas percentage, gross national income per capita, a new variable with values 'low' (income per capita \leq 5,000), 'middle' (5000<income per capita \leq 20,000), 'high' (income per capita $>$ 20,000) for gross national income per capita. The data set only with observations from region AMR and EUR. Present the SAS codes and results in the presentation.

QUES.4: Output

POPULATION BY COUNTRY, % URBAN, PER CAPITA GNI AND PER CAPITA GNI LEVEL FOR
"AMR" & "EUR" REGION
Date Created:- 11JUL18
PREP. BY PRIYANKA

Obs	COUNTRY_NAME	REGION_NAME	POPULATION	%POPULATION_URBAN	GROSS NATIONAL INCOME (PER CAPITA)	GNI PER CAPITA
61	LITHUANIA	EUR	3,431,033	66.60%	12610	MIDDLE
62	SLOVAKIA	EUR	5,400,908	58.00%	14370	MIDDLE
63	LUXEMBOURG	EUR	464,904	92.40%	61220	HIGH
64	MACEDONIA	EUR	2,034,060	59.70%	6480	MIDDLE
65	MOLDOVA	EUR	4,205,747	46.30%	1930	LOW
66	MAN	EUR	59,667,844	89.20%	31460	HIGH
67	MALTA	EUR	401,630	92.10%	18720	MIDDLE
68	MONACO	EUR	35,253	100.00%	.	NO DATA
69	NETHERLANDS	EUR	16,299,173	66.80%	31220	HIGH
70	NORWAY	EUR	4,620,275	80.50%	38550	HIGH
71	SERBIA	EUR	10,503,115	52.30%	.	NO DATA
72	POLAND	EUR	38,529,562	62.00%	12640	MIDDLE
73	PORTUGAL	EUR	10,494,502	55.60%	19250	MIDDLE
74	ROMANIA	EUR	21,711,472	54.70%	8190	MIDDLE
75	SAN MARINO	EUR	28,117	88.70%	.	NO DATA
76	SLOVENIA	EUR	1,966,814	50.80%	20730	HIGH
77	SPAIN	EUR	43,064,189	76.70%	25070	HIGH
78	SWEDEN	EUR	9,041,262	83.40%	29770	HIGH
79	SWITZERLAND	EUR	7,252,331	67.50%	35370	HIGH
80	UNITED KINGDOM	EUR	59,667,844	89.20%	31460	HIGH
81	UKRAINE	EUR	46,480,703	67.30%	6250	MIDDLE
82	CYPRUS	EUR	835,307	69.50%	22330	HIGH
83	ISRAEL	EUR	6,724,564	91.70%	23510	HIGH
84	KYRGYZSTAN	EUR	5,263,794	33.70%	1840	LOW
85	KAZAKHSTAN	EUR	14,825,105	55.90%	6980	MIDDLE
86	TAJIKISTAN	EUR	6,506,980	24.20%	1150	LOW
87	RUSSIA	EUR	143,201,572	73.30%	9620	MIDDLE
88	TURKEY	EUR	73,192,838	67.30%	7680	MIDDLE
89	TURKMENISTAN	EUR	4,833,266	45.80%	6910	MIDDLE
90	UZBEKISTAN	EUR	26,593,123	36.40%	1860	LOW

OBJECT:-4 Coding

*Project Problem:-4;

```
❑ PROC SQL;
  OPTION NODATE NONUMBER;
  CREATE TABLE SUBDEMOGRAPHICS as
  SELECT
    NAME LABEL = 'COUNTRY_NAME',
    REGION LABEL='REGION_NAME',
    POP LABEL= 'POPULATION',
    POPURBAN LABEL= '%POPULATION_URBAN',
    GNI LABEL='GROSS NATIONAL INCOME_(PER CAPITA)',
    case
      WHEN GNI =. THEN 'NO DATA'
      WHEN GNI <=5000 THEN 'LOW'
      WHEN GNI <=20000 THEN 'MIDDLE'
      WHEN GNI >20000 THEN 'HIGH'
    END AS PCILEVEL LABEL= 'GNI_PER CAPITA'
  FROM sashelp.Demographics
  WHERE REGION IN ('AMR' 'EUR');
  QUIT;

❑ PROC PRINT DATA=SUBDEMOGRAPHICS LABEL
  STYLE(HEADER)={BACKGROUNDCOLOR=PINK COLOR=GREEN JUST=C};
  VAR NAME region pop popUrban GNI PCILevel;
  TITLE 'POPULATION BY COUNTRY, % URBAN_POPULATION,
  PER CAPITA GNI AND PER CAPITA GNI LEVEL FOR "AMR" & "EUR"
  REGION';
  TITLE2 "Date Created:- &sysdate";
  TITLES3 'PREP. BY PRIYANKA';
  FOOTNOTE J=R 'Demographic Data Source:World Population Prospects - 2004 Rev.';
  RUN;

❑ DATA SUBDEMOGRAPHICS;
  LENGTH PERCAPITA $8.;
  SET sashelp.Demographics;
  WHERE REGION IN ("AMR" "EUR");
  IF GNI=. THEN percapita='Missing Value';
  IF GNI<=5000 THEN percapita='Low';
  IF 5000<GNI<=20000 THEN percapita='Middle';
  IF GNI>20000 THEN percapita='High';
  KEEP NAME GNI region pop popurban percapita ;
  RUN;

❑ PROC PRINT DATA=SUBDEMOGRAPHICS LABEL NOOBS;
  RUN;

❑ PROC FREQ DATA=SUBDEMOGRAPHICS;
  TABLE PERCAPITA;
  RUN;

❑ PROC SORT; BY REGION;
```



THANK YOU