Statistical Journal of the IAOS 36 (2020) S49–S62
DOI 10.3233/SJI-200693
IOS Press

Data science skills: Building partnership for efficient school curriculum delivery in Africa

Nureni Olawale Adeboye[a,*], Peter Osuolale Popoola[b] and Oluwatobi Nurudeen Ogunnusi[a]
[a]*Mathematics and Statistics Department, Federal Polytechnic, Ilaro, Ogun State, Nigeria*
[b]*Mathematics and Statistics Department, The Ibarapa Polytechnic, Eruwa, Oyo State, Nigeria*

Abstract. Data science is a concept to unify statistics, data analysis, machine learning and their related methods in order to analyze actual phenomena with data to provide better understanding. This article focused its investigation on acquisition of data science skills in building partnership for efficient school curriculum delivery in Africa, especially in the area of teaching statistics courses at the beginners' level in tertiary institutions. Illustrations were made using Big data of selected 18 African countries sourced from United Nations Educational, Scientific and Cultural Organization (UNESCO) with special focus on some macro-economic variables that drives economic policy. Data description techniques were adopted in the analysis of the sourced open data with the aid of R analytics software for data science, as improvement on the traditional methods of data description for learning and thus open a new charter of education curriculum delivery in African schools. Though, the collaboration is not without its own challenges, its prospects in creating self-driven learning culture among students of tertiary institutions has greatly enhanced the quality of teaching, advancing students skills in machine learning, improved understanding of the role of data in global perspective and being able to critique claims based on data.

Keywords: Big data, curriculum, data science, data description, statistics learning

1. Introduction

Data science is a "concept to unify statistics, data analysis, machine learning and their related methods" in order to "understand and analyze actual phenomena" with data. It employs techniques and theories drawn from many fields within the context of mathematics, statistics, computer science, and information science.

Data Science has spread its branches through several quintessential fields in modern day learning. It has emerged as a global phenomenon that has revolutionized industries and has increased their performances substantially [1]. Given the vast increase in the volume and complexity of data and the new technologies that have been developed to process and analyze this information, it can be argued that there is an increased need for statistical thinking in the context of working with data [2]. Key statistical reasoning topics that are critical for Data Scientists to know at a deep level include but are not limited to the following: developing clear statements of the problem/scientific research question; ensuring acquisition of high-quality data; understanding the process that produced the data, to provide proper context for analysis; allowing domain knowledge of the problem to guide both data collection and analysis; approaching modeling as a process that requires an overall strategy.

The modern day "romance" between Data Science and Statistics cannot be overemphasized (see Fig. 1). Statistics can be a powerful tool when performing the art of Data Science. From a high-level view, statistics is the use of mathematics to perform technical analysis of data. A basic visualization such as a bar chart might give some high-level information, but with statistics one gets to operate on the data in a much more information-driven and targeted way. The analysis involved helps to form concrete conclusions about our data rather than

*Corresponding author: Nureni Olawale Adeboye, Mathematics and Statistics Department, Federal Polytechnic, Ilaro P.M.B 50, Ilaro, Ogun State, Nigeria. Tel.: +234 8033348141; E-mail: nureni.adeboye@federalpolyilaro.edu.ng.

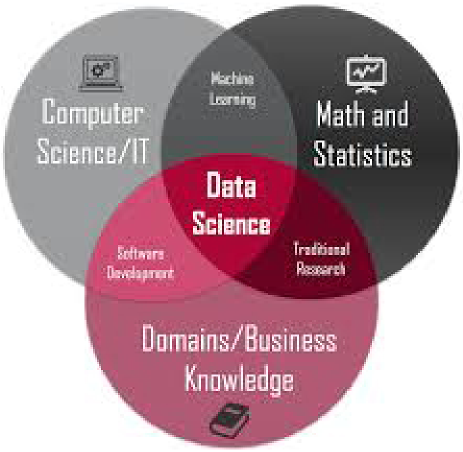

Fig. 1. The interactive disciplines of data science.

just guesstimating. Using statistics, we can gain deeper and more fine grained insights into how exactly our data is structured and based on that structure, optimally apply other data science techniques to get even more information [3].

Education is the key to shaping the lives of people. Since the dawn of civilization, humans have evolved through education and have developed mechanisms to improve education. In the 21st century, where data is omnipresent in every walk of life, education is no exception. With advancements in computing techniques, it is possible to imbibe all the information through powerful big-data platforms [4]. Various Schools have to keep themselves updated with the demands of the industry so as to provide appropriate courses to their students. Furthermore, it is a challenge for the Schools to keep up with the growth of industries. In order to accommodate this, Schools are using Data Science systems to analyze growing trends in the market [5]. Using various statistical measures and monitoring techniques, data science can be useful for analyzing the industrial patterns and help the course creators to imbibe useful topics. Furthermore, using predictive analytics, Schools can analyze demands for new skill sets and curate courses that address them [6].

The performance of students depends on the teachers. While there are many assessment techniques that have been used to assess the performance of teachers, it has been mostly manual in nature. With the breakthrough in data science, it is possible to keep track of the teacher performance. This is not only valid for recorded data but also real-time data. As a result, with real-time monitoring of teachers, rigorous data collection is possible, along with its analysis. Furthermore, we can store and manage unstructured data like student reviews on a big data platform.

1.1. Data science and statistics curriculum

A growing number of students are completing bachelor's degrees in statistics and entering the workforce as data analysts. In these positions, they are expected to understand how to use databases and other data warehouses, scrape data from Internet sources, program solutions to complex problems in multiple languages, and think algorithmically as well as statistically [7]. This increase in the number of undergraduates may help address the impending shortage of quantitatively trained workers. Statistics graduates at the bachelor's level often work as analysts, and as a result need training in statistical methods, statistical thinking and statistical practice; a foundation in theoretical statistics; increased skills in computing and data-related technologies; and the ability to communicate [6,7]. Computing skills to enable processing of large data sets are particularly relevant, as noted in the recent London Report on the Future of Statistics. Much of the statistics education literature focuses on the introductory statistics course and statistics before college. Given the relatively few decades since the establishment of undergraduate statistics programs, this is not surprising. While there has been impressive growth in the number of students taking introductory statistics, there has been a relative dearth of articles on the curriculum beyond the introductory course [8].

The digital age is having a profound impact on statistics and the nature of data analysis, and these changes necessitate revaluation of the training and education practices in statistics. Computing is an increasingly important and necessary aspect of a statistician's work, and needs to be incorporated into statistics [9]. Successful statisticians must be familiar with the computer, for they are expected to be able to access data from various sources, apply the latest statistical methodologies, and communicate their ïňĄndings to others in novel ways and via new media. In addition, researchers exploring new statistical methodology rely on computer experiments and simulation to explore the characteristics of methods as an aid to formalizing their mathematical framework [10–12].

Thus, for the field of statistics to have its greatest impact on policy and science, statisticians must seriously reflect on these major changes and their implications for statistics education. Faculty of science in African higher institutions needs to indicate to students that computing and data science is an important element of their statistics education, and it must be taught with an intellectual foundation that provides students with skills to reason about important computational tasks and continue to learn about new computational topics in statistics and Data science. Instead of teaching similar concepts with varying degrees of mathematical rigor, statisticians need to address what is missing from the curricula and take the lead in improving the level of students' data competence. It is our responsibility, as statistics educators, to ensure our students have the computational understanding, skills, and coniïňĄdence needed to actively and whole-heartedly participate in the computational arena.

Based on the discussion above, traditional statistics is the basis of data science, but there should be some improvement in the statistics curriculum. These changes are necessary in order to attract and prepare future statisticians, and to keep pace with the rapidly changing "big science" fields. As the practice of science and statistics research continues to change, its perspective and attitudes must also change so as to realize the field's potential and maximize the important influence that statistical thinking has on scientific endeavors.

2. Materials and methods

2.1. Materials

Social-economic panel data spanning between year 1999 and 2018, consisting of variables GDP at Purchasing Power Parity (PPP) per capita (constant 2011 international $), GNI per capita based on PPP and Official Exchange rates of sixteen Eq. (16) West African countries as published by United Nations Educational, Scientific and Cultural Organization (UNESCO), was used for data description and visualization in R-statistical software for data science. This made the dataset (named as ***social.csv***) to contain 320 rows and 4 columns. The data frame includes the following columns with description:

1. Variable ***Country*** relates to each of the West African countries as two letters abbreviation. A factor with levels: BJ, Benin; BF, Burkina Faso; CV, Cape Verde; GM, Gambia; GH, Ghana; GN, Guinea; GW, Guinea Bissau; CI, Cote d'Ivoire; LR, Liberia; ML, Mali; MR, Mauritania; NE, Niger; NG, Nigeria; SN, Senegal; SL, Sierra Leone; and TG, Togo was used to represent those countries as published by UNESCO.
2. Variable ***GDP*** at PPP per capita is the Gross Domestic Product adjusted for inflation. It relates to the total monetary or market value of all finished goods and services produced within countries borders in a specific period of time divided by the average (or mid-year) population for the same year.
3. Variable ***GNIPC based on PPP*** (US$) is referred to as the Gross National Income Per Capita based on the Purchasing Power Parity rates. It is the gross national income, converted to US dollars using the PPP rates.
4. Variable ***ER*** is shortened as Exchange Rate. It is the value of the selected West Africans currencies in relation to the United States' (US$) currency.

These variables were used to explain the data description techniques to the students, which also serves as a mean of driven their knowledge on the usefulness of socio-economic indicators.

2.2. Methods

Descriptive Statistics: Descriptive statistics is the first technique used to represent nearly every dataset as they form the foundations for more complicated computations. R sets of commands were generated for the statistics and used to calculate summary statistics, including mean, standard deviation, range, quartile and percentilepercentile as expressed in the following equations:

Arithmetic Mean: The arithmetic mean of observations $x_1, x_2, \dots x_n$ for ungrouped data is given by

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \quad i = 1, 2, \dots, n \tag{1}$$

For grouped data, we have

$$\bar{x} = \frac{1}{\sum_{i=}^{n} f_i}\sum_{i=1}^{n} f_i x_i \quad i = 1, 2, \dots, n \tag{2}$$

Where f_i is the frequency of each observations.

Median: The middle value after a set of observations $x_1, x_2, \dots x_n$ is arranged in order of magnitude is given by

$$x_{med} = \frac{1}{2}(n+1)th \text{ observation} \tag{3}$$

Equation (3) is used when the number of observation is odd. But when the number of observation is even, we have

$$x_{med} = \frac{1}{2}\left[\left(\frac{1}{2}nth\ observation\right) + \frac{1}{2}(n+1)th \text{ observation}\right] \tag{4}$$

For grouped observations with corresponding frequencies $f_1, f_2, \ldots, f_n$, we have

$$x_{med} = L_1 + \left[\frac{\frac{1}{2}N - \sum f^*}{f_m}\right] C \tag{5}$$

Where; L_1 is the lower class boundary of the median class; N is the total observations under consideration; $\sum f^*$ is the cumulative of the frequencies preceding the median class; f_m is the frequency of the median class. However, the median class is determined by the class to which $\frac{1}{2}n$ falls in the cumulative frequency column.

Variance: The variance of observations $x_1, x_2, \ldots x_n$ for ungrouped data is given by

$$s^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2 \quad i = 1, 2, \ldots, n \tag{6}$$

For ungrouped data, we have

$$s^2 = \frac{1}{\sum_{i=}^{n} f_i}\sum_{i=1}^{n} f_i(x_i - \bar{x})^2$$
$$i = 1, 2, \ldots, n \tag{7}$$

Square root of Eqs (6) and (7) give the standard deviation.

Range (R): Given observations $x_1, x_2, \ldots x_n$, the difference between the maximum and minimum value is referred to as the range. It is given as

$$R = \text{max } \textit{observation} - \text{min } \textit{observation} \tag{8}$$

Quartiles: this divides a given set of observations $x_1, x_2, \ldots x_n$ into four Eq. (4) equal parts given as

$$Q_i = L_{Q_i} + \left[\frac{\frac{1}{4}N - \sum f^*_{Q_i}}{f_{Q_i}}\right] C,$$
$$i = 1, 2, 3 \tag{9}$$

Where Q_i is the i^{th} quartile; $\sum f^*_{Q_i}$ is the cumulative frequencies preceding the i^{th} quartiles class; f_{Q_i} is the frequency of the i^{th} quartile class; C is the class interval.

Percentiles: This divide a given set of observations $x_1, x_2, \ldots x_n$ into hundred (100) parts, give as

$$P_i = L_{P_i} + \left[\frac{\frac{1}{100}N - \sum f^*_{P_i}}{f_{P_i}}\right] C,$$
$$i = 1, 2, \ldots, 99 \tag{10}$$

Where P_i is the i^{th} percentilepercentile; $\sum f^*_{P_i}$ is the cumulative frequencies preceding the i^{th} percentiles class; f_{P_i} is the frequency of the i^{th} percentile class; C is the class interval.

Moments: Given observations $x_1, x_2, \ldots x_n$, the r^{th} moment about the **origin** for grouped and ungrouped data is defined by

$$\mu_r = \frac{1}{n}\sum_{i=1}^{n} x^r; r = 1, 2, \ldots, n \tag{11}$$

$$\mu_r = \frac{\sum_{i=1}^{n} f x^r}{\sum_{i=1}^{n} f}; r = 1, 2, \ldots, n \tag{12}$$

However, the corresponding r^{th} moment about the mean for ungrouped data is defined by:

$$\mu_r = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^r \tag{13}$$

$$\mu_r = \frac{1}{\sum_{i=}^{n} f_i}\sum_{i=1}^{n} f_i(x_i - \bar{x})^r$$
$$i = 1, 2, \ldots, n \tag{14}$$

Equating $r = 1, 2, 3, 4, \ldots$ corresponds to first moment, second moment, third moment, forth moment and so on.

Skewness and Kurtosis: Skewness is the measure of departure of a curve from symmetry. The distribution of a set of data is symmetrical if the three measures of central tendencies coincide while Kurtosis is the measure of Peakedness. Students were exposed to how Skewness and Kurtosis of a curve can be measured using method of moments as given below:

$$\alpha_1 = \frac{\mu_1}{\sigma} = 0 \tag{15}$$

$$\alpha_2 = \frac{\mu_2}{\sigma^2} = \frac{\sigma^2}{\sigma^2} = 1 \tag{16}$$

$$\alpha_3 = \frac{\mu_3}{\sigma^3} = \frac{\mu_3}{\mu_2^{3/2}} = 1 \tag{17}$$

$$\alpha_4 = \frac{\mu_4}{\sigma^4} = \frac{\mu_4}{\mu_2} = 1 \tag{18}$$

If $\alpha_3 = 0$, then the distribution is symmetrical, but if $\alpha_3 < 0$, we have a negatively skewed curve and $\alpha_3 > 0$ indicates a positively skewed curve. α_4 measures the 4^{th} moment (Peakedness) of the dataset hereinafter referred to as Kurtosis. The criteria is to know if the curve is either Mesokurtic ($\alpha_4 = 3$), Platykurtic ($\alpha_4 < 3$) and Leptokurtic ($\alpha_4 > 3$).

Shapiro Wilk normality Test is a test of normality in frequents statistics. It tests the null hypothesis that a sample $x_1, x_2 \ldots x_n$ came from a normally distributed population. The test statistic is written as

$$W = \frac{\left[\sum_{i=1}^{n} a_i x_{(i)}\right]^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2} \tag{19}$$

where $x_{(i)}$ is the ith order statistic and the constants a_i are given by $\frac{m^T V^{-1}}{(m^T V^{-1} V^{-1} m)^{1/2}}$. $m = (m_1 \ldots m_n)^T$.

Table 1
Output of the first 15 observations of the w_africans dataset

	Country	period	GDPPC_PPP	GNIPC_PPP	ER
1	BJ	1999	1621.90	1260	615.47
2	BJ	2000	1666.47	1320	710.21
3	BJ	2001	1703.02	1380	732.40
4	BJ	2002	1728.70	1410	693.71
5	BJ	2003	1734.70	1450	579.90
6	BJ	2004	1757.90	1510	527.34
7	BJ	2005	1735.97	1540	527.26
8	BJ	2006	1752.96	1600	522.43
9	BJ	2007	1805.62	1690	478.63
10	BJ	2008	1841.19	1770	446.00
11	BJ	2009	1831.88	1770	470.29
12	BJ	2010	1818.78	1770	494.79
13	BJ	2011	1820.89	1820	471.25
14	BJ	2012	1855.94	1880	510.56
15	BJ	2013	1934.62	1990	493.90

And $m_1 \ldots m_n$ are the expected values of the order statistics of independent and identically distributed random variables sampled from the standard normal distribution, and V is the covariance matrix of those order statistics. The null hypothesis may be rejected if W is too small.

3. Results and discussion

The dataset was extracted in MS-excel and was saved as a "comma delimited (**social.csv**) file". Another object was created in R for the social.csv file named ***w_africans*** as used in exporting the data into the console using the command line:
w_africans<-read.csv("social.csv",header=T)

However, the ***w_africans*** dataset was inspected for correctness before commencing the analysis using the commands stated below and the output is as given in Table 1.
#Displaying the first 15 observations of the w_africans dataset

print(head(w_africans, n=15))

The nature of the columns (variables) in the ***w_africans*** dataset was also explored, using *ls* ***(DATAVAR)*** or ***names(DATAVAR)***, where ***DATAVAR*** represent the dataframe name to be explored using the commands given below, with the subsequent results.
#Dataset variable names can be viewed using names (dataset) or ls(dataset)
ls(w_africans)
[1] "Country" "ER" "GDPPC_PPP" "GNIPC_PPP"
#Viewing the number of rows and columns in the w_africans dataset; use ncol(dataset) and nrow(dataset)
ncol(w_africans); nrow(w_africans)
[1] 5
[1] 320

From the results output, the ***w_africans*** dataset contains 4 variables and 320 rows as explained earlier

#A more advanced way to view the structure of the dataset is by using str(DATAVAR)
str(w_africans) #Data structure
data.frame': 320 obs. of 5 variables:
$ Country: Factor w/16 levels "BF","BJ","CI",..:2 2 2 2 2 2 2 2 2 2...
$ period: int 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008...
$ GDPPC_PPP: num 1622 1666 1703 1729 1735...
$ GNIPC_PPP: int 1260 1320 1380 1410 1450 1510 1540 1600 1690 1770...
$ ER: num 615 710 732 694 580...

The ***w_africans*** data.frame includes 2 numeric variables, 2 integer variables and 1 categorical variable

The Mean value of each of the variables is computed using the commands:
#Calculate the mean of variable with mean(DATAVAR$VAR): mean of GDPPC_PPP variable
mean(w_africans$GDPPC_PPP, na.rm=TRUE)
[1] 2258.119
#mean of GNIPC_PPP variable
mean(w_africans$GNIPC na.rm=TRUE)
[1] 2117.962
#mean of ER variable
mean(w_africans$ER, na.rm=TRUE)
[1] 857.6926

Here, the average GDP at purchasing power parity per capita, GNI at purchasing power parity per capita and exchange rate (ER) for the 16 West African countries between years 1999 and 2018 is about $2258.12, $2117.962 and 857.6926 per US$ respectively.

Note: The **na.rm =TRUE** command the console to remove missing value in case there is one.

For the standard deviation, the following commands subsist; and the results represent the spread of the variables.
sd(w_africans$GDPPC_PPP, na.rm=TRUE)#Standard deviation of GDPPC_PPP
[1] 1331.402
> sd(w_africans$GNIPC_PPP, na.rm=TRUE)#Standard deviation of GNIPC_PPP
[1] 1341.855

sd(w_africans$ER, na.rm=TRUE)#Standard deviation of ER
[1] 1596.375

Continuing in the same terrain for the Range computation, minimum and maximum are computed on a single variable using the min(VAR) and max(VAR) formula. Students were taught how to calculate minimums and maximums using the codes below:
#Minimum and maximum GDP of the selected w_african countries
min(w_africans$GDP, na.rm=TRUE); max(w_africans$GDP, na.rm=TRUE)
[1] 754.86
[1] 6661.99

From the output, the minimum GDP at purchasing power parity per capita is $754.86 and the maximum is about $6,661.99. This indicated a large gap in GDP per capita taking distribution among the West African countries in response to their purchasing power parity into consideration.

> #Minimum and maximum GNIPC of the selected w_african countries
> min(w_africans$GNIPC_PPP, na.rm=TRUE); max(w_africans$GNIPC_PPP, na.rm=TRUE)
[1] 600
[1] 7330

It can be inferred that the Gross National Income at PPP per capital of all West Africa is between $600 and $7330 inclusive.
#Minimum and maximum ER of the selected w_african countries
min(w_africans$ER, na.rm=TRUE); max(w_africans$ER, na.rm=TRUE)
[1] 0.27
[1] 9088.32

It is evidenced within the studied periods that Ghana's economy has not been adversely affected by external forces as shown from their cedis minimum exchange rate to the US$ while the maximum exchange rate of 9088.32 is attributed to Guinea. We can infer that African countries as a nation is still developing and may take some time to meet up with other continents currency rates.

The command "range(VAR)" is used to summarize the minimums and maximums on individual variables. These computations are demonstrated in the following codes:
#Calculate the range of a variable with range(VAR)
range(w_africans$GDPPC_PPP, na.rm=TRUE)#Range of variable GDP
range(w_africans$GDPPC_PPP, na.rm=TRUE)#Range of variable GNDPPC_PPP
[1] 754.86 6661.99
range(w_africans$GNIPC_PPP, na.rm=TRUE)#Range of variable GNIPC_PPP
[1] 600 7330
range(w_africans$ER, na.rm=TRUE)#Range of variable ER
[1] 0.27 9088.32

Students have been taught that a quartile is a value computed from a collection of numeric measurements, showing observation's rank when compared to all other present observations. Quartile can also be alternatively expressed as a percentilepercentile, as it is identical but on a scale of 0 to 100. Thus, we used **quantile()** function to obtain quartile and percentile in R, with commands
quantile(VAR, prob=c(prob value1, prob value2, . . . , prob valuei))
#Calculate the 25th, 50th, 75th percentilepercentile for GDP per capita at PPP
quantile(w_africans$GDPPC_PPP, na.rm=TRUE, prob=c(0.25, 0.50, 0.75, 0.95))
25% 50% 75% 95%
1369.780 1728.700 2851.580 5361.187

From the output, it easily observed that 25% of average GDP at PPP per capita was $136.780 with median (50th percentilepercentile) of $1728.700; 75% was about $2851.580 and 95% of the African countries have about $5361.187. This may explain the wide gap in GDP growth of each West African countries since GDP per capita is correlated with GNI per capita.

#Calculate the 25th, 50th, 75th percentilepercentile for **GNIPC_PPP**
quantile(w_africans$GNIPC, na.rm=TRUE, prob=c(0.25, 0.50, 0.75, 0.95))
25% 50% 75% 95%
1195 1680 2625 5435
#Calculate the 25th, 50th, 75th percentilepercentile for **ER**
quantile(w_africans$ER, na.rm=TRUE, prob=c(0.25, 0.50, 0.75, 0.95))
25% 50% 75% 95%
83.060 494.040 591.740 4528.037

Students were also taught how to use ***summary(x)*** function, where *x* can be any number of objects, includ-

Table 2
Pooled descriptive statistics

Statistic	GDP per capita, PPP ($)	GNI per capita, PPP ($)	ER
Mean	2258.119	2117.965	857.693
Standard Deviation	1331.402	1341.855	1596.375
25th Percentile (Q_1)	1369.780	1195	83.060
50th Percentile (Q_2)	1728.700	1680	494.040
75th Percentile (Q_3)	2851.580	2625	591.740
95th Percentile	5361.187	5435	4528.037
Minimum	754.860	600	0.27
Maximum	6661.990	7330	9088.32

Source: Extracted from R-console output.

Table 3
Variables normality test

Moments	GDP per capita, PPP ($)	GNI per capita, PPP ($)	ER
Skewness	1.353	1.518	3.283
Kurtosis	4.227	4.940	13.810
Shapiro Wilk Test Statistics	0.848	0.840	0.502
P-value	0.000	0.000	0.0000

Source: Extracted from R-console output.

ing datasets, variables, and linear models to generate the descriptive statistics of the variables in the dataset. The code is written below for the ***w_africans*** dataset with the subsequent results presented below it.
> #Summarize the w_africans dataset using the command summary(x)
> print(summary(w_africans))

Country	period	GDPPC_PPP	GNIPC_PPP	ER
BF:20	Min.:1999	Min.:754.9	Min.:600	Min.:0.27
BJ:20	1st Qu.:2004	1st Qu.:1369.8	1st Qu.:1195	1st Qu.:83.06
CI:20	Median:2008	Median:1728.7	Median:1680	Median:494.04
CV:20	Mean:2008	Mean:2258.1	Mean:2118	Mean:857.69
GH:20	3rd Qu.:2013	3rd Qu.:2851.6	3rd Qu.:2625	3rd Qu.:591.74
GM:20	Max.:2018	Max.:6662.0	Max.:7330	Max.:9088.32
(Other):200		NA's:1	NA's:1	

The summary outputs provides the descriptive statistics of all objects in the sample dataset and is explicitly presented in Table 2. Further exploration was carried out on the data by checking their respective distributions through Skewness, kurtosis and further test such as the Shapiro wilk test of normality. These were done using the "**moments**" library in R. Students were taught how to load packages from R as library(). Details are as given below while the summary presented in Table 3:
library(moments)

skewness(w_africans$GDPPC_PPP, na.rm=T) #Skewness coefficient of GDP per capita at PPP
[1] 1.353004

skewness(w_africans$GNIPC_PPP, na.rm=T) #Skewness coefficient of GNIPC at PPP
[1] 1.517567

skewness(w_africans$ER, na.rm=T) #Skewness coefficient of ER
[1] 3.283139

kurtosis(w_africans$GDPPC_PPP, na.rm=T) #Kurtosis coefficient of GDP per capita at PPP
[1] 4.226773

kurtosis(w_africans$GNIPC_PPP, na.rm=T) #Kurtosis coefficient of GNIPC at PPP
[1] 4.940481

kurtosis(w_africans$ER, na.rm=T) #Kurtosis coefficient of ER
[1] 13.80796

shapiro.test(w_africans$GDP)#GDP test of Normality
Shapiro-Wilk normality test

data: w_africans$GDPPC_PPP
W = 0.84758, p-value $<$ 2.2e-16

shapiro.test(w_africans$GNIPC)#GNIPC test of Normality
Shapiro-Wilk normality test
data: w_africans$GNIPC_PPP
W = 0.83966, p-value $<$ 2.2e-16

shapiro.test(w_africans$ER)#ER test of Normality
Shapiro-Wilk normality test
data: w_africans$ER
W = 0.5022, p-value $<$ 2.2e-16

Positive coefficients of 1.353, 1.518, and 3.283 indicated that the econometric variables of GDP, GNIPC and ER is highly skewed to the right and may not be normally distributed. As the Kurtosis measure the fourth moments, selected West Africans exchange rate was found to be normally distributed (kurtosis $\approx$ 3) with other kurtosis of other variables $>$ 3, indicating a leptokurtic shape compared to a normal distribution. However, normality test of the data confirmed the non-normality of the data since its associated p-values are lower than 5% level of significance.

Quantile plots visualize the distribution of the data per variable and details generated by the below commands are as given in Figs 2–4 respectively

Table 4
Cross-section data description on average

S/n	Country	CODE	Mean GDP per capita PPP	Mean GNIPC PPP	Mean ER
1	Benin	BJ	1841.461 [141.8469]	1759.500 [330.621]	554.3915 [82.72886]
2	Burkina Faso	BF	1386.965 [213.250]	1320.000 [333.024]	555.261 82.9732
3	Cape Verde	CV	5355.335 [1009.555]	5039.500 [1414.874]	93.1725 [13.51637]
4	Cote D'Ivoire	CI	2913.830 [338.916]	2647.500 [614.524]	555.261 [82.9732]
5	Gambia	GM	1460.178 [40.394]	1349.500 [184.033]	29.855 [10.73796]
6	Ghana	GH	3031.057 [696.137]	2897.500 [969.063]	1.772 [1.346909]
7	Guinea	GN	1735.404 [226.013]	1593.500 [399.569]	5075.988 [2625.065]
8	Guinea Bissau	GW	1430.202 [72.702]	1360.500 [239.109]	555.261 [82.9732]
9	Liberia	LR	1137.824 [136.222]	970.526 [190.860]	71.5625 [24.41597]
10	Mali	ML	1794.605 [151.470]	1670.500 [321.943]	555.261 [82.9732]
11	Mauritania	MR	3348.436 [370.771]	3193.000 [627.259]	28.2255 [4.042962]
12	Niger	NE	823.119 [61.436]	779.500 [138.049]	555.261 [82.973]
13	Nigeria	NG	4565.789 [907.056]	4237.500 [1296.651]	158.823 [61.094]
14	Senegal	SN	2758.823 [263.357]	2614.500 [524.740]	555.261 [82.973]
15	Sierra Leone	SL	1204.011 [243.925]	1158.500 [342.856]	3822.465 [1756.515]
16	Togo	TG	1286.844 [226.013]	1238.500 [399.569]	555.261 [82.9732]

Values in parentheses [ ] represent standard deviation. Source: Extracted from R-console output.

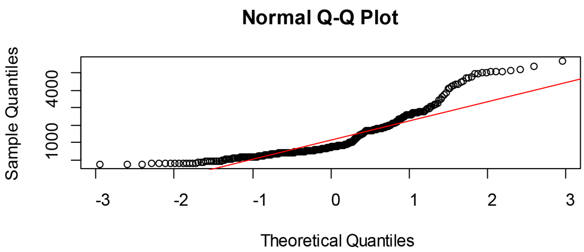

Fig. 2. Normal Q-Q plots of GDP at PPP per capita of some selected West African countries.

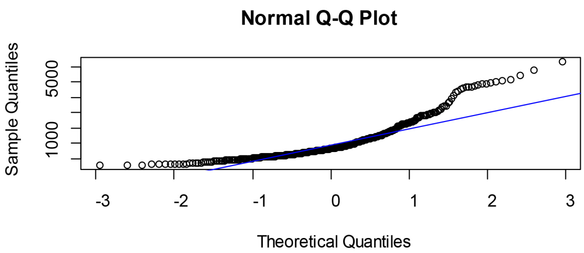

Fig. 3. Normal Q-Q plots of GNI at PPP per capita of some selected West African countries.

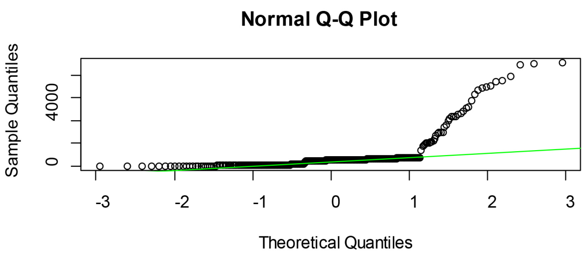

Fig. 4. Normal Q-Q plots of ER of some selected West African countries.

```
par(mfrow=c(2,2)) #Partitioning of plots space
#Quantile plot of GDP per capita at PPP rates
qqnorm(w_africans$GDPPC_PPP);qqline(w_africans$
GDPPC_PPP,col="red")
#Quantile plot of GNI per capita at PPP rates
qqnorm(w_africans$GNIPC_PPP);qqline(w_africans$
GNIPC_PPP,col="black")
#Quantile plot of Exchange rate
qqnorm(w_africans$ER);qqline(w_africans$ER,col=
"green")#Quantile plot of ER
```

The Figs 2–4 showed that the quantile plots of the selected variables do not lie on the theoretical normal line. Thus, the variables are not precisely normal but may not be too far off.

Students were also introduced to data splitting in R using dataframe_name[n:m,]. This method was used due to the fact that the data structure was paneled in nature with the first 20 observations on row-wise which represents republic of Benin followed by Burkina Faso, among others. The command line used is given below with the results output presented in Table 4.

```
benin_d<-w_africans[1:20,];benin_d #Extracted Benin republic variables from the panel structured data.
```

The data was further explored using **ExPanDaR** package in R. Average GDP, GNIPC and ER per cross sections (countries) were visualized from the Shiny app using simple bar chart presented in Figs 3–5 respectively.

```
library(ExPanDaR)
ExPanD(df=w_africans)
```

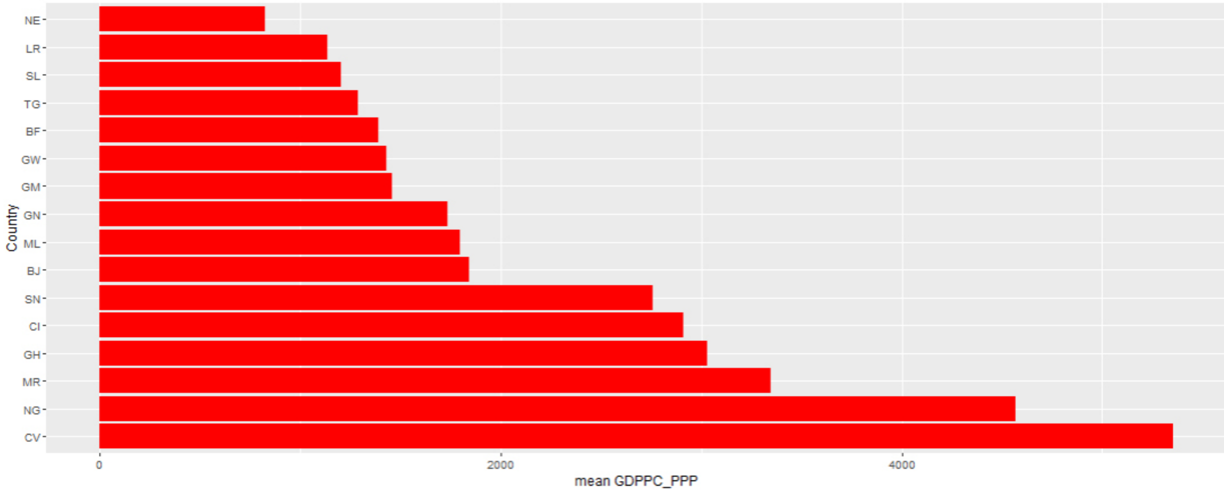

Fig. 5. Bar chart of average GDP per capita based on PPP rates of selected West African countries.

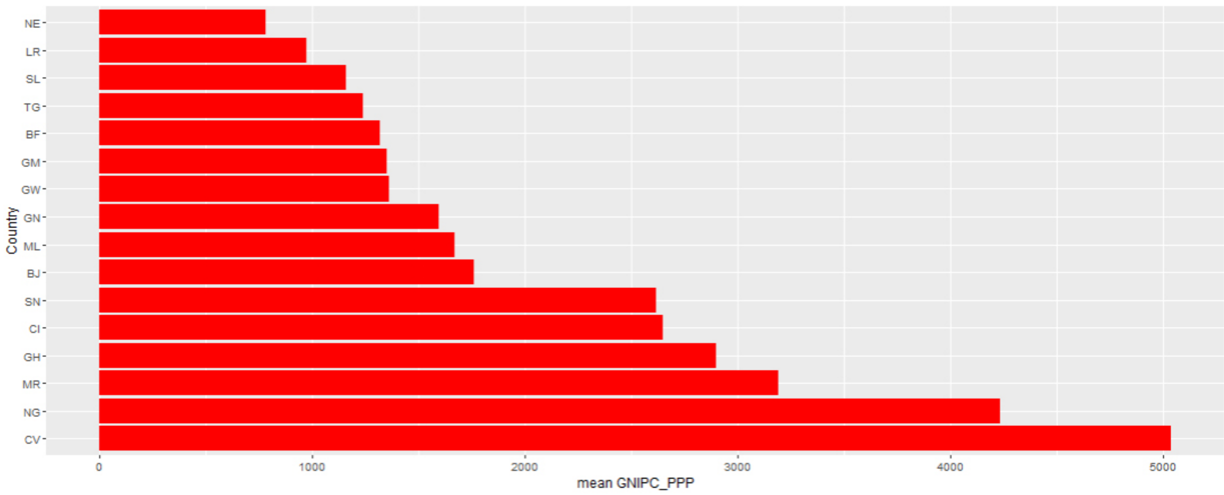

Fig. 6. Bar chart of average GNIPC based on PPP rates of selected West African countries.

The Figs 5–7 showed that Cape Verde (CV) recorded the highest average GDP (per capita) and GNI (per capita) taking into consideration purchasing power parity among the West African countries followed by Nigeria (NG). Cape Verde (CV) also has the highest average GNIPC at purchasing power parity rates and Ghana (GH) possess the strongest currency rate among other west African nations taking the US$ exchange rate into consideration. Niger (NE) recorded the lowest average GDP per capita and GNIPC at PPP and Guinea (GN) with the weakest currency rate within the selected timeframe. This can also be evidenced from Table 4 with an associated variability from the mean.

3.1. *Summary of findings*

This paper presented students learning experience on the introduction of data science skills for curriculum delivery in Africa using social-economic data extracted from UNESCO website. The interactive session helped students on how to use R software for analyzing for descriptive statistics, and appropriate interpretation of results based on the type of data used for analysis. This bridged the gap between the traditional method of data analysis and the conventional form especially in the area of big data. Findings from the analysis showed that economic growth varies from countries to countries as shown from the pictorial representation of data and respective spread of observation from the mean. However,

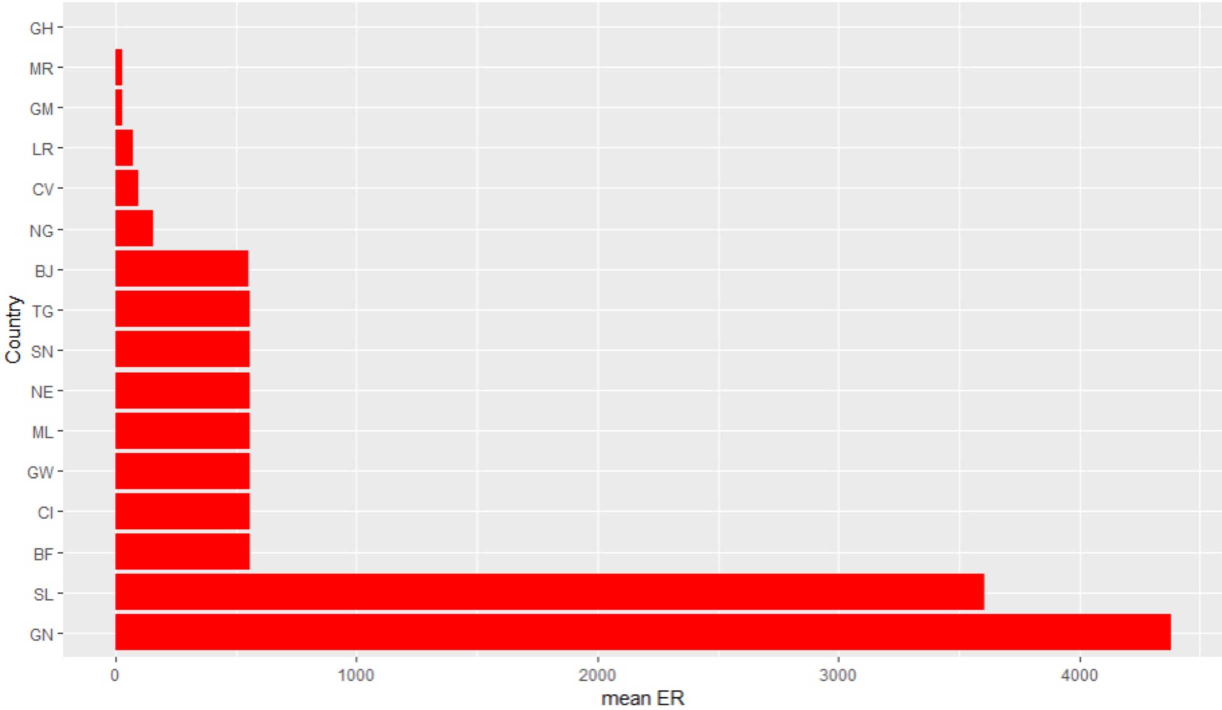

Fig. 7. Bar chart of average ER of selected West African countries.

this result is an indication that Cape Verde (CV) among other West African countries is better off in terms of their economic growth taking purchasing power parity into consideration. This indicated that Nigeria economic growth may be marred by inflation, resulting to the devaluation of her naira note in the international market, among other developing countries. Hence, West African countries in general are far from being developed compared to countries in Asia, America, and Europe to mention a few.

4. Conclusion

Introducing beginner students in statistics to data science is a vexatious task, especially in African countries where regular supply of power is a luxury and uninterrupted internet facilities are quite expensive and almost impossible. The developing nature of most Africa countries has created a paradoxical approach to achieving reasonable success in students' learning of data science. However, for the purpose of this research, great achievement was made in introducing the students to data description using R software for data science, thereby equipping them with a career in data analysis. From the beginning, students offering introductory statistics gain reasonable experience of what constitutes both the practical and conceptual aspects of the working life of a data scientist, as they were able to run simple codes on exploratory data analysis using the focused data. The students equally enhanced their knowledge in deducing reasonable inference from the output of data analysis. 200 level students were able to run with ease, R codes to estimate basic descriptive statistics within a 1 hour lecture period. The activities was carried out without much supervision on the part of the tutor. Comparison was made per member countries on their developmental rate taking their respective Gross Domestic Product, Gross National Income per capita, and Exchange Rate into consideration.

It is of the opinion that topics covered in data science courses can and should be brought into a variety of statistics courses at undergraduate level, while adequate facilities provided for its teaching and learning. Thus, key data science skills need to be introduced, reiterated, and reinforced throughout the undergraduate statistics curriculum.

Though, the exercise is not without its own challenges, but its prospects in creating self-driven learning culture among students of tertiary institutions has greatly enhance the quality of teaching, advancing students skills in machine learning, improved understanding of the role of data in global perspective and on the spot ability of the students to be able to critique claims based on data.

Acknowledgments

The authors are grateful to Federal Polytechnic Ilaro and the students of Mathematics & Statistics department for creating the enabling environments suitable for the data science activities carried out in this research.

References

[1] Jordan, M.I. and Mitchell, T.M. Machine learning: trends, perspectives, and prospects. *Science*, 2015, 349(6245), 255–260.
[2] Mayer-Schönberger, V. and Cukier, K. *Big Data: A Revolution That Will Transform How We Live, Work, and Think*, 2013. New York: Houghton Mifflin Harcourt.
[3] Provost, F. and Fawcett, T. Data science and its relationship to big data and data-driven decision making. *Big Data*, 2013, 1(1), 51–59.
[4] Kuhn, T.S. *The Structure of Scientific Revolutions*. 3rd ed. Chicago, IL: University of Chicago Press, 1996.
[5] Box, G.E.P. Science and statistics. *Journal of the American Statistical Association*, 2012, 71(356), 791–799. Reprint of original from 1962.
[6] Nolan, D. and Temple Lang, D. Computing in the statistics curricula. *The American Statistician*, 2010, 64, 97–107. doi: 10.1198/tast.2010.09132.
[7] Hardin, J., Hoerl, R., Horton, N.J. and Nolan, D. Data Science in Statistics Curricula: Preparing Students to "Think with Data". *The American Statistician*, 2014. doi: 10.1080/00031305.2015.1077729.
[8] Tukey, J.W. The future of data analysis. *Annals of Mathematical Statistics*, 1962, 33(1), 1–67. Moore, D.S., McCabe, G.P. and Craig, B.A. (2012). *Introduction to the Practice of Statistics*. New York: WH Freeman.
[9] National Science Foundation. Accelerating discovery in science and engineering through Petascale simulations and analysis (*PetaApps*), 2008. Posted July 28, 2008.
[10] Gershman, S.J., Horvitz, E.J. and Tenenbaum, J.B. Computational rationality: a converging paradigm for intelligence in brains, minds, and machines. *Science*, 2015, 349(6245), 273–278.
[11] Horton, N.J. and Hardin, J.S. Teaching the next generation of statistics students to 'think with data': special issue on statistics and the undergraduate curriculum, *The American Statistician*, 2015, 69, 259–265. doi: 10.1080/00031305.2015.1094283.
[12] Hoerl, J., Horton, J., Nolan, N.J., Baumer, D., Hall-Holt, D. and Ward, M.D. Data science in statistics curricula: preparing students to 'think with data', *The American Statistician*, 2015, 69, 343–353. doi: 10.1080/00031305.2015.1077729.

Appendix 1: Data

GDP per capita PPP, GNI per capita PPP, and Exchange Rate of selected 16 west African countries.

Country	Period	GDP per capita, PPP (2011 international $)	GNI per capita, PPP ($)	Exchange rate
BJ	1999	1621.9	1260	615.47
BJ	2000	1666.47	1320	710.21
BJ	2001	1703.02	1380	732.4
BJ	2002	1728.7	1410	693.71
BJ	2003	1734.7	1450	579.9
BJ	2004	1757.9	1510	527.34
BJ	2005	1735.97	1540	527.26
BJ	2006	1752.96	1600	522.43
BJ	2007	1805.62	1690	478.63
BJ	2008	1841.19	1770	446
BJ	2009	1831.88	1770	470.29
BJ	2010	1818.78	1770	494.79
BJ	2011	1820.89	1820	471.25
BJ	2012	1855.94	1880	510.56
BJ	2013	1934.62	1990	493.9
BJ	2014	2001.05	2100	493.76
BJ	2015	1987.14	2110	591.21
BJ	2016	2009.66	2160	592.61
BJ	2017	2069.29	2260	580.66
BJ	2018	2151.54	2400	555.45
BF	1999	1086.62	840	615.7
BF	2000	1075.4	850	711.98
BF	2001	1114.2	900	733.04
BF	2002	1129.74	930	696.99
BF	2003	1183.09	990	581.2
BF	2004	1200.42	1030	528.28
BF	2005	1266.36	1120	527.47
BF	2006	1305.92	1200	522.89
BF	2007	1338.84	1260	479.27
BF	2008	1393.7	1340	447.81
BF	2009	1392.2	1340	472.19
BF	2010	1423.38	1360	495.28
BF	2011	1472.72	1420	471.87
BF	2012	1521.45	1520	510.53
BF	2013	1562.3	1590	494.04
BF	2014	1582.33	1620	494.41
BF	2015	1596.33	1650	591.45
BF	2016	1642.48	1710	593.01
BF	2017	1696.23	1810	582.09
BF	2018	1755.59	1920	555.72
CV	1999	3472.6	2660	102.7
CV	2000	3896.96	3020	115.88
CV	2001	3915.16	3150	123.21
CV	2002	4053.37	3270	117.26
CV	2003	4157.15	3440	97.79
CV	2004	4513.97	3820	88.75
CV	2005	4759.13	4090	88.65
CV	2006	5071.86	4470	87.93
CV	2007	5768.87	5320	80.62
CV	2008	6078.55	5690	75.34
CV	2009	5929.44	5600	80.04
CV	2010	5943.35	5570	83.28
CV	2011	6102.41	5860	79.28
CV	2012	6090.55	5940	86.32
CV	2013	6061.31	6070	83.07
CV	2014	6021.63	6050	83.03
CV	2015	6007.22	6180	99.39
CV	2016	6214.08	6470	99.69
CV	2017	6387.1	6790	97.81

Country	Period	GDP per capita, PPP (2011 international $)	GNI per capita, PPP ($)	Exchange rate
CV	2018	6661.99	7330	93.41
GM	1999	1416.72	1060	11.4
GM	2000	1448.62	1110	12.79
GM	2001	1484.89	1150	15.69
GM	2002	1391.43	1080	19.92
GM	2003	1440.18	1160	28.53
GM	2004	1493.71	1240	30.03
GM	2005	1434.39	1230	28.58
GM	2006	1407.03	1240	28.07
GM	2007	1415.08	1290	24.87
GM	2008	1452.45	1360	22.19
GM	2009	1500.82	1410	26.64
GM	2010	1551.59	1470	28.01
GM	2011	1440.79	1390	29.46
GM	2012	1476.06	1460	32.08
GM	2013	1500.51	1520	35.96
GM	2014	1442.1	1490	41.73
GM	2015	1481.48	1540	42.51
GM	2016	1443.69	1530	43.88
GM	2017	1465.34	1580	46.61
GM	2018	1516.69	1680	48.15
GH	1999	2193.1	1670	0.27
GH	2000	2219.21	1710	0.54
GH	2001	2252.13	1790	0.72
GH	2002	2296.58	1860	0.79
GH	2003	2357.33	1940	0.87
GH	2004	2428.26	2050	0.9
GH	2005	2507.59	2210	0.91
GH	2006	2600.79	2370	0.92
GH	2007	2644.72	2480	0.94
GH	2008	2813.21	2690	1.06
GH	2009	2875.42	2770	1.41
GH	2010	3026.36	2920	1.43
GH	2011	3368.8	3260	1.51
GH	2012	3595.64	3480	1.8
GH	2013	3769.94	3830	1.95
GH	2014	3791.28	3880	2.9
GH	2015	3786.96	3990	3.67
GH	2016	3830.5	4060	3.91
GH	2017	4051.46	4340	4.35
GH	2018	4211.85	4650	4.59
GN	1999	1515.65	1150	1387.4
GN	2000	1518.52	1180	1746.87
GN	2001	1541.09	1210	1950.56
GN	2002	1588.79	1300	1975.84
GN	2003	1577.93	1230	1984.93
GN	2004	1583.62	1270	2243.93
GN	2005	1598.17	1290	3644.33
GN	2006	1582.66	1360	5148.75
GN	2007	1653.28	1470	4197.75
GN	2008	1682.66	1500	4601.69
GN	2009	1626.17	1450	4801.08
GN	2010	1666.49	1530	5726.07
GN	2011	1721.45	1600	6658.03
GN	2012	1783.67	1710	6985.83
GN	2013	1812.88	1780	6907.88
GN	2014	1836.56	1880	7014.12
GN	2015	1859.74	1930	7485.52
GN	2016	2007.34	2130	8959.72

Country	Period	GDP per capita, PPP (2011 international $)	GNI per capita, PPP ($)	Exchange rate
GN	2017	2213.46	2420	9088.32
GN	2018	2337.95	2480	9011.13
GW	1999	1365.77	1000	615.7
GW	2000	1410.92	1090	711.98
GW	2001	1411.49	1100	733.04
GW	2002	1367.12	1110	696.99
GW	2003	1343.98	1100	581.2
GW	2004	1349.35	1140	528.28
GW	2005	1374.03	1200	527.47
GW	2006	1372.44	1250	522.89
GW	2007	1383.12	1300	479.27
GW	2008	1392.52	1320	447.81
GW	2009	1403.55	1340	472.19
GW	2010	1430.97	1400	495.28
GW	2011	1506.7	1520	471.87
GW	2012	1442.15	1480	510.53
GW	2013	1450	1470	494.04
GW	2014	1425.77	1560	494.41
GW	2015	1474.24	1610	591.45
GW	2016	1526.81	1690	593.01
GW	2017	1576.75	1740	582.09
GW	2018	1596.36	1790	555.72
CI	1999	3132.64	2310	615.7
CI	2000	2989.15	2160	711.98
CI	2001	2922.03	2100	733.04
CI	2002	2810.19	2030	696.99
CI	2003	2714.01	1940	581.2
CI	2004	2690.74	2070	528.28
CI	2005	2679.79	2300	527.47
CI	2006	2662.33	2350	522.89
CI	2007	2650.49	2400	479.27
CI	2008	2657.67	2460	447.81
CI	2009	2682.04	2500	472.19
CI	2010	2673.01	2520	495.28
CI	2011	2495.5	2400	471.87
CI	2012	2696.19	2660	510.53
CI	2013	2864.05	2840	494.04
CI	2014	3038.84	3130	494.41
CI	2015	3225.19	3340	591.45
CI	2016	3395.09	3650	593.01
CI	2017	3564.6	3760	582.09
CI	2018	3733.05	4030	555.72
LR	1999			41.9
LR	2000	1317.87	930	40.9
LR	2001	1307.93	880	48.59
LR	2002	1325.38	900	61.75
LR	2003	910.1	610	59.38
LR	2004	916.49	650	54.91
LR	2005	940.16	700	57.1
LR	2006	981.89	780	58.01
LR	2007	1034.29	870	61.27
LR	2008	1063.37	930	63.21
LR	2009	1076.11	960	68.29
LR	2010	1101.48	980	71.4
LR	2011	1154.41	1090	72.23
LR	2012	1211.05	1120	73.51
LR	2013	1281.55	1200	77.52
LR	2014	1257.63	1190	83.89
LR	2015	1225.93	1190	86.19

Country	Period	GDP per capita, PPP (2011 international $)	GNI per capita, PPP ($)	Exchange rate
LR	2016	1176.19	1160	94.43
LR	2017	1175.64	1170	112.71
LR	2018	1161.18	1130	144.06
ML	1999	1508.48	1160	615.7
ML	2000	1465.76	1150	711.98
ML	2001	1642.35	1270	733.04
ML	2002	1643.04	1270	696.99
ML	2003	1738.13	1410	581.2
ML	2004	1710.11	1430	528.28
ML	2005	1763.9	1520	527.47
ML	2006	1786.31	1580	522.89
ML	2007	1788.03	1640	479.27
ML	2008	1812.05	1700	447.81
ML	2009	1835.97	1740	472.19
ML	2010	1875.19	1760	495.28
ML	2011	1877.89	1810	471.87
ML	2012	1808.01	1770	510.53
ML	2013	1796.77	1800	494.04
ML	2014	1868.31	1920	494.41
ML	2015	1922.43	2010	591.45
ML	2016	1974.31	2070	593.01
ML	2017	2019.44	2170	582.09
ML	2018	2055.62	2230	555.72
MR	1999	2922.44	2320	20.95
MR	2000	2833.93	2280	23.89
MR	2001	2813.65	2230	25.56
MR	2002	2755.18	2370	27.17
MR	2003	2839.11	2480	26.3
MR	2004	2918.42	2610	26.43
MR	2005	3090.86	2840	26.55
MR	2006	3570.52	3200	26.86
MR	2007	3567.26	3300	25.86
MR	2008	3503.27	3350	23.82
MR	2009	3367.49	3310	26.24
MR	2010	3426.47	3300	27.59
MR	2011	3483.52	3380	28.11
MR	2012	3578.1	3510	29.66
MR	2013	3685.7	3690	30.07
MR	2014	3779.09	3810	30.27
MR	2015	3722.7	3830	32.47
MR	2016	3690.24	3890	35.24
MR	2017	3696.35	4000	35.79
MR	2018	3724.41	4160	35.68
NE	1999	793.78	610	615.7
NE	2000	754.86	600	711.98
NE	2001	779.6	630	733.04
NE	2002	774.09	630	696.99
NE	2003	785.6	650	581.2
NE	2004	757.75	650	528.28
NE	2005	762.87	680	527.47
NE	2006	777.48	710	522.89
NE	2007	772.37	730	479.27
NE	2008	815.04	780	447.81
NE	2009	778.98	750	472.19
NE	2010	812.3	790	495.28
NE	2011	799.26	790	471.87
NE	2012	859.79	860	510.53
NE	2013	870.4	880	494.04
NE	2014	900.14	930	494.41
NE	2015	903.42	940	591.45
NE	2016	912.03	960	593.01
NE	2017	920.63	990	582.09
NE	2018	931.99	1030	555.72
NG	1999	2996.94	2270	92.34
NG	2000	3069.44	2230	101.7
NG	2001	3170.44	2440	111.23
NG	2002	3565.39	2760	120.58
NG	2003	3731.46	2910	129.22
NG	2004	3973.62	3190	132.89
NG	2005	4121.5	3390	131.27
NG	2006	4258.59	3830	128.65
NG	2007	4421.36	3990	125.81
NG	2008	4597	4220	118.55
NG	2009	4835.95	4450	148.9
NG	2010	5085.41	4710	150.3
NG	2011	5213.84	4920	153.86
NG	2012	5290.63	5130	157.5
NG	2013	5494.52	5420	157.31
NG	2014	5687.59	5810	158.55
NG	2015	5685.93	5910	192.44
NG	2016	5448.91	5760	253.49
NG	2017	5351.44	5710	305.79
NG	2018	5315.82	5700	306.08
SN	1999	2398.95	1840	615.7
SN	2000	2417.83	1890	711.98
SN	2001	2468.53	1980	733.04
SN	2002	2424.87	1970	696.99
SN	2003	2523.67	2100	581.2
SN	2004	2605.44	2230	528.28
SN	2005	2682.44	2370	527.47
SN	2006	2677.93	2450	522.89
SN	2007	2736.88	2570	479.27
SN	2008	2772.55	2660	447.81
SN	2009	2754.75	2640	472.19
SN	2010	2775.7	2690	495.28
SN	2011	2739.34	2700	471.87
SN	2012	2800.41	2810	510.53
SN	2013	2799.96	2850	494.04
SN	2014	2902.51	3010	494.41
SN	2015	3001.82	3140	591.45
SN	2016	3104.24	3260	593.01
SN	2017	3232.31	3460	582.09
SN	2018	3356.34	3670	555.72
SL	1999	875.35	660	1804.2
SL	2000	908.71	700	2092.13
SL	2001	820.7	650	1986.15
SL	2002	993.28	800	2099.03
SL	2003	1036.66	860	2347.94
SL	2004	1057.69	890	2701.3
SL	2005	1063.91	930	2889.59
SL	2006	1073.92	970	2961.91
SL	2007	1129.38	1120	2985.19
SL	2008	1162.41	1200	2981.51
SL	2009	1172.86	1230	3385.65
SL	2010	1208.05	1200	3978.09
SL	2011	1255.45	1240	4349.16
SL	2012	1413.88	1490	4344.04
SL	2013	1669.13	1720	4332.5

Country	Period	GDP per capita, PPP (2011 international $)	GNI per capita, PPP ($)	Exchange rate
SL	2014	1707.1	1760	4524.16
SL	2015	1326.21	1400	5080.75
SL	2016	1376.4	1330	6289.94
SL	2017	1403.79	1500	7384.43
SL	2018	1425.34	1520	7931.63
TG	1999	1282.72	970	615.7
TG	2000	1235.46	960	711.98
TG	2001	1182.2	940	733.04
TG	2002	1140.99	930	696.99
TG	2003	1167.5	970	581.2
TG	2004	1162.34	990	528.28
TG	2005	1145.91	1010	527.47
TG	2006	1161.06	1050	522.89
TG	2007	1156.06	1080	479.27
TG	2008	1170.78	1120	447.81
TG	2009	1202.52	1160	472.19
TG	2010	1241.92	1210	495.28
TG	2011	1286.47	1360	471.87
TG	2012	1334.66	1360	510.53
TG	2013	1379.4	1440	494.04
TG	2014	1423.55	1520	494.41
TG	2015	1467.25	1620	591.45
TG	2016	1501.12	1640	593.01
TG	2017	1529.52	1680	582.09
TG	2018	1565.46	1760	555.72

Source: Extracted from UIS.stat report (uis.unesco.org).