CONTINENTAL GDP AND LIFE EXPECTANCY ANALYSIS

By: Kikonyogo Steven

ABOUT THE DATA

This is an excerpt of data from gapminder package on life expectancy, GDP-percapita, Continent and population by Country by five year-interval group from 1952 to 2007.

AIM OF THE PROJECT

The aim is to explore, analyze and draw meaningful insights from the data through;

– Data transformation

– Data wrangling

– Data manipulation

– Data Visualization

– Correlation and Regression

– Hypothesis testing against a claim

(i) Loading required packages

```
if(!require(pacman)) install.packages("pacman")
```

```
## Loading required package: pacman
```

```
pacman::p_load(
               gt,
               car,
               dplyr,
               ggplot2,
               magrittr,
               janitor,
               flextable
               )
```

(ii) A glimpse about the data

```
df <- gapminder::gapminder
      dim(df)
```

```
## [1] 1704    6
```

```
      head(df, 5)
```

```
## # A tibble: 5 x 6
##   country     continent  year lifeExp      pop gdpPercap
##   <fct>       <fct>     <int>   <dbl>    <int>     <dbl>
## 1 Afghanistan Asia       1952    28.8  8425333      779.
## 2 Afghanistan Asia       1957    30.3  9240934      821.
## 3 Afghanistan Asia       1962    32.0 10267083      853.
## 4 Afghanistan Asia       1967    34.0 11537966      836.
## 5 Afghanistan Asia       1972    36.1 13079460      740.
```

```
      summary(df)
```

```
##         country        continent        year         lifeExp
##  Afghanistan:  12   Africa  :624   Min.   :1952   Min.   :23.60
##  Albania    :  12   Americas:300   1st Qu.:1966   1st Qu.:48.20
##  Algeria    :  12   Asia    :396   Median :1980   Median :60.71
##  Angola     :  12   Europe  :360   Mean   :1980   Mean   :59.47
##  Argentina  :  12   Oceania : 24   3rd Qu.:1993   3rd Qu.:70.85
##  Australia  :  12                  Max.   :2007   Max.   :82.60
##  (Other)    :1632
##       pop              gdpPercap
##  Min.   :6.001e+04   Min.   :   241.2
##  1st Qu.:2.794e+06   1st Qu.:  1202.1
##  Median :7.024e+06   Median :  3531.8
##  Mean   :2.960e+07   Mean   :  7215.3
##  3rd Qu.:1.959e+07   3rd Qu.:  9325.5
##  Max.   :1.319e+09   Max.   :113523.1
##
```

```
      str(df)
```

```
## tibble [1,704 x 6] (S3: tbl_df/tbl/data.frame)
##  $ country  : Factor w/ 142 levels "Afghanistan",..: 1 1 1 1 1 1 1 1 1 1 ...
##  $ continent: Factor w/ 5 levels "Africa","Americas",..: 3 3 3 3 3 3 3 3 3 3 ...
##  $ year     : int [1:1704] 1952 1957 1962 1967 1972 1977 1982 1987 1992 1997 ...
##  $ lifeExp  : num [1:1704] 28.8 30.3 32 34 36.1 ...
##  $ pop      : int [1:1704] 8425333 9240934 10267083 11537966 13079460 14880372 12881816 13867957 163
##  $ gdpPercap: num [1:1704] 779 821 853 836 740 ...
```

```
      unique(df$continent)
```

```
## [1] Asia     Europe   Africa   Americas Oceania
## Levels: Africa Americas Asia Europe Oceania
```

```
      unique(df$year)
```

```
##  [1] 1952 1957 1962 1967 1972 1977 1982 1987 1992 1997 2002 2007
```

```
      any(is.na(df))
```

```
## [1] FALSE
```

- *The are 1704 observations and 6 variables, 5 continents, 142 countries.The min GDP-percapita is $242 and max $113523.The min life expectancy is 23.6yrs and max 82yrs.*

(iii) Wrangling, Transformation and Manipulation

```
df1 <- df %>%
    mutate(gdp_status=case_when(
    gdpPercap < 1000 ~"lower income",
    gdpPercap >=1000 & gdpPercap < 4000 ~ "lower middle income",
    gdpPercap >= 4000 & gdpPercap < 12000 ~ "upper middle income",
    TRUE ~ "high income"))
    tabyl(df1$gdp_status)
```

```
##       df1$gdp_status   n   percent
##          high income 324 0.1901408
##         lower income 351 0.2059859
##  lower middle income 556 0.3262911
##  upper middle income 473 0.2775822
```

```
  high_life <- df1 %>%
  select(country,
         year,
         lifeExp,
         continent) %>%
  rename(High_life_Exp2007 = lifeExp) %>%
  filter(year ==2007) %>%
  arrange(desc(High_life_Exp2007))

  gt(head(high_life,10))
```

country	year	High_life_Exp2007	continent
Japan	2007	82.603	Asia
Hong Kong, China	2007	82.208	Asia
Iceland	2007	81.757	Europe
Switzerland	2007	81.701	Europe
Australia	2007	81.235	Oceania
Spain	2007	80.941	Europe
Sweden	2007	80.884	Europe
Israel	2007	80.745	Asia
France	2007	80.657	Europe
Canada	2007	80.653	Americas

```
   low_gdp <-  df1 %>%
  select(country,
         year,
```

```
        gdpPercap,
        continent) %>%
  rename(low_gdp_2002 = gdpPercap) %>%
  filter(year ==2002) %>%
  arrange(low_gdp_2002)
  gt(head(low_gdp,10))
```

country	year	low_gdp_2002	continent
Congo, Dem. Rep.	2002	241.1659	Africa
Burundi	2002	446.4035	Africa
Ethiopia	2002	530.0535	Africa
Liberia	2002	531.4824	Africa
Guinea-Bissau	2002	575.7047	Africa
Niger	2002	601.0745	Africa
Myanmar	2002	611.0000	Asia
Mozambique	2002	633.6179	Africa
Gambia	2002	660.5856	Africa
Malawi	2002	665.4231	Africa

```
df1%>%
  select(country,
         year,
         continent,
         gdpPercap) %>%
  filter(year ==2002 & continent =='Africa' & gdpPercap >5000) %>%
  arrange(desc(gdpPercap))
```

```
## # A tibble: 9 x 4
##   country            year continent gdpPercap
##   <fct>             <int> <fct>         <dbl>
## 1 Gabon              2002 Africa       12522.
## 2 Botswana           2002 Africa       11004.
## 3 Libya              2002 Africa        9535.
## 4 Mauritius          2002 Africa        9022.
## 5 South Africa       2002 Africa        7711.
## 6 Equatorial Guinea  2002 Africa        7703.
## 7 Reunion            2002 Africa        6316.
## 8 Tunisia            2002 Africa        5723.
## 9 Algeria            2002 Africa        5288.
```

```
df2 <- df %>%
  select(continent,
         year,
         lifeExp) %>%
  filter(year==2002) %>%
  group_by(continent) %>%
  summarise(meanlife_exp=mean(lifeExp)) %>%
  arrange(desc(meanlife_exp))
  flextable(print(df2,10))
```

```
## # A
```

```
## #   tibble:
## #   5 x 2
## # i 2
## #   more
## #   variables:
## #   continent <fct>,
## #   meanlife_exp <dbl>
```

continent	meanlife_exp
Oceania	79.74000
Europe	76.70060
Americas	72.42204
Asia	69.23388
Africa	53.32523

```
  dF2 <- df %>%
  select(continent,
         year,
         gdpPercap) %>%
  filter(year==2002) %>%
  group_by(continent) %>%
  summarise(meanGDP_percap=mean(gdpPercap)) %>%
  arrange(desc(meanGDP_percap))
  flextable(print(dF2,10))
```

```
## # A
## #   tibble:
## #   5 x 2
## # i 2
## #   more
## #   variables:
## #   continent <fct>,
## #   meanGDP_percap <dbl>
```

continent	meanGDP_percap
Oceania	26,938.778
Europe	21,711.732
Asia	10,174.090
Americas	9,287.677
Africa	2,599.385

```
data1 <-df %>%
  select(continent,
         year,
         pop,
```

```
        country) %>%
  group_by(continent) %>%
  filter(year==1997) %>%
  summarise(total_pop=sum(pop),
            countries=n_distinct(country)) %>%
  arrange(desc(total_pop))
  flextable(data1)
```

continent	total_pop	countries
Asia	3,383,285,500	33
Americas	796,900,410	25
Africa	743,832,984	52
Europe	568,944,148	30
Oceania	22,241,430	2

```
Avg_pop <- data1 %>%
  select(continent,
         total_pop,
         countries) %>%
  mutate(avg_pop_per_cntry=total_pop/countries)
  flextable(Avg_pop)
```

continent	total_pop	countries	avg_pop_per_cntry
Asia	3,383,285,500	33	102,523,803
Americas	796,900,410	25	31,876,016
Africa	743,832,984	52	14,304,480
Europe	568,944,148	30	18,964,805
Oceania	22,241,430	2	11,120,715

```
df3<- df %>%
  filter(continent=="Asia" & year>1990 & year<=2000) %>%
  summarise(Avglife.exp=mean(lifeExp),
            Avgpop=mean(pop),
            Avggdp= mean(gdpPercap))
  flextable(df3)
```

Avglife.exp	Avgpop	Avggdp
67.27886	98,736,026	9,236.892

```
data2 <- df1 %>%
  select( continent,
          pop,
          lifeExp ) %>%
```

```
group_by(continent) %>%
summarise(Total_pop=sum(pop),
          avglife= mean(lifeExp)) %>%
arrange(desc(avglife))
flextable(data2)
```

continent	Total_pop	avglife
Oceania	212,992,136	74.32621
Europe	6,181,115,304	71.90369
Americas	7,351,438,499	64.65874
Asia	30,507,333,901	60.06490
Africa	6,187,585,961	48.86533

```
data3 <- df %>%
  group_by(year,
           continent) %>%
  summarise(Year_avg.mean=mean(lifeExp),
            Year_pop=sum(pop))
  gt(head(data3,n=10))
```

continent	Year_avg.mean	Year_pop
1952		
Africa	39.13550	237640501
Americas	53.27984	345152446
Asia	46.31439	1395357351
Europe	64.40850	418120846
Oceania	69.25500	10686006
1957		
Africa	41.26635	264837738
Americas	55.96028	386953916
Asia	49.31854	1562780599
Europe	66.70307	437890351
Oceania	70.29500	11941976

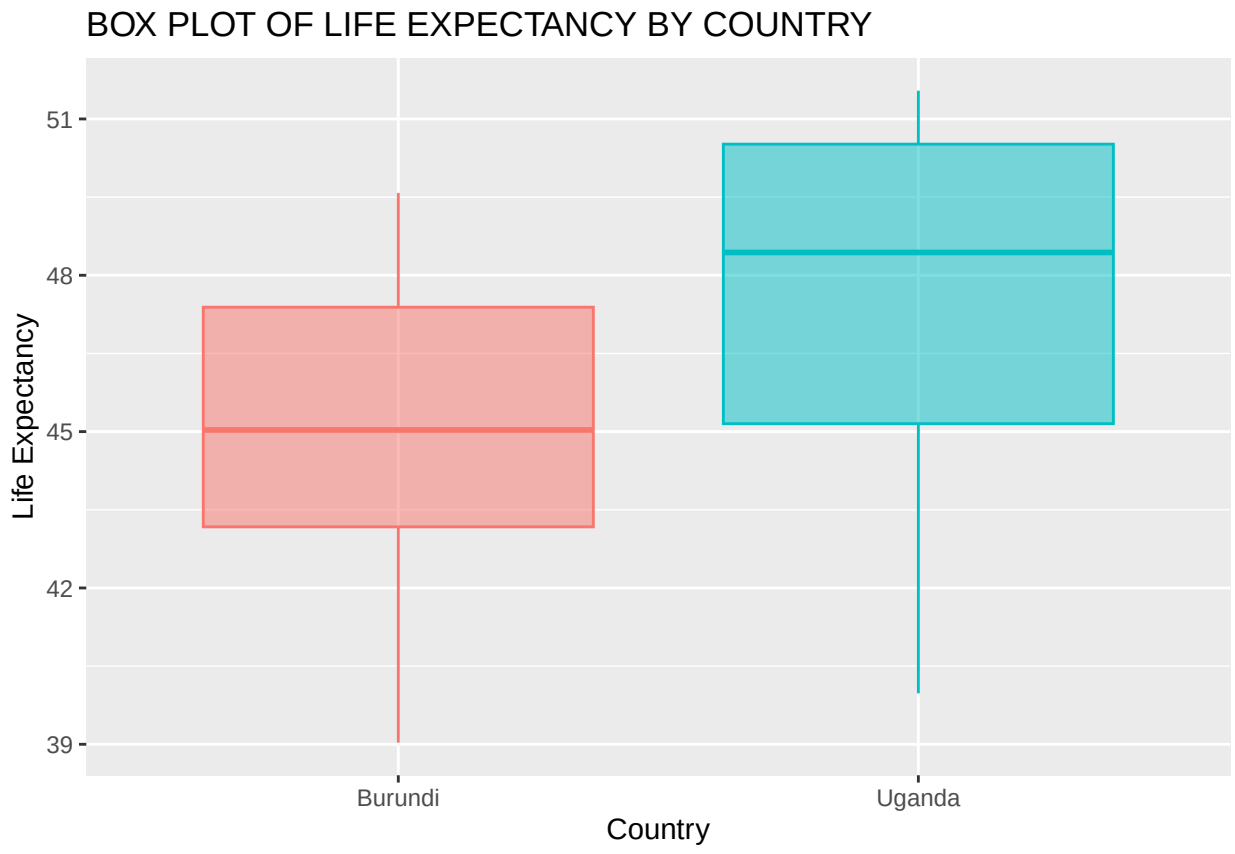
- *Majority of countries were middle income countries(33%) . From Asia, Japan had the highest life Expectancy in 2007(82.6yrs).From Africa, DRC had the lowest GDP-percapita in 2002($241).Oceania has the smallest (2 countries), Americas had 25 countries, Europe had 30 countries, Asia had 33 countries and Africa had the majority(52 countries).*
- *From Africa in 2002, only 9 countries had a GDP-percapita > $5000 with Gabon having the highest($12522) and lowest Algeria($5288).Among continents, Oceania had the highest mean life expectancy(79.4yrs) while Africa had the lowest(53.3yrs).The continent with the highest population wa Asia with average population per country being 102, 523,803 people while Oceania had the lowest population with average population per country 11,120,715.*
- *Between 1990 and 2000, Asia's mean life expectancy and GDP-percapita was 67.2yrs and $9,236.9 respectively.*

- *Among all continents, Asia had the highest total population and Oceania the lowest.On average, Oceania had the highest total life expectancy(74.3yrs)while Africa had the lowest(48.9yrs)*

Visualization, Hypothesis testing, Correlation and Regression

```
# Visualization of data spread between countries
data4 <- df %>%
  select(!c(continent,
            year,
            pop,
            gdpPercap)) %>%
  subset(country=="Burundi"|country=="Uganda")

data4 %>%
  group_by(country) %>%
  ggplot(data=data4,
         mapping=aes(x=country,
                     y=lifeExp,
                     col=country,
                     fill=country))+
  geom_boxplot(alpha=0.5,
               show.legend = FALSE)+
  labs(title="BOX PLOT OF LIFE EXPECTANCY BY COUNTRY",
       x="Country",
       y="Life Expectancy")
```



```r
# Test for equality of Variances
leveneTest(data4$lifeExp~data4$country)
```

```
## Levene's Test for Homogeneity of Variance (center = median)
##       Df F value Pr(>F)
## group  1  0.2182  0.645
##       22
```

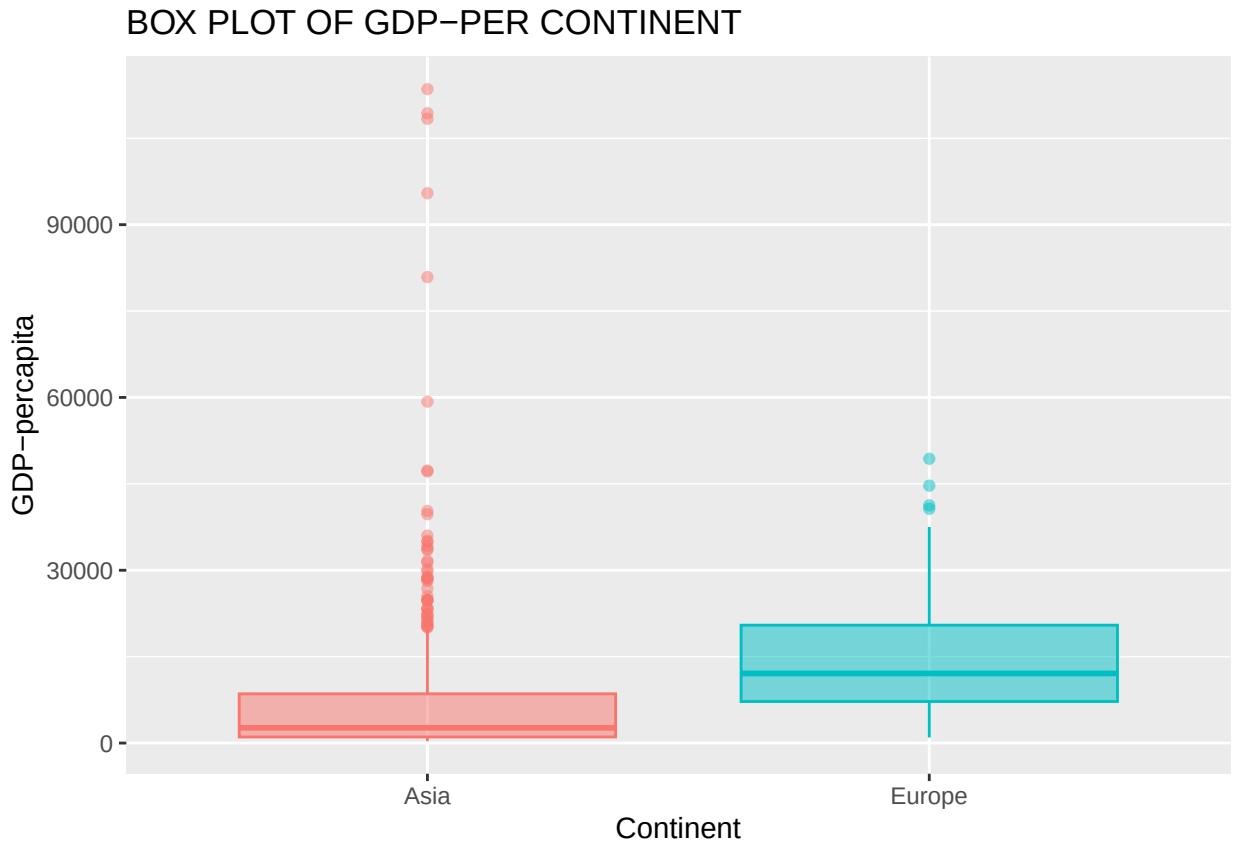
```r
# Sample t-test for testing equality of means between two groups
# H0: The means of two groups are equal
t.test(data=data4,
       lifeExp~country,
       alternative='two.sided',
       conf.level=0.95,
       var.eq=TRUE,
       paired= FALSE)
```

```
##
##  Two Sample t-test
##
## data:  lifeExp by country
## t = -1.9759, df = 22, p-value = 0.06083
## alternative hypothesis: true difference in means between group Burundi and group Uganda is not equal
## 95 percent confidence interval:
##  -5.7418375  0.1388375
## sample estimates:
## mean in group Burundi  mean in group Uganda
##              44.81733              47.61883
```

```r
data5 <- df %>%
  select(!c(country,
            year,
            lifeExp)) %>%
  subset(continent=="Europe"|continent=="Asia")
  print(data5)
```

```
## # A tibble: 756 x 3
##    continent      pop gdpPercap
##    <fct>        <int>     <dbl>
##  1 Asia       8425333      779.
##  2 Asia       9240934      821.
##  3 Asia      10267083      853.
##  4 Asia      11537966      836.
##  5 Asia      13079460      740.
##  6 Asia      14880372      786.
##  7 Asia      12881816      978.
##  8 Asia      13867957      852.
##  9 Asia      16317921      649.
## 10 Asia      22227415      635.
## # i 746 more rows
```

```
ggplot(data=data5,
       mapping=aes(x=continent,
                   y=gdpPercap,
                   col=continent,
                   fill =continent))+
geom_boxplot(alpha=0.5,
             show.legend = FALSE)+
labs(title="BOX PLOT OF GDP-PER CONTINENT",
     x="Continent",
     y="GDP-percapita")
```




```
leveneTest(data5$gdpPercap~data5$continent)
```

```
## Levene's Test for Homogeneity of Variance (center = median)
##        Df F value Pr(>F)
## group   1   1.009 0.3155
##       754
```

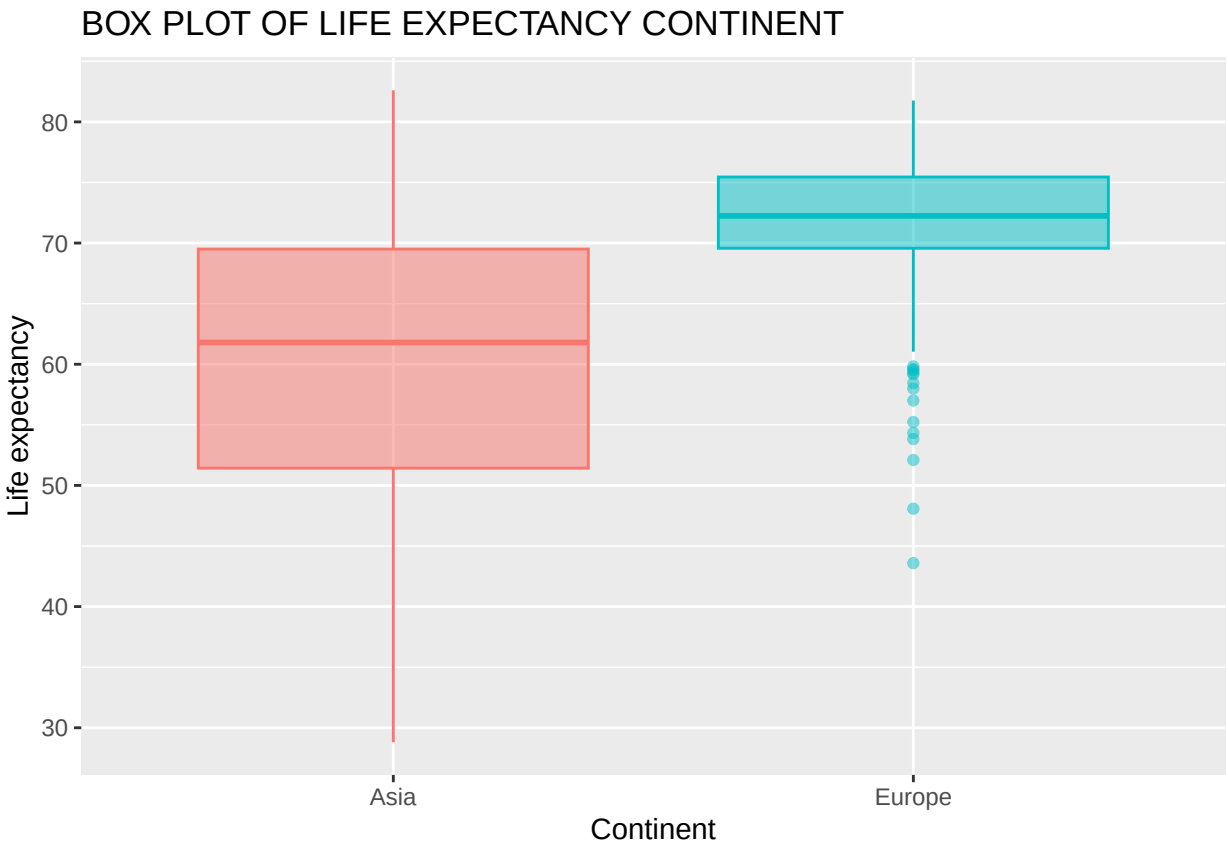
```
t.test(data=data5,
       gdpPercap~data5$continent,
       alternative='two.sided',
       conf.level=0.95,
       var.eq=TRUE,
       paired= FALSE)
```

```
##
##  Two Sample t-test
##
## data:  gdpPercap by data5$continent
## t = -7.4889, df = 754, p-value = 1.944e-13
## alternative hypothesis: true difference in means between group Asia and group Europe is not equal to
## 95 percent confidence interval:
##  -8288.862 -4845.788
## sample estimates:
##   mean in group Asia mean in group Europe
##              7902.15             14469.48
```

```
data6<- df%>%
  select(!c(country,
            year,
            gdpPercap)) %>%
  subset(continent=="Europe"|continent=="Asia")
  print(data6)
```

```
## # A tibble: 756 x 3
##    continent lifeExp      pop
##    <fct>       <dbl>    <int>
##  1 Asia         28.8  8425333
##  2 Asia         30.3  9240934
##  3 Asia         32.0 10267083
##  4 Asia         34.0 11537966
##  5 Asia         36.1 13079460
##  6 Asia         38.4 14880372
##  7 Asia         39.9 12881816
##  8 Asia         40.8 13867957
##  9 Asia         41.7 16317921
## 10 Asia         41.8 22227415
## # i 746 more rows
```

```
ggplot(data=data6,
       mapping=aes(x=continent,
                   y=lifeExp,
                   col=continent,
                   fill=continent))+
  geom_boxplot(alpha=0.5,
               show.legend = FALSE)+
  labs(title="BOX PLOT OF LIFE EXPECTANCY CONTINENT",
       x="Continent",
       y="Life expectancy")
```




```
leveneTest(data6$lifeExp~data6$continent)
```

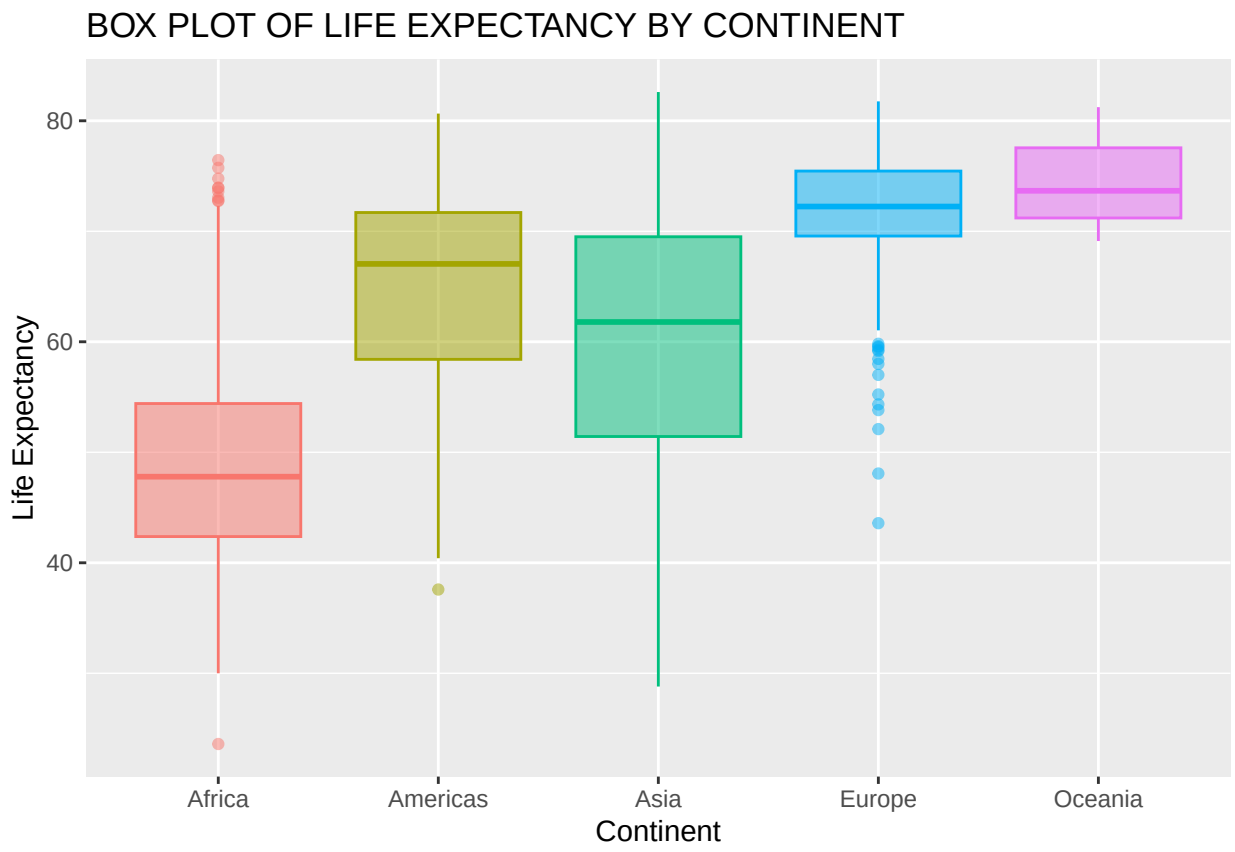
```
## Levene's Test for Homogeneity of Variance (center = median)
##        Df F value    Pr(>F)
## group   1  201.02 < 2.2e-16 ***
##       754
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
#H0: The mean life expectancy of Asia is greater than Europe
t.test(data=data6,
       lifeExp~continent,
       alternative='greater',
       conf.level=0.95,
       var.eq=FALSE,
       paired= FALSE)
```

```
##
##  Welch Two Sample t-test
##
## data:  lifeExp by continent
## t = -17.899, df = 565.16, p-value = 1
## alternative hypothesis: true difference in means between group Asia and group Europe is greater than
## 95 percent confidence interval:
##  -12.9285      Inf
```

```
## sample estimates:
##   mean in group Asia mean in group Europe
##             60.06490              71.90369
```

```
df1 %>%
  ggplot(mapping=aes(x=continent,
                     y=lifeExp,
                     col=continent,
                     fill=continent))+
  geom_boxplot(alpha=0.5,
               show.legend = FALSE)+
  labs(title="BOX PLOT OF LIFE EXPECTANCY BY CONTINENT",
       x="Continent",
       y="Life Expectancy")
```




```
# Testing for difference in mean between more than two groups
 ANOVA1 <- aov(lifeExp~continent,
               data=df1)

 # Undestanding where diffences lie
 summary(ANOVA1)
```

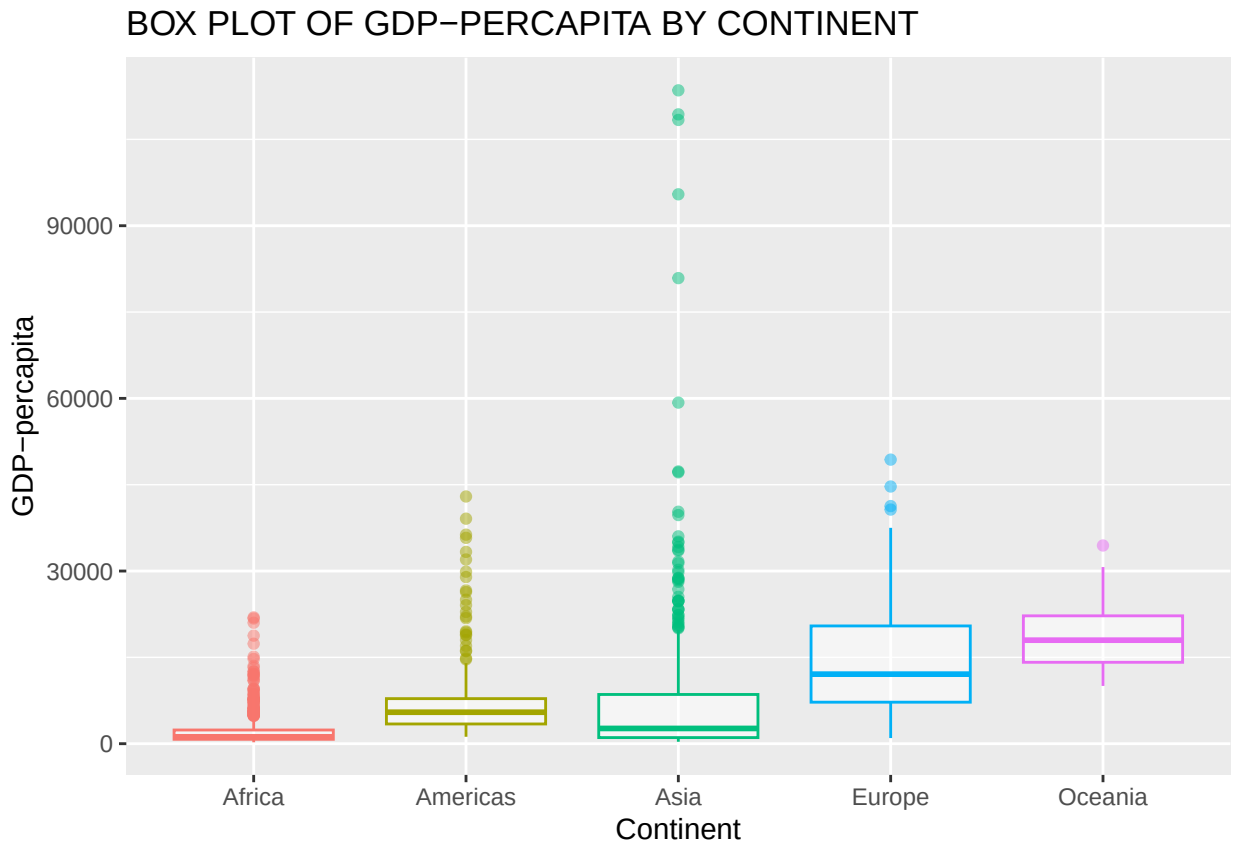
```
##               Df Sum Sq Mean Sq F value Pr(>F)
## continent      4 139343   34836   408.7 <2e-16 ***
## Residuals   1699 144805      85
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
TukeyHSD(ANOVA1)
```

```
##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##
## Fit: aov(formula = lifeExp ~ continent, data = df1)
##
## $continent
##                         diff       lwr       upr     p adj
## Americas-Africa  15.793407 14.022263 17.564550 0.0000000
## Asia-Africa      11.199573  9.579887 12.819259 0.0000000
## Europe-Africa    23.038356 21.369862 24.706850 0.0000000
## Oceania-Africa   25.460878 20.216908 30.704848 0.0000000
## Asia-Americas    -4.593833 -6.523432 -2.664235 0.0000000
## Europe-Americas   7.244949  5.274203  9.215696 0.0000000
## Oceania-Americas  9.667472  4.319650 15.015293 0.0000086
## Europe-Asia      11.838783 10.002952 13.674614 0.0000000
## Oceania-Asia     14.261305  8.961718 19.560892 0.0000000
## Oceania-Europe    2.422522 -2.892185  7.737230 0.7250559
```

```
df1%>%
  ggplot(mapping=aes(x=continent,
                     y=gdpPercap,
                     col=continent))+
  geom_boxplot(alpha=0.5,
               show.legend = FALSE)+
  labs(title="BOX PLOT OF GDP-PERCAPITA BY CONTINENT",
       x="Continent",
       y="GDP-percapita")
```




```
ANOVA2 <- aov(gdpPercap~continent,
              data=df1)
summary(ANOVA2)
```

```
##               Df    Sum Sq   Mean Sq F value Pr(>F)
## continent      4 3.799e+10 9.498e+09   126.6 <2e-16 ***
## Residuals   1699 1.275e+11 7.504e+07
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
TukeyHSD(ANOVA2)
```

```
##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##
## Fit: aov(formula = gdpPercap ~ continent, data = df1)
##
## $continent
##                         diff        lwr       upr     p adj
## Americas-Africa   4942.3558  3280.4806  6604.231 0.0000000
## Asia-Africa       5708.3958  4188.6340  7228.158 0.0000000
## Europe-Africa    12275.7210 10710.1623 13841.280 0.0000000
## Oceania-Africa   16427.8546 11507.4034 21348.306 0.0000000
## Asia-Americas      766.0401 -1044.5150  2576.595 0.7767582
```

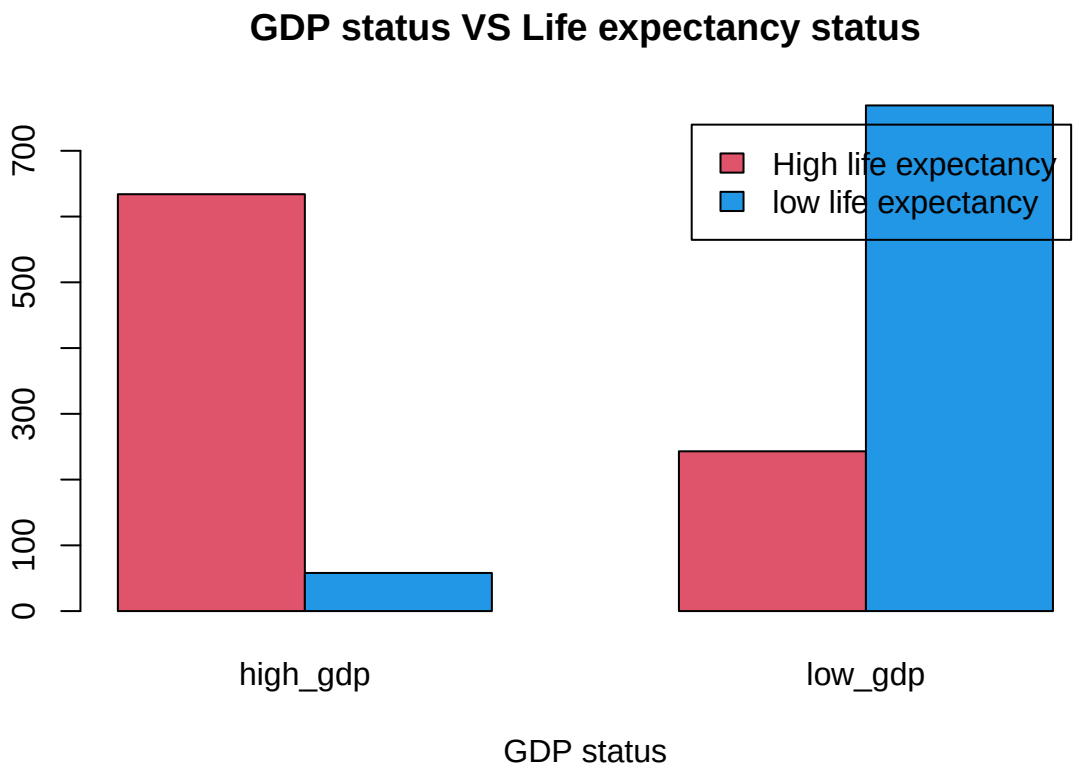
```
## Europe-Americas   7333.3652  5484.2011  9182.529 0.0000000
## Oceania-Americas 11485.4989  6467.6035 16503.394 0.0000000
## Europe-Asia       6567.3251  4844.7530  8289.897 0.0000000
## Oceania-Asia     10719.4588  5746.8215 15692.096 0.0000000
## Oceania-Europe    4152.1337  -834.6909  9138.958 0.1539474
```

```
# Testing for difference in means in categoricals
gdp_status <-df%>%
  mutate(GDP_status = ifelse(gdpPercap>5000,"high_gdp",
                             "low_gdp"),
 life_exp_status = ifelse(lifeExp>60,"high_life.exp",
                          "low_life.exp"))

tabyl(gdp_status$GDP_status)
```

```
##  gdp_status$GDP_status    n   percent
##              high_gdp  692 0.4061033
##               low_gdp 1012 0.5938967
```

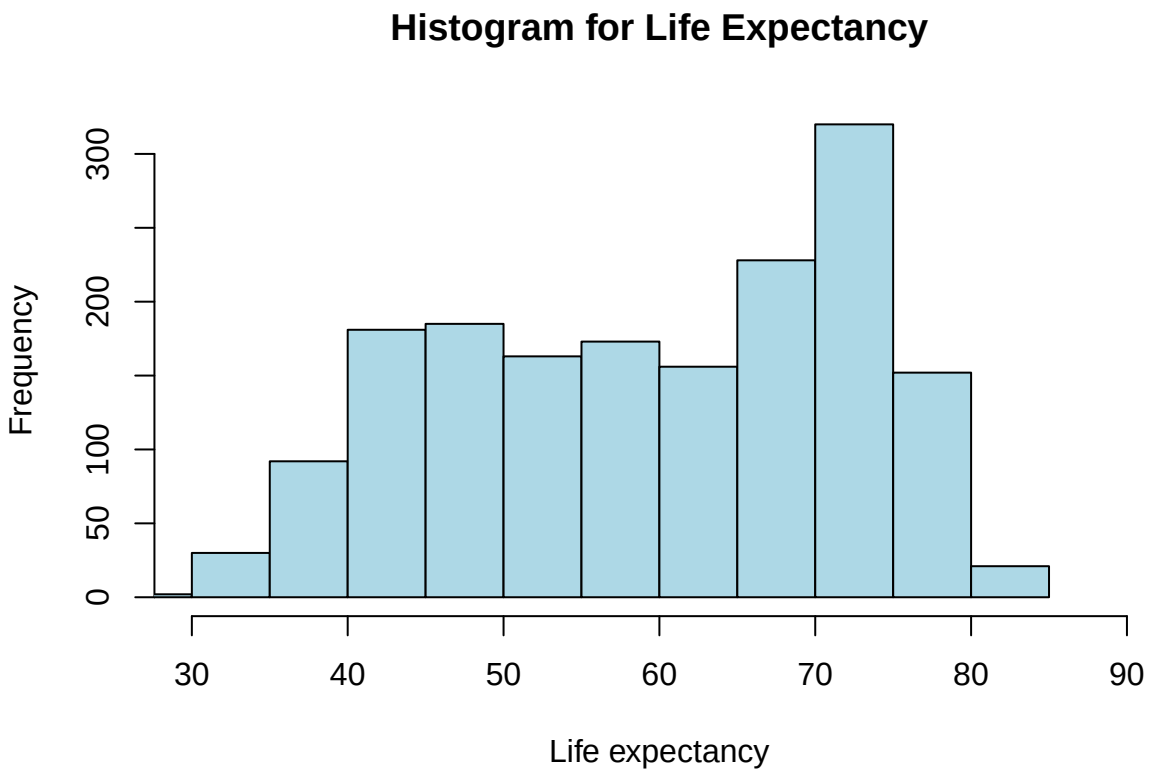
```
 TAB <-  table(gdp_status$life_exp_status,
               gdp_status$GDP_status)
 barplot(TAB,beside=T,
         legend.text = c('High life expectancy',
                         'low life expectancy'),
         main = 'GDP status VS Life expectancy status',
         xlab='GDP status',
         col=c(2,4))
```



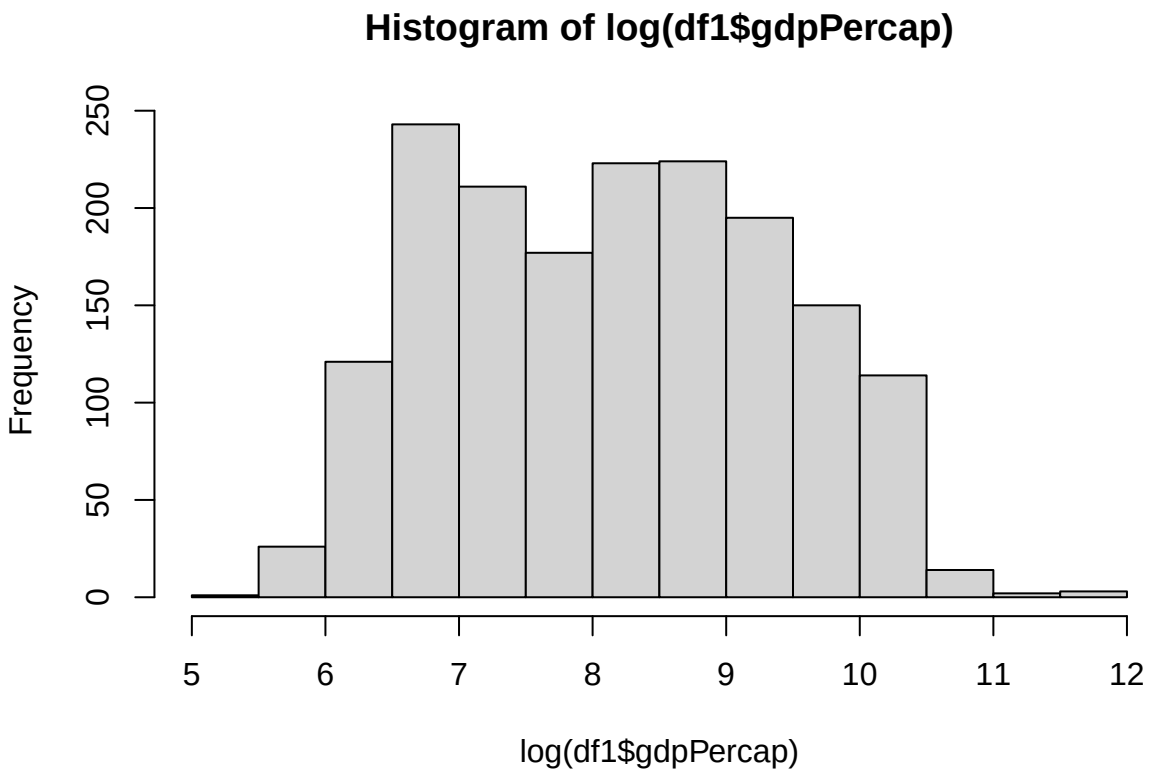

```
  chisq.test(TAB,
             correct=TRUE)
```

```
##
##  Pearson's Chi-squared test with Yates' continuity correction
##
## data:  TAB
## X-squared = 749.32, df = 1, p-value < 2.2e-16
```

```
  # Undestanding distribution of data
  hist(df1$lifeExp,
       breaks = 10,
       xlim=c(30,90),
       col='light blue',
       main  ='Histogram for Life Expectancy',
       xlab='Life expectancy')
```

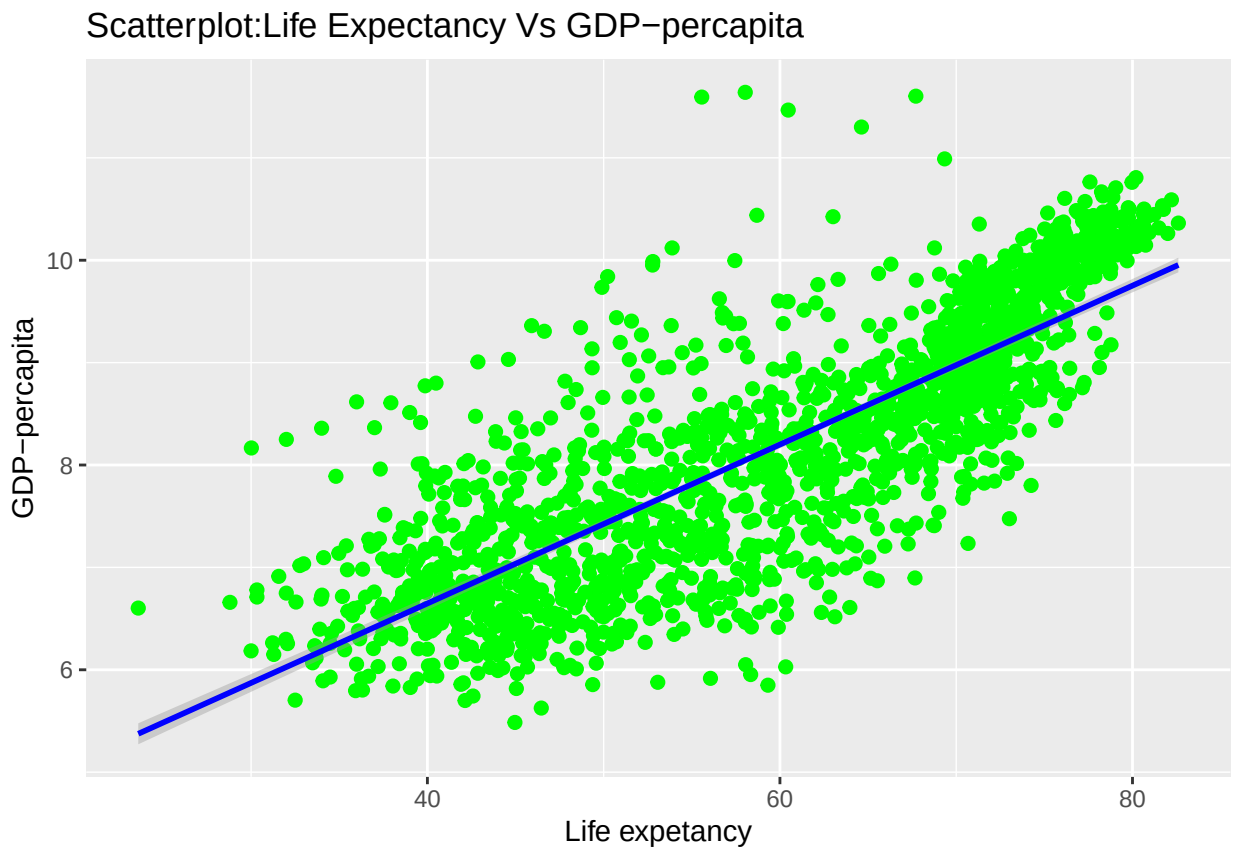



```
hist(log(df1$gdpPercap))
```



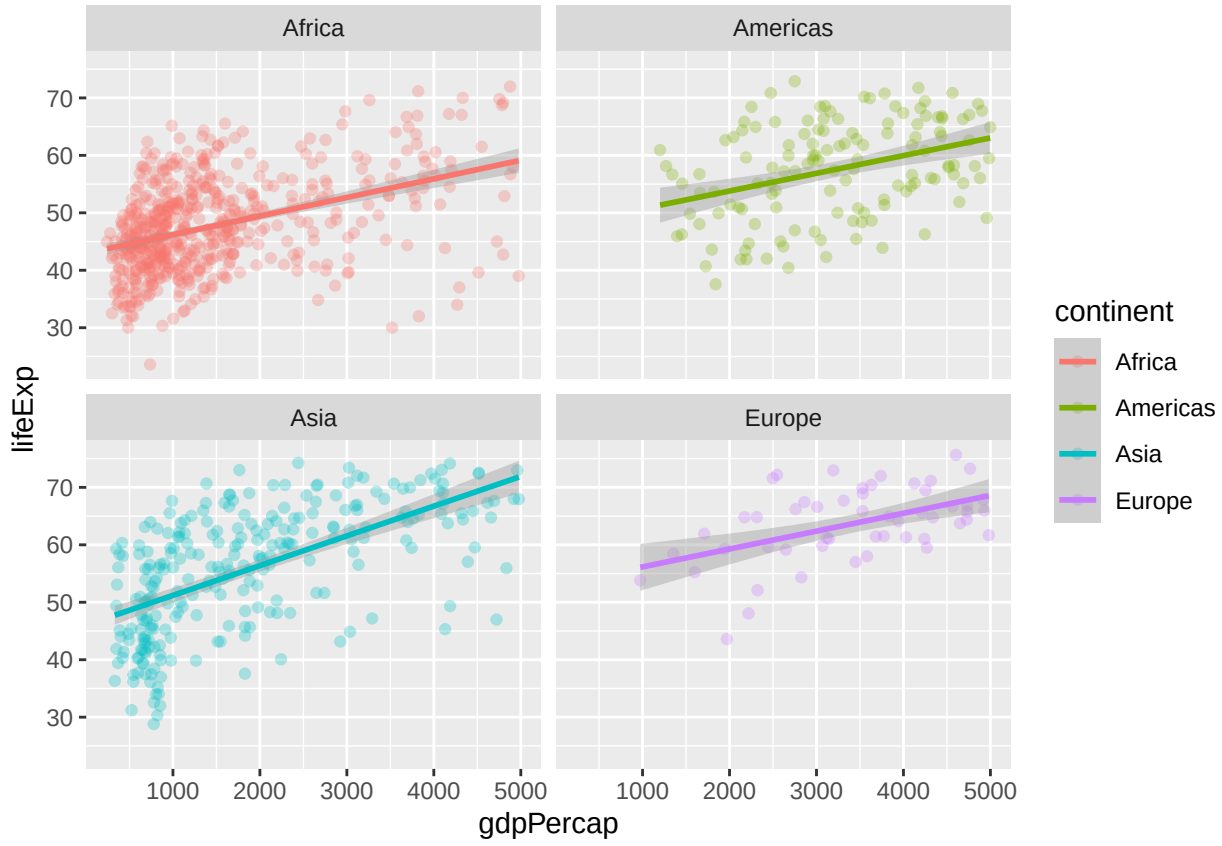

```
 ggplot(data=df1,
        mapping=aes(x=lifeExp,
                    y=log(gdpPercap)))+
     geom_point(col='green',
                size=2)+
     geom_smooth(method='lm',
                 col= 'blue')+
   labs(title="Scatterplot:Life Expectancy Vs GDP-percapita",
        x="Life expetancy",
       y="GDP-percapita")
```

```
## `geom_smooth()` using formula = 'y ~ x'
```




```
# Correlation and regression
df %>%
  filter(gdpPercap < 5000) %>%
  ggplot(aes(x=gdpPercap,
             y=lifeExp,
             col=continent))+
  geom_point(alpha=0.3)+
  geom_smooth(method=lm)+
  facet_wrap(~continent)
```

```
## `geom_smooth()` using formula = 'y ~ x'
```




```
cor(df$lifeExp,
    df$gdpPercap,
    method = 'pearson')
```

```
## [1] 0.5837062
```

```
lm(df$lifeExp ~ df$gdpPercap)
```

```
##
## Call:
## lm(formula = df$lifeExp ~ df$gdpPercap)
##
## Coefficients:
##  (Intercept)  df$gdpPercap
##    5.396e+01     7.649e-04
```

```
summary(lm(df$lifeExp ~ df$gdpPercap))
```

```
##
## Call:
## lm(formula = df$lifeExp ~ df$gdpPercap)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
```

```
## -82.754  -7.758   2.176   8.225  18.426
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  5.396e+01  3.150e-01  171.29   <2e-16 ***
## df$gdpPercap 7.649e-04  2.579e-05   29.66   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 10.49 on 1702 degrees of freedom
## Multiple R-squared:  0.3407, Adjusted R-squared:  0.3403
## F-statistic: 879.6 on 1 and 1702 DF,  p-value: < 2.2e-16
```

- *There is no significant difference in mean life expectancy between Uganda and Burundi(P=0.063).*
- *There is a significant difference in mean GDP-percapita between Asia and Europe(P<0.05)*
- *-here is no significant evidence to show that the life expectancy of Asia is greater than Europe(P=0.9)*
- *There is a significant difference in mean life expectancy continents(P<0.05).Pair-wise comparisons showed significant difference for the 9 pairs except between Oceania and Europe(P=0.72).*
- *There is a significant difference in mean GDP-percapita between continents(P<0.05).Pair-wise comparisons showed significant difference for the 8 pairs except Oceania-Europe(P=0.15) and Asia-America(P=0.78).*
- *There is a significant difference between GDP-percapital status levels and Life expectancy levels(p<0.05).*
- *There is a moderate positive correlation between GDP-percapita and life expectancy(r=0.58).*
- *The regression model is: Life exp=53.96+0.008GDP-percapita. The model can predict up to 34%.*