Kuis1_G64190069_Rizal Mujahiddan

Rizal Mujahiddan

2/19/2022

```
datasetku <- read.csv("Kuis UTS 1.csv")
print(names(datasetku))
```

```
##  [1] "wbcode"   "country"  "gdp50"    "gdp90"    "gdp95"    "lnd100km"
##  [7] "pop100km" "lnd100cr" "pop100cr" "dens65c"  "dens65i"  "dens95c"
## [13] "dens95i"  "landlock" "landlneu" "airdist"  "tropicar" "troppop"
## [19] "malfal66" "malfal94" "lhcpc"    "south"    "landarea" "open6590"
## [25] "newstate" "socialst" "lifex65"  "urbpop95" "wardum"   "pop95"
## [31] "zpolar"   "zboreal"  "zdestmp"  "zdestrp"  "zdrytemp" "zwettemp"
## [37] "zsubtrop" "ztropics" "zwater"   "eu"       "safri"    "sasia"
## [43] "transit"  "latam"    "eseasia"  "region"
```

```
print(datasetku$socialst)
```

```
##  [1] 1 0 0 0 0 1 0 0 1 0 0 0 0 0 0 1 0 0 1 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0
## [39] 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 1 0 0 0 0 0 0 1 0 0 1 0 0
## [77] 0 0 0 0 0 1 0 0 0 0 0 0 0 0 1 0 0 0 1
```

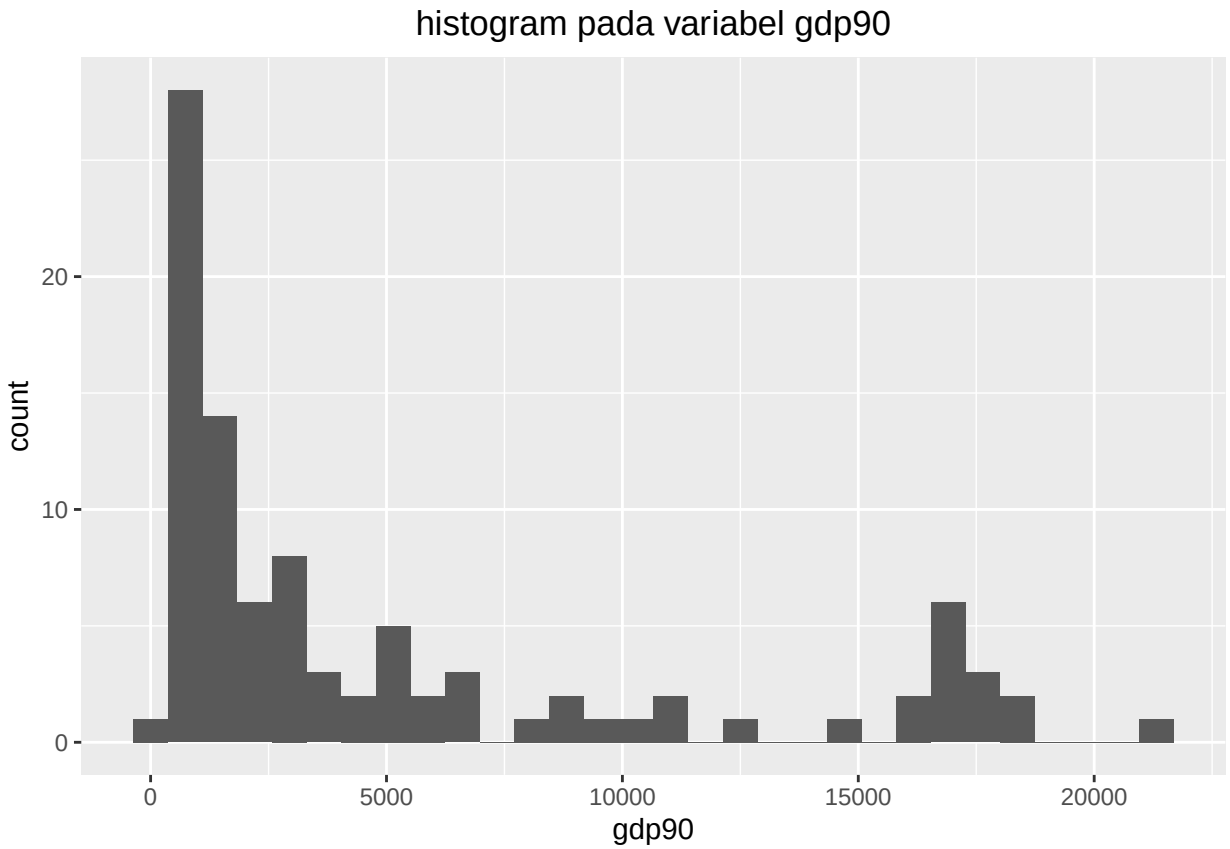
```
head(datasetku)
```

```
##   wbcode   country gdp50 gdp90      gdp95  lnd100km  pop100km lnd100cr pop100cr
## 1    AGO    Angola   986   654  2055.1880 0.1187576 0.2655837       14       30
## 2    ARG Argentina  4987  6581  9287.2120 0.1230889 0.1889420       20       30
## 3    AUT   Austria  3731 16792 21269.8500 0.0075848 0.0057634       62       71
## 4    BDI   Burundi   320   599   656.9504 0.0000000 0.0000000        0        0
## 5    BEL   Belgium  5346 16807 21695.3800 0.4892515 0.7138057       99       99
## 6    BEN     Benin  1087  1107  1206.5740 0.1089616 0.4924399       11       49
##      dens65c    dens65i   dens95c   dens95i landlock landlneu airdist tropicar
## 1   9.291969   3.462699  19.32299   7.20081        0        0    6830   1.0000
## 2  12.498500   7.530861  19.44353  11.71553        0        0    8570   0.0268
## 3  66.783080  88.049610  73.97482  97.53151        1        0     840   0.0000
## 4   0.000000 125.116800   0.00000 243.92520        1        1    6600   1.0000
## 5 420.711100 161.580900 451.02420 173.22310        0        0     190   0.0000
## 6  99.277840  12.513060 223.68160  28.19301        0        0    5040   1.0000
##   troppop malfal66 malfal94     lhcpc south landarea  open6590 newstate
## 1  0.7491        1        1  5.732350     1  1246700 0.0000000        2
## 2  0.0000        0        0  4.518682     1  2736690 0.0000000        0
## 3  0.0000        0        0  2.858482     0    82730 1.0000000        1
## 4  1.0000        1        1 -4.605170     1    25680 0.0000000        2
## 5  0.0000        0        0 -4.605170     0    32820 1.0000000        0
```

```
## 6  0.5141        1        1  1.053939     0   110620 0.0384615        2
##   socialst lifex65 urbpop95 wardum   pop95    zpolar   zboreal   zdestmp
## 1        1    36.0     31.0      1 10.7720 0.0000000 0.0000000 0.0221579
## 2        0    65.8     88.1      0 34.6650 0.0125004 0.0843879 0.2308541
## 3        0    69.9     64.3      0  8.0540 0.0228520 0.4829857 0.0000000
## 4        0    43.5      7.5      0  6.2640 0.0000000 0.0000000 0.0000000
## 5        0    70.9     97.0      0 10.1459 0.0000000 0.0000000 0.0000000
## 6        1    41.0     38.4      0  5.4750 0.0000000 0.0000000 0.0000000
##     zdestrp  zdrytemp  zwettemp  zsubtrop  ztropics    zwater eu safri sasia
## 1 0.0321074 0.0000000 0.0000000 0.8917290 0.0501317 0.0038740  0     1     0
## 2 0.0008983 0.3237756 0.0477848 0.2941360 0.0022708 0.0033919  0     0     0
## 3 0.0000000 0.0000000 0.4941623 0.0000000 0.0000000 0.0000000  1     0     0
## 4 0.0000000 0.0000000 0.0000000 0.9805556 0.0000000 0.0194444  0     1     0
## 5 0.0000000 0.0000000 0.9803160 0.0000000 0.0000000 0.0196840  1     0     0
## 6 0.0000000 0.0000000 0.0000000 0.0072161 0.9797114 0.0130726  0     1     0
##   transit latam eseasia                    region
## 1       0     0       0        Sub-Saharan Africa
## 2       0     1       0 Latin America & Caribbean
## 3       0     0       0            Western Europe
## 4       0     0       0        Sub-Saharan Africa
## 5       0     0       0            Western Europe
## 6       0     0       0        Sub-Saharan Africa
```

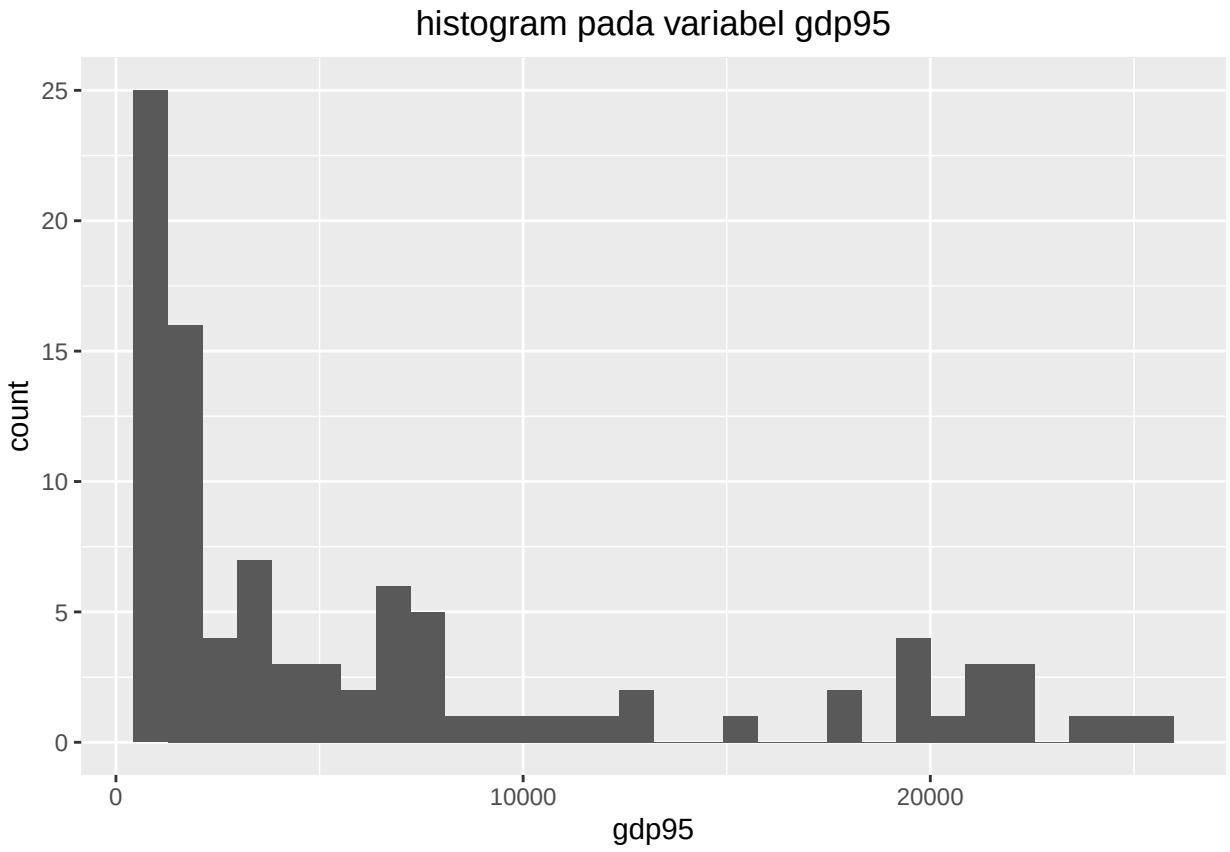
```
library(ggplot2)

ggplot(datasetku,aes(`gdp90`)) + geom_histogram() +
  labs(title="histogram pada variabel gdp90") +
  theme(plot.title = element_text(hjust=0.5))
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



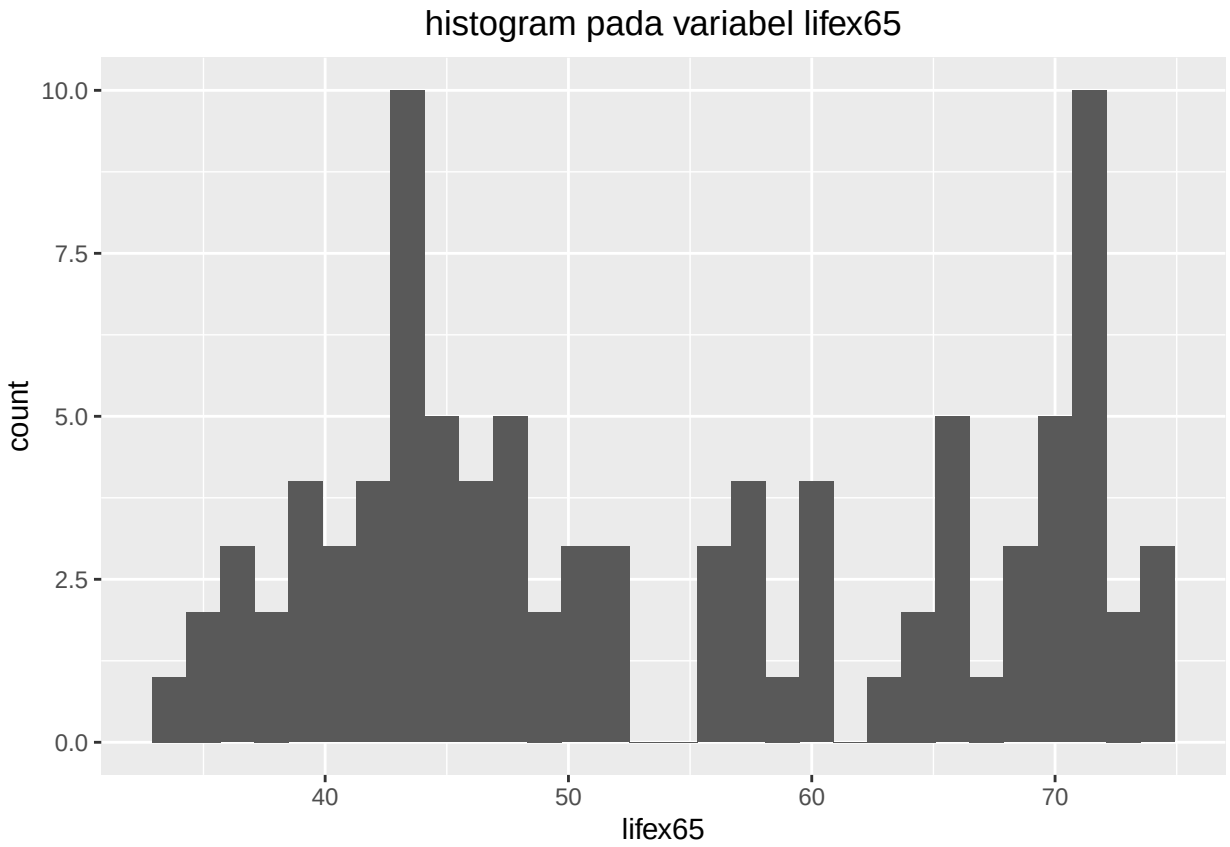

```
ggplot(datasetku,aes(`gdp95`)) + geom_histogram() +
  labs(title="histogram pada variabel gdp95") +
  theme(plot.title = element_text(hjust=0.5))
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



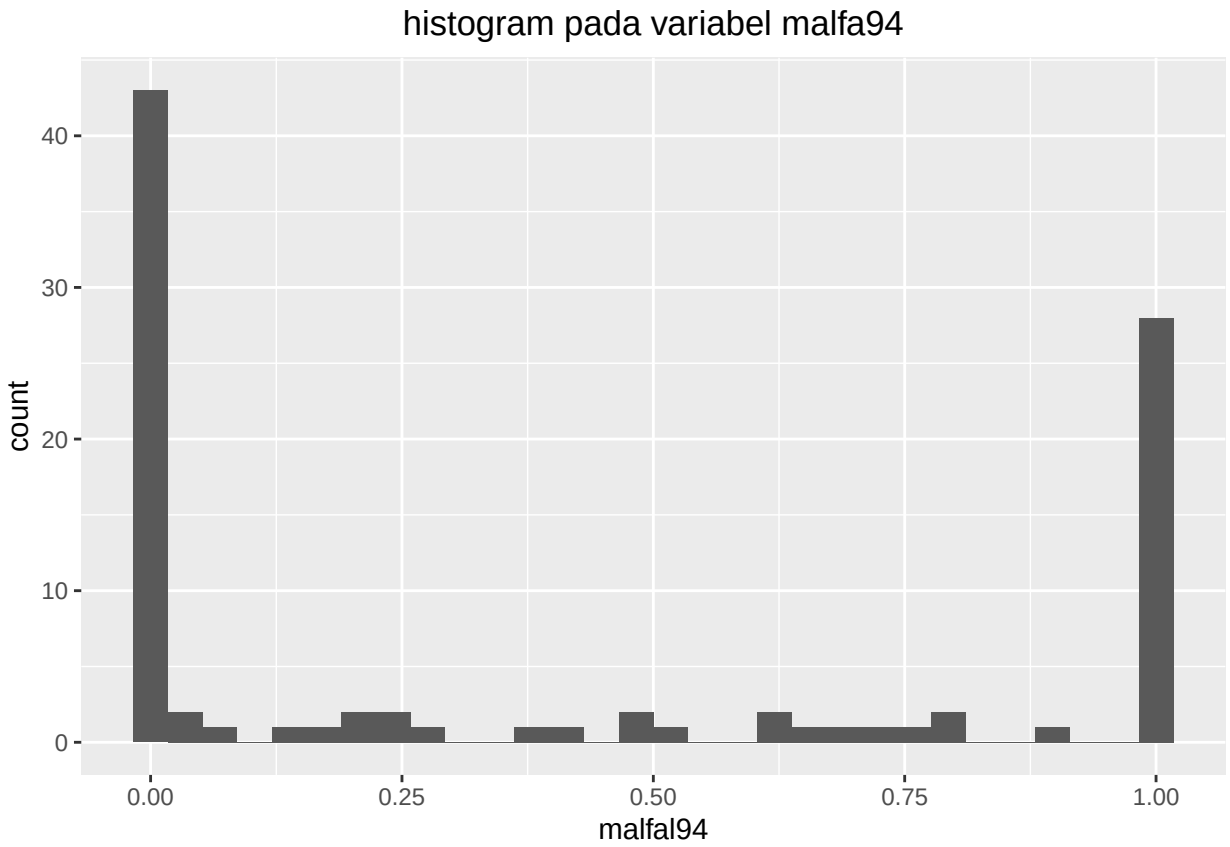

```
ggplot(datasetku,aes(`lifex65`)) + geom_histogram()+
  labs(title="histogram pada variabel lifex65") +
  theme(plot.title = element_text(hjust=0.5))
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



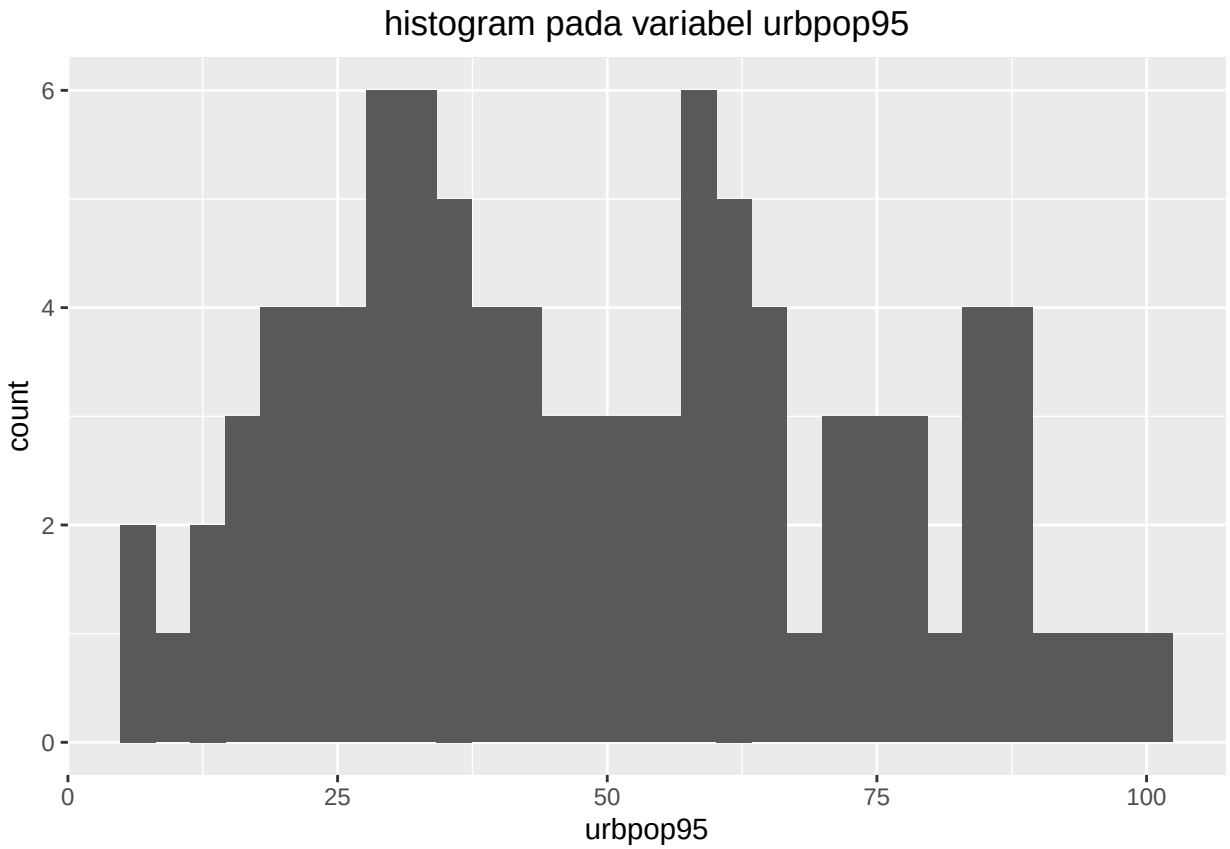

```
ggplot(datasetku,aes(`malfa194`)) + geom_histogram() +
  labs(title="histogram pada variabel malfa94") +
  theme(plot.title = element_text(hjust=0.5))
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```




```
ggplot(datasetku,aes(`urbpop95`)) + geom_histogram() +
  labs(title="histogram pada variabel urbpop95") +
  theme(plot.title = element_text(hjust=0.5))
```

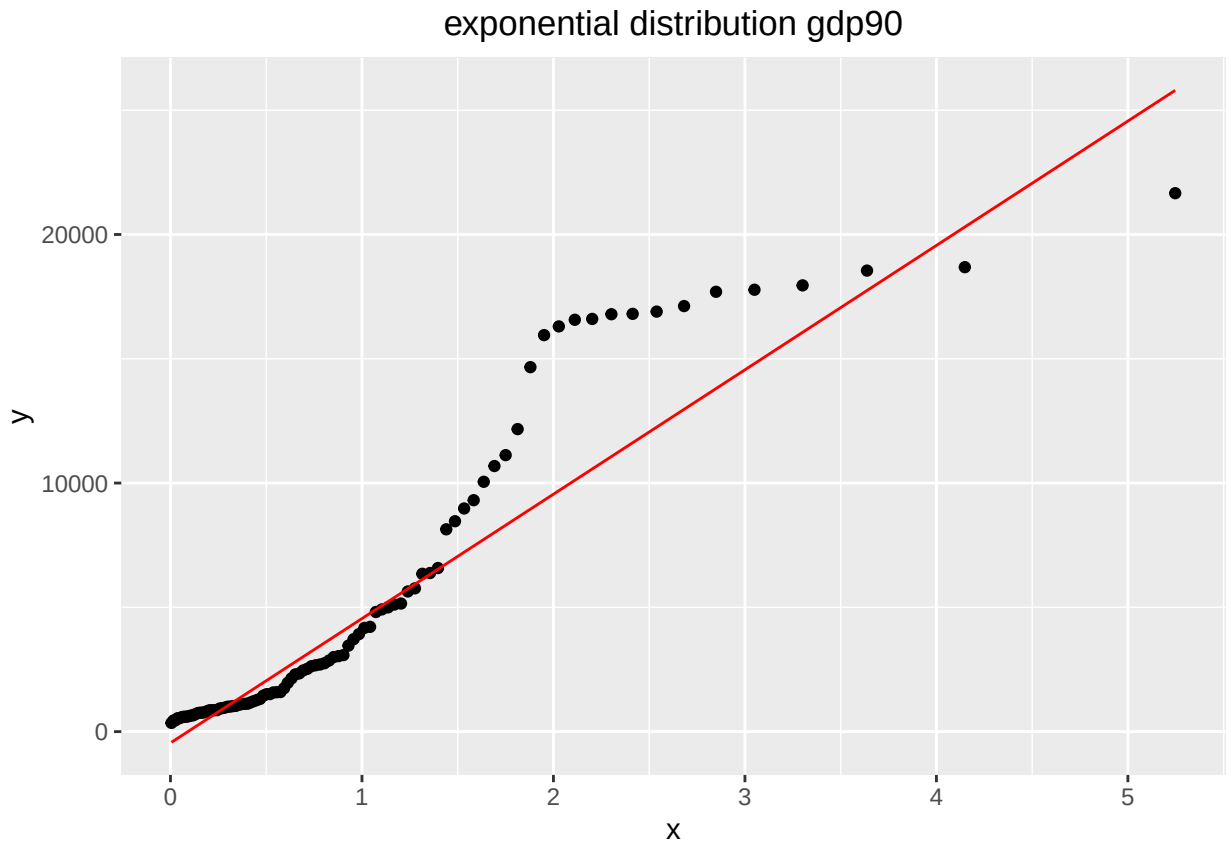
```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



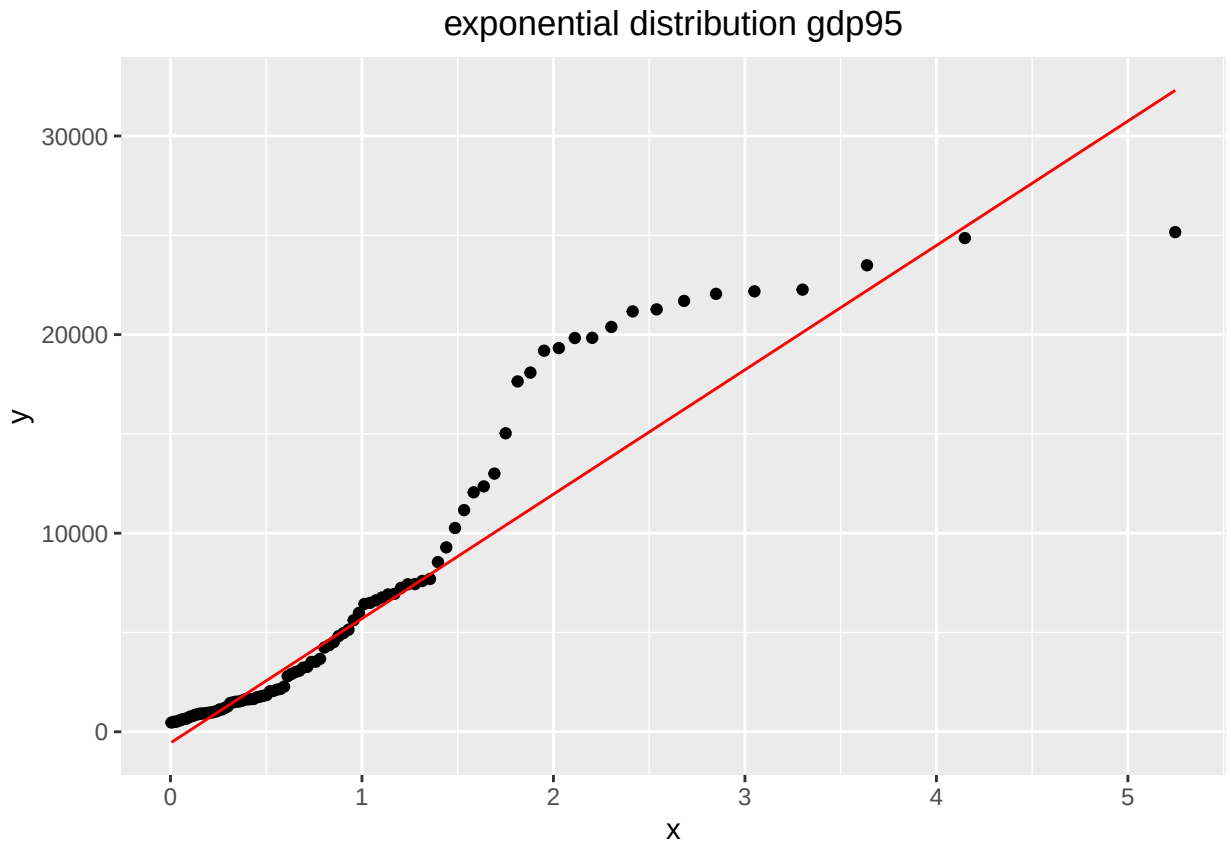
Penjelasan untuk gdp90 dan gdp95 itu jika ditinjau menyebar eksponensial sedangkan untuk lifex65 dan urbpop95 itu bersifat menyebar normal untuk malfal94 ini diperkirakan persebaran binomial ya

untuk saat ini penulis meninjau berdasarkan keseluruhan dahulu, baru dibagi berdasarkan subregion atau region

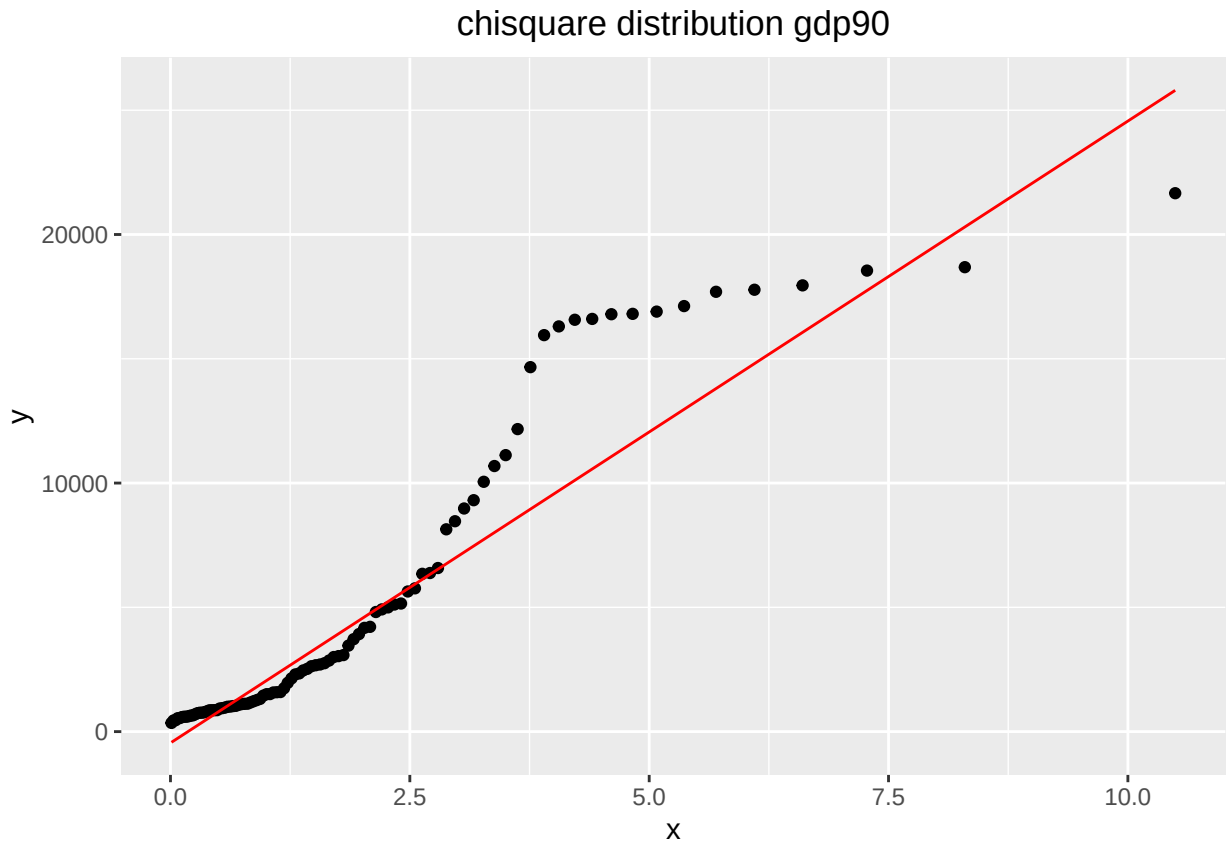
```
ggplot(datasetku,aes(sample = `gdp90`)) +
  geom_qq(distribution = qexp) +
  geom_qq_line(distribution = qexp,col='red')+
  labs(title="exponential distribution gdp90")+
  theme(plot.title = element_text(hjust=0.5))
```



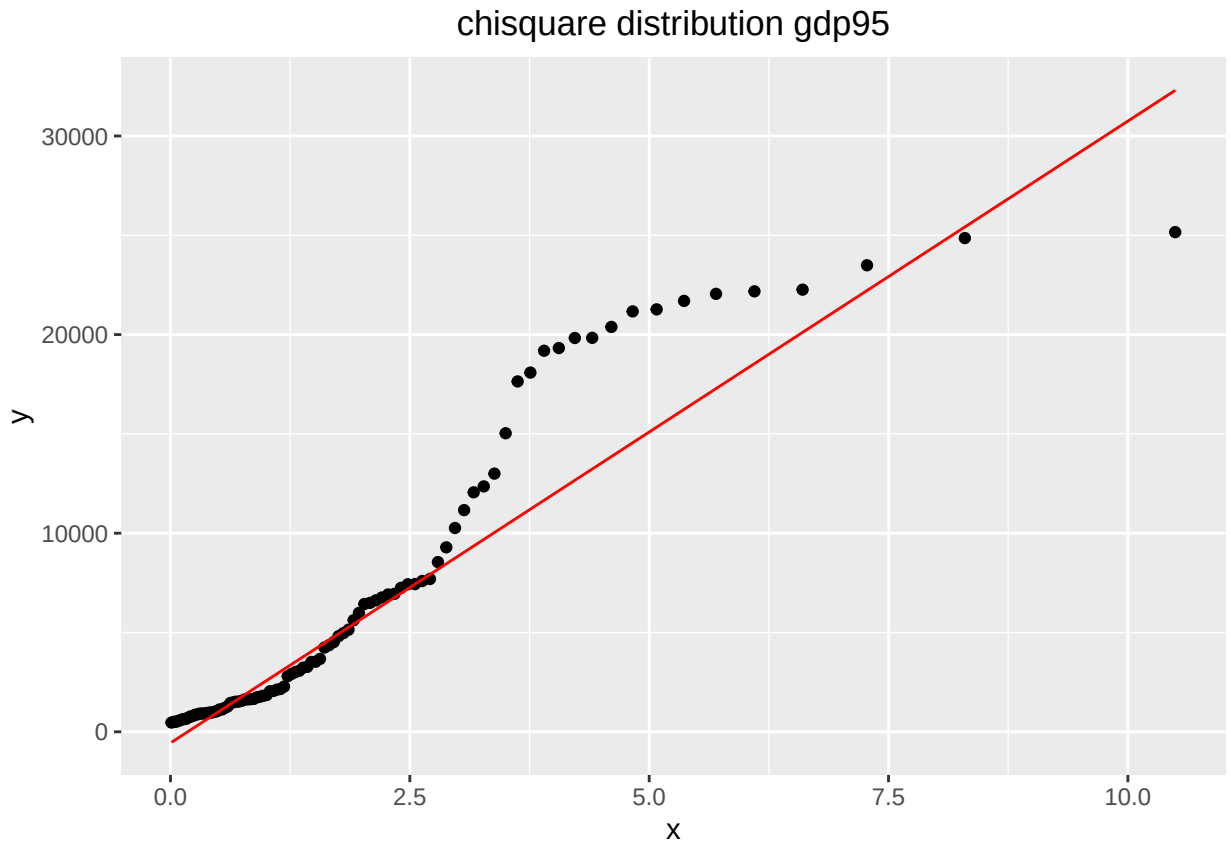

```
ggplot(datasetku,aes(sample = `gdp95`)) +
  geom_qq(distribution = qexp) +
  geom_qq_line(distribution = qexp,col='red')+
  labs(title="exponential distribution gdp95")+
  theme(plot.title = element_text(hjust=0.5))
```




```
der_chi <- 2

ggplot(datasetku,aes(sample = `gdp90`)) +
  geom_qq(distribution = qchisq,dparams = list(df=der_chi)) +
  geom_qq_line(distribution = qchisq,dparams = list(df=der_chi),col='red')+
  labs(title="chisquare distribution gdp90")+
  theme(plot.title = element_text(hjust=0.5))
```

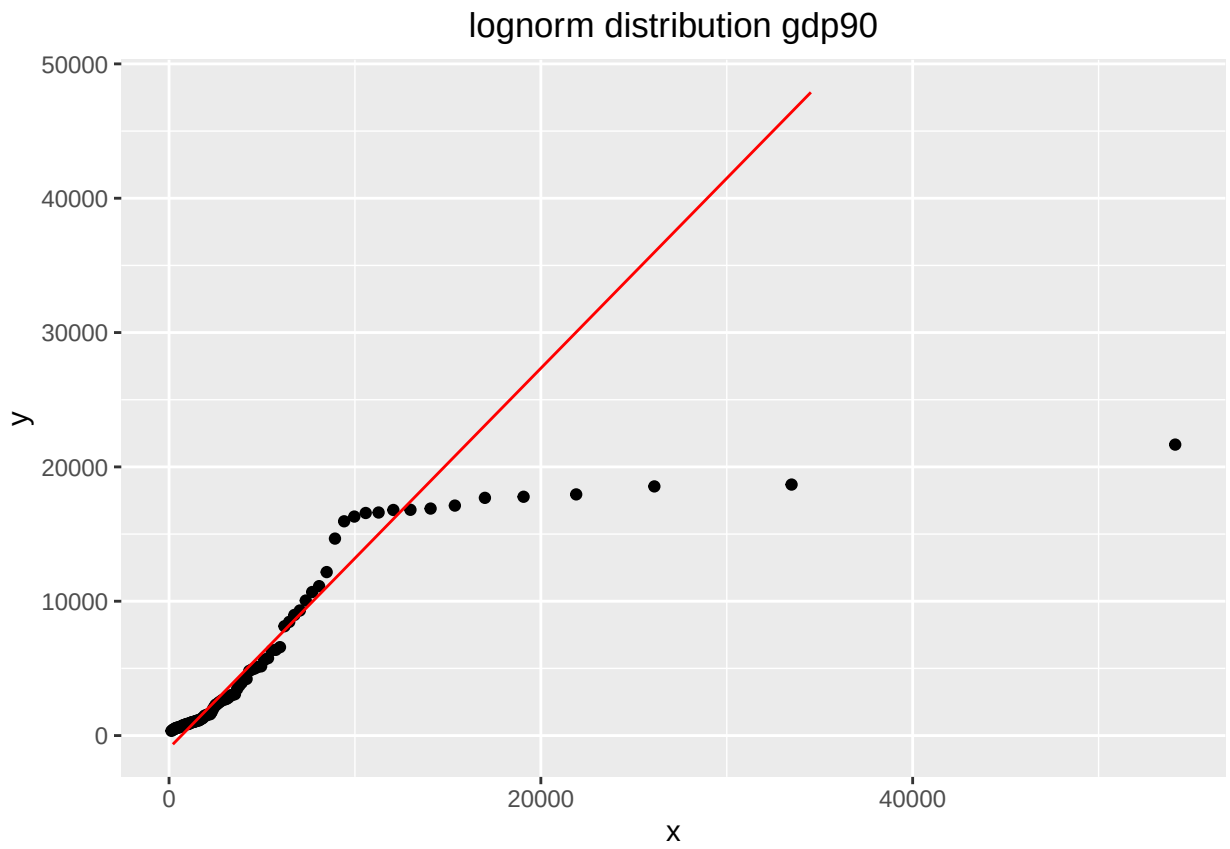



```
ggplot(datasetku,aes(sample = `gdp95`)) +
  geom_qq(distribution = qchisq,dparams = list(df=der_chi)) +
  geom_qq_line(distribution = qchisq,dparams = list(df=der_chi),col='red')+
  labs(title="chisquare distribution gdp95")+
  theme(plot.title = element_text(hjust=0.5))
```



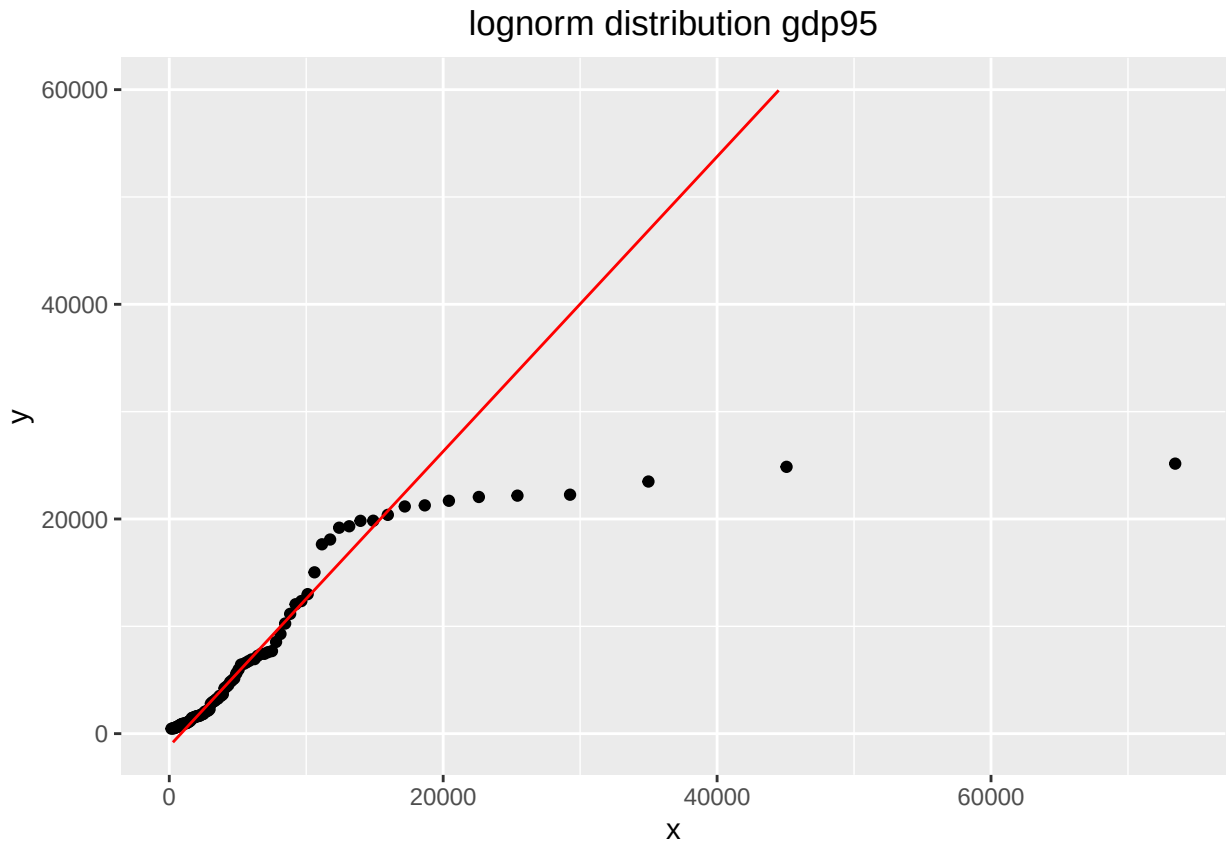

```
ggplot(datasetku,aes(sample = `gdp90`)) +
  geom_qq(distribution = qlnorm,
          dparams = list(meanlog=mean(log(datasetku$gdp90)),
                         sd = sd(log(datasetku$gdp90)))) +
  geom_qq_line(distribution = qlnorm,
               dparams = list(meanlog=mean(log(datasetku$gdp90))),
                         sd = sd(log(datasetku$gdp90)),col='red')+
  labs(title="lognorm distribution gdp90")+
  theme(plot.title = element_text(hjust=0.5))
```

```
## Warning: Ignoring unknown parameters: sd
```




```
ggplot(datasetku,aes(sample = `gdp95`)) +
  geom_qq(distribution = qlnorm,
          dparams = list(meanlog=mean(log(datasetku$gdp95)),
                         sd = sd(log(datasetku$gdp95)))) +
  geom_qq_line(distribution = qlnorm,
               dparams = list(meanlog=mean(log(datasetku$gdp95))),
                         sd = sd(log(datasetku$gdp95)),col='red')+
  labs(title="lognorm distribution gdp95")+
  theme(plot.title = element_text(hjust=0.5))
```

```
## Warning: Ignoring unknown parameters: sd
```



dites dengan berbagai ,kemungkinan yang terbaik antara lognorm, chisquare hingga exponensial , maka diperlukan ks.test atau kolmogorov sminorv test

```
print("lognorm")
```

```
## [1] "lognorm"
```

```
ks.test(datasetku$gdp90,"plnorm",mean(log(datasetku$gdp90)),sd(log(datasetku$gdp90)))
```

```
## Warning in ks.test(datasetku$gdp90, "plnorm", mean(log(datasetku$gdp90)), : ties
## should not be present for the Kolmogorov-Smirnov test

##
##  One-sample Kolmogorov-Smirnov test
##
## data:  datasetku$gdp90
## D = 0.11254, p-value = 0.1802
## alternative hypothesis: two-sided
```

dengan dilakukan lognormal, maka hasilnya yang cocok adalah lognormal ya ini untuk variable gdp90

```
print("lognorm")
```

```
## [1] "lognorm"
```

```
ks.test(datasetku$gdp95,"plnorm",mean(log(datasetku$gdp95)),sd(log(datasetku$gdp95)))
```

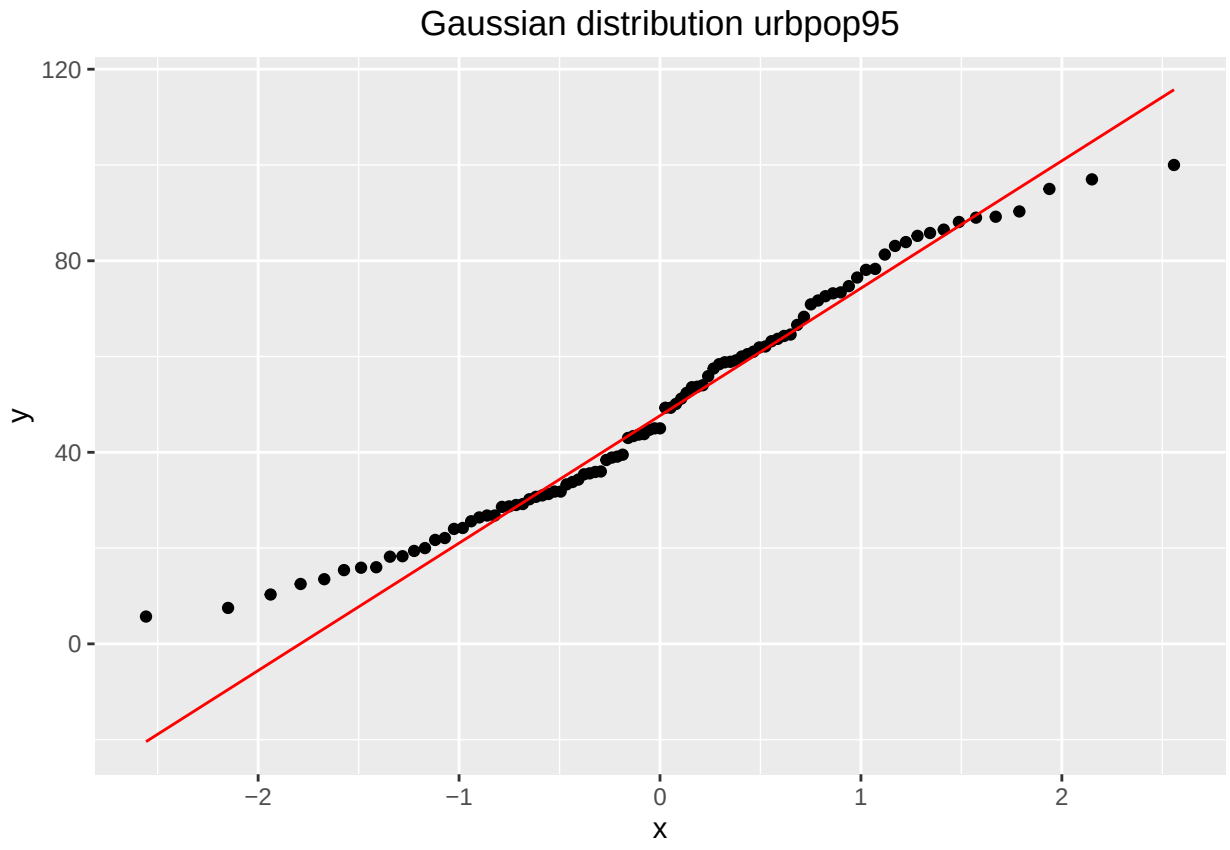
```
##
##  One-sample Kolmogorov-Smirnov test
##
## data:  datasetku$gdp95
## D = 0.098521, p-value = 0.2953
## alternative hypothesis: two-sided
```

begitu juga dengan, maka hasilnya yang cocok adalah persebaran lognormal ya ini untuk variable gdp95

```
mean(datasetku$urbpop95)
```

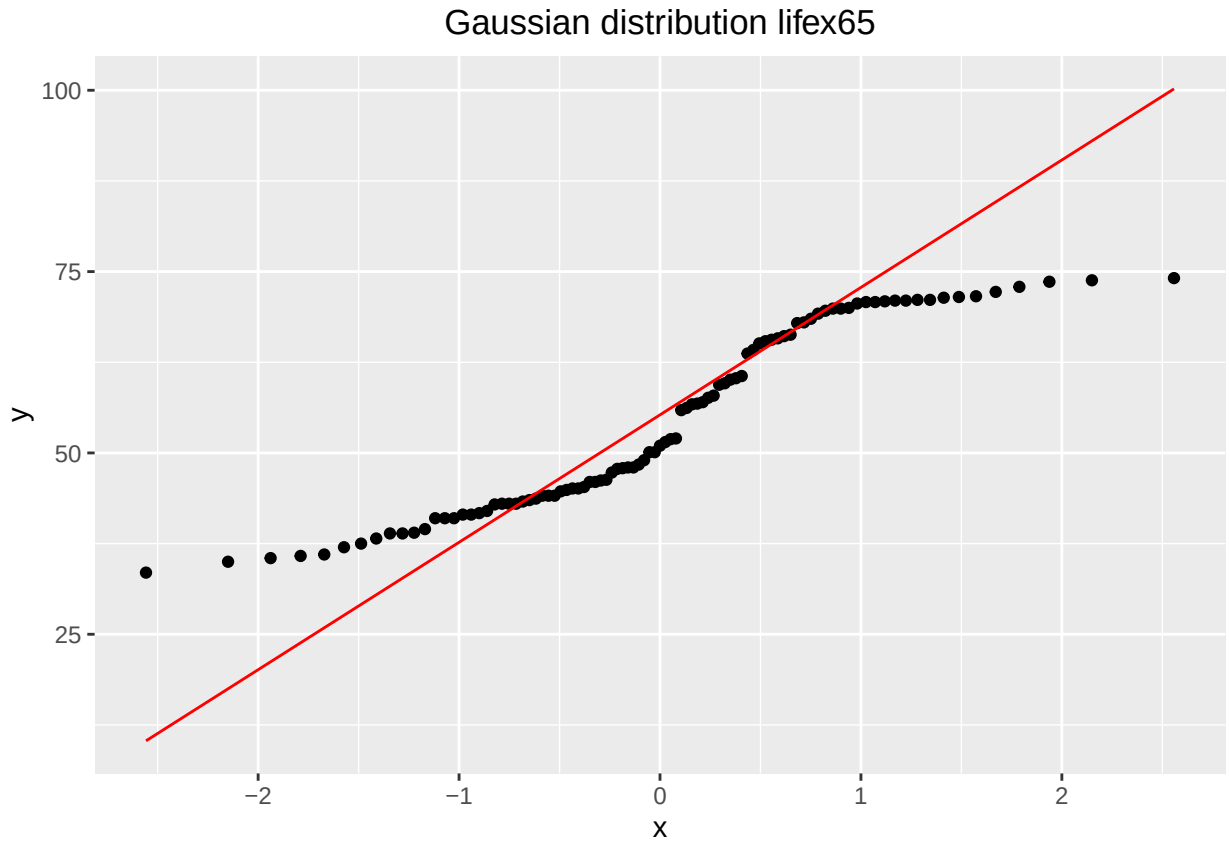
```
## [1] 49.09158
```

```
ggplot(datasetku,aes(sample = `urbpop95`)) +
  geom_qq() +
  geom_qq_line(col='red')+
  labs(title="Gaussian distribution urbpop95")+
  theme(plot.title = element_text(hjust=0.5))
```




```
ggplot(datasetku,aes(sample = `lifex65`)) +
  geom_qq() +
```

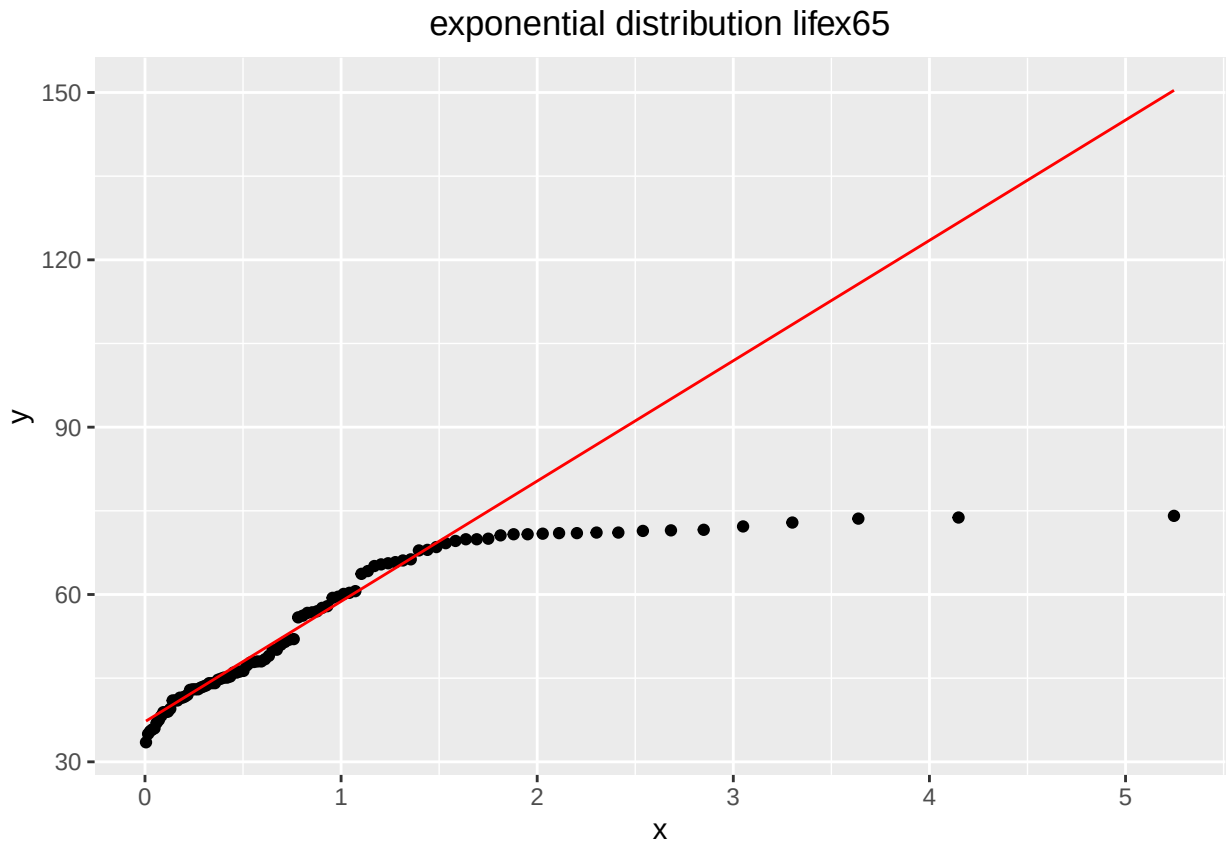
```
  geom_qq_line(col='red')+
  labs(title="Gaussian distribution lifex65")+
  theme(plot.title = element_text(hjust=0.5))
```



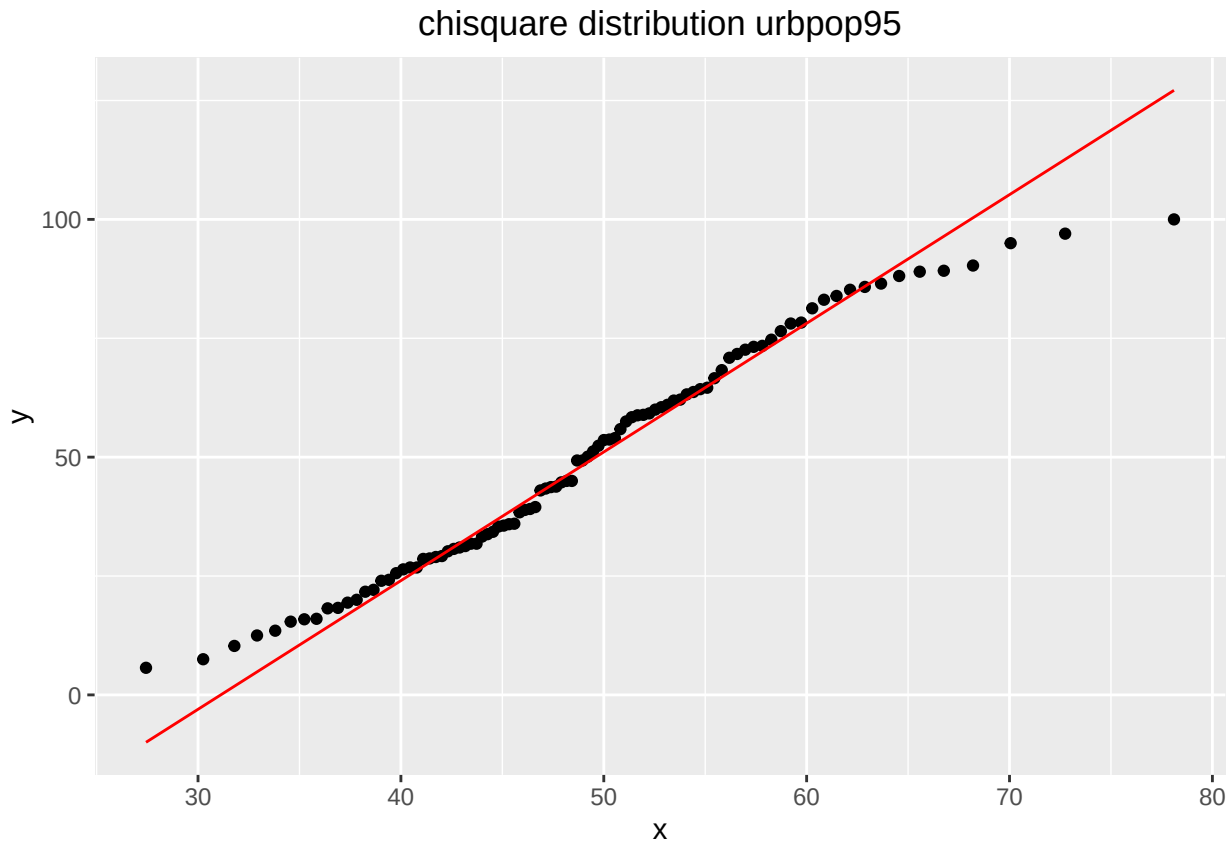
urbpop95 lebih cocok ke lognormal ya

```
rate1 <- mean(datasetku$urbpop95)
rate2 <- mean(datasetku$lifex65)

ggplot(datasetku,aes(sample = `urbpop95`)) +
  geom_qq(distribution = qexp) +
  geom_qq_line(distribution = qexp,col='red')+
  labs(title="exponential distribution urbpop95")+
  theme(plot.title = element_text(hjust=0.5))
```

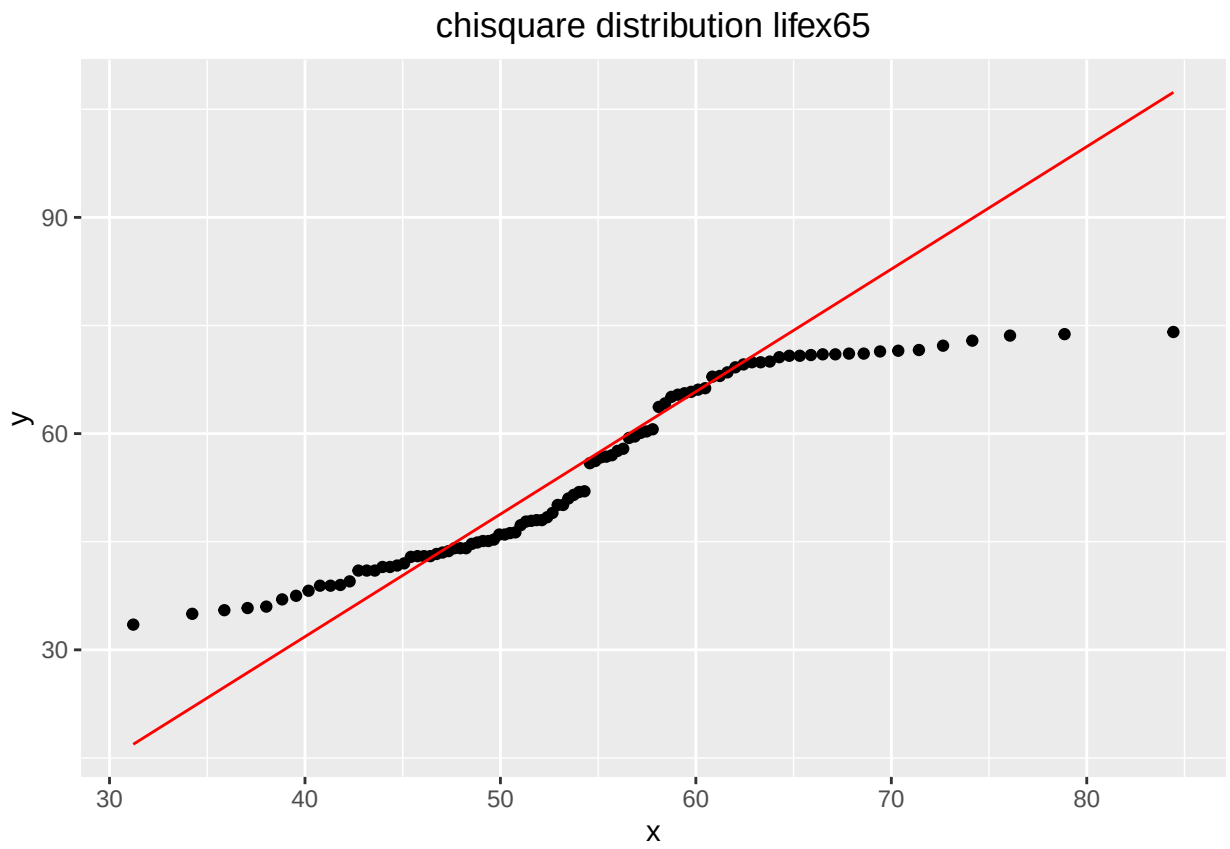

```
ggplot(datasetku,aes(sample = `lifex65`)) +
  geom_qq(distribution = qexp) +
  geom_qq_line(distribution = qexp,col='red')+
  labs(title="exponential distribution lifex65")+
  theme(plot.title = element_text(hjust=0.5))
```




```
ggplot(datasetku,aes(sample = `urbpop95`)) +
  geom_qq(distribution = qchisq,dparams = list(df=rate1)) +
  geom_qq_line(distribution = qchisq,dparams = list(df=rate1),col='red')+
  labs(title="chisquare distribution urbpop95")+
  theme(plot.title = element_text(hjust=0.5))
```



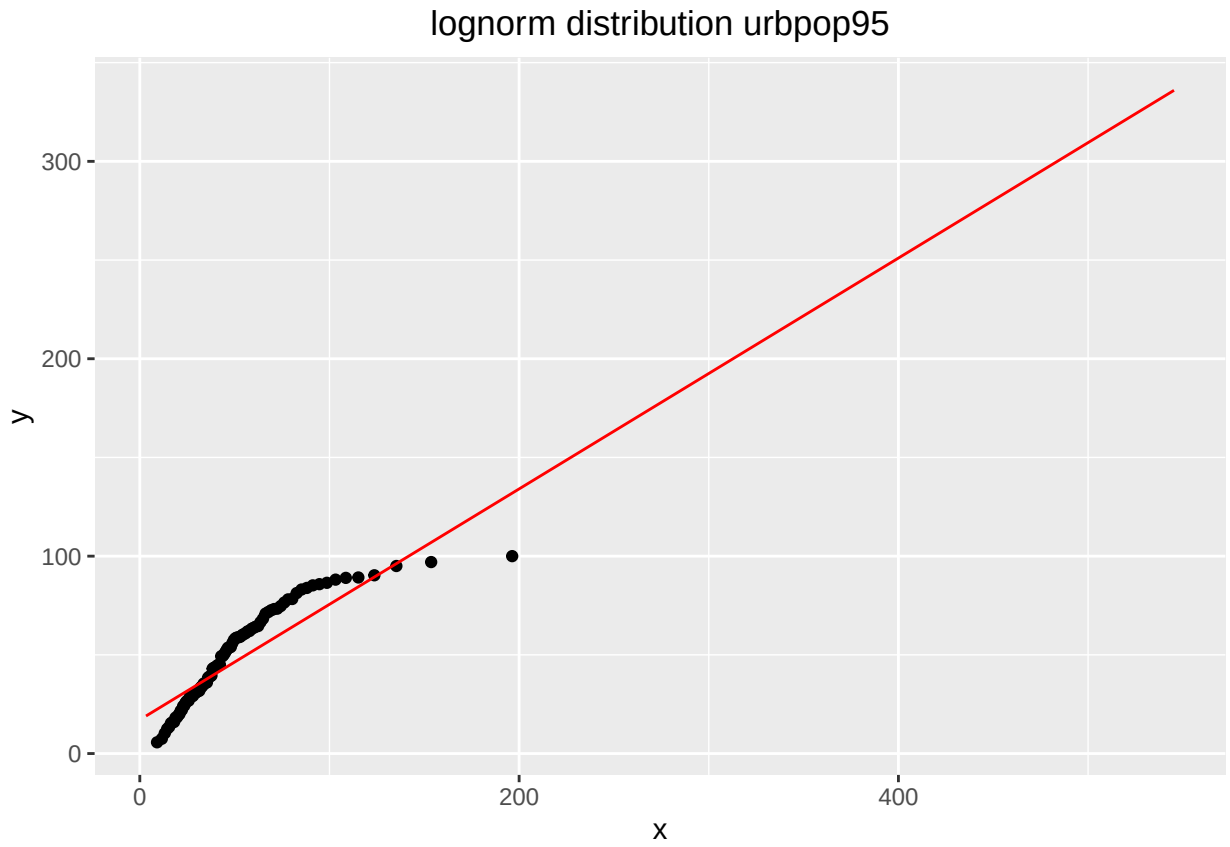

```
ggplot(datasetku,aes(sample = `lifex65`)) +
  geom_qq(distribution = qchisq,dparams = list(df=rate2)) +
  geom_qq_line(distribution = qchisq,dparams = list(df=rate2),col='red')+
  labs(title="chisquare distribution lifex65")+
  theme(plot.title = element_text(hjust=0.5))
```



urbpop lebih cocok juga chisquare juga yaaa

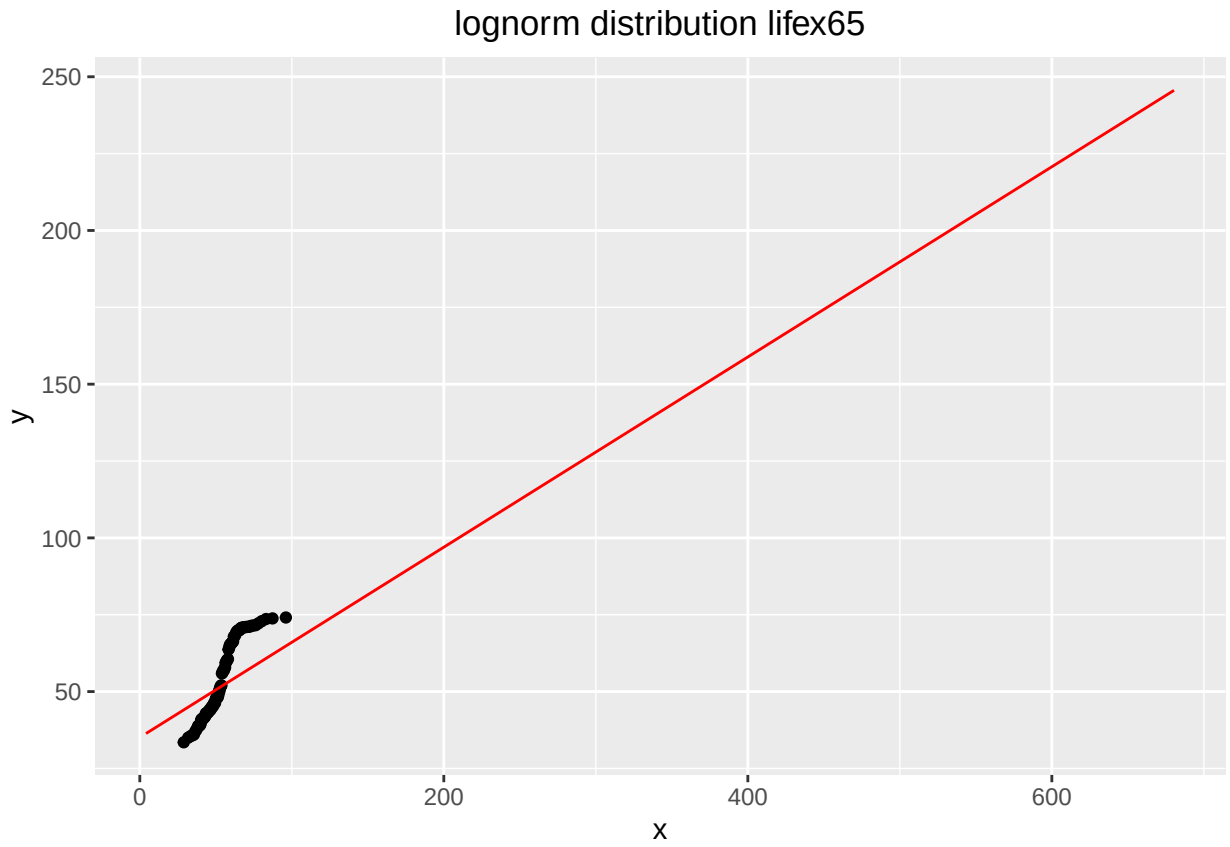
```
ggplot(datasetku,aes(sample = `urbpop95`)) +
  geom_qq(distribution = qlnorm,
          dparams = list(meanlog=mean(log(datasetku$urbpop95)),
                         sd = sd(log(datasetku$urbpop95)))) +
  geom_qq_line(distribution = qlnorm,
               dparams = list(meanlog=mean(log(datasetku$urbpop95))),
                         sd = sd(log(datasetku$urbpop95)),col='red')+
  labs(title="lognorm distribution urbpop95")+
  theme(plot.title = element_text(hjust=0.5))
```

```
## Warning: Ignoring unknown parameters: sd
```




```
ggplot(datasetku,aes(sample = `lifex65`)) +
  geom_qq(distribution = qlnorm,
          dparams = list(meanlog=mean(log(datasetku$lifex65)),
                         sd = sd(log(datasetku$lifex65)))) +
  geom_qq_line(distribution = qlnorm,
               dparams = list(meanlog=mean(log(datasetku$lifex65))),
                         sd = sd(log(datasetku$lifex65)),col='red')+
  labs(title="lognorm distribution lifex65")+
  theme(plot.title = element_text(hjust=0.5))
```

```
## Warning: Ignoring unknown parameters: sd
```



lognormal malah tidak cocok,

```
ks.test(datasetku$urbpop95,"pnorm",mean(datasetku$urbpop95),sd(datasetku$urbpop95))
```

```
## Warning in ks.test(datasetku$urbpop95, "pnorm", mean(datasetku$urbpop95), : ties
## should not be present for the Kolmogorov-Smirnov test
```

```
##
##  One-sample Kolmogorov-Smirnov test
##
## data:  datasetku$urbpop95
## D = 0.095978, p-value = 0.3457
## alternative hypothesis: two-sided
```

untuk urbpop95 ya cocok ke distribusi normal

```
ks.test(datasetku$lifex65,"plnorm",mean(log(datasetku$lifex65)),sd(log(datasetku$lifex65)))
```

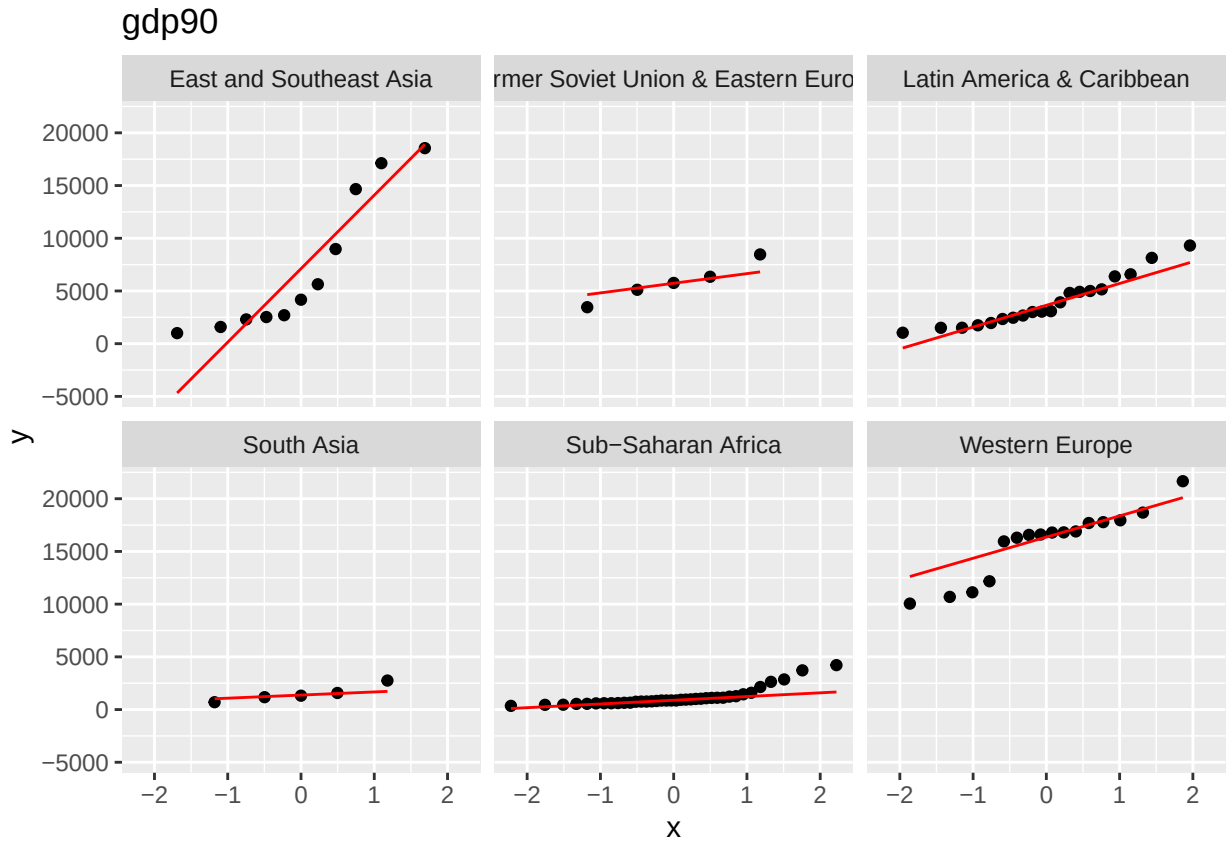
```
## Warning in ks.test(datasetku$lifex65, "plnorm", mean(log(datasetku$lifex65)), :
## ties should not be present for the Kolmogorov-Smirnov test
```

```
##
##  One-sample Kolmogorov-Smirnov test
##
```

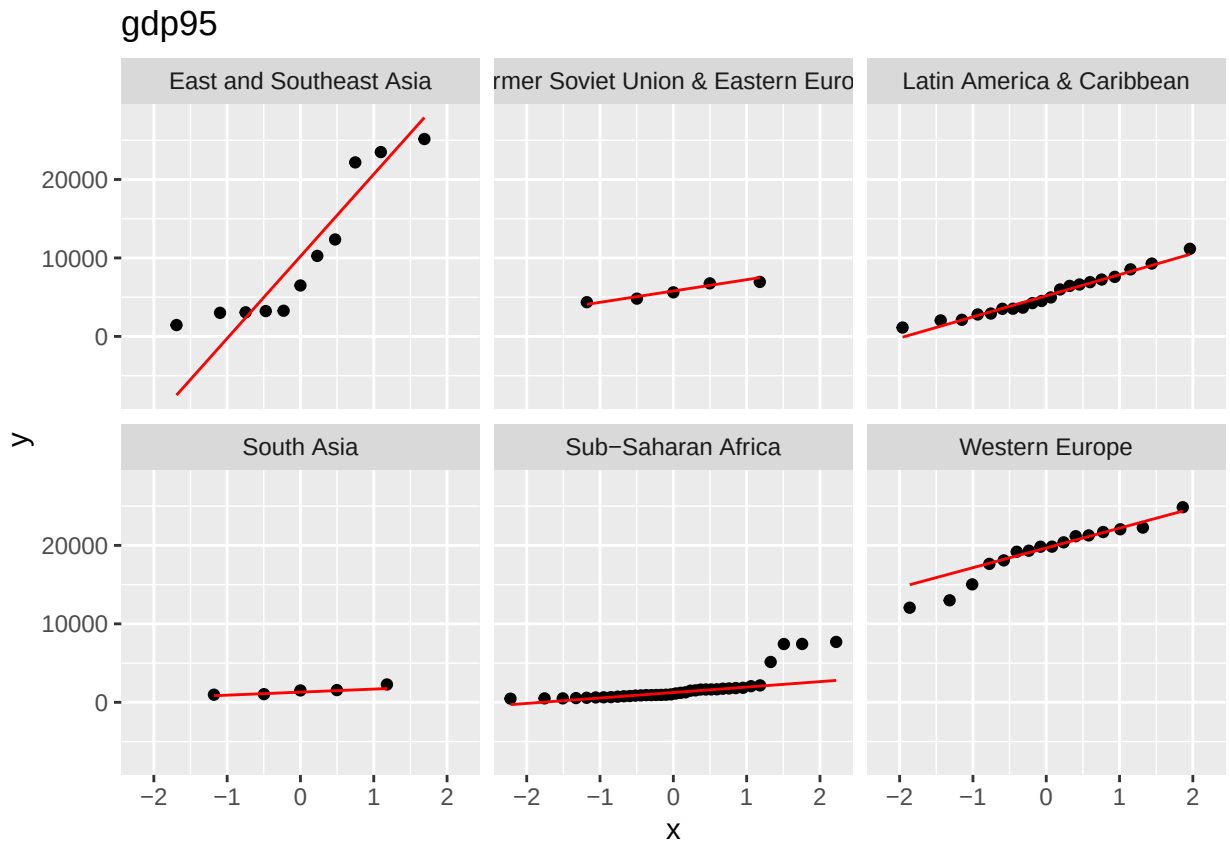
```
## data:  datasetku$lifex65
## D = 0.13175, p-value = 0.07389
## alternative hypothesis: two-sided
```

Disini lebih cocok digunakan lognormal dibandingkan yang lain yaaa

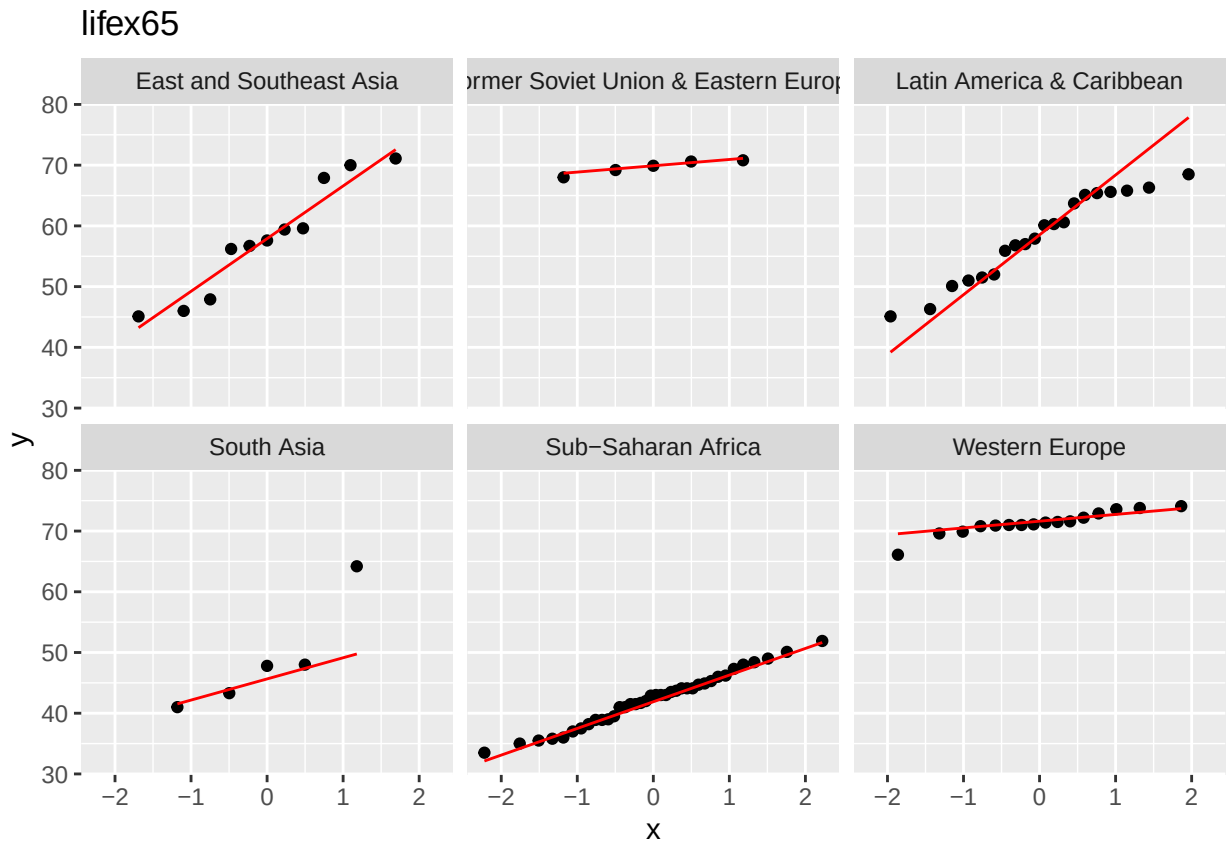
```
ggplot(datasetku,aes(sample = `gdp90`)) +
  geom_qq() +
  geom_qq_line(col='red') + facet_wrap(~`region`) + labs(title='gdp90')
```



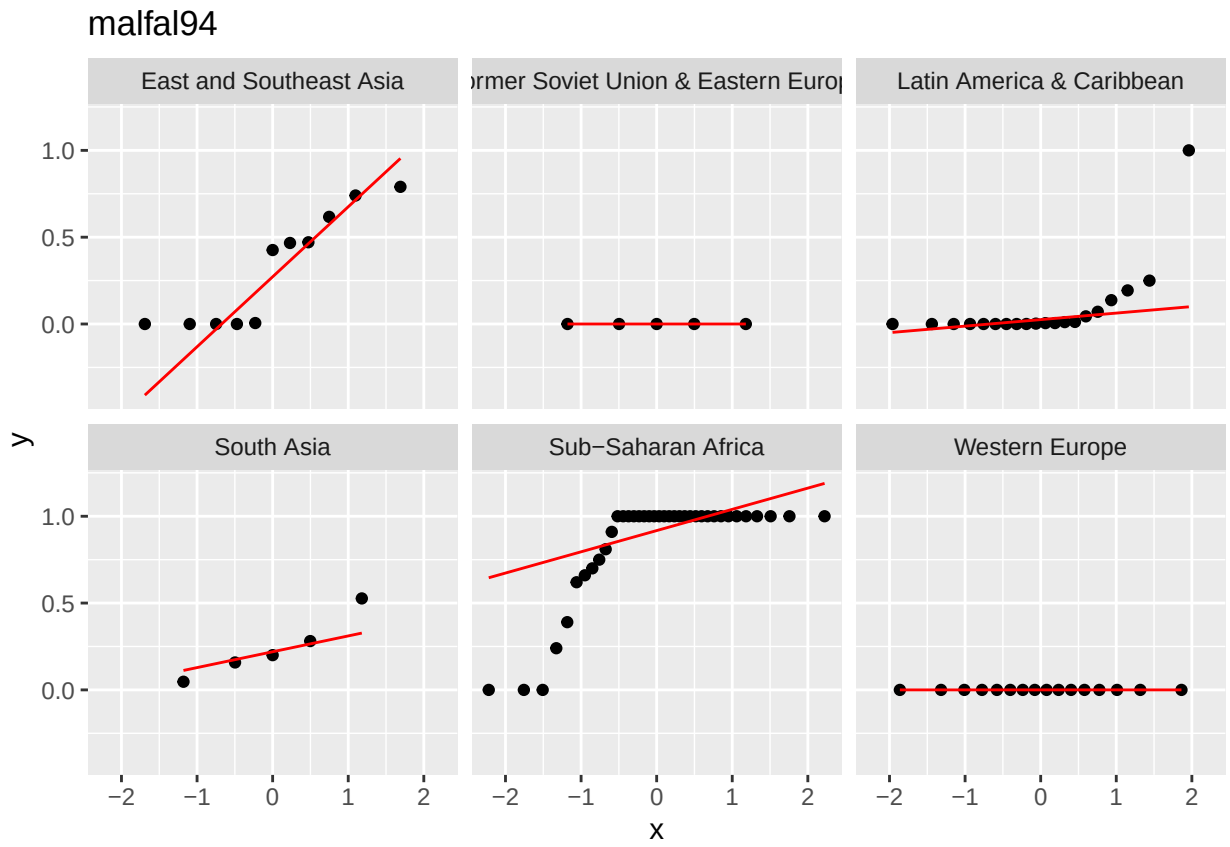

```
ggplot(datasetku,aes(sample = `gdp95`)) +
  geom_qq() +
  geom_qq_line(col='red') + facet_wrap(~`region`)+ labs(title='gdp95')
```



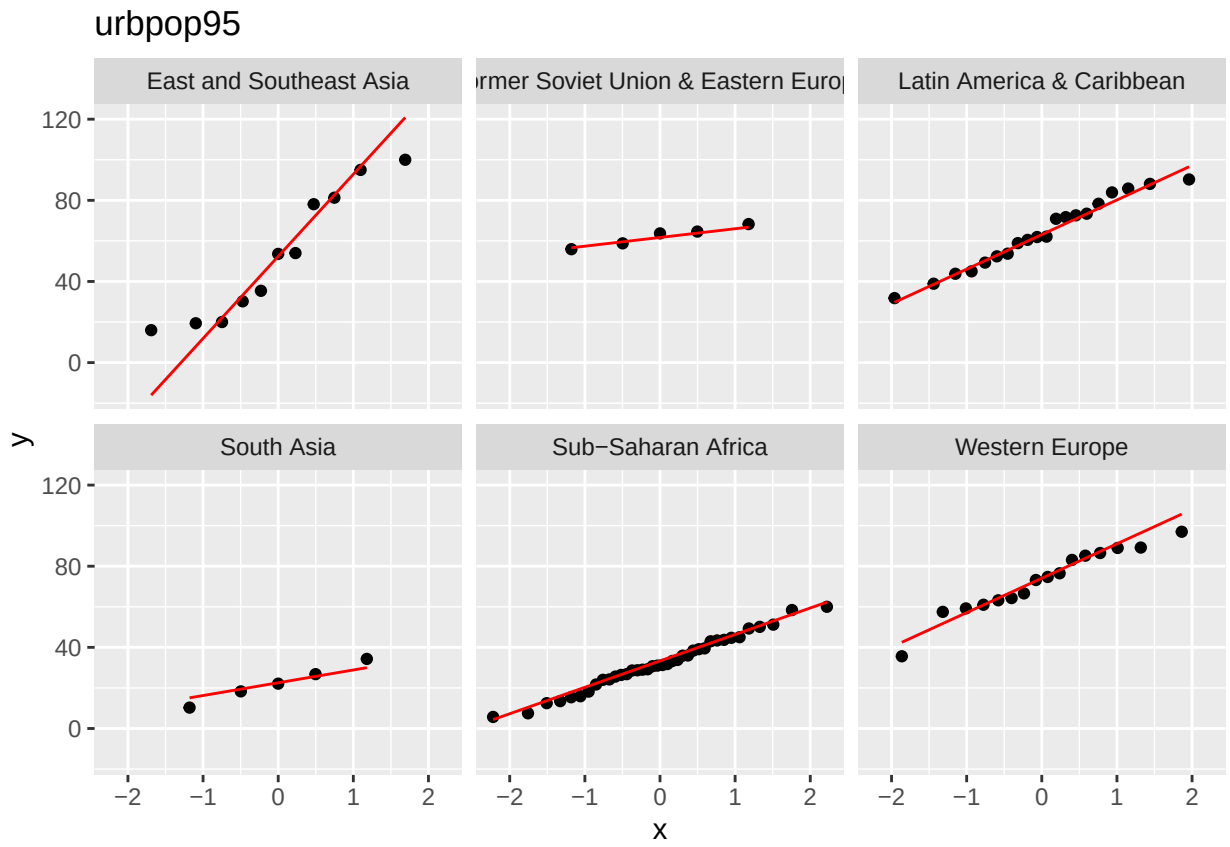

```
ggplot(datasetku,aes(sample = `lifex65`)) +
  geom_qq() +
  geom_qq_line(col='red') + facet_wrap(~`region`)+ labs(title='lifex65')
```




```
ggplot(datasetku,aes(sample = `malfal94`)) +
  geom_qq() +
  geom_qq_line(col='red') + facet_wrap(~`region`)+ labs(title='malfal94')
```

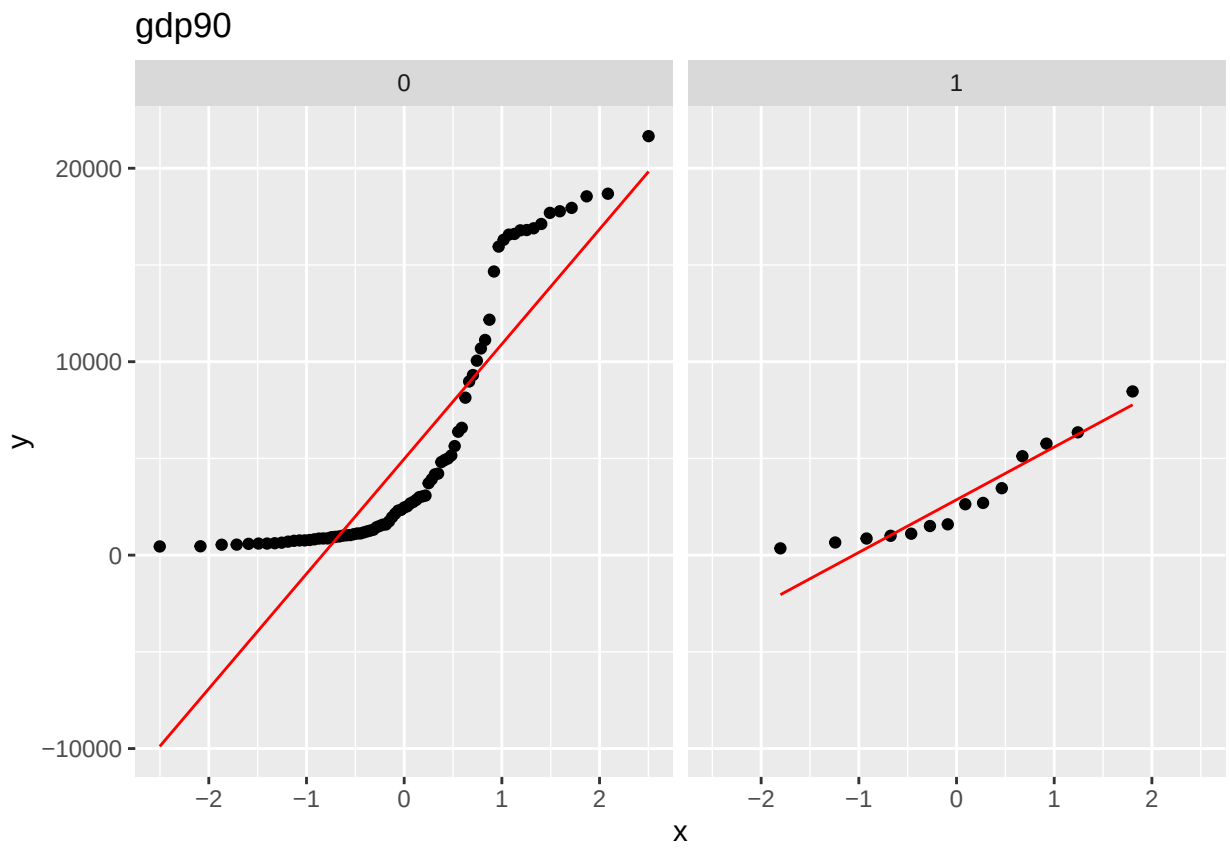



```
ggplot(datasetku,aes(sample = `urbpop95`)) +
  geom_qq() +
  geom_qq_line(col='red') + facet_wrap(~`region`)+ labs(title='urbpop95')
```

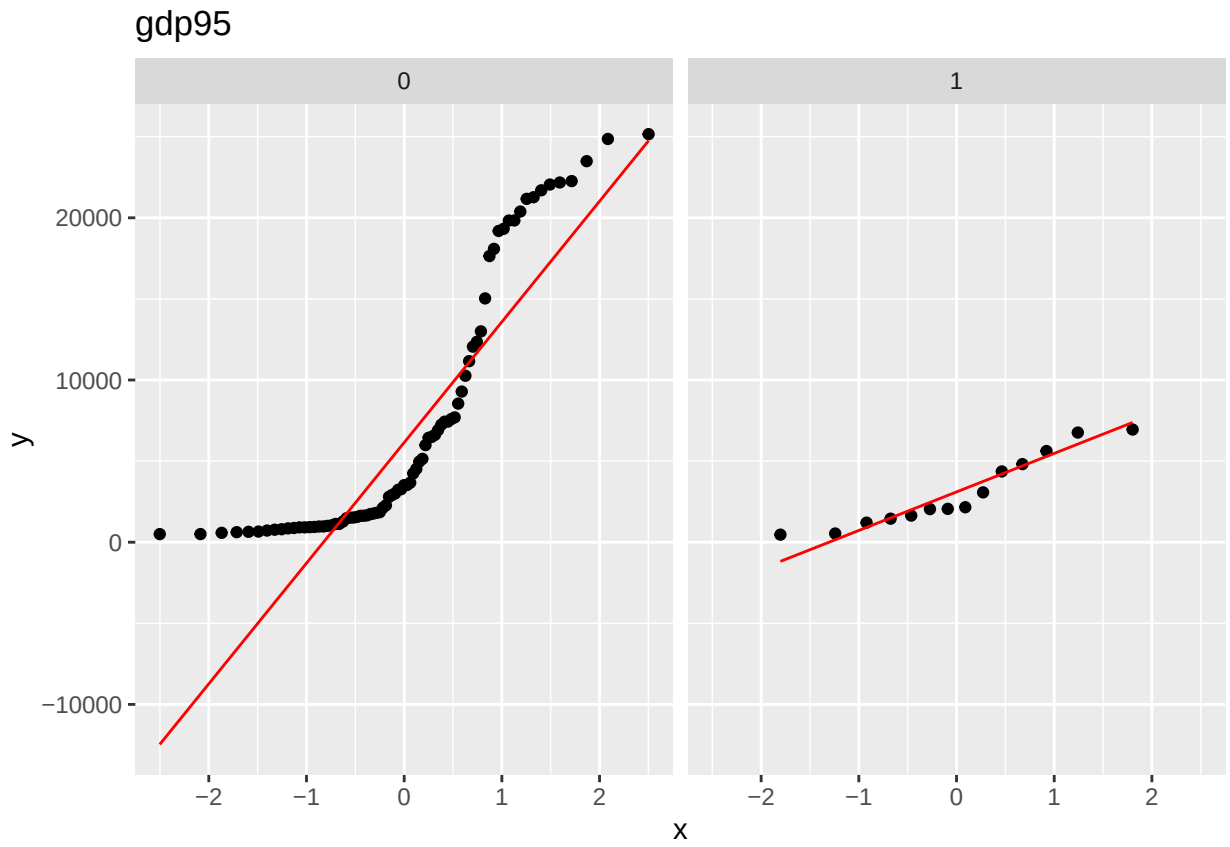


jika dilihat qqplot maka ada nilai yang menyebar normal semua dari kelima dan semua grid kecuali pada maltex94 pada africa yaaa, ada suatu kendala disini ya

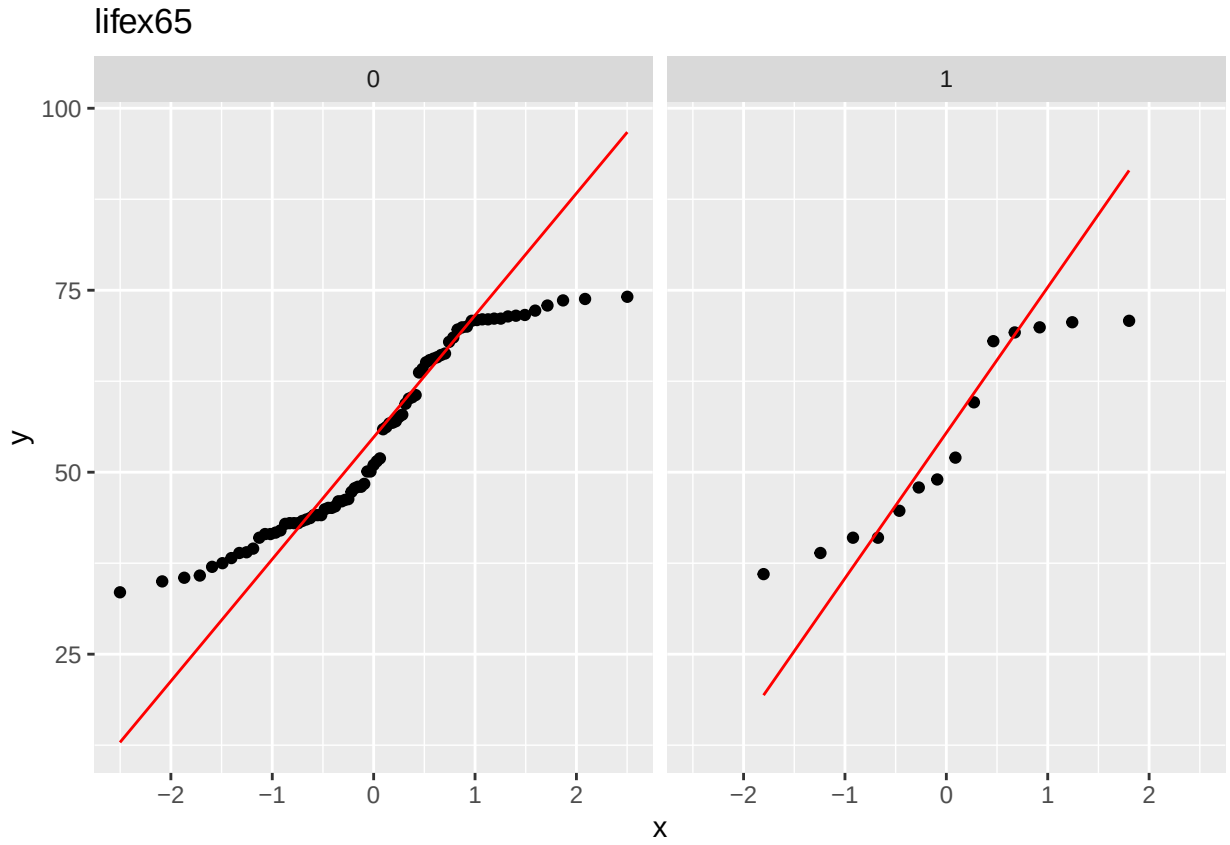
```
ggplot(datasetku,aes(sample = `gdp90`)) +
  geom_qq()+geom_qq_line(col='red') + facet_wrap(~`socialst`) + labs(title='gdp90')
```



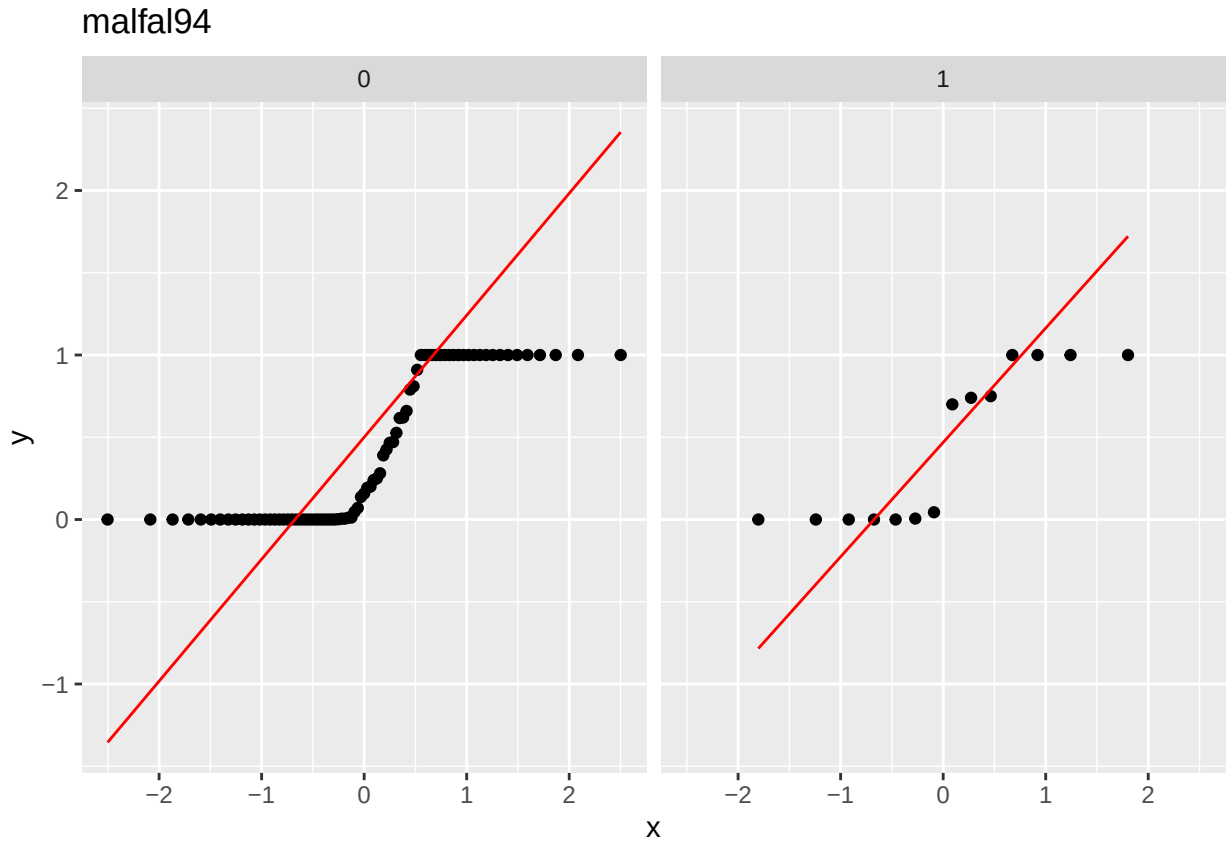

```
ggplot(datasetku,aes(sample = `gdp95`)) +
  geom_qq() +
  geom_qq_line(col='red') + facet_wrap(~`socialst`)+ labs(title='gdp95')
```



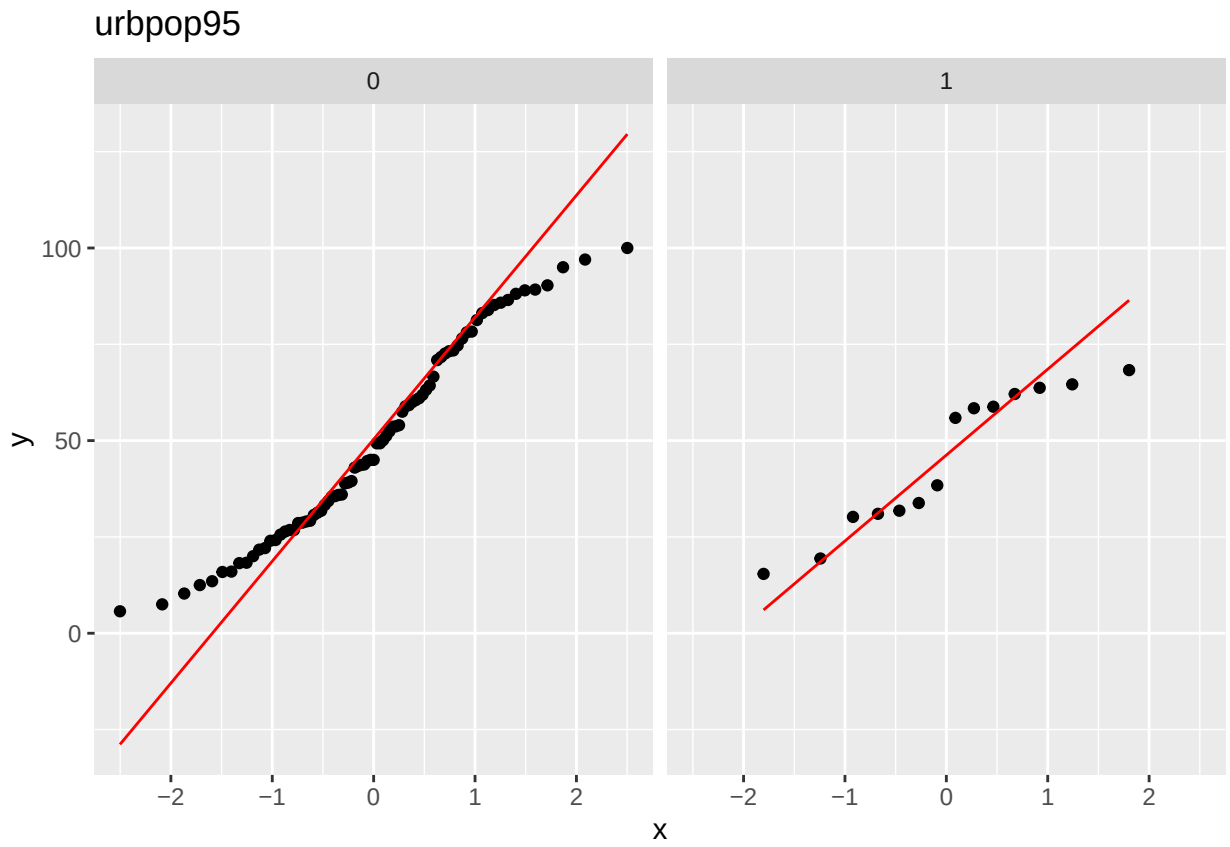

```
ggplot(datasetku,aes(sample = `lifex65`)) +
  geom_qq() +
  geom_qq_line(col='red') + facet_wrap(~`socialst`)+ labs(title='lifex65')
```




```
ggplot(datasetku,aes(sample = `malfal94`)) +
  geom_qq() +
  geom_qq_line(col='red') + facet_wrap(~`socialst`)+ labs(title='malfal94')
```

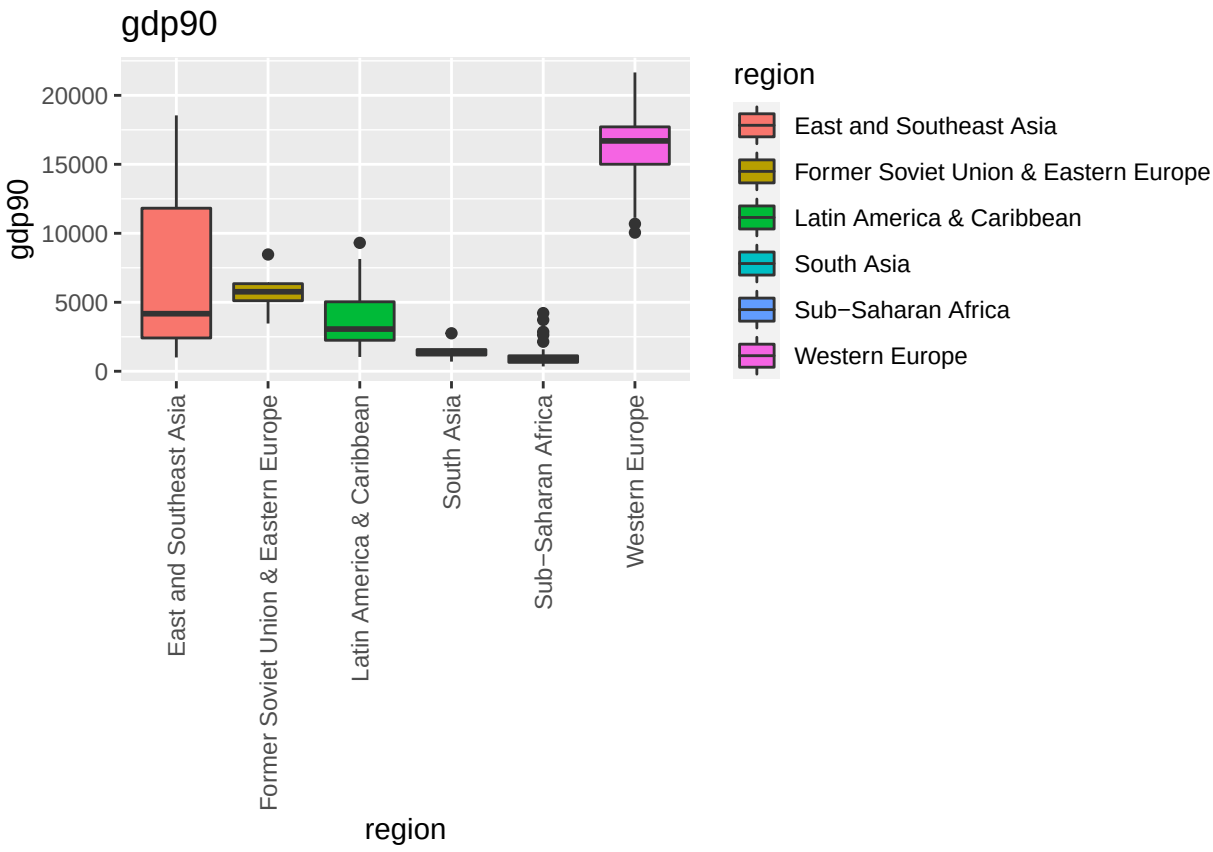



```
ggplot(datasetku,aes(sample = `urbpop95`)) +
  geom_qq() +
  geom_qq_line(col='red') + facet_wrap(~`socialst`)+ labs(title='urbpop95')
```

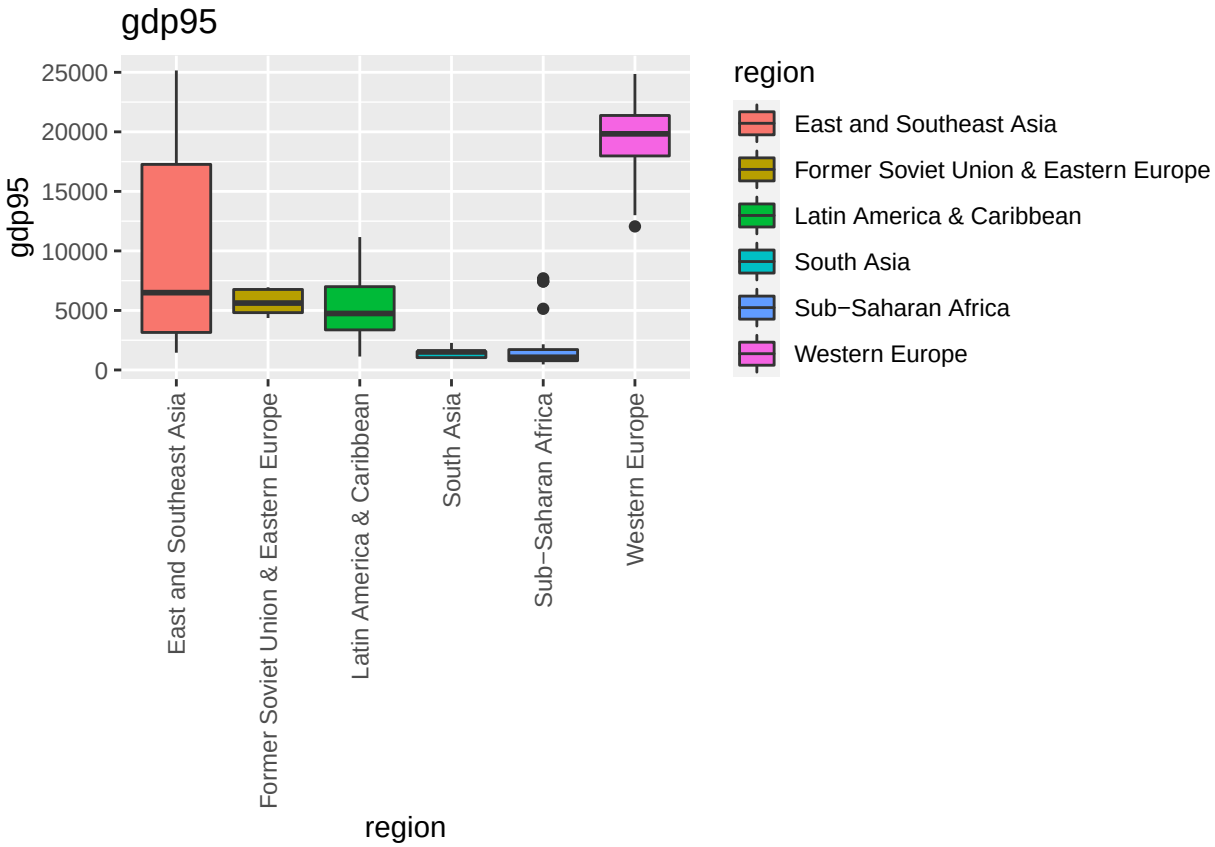


hanya pada urbpop95 mendekati normal yaaa, sehingga bisa ditinjau ulang , dikarenakan distribusi normal seharusnya dalam kejadian di statistik

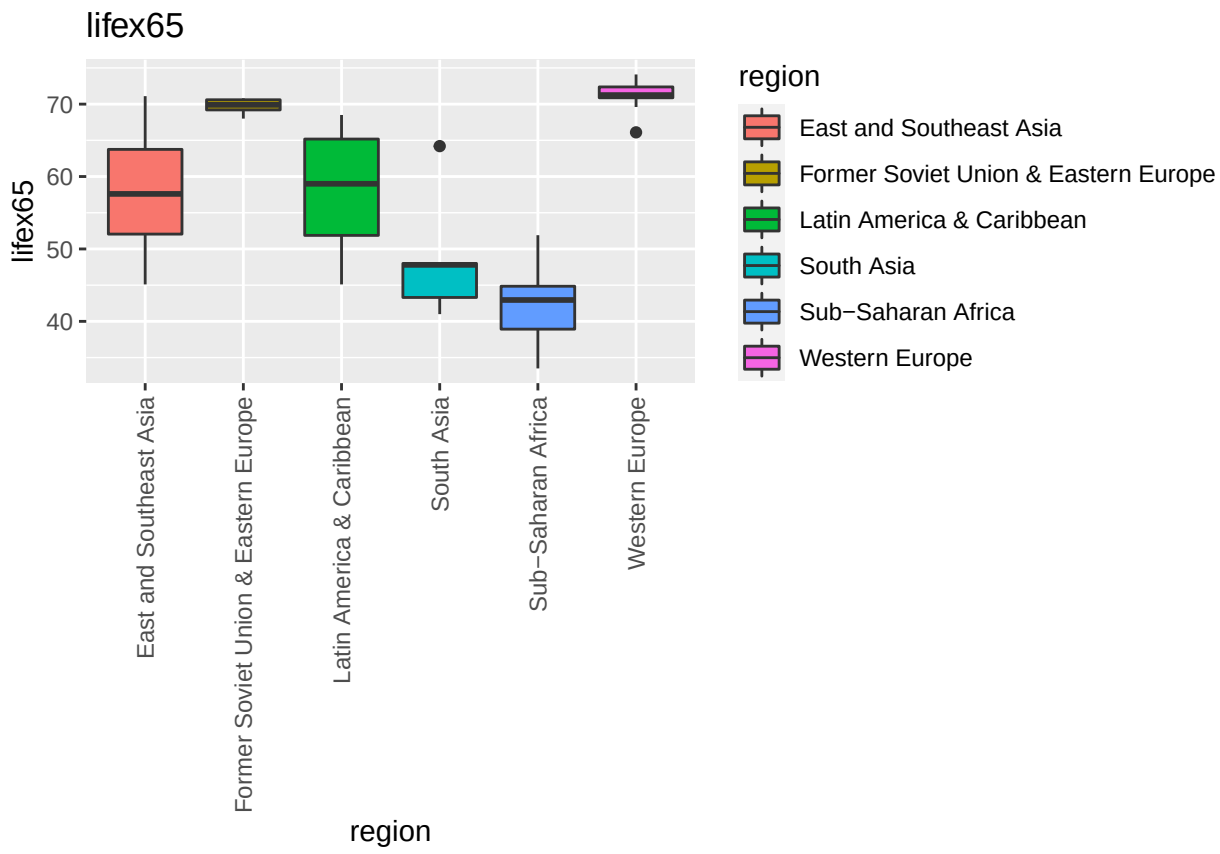
```
ggplot(datasetku,aes(x = `gdp90`,y=`region`,fill=`region`)) +
  geom_boxplot()  + labs(title='gdp90')+coord_flip() +
  theme(axis.text.x = element_text(angle = 90, vjust = 0.5, hjust=1))
```



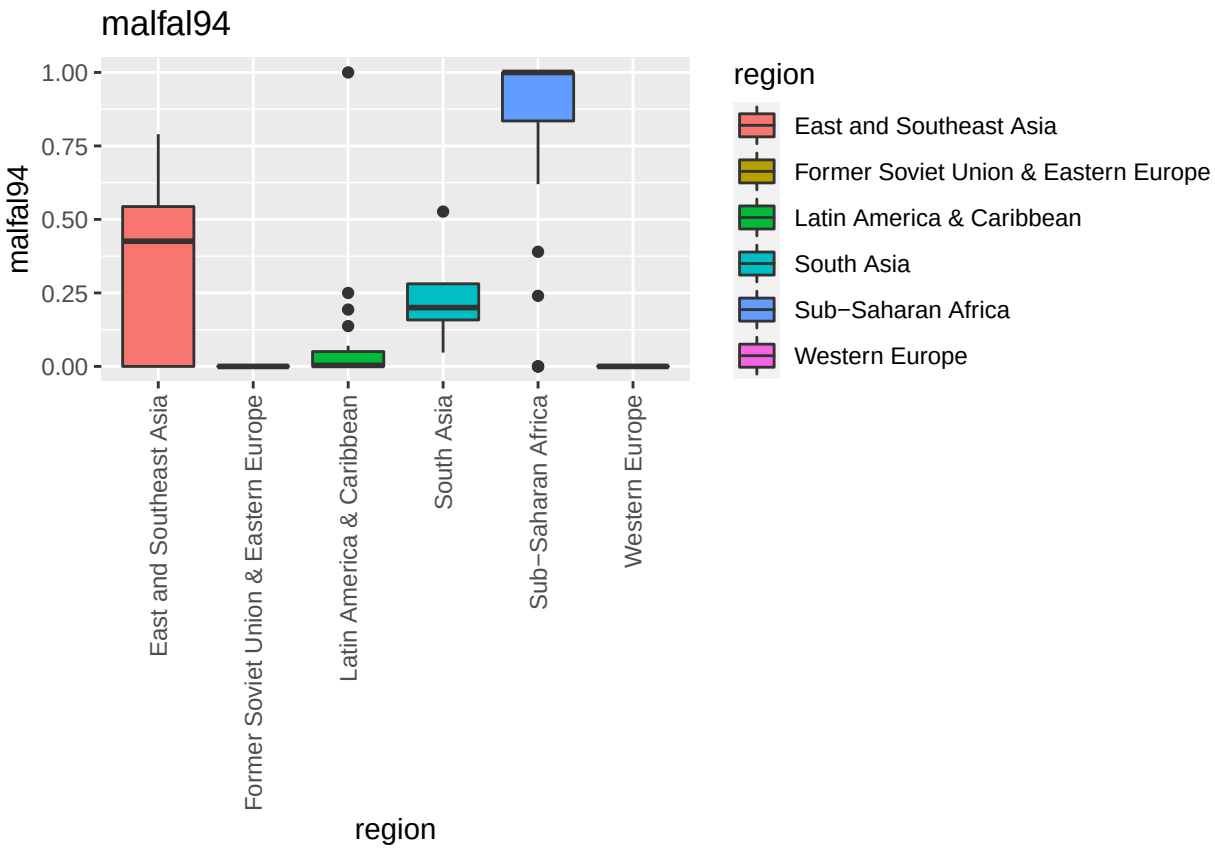

```
ggplot(datasetku,aes(x = `gdp95`,y=`region`,fill=`region`)) +
  geom_boxplot()  + labs(title='gdp95')+coord_flip()+
  theme(axis.text.x = element_text(angle = 90, vjust = 0.5, hjust=1))
```



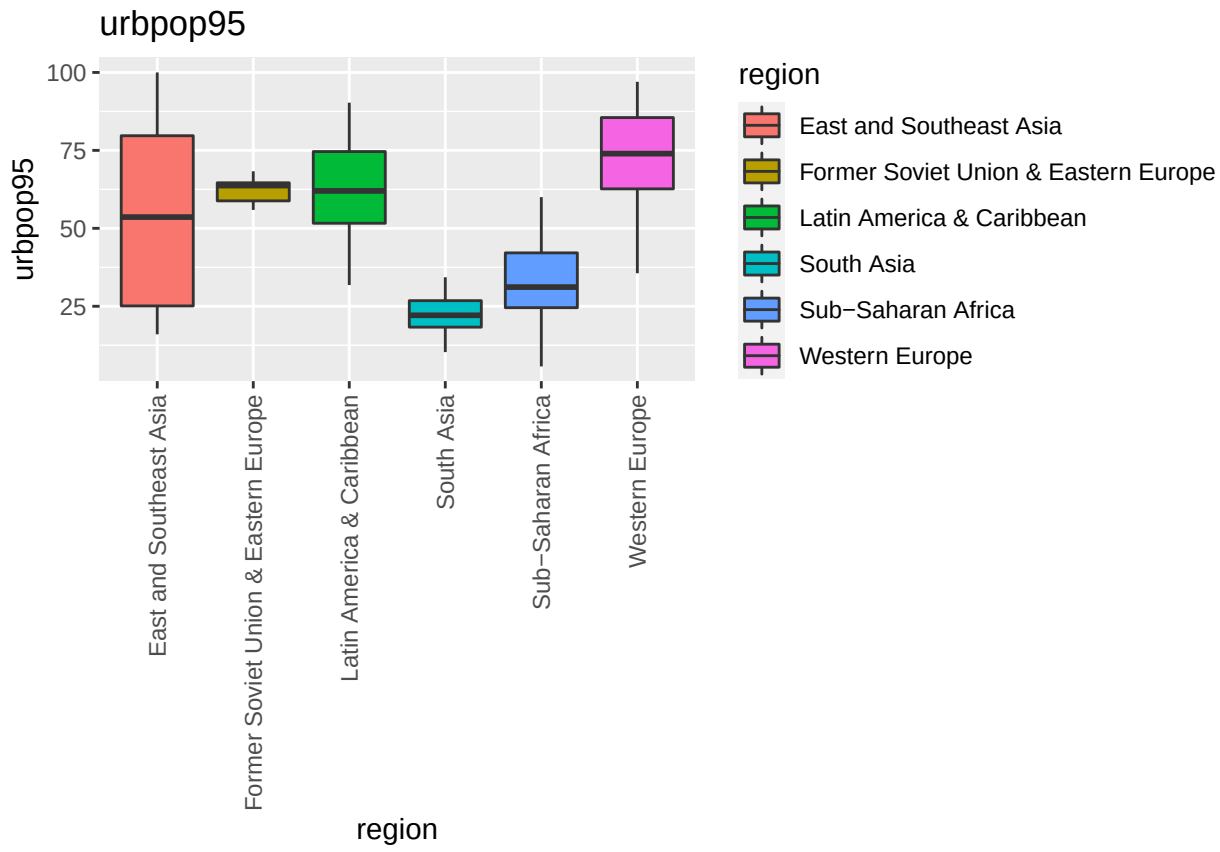

```
ggplot(datasetku,aes(x = `lifex65`,y=`region`,fill=`region`)) +
  geom_boxplot()  + labs(title='lifex65')+coord_flip()+
  theme(axis.text.x = element_text(angle = 90, vjust = 0.5, hjust=1))
```




```
ggplot(datasetku,aes(x = `malfal94`,y=`region`,fill=`region`)) +
  geom_boxplot()  + labs(title='malfal94')+coord_flip()+
theme(axis.text.x = element_text(angle = 90, vjust = 0.5, hjust=1))
```




```
ggplot(datasetku,aes(x = `urbpop95`,y=`region`,fill=`region`)) +
  geom_boxplot()  + labs(title='urbpop95')+coord_flip()+
theme(axis.text.x = element_text(angle = 90, vjust = 0.5, hjust=1))
```



pencilan di gdp 90 dibanding gdp 95, urbpop 95 bagus nah ini kenapa ,

ini dikarenakan waktu jadi kurang lengkap , mohon maafkan saya apabila kurang lengkap