

# World Happiness Report 2023

Mia Gmiza, Gabriijela Perković, Matija Roginić, Erika Tomakić

2023-12-06

## Uvod

\*dodati opis naseg zadatka

## Deskriptivna analiza

Učitavanje podataka.

```
opis_var = read.csv("datasets/opis_varijabli.csv")
WHR_22 = read.csv("datasets/WHR_2022.csv")
WHR_22 = head(WHR_22, -1) # preskacem zadnji red jer je "xx"
WHR_23 = read.csv("datasets/WHR_2023.csv")
```

Podatci za 2022. godinu sastoje se od 146 država i dvije varijable. Podatci za 2023. godinu sastoje se od 137 država i 15 varijabli.

```
cat("Varijable za 2022. godinu:\n")
```

```
## Varijable za 2022. godinu:
```

```
names(WHR_22)
```

```
## [1] "Country" "Happiness.score"
```

```
cat("Varijable za 2023. godinu:\n")
```

```
## Varijable za 2023. godinu:
```

```
names(WHR_23)
```

```
## [1] "Country.name"
## [2] "Regional.indicator"
## [3] "Ladder.score"
## [4] "GDP.per.capita"
## [5] "Social.support"
## [6] "Healthy.life.expectancy"
## [7] "Freedom.to.make.life.choices"
## [8] "Generosity"
## [9] "Perceptions.of.corruption"
## [10] "Alcohol.consumption.Both.Sexes..L.year."
## [11] "Alcohol.consumption.Male..L.year."
## [12] "Alcohol.consumption.Female..L.year."
## [13] "Crime.rate.Crime.Index"
## [14] "Healthcare.Legatum.Prosperty.Index.Health.Score"
## [15] "Gini.Coefficient...World.Bank"
```

```

any(is.na(WHR_22))

## [1] FALSE
cat("U podatcima za 2022. godinu nema nedostajućih vrijednosti.\n")

## U podatcima za 2022. godinu nema nedostajućih vrijednosti.
any(is.na(WHR_23))

## [1] TRUE
cat("U podatcima za 2023. godinu ima nedostajućih vrijednosti.\n")

## U podatcima za 2023. godinu ima nedostajućih vrijednosti.
for (col_name in names(WHR_23)) {
  if (sum(is.na(WHR_23[,col_name])) > 0){
    cat('Ukupno nedostajućih vrijednosti za varijablu ', col_name, ': ', sum(is.na(WHR_23[,col_name])), '\n')
  }
}

## Ukupno nedostajućih vrijednosti za varijablu Healthy.life.expectancy : 1
## Ukupno nedostajućih vrijednosti za varijablu Alcohol.consumption.Both.Sexes..L.year. : 6
## Ukupno nedostajućih vrijednosti za varijablu Alcohol.consumption.Male..L.year. : 6
## Ukupno nedostajućih vrijednosti za varijablu Alcohol.consumption.Female..L.year. : 6
## Ukupno nedostajućih vrijednosti za varijablu Crime.rate.Crime.Index : 24
## Ukupno nedostajućih vrijednosti za varijablu Healthcare.Legatum.Prosperty.Index.Health.Score : 2
## Ukupno nedostajućih vrijednosti za varijablu Gini.Coefficient...World.Bank : 10

```

### 3. Postoje li razlike u kvaliteti zdravstvene skrbi među različitim regijama?

Prvo je potrebno provjeriti jesu li podatci normalno distribuirani. To ćemo napraviti analitički i grafički. Analitičku provjeru čini Kolmogorov-Smirnov test. Postavljamo hipoteze:

H0: podatci su normalno distribuirani

H1: podatci nisu normalno distribuirani

Prisjetimo se, varijabla "Healthcare.Legatum.Prosperty.Index.Health.Score" ima dvije nedostajuće vrijednosti. Jedan od načina na koji se to može riješiti je da svedemo nedostajuće vrijednosti na srednju vrijednost te varijable za pripadnu regiju. U ovom slučaju ne možemo raditi takvu procjenu zato što pitanje kojim se bavimo ovisi o regijama. Nedostajuće vrijednosti su za Kosovo i Palestinu, države kojima to nije jedini nedostajući podatak. Prema tome, jednostavno ćemo te dvije države ukloniti iz daljnje procjene zdravstvene skrbi.

```

data3 <- WHR_23[!is.na(WHR_23$Healthcare.Legatum.Prosperty.Index.Health.Score), ]
names(data3)[names(data3) == "Healthcare.Legatum.Prosperty.Index.Health.Score"] <- "Health.Score"
names(data3)[names(data3) == "Regional.indicator"] <- "Region"
regions <- unique(data3$Region)

ks_results <- list()

# KS test se radi za svaku regiju
for (region in regions) {
  data_region <- data3$Health.Score[data3$Region == region]
  ks_result <- ks.test(data_region, "pnorm", mean = mean(data_region), sd = sd(data_region))
}

```

```

ks_results[[region]] <- ks_result
}

## Warning in ks.test(data_region, "pnorm", mean = mean(data_region), sd =
## sd(data_region)): ties should not be present for the Kolmogorov-Smirnov test

for (k in names(ks_results)) {
  cat(k, ":\n")
  print(ks_results[[k]])
  cat("\n")
}

## Western Europe :
##
## One-sample Kolmogorov-Smirnov test
##
## data: data_region
## D = 0.11197, p-value = 0.9397
## alternative hypothesis: two-sided
##
##
## Middle East and North Africa :
##
## One-sample Kolmogorov-Smirnov test
##
## data: data_region
## D = 0.098219, p-value = 0.9982
## alternative hypothesis: two-sided
##
##
## North America and ANZ :
##
## One-sample Kolmogorov-Smirnov test
##
## data: data_region
## D = 0.34591, p-value = 0.6191
## alternative hypothesis: two-sided
##
##
## Central and Eastern Europe :
##
## One-sample Kolmogorov-Smirnov test
##
## data: data_region
## D = 0.092969, p-value = 0.9968
## alternative hypothesis: two-sided
##
##
## Latin America and Caribbean :
##
## One-sample Kolmogorov-Smirnov test
##
## data: data_region
## D = 0.10035, p-value = 0.9805
## alternative hypothesis: two-sided

```

```
##
##
## Southeast Asia :
##
## One-sample Kolmogorov-Smirnov test
##
## data: data_region
## D = 0.21607, p-value = 0.7185
## alternative hypothesis: two-sided
##
##
## East Asia :
##
## One-sample Kolmogorov-Smirnov test
##
## data: data_region
## D = 0.35303, p-value = 0.3569
## alternative hypothesis: two-sided
##
##
## Commonwealth of Independent States :
##
## One-sample Kolmogorov-Smirnov test
##
## data: data_region
## D = 0.19633, p-value = 0.8785
## alternative hypothesis: two-sided
##
##
## Sub-Saharan Africa :
##
## One-sample Kolmogorov-Smirnov test
##
## data: data_region
## D = 0.089175, p-value = 0.9346
## alternative hypothesis: two-sided
##
##
## South Asia :
##
## One-sample Kolmogorov-Smirnov test
##
## data: data_region
## D = 0.18957, p-value = 0.9535
## alternative hypothesis: two-sided
```

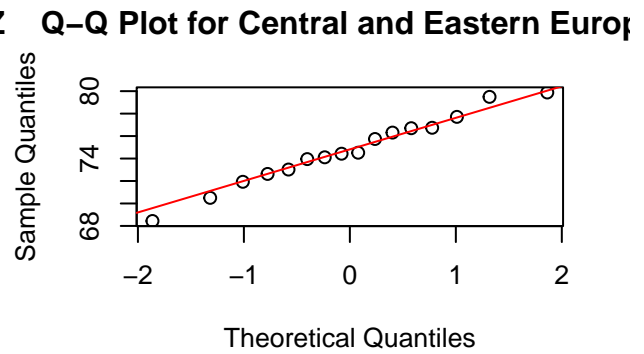
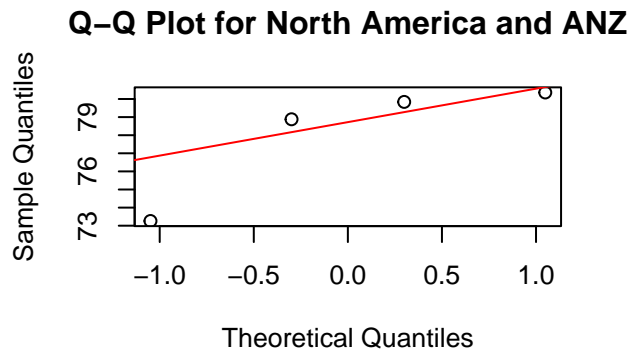
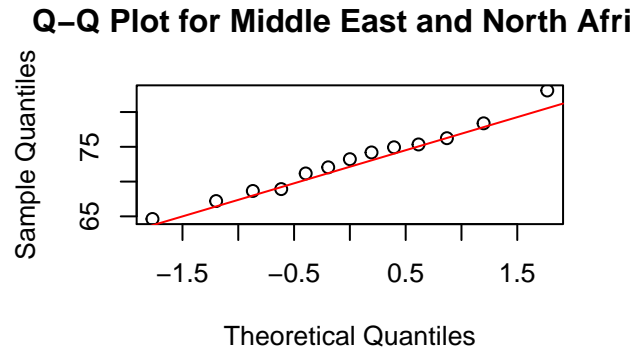
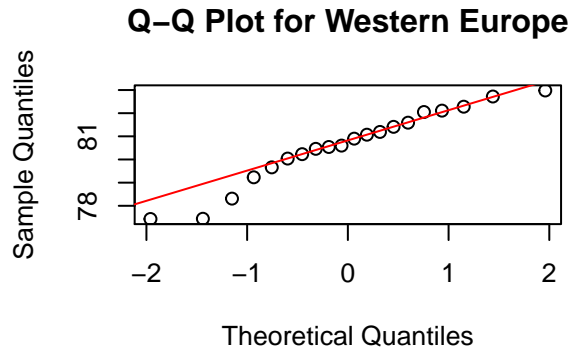
Za svaku regiju dobivamo veliku p-vrijednost, što znači da ne možemo odbaciti hipotezu  $H_0$  i zaključujemo da su podatci normalno distribuirani.

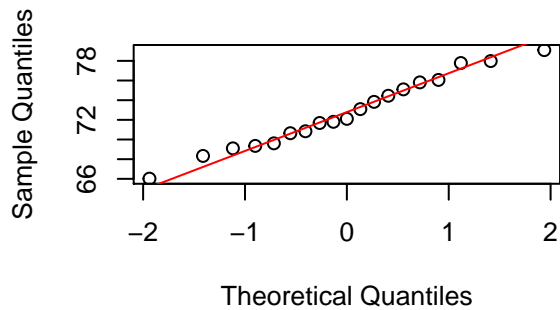
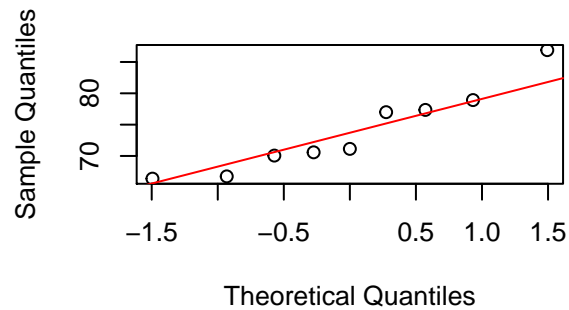
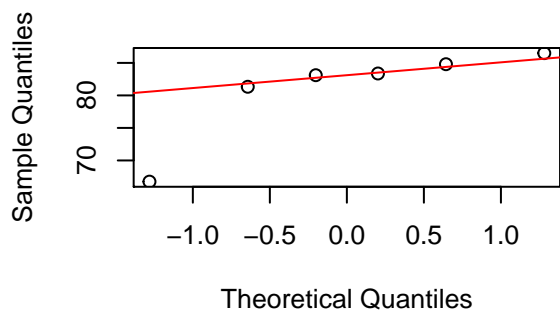
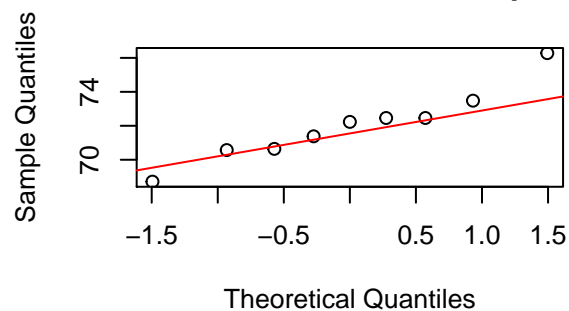
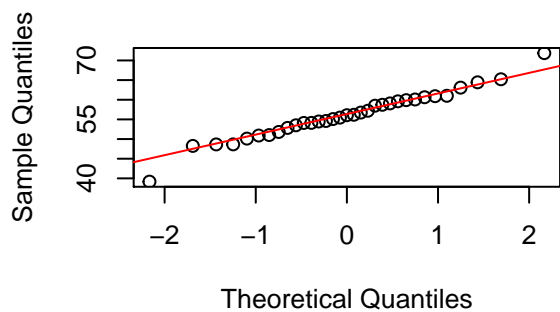
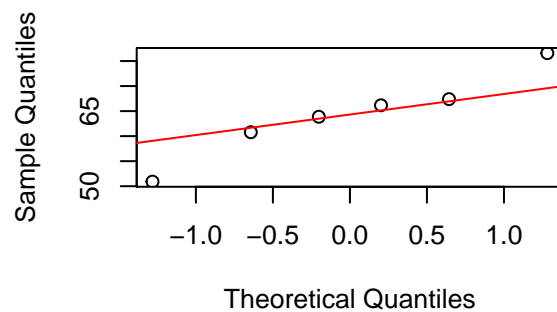
P-vrijednost ovisi o veličini uzorka pa ćemo se koristiti i grafičkom provjerom. Veći uzorak rezultira manjom p-vrijednošću.

```
library(ggplot2)

par(mfrow = c(2, 2)) # 2 rows, 2 columns
```

```
for (region in regions) {
  data_region <- data3$Health.Score[data3$Region == region]
  qqnorm(data_region, main = paste("Q-Q Plot for", region))
  qqline(data_region, col = "red")
}
```



**Q-Q Plot for Latin America and Caribbe****Q-Q Plot for Southeast Asia****Q-Q Plot for East Asia****Q Plot for Commonwealth of Independent****Q-Q Plot for Sub-Saharan Africa****Q-Q Plot for South Asia**

Općenito, što su podaci udaljeniji od crvene linije, to su manje normalno distribuirani. U našem slučaju se iz grafičkog prikaza može vidjeti da se uglavnom radi o normalnoj distribuciji, osim nekih stršećih vrijednosti.

```
model <- aov(Health.Score ~ Region, data = data3)
summary(model)
```

```
##              Df Sum Sq Mean Sq F value Pr(>F)
## Region        9  10572   1174.7   49.95 <2e-16 ***
## Residuals    125    2940     23.5
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
TukeyHSD(model)
```

```
##      Tukey multiple comparisons of means
##      95% family-wise confidence level
##
## Fit: aov(formula = Health.Score ~ Region, data = data3)
##
```

## \$Region	
##	diff
## Commonwealth of Independent States-Central and Eastern Europe	-2.7329167
## East Asia-Central and Eastern Europe	6.2187500
## Latin America and Caribbean-Central and Eastern Europe	-1.9846711
## Middle East and North Africa-Central and Eastern Europe	-1.8324038
## North America and ANZ-Central and Eastern Europe	3.3287500
## South Asia-Central and Eastern Europe	-10.4745833
## Southeast Asia-Central and Eastern Europe	-0.8629167
## Sub-Saharan Africa-Central and Eastern Europe	-18.6368561
## Western Europe-Central and Eastern Europe	5.8557500
## East Asia-Commonwealth of Independent States	8.9516667
## Latin America and Caribbean-Commonwealth of Independent States	0.7482456
## Middle East and North Africa-Commonwealth of Independent States	0.9005128
## North America and ANZ-Commonwealth of Independent States	6.0616667
## South Asia-Commonwealth of Independent States	-7.7416667
## Southeast Asia-Commonwealth of Independent States	1.8700000
## Sub-Saharan Africa-Commonwealth of Independent States	-15.9039394
## Western Europe-Commonwealth of Independent States	8.5886667
## Latin America and Caribbean-East Asia	-8.2034211
## Middle East and North Africa-East Asia	-8.0511538
## North America and ANZ-East Asia	-2.8900000
## South Asia-East Asia	-16.6933333
## Southeast Asia-East Asia	-7.0816667
## Sub-Saharan Africa-East Asia	-24.8556061
## Western Europe-East Asia	-0.3630000
## Middle East and North Africa-Latin America and Caribbean	0.1522672
## North America and ANZ-Latin America and Caribbean	5.3134211
## South Asia-Latin America and Caribbean	-8.4899123
## Southeast Asia-Latin America and Caribbean	1.1217544
## Sub-Saharan Africa-Latin America and Caribbean	-16.6521850
## Western Europe-Latin America and Caribbean	7.8404211
## North America and ANZ-Middle East and North Africa	5.1611538
## South Asia-Middle East and North Africa	-8.6421795
## Southeast Asia-Middle East and North Africa	0.9694872
## Sub-Saharan Africa-Middle East and North Africa	-16.8044522
## Western Europe-Middle East and North Africa	7.6881538
## South Asia-North America and ANZ	-13.8033333
## Southeast Asia-North America and ANZ	-4.1916667
## Sub-Saharan Africa-North America and ANZ	-21.9656061
## Western Europe-North America and ANZ	2.5270000
## Southeast Asia-South Asia	9.6116667
## Sub-Saharan Africa-South Asia	-8.1622727
## Western Europe-South Asia	16.3303333
## Sub-Saharan Africa-Southeast Asia	-17.7739394
## Western Europe-Southeast Asia	6.7186667
## Western Europe-Sub-Saharan Africa	24.4926061
##	lwr
## Commonwealth of Independent States-Central and Eastern Europe	-9.2427398
## East Asia-Central and Eastern Europe	-1.2604673
## Latin America and Caribbean-Central and Eastern Europe	-7.2859134
## Middle East and North Africa-Central and Eastern Europe	-7.6661531
## North America and ANZ-Central and Eastern Europe	-5.4050942
## South Asia-Central and Eastern Europe	-17.9538007

## Southeast Asia-Central and Eastern Europe	-7.3727398
## Sub-Saharan Africa-Central and Eastern Europe	-23.3963580
## Western Europe-Central and Eastern Europe	0.6154435
## East Asia-Commonwealth of Independent States	0.7173194
## Latin America and Caribbean-Commonwealth of Independent States	-5.5738548
## Middle East and North Africa-Commonwealth of Independent States	-5.8743240
## North America and ANZ-Commonwealth of Independent States	-3.3269338
## South Asia-Commonwealth of Independent States	-15.9760140
## Southeast Asia-Commonwealth of Independent States	-5.4950241
## Sub-Saharan Africa-Commonwealth of Independent States	-21.7791962
## Western Europe-Commonwealth of Independent States	2.3175746
## Latin America and Caribbean-East Asia	-15.5198306
## Middle East and North Africa-East Asia	-15.7621401
## North America and ANZ-East Asia	-12.9749746
## South Asia-East Asia	-25.7136088
## Southeast Asia-East Asia	-15.3160140
## Sub-Saharan Africa-East Asia	-31.7895470
## Western Europe-East Asia	-7.6353786
## Middle East and North Africa-Latin America and Caribbean	-5.4712359
## North America and ANZ-Latin America and Caribbean	-3.2814141
## South Asia-Latin America and Caribbean	-15.8063218
## Southeast Asia-Latin America and Caribbean	-5.2003460
## Sub-Saharan Africa-Latin America and Caribbean	-21.1515200
## Western Europe-Latin America and Caribbean	2.8352268
## North America and ANZ-Middle East and North Africa	-3.7719672
## South Asia-Middle East and North Africa	-16.3531657
## Southeast Asia-Middle East and North Africa	-5.8053496
## Sub-Saharan Africa-Middle East and North Africa	-21.9204541
## Western Europe-Middle East and North Africa	2.1220574
## South Asia-North America and ANZ	-23.8883079
## Southeast Asia-North America and ANZ	-13.5802671
## Sub-Saharan Africa-North America and ANZ	-30.2372975
## Western Europe-North America and ANZ	-6.0303847
## Southeast Asia-South Asia	1.3773194
## Sub-Saharan Africa-South Asia	-15.0962137
## Western Europe-South Asia	9.0579547
## Sub-Saharan Africa-Southeast Asia	-23.6491962
## Western Europe-Southeast Asia	0.4475746
## Western Europe-Sub-Saharan Africa	20.0652302
##	upr
## Commonwealth of Independent States-Central and Eastern Europe	3.7769065
## East Asia-Central and Eastern Europe	13.6979673
## Latin America and Caribbean-Central and Eastern Europe	3.3165713
## Middle East and North Africa-Central and Eastern Europe	4.0013455
## North America and ANZ-Central and Eastern Europe	12.0625942
## South Asia-Central and Eastern Europe	-2.9953660
## Southeast Asia-Central and Eastern Europe	5.6469065
## Sub-Saharan Africa-Central and Eastern Europe	-13.8773541
## Western Europe-Central and Eastern Europe	11.0960565
## East Asia-Commonwealth of Independent States	17.1860140
## Latin America and Caribbean-Commonwealth of Independent States	7.0703460
## Middle East and North Africa-Commonwealth of Independent States	7.6753496
## North America and ANZ-Commonwealth of Independent States	15.4502671
## South Asia-Commonwealth of Independent States	0.4926806

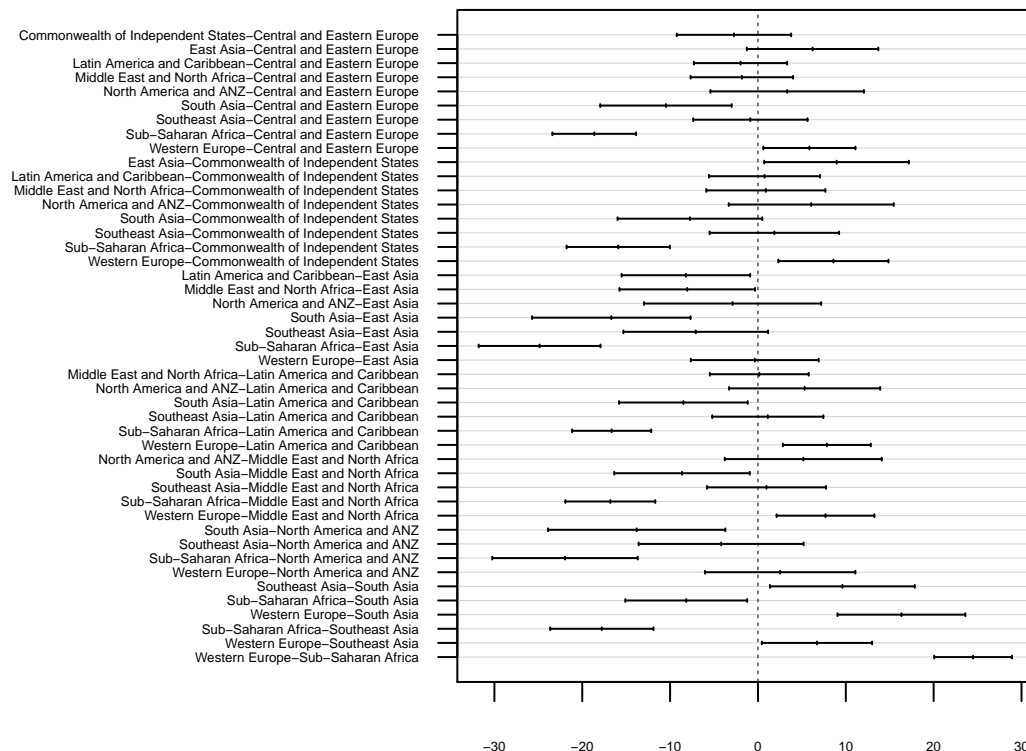


## Southeast Asia-Commonwealth of Independent States	9.2350241
## Sub-Saharan Africa-Commonwealth of Independent States	-10.0286826
## Western Europe-Commonwealth of Independent States	14.8597587
## Latin America and Caribbean-East Asia	-0.8870115
## Middle East and North Africa-East Asia	-0.3401676
## North America and ANZ-East Asia	7.1949746
## South Asia-East Asia	-7.6730578
## Southeast Asia-East Asia	1.1526806
## Sub-Saharan Africa-East Asia	-17.9216651
## Western Europe-East Asia	6.9093786
## Middle East and North Africa-Latin America and Caribbean	5.7757704
## North America and ANZ-Latin America and Caribbean	13.9082562
## South Asia-Latin America and Caribbean	-1.1735027
## Southeast Asia-Latin America and Caribbean	7.4438548
## Sub-Saharan Africa-Latin America and Caribbean	-12.1528500
## Western Europe-Latin America and Caribbean	12.8456153
## North America and ANZ-Middle East and North Africa	14.0942749
## South Asia-Middle East and North Africa	-0.9311932
## Southeast Asia-Middle East and North Africa	7.7443240
## Sub-Saharan Africa-Middle East and North Africa	-11.6884503
## Western Europe-Middle East and North Africa	13.2542503
## South Asia-North America and ANZ	-3.7183587
## Southeast Asia-North America and ANZ	5.1969338
## Sub-Saharan Africa-North America and ANZ	-13.6939146
## Western Europe-North America and ANZ	11.0843847
## Southeast Asia-South Asia	17.8460140
## Sub-Saharan Africa-South Asia	-1.2283318
## Western Europe-South Asia	23.6027119
## Sub-Saharan Africa-Southeast Asia	-11.8986826
## Western Europe-Southeast Asia	12.9897587
## Western Europe-Sub-Saharan Africa	28.9199820
##	p adj
## Commonwealth of Independent States-Central and Eastern Europe	0.9388863
## East Asia-Central and Eastern Europe	0.1939504
## Latin America and Caribbean-Central and Eastern Europe	0.9701341
## Middle East and North Africa-Central and Eastern Europe	0.9910529
## North America and ANZ-Central and Eastern Europe	0.9664679
## South Asia-Central and Eastern Europe	0.0006008
## Southeast Asia-Central and Eastern Europe	0.9999918
## Sub-Saharan Africa-Central and Eastern Europe	0.0000000
## Western Europe-Central and Eastern Europe	0.0159694
## East Asia-Commonwealth of Independent States	0.0217443
## Latin America and Caribbean-Commonwealth of Independent States	0.9999969
## Middle East and North Africa-Commonwealth of Independent States	0.9999916
## North America and ANZ-Commonwealth of Independent States	0.5449593
## South Asia-Commonwealth of Independent States	0.0842806
## Southeast Asia-Commonwealth of Independent States	0.9981865
## Sub-Saharan Africa-Commonwealth of Independent States	0.0000000
## Western Europe-Commonwealth of Independent States	0.0008872
## Latin America and Caribbean-East Asia	0.0153549
## Middle East and North Africa-East Asia	0.0331366
## North America and ANZ-East Asia	0.9954342
## South Asia-East Asia	0.0000011
## Southeast Asia-East Asia	0.1582095

```
## Sub-Saharan Africa-East Asia 0.0000000
## Western Europe-East Asia 1.0000000
## Middle East and North Africa-Latin America and Caribbean 1.0000000
## North America and ANZ-Latin America and Caribbean 0.6064486
## South Asia-Latin America and Caribbean 0.0101559
## Southeast Asia-Latin America and Caribbean 0.9999011
## Sub-Saharan Africa-Latin America and Caribbean 0.0000000
## Western Europe-Latin America and Caribbean 0.0000663
## North America and ANZ-Middle East and North Africa 0.6946841
## South Asia-Middle East and North Africa 0.0154306
## Southeast Asia-Middle East and North Africa 0.9999841
## Sub-Saharan Africa-Middle East and North Africa 0.0000000
## Western Europe-Middle East and North Africa 0.0007664
## South Asia-North America and ANZ 0.0008969
## Southeast Asia-North America and ANZ 0.9126184
## Sub-Saharan Africa-North America and ANZ 0.0000000
## Western Europe-North America and ANZ 0.9942937
## Southeast Asia-South Asia 0.0094310
## Sub-Saharan Africa-South Asia 0.0084717
## Western Europe-South Asia 0.0000000
## Sub-Saharan Africa-Southeast Asia 0.0000000
## Western Europe-Southeast Asia 0.0254279
## Western Europe-Sub-Saharan Africa 0.0000000
```

```
with(par(mai=c(0.5,2.5,0.5,1)),{plot(TukeyHSD(model), las=1,cex.axis=0.4)})
```

## 95% family-wise confidence level



Stupac “diff” pokazuje razliku u srednjoj vrijednosti između dvije promatrane grupe, dok su “lwr” i “upr” granice intervala pouzdanosti. Gledamo p-vrijednosti Tukey HSD testa. Ako je p-vrijednost manja od 0.05,

imamo par grupa koje se statistički značajno razlikuju.

Ako gledamo samo Europu, jednu grupu čini zapadni dio, dok drugu grupu čine centralni i istočni dio. Razlika srednjih vrijednosti iznosi 5.86, što se ne čini puno, ali se pokazalo statistički značajno s p-vrijednošću 0.016.

Unutar same Azije postoje velike razlike razvijenosti zdravstva, s istočnom Azijom koja je značajno bolja od južne Azije. U istočnoj Aziji indeks razvijenosti zdravstva iznosi više od 80, s iznimkom Mongolije čiji je indeks 66.74.

Najmanje zdravstveno razvijena pokazala se regija Sub-Saharan Africa čiji rezultati odskakuju od svih ostalih regija. Najviše odstupa od istočne Azije s razlikom srednjih vrijednosti 24.86 i zapadne Europe s razlikom 24.49.

Najbolje zdravstveno razvijene regije su istočna Azija i zapadna Europa, te odmah uz njih Sjeverna Amerika, Australija i Novi Zeland.