

**ŠTUDIJSKO LETO 2018/2019**

**STATISTIČNA NALOGA**

**Predmet: *Statistika 1*, 1. letnik (RST)**

**Nosilka predmeta: doc.dr. Nuša Erman**

|  |  |
| --- | --- |
| **PRIIMEK IN IME:** |  |
| **VPISNA ŠTEVILKA:** |  |

1. **Opisovanje in urejanje podatkov**
2. **Opredelitev populacije, statistične enote in statističnih spremenljivk**

Opredelitev populacije:

|  |  |
| --- | --- |
| * stvarna opredelitev: |  |
| * krajevna opredelitev: |  |
| * časovna opredelitev: |  |
|  |  |
| Statistična enota: |  |

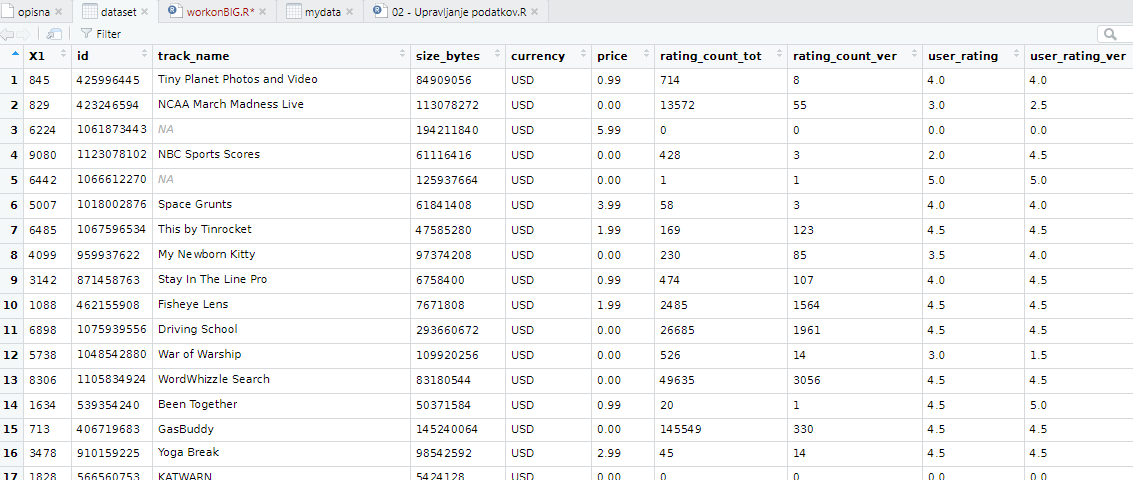
|  |  |  |  |
| --- | --- | --- | --- |
| Statistične spremenljivke: | | | |
| ***Spremenljivka*** | ***Opis*** | ***Vrsta (podvrsta)*** | ***Tip merjenja*** |
| ID | App ID |  |  |
| track\_name | App Name |  |  |
| size\_bytes | Size (in Bytes) |  |  |
| currency | Currency Type |  |  |
| price | Price amount |  |  |
| rating\_count\_tot | User Rating counts (for all version) |  |  |
| rating\_count\_ver | User Rating counts (for current version) |  |  |
| user\_rating | Average User Rating value (for all version) |  |  |
| user\_rating\_ver | Average User Rating value (for current version) |  |  |
| ver | Latest version code |  |  |
| cont\_rating | Content Rating |  |  |
| prime\_genre | Primary Genre |  |  |
| sup\_devices.num | Number of supporting devices |  |  |
| ipadSc\_urls.num | Number of screenshots showed for display |  |  |
| lang.num | Number of supported languages |  |  |
| vpp\_lic | Vpp Device Based Licensing Enabled |  |  |

1. **Prikaz podatkov**

Uporabljen ukaz v programskem okolju R:

|  |
| --- |
| dataset <- read\_csv("AppleStoreData-sample.csv") |

Slika 1: Prikaz podatkov (okno Data)

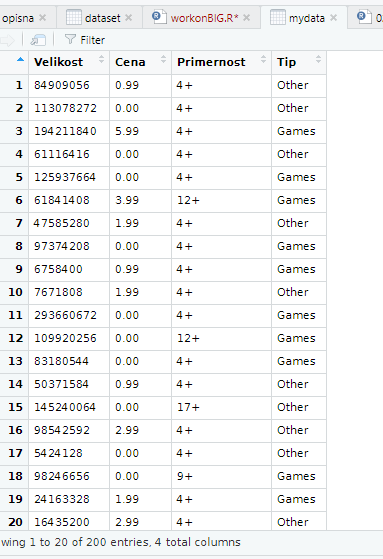


1. **Priprava podatkov**

Uporabljeni ukazi v programskem okolju R:

|  |
| --- |
| interest<-c("size\_bytes","price","cont\_rating","prime\_genre")  mydata<-dataset[interest]  names(mydata)[1]<-"Velikost"  names(mydata)[2]<-"Cena"  names(mydata)[3]<-"Primernost"  names(mydata)[4]<-"Tip" |

Slika 2: Prikaz podatkov za analizo (okno Data)



1. **Analiza številskih spremenljivk**
2. **Opisna statistika številskih spremenljivk**

Izpis 1: Opisna statistika številskih spremenljivk

velikost

cena



|  |
| --- |
| > describe(copydata$Velikost)  mean sd median trimmed mad min max range skew kurtosis se  153.52 215.57 101.5 111 73.71 1.74 1909.81 1908.07 4.67 28.55 15.24  > describe(copydata$Cena)  mean sd median trimmed mad min max range skew kurtosis se  1.49 2.31 0 1.01 0 0 14.99 14.99 2.54 9.26 0.16 |

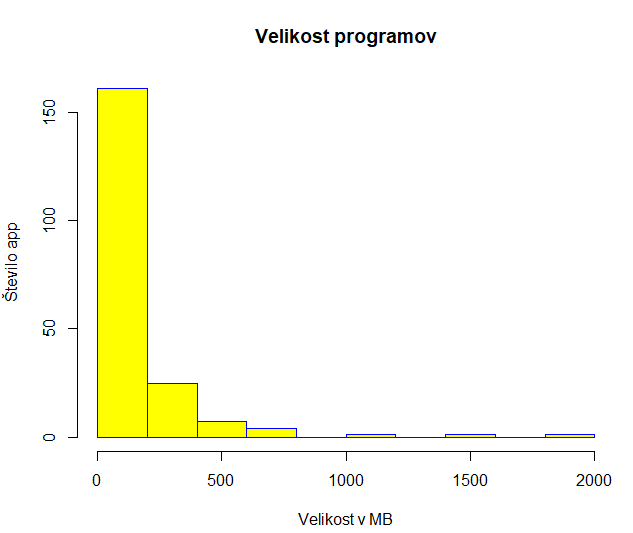
Interpretacija:

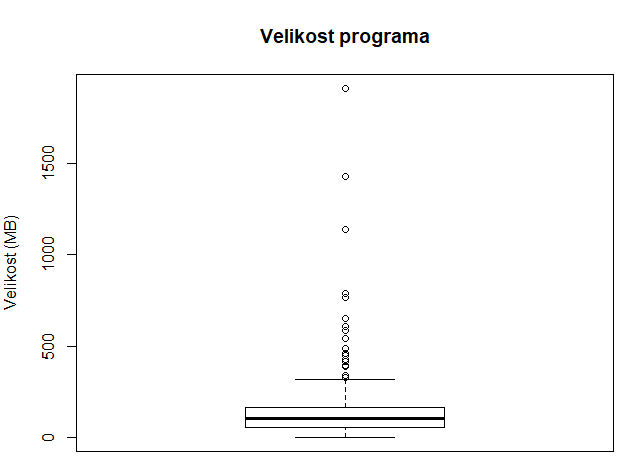
1. **Grafični prikaz številskih spremenljivk**

Uporabljen ukaz v programskem okolju R:

|  |
| --- |
| hist(copydata$Velikost, xlab="Velikost v MB", ylab="Število app", col = "yellow", border = "blue", main = "Velikost programov")  boxplot(copydata$Velikost, notch = FALSE, ylab = "Velikost (MB)", main="Velikost programa") |

Slika 3: Histogram in okvir z ročaji spremenljivke Kliknite tukaj, če želite vnesti besedilo.

****

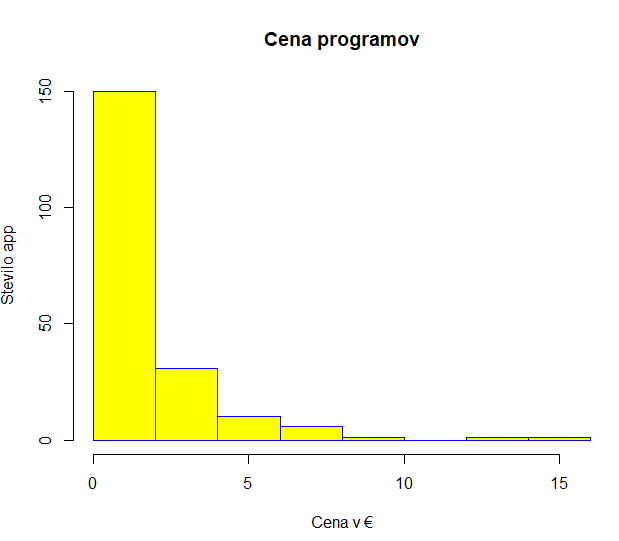


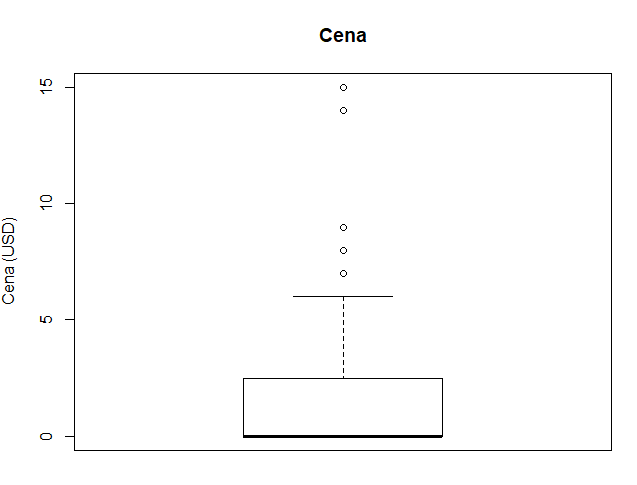
Interpretacija:

Uporabljen ukaz v programskem okolju R:

|  |
| --- |
| hist(copydata$Cena, xlab="Cena v €", ylab="Število app", col = "yellow", border = "blue", main = "Cena programov") |

Slika 4: Histogram in okvir z ročaji spremenljivke Kliknite tukaj, če želite vnesti besedilo.

****

****

Interpretacija:

1. **Razsevni grafikon**

Uporabljen ukaz v programskem okolju R:

|  |
| --- |
|  |

Slika 5: Razsevni grafikon spremenljivk Kliknite tukaj, če želite vnesti besedilo.

Interpretacija:

1. **Povezanost številskih spremenljivk**

Izpis 2: Pearsonov korelacijski koeficient med spremenljivkama Kliknite tukaj, če želite vnesti besedilo.

|  |
| --- |
|  |

Interpretacija:

1. **Odvisnost med številskima spremenljivkama**

|  |  |
| --- | --- |
| Odvisna spremenljivka (*Y*): | Cena programa v Dolarjih |
| Neodvisna spremenljivka (*X*): | Velikost programa v Megabajtih |

Izpis 3: Enačba regresijske premice odvisnosti spremenljivke Kliknite tukaj, če želite vnesti besedilo. od spremenljivke Kliknite tukaj, če želite vnesti besedilo.

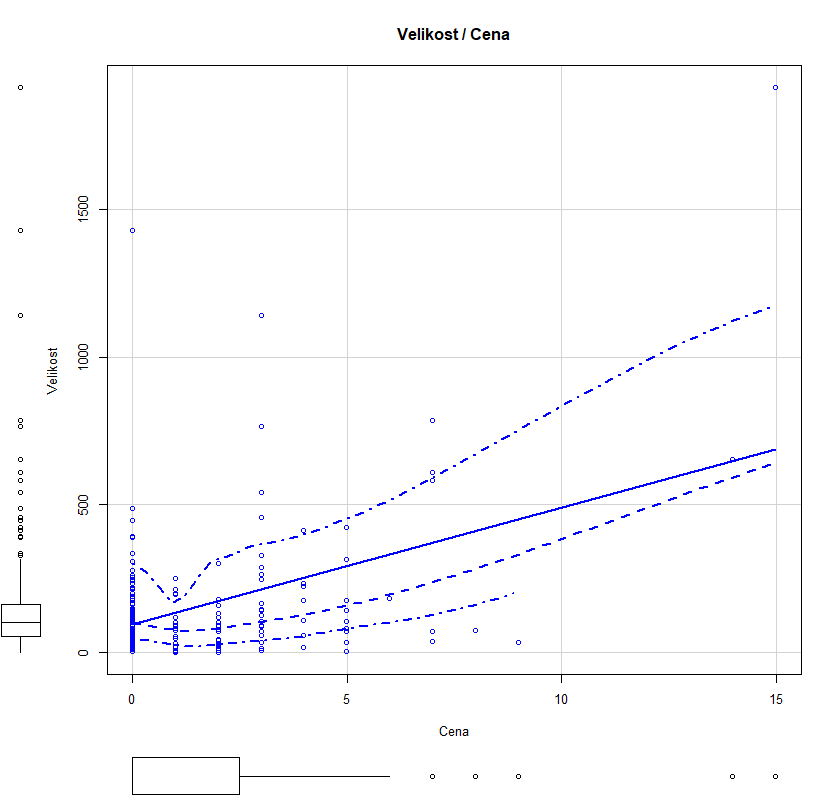
|  |
| --- |
| lm(formula = x ~ y)  Coefficients:  (Intercept) y  0.787088 0.004548  Cena = 0.78 + 0.004548 \* Velikost |

Interpretacija:

Uporabljen ukaz za izris razsevnega grafikona z vrisano regresijsko premico v programskem okolju R:

|  |
| --- |
| install.packages('car',repos='http://cran.us.r-project.org')  library(car)  scatterplot(Velikost ~ Cena, data = copydata) |

Slika 6: Razsevni grafikon spremenljivk Kliknite tukaj, če želite vnesti besedilo. z vrisano regresijsko premico



1. **Kakovost regresijskega modela**

Izpis 4: Značilnosti regresijskega modela (odvisna spremenljivka: Kliknite tukaj, če želite vnesti besedilo., neodvisna spremenljivka: Kliknite tukaj, če želite vnesti besedilo.)

|  |
| --- |
|  |

Interpretacija:

1. **Analiza opisnih spremenljivk**
2. **Opisna statistika opisnih spremenljivk**

Izpis 5: Predstavitev opisnih spremenljivk

|  |
| --- |
|  |

Interpretacija:

1. **Grafični prikaz opisnih spremenljivk**

Uporabljen ukaz v programskem okolju R:

|  |
| --- |
|  |

Slika 7: Kliknite tukaj, če želite vnesti besedilo. spremenljivke Kliknite tukaj, če želite vnesti besedilo.

Interpretacija:

Uporabljen ukaz v programskem okolju R:

|  |
| --- |
|  |

Slika 8: Kliknite tukaj, če želite vnesti besedilo. spremenljivke Kliknite tukaj, če želite vnesti besedilo.

Interpretacija:

1. **Kontingenčna tabela**

Izpis 6: Kontingenča tabela spremenljivk Kliknite tukaj, če želite vnesti besedilo.

|  |
| --- |
|  |

Interpretacija:

1. **Dvorazsežna struktura**

Odgovor:

1. **Prikaz dvorazsežne strukture**

Izpis 7: Kliknite tukaj, če želite vnesti besedilo. spremenljivk Kliknite tukaj, če želite vnesti besedilo.

|  |
| --- |
|  |

Uporabljeni ukazi za prikaz dvorazsežne strukture v programskem okolju R:

|  |
| --- |
|  |

Slika 9: Struktura za Kliknite tukaj, če želite vnesti besedilo. po Kliknite tukaj, če želite vnesti besedilo.

1. **Tabela pričakovanih frekvenc**

Izpis 8: Tabela pričakovanih frekvenc spremenljivk Kliknite tukaj, če želite vnesti besedilo.

|  |
| --- |
|  |

1. **Povezanost opisnih spremenljivk**

Izpis 9: Hi-kvadrat in vrednosti kontingenčnih koeficientov med spremenljivkama Kliknite tukaj, če želite vnesti besedilo.

|  |
| --- |
|  |

Interpretacija:

1. **Glavne ugotovitve**

Odgovor: