

# Domača naloga

## - Rdeče vino -

STATISTIKA 2024/25

V datoteki *winequality-red* so zbrani naslednji podatki za izbrana rdeča vina: kislost (angl. fixed acidity), hlapna kislost (angl. volatile acidity), citronska kislina (angl. citric acid), sladkor (angl. residual sugar), kloridi (angl. chlorides), prost žveplov dioksid (free sulfur dioxide), celoten žveplov dioksid (angl. total sulphur dioxide), gostota (angl. density), pH, sulfati (angl. sulphates), odstotek alkohola ter kvaliteta vina, kot so jo ocenili preizkuševalci vina.

Vaša naloga je pripraviti poročilo, v katerem boste odgovorili na vseh pet spodaj zastavljenih vprašanj. Za analizo podatkov uporabite program R, poročilo pa naj pri vsakem od vprašanj vsebuje (1) opis problema in statističnih metod uporabljenih za analizo problema ter (2) rezultate analize in njihovo interpretacijo.

Poročilo naj vsebuje imena vseh sodelujočih v raziskavi, eden od članov skupine pa poročilo odda v e-učilnici. Namen domače naloge je (poleg prikaza poznavanja statističnih metod in uporabe programa R) spodbuditi ekipno delo in sodelovanje med študenti. V kolikor pa vam je zaradi kateregakoli razloga ljubše samostojno delo, pa lahko seveda na izbrani temi delate sami (v tem primeru prosim vašo odločitev sporočite predavateljici in obvestite ostale študente v skupini).

Vprašanja:

I. [5t] Grafično prikažite porazdelitev spremenljivk (a) kislost, (b) sladkor ter (c) celoten žveplov dioksid. Grafično prikažite še porazdelitev kvalitete vseh vin. Narišite še histograme, ki prikazujejo porazdelitev kislosti glede na kvaliteto vina. Interpretirajte rezultate.

II. [5t] Izpišite kvartile in standardni odklon za spremenljivko "kislost" ter narišite okvir z ročaji. Enako naredite še za spremenljivki "sladkor" in "celoten žveplov dioksid". Interpretirajte rezultate.

III. [5t] Zanima nas, ali je povprečna kislost vin, ki so dobila oceno 5, višja kot povprečne kislosti vin z oceno 6. Katerega od statističnih testov bi uporabili in kakšen je vaš zaključek? Z ustreznim statističnim testom preverite še hipotezo, da je povprečna vsebnost sladkorja vseh izbranih vin večja od 2,3. Interpretirajte rezultate.

IV. [5t] Za vsako od dvanajstih zgoraj omenjenih spremenljivk preverite, ali so podatki skladni z normalno porazdelitvijo. Če da, podajte parametre teh porazdelitev.

V. [5t] Po lastni izbiri postavite vsaj še eno raziskovalno vprašanje in nanj natančno odgovorite.