

Uvod u analizu vremenskih nizova podataka korištenjem programskog jezika R

Filjević, Mislav

Undergraduate thesis / Završni rad

2024

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **University of Zagreb, Faculty of Organization and Informatics / Sveučilište u Zagrebu, Fakultet organizacije i informatike**

Permanent link / Trajna poveznica: <https://urn.nsk.hr/urn:nbn:hr:211:159822>

Rights / Prava: [Attribution 3.0 Unported/Imenovanje 3.0](#)

Download date / Datum preuzimanja: **2024-11-13**



Repository / Repozitorij:

[Faculty of Organization and Informatics - Digital Repository](#)



SVEUČILIŠTE U ZAGREBU
FAKULTET ORGANIZACIJE I INFORMATIKE
VARAŽDIN

Mislav Filjević

UVOD U ANALIZU VREMENSKIH NIZOVA
PODATAKA KORIŠTENJEM
PROGRAMSKOG JEZIKA R

ZAVRŠNI RAD

Varaždin, 2024.

SVEUČILIŠTE U ZAGREBU
FAKULTET ORGANIZACIJE I INFORMATIKE
V A R A Ź D I N

Mislav Filjević

Matični broj: 0016154857

Studij: Informacijski i poslovni sustavi

**UVOD U ANALIZU VREMENSKIH NIZOVA PODATAKA KORIŠTENJEM
PROGRAMSKOG JEZIKA R**

ZAVRŠNI RAD

Mentor/Mentorica:

Doc. dr. sc. Petra Žugec

Varaždin, rujan 2024.

Mislav Filjević

Izjava o izvornosti

Izjavljujem da je moj završni rad izvorni rezultat mojeg rada te da se u izradi istoga nisam koristio drugim izvorima osim onima koji su u njemu navedeni. Za izradu rada su korištene etički prikladne i prihvatljive metode i tehnike rada.

Autor/Autorica potvrdio/potvrdila prihvaćanjem odredbi u sustavu FOI-radovi

Sažetak

Rad se bavi analizom vremenskih nizova cijena bitcoina koristeći metode autokorelacije, djelomične autokorelacije i ARIMA modela. Vremenski nizovi su definirani kao niz podataka prikupljenih u vremenskim intervalima, a ciljevi rada uključuju razumijevanje dinamike cijena bitcoina i predviđanje budućih kretanja. Teorijski dio objašnjava osnovne koncepte vremenskih nizova, a analiza se fokusira na dnevne i tjedne cijene bitcoina u periodu od 1. siječnja 2020. do 1. siječnja 2024. godine. Rad koristi R programski jezik s paketima forecast, ggplot2 i xts za analizu i vizualizaciju podataka. Zaključak rada je da su vremenski nizovi korisni za predviđanje kretanja cijena na temelju povijesnih podataka, te da ARIMA model pruža korisne informacije.

Ključne riječi: vremenski nizovi, bitcoin, ARIMA model, autokorelacija, predikcija, R, analiza

Sadržaj

1. Uvod	1
2. Vremenski nizovi	2
3. Analiza vremenskih nizova	3
3.1. Primjeri analize vremenskih nizova	3
4. Paketi za rad s vremenskim nizovima u R-u	5
5. Vizualizacija vremenskih nizova	7
6. Analiza sezonalnosti i trenda	12
6.1. Autokorelacija i djelomična autokorelacija	13
7. Korelacija između dva vremenska niza	18
8. ARIMA model	20
8.1. Stacionarnost vremenskog niza	21
8.2. Komponente ARIMA modela	22
8.3. Izrada ARIMA modela	22
8.4. Predikcija tjednih cijena bitcoina	24
9. Zaključak	26
Popis literature	28
Popis slika	29
1. Prilog 1	30
2. Prilog 2	31

1. Uvod

Rad je organiziran u nekoliko glavnih poglavlja, s dodatnim podnaslovima za detaljniju obradu svakog dijela teme. Obuhvatit će se teorijski dio s prikazom praktičkog dijela.

Analiza vremenskih nizova igra ključnu ulogu u raznim područjima kao što su financije, ekonomija, meteorologija i tako dalje. Vremenski niz je ukratko slijed informacija koji pridružuje vremensko razdoblje svakoj vrijednosti. Vrijednost se može odnositi na gotovo svaki mjerljiv događaj koji djeluje ili se mijenja tijekom vremena kao što su cijene, vlažnost ili brojevi ljudi. Međutim, u analizi vremenskih nizova, podatkovne točke snimljene su u redovitim vremenskim trenutcima, a ne povremeno ili nasumično. Ova vrsta analize omogućuje analitičarima da otkriju kako se varijable mijenjaju kroz vrijeme, odnosno vrijeme postaje ključna varijabla koja pokazuje kako se podaci mijenjaju kroz različite vremenske točke te kako se te promjene odražavaju u kasnijim konačnim rezultatima.

Cilj ovog rada je analizirati vremenske nizove cijena bitcoina koristeći metode autokorelacije, parcijalne autokorelacije i ARIMA modela. Kroz ovu analizu, koristit će se dnevni i tjedni podaci o prvim cijenama i volumenu trgovanja bitcoina od 1. siječnja 2020. do 1. siječnja 2024. godine. Rezultati ove analize bit će korisni za bolje razumijevanje dinamike tržišta bitcoina i za izradu pouzdanih predikcijskih modela.

Dakle, analizirat će se stacionarnost vremenskog niza korištenjem proširenog Dickey-Fuller testa (engl. *Augmented Dickey-Fuller - ADF*), detaljno ćemo objasniti model autoregresivnog integriranog pomičnog prosjeka (engl. *Autoregressive integrated moving average - ARIMA*), kreirati ARIMA model koristeći funkciju `auto.arima`. Na kraju, napraviti ću predikcije za buduće vrijednosti cijene bitcoina na temelju izrađenog ARIMA modela. Kroz cijeli rad, koristit ću literaturu i izvore kako bi potkrijepio metode i zaključke.

2. Vremenski nizovi

Vremenski niz je niz podataka

$$x_{t_1}, x_{t_2}, \dots, x_{t_n},$$

prikupljenih u uzastopnim vremenskim trenucima $T_0 = \{t_1, \dots, t_n\}$, $t_1 < t_2 < \dots < t_n$ i $x_{t_i} \in \mathbb{R}$, $i = 1, \dots, n$. Zbog načina na koji prikupljamo podatke, skup vremenskih trenutaka T_0 će najčešće biti konačan skup pa govorimo o vremenskom nizu u diskretnom vremenu. Moguće je da vrijednost nekog procesa poznajemo u svakom trenutku pa u tom slučaju govorimo o vremenskom nizu u neprekidnom vremenu (npr. broj ljudi u trgovini tijekom radnog dana može biti poznat u svakom trenutku) [1]. U ovom radu cijena bitcoina smatra se diskretnim vremenskim nizom jer se podaci o cijenama bilježe u specifičnim, odvojenim vremenskim trenucima.

Dakle, vremenski niz je slijed informacija koji pridružuje vremensko razdoblje svakoj vrijednosti. Vrijednost može biti gotovo bilo što mjerljivo što na neki način ovisi o vremenu, poput cijena, vlažnosti ili broja ljudi. Sve dok su vrijednosti koje bilježimo jednoznačne, bilo koji medij se može mjeriti vremenskim nizovima [2].

Postoji nekoliko ključnih karakteristika vremenskih nizova koje su važne za analizu i interpretaciju vremenskih nizova. Dvije glavne karakteristike vremenskih nizova su vremenski period i frekvencija. Ove karakteristike pružaju osnovne informacije o strukturi i prirodi podataka te omogućavajući da bolje razumijemo dinamiku promjena koje se odvijaju kroz vrijeme. Razumijevanje ovih karakteristika pomaže u pravilnom pristupu analizi vremenskih nizova i u primjeni odgovarajućih metoda za predviđanje i modeliranje budućih vrijednosti.

Prvo, prema [2], nema ograničenja u pogledu ukupnog trajanja vremenskog niza. To može biti minuta, dan, mjesec ili čak stoljeće. Sve što je potrebno su početna i završna točka. Naravno, obično postoji više vremenskih trenutaka između tih točaka. Na primjer, ako su podaci prikupljeni jednom dnevno od 1. siječnja 2000. do 31. prosinca 2009., razmak od jednog dana predstavlja vremenski trenutak, dok je skup vremenskih trenutaka cijelo desetljeće.

Frekvencija skupa podataka označava koliko često se bilježe vrijednosti skupa podataka [2]. Da bi analiza vremenskog niza bila smisljena, svi vremenski trenutci moraju biti jednaki i jasno definirani. To rezultira ekvidistantnim rasporedom vremenskih trenutaka, što pokazuje pravilnost u bilježenju podataka. Frekvencija se može kretati od nekoliko milisekundi do nekoliko desetljeća, ali najčešće se pojavljuju dnevni, mjesečni, kvartalni i godišnji podaci[1].

3. Analiza vremenskih nizova

Analiza vremenskih nizova ima ključnu ulogu u raznim područjima, posebno u kontekstu upravljanja, gdje je često potrebno predvidjeti određene vrste ishoda. Primjeri takvih ishoda mogu uključivati prodaju, razinu zadovoljstva kupaca, sposobnost tvrtke da postigne ciljane razine troškova ili uspješno izvršavanje projekata. U svim tim situacijama, korištenje povijesnih podataka pomaže u izradi predviđanja. Prema [3] analiza vremenskih nizova je dio prediktivne analize i predstavlja učinkovit alat za kvantificiranje utjecaja upravljačkih odluka na buduće rezultate.

Dakle, u analizi vremenskih nizova podatkovne točke se bilježe u dosljednim vremenskim trenutcima, umjesto da se bilježe povremeno ili nasumično. Ova vrsta analize nije samo čin prikupljanja podataka tijekom vremena; ona omogućuje analitičarima da otkriju kako se varijable mijenjaju kroz vrijeme. Drugim riječima, vrijeme postaje ključna varijabla koja pokazuje kako se podaci mijenjaju kroz različite vremenske točke, te kako se te promjene manifestiraju u konačnim rezultatima.

Analiza vremenskih nizova tako pruža dodatni izvor informacija i jasno definiran redoslijed ovisnosti među podacima. Za što pouzdaniju analizu potrebno je imati što veći broj podatkovnih točaka kako bi se osigurala dosljednost i pouzdanost rezultata. Opsežan skup podataka omogućava postizanje reprezentativne veličine uzorka, što pomaže u izbjegavanju pogrešaka zbog izvanrednih vrijednosti.

3.1. Primjeri analize vremenskih nizova

Analiza vremenskih nizova koristi se za nestacionarne podatke – one koje se neprestano mijenjaju tijekom vremena. Industrije kao što su financije, maloprodaja i ekonomija često primjenjuju analizu vremenskih nizova jer se parametri poput valute i prodaje stalno mijenjaju. Analiza tržišta dionica predstavlja izvrstan primjer praktične primjene analize vremenskih nizova, posebice kada se koriste algoritmi za automatsko trgovanje. Slično tome, analiza vremenskih nizova je idealna za predviđanje vremenskih promjena, pomažući meteorolozima u predviđanju vremenskih uvjeta, od kratkoročnih prognoza do dugoročnih klimatskih promjena [3].

Primjeri primjene analize vremenskih nizova uključuju:

- vremenski podaci
- mjerenja količine padalina
- očitavanje temperature
- praćenje srčanog ritma (EKG)
- praćenje moždane aktivnosti (EEG)

- tromjesečna prodaja
- cijene dionica
- automatizirano trgovanje dionicama
- industrijske prognoze
- kamatne stope.

4. Paketi za rad s vremenskim nizovima u R-u

R, jedan od najpopularnijih programskih jezika za analizu podataka, zahvaljujući svojoj fleksibilnosti, bogatstvu funkcionalnosti i otvorenosti koda ima nekoliko paketa napisanih posebno za rad s vremenskim nizovima tako da se može izvršiti širok raspon analiza i manipulacija ovim tipom podataka. Ovdje su ključni paketi koji se često koriste u analizi vremenskih nizova:

zoo

Paket zoo je osnovni alat za rad s vremenskim nizovima, posebno s neregularnim nizovima. Neregularni vremenski nizovi su nizovi u kojima vremenski trenutci nisu jednaki. Takvi nizovi su česti u situacijama kada podaci nisu kontinuirano prikupljeni. Zoo je fleksibilan i omogućava rad s različitim formatima datuma i vremena, pružajući stabilan temelj za napredne analize. Njegova integracija s xts paketom omogućava proširene funkcionalnosti i poboljšava učinkovitost rukovanja podacima [4].

xts (eXtensible Time Series)

Paket xts pruža moćan alat za prikaz i manipulaciju vremenskim nizovima u R-u. Temeljen na zoo paketu, xts omogućava jedinstveno rukovanje različitim vremenski baziranim klasama podataka u R-u. Njegove funkcionalnosti uključuju indeksiranje, poravnavanje i podskupove podataka, što ga čini idealnim za kompleksne analize vremenskih nizova. Također omogućava lako pretvaranje podataka u druge vremenske klase i podržava napredne metode za upravljanje nedostajućim vrijednostima [4].

stats

Paket stats pruža širok spektar statističkih metoda koje omogućuju analitičarima da kvantificiraju nesigurnost i donose informirane odluke na temelju podataka. Korištenjem tih metoda, moguće je analizirati obrasce, trendove i sezonske varijacije u vremenskim nizovima [5].

forecast

Forecast omogućava korištenje povijesnih podataka za predviđanje budućih vrijednosti vremenskog niza. Prognoza se temelji na identificiranju obrazaca i trendova u povijesnim podacima te korištenju tih informacija za predviđanje budućih događaja. Metode prognoze mogu uključivati statističke modele, kao što su ARIMA i eksponencijalno izgladivanje, kao i napredne tehnike strojnog učenja [6].

Osim gore navedenih paketa za rad s vremenskim nizovima u R-u, postoji niz drugih paketa koji omogućuju fleksibilnije upravljanje i analizu podataka. U **fts** paketu, vremenski nizovi su predstavljeni kao matrica, dok su datumi pohranjeni kao atributi te matrice, što omogućuje efikasno rukovanje podacima u ovom formatu. **lrts** iz **tseries** paketa pruža podršku za rad s nepravilnim vremenskim nizovima, dok **timeSeries** paket nudi alate za analizu financijskih vremenskih nizova. **Tsibble** paket, dio tidyverse okruženja, omogućuje rad s vremenskim nizovima na način koji je usklađen s principima tidyverse-a, što ga čini korisnim za analizu podataka po presjeku. **Tsbox** omogućuje jednostavno pretvaranje između različitih formata vremenskih nizova, olakšavajući integraciju i analizu podataka unutar različitih R paketa [7].

5. Vizualizacija vremenskih nizova

Za skup podataka je odabrani su skupovi podataka s dnevnim, odnosno tjednim cijenama bitcoina u rasponu od 01.01.2020. do 01.01.2024. [8]. Dakle, za svaki dan i tjedan u danom vremenskom intervalu od četiri godine imamo podatak o cijeni bitcoina. Preciznije, za svaki dan i tjedan postoji cijena s kojom je započeo dan (prva cijena) te volumen trgovanja toga dana, odnosno tjedna. Volumen označava ukupnu količinu trgovine bitcoinom za određeni vremenski period. To znači ukupnu količinu kupljenih i prodanih bitcoina ukupno tokom promatranog perioda. Sve je izraženo u jedinici USD.

Kako vizualizirati vremenske nizove? Prvo, postavlja se okruženje i učitavaju potrebne biblioteke. U nastavku su navedene linije koda koje su potrebne za pripremu okruženja.

```
# knitr postavke opcija
knitr::opts_chunk$set(
  warning=FALSE,
  message=FALSE,
  echo = TRUE)

# Učitavanje potrebnih biblioteka
library(xts)
library(zoo)
library(ggplot2)
library(stats)
library(forecast)
library(tseries)
```

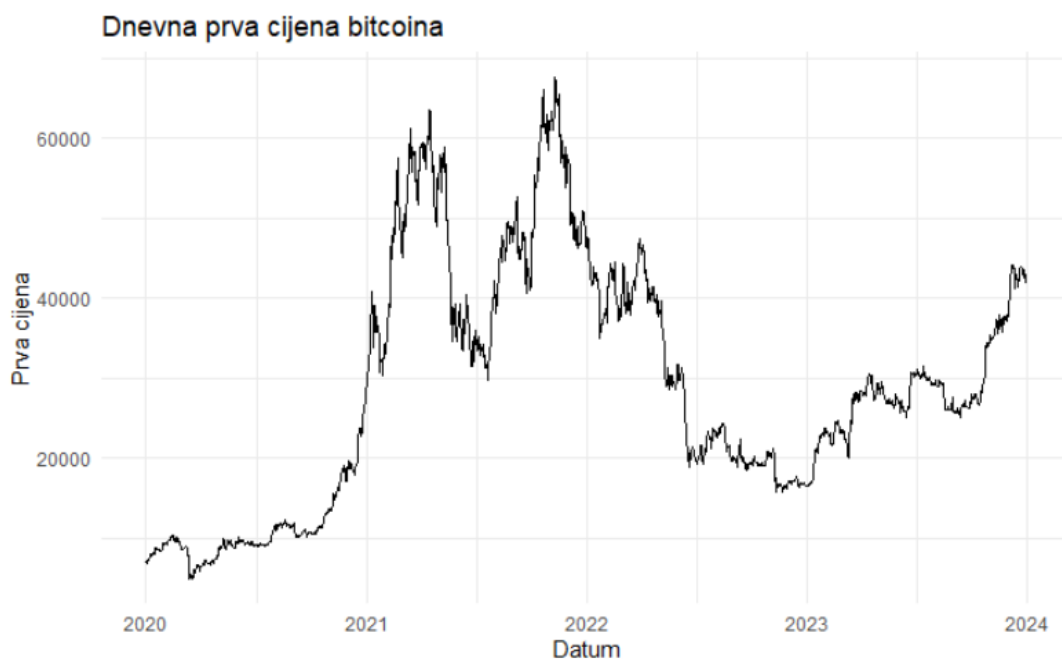
Zatim slijedi učitavanje podataka iz CSV datoteka.

```
# Učitavanje podataka iz CSV datoteka
df <- read.csv("BTC-USD.csv")
df$Date <- as.Date(df$Date)
df2 <- read.csv("BTCweekly.csv")
df2$Date <- as.Date(df2$Date)
```

Nastavlja se kreiranje grafikona koji će prikazivati dnevne prve cijene bitcoina. Prve cijena bitcoina predstavlja cijenu na početku tog dana. Upotrijebit ćemo ggplot2 paket za generiranje linijskog grafikona. Ggplot2 je paket za R koji omogućuje stvaranje vizualizacija podataka koristeći konceptualni okvir temeljen na "gramatici grafike". Ovaj paket omogućuje da definiramo kako će se podaci prikazivati na grafovima [9].

```
# Kreiranje grafikona za dnevne prve cijene bitcoina
ggplot(df, aes(x = Date, y = Open)) +
  geom_line() +
  labs(title = "Dnevna prva cijena bitcoina", x = "Datum",
        y = "Prva cijena") +
  theme_minimal()
```

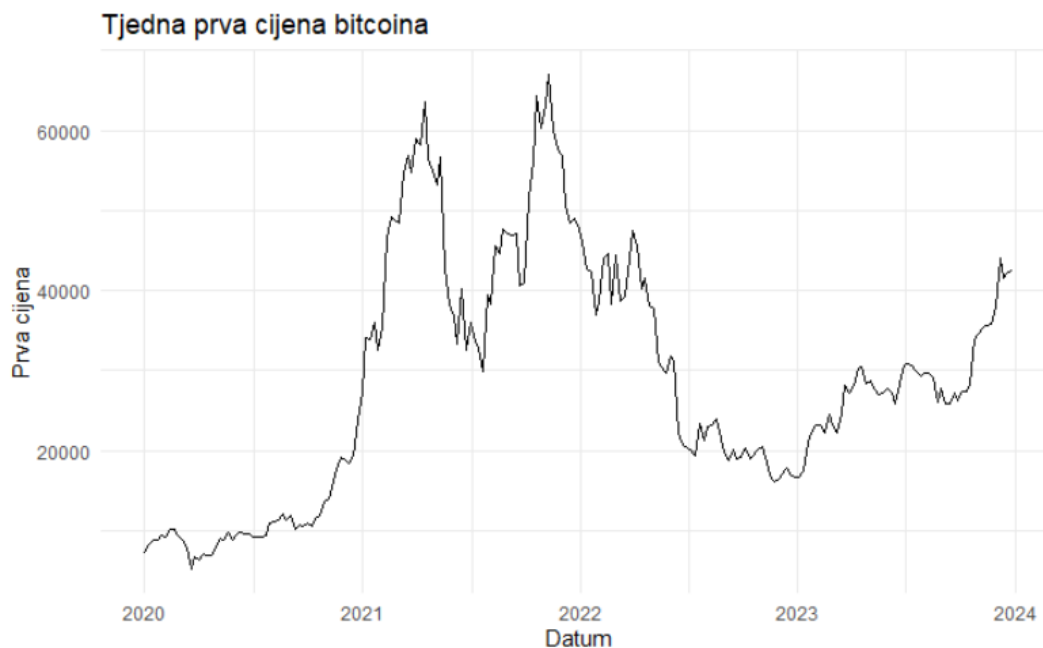
Izvršenje ovog koda u R-u koji koristi paket ggplot2 za kreiranje grafikona rezultira grafikonom koji prikazuje kako su se dnevne prve cijene bitcoina mijenjale od 01.01.2020. do 01.01.2024. godine. Vizualizacija izgleda ovako:



Slika 1: Dnevna prva cijena (Izvor: vlastiti rad, 2024.)

Slika 1 prikazuje graf dnevne prve cijene bitcoina. Najviše vrijednosti su zabilježene 2021. godine, dok se od 2022. do 2023. godine vidi jasan trend pada cijena. Međutim, unatoč oscilacijama, nakon tog perioda primjetan je ponovni rast. Ovdje se jasno vidi tipičan graf vremenskog niza, koji pokazuje kako se podaci mijenjaju kroz vrijeme.

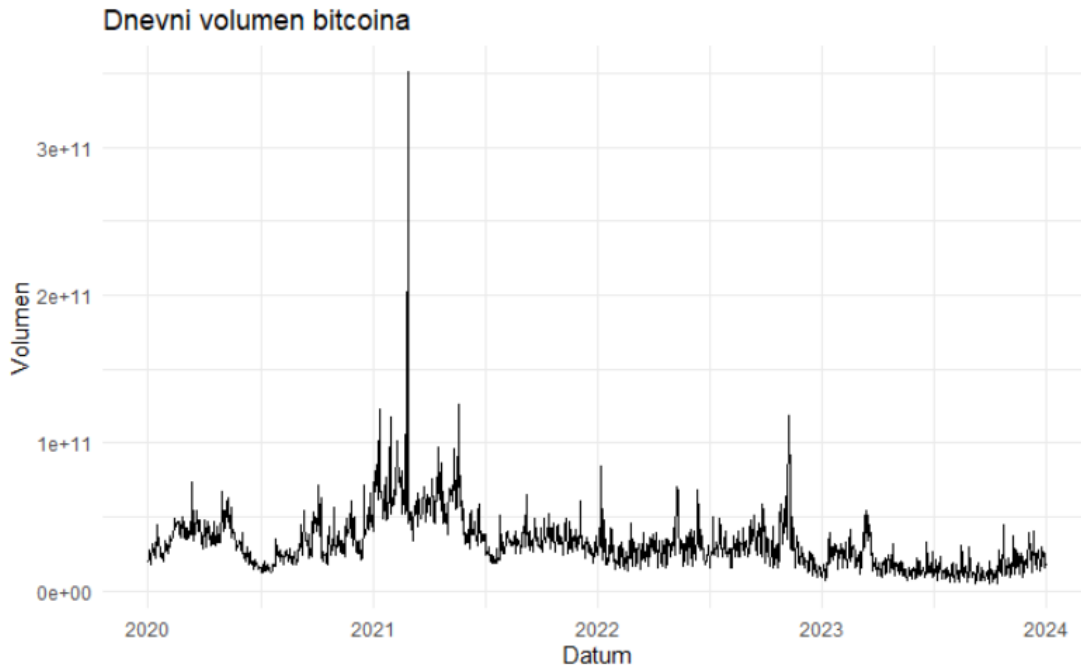
Nastavljamo s prikazom ostalih grafova, uključujući tjednu prvu cijenu, dnevni volumen i tjedni volumen. Kod za generiranje ovih grafova je gotovo identičan prethodnome pa ga nećemo ovdje navoditi (može se pronaći u priloženoj rmd datoteci). Za vizualizacija tjednih prvih cijena bitcoina koristit će se sličan pristup kao i za dnevne prve cijene .



Slika 2: Tjedna prva cijena (Izvor: vlastiti rad, 2024.)

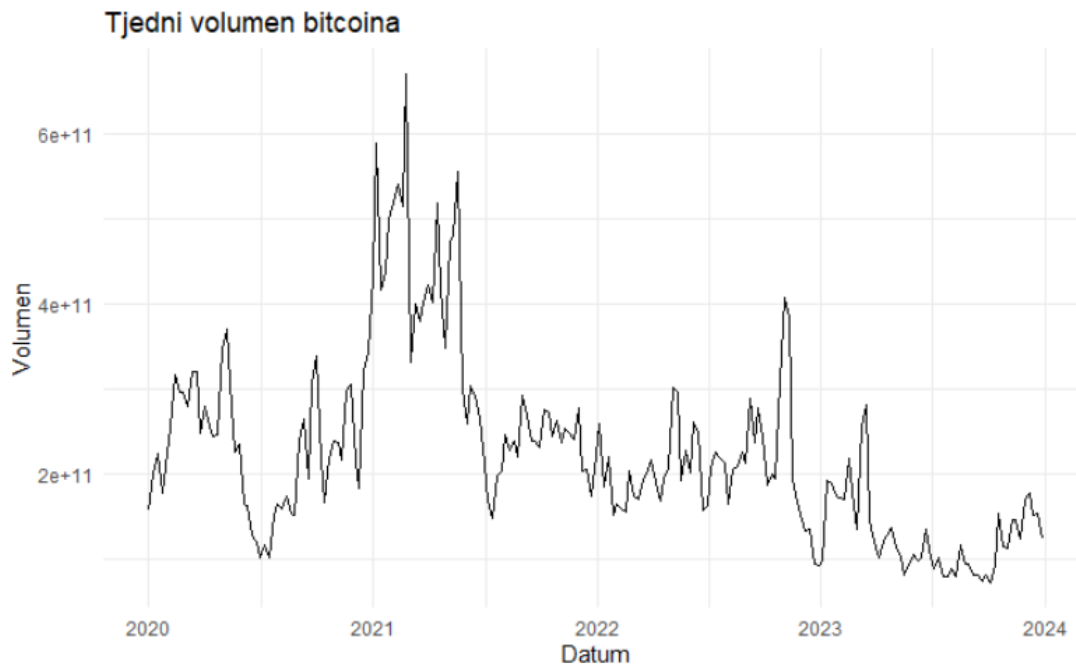
Slika 2 prikazuje tjedne prve cijene. Usporedimo sa pretodnim dnevnim grafom. Iako se temeljni trend rasta cijena i dalje vidi, linija je "glađa" u odnosu na dnevni graf zbog tjednih podataka. Time je na tjednom grafu lakše uočiti dugoročne trendove i smjer kretanja cijena bez utjecaja svakodnevnih oscilacija. Ovo "zaglađivanje" linije pomaže u boljem prepoznavanju osnovnog trenda, što je korisno za analize s dugoročnijom perspektivom.

Oba grafa pokazuju sličan uzorak kretanja cijena, ali tjedni graf nudi stabilniju sliku ukupnog trenda, dok dnevni graf ističe kratkoročne oscilacije.



Slika 3: Dnevni volumen trgovanja(Izvor: vlastiti rad, 2024.)

Slika 3 prikazuje dnevni volumen trgovanja bitcoina od 2020. do 2024. godine. Volumen trgovanja odnosi se na ukupnu količinu bitcoina koja je kupljena ili prodana u tome vremenskom periodu što pruža uvid u razinu aktivnosti na tržištu. Na grafu su jasno vidljivi značajni skokovi volumena tijekom 2021. godine. To razdoblje rezultira s naglim rastom cijene bitcoina, kada je dosegao rekordne vrijednosti. Zatim, nakon vrhunca u 2021., volumen trgovanja opada, iako postoje povremeni skokovi u 2022. i 2023. godini. To bi moglo ukazivati na stabilizaciju tržišta ili smanjenje interesa ulagača u odnosu na prethodne godine.



Slika 4: Tjedni volumen trgovanja (Izvor: vlastiti rad, 2024.)

Slika 4 prikazuje tjedni volumen trgovanja bitcoina. Sličan je prethodnom grafu, ali ovdje se promatra volumen na tjednoj razini, što može pružiti širi pregled. Graf pokazuje slične vrijednosti kao i dnevni, s naglašenim skokovima tijekom 2021. godine. Ipak, tjedni volumen pokazuje nešto glađe prijelaze između skokova i padova, što je uobičajeno kada se podaci objedine na tjednoj razini.

Kao i na dnevnom grafu, najveći volumeni trgovanja bilježi se 2021. godine, ali u tjednim podacima vidimo nešto više kontinuiteta između perioda visoke aktivnosti, dok dnevni graf pokazuje nagle oscilacije.

Manji volumen nakon 2021. godine. Tjedni graf jasno prikazuje stabilizaciju nakon 2021. godine, gdje volumen opada, iako su povremeni skokovi prisutni i u 2022. i 2023. godini. Također, u tjednom prikazu ove oscilacije izgledaju manje dramatično nego u dnevnom prikazu.

Oba grafa pokazuju slične tržišne trendove, ali iz različitih perspektiva. Dnevni graf je korisniji za analizu kratkoročnih promjena, dok tjedni graf pruža stabilniji prikaz dugoročnih tržišnih obrazaca. Iz oba grafa možemo zaključiti da su najaktivniji periodi trgovanja bili u 2021., s kasnijim opadanjem interesa i volumenom koji se stabilizirao na nižim razinama u narednim godinama.

6. Analiza sezonalnosti i trenda

U ovom poglavlju analizirat će se sezonalnost i trend u vremenskim nizovima cijena bitcoina koristeći metode autokorelacije (engl. *Autocorrelation function - ACF*) i djelomične autokorelacije (engl. *Partial autocorrelation function - ACF*). Prema [10] Ovi statistički alati omogućuju identifikaciju ponavljajućih obrazaca i povezanosti unutar vremenskog niza podataka. Ponavljajući obrasci, poput sezonskih ciklusa ili trendova, mogu se otkriti analizom autokorelacijskih funkcija. Na primjer, sezonalnost podrazumijeva da se vrijednosti niza ponavljaju u redovitim intervalima, dok trendovi ukazuju na dugoročne promjene u smjeru cijena. Što se tiče odnosa unutar podataka, ACF i PACF pomažu identificirati u kojoj mjeri su prethodne vrijednosti na različitim vremenskim odmacima (lagovima) povezane s trenutnom vrijednošću. Na taj način možemo razumjeti koliko su prošle cijene relevantne za predviđanje budućih cijena.

Autokorelacija je ključna mjera koja pokazuje koliko je trenutna vrijednost vremenskog niza povezana sa svojim prošlim vrijednostima. Kroz autokorelacijsku funkciju (ACF) možemo vizualizirati stupanj povezanosti između trenutne vrijednosti i vrijednosti koje su prethodile za različite vremenske intervale. Ova funkcija omogućuje da prepoznamo ponavljajuće obrasce, odnosno sezonalnost i trendove, koji su prisutni u podacima. ACF je ključan alat za analizu vremenskih nizova zbog svoje sposobnosti otkrivanja ovih obrazaca [11].

Djelomična autokorelacija (PACF), s druge strane, mjeri direktnu povezanost između trenutne vrijednosti i njezinih prošlih vrijednosti, dok eliminira utjecaj međusobno povezanih vrijednosti između njih. Ova metoda omogućuje dublji uvid u strukturu podataka, pomažući u prepoznavanju značajnih direktnih veza između različitih vremenskih točaka [10].

Drugim riječima, ACF pokazuje kako svaka vrijednost u nizu ovisi o prošlim vrijednostima, uključujući sve međusobne odnose između tih vrijednosti. Međutim, ACF ne razlikuje izravne veze od neizravnih. Na primjer, ako trenutna vrijednost ovisi o vrijednosti dva vremenska trenutka unatrag, ACF neće moći jasno razlikovati tu direktnu povezanost od one koja je posredovana vremenskom točkom između njih. Djelomična autokorelacijska funkcija (PACF) koristi se za rješavanje ovog problema. Ona mjeri samo direktnu povezanost između trenutne vrijednosti i prošlih vrijednosti, eliminirajući utjecaj međusobno povezanih vremenskih točaka. Na primjer, PACF će jasno pokazati povezanost između trenutne vrijednosti i vrijednosti dva vremenska trenutka unatrag, bez utjecaja vrijednosti koje su u međuvremenu zabilježene. Zbog toga PACF omogućuje jasniju sliku o tome koliko daleke prošle vrijednosti izravno utječu na trenutnu vrijednost.

Kroz ovu analizu, koristit će se dnevni i tjedni podaci o prvim cijenama i volumenu trgovanja bitcoina od 1. siječnja 2020. do 1. siječnja 2024. godine. Podaci uključuju cijene i volumen trgovanja, što omogućuje sveobuhvatnu analizu sezonalnosti i trendova. Rezultati ove analize bit će korisni za bolje razumijevanje dinamike tržišta bitcoina i za izradu pouzdanih predikcijskih modela.

6.1. Autokorelacija i djelomična autokorelacija

U ovom podpoglavlju, detaljno će se analizirati autokorelacija (ACF) i djelomičnu autokorelacija (PACF) vremenskih nizova cijena bitcoina. Koristit će se dnevni i tjedni podaci o prvim cijenama bitcoina od 1. siječnja 2020. do 1. siječnja 2024. godine.

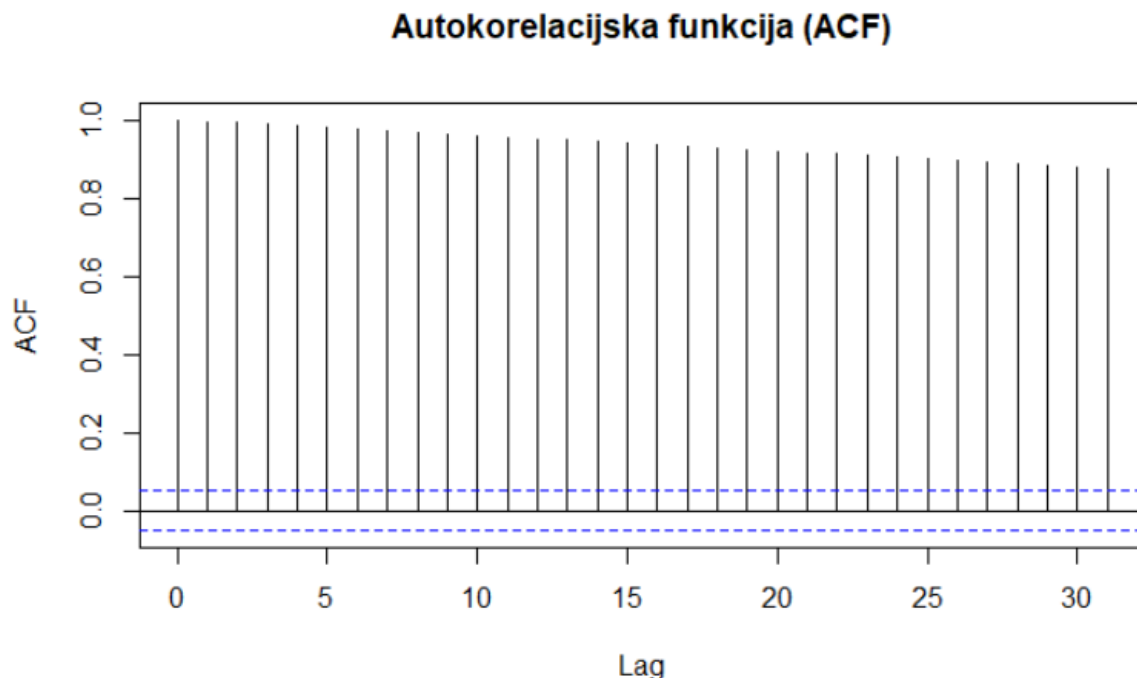
Sljedeći kod u R-u koristi funkcije `acf` i `pacf` za kreiranje autokorelacijskih i parcijalnih autokorelacijskih grafova.

```
# Kreiranje vremenskih nizova za cijene bitcoina
bitcoin_ts <- ts(df$Open, frequency = 1)
acf(bitcoin_ts, main = "Autokorelacijska funkcija (ACF)")

bitcoin_ts2 <- ts(df2$Open, frequency = 1)
acf(bitcoin_ts2, main = "Autokorelacijska funkcija (ACF)")

pacf(bitcoin_ts, main = "Djelomična autokorelacijska funkcija (PACF)")
pacf(bitcoin_ts2, main = "Djelomična autokorelacijska funkcija (PACF)")
```

Slijede rezultati analiza ACF i PACF te analiza svakog grafa pojedinačno radi identificiranja sezonalnosti i trendova u podacima.



Slika 5: Autokorelacijska funkcija za dnevnu cijenu bitcoina (Izvor: vlastiti rad, 2024.)

Slika 5 prikazuje grafikon autokorelacijsku funkciju (ACF) za dnevnu cijenu bitcoina. Autokorelacijska funkcija pokazuje stupanj povezanosti trenutne vrijednosti vremenskog niza s njegovim prošlim vrijednostima na različitim vremenskim zaostacima (engl. *lagovima*).

Vremenski zaostaci predstavljaju broj vremenskih perioda unatrag u kojem se mjeri povezanost trenutnih vrijednosti vremenskog niza s prošlim vrijednostima. Na primjer, lag 1 označava korelaciju između trenutne vrijednosti i vrijednosti iz prethodnog vremenskog trenutka (npr. jučerašnja cijena bitcoina u odnosu na današnju), lag 2 mjeri korelaciju između trenutne vrijednosti i vrijednosti prije dva vremenska perioda i tako dalje.

U grafu ACF-a, svaka vertikalna linija prikazuje vrijednost autokorelacije za specifičan vremenski zaostatak. Visina linije ukazuje na snagu korelacije, dok se horizontalna os odnosi na broj vremenskih perioda zaostataka. Ako je autokorelacija visoka za određeni lag, to znači da je trenutna vrijednost niza jako povezana s vrijednošću niza iz tog prethodnog perioda.

Ključni uvidi:

Visoka početna autokorelacija:

Slika 5 prikazuje visoke autokorelacijske vrijednosti na početku, a zatim postepeno opadaju. Početna vrijednost autokorelacije (na lagu 0) iznosi 1, što je očekivano jer je to korelacija niza sa samim sobom.

Postepeni pad autokorelacije:

Autokorelacija se smanjuje kako se povećava broj vremenskog zaostatka. Ovaj postepeni pad sugerira da postoji dugoročna povezanost između dnevnih cijena bitcoina, ali ta povezanost slabi kako se udaljavamo od trenutne vrijednosti.

Značajne autokorelacije:

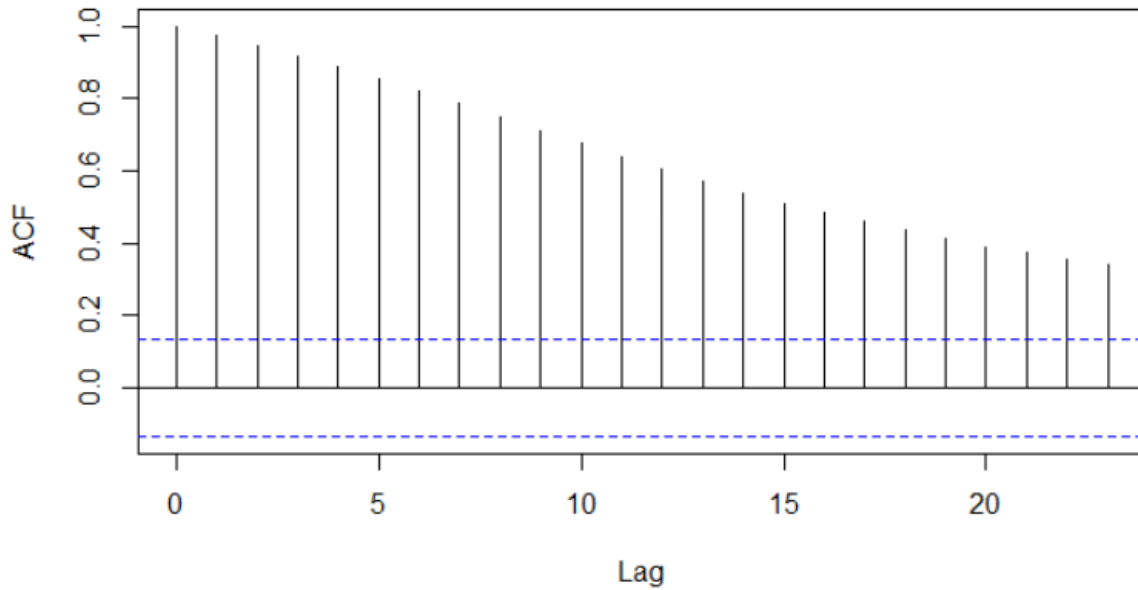
Plave isprekidane linije na grafikonu predstavljaju granice značajnosti. Ako autokorelacijske vrijednosti prelaze ove granice, možemo zaključiti da su te autokorelacije statistički značajne. U ovom slučaju, autokorelacijske vrijednosti za sve lagove prikazane na grafikonu su značajno iznad granica značajnosti, što ukazuje na jaku povezanost između trenutnih i prošlih vrijednosti cijena bitcoina.

Mogući trend:

Postepeni pad autokorelacija također može ukazivati na prisutnost trenda u podacima. Ako postoji trend, to znači da se niza ne vraća brzo na svoju srednju vrijednost, što rezultira visokim autokorelacijama na većem broju lagova.

Grafikon ACF za dnevnu cijenu bitcoina pokazuje visoku početnu autokorelaciju koja postepeno opada s povećanjem broja lagova. Sve autokorelacije prikazane na grafikonu su statistički značajne, što ukazuje na jaku povezanost između trenutnih i prošlih vrijednosti cijena. Ovi rezultati sugeriraju da postoji dugoročna povezanost i mogući trend u nizu, što je važno razmotriti prilikom modeliranja i predviđanja budućih vrijednosti cijena bitcoina.

Autokorelacijska funkcija (ACF)

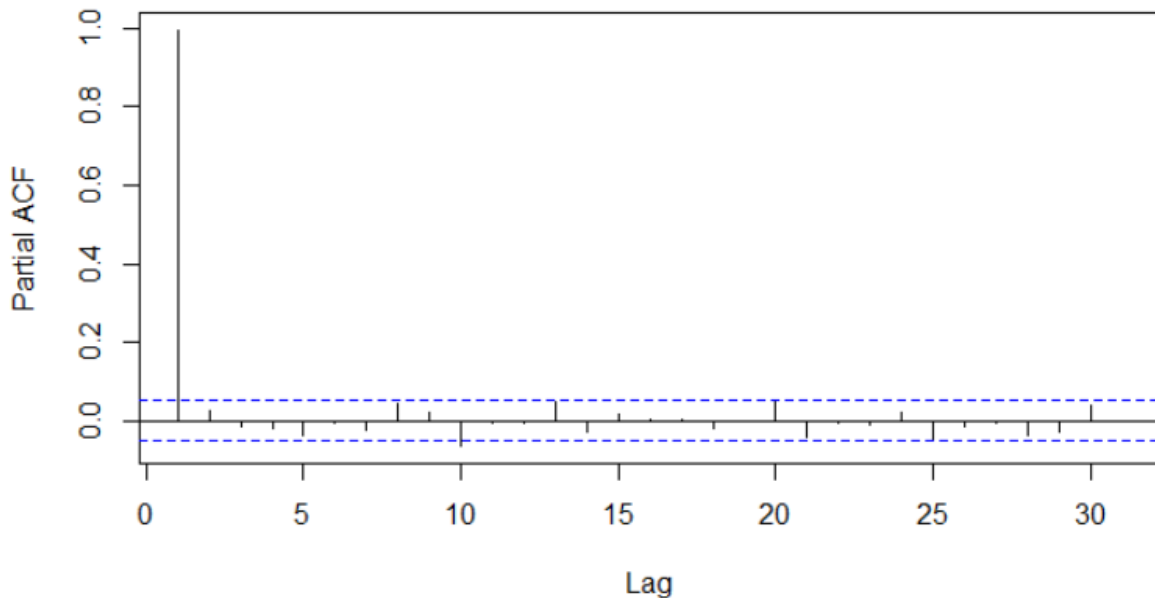


Slika 6: Autokorelacijska funkcija za tjednu cijenu bitcoina (Izvor: vlastiti rad, 2024.)

Slika 6 prikazuje grafikon autokorelacijske funkcije (ACF) za tjedne cijene bitcoina. Prikazuje visoku početnu autokorelaciju koja postepeno opada s povećanjem broja lagova. Ovaj obrazac sugerira da postoji značajna povezanost između trenutnih i prošlih vrijednosti cijena bitcoina na tjednoj razini, ali ta povezanost slabi kako se povećava vremenski razmak.

Uspoređujući s prethodnim grafom ACF za dnevne cijene bitcoina, može se primijetiti sličan trend postepenog opadanja autokorelacije. Oba grafa pokazuju da postoje dugoročne veze između vrijednosti niza, što je važan pokazatelj prisutnosti trenda. Međutim, tjedni podaci pokazuju stabilniji i sporiji pad autokorelacije, što može ukazivati na dulji period povratka na srednju vrijednost.

Djelomična autokorelacijska funkcija (PACF)



Slika 7: Djelomična autokorelacijska funkcija za dnevnu cijenu bitcoina (Izvor: vlastiti rad, 2024.)

Slika 7 prikazuje djelomičnu autokorelacijsku funkciju (PACF) za dnevnu cijenu bitcoina. PACF pokazuje stupanj povezanosti trenutne vrijednosti vremenskog niza s njegovim prošlim vrijednostima, eliminirajući utjecaj međusobno povezanih vrijednosti između njih.

Ključni uvidi:

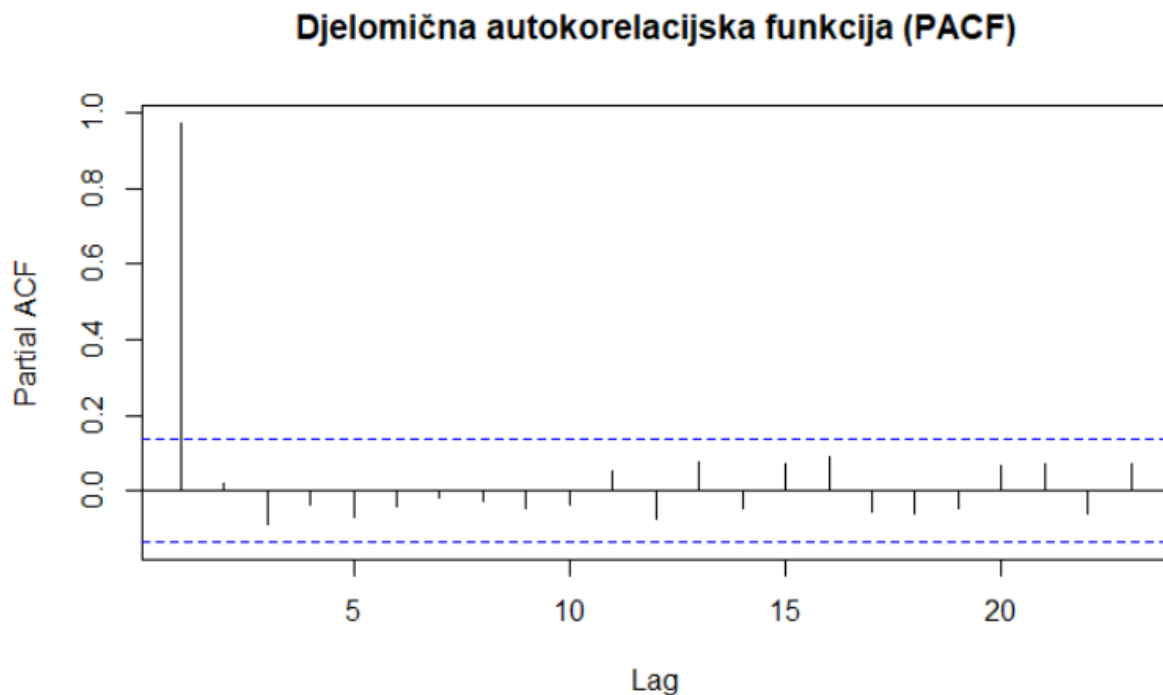
Značajna autokorelacija na prvom lagu:

Na grafikonu je vidljivo da je autokorelacija značajna samo na prvom lagu (lag 1). To znači da je trenutna vrijednost dnevne cijene bitcoina najviše povezana sa svojom neposredno prethodnom vrijednošću. Vrijednosti za sve ostale lagove su unutar granica značajnosti (prikazane plavim isprekidanim linijama), što ukazuje na to da nema značajnih direktnih povezanosti između trenutne vrijednosti i bilo koje druge prethodne vrijednosti osim prve.

Eliminacija utjecaja posrednih vrijednosti:

PACF eliminira utjecaj posrednih vrijednosti, što znači da su prikazane korelacije očišćene od efekata koje imaju međusobno povezane vrijednosti. U ovom slučaju, nakon prvog laga, sve druge vrijednosti postaju neznatne.

Dakle, grafikon PACF za dnevnu cijenu bitcoina pokazuje značajnu autokorelaciju samo na prvom lagu, dok su sve ostale vrijednosti unutar granica značajnosti. Ovi rezultati sugeriraju da trenutna vrijednost dnevne cijene bitcoina ima jaku povezanost samo s neposredno prethodnom vrijednošću, a ostale prethodne vrijednosti nemaju značajan direktan utjecaj.



Slika 8: Djelomična autokorelacijska funkcija za tjednu cijenu bitcoina (Izvor: vlastiti rad, 2024.)

Uspoređujući s PACF grafom za dnevne cijene bitcoina, može se primijetiti sličan obrazac u grafikonu djelomičnes autokorelacijske funkcije za tjednu cijenu (Slika 8). U oba slučaja, autokorelacija je značajna samo na prvom lagu, dok su sve ostale vrijednosti unutar granica značajnosti. To znači da za oba vremenska niza (dnevne i tjedne cijene bitcoina), trenutna vrijednost ima značajnu povezanost samo s neposredno prethodnom vrijednošću.

Iz grafova cijena bitcoina kroz godinu dana, i iz grafova ACF-a može se utvrditi da cijena bitcoina ne pokazuje sezonalnost. Kada bi sezonalnost postojala, tada ne bi bilo konstantnog pada u ACF-u s povećanjem razlika; isto is PACF-om. Postojali bi rastovi na određenim razlikama koji bi ukazivali na to da postoji veza između cijena u određenim tjednim razlikama, tj. postojala bi korelacija između promatranog tjedna i nekog n-tog tjedna iza tog tjedna. Drugim riječima, tada bi to pokazivalo sezonalnost ovog vremenskog niza.

7. Korelacija između dva vremenska niza

U ovom poglavlju analizirat će se korelacija između tjedne cijene bitcoina i volumena trgovanja koristeći Pearsonov koeficijent korelacije. Pearsonov koeficijent korelacije označava se slovom r . Formula za Pearsonov koeficijent korelacije r dana je na sljedeći način:

$$r = \frac{n(\sum xy) - (\sum x)(\sum y)}{\sqrt{[n \sum x^2 - (\sum x)^2][n \sum y^2 - (\sum y)^2]}}$$

Gdje je:

- x = vrijednosti u prvom skupu podataka
- y = vrijednosti u drugom skupu podataka
- n = ukupan broj vrijednosti.

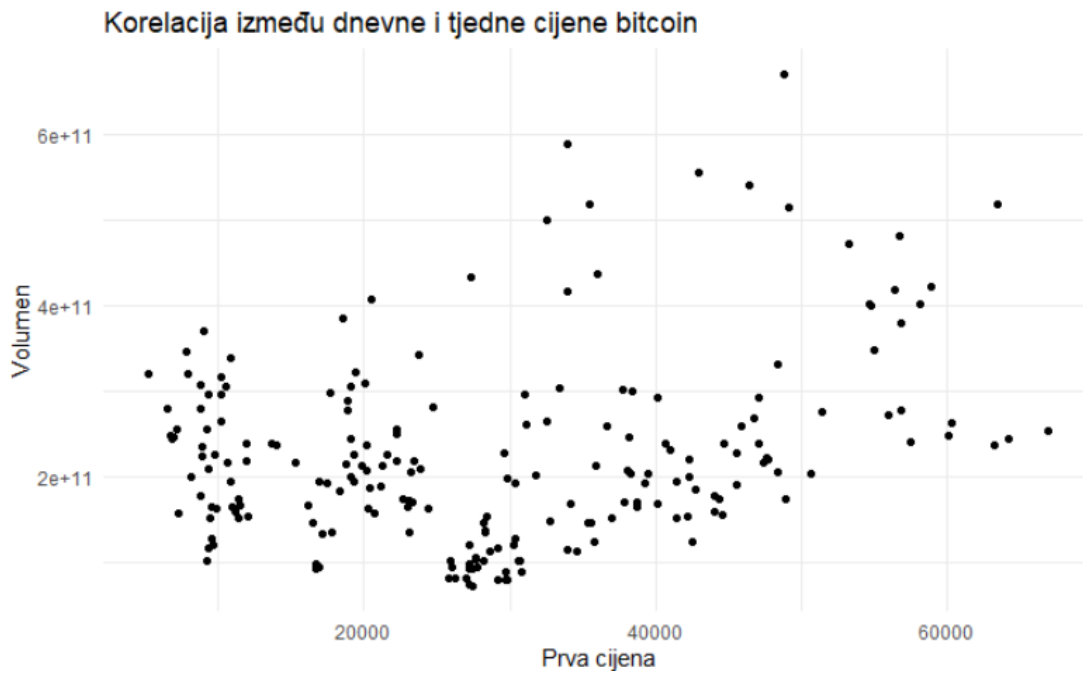
Pearsonov koeficijent korelacije r poprima vrijednosti od -1 do $+1$. Vrijednosti -1 ili $+1$ označavaju savršen linearni odnos između dvije varijable, dok vrijednost 0 označava nepostojanje linearnog odnosa. Koeficijenti korelacije koji se razlikuju od 0 , ali nisu -1 ili $+1$, isto tako označavaju linearni odnos, iako ne savršen linearni odnos [12].

Pretpostavlja se da postoji pozitivna korelacija između cijene bitcoina i volumena trgovanja, to jest očekujemo da će veća cijena bitcoina uzrokovati veći volumen trgovanja. Ova pretpostavka se temelji na logici da više ljudi trguje kada cijena raste, što dovodi do povećanja volumena trgovanja. Korelaciju možemo provjeriti vizualno pomoću dijagrama raspršenja (engl. *scatter plot*), a kvantitativno izračunavanjem Pearsonovog koeficijenta korelacije.

Vizualizacija

Prvo, prikazat će se dijagram raspršenja koji prikazuje odnos između tjednih prvih cijena i volumena trgovanja bitcoina.

```
# Prikaz korelacije između dnevne i tjedne cijene bitcoina
ggplot(df2, aes(x = Open, y = Volume)) +
  geom_point() +
  labs(title = "Korelacija između dnevne i tjedne cijene bitcoin",
x = "Prva cijena", y = "Volumen") +
  theme_minimal()
```

Slika 9: Prikaz korelacije između dnevne i tjedne cijene bitcoina (Izvor: vlastiti rad, 2024.)

Slika 9 prikazuje dijagram raspršenja na kojem se može uočiti postoji li vidljiva linearna povezanost između prvih cijena i volumena trgovanja. Vizualno se može uočiti tendencija da veće prve cijene prate veći volumeni trgovanja, što sugerira postojanje pozitivne korelacije.

Kvantitativna analiza

```
# Izračun Pearsonove korelacije između prve cijene i volumena
pearson_cor <- cor.test(df2$Open, df2$Volume, method = "pearson")
print(paste("Pearsonov koeficijent korelacije:",
round(pearson_cor$estimate, 2)))
print(paste("P-vrijednost:", format(pearson_cor$p.value,
scientific = TRUE)))
```

Ovaj kod u R-u koristi funkciju `cor.test` za izračun Pearsonovog koeficijenta korelacije između prve cijene i volumena trgovanja.

Rezultati analize su sljedeći:

Pearsonov koeficijent korelacije: 0.28

P-vrijednost: 0.00004875447

Pearsonov koeficijent korelacije od 0.28 ukazuje na umjerenu pozitivnu korelaciju između tjednih prvih cijena bitcoina i volumena trgovanja. To znači da postoji tendencija da se veći volumeni trgovanja povezuju s višim cijenama, iako ova korelacija nije izrazito jaka. Niska p-vrijednost (0.00004875447) ukazuje na to da je korelacija statistički značajna, što znači da je malo vjerojatno da je ova povezanost nastala slučajno. Ovi rezultati podržavaju početnu hipotezu da se povećanje cijene bitcoina povezuje s povećanjem volumena trgovanja.

8. ARIMA model

U analizi vremenskih nizova, ARIMA model jedan je od najčešće korištenih i najmoćnijih alata za predviđanje budućih vrijednosti. Prema [13] ARIMA model kombinira tri osnovne komponente: autoregresiju (engl. *Autoregression* - AR), integraciju (engl. *Integration* - I) i pomični prosjek (engl. *Moving Average* - MA). Ovaj model je posebno koristan za vremenske nizove koji pokazuju određeni stupanj nestacionarnosti i potrebu za uklanjanjem trendova ili sezonalnosti kako bi se postigla stacionarnost niza. Autoregresijski (AR) dio modela koristi povezanost između trenutne vrijednosti i njezinih prošlih vrijednosti. Integracijski (I) dio modela odnosi se na diferenciranje vremenskog niza kako bi se uklonila nestacionarnost. Diferenciranje je proces oduzimanja trenutne vrijednosti od prethodne vrijednosti u nizu. Ako niz nije stacionaran, može biti potrebno višestruko diferenciranje kako bi se uklonili eventualni trendovi i sezonalnosti. Pomični prosjek (MA) dio modela koristi prošle pogreške predviđanja kako bi poboljšao točnost predikcije.

Prema [14] ARIMA modeli imaju primjenu u ovim područjima:

Ekonomija i financije: ARIMA modeli koriste se za modeliranje i predviđanje ekonomskih i financijskih vremenskih nizova, kao što su cijene dionica, tečajevi i stope inflacije. Ti se modeli mogu koristiti za prepoznavanje trendova, ciklusa i drugih obrazaca u podacima, što može pomoći u donošenju informiranih odluka o ulaganju.

Prodaja i marketing: Isto tako koriste se za predviđanje prodaje i potražnje za proizvodima i uslugama. Analizirajući prošle podatke o prodaji, tvrtke mogu koristiti ARIMA modele za prepoznavanje sezonskih trendova i drugih obrazaca, što im može pomoći u donošenju boljih odluka o upravljanju zalihama, planiranju proizvodnje i cijenama.

Energija: ARIMA modeli koriste se za predviđanje potražnje za energijom, kao što je potrošnja električne energije i plina. Ovi modeli mogu pomoći komunalnim poduzećima da učinkovitije planiraju svoje proizvodne i distribucijske aktivnosti, a također pomažu vladama da planiraju buduće energetske potrebe.

Zdravstvo: ARIMA modeli koriste se u medicinskim istraživanjima za modeliranje i predviđanje izbijanja bolesti, prijema u bolnicu i drugih vremenskih nizova podataka povezanih sa zdravljem. Ovi modeli mogu pomoći službenicima javnog zdravstva u planiranju budućih zdravstvenih potreba i učinkovitijoj raspodjeli resursa.

Vremenska prognoza: ARIMA modeli se koriste u vremenskoj prognozi za modeliranje i prognozu različitih vremenskih varijabli, kao što su temperatura, oborine i ostalo. Ovi modeli mogu pomoći meteorolozima da naprave točna predviđanja vremena, koja se mogu koristiti za zaštitu ljudi i imovine od ekstremnih vremenskih događaja.

Općenito, ARIMA modeli korisni su kad god postoji potreba za modeliranjem i predviđanjem vremenskih nizova podataka te za prepoznavanje trendova, ciklusa i drugih obrazaca u podacima.

Cilj ovog poglavlja je pokazati kako koristiti ARIMA model za predviđanje tjedne cijene bitcoina. Provest će se ADF test za provjeru stacionarnosti niza, izraditi ARIMA model koristeći funkciju `auto.arima`, provjeriti stacionarnost reziduala modela i napraviti predikcije za buduće vrijednosti cijene bitcoina.

8.1. Stacionarnost vremenskog niza

Stacionarnost je ključni koncept u analizi vremenskih nizova. Stacionarni vremenski niz je onaj čija svojstva ne ovise o vremenu u kojem se niz promatra [15]. Dakle, ako je vremenski niz stacionaran, onda njegove statističke karakteristike, kao što su srednja vrijednost i varijanca, ostaju nepromijenjene kroz vrijeme. Prema [16] prije nego što se primijeni ARIMA model, mora se provjeriti stacionarnost niza. Mnoge statističke metode i modeli, uključujući ARIMA, zahtijevaju da vremenski niz bude stacionaran kako bi bili učinkoviti i precizni u predviđanjima. ARIMA modeli pretpostavljaju da su obrasci unutar podataka stabilni i ne mijenjaju se kroz vrijeme. Ako niz nije stacionaran, može biti potrebno primijeniti transformacije poput diferenciranja kako bi se postigla stacionarnost.

Da bi se utvrdilo je li vremenski niz stacionaran, koristi se ADF test. ADF test je proširenje Dickey-Fuller testa koji uključuje dodatne lagove diferenciranih varijabli kako bi se riješila autokorelacija u rezidualima. Da bi se utvrdilo je li vremenski niz stacionaran, koristi se Augmented Dickey-Fuller (ADF) test. ADF test je proširenje Dickey-Fuller testa koji poboljšava osnovnu metodu dodavanjem dodatnih vremenskih zaostataka diferenciranih varijabli. To znači da se u regresijskom modelu koristi više prošlih vrijednosti niza kako bi se obuhvatili svi mogući utjecaji prošlih vrijednosti na trenutne, što omogućava precizniju procjenu stacionarnosti niza. Diferenciranje varijabli je ključni korak u ovom testiranju. Diferenciranje se postiže oduzimanjem prethodne vrijednosti od trenutne ili od vrijednosti u nekom drugom vremenskom trenutku. Na taj način dolazi do uklanjanja trenda ili sezonalnosti iz niza. Nadalje, ako su reziduali, odstupanja između stvarnih vrijednosti i onih predviđenih modelom, povezani s prošlim vrijednostima (autokorelirani), to može značiti da model nije u potpunosti obuhvatio sve obrasce u podacima. U kontekstu ADF testa, cilj je osigurati da reziduali budu neovisni i da ne pokazuju značajnu autokorelaciju. Ako reziduali pokazuju autokorelaciju, to može ukazivati na nedostatke u modelu, sugerirajući da još uvijek postoji struktura u podacima koja nije obuhvaćena modelom. Idealno, reziduali bi trebali biti nasumični i neovisni, čime se osigurava valjanost modela i rezultata testiranja [17].

ADF test postavlja sljedeće hipoteze:

Nulta hipoteza (H_0): Vremenski niz je nestacionaran.

Alternativna hipoteza (H_1): Vremenski niz je stacionaran.

Ako je p-vrijednost testa manja od odabrane razine značajnosti (obično 0.05), odbacujemo nultu hipotezu o nestacionarnosti niza.

Slijede linije koda koje prikazuju kako u RStudiju pozivamo ADF test na odabranom skupu podataka.

```
# Prošireni Dickey-Fuller test za stacionarnost
adf_result <- adf.test(df2$Open)
print(adf_result)
```

Augmented Dickey-Fuller Test

```
data: df2$Open
Dickey-Fuller = -2.0317, Lag order = 5, p-value = 0.5627
alternative hypothesis: stationary
```

Slika 10: Ispis iz R-a. Rezultat ADF testa (Izvor: vlastiti rad, 2024.)

Rezultat provedenog ADF testa govori da se na temelju p-vrijednosti ne može odbaciti nulta hipoteza, što znači da promatrani vremenski niz (tjedna cijena bitcoina) može biti nestacionaran. To se može iščitati iz p-vrijednosti koja je 0.5627, što je više od 0.05. Ovo potvrđuje ono što se moglo vidjeti i vizualno iz grafa tjednih cijena bitcoina; naime, možemo vidjeti da taj niz cijena upućuje na nestacionarnost, pogotovo u zadnjih par mjeseci kada je cijena bitcoina naglo skočila. Iako ARIMA model zahtijeva stacionarni niz, nestacionarni niz se može koristiti u ARIMA modelu jer model uključuje integracijski dio koji omogućuje diferenciranje niza kako bi se postigla stacionarnost. Više o tome u sljedećem poglavlju.

Dakle, zaključuje se da postoji trend rasta cijene bitcoina. Kada bi ovaj vremenski niz koji prikazuje godinu dana vremena tjedne cijene bitcoina bila stacionarna, tada ne bi postojao trend zato što bi vremenski niz imao konstantan prosjek i varijancu. Graf i ADF test upućuju na pretpostavku da trend rasta postoji.

8.2. Komponente ARIMA modela

Pream [14] ARIMA model kombinira tri osnovne komponente: p - autoregresiju (AR), d - integraciju (I) i q - pomični prosjek (MA). Svaka od ovih komponenti igra ključnu ulogu u modeliranju složenih obrazaca u vremenskim nizovima. Slijedi kratki opis svake od ovih komponenti i njihov značaj u ARIMA modelu.

Autoregresija (AR) je komponenta modela koja opisuje kako trenutna vrijednost vremenskog niza ovisi o njegovim prethodnim vrijednostima.

Integracija (I) se odnosi na diferenciranje vremenskog niza kako bi se postigla stacionarnost, što će se koristiti u ovom slučaju.

Pomični prosjek (MA) je komponenta modela koja opisuje kako trenutna vrijednost vremenskog niza ovisi o prošlim pogreškama predviđanja.

8.3. Izrada ARIMA modela

Koristi se funkcija `auto.arima` u R-u za pronalaženje najboljeg ARIMA modela s optimalnim p , d i q vrijednostima za tjedne cijene bitcoina. Funkcija `auto.arima` iz R paketa `forecast` automatski pronalazi najbolji ARIMA model za zadani vremenski niz koristeći kriterij za optimizaciju kao što su u ovom slučaju Akaikeov informacijski kriterij (engl. *Akaike Information*

Criterion - AIC) te Bayesov informacijski kriterij (engl. *Bayesian Information Criterion - BIC*). Dakle, funkcija `auto.arima` isprobava različite kombinacije parametara ARIMA modela (p, f, q) i koristi kriterije poput AIC-a i BIC-a kako bi pronašla model koji najbolje odgovara zadanim podacima.[15].

```
# Kreiranje ARIMA modela za prve cijene bitcoina
btcmodel <- auto.arima(df2$Open, ic = "aic", trace = TRUE)
print(btcmodel)
```

Ovaj kod u R-u koristi funkciju `auto.arima` iz paketa `forecast` za automatsko kreiranje ARIMA modela za varijablu `Open` iz skupa podataka koja predstavlja prve cijene bitcoina.

```
Fitting models using approximations to speed things up...
```

```
ARIMA(2,1,2) with drift      : 3922.245
ARIMA(0,1,0) with drift     : 3916.682
ARIMA(1,1,0) with drift     : 3919.377
ARIMA(0,1,1) with drift     : 3918.486
ARIMA(0,1,0)                : 3915.323
ARIMA(1,1,1) with drift     : 3920.001
```

```
Now re-fitting the best model(s) without approximations...
```

```
ARIMA(0,1,0)                : 3931.376
```

```
Best model: ARIMA(0,1,0)
```

```
Series: df2$Open
ARIMA(0,1,0)
```

```
sigma^2 = 9373108: log likelihood = -1964.69
AIC=3931.38  AICc=3931.4  BIC=3934.71
```

Slika 11: Ispis iz R-a. Rezultat funkcije `auto.arima`(Izvor: Vlastiti rad, 2024.)

Funkcija `auto.arima` pronalazi najbolji ARIMA model s najboljim p, f i q vrijednostima. Odabran je model ARIMA (0,1,0). ARIMA(0,1,0) model je odabran ne samo zbog svoje jednostavnosti, već i zbog toga što ima najniže AIC i BIC vrijednosti među testiranim modelima, što ga čini najboljim kandidatom. AIC vrijednost za ARIMA(0,1,0) iznosi 3931.38, dok je BIC vrijednost 3934.71. Ove vrijednosti pokazuju da je odabrani model optimalan, uzimajući u obzir kompromis između preciznosti modela i njegove složenosti.

Model ARIMA(0,1,0) označava:

$p = 0$: Nema autoregresijske komponente. Trenutna vrijednost ne ovisi o prošlim vrijednostima.

$d = 1$: Jedno diferenciranje je provedeno kako bi se postigla stacionarnost niza. Ovo diferenci-

ranje uklanja trendove i čini niz stacionarnim.

$q = 0$: Nema komponente pomičnog prosjeka. Trenutna vrijednost ne ovisi o prošlim pogreškama predviđanja.

8.4. Predikcija tjednih cijena bitcoina

Slijedi izrada modela kojim će se izračunati tjedna cijena bitcoina za sljedećih 12 tjedana uz zadanu 95 postotnu pouzdanost.

```
# Prognoza cijena bitcoina za sljedećih 12 tjedana
btcforecast <- forecast(btcmodel, level = c(95), h = 12)
print(btcforecast)
plot(btcforecast)
```

	Point Forecast <0.01>	Lo 95 <0.01>	Hi 95 <0.01>
210	42518.47	36517.93	48519.00
211	42518.47	34032.43	51004.51
212	42518.47	32125.24	52911.70
213	42518.47	30517.40	54519.54
214	42518.47	29100.87	55936.07
215	42518.47	27820.22	57216.72
216	42518.47	26642.55	58394.39
217	42518.47	25546.40	59490.54
218	42518.47	24516.87	60520.07
219	42518.47	23543.11	61493.82
220	42518.47	22616.95	62419.99
221	42518.47	21732.01	63304.93

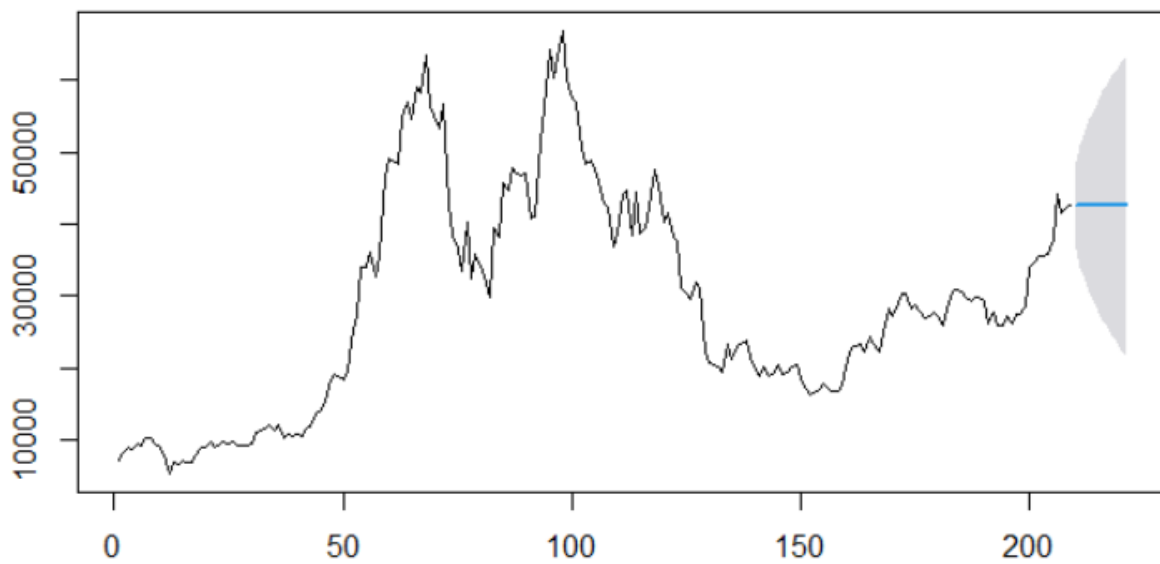
Slika 12: Ispis iz R-a. Prikaz tjednih cijena bitcoina za sljedećih 12 tjedana (Izvor: vlastiti rad, 2024.)

Ispis prikazuje predviđene tjedne cijena bitcoina za sljedećih 12 tjedana, njihovu donju i gornju vrijednost, dakle to označava interval cijena za svaki sljedeći tjedan koji je izračunat uz zadanu pouzdanost od 95

Grafički to izgleda ovako:

Plava linija prikazuje predviđenu tjednu cijenu bitcoina. Siva sjena oko te prave linije označava interval u kojem cijena može biti (unutar 95 postotne pouzdanosti).

Forecasts from ARIMA(0,1,0)



Slika 13: Predviđanje tjedne cijene bitcoina za sljedećih 12 tjedana(Izvor: Vlastiti rad, 2024.)

9. Zaključak

U ovom radu analizirani su vremenski nizovi cijena bitcoina koristeći metode autokorelacije (ACF), djelomične autokorelacije (PACF) i ARIMA modela. Predmet rada bio je razumjeti dinamiku cijena bitcoina i predvidjeti buduća kretanja na temelju povijesnih podataka od 1. siječnja 2020. do 1. siječnja 2024. godine.

Korišteni su sljedeći programski alati:

R: Glavni alat za analizu podataka i izradu modela. Paketi poput forecast, ggplot2, xts, zoo, stats i tseries omogućili su sveobuhvatnu analizu i vizualizaciju podataka. Osim toga, za izradu i formatiranje ovog rada korišten je i Overleaf, online aplikacija za pisanje i uređivanje dokumenata u LaTeX-u.

Rezultati analize pokazali su sljedeće:

Stacionarnost: Pomoću ADF testa utvrđeno je da vremenski niz tjedne cijene bitcoina nije stacionarna, ali je postala stacionarna nakon diferenciranja.

ARIMA model: Optimalni model za predikciju tjednih cijena bitcoina bio je ARIMA(0,1,0). Ovaj model je koristio jedno diferenciranje za postizanje stacionarnosti, bez autoregresijske ili komponente pomičnog prosjeka.

Predikcija: Predikcija budućih vrijednosti za sljedećih 12 tjedana pokazala je pouzdane intervale unutar 95 postotne pouzdanosti, što znači da model može pružiti korisne informacije za buduće kretanje cijena.

Također je analizirana korelacija između tjedne cijene bitcoina i volumena trgovanja koristeći Pearsonov koeficijent korelacije. Rezultati su pokazali umjerenu pozitivnu korelaciju, što sugerira da se veći volumeni trgovanja povezuju s višim cijenama. Ova povezanost bila je statistički značajna.

Dakle, ovaj rad potvrđuje da su metode analize vremenskih nizova, posebno ARIMA model, učinkovite za razumijevanje i predviđanje cijena bitcoina. Rezultati mogu biti korisni za investitore, analitičare i sve one koji žele donositi informirane odluke na temelju povijesnih podataka i naprednih statističkih metoda.

Popis literature

- [1] D. Grahovac, *Analiza vremenskih nizova - Dijelovi predavanja*, Posljednja promjena: 24. veljače 2022., 2022.
- [2] 3. D. Science, *Time Series Data - Time Series Analysis Tutorial*, Pristupano: 2024-09-06, 2023. adresa: <https://365datascience.com/tutorials/time-series-analysis-tutorials/time-series-data/>.
- [3] Tableau, *What is Time Series Analysis?* Pristupano: 2024-09-06, 2023. adresa: <https://www.tableau.com/learn/articles/time-series-analysis#definition>.
- [4] CRAN, *xts: eXtensible Time Series*, Pristupano: 2024-09-06, 2023. adresa: <https://cran.r-project.org/web/packages/xts/index.html>.
- [5] SpringerLink, *Time Series Forecasting: Core Concepts and Definitions*, <https://link.springer.com/article/10.1007/s00180-023-01234-5>, Pristupano: 2024-09-08, 2023.
- [6] SpringerLink, *Time Series Forecasting with Statistical, Machine Learning, and Deep Learning Methods: Past, Present, and Future*, <https://link.springer.com/article/10.1007/s00180-023-01235-6>, Pristupano: 2024-09-08, 2023.
- [7] RC2E, *Time Series Analysis*, Pristupano: 2024-09-06, 2024. adresa: <https://rc2e.com/timeseriesanalysis>.
- [8] Y. Finance, *Bitcoin (BTC-USD) Historical Data*, Pristupano: 2024-09-06, 2024. adresa: <https://finance.yahoo.com/quote/BTC-USD/history/>.
- [9] H. Wickham, *ggplot2: Create Elegant Data Visualisations Using the Grammar of Graphics*, Accessed: 2024-09-06, 2023. adresa: <https://www.rdocumentation.org/packages/ggplot2/versions/3.5.0>.
- [10] J. Frost, *Autocorrelation and Partial Autocorrelation in Time Series Forecasting*, Pristupano: 2024-09-06, 2023. adresa: <https://statisticsbyjim.com/time-series/autocorrelation-partial-autocorrelation/>.
- [11] J. Brownlee, *A Gentle Introduction to Autocorrelation and Partial Autocorrelation*, Pristupano: 2024-09-06, 2023. adresa: <https://machinelearningmastery.com/gentle-introduction-autocorrelation-partial-autocorrelation/>.
- [12] BYJU's, *Pearson Correlation Formula*, Pristupano: 2024-09-06, 2024. adresa: <https://byjus.com/pearson-correlation-formula/>.

- [13] Unknown, *Time Series Analysis*, Pristupano: 2024-09-06, n.d. adresa: <https://pzs.dstu.dp.ua/DataMining/times/bibl/TimeSeries.pdf>.
- [14] D. Camp, *ARIMA Model - Complete Guide to Time Series Forecasting*, Pristupano: 2024-09-06, 2024. adresa: <https://databasecamp.de/en/statistics/arima-en>.
- [15] R. J. Hyndman, *Stationarity and differencing*, Pristupano: 2024-09-07, 2024. adresa: <https://otexts.com/fpp2/stationarity.html>.
- [16] M. L. Plus, *ARIMA Model for Time Series Forecasting in Python*, Pristupano: 2024-09-14, 2024. adresa: <https://www.machinelearningplus.com/time-series/arima-model-time-series-forecasting-python/>.
- [17] M. L. Plus, *Augmented Dickey-Fuller Test*, Pristupano: 2024-09-14, 2024. adresa: <https://www.machinelearningplus.com/time-series/augmented-dickey-fuller-test/>.

Popis slika

1.	Dnevna prva cijena (Izvor: vlastiti rad, 2024.)	8
2.	Tjedna prva cijena (Izvor: vlastiti rad, 2024.)	9
3.	Dnevni volumen trgovanja(Izvor: vlastiti rad, 2024.)	10
4.	Tjedni volumen trgovanja (Izvor: vlastiti rad, 2024.)	11
5.	Autokorelacijska funkcija za dnevnu cijenu bitcoina (Izvor: vlastiti rad, 2024.) . . .	13
6.	Autokorelacijska funkcija za tjednu cijenu bitcoina (Izvor: vlastiti rad, 2024.) . . .	15
7.	Djelomična autokorelacijska funkcija za dnevnu cijenu bitcoina (Izvor: vlastiti rad, 2024.)	16
8.	Djelomična autokorelacijska funkcija za tjednu cijenu bitcoina (Izvor: vlastiti rad, 2024.)	17
9.	Prikaz korelacije između dnevne i tjedne cijene bitcoina (Izvor: vlastiti rad, 2024.)	19
10.	Ispis iz R-a. Rezultat ADF testa (Izvor: vlastiti rad, 2024.)	22
11.	Ispis iz R-a. Rezultat funkcije auto.arima(Izvor: Vlastiti rad, 2024.)	23
12.	Ispis iz R-a. Prikaz tjednih cijena bitcoina za sljedećih 12 tjedana (Izvor: vlastiti rad, 2024.)	24
13.	Predviđanje tjedne cijene bitcoina za sljedećih 12 tjedana(Izvor: Vlastiti rad, 2024.)	25

1. Prilog 1

U ovom prilogu se nalazi naziv glavne .tex datoteke korištene za izradu završnog rada:

- `zavrsni_rad_uvod_u_analizu_vremenskih_nizova.tex`.

2. Prilog 2

U ovom prilogu se nalaze nazivi RMarkdown i HTML datoteka korištenih za zradu završnog rada:

- `zavrsni_rad_uvod_u_analizu_vremenskih_nizova.Rmd`
- `zavrsni_rad_uvod_u_analizu_vremenskih_nizova.html`.