



UNIVERZITET U NOVOM SADU
PRIRODNO - MATEMATIČKI FAKULTET
DEPARTMAN ZA MATEMATIKU I INFORMATIKU



Silvija Sokol

Matematički modeli za analizu tržišta nekretnina

Master rad

Mentor
prof. dr Zorana Lužanin

Novi Sad, 2023. godine

Sadržaj

Zahvalnica	1
Uvod	2
1 Tržište nekretnina kao ekonomska pojava	3
1.1 Karakteristike tržišta nekretnina i faktori koji ga oblikuju	4
1.2 Boom and Bust ciklus	5
1.3 Segmentacija tržišta nekretnina	6
1.4 Tržište nekretnina u Republici Srbiji	7
1.5 Matematički modeli za modeliranje cene nekretnina	8
1.6 Koraci prilikom formiranja modela	9
2 Autoregresivni modeli integrisanih pokretnih preseka: ARIMA	10
2.1 Osnovni pojmovi	10
2.2 Autoregresivni model: AR	12
2.3 Model pokretnih preseka: MA	13
2.4 Autoregresivni model pokretnih preseka: ARMA	13
2.5 Autoregresivni modeli integrisanih pokretnih preseka: ARIMA	14
2.6 Sezonski autoregresivni modeli integrisanih pokretnih preseka: SARIMA	15
2.7 Određivanje reda modela: p, d, q, s, P, D, Q	15
2.8 Ocena parametara	17
2.9 Izbor modela	18
2.10 Provera adekvatnosti modela	18
2.11 Kvalitet predikcija	20
3 Hedonistička regresija	22
3.1 Razvoj hedonističkog modela	22
3.2 Hedonistička regresija na tržištu nekretnina	24
3.3 Hedonistički model vrednovanja cena nekretnina	26
3.4 Ocena nepoznatih parametara	28
3.5 Analiza reziduala	29
3.5.1 Normalnost	29
3.5.2 Homoskedastičnost	29
3.5.3 Autokorelacija	30
3.5.4 Multikolinearnost	30
3.5.5 Uticajne tačke i outlier-i	31
3.6 Evaluacija modela	31
3.7 Formiranje hedonističkog indeksa cena nekretnina	32
4 Rezultati empirijskog istraživanja	33
4.1 Rezultati primene ARIMA modela na realnim podacima	33
4.2 Rezultati primene hedonističkog modela na realnim podacima	42
Zaključak	52
Literatura	53
Biografija	55

Zahvalnica

Zahvaljujem se dr Zorani Lužanin na saradnji, strpljenju i savetima tokom izrade ovog rada, kao i članovima komisije dr Sanji Rapajić i dr Andreji Tepavčević na komentarima koji su pomogli poboljšanju ovog rada.

Takođe sam neizmerno zahvalna svojoj porodici na beskrajnoj podršci koju su mi pružili tokom studiranja.

*Silvija Sokol,
Novi Sad, 2023. godine*

Uvod

Tržište nekretnina predstavlja vitalni segment globalne ekonomije, njegov značaj seže u različite sfere i time oslikava makroekonomsku stabilnost i društveno blagostanje. U ovom kontekstu, razumevanje dinamike tržišta nekretnina zauzima važno mesto kao predmet analize među akademskom i stručnom javnosti.

U naporima da se predvidi i analizira ponašanje ovog kompleksnog tržišta, razvijene su različite metodologije, od kojih se posebno ističu matematički modeli. Ovaj rad je posvećen istraživanju efikasnosti dva takva modela: ARIMA (AutoRegressive Integrated Moving Average) i hedonističkog modela. ARIMA model, kao predstavnik modela vremenskih serija, koristi se za analizu i predikciju trendova na tržištu nekretnina. Sa druge strane, hedonistički model pruža detaljno sagledavanje elemenata vrednosti nekretnina, uključujući karakteristike samih objekata, lokalnu infrastrukturu i druge faktore koji utiču na tržišnu vrednost.

Predmet istraživanja ovog rada je analiza tržišta nekretnina kroz primenu ARIMA i hedonističkog modela. Kroz detaljno istraživanje, rad će se fokusirati na evaluaciju efikasnosti ovih modela u predviđanju i analiziranju ponašanja tržišta nekretnina. Istaći će se osnovne karakteristike ARIMA modela u analizi trendova tržišta i mogućnosti hedonističkog modela u oceni faktora koji utiču na vrednost nekretnina. Poseban akcenat će biti stavljen na primenu ovih modela u R Studio okruženju, gde će se realni podaci koristiti za empirijsku evaluaciju performansi modela.

Cilj rada je evaluacija i poređenje efikasnosti ARIMA i hedonističkog modela u analizi i predviđanju dinamike tržišta nekretnina. Kroz empirijsko istraživanje, rad će pokušati da odgovori na pitanja o preciznosti, pouzdanosti i praktičnosti ovih modela u realnim tržišnim uslovima. Osim toga, cilj je i demonstracija primene i adaptacije ovih modela u R Studio okruženju, istraživanje njihove fleksibilnosti i mogućnosti za doprinos boljem razumevanju i analizi tržišta nekretnina.

Ovaj rad je obrađen kroz četiri strukturalne celine.

Prva celina je posvećena pojmovnoj analizi tržišta nekretnina kao ekonomske pojave. Ovde se razmatraju osnovne karakteristike, faktori i dinamika koji oblikuju tržište nekretnina. Pored globalnog osvrta, poseban akcenat je stavljen na analizu tržišta nekretnina u Republici Srbiji, kako bi se stekao uvid u specifičnosti, izazove i prilike koje karakterišu domaće tržište.

Drugi deo se fokusira se na teoriju ARIMA modela. U ovoj sekciji, su sagledani osnovni principi, metodologija i aplikacija ovog popularnog modela vremenskih serija. Detaljno su objašnjeni matematički koncepti, parametri i procedura modelovanja

Treća celina je usmerena na teorijsko sagledavanje hedonističkog modela za procenu cena nekretnina. Ovde se razmatraju osnovni postulati, varijable i faktori koji se uključuju u model, kao i metodologija za empirijsku primenu.

U **četvrtoj celini** predstavljen srž rada i obuhvata rezultate empirijskog istraživanja. U ovom segmentu, prikazani su koraci u izgradnji i analizi ARIMA i hedonističkog modela kroz R Studio program. Detaljno su analizirani rezultati, interpretirane su implikacije modela i njihova efikasnost u predviđanju i analiziranju tržišta nekretnina.

1 Tržište nekretnina kao ekonomska pojava

Opstanak ljudske vrste je oduvek bio uslovljen sigurnim utočištem. Kroz istoriju, osnovna uloga utočišta je ostala nepromenjena, a važnost se ogleda u ljudskoj potrebi za osećajem sigurnosti. Fizički oblik utočišta se menjao sa razvojem civilizacije. Počevši od praistorije, kada su skrovišta predstavljala primitivne strukture građene od osnovnih prirodnih materijala, sve do modernog doba koje uključuje korišćenje naprednih i sofisticiranih građevinskih materijala i tehnologija. U davnim vremenima su skrovišta bila ništa više od običnih pećina ili sličnih prirodnih formacija, kao i jednostavnih građevina od drveta, kože i drugih prirodnih materijala. Kasnije, sa razvojem ljudske civilizacije, ona poprimaju kompleksniji oblik. Za izgradnju se koriste trajniji i efikasniji materijali, a nove građevine su smišljene da zadovolje više ljudskih prohteva u pogledu stanovanja. Osim tehničkog napretka, razvoj utočišta utiče i na društvene i kulturne promene. Dizajn i funkcionalnost domova variraju u različitim kulturama i vremenima, od tradicionalnih nasleđa do savremenih arhitektonskih trendova. Može se primetiti da razvoj utočišta predstavlja dinamičan proces koji oslikava kontinuirane napore ljudi da stvore okruženje koje odgovara njihovim potrebama i vrednostima. Taj razvoj se ne odnosi samo na promene u građevinskim tehnikama, već definiše i duboko ukorenjene težnje za boljim načinom života i stvaranjem domova koji pružaju sigurnost, komfor i identitet.

Sredinom 20. veka Abraham Maslov, istaknuti američki psiholog i mislilac, je ustanovio osnovne ljudske potrebe koje se međusobno nadopunjuju, tako što ih je prikazao hijerarhijski pomoću piramide[1]. Na samom vrhu te piramide, simbolizujući najviši nivo ljudskog dostignuća, nalaze se potrebe za samorealizacijom i duhovnim razvojem. Da bi se zadovoljile potrebe sa viših nivoa piramide, neophodno je prethodno zadovoljiti potrebe smeštene u njenoj osnovi. Zanimljivo je da je utočište svrstano neposredno iza osnovnih fizioloških potreba čime se dokazuje značaj istog u ljudskom životu i razvoju. Utočište se ne svodi samo na fizičko sklonište od nepovoljnih vremenskih uslova ili potencijalnih pretnji, već ima dublje značenje. Osećaj sigurnosti koji pruža utočište omogućuje ljudima da se fokusiraju na druge aspekte života i da razvijaju međuljudske odnose. Ovaj nivo sigurnosti i zaštite je osnova na kojoj se gradi potencijal za više nivoe razvoja i samorealizacije. Maslovljeva hijerarhija nas time podseća da su osnovne potrebe međusobno povezane i da svaka igra bitnu ulogu u oblikovanju ljudskog razvoja. Potreba za utočištem nije samo praktična, ona ima duboko ukorenjen uticaj na našu sposobnost da ostvarimo svoj puni potencijal i postignemo više nivoe zadovoljstva i ispunjenja.

Pojam utočišta je analiziran zbog njegove neosporive veze sa konceptom tržišta nekretnina. Ljudska potreba za utočištem, društvene promene, kao i ekonomski faktori su konstitutivni elementi koji uslovljavaju razvoj i funkcionisanje tržišta nekretnina. Upravo ti elementi dokazuju kompleksnu prirodu tržišta nekretnina. Oblik tržišta nekretnina, kakav danas poznajemo vuče korene iz 20. veka. Prva organizacija čiji je zadatak bio da poveže posrednike sa kupcima i prodavcima u SAD-u je osnovana 1908 godine (National Association of Real Estate Exchanges).

Nekretnine, nepokretnosti ili nepokretna imovina u pravnom smislu označavaju deo materijalne imovine fizičkog ili pravnog lica koje se ne mogu pokretati sa svog mesta bez oštećenja ili uništenja. Nepokretnostima se smatraju: zemljište (poljoprivredno, građevinsko, šume i šumsko zemljište), objekti (komercijalne, stambene, stambeno-poslovne, ekonomske i dr.) i drugi građevinski objekti, kao i posebni delovi zgrada (stanovi, poslovni prostori, garaže i garažna mesta) na kojima može postojati zasebno pravo svojine[2]. Stambeni objekti, kao oblik nepokretnosti, će biti fokus ovog rada.

Tržište nekretnina je ekonomska platforma na kojoj se kupuju, prodaju ili iznajmljuju različite vrste nepokretnosti, uključujući kuće, stanove, zemljišta i komercijalne objekte. Drugim rečima to je prostor gde se susreću ponuda i tražnja određujući cene i uslove transakcija. Ova ekonomska interakcija obuhvata fizičke nekretnine, kao i prava na njihovu upotrebu, posedovanje i razmenu. Ono uključuje složen odnos tržišnih činilaca. Tržište nekretnina je dinamično i podložno uticajima ekonomske situacije, urbanog razvoja, demografskih promena i drugih faktora. Glavni akteri na tržištu uključuju vlasnike nekretnina, investitore, agente za nekretnine, banke i državne agencije.

1.1 Karakteristike tržišta nekretnina i faktori koji ga oblikuju

Da bi se tržište nekretnina bolje razumelo kao ekonomska pojava, bitno je sagledati njegove glavne karakteristike koje ga čine drugačijim u odnosu na druga tržišta. Očigledne karakteristike ovog tržišta su niska likvidnost i visoki kapitalni zahtevi. Prva karakteristika je povezana sa činjenicom da se nekretnine relativno sporo prodaju, što dovodi do drugačije dinamike cena u poređenju sa drugim dobrima. Druga pomenuta karakteristika, visoki kapitalni zahtevi, znači da nekretnine obično iziskuju značajna finansijska sredstva, uključujući stambene kredite, poreze i redovne troškove održavanja. Pored niske likvidnosti i visokih kapitalnih zahteva, treba napomenuti i dug vremenski period kao karakteristiku tržišta nekretnina. Kupovina nekretnine se obično planira na duže staze, što znači da je potrebno analizirati dugoročne trendove.

Posedovanje nekretnine često predstavlja značajan izvor bogatstva u mnogim delovima sveta, a izgradnja istih uslovlja razvoj ekonomske aktivnosti i smanjenje nezaposlenosti. U globalnom kontekstu, stambeni sektor predstavlja značajan deo sveukupne ekonomske aktivnosti, a promene na tržištu nekretnina mogu izazvati šire implikacije na ekonomiju. Drugim rečima, kako na tržište nekretnina utiče mnogo faktora, ono predstavlja ključnu ulogu u globalnoj ekonomiji. Neki od bitnih faktora koji oblikuju tržište nekretnina predstavljaju sledeći[3]:

1. *Ponuda i potražnja*: odnos između dostupnih nekretnina za prodaju i interesovanja kupaca ima direktan uticaj na formiranje cena i tržišne dinamike.
2. *Ekonomska situacija*: opšte makroekonomske promene poput inflacije, stope nezaposlenosti, BDP-a i sličnih, koje imaju direktan uticaj na sposobnost ljudi da kupuju i investiraju u nekretnine.
3. *Demografske promene*: prirodni priraštaj i migracije posredno utiču na potražnju za određenim vrstama nekretnina i lokacijama.
4. *Stanje tržišta rada*: Lokalna ekonomska situacija, industrijska aktivnost i zapošljavanje uslovljavaju potrebu za stambenim i komercijalnim prostorom u određenim oblastima.
5. *Kamatne stope*: nivo istih na kredite može direktno uticati na sposobnost i odluku kupaca da finansiraju kupovinu nekretnina.
6. *Urbanizacija i planiranje*: razvoj infrastrukture kao i urbanistički planovi i regulative značajno doprinose atraktivnosti određenih lokacija sa aspekta tržišta nekretnina.
7. *Investicioni trendovi*: kupovina nekretnine kao oblik ulaganja utiče na cene i dinamiku na tržištu.
8. *Društvene promene*: promene u načinu života, porodičnim strukturama kao kulturološke preference mogu oblikovati tržište nekretnina i uticati na odluke o kupovini odnosno prodaji.
9. *Stanje građevinske industrije*: cene i dostupnost građevinskog materijala, kao i kapacitet građevinskih kompanija utiču na gradnju novih nekretnina, što direktno oblikuje ponudu na tržištu.

Pored pomenutih, postoji još mnoštvo faktora koji u manjoj ili većoj meri utiču na tržište nekretnina. Promene u ekonomskim uslovima, demografskim trendovima, urbanom razvoju, kamatnim stopama i pravnim regulativama direktno utiču na dinamiku tržišta. Cena nekretnina često igra ključnu ulogu u ovom procesu, odražavajući ravnotežu između ponude i potražnje. Može se reći da su ponuda i potražnja za različitim tipovima nekretnina ključni elementi tržišta nekretnina. Upravo ti elementi oblikuju cene, promet i investicije. Tržište nekretnina ima važnu ulogu u ekonomiji jer povezuje potrebe stanovništva za stambenim i komercijalnim prostorom, kao i resursima koji su dostupni za izgradnju i razvoj.

Ekonomska stabilnost države je pod direktnim uticajem promena koje se dešavaju na tržištu nekretnina. Svaka promena doprinosi visini bruto domaćeg proizvoda (BDP-a). Shodno tome možemo zaključiti da tržište nekretnina oslikava stanje države. Precizno praćenje trendova na tržištu nekretnina je od ključne važnosti za razumevanje stanja tržišta.

1.2 Boom and Bust ciklus

Posmatrajući dinamiku cena i aktivnosti na tržištu nekretnina uočava se ciklično kretanje tokom vremena. Dužina trajanja ciklusa zavisi od opštih ekonomskih okolnosti. Ciklusi se sastoje iz četiri faze:

1. Faza rasta/ ekspanzije: Fazu rasta karakteriše rast cena nekretnina, visoka tražnja i masovno ulaganje u nekretnine. Ova faza je uzrokovana raznim faktorima, kao što su makroekonomska situacija, urbanizacija ili snažan ekonomski rast, povoljni uslovi za kupovinu i niske kamatne stope.
2. Faza vrhunca: Trenutak kada cene nekretnina dostižu svoj najviši nivo, dok potražnja počinje da se smanjuje. Visoke cene utiču na povišenu spekulativnu aktivnost na tržištu jer se investitori nadaju profitu usled visokih cena.
3. Faza pucanja/ kraha: Nakon prekomernog rasta cena dolazi do drastičnog pada ili pucanja mehura. Uslovi za nastanak ove faze su prevelika ponuda na tržištu, primetan rast kamatnih stopa što dovodi do smanjenja tražnje. Inflacija ima indirektan uticaj, ali nikako zanemarljiv.
4. Faza oporavka: U poslednjoj fazi ciklusa cene nekretnina započinju blagi rast i tržište polako počinje da se stabilizuje što stvara nove investicione prilike po pristupačnijim cenama.

Mehur postoji ako cena imovine naglo raste i zatim naglo pada. Pod naglim rastom se smatra desetostruki rast u periodu do godinu dana. Specifičnost kod tržišta nekretnina je da se naglim rastom cena smatra najmanje dva puta veća cena u periodu od pet godina ili rast od 50% u periodu od tri godine. Određivanje vremenskog perioda u kojem se ispituje prisustvo mehura, kao i postavljanje optimalne granice pada cene nije jednostavan zadatak. Pojava mehura nije uzrokovana samo jednim faktorom. Za svaki mehur postoji skup uzročnih faktora koji se mogu međusobno preklapati, ali takođe zasebni događaji mogu uticati na njihov nastanak[4].

Mogu se uočiti tri šablona po kojem se mehuri na tržištu nekretnina javljaju:

- Čist spekulativni mehur: kupci kupuju danas sa namerom da prodaju po većoj ceni.
- Mehur usled iracionalnog očekivanja: kupci kupuju danas kako ne bi kupovali po većoj ceni u budućnosti.
- Mehur usled iracionalnog ponašanja institucija: odobravanje rizičnih kredita.

Citat koji se može povezati i na zanimljiv način opisati zašto nastaje mehur na tržištu:

*“It ain’t what you don’t know that gets you into trouble.
It’s what you know for sure that just ain’t so.”*

Mark Twain

Iako naizgled kompleksna i zbunjujuća rečenica, njome se precizno ukazuje na srž problema.

Izreka opisuje sklonost ljudi da svoja subjektivna uverenja tretiraju kao objektivne činjenice, iako su samo lična mišljenja koja mogu biti pogrešna. Ovakva verovanja rezultiraju pogoršanjem situacije.

Ukoliko se sumnja na opasnost od pucanja mehura, postavlja se pitanje kako reagovati da bi negativni uticaj na ekonomiju bio blaži i koje preventivne mere mogu ublažiti negativne posledice. Treba razmotriti da li bi preventivno reagovanje koštalo više u odnosu na troškove saniranja posledica pucanja mehura. Veliku ulogu u ovoj priči ima način finansiranja kupovina. Transakcije na tržištu nekretnina najčešće uključuju pozajmljivanje (time dolazi do visokog stepena zaduženosti) u odnosu na druga tržišta. Samim tim je makroekonomska stabilnost ugroženiya. U toku faze rasta mehura, kriterijumi za sticanje kredita su ublaživani, a dužnici su se počeli oslanjati na porast cena nekretnina kako bi mogli da izmire svoja dugovanja. Jačina pucanja, odnosno težina sanacije, zavisi od visine leveridža. Uvođenjem oštrijih monetarnih mera (pooštravanje

uslova za dobijanje kredita i uvećanje kamatnih stopa) smanjuje se osetljivost finansijskog sektora. Iz iskustava se zaključuje da saniranje posledica pucanja mehura iziskuje znatno više troškove i da je efektivnost politike ograničena te se favorizuje preduzimanje preventivnih mera. Glavni ciljevi mera su prevencija mehura i povezanog povećanja leveridža, zatim ublažavanje pucanja mehura povećavanjem elastičnosti finansijskog sektora[5]. Matematičko modeliranje cena ima ključnu ulogu u ovom kontekstu, jer pruža poboljšano i dublje razumevanje mehura putem predikcija.

1.3 Segmentacija tržišta nekretnina

Ključna strategija za poboljšano razumevanje preferencija i potreba potrošača u kontekstu nekretnina jeste primena segmentacije tržišta. Ovaj pristup podrazumeva klasifikaciju potrošača u grupe na osnovu sličnih potreba i želja. Analogno drugim kompleksnim tržištima, tržište nekretnina se može razložiti na manje grupe ili segmente prema specifičnim kriterijumima. Segmentacija omogućava dublje razumevanje potreba potrošača i stvara osnovu za kreiranje izuzetne ponude s ciljem postizanja konkurentske prednosti na tržištu[6].

Nekretnine se mogu razvrstati u grupe prema:

1. *Vrsti*: stambene, komercijalne, stambeno-poslovne
2. *Lokaciji*: gradski naspram ruralnih područja, delovi grada, ulice
3. *Nameni*: stanovanje, izdavanje, vršenje delatnosti
4. *Vrednosti*: luksuzne, srednje klasne, pristupačno stanovanje

Svaka od navedenih grupa može se dodatno segmentirati u skladu s preferencijama potrošača. Pored ovih kategorija, postoji i mogućnost grupisanja nekretnina na druge načine.

U analizi podataka, posmatranje podataka kao jedne celine može otežati prepoznavanje veza, korelacija i trendova. Kada se suočimo s izazovima analize obimnih podataka, segmentacija postaje koristan alat koji olakšava identifikaciju osnovnih uzroka i razvoj ciljanih rešenja. Segmentacija u analizi tržišta predstavlja efikasnu tehniku koja pomaže dekompoziciju kompleksnih setova podataka na manje, lakše prepoznatljive podskupove, u kojima je moguće identifikovati obrasce i trendove. Strateški pristup segmentaciji nije fiksni, već se kontinuirano usavršava, prilagođava i menja u skladu s novim podacima i tržišnim uslovima. Razdvajanje na segmente može značajno poboljšati preciznost prognoza i razvoj prediktivnih modela.

1.4 Tržište nekretnina u Republici Srbiji

Tržište nekretnina u Srbiji je doživelo ekspanziju prethodnih godina i beleži primetan rast broja prometovanih nekretnina i uvećanje obima novčanih sredstava. Značajano uvećenje vrednosti tržišta i broja realizovanih transakcija se dogodilo nakon pandemijske krize Covid-19. Na osnovu podataka koji se objavljuju na sajtu Republičkog Geodetskog zavoda, vrednost tržišta nekretnina u Srbiji beleži relativan rast od 46% tokom 2021. godine u odnosu na prethodnu godinu. Ovu godinu karakteriše i broj kupoprodaja koji je za 28% veći. Treba napomenuti da je tokom pandemije vreme provedeno u prirodi postalo atraktivno za mnoge, što je dalje uslovlilo povećanje tražnje za kućama i vikendicama. U 2022. godini broj kupoprodaja nije znatno skočio, prisutan je rast od samo 2% ali je praćeno drastičnim porastom vrednosti nepokretnosti koje je dovelo do uvećanja vrednosti tržišta za 22%[7]. Nesrazmerno kretanje broja transakcija nepokretnostima i njihove vrednosti oslikava globalne uticaje. Pandemijska kriza je imala veliki uticaj na svetsku ekonomiju. Rezultat toga su izuzetno veće cene građevinskog materijala, što je navelo investitore da podignu prodajne cene nepokretnosti. Druga dešavanja su izazvala migracije u svetu čime je i broj imigranata u Srbiji znatno porastao. Ove promene su izazvala daleko veću tražnju za stambenim prostorom, što je vlasnicima nekretnina dalo prostora za malverzacije cena.

Ruralni regioni beleže porast u broju prometovanih nekretnina i vrednosti, ali su za nijansu atraktivnija urbana mesta. Najveći akcenat, kada je u pitanju razvoj tržišta, je na glavnom gradu Beogradu, Novom Sadu, kao najvećem gradu autonomne pokrajne Vojvodine i Nišu koji je lociran na jugu Srbije. Ova tri grada najbolje oslikavaju promene koje su se desile na tržištu nekretnina.

Kao što je napomenuto, podaci o promenama na tržištu nekretnina na prostoru republike Srbije je dostupan na sajtu Republičkog geodetskog zavoda[7]. Zavod objavljuje polugodišnje i godišnje izveštaje koji pružaju detaljne informacije o aktuelnom stanju i promenama u posmatranom periodu na tržištu nepokretnosti. Ovim korakom je obezbeđena transparentnost tržišta. Transparentnost doprinosi stabilnosti tržišta i onemogućuje fluktuacije, drugim rečima potpomaže samoregulaciju tržišta. Transparentnost pruža dodatnu sigurnost učesnicima na tržištu jer omogućuje donošenje informisanih odluka. Sistem za prikupljanje podataka sa tržišta nepokretnosti na teritoriji Republike Srbije je pokrenut zbog masovne procene¹.

Uzurban razvoj tržišta nepokretnosti na teritoriji Republike Srbije se može pratiti na portalu gradnja.rs, gde su prikazana aktuelna gradilišta. Ovaj vid prikaza gradilišta je prilično zanimljiv za sagledavanje celokupne slike o aktivnostima građevinskog sektora. Gradilišta su prikazana na mapi sa tačnom lokacijom i mogu biti od velike koristi za pojedince koji razmišljaju o kupovini stambenog prostora ili ulaganju, investitore i analitičare.

Osvrt na arhitektonske i građevinske trendove na tržištu nekretnina u Srbiji je usaglašen sa svetskim tokovima. Arhitektonski stilovi i dizajn prate širok spektar promena. Počev od tehnološkog napretka do načina života pojedinaca. Nije strano da je život postao užurbaniji u protekloj deceniji, kao i da su se društvene vrednosti promenile. Trendovi i preferencije potrošača oblikuju izgled i funkcionalnost građevinskih objekata. Savremena arhitektura i građevinarstvo u stopu prati tehnološki napredak. Dizajn i funkcionalnost modernih građevina je integrisan pametnim sistemima za upravljanje osvetljenjem, temperaturom, sigurnošću i komunikacijom. Kao takve dobijaju naziv pametne kuće, zgrade, vile. Dominantan stil je zasnovan na minimalizmu, jednostavnim linijama i geometrijskim oblicima kojima se izražavaju individualnost i kreativnost sa ciljem izgradnje jedinstvene i atraktivne građevine. Savremene građevine se pažljivo projektuju kako bi ispratile zahteve društvenih promena kako bi se dizajnom postigao efekat multifunkcionalnosti. Pored navedenog, ekološka osvešćenost ostavlja trag i na građevinski sektor. Pažnja je usmerena na održivost i smanjenje negativnog uticaja na okolinu. Koriste se solarni paneli, reciklirani materijali, energetska efikasna izolacija.

¹ *Masovna procena*, po međunarodnim i evropskim standardima, definisana je kao proces određivanja vrednosti nepokretnosti na određeni datum, korišćenjem standardizovanih metoda nad svim podacima na tržištu nepokretnosti koje se posmatra, kao i statističko ispitivanje tih podataka u okviru posmatranog tržišta.

Druga polovina 20. veka u gradovima bivše Socijalističke Federativne Republike Jugoslavije je obeležena urbanističkim razvojem i velikim infrastrukturnim projektima. Tačnije, ovaj period se može posmatrati kroz nagli rast urbanih sredina i planirani razvoj gradskih četvrti. U urbanim sredinama dominira gradnja stambenih blokova koji su praćeni funkcionalnim dizajnom i bili su osmišljeni da zadovolje potrebe stanovništva. To je značilo da su gradske četvrti razvijene planski tako da sadrže škole, javne institucije i komercijalne objekte koji su omogućili olakšan tok života. Takav vid urbanog razvoja se delom preslikao i na razvoj savremenog tržišta nekretnina u Srbiji. Novogradnja obuhvata izgradnju pojedinačnih stambenih zgrada, smeštenih na dostupnim parcelama u već razvijenim delovima grada i kompleksa koji predstavljaju zasebne gradske celine.

Primer jednog takvog kompleksa je smešten na teritoriji glavnog grada, na obali reke Save i nazvan je Beograd na vodi (eng. Belgrade Waterfront). Ovaj kompleks je predstavljen kao primer luksuznog naselja. Sastoji se od 10.000 stanova smeštenih u 14 zgrada. Kula Beograd doseže visinu od 168m i sastoji se od 42 sprata. Panoramski pogled sa kule se prostire na tvrđavu Kalemegdan i reku Savu.

Ovakvi projekti igraju ključnu ulogu u transformaciji urbanih sredina i pružaju moderna rešenja za stambeni prostor, ali takođe često izazivaju i diskusije o urbanom planiranju, infrastrukturnim potrebama i ekološkim pitanjima.

Uprkos izazovima i kontroverzama, tržište nekretnina u Srbiji nastavlja da se razvija i prilagođava potrebama stanovništva, odražavajući promene u urbanističkom pejzažu i društvenim preferencijama.

1.5 Matematički modeli za modeliranje cene nekretnina

Matematički modeli za predikciju cena nekretnina zauzimaju bitno mesto na tržištu nekretnina. Oni omogućuju akterima na tržištu da donose racionalne i informisane odluke o ulaganjima. Tačno određivanje cena u budućnosti predstavlja poseban izazov što je uslovalo razvoj mnogih modela za predikciju. Modeli koriste različite matematičke i statističke tehnike kako bi predvideli cene nekretnina na osnovu niza relevantnih faktora. Pritom treba naglasiti da izbor modela zavisi od dostupnosti podataka, specifičnosti tržišta nekretnina i željenog nivoa preciznosti. Priroda dostupnih podataka ima važnu ulogu u izboru adekvatnog modela za modeliranje veze između varijabli. Podaci mogu biti kvantitativni ili numerički i kvalitativni ili deskriptivni. U cilju postizanja relevantnih rezultata često je kombinovanje različitih modela u okviru jedne studije. Tačnost ovih modela zavisi od kvaliteta i količine dostupnih podataka za analizu.

U ovom radu su predstavljena dva modela. ARIMA model, kao model vremenskih serija je korišćen kao alat za formiranje prosečne cene stana novogradnje po kvadratnom metru. Vremenska serija cena prikazuje kako se cena kreće tokom vremena, ne uzimajući u obzir ostale faktore i ekonomske uticaje koji utiču na njeno oblikovanje. Samim tim ovako definisana cena pokazuje kako se ona kreće na makroekonomskom nivou. Pored toga, posmatran je broj izdatih građevinskih dozvola za izgradnju stanova. Ova serija nam pokazuje dinamiku razvoja tržišta, čime dobijamo sliku o njegovom rastu. Sa druge strana, kada je reč o cenama na mikroekonomskom nivou, moramo uzeti u obzir karakteristike nepokretnosti koje na nju utiču. Za tu svrhu analiziran je hedonistički model za procenu cene nepokretnosti.

1.6 Koraci prilikom formiranja modela

Proces formiranja modela je šematski prikazan na slici 1:



Slika 1: Koraci u formiranju modela

Prikupljanje podataka. Prikupljanje relevantnih podataka o nekretninama, kao što su tip stambene jedinice, površina, spratnost, broj soba, broj kupatila, posjedovanje terase, garažnog ili parkig mesta, garaže, bazena, lokaciju (ulica i mesto), tip naselja (da li je ruralno ili urbano mesto), udaljenost od javnih institucija, škola, centra grada ili najbližeg grada ukoliko je ruralna sredina u pitanju, blizina javnog prevoza, ekološki kvaliteti okruženja, blizina parka i mnoge druge karakteristike.

Manipulacija podacima. Prikupljeni podaci se nalaze u sirovom obliku i potrebno je izvršiti njihovu obradu kako bi se dostigao kvalitetan skup za formiranje modela. Ovaj korak podrazumeva proveru koja uključuje izbacivanje neželjenih podataka, formatiranje, menjanje tipova podataka, utvrđivanje nedostajućih podataka.

Istraživačka analiza. Vizualizacija podataka pomoću grafičkih prikaza kako bi se utvrdile korelacija između podataka. Analizirani podaci predstavljaju promenljive modela. Numerički podaci se grafički prikazuju pomoću matričnih grafičkih prikaza (eng. scatter plot matrices). Veza između numeričkih i kategorijalnih podataka se prikazuje pomoću trakastog grafikona (eng. bar plot) ili pomoću box plot grafikona.

Priprema podataka. Kategorijalne promenljive se transformišu u dummy varijable sa ciljem unapređenja prediktivne moći modela.

Podela podataka. Potrebno je izvršiti podelu podataka na dva skupa, trening skup (eng. training set) i testing skup (eng. test set). Trening skup se koristi za testiranje varijabli i formiranje modela, dok se testing skup koristi za testiranje modela. Podaci se najčešće dele u razmeri 70:30, 75:25, 80:20.

Konstrukcija modela. Izbor odgovarajućeg skupa varijabli (atributa ili prediktora) je ključan korak u procesu modeliranja. Prilikom konstrukcije modela svrha je uključiti sve varijable koje su statistički značajne i u sebi nose neku informaciju, dok istovremeno treba izbeći varijable koje unose šum ili nisu relevantne i ne doprinose prediktivnoj moći. Postoji nekoliko načina selekcije modela, izbor zavisi od prirode podataka, karakteristika posmatrane pojave i cilja analize.

Validacija modela. Korišćenjem različitih metrika se utvrđuje preciznost modela tako što se analizira distribucija reziduala. Procena se izvodi na testing skupu i određuje se koliko dobro model predviđa njihove vrednosti. Metrike koje se najčešće koriste za utvrđivanje preciznosti modela su: srednja kvadratna greška (eng. mean squared error), koren srednje kvadratne greške (eng. root mean square error), R^2 (eng. R squared), prilagođeni R^2 (eng. adjusted R squared), F statistika. Upoređivanje modela se može vršiti pomoću kriterijuma AIC (eng. Akaike's Information Criterion) i BIC (eng. Bayesian's Information Criterion). Model se smatra preciznim i adekvatnim koliko je središte raspodele reziduala skoncentrisano oko 0, ukoliko su vrednosti R^2 i Adj. R^2 podjednake.

Izračunavanje predikcija. Primena modela na novim podacima koji nisu bili deo trening i testing skupa.

2 Autoregresivni modeli integrisanih pokretnih proseka: ARIMA

Autoregresivni modeli integrisanih pokretnih proseka, kao modeli vremenskih serija, su matematički modeli za analizu i prognozu vremenskih serija podataka. Modeliranje pomoću ARIMA modela je efikasno za predikciju vrednosti tokom vremena uzimajući u obzir prisustvo sezonskih trendova i šablona. Ovaj model se odlikuje mogućnošću efikasnog modeliranja i analize složenih vremenskih serija, posebno onih sa sezonskim i nestacionarnim karakteristikama. ARIMA model je široko primenljiv u različitim disciplinama, uključujući ekonomiju, finansije, inženjering i ekologiju, što doprinosi sveobuhvatnom razumevanju dinamike promenljivih tokom vremena.

Def. Vremenska serija $\{r_t\}_{t=1,\dots,T}$ je hronološki uređeni niz vrednosti pri čemu t predstavlja vremenski indeks.

Drugim rečima, vremenska serija su podaci prikupljeni tokom diskretnih vremenskih intervala.

ARIMA² model je nastao kao kombinacija teorije vremenskih serija i statistike. Sačinjavaju ga tri dela: Autoregresivni model AR³, model pokretnih proseka MA⁴ i integracija I⁵. AR model se temelji na ideji da trenutna vrednost vremenske serije zavisi od njenih istorijskih vrednosti do određenog trenutka. Sledeća komponenta, MA model, je zasnovana na ideji da trenutna vrednost vremenske serije zavisi od prethodnih slučajanih uticaja, belih šumova. Kombinacijom pomenutih komponenti je razvijen ARMA model, prethodnik ARIMA-e. Dodavanjem treće komponente, integracije (I) nastao je ARIMA model i rešen je problem nestacionarnosti sa kojim se suočavao ARMA model.

U nastavku sledi osvrt na pojmove bitne za razvoj ARIMA modela, kao i detaljniji opis sastavnih komponenti modela. Način formiranja ARIMA modela i provera adekvatnosti, kao i testovi koji su razvijeni za testiranje određenih hipoteza.

2.1 Osnovni pojmovi

Formiranje ARIMA modela se zasniva na osnovnoj pretpostavci o stacionarnosti vremenske serije koja je predmet analize. Ova pretpostavka doprinosi preciznijim prognozama modela. Postoje dve vrste stacionarnosti: striktna i slaba stacionarnost.

Def. Striktne stacionarnost: Vremenska serija $\{r_t\}_{t=1,\dots,T}$ je striktno stacionarna ako $(r_{t_1}, \dots, r_{t_k})$ i $(r_{t_1+t}, \dots, r_{t_k+t})$ za $k \in \mathbb{N}$ imaju istu raspodelu $\forall t$, pri čemu $t_1, \dots, t_k \in \mathbb{N}$.

Def. Slaba stacionarnost: Vremenska serija $\{r_t\}_{t=1,\dots,T}$ je slabo stacionarna ako su $E(r_t)$ i $cov(r_t, r_{t-l})$ nezavisni od vremena t , $\forall l \in \mathbb{Z}$.

Def. Vremenska serija $\{a_t\}_{t=1,\dots,T}$ se naziva beli šum ukoliko je $\{a_t\}_{t=1,\dots,T}$ nezavisna i identično distribuirana (i.i.d. - Independent and identically distributed), odnosno da $\{a_t\} \sim \mathcal{N}(0, \sigma^2)$.

Definicija implicira sledeće:

1. $E(a_t) = 0$
2. $var(a_t) = \sigma^2$

²ARIMA (eng. Autoregressive Integrated Moving Average)

³AR (eng. Autoregressive)

⁴MA (eng. Moving Average)

⁵I (eng. Integrated)

$$3. E(a_t, a_{t-i}) = cov(a_t, a_{t-i}) = 0$$

Prethodno znači da je $a_{t=1, \dots, T}$ niz nekorelisanih slučajnih promenljivih sa istom raspodelom i konačnim očekivanjem i disperzijom.

Dickey-Fuller test se koristi za testiranje stacionarnosti vremenske serije. Ovaj test, posebno *Augmented Dickey-Fuller* (ADF) test, testira nultu hipotezu da postoji jedinični koren u vremenskoj seriji, što ukazuje na nestacionarnost. Alternativna hipoteza je da vremenska serija nema jedinični koren, tj. da je stacionarna.

Test statistika za Dickey-Fuller test se izračunava na osnovu procenjene vrednosti jediničnog korena u autoregresijskom modelu. Test statistika (ADF) je data formulom:

$$ADF = \frac{\hat{\gamma}}{SE(\hat{\gamma})},$$

gde je:

$\hat{\gamma}$ - procenjena vrednost koeficijenta autoregresijskog modela koji predstavlja jedinični koren,
 $SE(\hat{\gamma})$ - standardna greška procenjene vrednosti $\hat{\gamma}$.

Lag operator L je alat kojim se vremenske serije pojednostavljaju sa matematičkog gledišta. Operator se definiše na sledeći način:

$$Lr_t = r_{t-1}$$

Iz definicije sledi da je

$$L^2 r_t = L(Lr_t) = Lr_{t-1} = r_{t-2}$$

pa možemo zaključiti da važi sledeće:

$$\begin{aligned} L^P r_t &= r_{t-P} \\ (L^P)^Q r_t &= L^{PQ} r_t = r_{t-PQ} \\ (L^P)(L^Q) r_t &= L^P r_{t-Q} = L^{P+Q} r_t = r_{t-P-Q} \end{aligned}$$

Poseban slučaj lag operatora je **operator diferenciranja** koji se definiše na sledeći način za diferenciranje prvog reda:

$$(1 - L)r_t = r_t - r_{t-1} = \Delta r_t$$

Diferenciranje drugog reda:

$$\begin{aligned} \Delta^2 r_t &= \Delta(\Delta r_t) = \Delta(r_t - r_{t-1}) \\ &= (r_t - r_{t-1}) - (r_{t-1} - r_{t-2}) \\ &= r_t - 2r_{t-1} + r_{t-2} \\ &= r_t - 2Lr_t + L^2 r_t \\ &= (1 - L)^2 r_t \end{aligned}$$

Generalno važi:

$$\Delta^d r_t = (1 - L)^d r_t$$

2.2 Autoregresivni model: AR

Autoregresivni model reda p , $AR(p)$:

$$r_t = \phi_0 + \sum_{i=1}^p \phi_i r_{t-i} + a_t \quad p \in \mathbb{N}_0$$

Pretpostavićemo da je $\{a_t\}_{t \in \mathbb{N}}$ beli šum.

Uslovna raspodela:

$$E(r_t | r_{t-1}, \dots, r_1) = \phi_0 + \sum_{i=1}^p \phi_i r_{t-i}$$

$$D(r_t | r_{t-1}, \dots, r_1) = D(\phi_0 + \sum_{i=1}^p \phi_i r_{t-i} + a_t | r_{t-1}, \dots, r_1) = D(a_t) = \sigma_a^2$$

Bezuslovna raspodela:

$$E(r_t) = \phi_0 + \sum_{i=1}^p \phi_i E(r_{t-i})$$

Ako pretpostavimo slabu stacionarnost:

$$E(r_t) = \frac{\phi_0}{1 - \sum_{i=1}^p \phi_i}$$

to bi značilo da $\sum_{i=1}^p \phi_i \neq 1$.

Autovarijansna funkcija $AR(p)$ modela je:

$$\gamma_l = \begin{cases} \phi_1 \gamma_1 + \dots + \phi_p \gamma_p + \sigma_a^2, l=0 \\ \phi_1 \gamma_{l-1} + \dots + \phi_p \gamma_{l-p}, l > 0 \end{cases} \quad (1)$$

Autokorelaciona funkcija $AR(p)$ modela:

$$\rho_l = \frac{\gamma_l}{\gamma_0} = \begin{cases} 1, l=0 \\ \phi_1 \rho_{l-1} + \dots + \phi_p \rho_{l-p}, l > 0 \end{cases} \quad (2)$$

Primenom lag operatora, model dobija sledeći oblik:

$$r_t = \phi_0 + \sum_{i=1}^p \phi_i L^i r_t + a_t = \phi_0 (1 - \sum_{i=1}^p \phi_i L^i) + a_t \quad p \in \mathbb{N}_0$$

$$E(r_t) = \hat{r}_t = \phi_0 (1 - \sum_{i=1}^p \phi_i L^i)^{-1} = \frac{\phi_0}{\phi(L)}$$

Gde je:

$$\phi(L) = 1 - \sum_{i=1}^p \phi_i L^i$$

Ako su L_i karakteristični koreni jednačine $\phi(L) = 0$, onda to implicira da $\phi_i \neq 1$ kako bi vremenska serija bila stacionarna. Takođe može se reći da nema $\phi(L) = 0$ jedinični koren ili da rešenja jednačine leže van jediničnog kruga.

2.3 Model pokretnih preseka: MA

Ukoliko krenemo od beskonačnog autoregresivnog modela $AR(\infty)$:

$$r_t = \phi_0 + \sum_{i=1}^{\infty} \phi_i r_{t-i} + a_t \quad p \in \mathbb{N}_0$$

zamenom $\phi_i = -\theta_k^i$ $i \in \mathbb{N}$. Važi da je $\phi_i \neq 1$ što implicira da su $\theta_i \neq 1$. Sprovođenjem rekurzije u q koraka $k = 1, \dots, q$ dolazimo do modela pokretnih preseka reda q , $MA(q)$:

$$r_t = C_0 + a_t - \sum_{i=1}^q \theta_i a_{t-i} \quad q \in \mathbb{N}_0$$

Procesi AR i MA su invertibilni (ekvivalentni) procesi uz pretpostavku $\phi_i = -\theta_k^i$ $i \in \mathbb{N}$.

Autokovarijansna funkcija $MA(q)$ modela:

$$\gamma_l = \begin{cases} \sigma_a^2(\theta_0^2 + \theta_1^2 + \dots + \theta_q^2), l=0 \\ \sigma_a^2(\theta_l \theta_0 + \theta_{l+1} \theta_1 + \dots + \theta_q \theta_{q-l}), 0 < l \leq q \\ 0, l > q \end{cases} \quad (3)$$

Autokorelaciona funkcija $MA(q)$ modela:

$$\rho_l = \frac{\gamma_l}{\gamma_0} = \begin{cases} 1, l=0 \\ \frac{\theta_l \theta_0 + \theta_{l+1} \theta_1 + \dots + \theta_q \theta_{q-l}}{\theta_0^2 + \theta_1^2 + \dots + \theta_q^2}, 0 < l \leq q \\ 0, l > q \end{cases} \quad (4)$$

Primenom lag operatora, model dobija sledeći oblik:

$$r_t = C_0 - \sum_{i=0}^q \theta_i L^i a_t \quad q \in \mathbb{N}_0$$

Što se može zapisati na sledeći način:

$$r_t = C_0 - \theta(L)a_t$$

Ako krenemo od procesa MA i pretpostavke da su procesi AR i MA invertibilni (ekvivalentni) procesi, rešavanjem jednačine $\theta(L) = 0$, da bi prethdo važi, karakteristični koreni ove jednačine L^i moraju da budu van jediničnog kruga, što implicira da $\theta_i \neq 1$.

2.4 Autoregresivni model pokretnih preseka: ARMA

Autoregresivni model pokretnih preseka reda p, q , $ARMA(p, q)$:

$$r_t = \phi_0 + \sum_{i=1}^p \phi_i r_{t-i} + a_t - \sum_{i=1}^q \theta_i a_{t-i} \quad p, q \in \mathbb{N}_0$$

ARMA model predstavlja kombinaciju prikazanih modela AR i MA pri čemu važi:

ARMA(p,0)=AR(p) model

ARMA(0,q)=MA(q) model

Autokovarijansna funkcija ARMA(p,q) modela:

$$\gamma_l = \phi_1 \gamma_{l-1} + \dots + \phi_p \gamma_{l-p}, l > q$$

Autokorelaciona funkcija ARMA(p,q) modela:

$$\rho_l = \phi_1 \rho_{l-1} + \dots + \phi_p \rho_{l-p}, l > q$$

Primenom lag operatora dobijamo:

$$r_t = \phi_0 + \sum_{i=1}^p \phi_i L^i r_t + a_t - \sum_{i=0}^q \theta_i L^i a_t$$

što možemo zapisati na sledeći način:

$$\phi(L)r_t = \phi_0 + \theta(L)a_t$$

gde su:

$$\phi(L) = 1 - \phi_1 L^1 + \dots + \phi_p L^p$$

$$\theta(L) = 1 - \theta_1 L^1 + \dots + \theta_q L^q$$

Stacionarnost autoregresinog dela, AR procesa se postiže rešavanjem $\phi(L) = 0$, dok se invertibilnost procesa MA postiže rešavanjem $\theta(L) = 0$. Karakteristični koreni ovih jednačina moraju da budu van jediničnog kruga.

2.5 Autoregresivni modeli integriranih pokretnih proseka: ARIMA

Praktična primena autoregresivnih modela je implicirala potrebu za formiranjem modela koji će omogućiti analizu nestacionarnih vremenskih serija. Mnoge ekonomske varijable podležu trendovima tokom vremena što ukazuje na narušenu pretpostavku o stacionarnosti. Razvoj ARIMA modela je nastao kao unapređena verzija ARMA modela.

Vremeska serija r_t prati ARIMA (p,d,q) model ukoliko priraštaji $C_t=r_t - r_{t-1}$ prate ARIMA (p,d-1,q) model. Model se diferencira sve dok se ne dostigne situacija u kojoj priraštaji prate ARIMA (p,0,q) model, odnosno ARMA(p,q). Ukoliko vremensku seriju zapišemo preko lag operatora

$$r_t = (1 - L)^d r_t$$

dobijamo da ukoliko r_t prati ARIMA (p,d,q) onda $(1 - L)^d r_t$ prati ARMA(p,q) proces.

Model se može zapisati pomoću lag operatora:

$$\phi(L)(1 - L)^d r_t = \phi_0 + \theta(L)a_t$$

U zavisnosti od reda diferenciranja konstanta ϕ_0 se interpretira na različite načine. U slučaju stacionarne vremenske serije, kada je $d = 0$, konstanta predstavlja srednju vrednost vremenske serije. U slučaju nestacionarne vremenske serije, kada je $d > 0$, konstanta predstavlja deterministički trend.

Najpoznatiji test za testiranje stacionarnosti vremenske serije je Dickey-Fuller test (ADF test).

Hipoteze koje se testitaju su:

$$H_0(\phi = 1)$$

$$H_1(\phi < 1)$$

Nulta hipoteza kaže da ne postoji jedinični koren ili da je vremenska serija stacionarna. Na nivou značajnosti α , ukoliko je vrednost $p < \alpha$ prihvatamo nultu hipotezu[8].

2.6 Sezonski autoregresivni modeli integriranih pokretnih proseka: SARIMA

SARIMA model, što je skraćenica od Seasonal AutoRegressive Integrated Moving Average, je model vremenskih serija koji generalizuje jednostavniji ARIMA model dodavanjem sezonske komponente. SARIMA modeli su korisni za modelovanje podataka vremenskih serija koji pokazuju sezonske obrasce.

SARIMA model se predstavlja notacijom SARIMA(p,d,q)(P,D,Q) s , gde su:
 p, d, q redovi nesezonske autoregresivne, diferencijalne i komponente pokretnog proseka, respektivno.
 P, D, Q redovi sezonske autoregresivne, diferencijalne i komponente pokretnog proseka, respektivno.
 s je sezonski period.

SARIMA model se takođe može zapisati preko lag operatora:

$$(1 - \phi(L^s)\phi(L))(1 - L)^d(1 - L^s)^D r_t = (1 + \theta(L^s)\theta(L))t$$

Komponente se mogu razložiti na nesezonske i sezonske:

Nesezonske ARIMA komponente:

Autoregresivna komponentna, AR(p)

$$\phi(L) = 1 - \phi_1 L^1 - \dots - \phi_p L^p$$

Integrirana komponenta, I(d)

$$(1 - L)^d$$

Komponentna pokretnog proseka, MA(q)

$$\theta(L) = 1 - \theta_1 L^1 - \dots - \theta_q L^q$$

Sezonske komponente:

Sezonska autoregresivna komponentna, SAR(p)

$$\phi(L^s) = 1 - \phi_1 L^s - \dots - \phi_p L^{ps}$$

Sezonska integrirana komponenta, SI(d)

$$(1 - B^s)^d$$

Sezonska komponentna pokretnog proseka, SMA(q)

$$\theta(L^s) = 1 - \theta_1 L^s - \dots - \theta_q L^{qs}$$

2.7 Određivanje reda modela: p, d, q, s, P, D, Q

Određivanje reda modela je veoma značajno za formiranje modela i efikasne prediktivne sposobnosti. Parametar integracije modela, d predstavlja broj diferenciranja sa ciljem postizanja stacionarne vremenske serije.

Za procenu parametara modela se koriste:

Funkcija autokorelacija (eng. Autocorrelation function, ACF)

Ove funkcije ρ_l smo prethodno definisali za svaki model.

Funkcija parcijalne autokorelacije (eng. Partial autocorrelation function, PACF)

U statistici se često susrećemo s situacijom u kojoj korelacija između dve varijable može biti posledica njihove povezanosti s trećom varijablom unutar modela. U takvim situacijama, koristi se koncept parcijalne korelacije, koji omogućava merenje stepena korelacije između dve varijable, dok se kontroliše ili eliminiše uticaj ostalih varijabli u modelu. Konkretno, parcijalni koeficijent korelacije izračunava se tako što se izoluje i analizira veza između dve odabrane varijable, neutrališući efekte svih ostalih varijabli.

Neka je dat AR(p) proces:

$$r_t = \phi_0 + \sum_{i=1}^p \phi_{ii} r_{t-i} + a_t$$

Upotrebom Yule-Walkers sistema, proces će izgledati:

$$\begin{bmatrix} \rho_1 \\ \cdot \\ \cdot \\ \rho_k \end{bmatrix} = \begin{bmatrix} 1 & \cdot & \cdot & \cdot & \rho_{k-1} \\ \cdot & & & & \\ \cdot & & & & \\ \rho_{k-1} & \cdot & \cdot & \cdot & 1 \end{bmatrix} \begin{bmatrix} \phi_{k1} \\ \cdot \\ \cdot \\ \phi_{kk} \end{bmatrix}$$

Primenom Kramerovog pravila računamo parcijale izvode koji će izgledati:

$$\phi_{kk} = \frac{\begin{vmatrix} 1 & \cdot & \cdot & \cdot & \rho_1 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \rho_{k-1} & \cdot & \cdot & \cdot & \rho_k \end{vmatrix}}{\begin{vmatrix} 1 & \cdot & \cdot & \cdot & \rho_{k-1} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \rho_{k-1} & \cdot & \cdot & \cdot & 1 \end{vmatrix}}$$

Rešavanje ovog sistema jednačina je kompleksno i komplikovano. Kako bi olakšao i skratio ovaj postupak Durbin je predložio da se vrednosti ϕ_{kk} rekursivno izračunavaju na sledeći način:

$$\begin{aligned} \phi_{kk} &= \frac{\rho_k - \sum_{i=1}^p \phi_{k-1} \rho_{k-i}}{1 - \sum_{i=1}^k \phi_{k-1} \rho_{k-i}} \\ \phi_{kj} &= \phi_{k-1,j} - \phi_{kk} \phi_{k-1,k-j} \end{aligned}$$

pri tome je $0 < j \leq k - 1$.

Osobine parcijalne autokorelacione funkcije:

1. $PACF(p) = \hat{\phi}_{pp} \Rightarrow \phi_p$ pri čemu je $\phi_p \neq 0$
2. $PACF(j) = \hat{\phi}_{jj} \Rightarrow 0$ za $j > p$

Ove osobine parcijalne autokorelaciona funkcija AR(p) procesa ukazuju da će funkcija biti odsečena posle p lagova, odnosno da će procenjeni parametar biti statistički jedan nuli. Ukoliko je procenjeni parametar statistički jednak nuli nema razloga da se uključuje u model.

Autoregresivni modeli i modeli pokretnih proseka su međusobno povezani i mogu se izvesti jedan iz drugog. Ukoliko pođemo od beskonačnog autoregresivnog modela $AR(\infty)$ rekursivnim postupkom možemo da ga svedemo na konačan model pokretnih proseka $MA(q)$. Važi i obrnuto, ukoliko pođemo od beskonačnog modela pokretnih proseka $MA(\infty)$ rekursivnim postupkom možemo da ga svedemo na konačan autoregresivni model $AR(p)$. Na ovaj način je ispunjen uslov invertibilnosti, odnosno ekvivalenciji ovih procesa. Njihova povezanost se oslikava na ponašanje autokorelacione i parcijalne autokorelacione funkcije što se može prikazati tabelarno na sledeći način (tabela 1):

Model	ACF	PACF
AR(p)	eksponencijalno opada	nula kada je l veće od p
MA(q)	nula kada je l veće od q	eksponencijalno opada
ARMA(p,q)	eksponencijalno opada posle lag max(0,q-p)	eksponencijalno opada posle lag max(0,p-q)

Tabela 1: Ponašanje funkcija ACF i PACF

Parametar D predstavlja broj diferenciranja kako bi vremenska serija postala stacionarna na sezonskom nivou. Parametar s predstavlja sezonalnost i moguće ga je uočiti vizuelnim pregledom vremenske serije, kao i pregledom grafičkih prikaza funkcija ACF i PACF. Parametri P i Q predstavljaju sezonske "šiljke" na grafičkim prikazima PACF i ACF, respektivno.

2.8 Ocena parametara

Ocena nepoznatih parametara modela vremenskih serija se može vršiti primenom uslovne funkcije maksimalne verodostojnosti. Ideja ove funkcije je da na osnovu odabranog uzorka (realizovane vrednosti vremenske serije na kojima se formira model) odrede vrednosti nepoznatih parametara koje daju najveću verovatnoću da taj uzorak bude odabran. Neka je x vektor nepoznatih parametara modela:

$$x = (\theta_q, \dots, \theta_1, \phi_0, \sigma_a, \phi_p, \dots, \phi_1) \in \mathbb{R}^{p+q+2}$$

Pretpostavimo da je a_t Gausov beli šum i da $r_t | F_{t-1} : N(\mu_t, \sigma_t^2)$

$$\begin{aligned} \mu_t = E(r_t | F_{t-1}) = \hat{r}_{t-1}(1) &= \phi_0 + \sum_{i=1}^p \phi_i r_{t-i} - \sum_{i=1}^q \theta_i a_{t-i} \Rightarrow r_t - \mu_t = a_t \\ \sigma_t^2 = D(r_t | F_{t-1}) &= \sigma_a^2 \end{aligned}$$

Funkcija gustine ima oblik:

$$\mathcal{L}(x) = \left(\frac{1}{\sqrt{2\pi\sigma_a^2}} \right)^T \prod_{t=1}^T e^{-\frac{(r_t - \phi_0 + \theta_1 a_{t-1} + \dots + a_{t-q} + \phi_1 r_{t-1} + \dots + \phi_{t-p} r_{t-p})^2}{2\sigma_a^2}}$$

Kako je

$$a_t = r_t - \phi_0 + \theta_1 a_{t-1} + \dots + a_{t-q} + \phi_1 r_{t-1} + \dots + \phi_{t-p} r_{t-p}$$

Funkcija gustine je:

$$\mathcal{L}(x) = \left(\frac{1}{\sqrt{2\pi\sigma_a^2}} \right)^T \prod_{t=1}^T e^{-\frac{a_t^2}{2\sigma_a^2}}$$

Iz osobine logaritamske funkcije da je monotono rastuća sledi da obe funkcije L i $\ln L$ dostižu maksimum za iste vrednosti parametara i ova osobina nam dozvoljava da posmatramo logaritmovanu funkciju gustine:

$$\ln(\mathcal{L}(x)) = -\left(\frac{T}{2}\ln(2\pi) + \frac{T}{2}\ln\sigma_a^2 + \sum_{t=1}^T \frac{a_t^2}{2\sigma_a^2}\right)$$

$$f(x) = -\ln(\mathcal{L}(x)) = \frac{T}{2}\ln(2\pi) + \frac{T}{2}\ln\sigma_a^2 + \sum_{t=1}^T \frac{a_t^2}{2\sigma_a^2}$$

Rešavanje problema maksimizacije se svodi na rešavanje problema minimizacije:

$$\max_x \ln \mathcal{L}(x) = \min_x f(x)$$

Za pronalazak nepoznatih parametara dovoljno je da pronađemo stacionarne tačke funkcije tako što se parcijalni izvodi prvog reda izjednačavaju sa 0.

$$\frac{\partial \ln \mathcal{L}}{\partial x} = 0$$

2.9 Izbor modela

Prilikom selekcije parametara u cilju boljeg fitovanja modela postoji rizik od formiranja isuviše kompleksnih modela. Izbor optimalnog modela u analizi podataka ključan je jer omogućava preciznije prognoze i efikasno iskorišćavanje resursa. Neki od kriterijuma za upoređivanje modela su:

AIC kriterijum (eng. Akaike Information Criterion):

$$AIC = -2\ln(L) + 2k$$

BIC kriterijum (eng. Bayesian Information Criterion):

$$BIC = -2\ln(L) + 2k\ln(n)$$

U navedenim kriterijumima, L predstavlja vrednost funkcije maksimalne verodostojnosti, k predstavlja broj parametara modela, n je veličina serije. Prilikom poređenja ovih vrednosti bolji model je onaj koji ima manju vrednost navedenih kriterijuma. Na osnovu formula vidimo da će vrednost kriterijuma biti veća ukoliko povećamo broj parametara modela, na ovaj način kriterijumi kažnjavaju kompleksnost modela. AIC kriterijum kažnjava samo kompleksnost modela, dok BIC kriterijum pored kompleksnosti uzima u obzir i obim serije.

2.10 Provera adekvatnosti modela

Faza evaluacije adekvatnosti modela se zasniva na analizi reziduala modela. Očekuje se da se reziduali ponašaju kao beli šum, što podrazumeva da reziduali nisu korelisani i da prate normalnu raspodelu sa očekivanjem 0 i konstantnom varijansom.

Najpoznatiji testovi za ispitivanje serijske korelacije ili autokorelacije su *Box-Pierce* test statistika i *Ljung-Box* test statistika. Ovim testovima se testira nulta hipoteza koja pretpostavlja da su svi koeficijenti korelacije jednaki nuli protiv alternativne hipoteze koja kaže da postoji bar jedan koeficijent korelacije različit od nule:

$$H_0(\rho_1 = \dots = \rho_k = 0)$$

$$H_1(\exists i, i \in \{1, \dots, k\} : \rho_i \neq 0)$$

Box – Pierce test statistika:

$$Q = T \sum_{j=1}^k \rho_j^2$$

Ljung – Box test statistika:

$$Q' = T(T+2) \sum_{j=1}^k \frac{\rho_j^2}{T-j}$$

Pri čemu je

$$\rho_j = \frac{\sum_{t=j+1}^T e_t e_{t-j}}{\sum_{t=1}^T e_t^2}$$

Test statistike Q i Q' asimptotski prate $\chi^2(k)$.

Normalna raspodela reziduala ukazuje da model pravilno hvata strukturu podataka. Normalnost reziduala utiče na tačnost zaključaka o parametrima modela. Neke od test statistika koje se koriste za testiranje normalnosti reziduala su *Shapiro-Wilk* i *Anderson-Darling*. Ovim test statistikama se testira nulta hipoteza H_0 koja tvrdi da reziduali prate normalnu raspodelu protiv alternativne hipoteze H_1 koja tvrdi da reziduali

ne prate normalnu raspodelu.

Shapiro-Wilk test statistika se primenjuje najčešće na manje uzorke ($n \leq 50$).

Test statistika Shapiro-Wilk testa za proveru normalnosti je data formulom:

$$W = \frac{(\sum_{i=1}^n a_i x_{(i)})^2}{\sum_{i=1}^n (x_i - \bar{x})^2}$$

gde su $x_{(1)}, x_{(2)}, \dots, x_{(n)}$ redosled podataka sortiranih od najmanjeg do najvećeg, a_i su koeficijenti koji zavise od očekivanih vrednosti standardne normalne distribucije, a n je ukupan broj opservacija.

Anderson-Darling test statistika se primenjuje na veće uzorke ($n \geq 50$).

Test statistika Anderson-Darling testa je data formulom:

$$A^2 = -n - \frac{1}{n} \sum_{i=1}^n (2i-1) (\ln(Z_{(i)}) + \ln(1 - Z_{(n+1-i)}))$$

gde su $Z_{(i)}$ vrednosti kumulativne distribucijske funkcije standardne normalne distribucije evaluirane u tačkama podataka, a n je ukupan broj opservacija.

Kumulativna distributivna funkcija (CDF) za slučajnu varijablu X definisana je kao verovatnoća da X uzima vrednost manju ili jednaku x :

$$F(x) = P(X \leq x).$$

Za neprekidne distribucije, CDF se može izraziti integralom gustine verovatnoće (PDF):

$$F(x) = \int_{-\infty}^x f(t) dt,$$

gde je $f(t)$ PDF.

PDF standardne normalne distribucije je

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}.$$

CDF standardne normalne distribucije, označena sa $\Phi(x)$, je

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{1}{2}t^2} dt.$$

CDF u Anderson-Darling Testu

U Anderson-Darling testu, $F(x)$ je CDF teorijske distribucije sa kojom upoređujemo distribuciju našeg uzorka. Ako testiramo za normalnost, koristimo CDF standardne normalne distribucije.

Vrednosti $Z_{(i)}$ se računaju kao

$$Z_{(i)} = F(x_{(i)}),$$

gde je $x_{(i)}$ i-ti sortirani podatak iz uzorka. To daje verovatnoću da slučajna varijabla iz teorijske distribucije ima vrednost manju ili jednaku $x_{(i)}$.

2.11 Kvalitet predikcija

Standardne mere za proveru kvaliteta predikcija su statistički pokazatelji koji nam pružaju sažet pregled o tome koliko su tačne predikcije koje daje odabrani model. Osnovne metrike za utvrđivanje performansi modela su:

Srednja apsolutna greška (eng. Mean Absolute error)

$$MAE = \frac{1}{n} \sum_1^n |e_t|$$

Srednja apsolutna procentualna greška (eng. Mean Absolute Percentage error)

$$MAPE = \frac{1}{n} \sum_1^n \left| \frac{e_t}{r_t} \right|$$

Srednja kvadratna greška (eng. Mean Squared error)

$$MSE = \frac{1}{n} \sum_1^n e_t^2$$

Koren srednje kvadratne greške (eng. Root Mean Squared error)

$$RMSE = \sqrt{\frac{1}{n} \sum_1^n e_t^2}$$

Manja vrednost navedenih grešaka implicira bolju prediktivnu moć modela.

Pored navedenih mera primenjuje se i R^2 (eng. R squared) i prilagođeni R^2 (eng. Adjusted R squared). Ove statističke mere kvantifikuju u kojoj meri se varijansa u zavisnoj promenljivoj može objasniti varijansom nezavisnih promenljivih. Prilikom poređenja ovih vrednosti između različitih modela, veća vrednost ukazuje na bolji model.

Da bismo razumeli zašto dolazi do odstupanja modela, najpre ćemo definisati:

$$\sum_{i=1}^n (r_i - \bar{r}_i)^2 = \sum_{i=1}^n (\hat{r}_i - \bar{r}_i)^2 + \sum_{i=1}^n (r_i - \hat{r}_i)^2$$

Pri čemu su: $\{r_i\}$ realizovane vrednosti, \bar{r}_i je srednja vrednost realizovanih vrednosti, $\{\hat{r}_i\}$ su procenjene vrednosti. Jednačinu možemo zapisati:

$$SST = SSR + SSE$$

pri čemu je:

$$SST = \sum_{i=1}^n (r_i - \bar{r}_i)^2$$

$$SSR = \sum_{i=1}^n (\hat{r}_i - \bar{r}_i)^2$$

$$SSE = \sum_{i=1}^n (r_i - \hat{r}_i)^2$$

Sada možemo da definišemo:

$$R^2 = \frac{SSR}{SST}$$

$$\bar{R}^2 = 1 - \frac{n-1}{n-(k+1)} \frac{SSE}{SST}$$

\bar{R}^2 je statistička mera koja predstavlja udeo varijanse zavisne promenljive koji se može predvideti iz nezavisnih promenljivih u regresionom modelu. Ovaj koeficijent pruža meru koliko dobro posmatrani rezultati odgovaraju modelu. Kreće se u intervalu $[0, 1]$ i više vrednosti ukazuju na bolje prilagođavanje. Međutim, visoka vrednost \bar{R}^2 ne znači nužno da je model najbolji ili da ima validne prediktore. On samo govori koliko dobro model objašnjava varijabilnost zavisne promenljive. Dodavanjem više nezavisnih promenljivih vrednost \bar{R}^2 se može povećati čak i ako dodatne promenljive nemaju značajan uticaj na zavisnu promenljivu. Sa druge strane, prilagođeni \bar{R}^2 uzima u obzir broj nezavisnih promenljivih u modelu i kažnjava dodavanje irelevantnih promenljivih koje ne poboljšavaju model značajno. Prilagođeni \bar{R}^2 može da ima i negativnu vrednost[9].

Interpretacija R^2 :

$R^2 = 1$: Model savršeno objašnjava varijaciju u zavisnoj promenljivoj (mada, u praksi, savršen R^2 može ukazivati na overfitting).

$R^2 = 0$: Model ne objašnjava nimalo varijaciju u zavisnoj promenljivoj; model nije bolji od modela koji koristi samo srednju vrednost zavisne promenljive za predviđanje.

$0 < R^2 < 1$: Model delimično objašnjava varijaciju u zavisnoj promenljivoj; veće vrednosti R^2 ukazuju na bolje objašnjenje varijacije.

3 Hedonistička regresija

Hedonistička regresija se često primenjuje u ekonomiji za procenu vrednosti imovine. Svrha hedonističkog modela je konstrukcija cene imovine analizom njenih karakteristika. Pristup podrazumeva razlaganje imovine na njene karakteristike ili atribute za koje se smatra da utiču na njenu vrednost i da vrednost imovine jeste suma vrednosti njenih karakteristika[10]. Pored osnovnih atributa, na vrednost imovine mogu uticati i drugi faktori kao što su preferencije potrošača, ekonomski uslovi i dinamika tržišta. Postoji nekoliko osnovnih funkcionalnih oblika hedonističkog pristupa vrednovanja, kao što su linearni, polu-logaritamski i logaritamski oblici. Odabir odgovarajućeg funkcionalnog oblika se ističe kao jedan od ključnih empirijskih izazova u hedonističkom modeliranju cena.

3.1 Razvoj hedonističkog modela

Hedonizam predstavlja filozofski pravac koji svrhu života i ljudskog delovanja vidi u zadovoljstvu i uživanju. Sama reč hedonizam potiče od grčke reči *hedone* što znači užitek. Pored užitka, reč hedonizam opisuje širok spektar emocija i iskustava. Kao najbitnije emocije relevantne za temu ovog rada mogu se izdvojiti zadovoljstvo, radost i spokoj. Njihovim prisustvom isključeni su osećaji zabrinutosti, straha i očaja. Uvidevši da ljudi ponekad donose iracionalne odluke zasnovane na emocijama ekonomisti su se posvetili proučavanju ponašanja potrošača. Istraživanje je sprovedeno sa ciljem boljeg razumevanja šta motiviše i na koji način će se pojedinac ponašati prilikom donošenja ekonomske odluke. Želja za saznanjem kako emocionalni i psihološki aspekti utiču na donošenje ekonomske odluke navela je istraživače da dublje ispituju povezanost ekonomije sa psihologijom. Time je koncept hedonističkog modela nastao kao kombinacija ekonomskih teorija vrednovanja nekretnina i principa hedonizma.

Prve primene hedonističkog stila vrednovanja se nisu odnosile na nekretnine, iako je danas hedonistički pristup na ovom polju najviše zastupljen. Hedonističko modeliranje je počelo da se oblikuje početkom 1920ih godina. Najranije primene ovakvog pristupa datiraju iz 1922 i 1926 godine kada su Hass i Wallace primenili hedonistički metod za procenu vrednosti poljoprivrednog zemljišta. Hass je u svom istraživanju radio na izračunavanju cene zemljišta po jutru uzimajući u obzir godinu prodaje, veličinu grada i tip puta ali nije uveo naziv hedonistički[11]. Court je 1939 godine formirao hedonistički indeks za primenu u automobilskoj industriji i prvi je upotrebio naziv hedonistički model. Courtova analiza se smatra prvim formalnim doprinosom formiranja hedonističkog modela i mnogi naučnici ga smatraju ocem hedonističke regresije. Teorija je zasnovana na tvrdnji da se korisnost određenog dobra ogleda u njegovim karakteristikama. Na taj način korisnost dobra je predstavljena kao agregatna korisnost pojedinačnih korisnosti koje pružaju njegove karakteristike. Kupovina se zasniva na odabiru najpogodnijeg dobra iz skupa dobara koje dele zajedničke osobine[12]. Narednih godina mnogi naučnici su se posvetili izučavanju hedonističkog pristupa vrednovanja i ostavili svoj izuzetan doprinos razvoju ovog modela. Revolucionarna teorija hedonističke korisnosti je prvi pokušaj stvaranja teorijske osnove hedonističkog modeliranja od strane Lancastera 1966 godine koji je među karakteristike uključio one koje pružaju korisnost u donošenju odluka potrošača. Model implicira da dobra pripadaju istoj kategoriji i da se sva ili neka od tih dobara konzumiraju u različitim kombinacijama, uzimajući u obzir budžet potrošača. Lancasterova teorija takođe podrazumeva da postoji linearni odnos između cene proizvoda i njegovih svojstava ili karakteristika. Ideja je da se korisnost dobra zavisi od njegovih karakteristika i na osnovu njih se dobra mogu razvrstati u grupe. Korisnost potrošača ne zavisi od količine različitih dobara nego od korisnosti karakteristika koje određeno dobro pruža[13]. Pristup je fokusiran na stranu tražnje na tržištu i više je prilagođen potrošačkim dobrima.

Inspirisan istraživanjima autora Houthakker (1952), Becker (1965), Muth (1966), i Lancaster (1966), Rosen je 1974 godine[14] integrisao hedonistički pristup sa standardnim ekonomskim teorijama. Za razliku od ranijih radova o hedonističkom pristupu, koji su koristili linearne modele, Rosen ističe da veza između cena i atributa može biti nelinearna. Rosenov model simulira kompetitivno tržište i istovremeno predviđa funkcije ponude i tražnje. Model podrazumeva da se cene implicitnih karakteristika mogu izvesti na osnovu

tržišne ponude i tražnje. Ovaj pristup se sastoji od dva koraka. Prvo, implicira da se dobro može razložiti na sastavne karakteristike od kojih svaka od njih jednako doprinosi njegovoj vrednosti. Upotrebom definisanih karakteristika u hedonističkoj funkciji primenjenoj na ravnotežne cene koje oslikavaju ponašanje potrošača i proizvođača na konkurentnom tržištu dobijaju se implicitne marginalne vrednosti karakteristika. U drugom koraku se procenjuje ponašanje potrošača na osnovu implicitnih cena i potrošačevih socioekonomskih karakteristika. Socioekonomske karakteristike obuhvataju različite aspekte društva i ekonomije koji oslikavaju živote ljudi. U njih se ubrajaju faktori kao što su obrazovanje, zaposlenost, primanja, siromaštvo, rasna i rodna pripadnost. Visina prihoda, kao socio ekonomska karakteristika, direktno je uključena u budžetska ograničenja potrošača, čime se marginalna spremnost potrošača da plati određenu implicitnu karakteristiku može promeniti u zavisnosti od promene prihoda. Drugim rečima, Rosen tvrdi da cena dobra ne zavisi samo od tržišne ponude i tražnje za određenim dobrom, nego na cenu veliki uticaj imaju i njegove karakteristike. Bartik (1987)[15] nije bio saglasan sa Rosenovim pristupom pri proceni hedonističkog modela cena. On je tvrdio da razumevanje cena i atributa proizvoda u realnom svetu zahteva uzimanje u obzir međusobnih veza između cena i atributa, kao i nelinearnih efekata budžetskih ograničenja. Cena i količina predstavljaju endogene faktore koji su međusobno povezani i zavise jedan od drugog, čime se ne mogu posmatrati kao spoljni uticaji. Budžetsko ograničenje odnosi se na ograničenje potrošača u pogledu toga koliko novca može potrošiti na proizvode. Nelinearno budžetsko ograničenje implicira da velike promene u ceni proizvoda ili količini atributa mogu prouzrokovati neujednačene promene u potrošačkom izboru ili potrošnji. Drugim rečima, specifične okolnosti mogu izazvati nelinearne reakcije potrošača na promenu cene. Shodno tome, Bartik smatra da je suvišno modelirati ponudu na tržištu.

Freeman je 1979 godine unapredio metod vrednovanja primenom uticaja ekoloških faktora[16]. Ovo proširenje modela ima značajan uticaj na procenu cena nepokretnosti jer na njihovu vrednost značajno utiču ekološki kvalitet lokacije.

Sposobnost hedonističkih modela da analiziraju složen odnos između cena nekretnina i njenih karakteristika ga čini važnim činiocem na tržištu nekretnina. Pored aktera na tržištu koji ga koriste za ispitivanje ekonomskog uticaja nekretnina na tržište, model je pogodan za donosiocima odluka koji oblikuju urbanistički razvoj. S obzirom na kontinuirani napredak tehnologije i izvora podataka, možemo očekivati unapređenja i inovacije u oblasti hedonističkog modeliranja. Očekuje se da će dalji razvoj doprineti još većoj preciznosti i praktičnoj primeni ovih modela.

3.2 Hedonistička regresija na tržištu nekretnina

Hedonistički pristup ekonomskim analizama je omogućen hedonističkom regresijom. To je statistički i ekonometrijski pristup korišten u oblasti nekretnina radi procene vrednosti stambenih objekata na osnovu njihovih različitih karakteristika. Hedonistički pristup se zasniva na pretpostavci da se svaka heterogena imovina može posmatrati kao skup njenih kvantitativnih i kvalitativnih obeležja. Kada je reč o nekretninama obeležja mogu da budu strukturalna ili lokacijska. Procenom tržišne vrednosti svakog pojedinačnog obeležja, koje predstavljaju implicitne vrednosti, se formira vrednost imovine kao zbir implicitnih vrednosti. Obeležjima se ne može zasebno trgovati i njihov marginalni doprinos vrednosti nekretnine je implicitno određen ponudom i tražnjom za nekretninom.

Imovinu stanovništva u najvećem delu čini nepokretna imovina. Prilikom donošenja odluke o ulaganju novčanih sredstava, većina stanovništva se odlučuje na investiranje u nekretninu jer je nepokretna imovina široko prihvaćena kao najsigurniji vid ulaganja i najbolji pokazatelj kredibiliteta i solventnosti. Kupovine se najčešće realizuju zaduženjem kod banaka, samim tim promene kamatnih stopa znatno utiču na vrednost imovine stanovništva, kao i na njihove prihode i rashode. Otplata kredita i izmirenje zakupa nepokretnosti predstavljaju najveće izdatke domaćinstva. Samim tim promene na tržištu nekretnina imaju značajan uticaj na finansijsku stabilnost države. Praćenje promena na tržištu nekretnina je veoma značajno jer njihovo kretanje je pokazatelj prekomernog rasta ili pada cena nepokretnosti. Prekomeran rast cena može nekretnine učiniti nepristupačnim za većinu stanovništva, u tom slučaju pojavila bi se jasna granica između bogatih koju mogu da priušte kupovinu i siromašnih koji su prinudeni na iznajmljivanje ili u najgorem slučaju da ih dovede do beskućništva.

Ulaganje u stambeni prostor se smatra značajnim korakom u životu svakog pojedinca. Imajući to na umu možemo zaključiti da je donošenje takve odluke emotivan i zahtevan potez u ljudskom životu. Odluka se zasniva na mnogobrojnim kriterijumima koji se tiču celokupnog toka života pojedinca. Motivacija za kupovinu je najčešće osnivanje ili proširenje porodice, kao i mogućnost pasivnog prihoda koji se ostvaruje kroz izdavanje nekretnine. Još jedan od razloga za kupovinu nekretnine je očekivanje rasta cene i prilika za zaradom od prodaje iste. Postoji nekoliko faktora koji utiču na donošenje investicione odluke o kupovini nekretnine. Jedan od bitnih faktora je trenutni odnos visine kirije i rate za kredit. Ovaj odnos meri Price-to-Rent ratio koji pokazuje da li tržište favorizuje kupovinu ili iznajmljivanje. Razumno je pretpostaviti da, ukoliko su cene nekretnina visoke, pojedinac će se pre odlučiti za iznajmljivanje. Optimalnim rokom za otplatu kredita se smatra period od 15 godina. Drugi bitan faktor je odabir najpogodnije investicione prilike u zavisnosti od visine prihoda domaćinstva, sa obzirom da treba razmotriti druge životne troškove kao što su hrana, prevoz i zdravstvene potrebe. Kako bi se ustanovio nivo priuštivosti nekretnine koristi se Price-to-Income ratio. Njime se može definisati da li se odabrana nekretnina uklapa u pretpostavku da 30% mesečnih prihoda odlazi na troškove kupovine.

Ljudska jedinstvenost se odražava i na njihove preferencije kada je u pitanju odabir nekretnine. Glavni faktor kojim se definišu neophodne karakteristike nekretnine jeste njena namena u budućnosti. Zavisno od namena određuju se osobine koje nekretnina mora da poseduje. Demografska slika domaćinstva, kao što su bračni status i veličina porodice značajno utiče na njihove preferencije. Udaljenost škola i parkova je bitan faktor za porodice sa decom. Osim demografske slike, odluka se bazira na udaljenosti od radnog mesta u zavisnosti od načina transporta do istog. U univerzitetskim gradovima, preferencija za kupovinu može biti blizina univerziteta, ukoliko je izdavanje motivacija za kupovinu. Samci i mladi ljudi koji nemaju striktno definisane obaveze će se odlučiti na nekretnine bliže centru grada. Urbanost dela grada takođe igra veliku ulogu, bilo da ga pojedinac kupuje za sebe ili je cilj izdavanje.

Razvijena su dva načina za merenje i analiziranje promena u ceni na tržištu nekretnina. Praćenje promene kvaliteta određene stambene jedinice tokom vremena, drugim rečima posmatranje ponovljenih prodaja. Na taj način se najbolje mogu uočiti principi po kojima se cena menja tokom vremena. Nedostatak ovog pristupa je veličina uzorka pogodnog za analizu jer se pojedine nekretnine na tržištu ne pojavljuju često. Drugi

nedostatak ovog modela je uzorak koje ne oslikava lokalno tržište na najbolji način, naročito ukoliko se posmatra kraći vremenski period. Hedonistički metod je drugi način praćenja promena.

Princip vrednovanja primenom hedonističke regresije se zasniva na ideji da se svako heterogeno dobro može opisati preko skupa njegovih karakteristika. Hedonistička regresija omogućuje da se karakteristike, kojima se detaljno prikazuju osobine dobra, iskažu u numeričkom obliku. Svaka karakteristika doprinosi ukupnoj vrednosti dobra u zavisnosti od njene važnosti. Preferencija krajnjeg kupca određuje koliki uticaj određena karakteristika ima na vrednost dobra. Prilikom kupovine nekretnine kupac teži da uskladi svoje preferencije sa ponudom na tržištu. Može se reći da želi da maksimizira korisnost koju kupovina nekretnine donosi.

Jedna od primena hedonističkog metoda je da se proceni spremnost na kupovinu zbog određenih karakteristika ili marginalni trošak njihove proizvodnje. Tema ovog rada je da se odredi spremnost na kupovinu.

U regresionoj jednačini komponente i karakteristike predstavljaju nezavisne promenljive. Kako ne postoji tržište za svaku karakteristiku jer se njima ne može trgovati zasebno, njihovi marginalni doprinosi se implicitno određuju na osnovu ponude i tražnje za nekretninom.

Nekretnine se smatraju izuzetno heterogenim proizvodom jer ne postoje dve identične na tržištu. Imajući u vidu raznolikost kvalitativnih i kvantitativnih faktora koji je definišu može se reći da je svaka nekretnina jedinstvena na svoj način. Hedonistički pristup se uvek zasniva na istoj ideji. Kompleksno dobro se razlaže na sopstvene osobine i karakteristike, procenjuju se vrednosti svake od njih i formira se vrednost dobra kao zbir njegovih obeležja.

Nepokretnost se može opisati kroz tri grupe karakteristika:

1. *Fizičke ili strukturalne karakteristike*: površina, broj soba, broj kupatila, raspored prostorija, godina izgradnje, arhitektonski stil i dizajn, sprat na kojem se nalazi, ukupna spratnost zgrade (ukoliko je stan u pitanju), tip nekretnine, da li poseduje terasu, bazen, parking mesto ili garažu;
2. *Lokacijske karakteristike*: geografske i demografske karakteristike okruženja, tip naselja, deo grada ili ulica u kojoj se nalazi, blizina objekata od važnosti (zdravstvene ustanove, pošta, banka, škole, marketa), udaljenost od posla, blizina glavnog puta, autoputa, prometnost ulice, udaljenost od centra grada, dostupnost javnog prevoza;
3. *Ekološke karakteristike*: Procenat zelenila, udaljenost od parka ili rekreativnog centra, nivo buke, kvalitet vode, zagađenost vazduha

Uspešnost procene primenom hedonističkog modela zavisi od kvaliteta dostupnih podataka. Za najpreciznije rezultate analize podaci moraju da budu detaljni. Ukoliko ne postoje adekvatno prikupljeni podaci, primena ovog modela je komplikovana i skupa za sprovođenje.

Koncept hedonističke regresije se zasniva na izračunavanju cene nekretnine ili tražnje za nekretninama tokom nekog vremenskog perioda u zavisnosti od karakteristika koje je detaljnije opisuju. Možemo reći da je osnovna ideja da se cena proizvoda ili usluge može razložiti na njegove različite komponente ili attribute, pri čemu svaki doprinosi ukupnoj vrednosti u očima potrošača. Zavisna promenljiva može biti tražnja ili cena nekretnine. Nezavisne promenljive su karakteristike nekretnina i mogu biti numeričke ili kategorijalne varijable. Kategorijalne varijable se u regresijskoj analizi koriste za formiranje različitog tretmana obeležja. Postoje tri vrste kategorijalnih varijabli: ordinalne, nominalne i binarne. Za ocenjivanje koeficijenata modela se najčešće koristi metod najmanjih kvadrata. Metod najmanjih kvadrata je statistička tehnika kojom se opisuje linearna veza između zavisne i nezavisnih promenljivih sa ciljem da se pronađe funkcija koja minimizuje odstupanje od stvarnih podataka. Matematički rečeno metod najmanjih kvadrata minimizuje sumu kvadrata reziduala. Pre primene ove metode potrebno je ispitati normalnost grešaka i homoskedastičnost. Prvi korak prilikom određivanja vrednosti je vrednovanje sastavnih komponenti koje predstavljaju implicitne vrednosti. Nakon određivanja vrednosti obeležja, cena dobra ili tražnja za posmatranim dobrom se formira kao njihova suma. Kao što je ranije navedeno ne postoji tržište za svaku od sastavnih komponenti. U tom slučaju posmatra se korisnost koju krajnji kupac ima od pojedinačne komponente i koliko je spreman da plati za istu.

3.3 Hedonistički model vrednovanja cena nekretnina

U zavisnosti od predmeta procene, razlikujemo dva tipa modela hedonističke regresije. Aditivni model je model hedonističke regresije čiji je predmet procene cena nekretnine. Ukoliko je ipak predmet procene elastičnost onda je u pitanju log model hedonističke regresije.

Hedonistička jednačina se definiše na sledeći način:

$$P = f(S, L, E, t, \epsilon)$$

Pri čemu je:

P - cena nepokretnosti ili tražnja za nepokretnostima

S - strukturalne karakteristike

L - lokacijske karakteristike

E - ekološke karakteristike

t - vremenska promenljiva

ϵ - slučajna greška (beli šum)

Glavna karakteristika nezavisnih promenljivih koje se koriste za formiranje hedonističkih modela jeste što su mnoge karakteristike kategorijalne. Postoje tri vrste kategorijalnih promenljivih, binarne, ordinalne i nominalne. Binarne, kao što samo ima kaže imaju samo dve vrednosti. Ordinalne i nominalne kategorijalne promenljive imaju više od dve vrednosti. Razlika između njih je što kod ordinalnih postoji jasan poredak, dok kod nominalnih poredak nije izražen. Sve kategorijalne karakteristike u hedonističkom modelu se predstavljaju pomoću indikator varijabli. Indikator varijable su binarne varijable koje uzimaju samo dve vrednosti 0,1 i koriste se kod karakteristika koje mogu imati samo dve inicijalne vrednosti. Na primer, da li kuća poseduje bazen, binarna varijabla će uzeti vrednost 1 ukoliko je odgovor da, u suprotnim uzima vrednost 0. U hedonističkom pristupu kategorijalne varijable koje uzimaju više od dve vrednosti se takođe predstavljaju uz pomoć indikator varijabli tako što ćemo u model uvesti onoliko varijabli koliko kategorijalna promenljiva uzima različitih vrednosti. Na primer, ukoliko posmatramo opremljenost nekretnine, vrednost kategorijalne varijable može da bude nameštena, polunameštena i nenameštena. U model u tom slučaju uvodimo tri indikator varijable. Indikator varijable su poznate kao dummy varijable koje se definišu na sledeći način:

$$D_{i,t} = \begin{cases} 1, & \text{ako nekretnina poseduje posmatranu karakteristiku} \\ 0, & \text{ako nekretnina ne poseduje posmatranu karakteristiku} \end{cases} \quad (5)$$

Pored nezavisnih varijabli koje bliže opisuju nekretnine, u model se može uključiti i vremenska promenljiva (eng. Time Dummy variable). Vremenske promenljive se takođe predstavljaju pomoću indikator varijable. Vremensku promenljivu možemo da tumačimo na sledeći način:

$$D_{i,t} = \begin{cases} 1, & \text{ako se transakcija realizovala u periodu } t \\ 0, & \text{inače} \end{cases} \quad (6)$$

Uopšteni oblik hedonističkog modela se može zapisati:

$$P_{i,t} = \beta_0 + \sum_{k=1}^K \beta_{i,k} x_{i,k} + \sum_{j=1}^L \gamma_{i,j} y_{i,j} + \sum_{t=1}^T \delta_t D_{i,t} + \epsilon_{i,t}$$

Iz hedonističke jednačine vidimo da cenu nekretnine možemo izraziti kao zbir njenih karakteristika iz posmatranog skupa obeležja, kategorijalnih promenljivih i vremenskih dummy promenljivih.[17]

Pri čemu:

$P_{i,t}$ predstavlja ostvarenu prodajnu cenu stambene jedinice i u vremenu t pri čemu $i = 1, \dots, n$ i $t = 1, \dots, T$
 $x_{i,k}$ predstavlja karakteristiku j stambene jedinice i , dok $i = 1, \dots, n$, $k = 1, \dots, K$
 $\beta_{i,k}$ su parametri koji stoje uz numeričke varijable, $i = 1, \dots, n$, $k = 1, \dots, K$
 $y_{i,j}$ predstavlja indikator varijable j stambene jedinice i , dok $i = 1, \dots, n$, $j = 1, \dots, L$
 $\gamma_{i,j}$ su parametri koji stoje uz indikator varijable, $i = 1, \dots, n$, $j = 1, \dots, L$
 $D_{i,t}$ su vremenske indikator varijable nekretnine i u periodu t , dok $i = 1, \dots, n$, $t = 1, \dots, T$ $\epsilon_{i,t}$ je slučajna greška u periodu t povezana sa nekretninom i

Ukoliko je model log linearan, cena se izračunava na sledeći način:

$$\hat{P}_{i,t} = e^{\hat{\beta}_0 + \sum_{k=1}^K \hat{\beta}_{i,k} x_{i,k} + \hat{\gamma} + \hat{\delta}}$$

Hedonistički modeli se oslanjaju na nekoliko osnovnih pretpostavki:

1. $\epsilon_i : \mathcal{N}(0, \sigma^2)$, $i = 1, \dots, n$
2. $E(\epsilon_i) = 0$, $i = 1, \dots, n$
3. $var(\epsilon_i) = \sigma^2$, $i = 1, \dots, n$
4. $cov(\epsilon_i, \epsilon_j) = 0$, $i, j = 1, \dots, n$ i $i \neq j$
5. $n > K$
6. $cov(x_{i,j}, x_{i,k}) = 0$, $j, k = 1, \dots, K$ i $j \neq k$

Hedonistički model može da bude linearni i log-linearni model. Formulacija modela je ista, samo je tumačenje drugačije. U linearnom modelu imamo direkto povećanje cene, dok u log-linearnom modelu posmatramo logaritamsku cenu te promene nezavisnih promenljivih tumačimo kao procentualni uticaj na cenu.

Linearni model je statistička metoda koja se koristi za modeliranje linearnog odnosa između zavisne promenljive i jedne ili više nezavisnih promenljivih. Primena linearnog oblika hedonističkog modela vrednovanja je najadekvatnija na tržištima na kojima su kupci i proizvođači heterogeni. Tržište nekretnina karakteriše izuzetna heterogenost. Kupci kao učesnici na tržištu mogu imati različite preferencije prilikom donošenja odluke o kupovini, dok svaka nekretnina ima različite karakteristike. Zavisna promenljiva je ona koju model pokušava da predvidi korišćenjem jedne ili više nezavisnih promenljivih. U zavisnosti od broja nezavisnih promenljivih razlikujemo jednostruku i višestruku linearnu regresiju. U osnovi linearne regresije leži pretpostavka da postoji linearna veza između zavisne i nezavisnih promenljivih, što znači da se promena u vrednostima nezavisnih promenljivih proporcionalno odražava u promeni vrednosti zavisne promenljive.

Log - linearni model omogućuje da vrednosti određene karakteristike variraju proporcionalno sa vrednošću drugih karakteristika, što mu obezbeđuje prednost u odnosu na svoj linearni pandan. Takođe posmatranjem logaritamske vrednosti umanjuje se rizik od heteroskedastičnosti.

Jednačina hedonističke regresije se može zapisati u matricinom obliku:

$$P = Z\omega + \epsilon$$

Z je matrica svih nezavisnih promenljivih dimenzija $n \times (K+L+T+1)$
 $\omega = (\beta_0, \beta_1, \dots, \beta_K, \gamma_1, \dots, \gamma_L, \delta_1, \dots, \delta_T)$ vektor dimenzije $(K+L+T+1) \times 1$
 ϵ vektor dimenzije $n \times 1$

$$\begin{bmatrix} P_1 \\ P_2 \\ \vdots \\ P_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{12} & \dots & x_{1K} & y_{11} & \dots & y_{1L} & D_{11} \\ 1 & x_{21} & x_{22} & \dots & x_{2K} & y_{21} & \dots & y_{2L} & D_{21} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots & \ddots \\ 1 & x_{n1} & x_{n2} & \dots & x_{nK} & y_{n1} & \dots & y_{nL} & D_{n1} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_K \\ \gamma_1 \\ \vdots \\ \gamma_L \\ \delta_1 \\ \vdots \\ \delta_T \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{bmatrix}.$$

3.4 Ocena nepoznatih parametara

Nepoznati parametri[18] $(\beta_0, \beta_1, \dots, \beta_K, \gamma_1, \dots, \gamma_L, \delta_1, \dots, \delta_T)$, $x \in \mathbb{R}^{K+L+T+1}$ hedonističkog modela se najčešće ocenjuju metodom najmanjih kvadrata (Ordinary Least Square, OLS).

Metod najmanjih kvadrata se zasniva na ideji o minimizaciji sume kvadrata reziduala:

$$\min_x \sum_{i=1}^n (P_{i,t} - E(P_{i,t}))^2 = Q(x)$$

Rešavanjem sistema jednačina dobijenih preko parcijalnih izvoda prvog reda dobijamo ocene parametara modela $(\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_K, \hat{\gamma}_1, \dots, \hat{\gamma}_L, \hat{\delta}_1, \dots, \hat{\delta}_T)$

$$\frac{\partial Q}{\partial \beta_0} = -2 \sum_{i=1}^n (P_{i,t} - \hat{P}_{i,t}) = 0$$

$$\frac{\partial Q}{\partial \beta_{i,k}} = -2 \sum_{i=1}^n x_{i,k} (P_{i,t} - \hat{P}_{i,t}) = 0$$

$$\frac{\partial Q}{\partial \gamma_{i,j}} = -2 \sum_{i=1}^n y_{i,j} (P_{i,t} - \hat{P}_{i,t}) = 0$$

$$\frac{\partial Q}{\partial \delta_t} = -2 \sum_{i=1}^n D_{i,t} (P_{i,t} - \hat{P}_{i,t}) = 0$$

3.5 Analiza reziduala

Prema teoremi Gaus-Markova ocene parametara primenom metode najmanjih kvadrata su najefikasnije ukoliko imaju sledeće osobine:

1. $E(\epsilon_i) = 0 \quad i = 1, \dots, n$
2. $D(\epsilon_i) = \sigma^2 \quad i, j = 1, \dots, n$
3. $cov(\epsilon_i, \epsilon_j) = 0$ za $i \neq j$ i $i, j = 1, \dots, n$

Hedonistički model mora da ispunjava neke osnovne pretpostavke. Narušavanje pretpostavki utiče na tačnost interpretacije rezultata modela.

3.5.1 Normalnost

Osnovna pretpostavka na koju se oslanja hedonistički model jeste da reziduali prate normalnu raspodelu. Normalnost reziduala se proverava kao i kod ARIMA modela primenom *Anderson-Darling* test statistike i *Shapiro-Wilk* test statistike. Nulta hipoteza ovih testova je da reziduali prate normalnu raspodelu. Ukoliko je vrednost $p > \alpha$ gde je α nivo značajnosti nemamo razlog da odbacimo nultu hipotezu.

3.5.2 Homoskedastičnost

Homoskedastičnost se odnosi na situaciju u kojoj je varijansa grešaka (reziduala) u regresionom modelu konstantna za svaku opservaciju. Može se reći i da je varijansa stabilna. Jednostavnije rečeno, to znači da raspršenost reziduala ostaje uniformna bez obzira na vrednosti nezavisnih promenljivih. Homoskedastičnost, ili homogenost varijanse, podrazumeva da je varijansa grešaka ista za sve vrednosti nezavisnih promenljivih. Ovo stanje osigurava da raspršenost reziduala bude dosledna duž celog opsega vrednosti prediktorskih promenljivih. Nasuprot tome heteroskedastičnost predstavlja kršenje pretpostavke o konstantnoj varijansi grešaka u regresionom modelu. Ako se raspršenost reziduala sistematski širi ili sužava, to sugeriše na heteroskedastičnost. Takođe, možemo reći da ukoliko posmatranjem grafičkog prikaza uočimo da položaji reziduala obrazuju oblik koji podseća na levak, to ukazuje na prisustvo heteroskedastičnosti. Priroda pojave heteroskedastičnosti najčešće nije poznata. Heteroskedastičnost reziduala je česta pojava kod podataka preseka.

Hipoteze koje se postavljaju su:

$$H_0(\text{var}(\epsilon_i) = \sigma^2)$$

$$H_1(\text{var}(\epsilon_i) = \sigma_i^2)$$

Nulta hipoteza kaže da su reziduali homoskedastični, odnosno da nema razloga da pretpostavimo suprotno. Ukoliko je vrednost $p < \alpha$ pri čemu je α posmatran nivo značajnosti.

Najpoznatiji testovi za ispitivanje prisustva heteroskedastičnosti su:

Goldfeld – Quandt test (gqtest)

Test statistika za Goldfeld-Quandt test se računa kao odnos varijansi reziduala dva poduzorka. Ako je s_1^2 varijansa reziduala prvog poduzorka i s_2^2 varijansa reziduala drugog poduzorka, test statistika F je

$$F = \frac{s_1^2}{s_2^2}.$$

Pod nultom hipotezom homoskedastičnosti, F sledi F-distribuciju sa $n_1 - p$ i $n_2 - p$ stepenima slobode, gde su n_1 i n_2 broj opservacija u dva poduzorka, a p je broj parametara u regresijskom modelu.

Breusch – Pagan – Godfrey test (bptest)

Breusch-Pagan-Godfrey test se koristi za testiranje heteroskedastičnosti u rezidualima linearnog regresijskog modela. Test ispituje da li varijansa reziduala varira sa nivoom nezavisnih promenljivih, što može ukazivati na heteroskedastičnost.

Test statistika za Breusch-Pagan-Godfrey test se izračunava na osnovu korelacije između kvadrata reziduala iz originalnog regresijskog modela i predviđenih vrednosti nezavisne promenljive. Ako su e_i reziduali iz originalnog modela, test statistika LM (Likelihood Ratio Statistic) je

$$LM = nR^2,$$

gde je n broj opservacija, a R^2 je koeficijent determinacije, predstavlja meru koja pokazuje koliki deo varijanse zavisne promenljive je objašnjen varijansom nezavisne promenljive (ili nezavisnih promenljivih) u regresijskom modelu.

Pod nultom hipotezom homoskedastičnosti, LM statistika asimptotski sledi χ^2 distribuciju sa k stepenima slobode, gde je k broj nezavisnih promenljivih uključenih u test regresiju.

Neki autori, kao što su Bickel(1978) i Koenker(1981) tvrde da Breusch-Pagan-Godfrey test nije precizan kada podaci nemaju normalnu distribuciju.

3.5.3 Autokorelacija

Autokorelacija ili serijska korelacija, odnosi se na korelaciju između uzastopnih posmatranja u vremenskom nizu ili reziduala regresionog modela. Jednostavno rečeno, ovom funkcijom se utvrđuje da li postoji sistematski obrazac u rezidualima koji ukazuje na prisustvo korelacije između posmatranja u različitim trenucima vremena.

Reziduali nisu korelisani ukoliko je za $i \neq j$ $E(\epsilon_i, \epsilon_j) = 0$. Narušavanjem ove pretpostavke se zaključuje da postoji korelacija među rezidualima te da reziduali prate prepoznatljiv obrazac.

Identifikovanje postojanja autokorelacije među rezidualima se proverava pomoću *Durbin-Watson* testa. Ovo je metoda statističke analize koja je osmišljena i razvijena sredinom 20. veka. Primenjuje se za testiranje reziduala dobijenih pomoću metode najmanjih kvadrata. Nulta hipoteza ovog testa tvrdi da nema autokorelacije među rezidualima, dok alternativna hipoteza sugerise postojanje autokorelacije među rezidualima.

Test statistika ima oblik:

$$DW = \frac{\sum_{i=2}^n (\epsilon_i - \epsilon_{i-1})^2}{\sum_{i=1}^n \epsilon_i^2}$$

3.5.4 Multikolinearnost

Multikolinearnost nastaje kada postoji visoka korelacija između dve ili više nezavisnih promenljivih u regresionom modelu. Ne odnosi se direktno na zavisnu promenljivu, već se odnosi na odnose između nezavisnih promenljivih. Prisustvo multikolinearnosti otežava identifikaciju pojedinačnog doprinosa svake nezavisne promenljive zavisnoj promenljivoj. Multikolinearnost se može detektovati različitim metodama. Posmatranjem matrice korelacije između nezavisnih promenljivih utvrđuju se koeficijenti korelacije. Ukoliko su koeficijenti blizu 1 ili -1, sugerisu na multikolinearnost. VIF (faktor inflacije varijanse) kvantifikuje koliko se povećava varijansa procenjenog koeficijenta regresije ako su nezavisne promenljive korelirani. Vrednost VIF veća od 10 često se smatra indikatorom multikolinearnosti. Uključivanjem velikog broja nezavisnih promenljivih u odnosu na veličinu posmatranog uzorka može dovesti do nastanka multikolinearnosti. Takođe uzrok može biti ukoliko dve ili više nezavisnih promenljivih predstavljaju istu posmatranu karakteristiku, kao i ako su dve ili više nezavisnih promenljivih skoro savršeno linearno povezane. Važno je napomenuti da je multikolinearnost

svojestvo uzorka, a ne populacije.

Pretpostavimo da imamo k nezavisnih promenljivih X_k . Za svaku od k promenljivih formiramo model linearne regresije i izrazimo R_k^2 .

Faktor inflacije varijanse se izračunava:

$$VIF = \frac{1}{1 - R_k^2} \geq 1$$

Ukoliko je vrednost R_k^2 blizu 1, što ukazuje na visoku zavisnost modela koji odgovara nezavisnoj promenljivoj X_k , VIF vrednost će biti visoka. Ukoliko je $VIF > 5$ postoji samo sumnja, a za $VIF > 10$ možemo sa sigurnošću da tvrdimo da postoji multikolinearnost.

3.5.5 Uticajne tačke i outlier-i

Outlier je posmatranje koje se nalazi na značajnom rastojanju od ostalih vrednosti u slučajnom uzorku iz posmatrane populacije. Jednostavnije rečeno, to je podatak koji značajno odstupa od ostalih podataka. Outlieri mogu nastati iz različitih razloga, kao što su greške u merenju, eksperimentalne greške ili prirodna varijacija u podacima. Outlieri mogu imati značajan uticaj na statističke analize i modele, potencijalno iskrivljujući rezultate i dovodeći do netačnih zaključaka. Detektovanje i rukovanje outlierima ključni su koraci u analizi podataka kako bi se osigurala robustnost statističkih analiza. Uticajna tačka je podatak koji ima značajan uticaj na procene parametara modela. Drugim rečima, isključivanje uticajne tačke iz posmatranog skupa izaziva značajnu promenu parametara modela. Uticajne tačke mogu izvršiti veliki uticaj na model, često zbog njihovih ekstremnih vrednosti ili strateškog položaja u odnosu na ostale podatke. Iako i outlieri i uticajne tačke predstavljaju podatke koji se značajno razlikuju od norme, outlieri se obično razmatraju u kontekstu univarijantnih distribucija podataka, dok se uticajne tačke odnose na njihov uticaj na statističke modele, posebno u regresionoj analizi.

Neke od najpoznatijih metoda za identifikovanje outlier-a i uticajnih tačaka su Cook's rastojanje, Leverage metod, Dffits, Dfbeta's, Atkinson's.

Cook's rastojanje se koristi za identifikovanje uticajnih tačaka koje mogu znatno uticati na procenjene parametre modela.

$$C_i = \frac{\sum_{j=1}^n (\hat{Y}_j - \hat{Y}_{j(i)})^2}{n \sum_{i=1}^n (Y_i - \hat{Y}_i)^2}$$

Pri čemu je:

Y_j stvarna vrednost

\hat{Y}_j predviđenja vrednost za j posmatranje

$\hat{Y}_{j(i)}$ predviđenja vrednost za j posmatranje kada se isključi tačka i

p je broj parametara modela

Što je Cook's rastojanje veće za neko posmatranje i , smatra se da je to posmatranje uticajnije. Obično se posmatranja kod kojih Cook's rastojanje prelazi određeni prag smatraju uticajnim. Najčešće ovaj prag iznosi $\frac{4}{n}$ gde je n broj opservacija.

3.6 Evaluacija modela

Upoređivanjem vrednosti kriterijuma AIC i BIC pomenutih ranije u radu (2.9 Izbor modela) kao i poređenjem grešaka modela i vrednosti R^2 i $Adj.R^2$ navedenih u (2.11 Kvalitet predikcija) odabire se najbolji model.

3.7 Formiranje hedonističkog indeksa cena nekretnina

Značajan uticaj promena cena nepokretnosti na finansijsku stabilnost ukazuje na potrebu za pouzdanim pokazateljem koji će pratiti njihovu dinamiku.

Konstruisanje indeksa cena nekretnina, u odnosu na indeks potrošačkih cena je komplikovaniji zadatak jer nije moguće neprekidno pratiti kretanje cene određene nekretnine tokom vremena jer se svaka pojedinačna nekretnina prometuje jako retko. Kako je ranije spomenuto, nekretnine predstavljaju izuzetno heterogene proizvode. Jedinствене specifikacije komplikuju njihovo međusobno poređenje i otežavaju formiranje indeksa. Pored toga, važnu ulogu imaju kvalitetni podaci na osnovu kojih se gradi indeks. Primena koncepta hedonističkog indeksa je izazovan zadatak na tržištima koja nemaju adekvatnu bazu podataka koja uključuje informacije vezane za samu transakciju kao i podatke o karakteristikama samih nepokretnosti.

U svom radu iz 2016. godine, Mark Silver je istakao da je uloga hedonističkih indeksa da prate promenu u kvalitetu proizvoda tokom vremena koja utiče na promenu cene posmatranog proizvoda.

Specifičnost nekretnina, kao i činjenica da se pojedinačnom nekretninom ne prometuje često, onemogućavaju praćenje promena kvaliteta tokom vremena.

Indeksi se mogu formirati korišćenjem prosečne ili medijalne cene stambene jedinice. Kreiranjem indeksa zasnovanog na prosečnoj ceni stambene jedinice odlikuje jednostavnost konstrukcije i interpretacije. Nedostatak je što može da se javi problem pristrasnosti ukoliko su među podacima prisutni *outlier-i*. U određenom periodu su prometovane luksuzne nekretnine ili ipak nekretnine koje su prometovane za nisku vrednost. Imajući to na umu potrebno je eliminisati pristrasnost i izvršiti stratifikaciju podatka u odnosu na kvalitet. Sa druge strane, indeksi zasnovani na medijalnoj ceni su manje osetljivi na prisutnost *outlier-a*.

Hedonistički metod konstruisanja indeksa zahteva skup proizvoda koji su opisani pomoću detaljnih osobina pa je samim tim njegovo formiranje zahteva izvesna ulaganja prilikom prikupljanja potrebnih podataka. Ukoliko dođe do promene kvaliteta nekih od proizvoda unutar posmatranog skupa, indeks gubi na efikasnosti predviđanja tokom vremena.

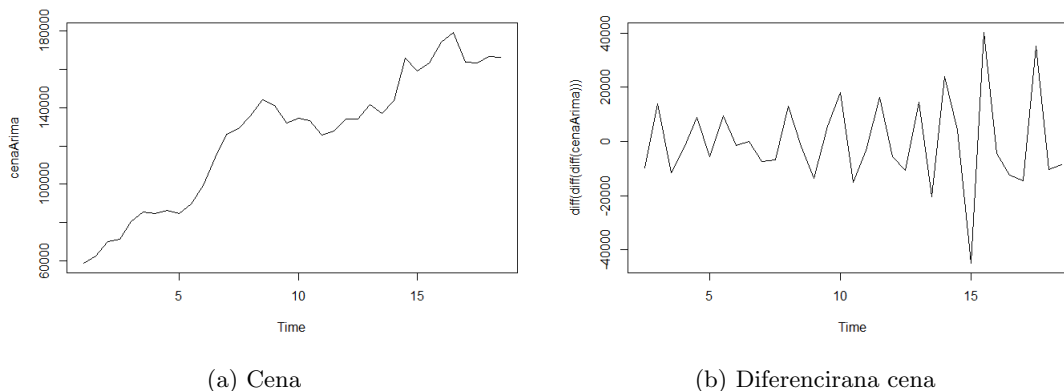
4 Rezultati empirijskog istraživanja

U okviru empirijskog istraživanja, fokusirali smo se na primenu ARIMA i hedonističkog modela za analizu realnih podataka. Koristeći R Studio, sofisticiran alat za statističku analizu i grafički prikaz podataka, omogućeno je dublje razumevanje kompleksnosti i dinamike podataka koji su predmet istraživanja. Ovi modeli su pažljivo konfigurisani i testirani kako bi osigurali preciznost i pouzdanost rezultata

4.1 Rezultati primene ARIMA modela na realnim podacima

ARIMA je formiran na hronološkim podacima o prosečnim cenama stanova novogradnje po kvadratnom metru. Uzor za izgradnju ovog modela odabran je temeljem pregleda i analize više stručnih radova posvećenih ovoj specifičnoj oblasti[19][20][21]. Podaci su preuzeti sa sajta Republičkog zavoda za statistiku[22]. Serija podataka analizirana u modelu obuhvata period od prve polovine 1994 do prve polovine 2023 i podaci su iskazani na polugodišnjim intervalima. Vremenska serija obuhvata 59 opservacija. Pritom se stanom smatra svaka građevinski povezana celina namenjena za stanovanje, koja se sastoji od jedne ili više soba sa odgovarajućim pomoćnim prostorijama(kuhinja, ostava, predsoblje, kupatilo i slično) ili bez pomoćnih prostorija, i ima jedan ili više posebnih ulaza direktno iz hodnika, stepeništa, dvorišta ili sa ulice. Prosečne cene koje su objavljene po jednom kvadratnom metru rezultat su različitih karakteristika izgradnje i razlike u mesecu u kome je sklopljen kupoprodajni ugovor na jednoj teritoriji. Prosečne cene koje su analizirane ARIMA modelom, su obračunate na republičkom nivou i u cenu je uračunat PDV, pritom cene nisu prilagođene inflaciji. Republički zavod za statistiku[22] preuzima podatke za polugodišnje istraživanje o cenama stanova novogradnje iz administrativnog izvora podataka, Registra cena nepokretnosti koji vodi Republički geodetski zavod[7].

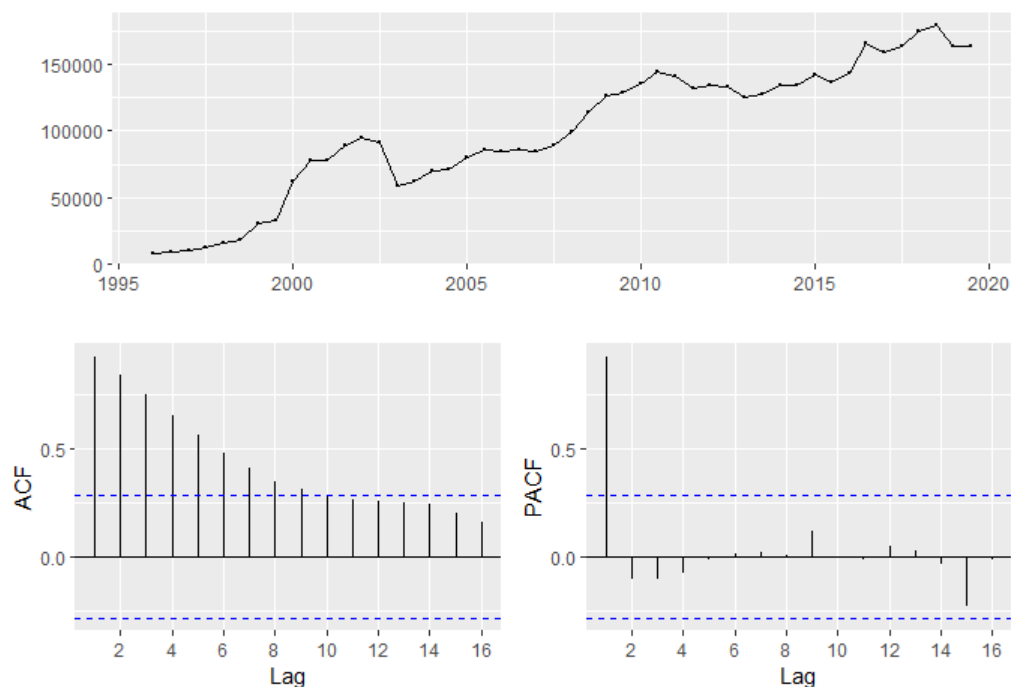
Najpre je prikazana na grafiku 1 vremenska serija prosečnih cena stanova po kvadratu i testirana je stacionarnost vremenske serije primenom Augmented Dickey-Fuller testa. Nulta hipoteza ovog testa kaže da serija nije stacionarna. Vremenska serija cena je tri puta diferencirana kako bi se dostigla stacionarnost.



Grafik 1: Stacionarnost vremenske serije cena

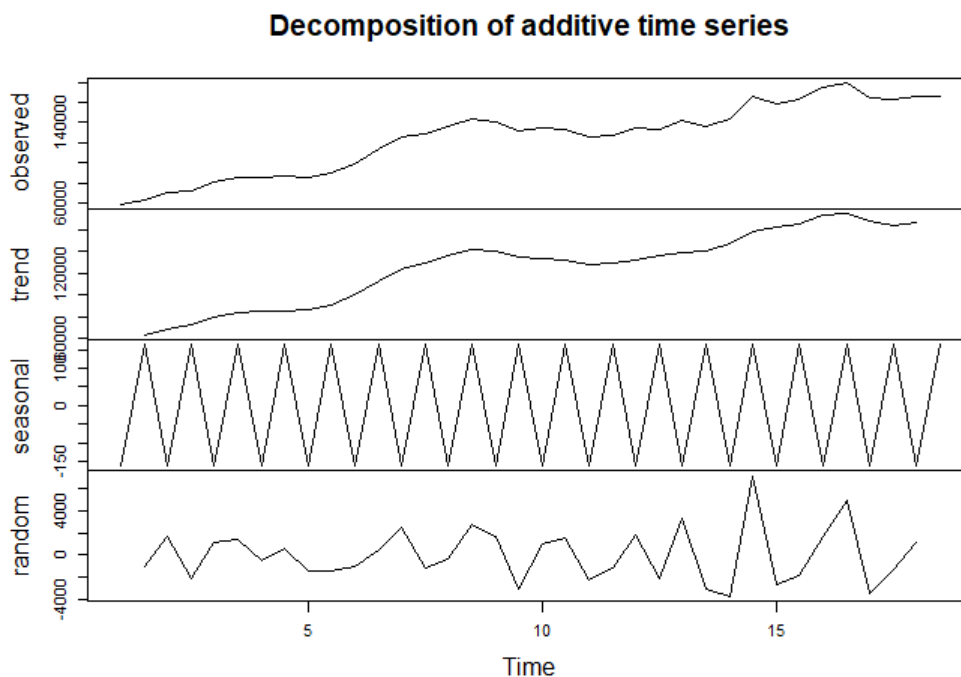
Na prikazu kretanja vremenske serije tokom određenog perioda (grafik 2), možemo posmatrati autokorelacionu i parcijalnu autokorelacionu funkciju kako bismo bolje razumeli svojstva ove vremenske serije. Autokorelaciona funkcija meri korelaciju između vrednosti u vremenskoj seriji i njenih prošlih lagova (prethodnih vrednosti). Na grafiku autokorelacione funkcije, primetno je da vrednosti eksponencijalno opadaju kako se pomeramo kroz prošle lagove i da padaju ispod praga značajnosti nakon 9 lagova. Ovo ukazuje na to da postoji statistički značajna autokorelacija samo za prvih nekoliko lagova, dok se dalje u prošlosti ta korelacija smanjuje i postaje neznačajna. Parcijalna autokorelaciona funkcija fokusira se na merenje korelacije između trenutne vrednosti i prošlih lagova, eliminišući uticaj međuvremenanja (lagova) između njih. Na

grafiku parcijalne autokorelacione funkcije, primetno je da je samo prvi lag statistički značajan jer izlazi van intervala poverenja. Ovo ukazuje na to da postoje značajne parcijalne autokorelacije samo za prvi lag



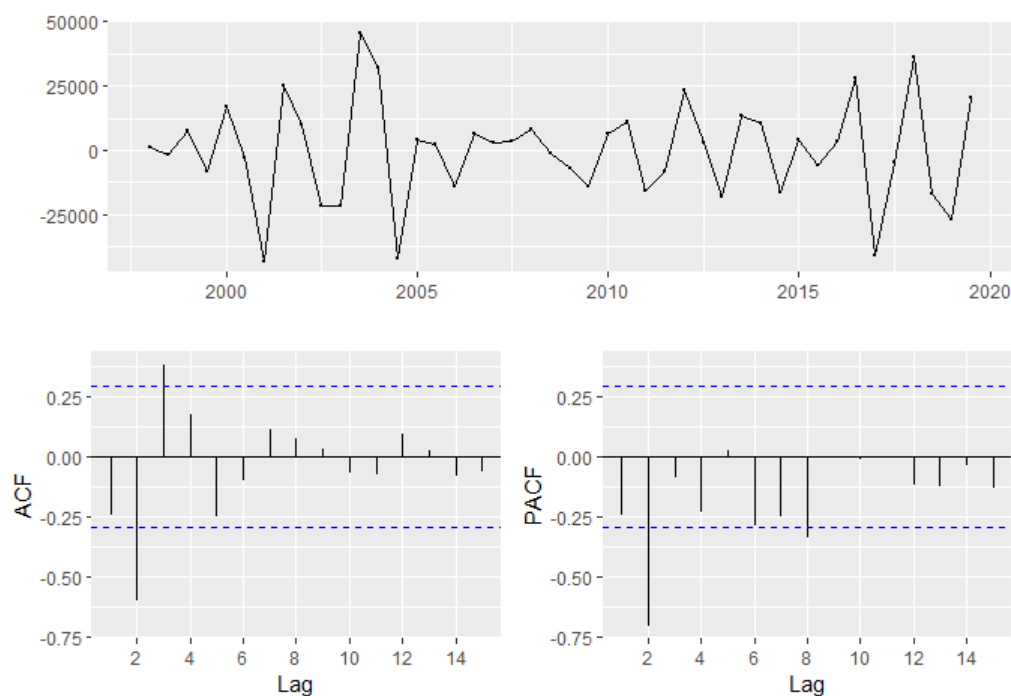
Grafik 2: Vremenska serija prosečna cena sa ACF i PACF funkcijama

Korišćenjem ugrađene funkcije `decompose` izvršena je dekompozicija vremenske serije na tri osnovna komponentna dela: trend, sezonalnost i slučajnu grešku. (grafik 3) Analiziranjem ovih komponenti, možemo dobiti dublji uvid u strukturu vremenske serije prosečnih cena po kvadratu. Vremenska serija ima uzlazni trend koji se postepeno povećava tokom vremena, uz male fluktuacije. Ovaj trend može biti posledica različitih faktora kao što su inflacija, rast tržišta ili dugoročne promene u ponudi i potražnji. Razumevanje ovog trenda je ključno za predviđanje budućih cena i donošenje dugoročnih odluka. Primetimo da vremenska serija ima sezonski deo. Sezonalnost se odnosi na ponavljajuće obrasce ili cikluse koji se javljaju u vremenskoj seriji u određenim vremenskim intervalima. Identifikacija sezonskih obrazaca može biti korisna za pravilno planiranje i upravljanje resursima, posebno u sektorima koji su podložni sezonskim fluktuacijama, kao što su turizam ili poljoprivreda. Komponenta greške predstavlja sve neregularne, nepredvidive fluktuacije u podacima koje nisu povezane ni sa trendom ni sa sezonskim oblicima. Razumevanje slučajne greške je važno kako bismo ocenili koliko dobro naši modeli mogu predviđati buduće vrednosti.



Grafik 3: Dekompozicija vremenske serije cena

Nakon dekompozicije vremenske serije, sledeći korak je postizanje stacionarnosti. U ovom slučaju, vremenska serija pokazuje sezonski karakter, pa je potrebno preduzeti odgovarajuće korake kako bismo postigli stacionarnost. Vremenska serija je sezonski diferencirana počevši od lag=2 jer je serija polugodišnja. Serija je sezonski diferencirana jednom što implicira da je parametar $D = 1$. Kako sezonskim diferenciranjem nije dostignuta stacionarnost serije, pored sezonskog diferenciranja izvršeno je i diferenciranje serija dva puta i time je određen parametar $d = 2$. Posle izvršenih diferenciranja, ispitivanje stacionarnosti je izvršeno primenom Dickey- Fuller testa, kojim je potvrđena stacionarnost vremenske serije. Možemo primetiti da je stacionarnost vremenske serije dostignuta diferenciranjem tri puta, kada sezonalnost nije razmatrana. Grafički je prikazana stacionarna vremenska serija, kao i funkcije autokorelacije i parcijalne autokorelacije. Na grafičkom prikazu (grafik 4) autokorelacione funkcije, može se uvideti da su vrednosti autokorelacije za lag=2 i lag=3 statistički značajno različite od nule. Istovremeno se grafičkim prikazom uviđa da je vrednost funkcije parcijalne autokorelacije za lag=2 statistički značajno različita od nule. Time se zaključuje da postoji određeni obrazac u fluktuacijama prosečnih cena koji se ponavlja na svaka dva ili tri vremenska perioda.



Grafik 4: Diferencirana vremenska serija sa funkcijama ACF i PACF

Na osnovu vrednosti autokorelacione funkcije i parcijalne autokorelacione funkcije formirano je šest modela za procenu prosečne cene stana po kvadratnom metru. Modeli su formirani sa različitim parametrima (p,d,q) (P,D,Q) . Parametri modela su $(p, 2, 0)$ $(P, 1, 0)$ za nesezonske parametre $p \in \{1, 2, 3\}$ i sezonske parametre $P \in \{1, 2\}$. Definisani su različiti $SARIMA(p, 2, 0)(P, 1, 0)$ modeli za $p \in \{1, 2, 3\}$ i $P \in \{1, 2\}$. Zatim se izračunava Akaike informacioni kriterijum (AIC) i Bayesian informacioni kriterijum (BIC) za svaki model, a modeli se rangiraju na osnovu AIC i BIC vrednosti, počevši sa modelom sa najnižom vrednošću.

Reziduali su testirani za svaki model. Normalnost reziduala je testirana primenom Shapiro-Wilk test statistike. Kod modela $ARIMA(2,2,0)(2,1,0)$ je jedino narušena pretpostavka o normalnosti reziduala. Kod ostalih modela, na osnovu p vrednosti zaključujemo da reziduali prate normalnu raspodelu. Autokorelacija je proverena primenom Box-Pierce testa i Ljung-Box testa. Nulta hipoteza ovih testova kaže da ne postoji autokorelacija između reziduala. Na osnovu p vrednosti koja je veća od nivoa značajnosti od 5%, kod svih modela je prihvaćena alternativna hipoteza o odsustvu autokorelacije reziduala. Rezultati testova su prikazani u tabeli 2.

Test		(1,2,0)(1,1,0)	(2,2,0)(1,1,0)	(2,2,0)(2,1,0)	(3,2,0)(1,1,0)	(3,2,0)(2,1,0)	(1,2,0)(2,1,0)
Shapiro-Wilk	t statistika	0,99151	0,97205	0,93043	0,97127	0,95384	0,97492
	p	0,9844	0,3037	0,007075	0,2837	0,05696	0,3885
Box-Pierce	t statistika	13,964	12,135	8,8125	12,208	5,4252	10,79
	p	0,1746	0,2761	0,55	0,2714	0,861	0,3741
Box-Ljung	t statistika	16,067	14,391	10,277	14,49	6,3726	12,546
	p	0,09773	0,1559	0,4165	0,1518	0,783	0,2502

Tabela 2: Rezultati testova reziduala sprovedinih na modelima

Tabela 3 prikazuje rezultate SARIMA modela za prognoziranje polugodišnjih cena po kvadratnom metru. AIC (Akaike-ov informacijski kriterijum), AICc (korigovani AIC) i BIC (Bajesov informacijski kriterijum) su često korišćene metrike za izbor modela. Niže vrednosti ovih kriterijuma ukazuju na bolje prilagođene modele. Model 3, ARIMA(2,2,0)(2,1,0)[2], se ističe kao model sa najnižim vrednostima AIC, AICc i BIC, što ukazuje na njegovu izuzetnu prilagodljivost podacima među navedenim modelima. Model 5, ARIMA(3,2,0)(2,1,0)[2], takođe pokazuje snažno performanse sa niskim vrednostima AIC i BIC, ali nešto višim od Modela 3. Model 2, ARIMA(2,2,0)(1,1,0)[2], takođe se dobro pokazao, ali je nešto manje povoljan u poređenju sa Modelima 3 i 5.

Model	AIC	AICc	BIC
ARIMA(1,2,0)(1,1,0)[2]	993,7	994,36	998,76
ARIMA(2,2,0)(1,1,0)[2]	966,24	967,27	973,38
ARIMA(2,2,0)(2,1,0)[2]	964,61	966,19	973,53
ARIMA(3,2,0)(1,1,0)[2]	968,2	969,78	977,12
ARIMA(3,2,0)(2,1,0)[2]	963,24	965,51	973,95
ARIMA(1,2,0)(2,1,0)[2]	969,52	970,54	976,65

Tabela 3: Vrednosti kriterijuma: AIC i BIC

Sledeći deo je posvećen formiranju vremenske serije prosečnih cena stanova novogradnje po kvadratnom metru bez uzimanja u obzir uticaj sezonalnosti. Formirani su modeli za dva različita in-sample perioda (training skupovi): I 1997 - II 2017 i I 2003 - II 2017. Modeli su testirani na out-of-sample skupu koji predstavlja podatke iz perioda I 2018 - I 2023. Kako bi se stabilizovala varijansa i linearizovali trendovi, cene su posmatrane u logaritamskom obliku. Izvršen je neophodan broj diferenciranja kako bi se postigla stacionarnost vremenske serije koja je proverena primenom Augmented Dickey Fuller testa čija alternativna hipoteza tvrdi da je vremenska serija stacionarna. Stracionarnost vremenske serije je dobijena različitim brojem diferenciranja, u zavisnosti od perioda na kojem je formiran model. Vremenska serija koja obuhvata period od 2003 do 2017 godine je diferencirana tri puta, odnosno $d = 3$, dok je vremenska serija perioda 1997 do 2017 godine diferencirana dva puta, $d = 2$ kako bi se postigla stacionarnost. Na rezidualima su sprovedeni testovi normalnosti i autokorelacije.

Prema vrednostima AIC (Akaike informacioni kriterijum) i BIC (Bayesian informacioni kriterijum) (tabela 4), najoptimalniji model je ARIMA(2,3,2) koji je formiran na osnovu podataka iz perioda od 2003. godine do 2017. godine. Ovi kriterijumi se koriste za procenu kvaliteta različitih statističkih modela, pri čemu manje vrednosti AIC i BIC ukazuju na bolje modele. ARIMA(2,3,2) je model koji daje najmanje vrednosti AIC i BIC, što znači da je optimalan prema oba kriterijuma.

Period	Model	sigma2	AIC	BIC	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
2003 - 2017	ARIMA(2,3,0)	0,00542	-58,25	-54,36	-0,00261	0,06721	0,05348	-0,02218	0,45813	1,00310	-0,11951
	ARIMA(2,3,1)	0,00396	-62,63	-57,44	0,00070	0,05631	0,04707	0,00544	0,40377	0,88297	0,01829
	ARIMA(2,3,2)	0,00343	-65,04	-58,56	-0,00135	0,05125	0,04443	-0,01167	0,38090	0,83334	-0,07932
	ARIMA(2,3,4)	0,00351	-61,77	-52,70	0,00050	0,04959	0,04166	0,00391	0,35722	0,78143	-0,02997
1997 - 2017	ARIMA(0,2,1)	0,02358	-32,64	-29,26	-0,01577	0,14796	0,09605	-0,13669	0,86298	0,93049	-0,07823
	ARIMA(5,2,1)	0,02504	-25,24	-13,42	-0,02680	0,14237	0,09237	-0,23166	0,82884	0,89481	-0,08274
	ARIMA(6,2,0)	0,02016	-32,14	-20,32	-0,01190	0,12776	0,09374	-0,10150	0,83341	0,90812	-0,06360
	ARIMA(6,2,1)	0,02069	-30,39	-16,88	-0,01432	0,12750	0,09421	-0,12338	0,83820	0,91266	-0,03527

Tabela 4: Vrednosti kriterijuma: AIC i BIC sa iznosima grešaka modela

U narednoj tabeli (tabela 5) je prikazano stvarno kretanje tržišta za periode za koje je izvršena predikcija cene, kao i rezultati predikcije svakog od posmatranih modela. Modeli koji su formirani na periodu 1997 – 2017 godine imaju mnogo niže vrednosti AIC i BIC i greške ovih modela su znatno veće. Međutim, modeli ARIMA(6,2,0) i ARIMA(6,2,1) bolje predviđaju pravac kretanja cena.

Polugodište:		I 2018	II 2018	I 2019	II 2019	I 2020	II 2020	I 2021	II 2021	I 2022	II 2022	I 2023
2003 - 2017	Kretanje tržišta:	Rast	Rast	Pad	Pad	Rast	Pad	Rast	Pad	Rast	Rast	Pad
	ARIMA(2,3,0)	Rast	Pad	Pad	Rast	Pad	Pad	Pad	Pad	Pad	Pad	Pad
	ARIMA(2,3,1)	Rast	Rast	Rast	Rast	Rast	Rast	Rast	Rast	Rast	Rast	Rast
	ARIMA(2,3,2)	Rast	Pad	Rast	Rast	Pad	Rast	Pad	Rast	Rast	Pad	Rast
	ARIMA(2,3,4)	Rast	Pad	Rast	Rast	Rast	Rast	Rast	Rast	Rast	Rast	Rast
1997 - 2017	ARIMA(0,2,1)	Rast	Rast	Rast	Rast	Rast	Rast	Rast	Rast	Rast	Rast	Rast
	ARIMA(5,2,1)	Rast	Rast	Rast	Rast	Rast	Pad	Rast	Rast	Rast	Rast	Rast
	ARIMA(6,2,0)	Rast	Rast	Pad	Pad	Rast	Pad	Rast	Rast	Rast	Rast	Pad
	ARIMA(6,2,1)	Rast	Rast	Pad	Pad	Rast	Pad	Rast	Rast	Rast	Rast	Pad

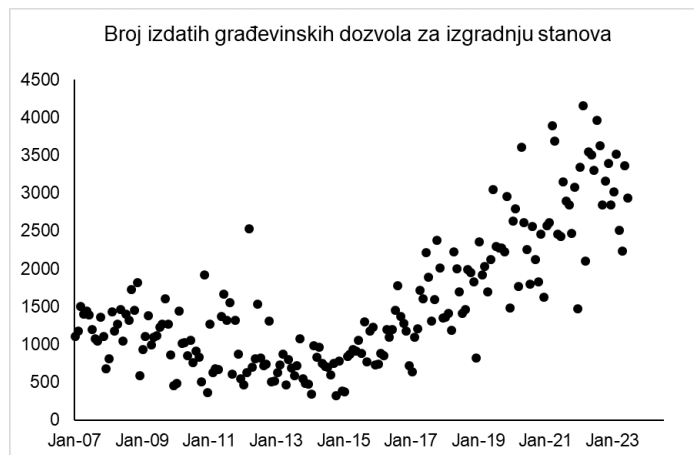
Tabela 5: Pravac kretanja tržišta

Sumirajmo prethodno prikazane rezultate. Na osnovu priloženih informacija o kretanju tržišta i rezultatima predikcija različitih modela, možemo izvesti nekoliko ključnih zaključaka. Prvo, primetno je da modeli koji su formirani u periodu 1997-2017. godine imaju niže vrednosti AIC i BIC, što ukazuje na bolje prilagođavanje podacima. Međutim, važno je napomenuti da ovi modeli imaju znatno veće greške u predikciji cena, što ukazuje na ograničenja ovih metrika kao samostalnih indikatora tačnosti modela. Drugo, modeli ARIMA(6,2,0) i ARIMA(6,2,1) se izdvajaju po tome što bolje predviđaju pravac kretanja cena. To je ključno za investitore i trgovce na tržištu, jer tačno predviđanje pravca kretanja često ima veći uticaj na donošenje odluka od apsolutnih cena. Treće, analizom kvadratne greške uočavamo da je model ARIMA(2,3,2) imao najveći broj predikcija cena koje su bile bliske stvarnim vrednostima. To sugerise da je ovaj model možda najpouzdaniji kada je reč o tačnosti predikcija.

	Polugodište:	I 2018	II 2018	I 2019	II 2019	I 2020	II 2020	I 2021	II 2021	I 2022	II 2022	I 2023
	Stvarna log cena	12,06899	12,09732	12,00744	12,00269	12,02433	12,02101	12,09989	12,08808	12,21101	12,29210	12,27381
2003 - 2017	ARIMA(2,3,0)	12,07745	12,02023	12,01629	12,02916	11,95144	11,91276	11,87369	11,77129	11,69547	11,60997	11,47954
	ARIMA(2,3,1)	12,07855	12,09115	12,11335	12,15693	12,18156	12,20439	12,23485	12,25951	12,28146	12,30533	12,32705
	ARIMA(2,3,2)	12,05542	12,02927	12,06912	12,08286	12,07780	12,10747	12,10698	12,11319	12,12826	12,12507	12,13258
	ARIMA(2,3,4)	12,05111	12,04165	12,09041	12,09811	12,11566	12,14393	12,14801	12,17192	12,18260	12,19306	12,20863
1997 - 2017	ARIMA(0,2,1)	12,03079	12,05716	12,08352	12,10988	12,13624	12,16260	12,18896	12,21532	12,24168	12,26804	12,29440
	ARIMA(5,2,1)	12,04252	12,08249	12,10447	12,10525	12,16987	12,16716	12,21339	12,22271	12,26080	12,28718	12,30202
	ARIMA(6,2,0)	12,00741	12,04568	12,02927	12,00698	12,09036	12,07567	12,10712	12,11860	12,15412	12,17661	12,15599
	ARIMA(6,2,1)	12,01332	12,04939	12,03789	12,00618	12,08381	12,08079	12,10823	12,12680	12,16263	12,19792	12,17881

Tabela 6: Vrednosti predviđene modelima po polugodištima

Podaci o broju građevinskih dozvola za izgradnju stanova su prikupljeni iz baze podataka Republičkog zavoda za statistiku[22]. (grafik 5) Građevinskim dozvolama se smatraju odobrenja za izgradnju koje se izdaju na zahtev investitora kako bi se omogućio početak građevinskih radova. Ovi podaci su prikupljeni radi istraživanja budućih trendova u građevinskoj aktivnosti i dostupni su od 2007. godine, objavljuju se na mesečnom nivou. U ovom istraživanju je analiziran kompletan vremenski period od januara 2007. do jula 2023. godine, obuhvatajući ukupno 199 posmatranja. Skup podataka je podeljen na dva podskupa. Trening skup obuhvata period od januara 2007. do decembra 2021. godine i sadrži 180 posmatranja, dok testing skup obuhvata period od januara 2022. do jula 2023. godine i sastoji se od 19 posmatranja. Ovaj pristup omogućava analitičarima da obuče modele za prognoziranje na osnovu trening skupa podataka i testira njihovu tačnost na testing skupu, što je koristan metod za procenu performansi modela.



Grafik 5: Broj građevinskih dozvola

Model	sigma	AIC	BIC	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
ARIMA(4,2,0)	341397	1285,56	1297,60	-4,92425	563,03790	444,83820	-5,35163	26,60704	1,04163	-0,07514
ARIMA(4,2,4)	196866	1255,61	1277,28	-13,40661	416,44910	310,58820	-7,29874	19,83404	0,72727	-0,01447
ARIMA(5,2,3)	209068	1259,62	1281,28	-13,68540	429,16120	323,49440	-7,62132	20,61110	0,75749	-0,00921
ARIMA(8,2,2)	194596	1258,00	1284,48	-4,45232	408,40730	313,43830	-6,60194	19,54659	0,73394	-0,02775

Tabela 7: Vrednosti AIC i BIC kriterijuma i iznosi grešaka

Tabela 7 prikazuje performanse ARIMA modela koristeći različite parametre. Model ARIMA(4,2,0) ima najveću vrednost sigma od 341,397, dok su njegove vrednosti RMSE i MAE znatno veće od drugih modela, što ukazuje na manje precizno prilagođavanje. Nasuprot tome, model ARIMA(4,2,4) pokazuje nižu vrednost sigma od 196,866 i ima poboljšane performanse sa RMSE od 416,44910 i MAE od 310,58820. Modeli ARIMA(5,2,3) i ARIMA(8,2,2) imaju sigma vrednosti od 209,068 i 194,596, respektivno. Oba ova modela imaju relativno uporedive vrednosti RMSE i MAE, što sugerise slične nivoe preciznosti. Međutim, svi modeli imaju negativne vrednosti ME, MPE i ACF1, što može ukazivati na potencijalne pristrasnosti ili probleme sa autokorelacijom. Kada se porede na osnovu AIC i BIC kriterijuma, koji su mere prilagođenosti modela, model ARIMA(4,2,4) se izdvaja sa najnižim vrednostima, što sugerise da bi mogao biti najbolje prilagođeni model među datim opcijama.

Dodatno, vredno je napomenuti da negativne vrednosti ME (srednje greške) i MPE (procentualne srednje greške) mogu ukazivati na sistematsko potcenjivanje modela, dok negativne vrednosti ACF1 (autokorelacija prvog reda) mogu sugerisati postojanje autokorelacije u rezidualima. Ovo može biti znak da model nije uspeo da uhvati svu strukturu vremenske serije, što može uticati na pouzdanost prognoza.

Tabela 8 prikazuje stvarno kretanje tržišta i predviđanje kretanja modela. Možemo videti da pad u oktobru 2022, maju 2023 i julu 2023 nijedan model nije predvideo.

Mesec:	jan.22	feb.22	mar.22	apr.22	maj.22	jun.22	jul.22	avg.22	sep.22	okt.22	nov.22	dec.22	jan.23	feb.23	mar.23	apr.23	maj.23	jun.23	jul.23
Kretanje tržišta:	Pad	Rast	Rast	Pad	Rast	Pad	Pad	Rast	Pad	Pad	Rast	Rast	Pad	Rast	Rast	Pad	Pad	Rast	Pad
ARIMA(4,2,0)	Rast	Pad	Rast	Rast	Rast	Pad	Pad	Rast	Rast	Rast	Pad	Rast	Rast	Rast	Rast	Rast	Rast	Rast	Rast
ARIMA(4,2,4)	Pad	Rast	Pad	Pad	Rast	Pad	Rast	Pad	Rast	Rast	Pad	Rast	Pad	Rast	Pad	Pad	Rast	Pad	Rast
ARIMA(5,2,3)	Pad	Pad	Rast	Pad	Rast	Pad	Rast	Rast	Rast	Rast	Rast	Rast	Rast	Rast	Rast	Rast	Rast	Rast	Rast
ARIMA(8,2,2)	Pad	Rast	Rast	Pad	Rast	Pad	Rast	Pad	Rast	Rast	Pad	Rast	Pad	Rast	Rast	Rast	Rast	Rast	Rast
ARIMA(6,0,0)	Pad	Pad	Rast	Pad	Rast	Pad	Rast	Rast	Pad	Rast	Pad	Rast	Pad	Pad	Rast	Pad	Rast	Pad	Rast

Tabela 8: Pravac kretanja tržišta prema broju građevinskih dozvola

U tabeli 9 ispod su navede predikcije modela. Na osnovu vizuelne analize se može zaključiti sledeće: Model ARIMA(4,2,0) generalno predviđa veće vrednosti u odnosu na stvarne, posebno u prvim mesecima. Predviđanja ovog modela su često iznad stvarne vrednosti "Broj dozvola".

Model ARIMA(4,2,4) ima varijacije u predviđanjima, gde su neka predviđanja bliska stvarnim vrednostima, dok su druga dosta različita. Na primer, za jun.22, predviđanje je niže od stvarne vrednosti, dok je za jul.22 znatno više.

Model ARIMA(5,2,3) daje predviđanja koja su generalno bliska stvarnim vrednostima, mada sa nekim odstupanjima. U većini meseci, predviđanja ovog modela su nešto veća od stvarnih vrednosti.

Model ARIMA(8,2,2) ima slične karakteristike kao i ARIMA(5,2,3) model, sa nekim predviđanjima koja su vrlo bliska stvarnim vrednostima, ali sa odstupanjima u određenim mesecima.

Mesec:	jan.22	feb.22	mar.22	apr.22	maj.22	jun.22	jul.22	avg.22	sep.22	okt.22	nov.22	dec.22	jan.23	feb.23	mar.23	apr.23	maj.23	jun.23	jul.23
Broj dozvola:	1489	3341	4157	2097	3537	3496	3296	3959	3624	2836	3153	3387	2836	3012	3510	2500	2224	3361	2935
ARIMA(4,2,0)	3110	2909	2997	3027	3244	3168	3160	3275	3317	3383	3373	3433	3505	3534	3580	3613	3673	3720	3755
ARIMA(4,2,4)	2771	3050	2961	2911	3179	2874	3237	3017	3140	3247	3046	3388	3103	3358	3303	3250	3497	3229	3548
ARIMA(5,2,3)	3036	2907	3032	2921	3092	3029	3105	3113	3138	3183	3191	3238	3252	3289	3313	3342	3372	3397	3428
ARIMA(8,2,2)	2874	2894	3041	2905	3105	3031	3152	3110	3115	3203	3165	3245	3229	3277	3308	3326	3379	3383	3430

Tabela 9: Predikcije broja građevinskih dozvola

4.2 Rezultati primene hedonističkog modela na realnim podacima

Kao što je u prethodnom poglavlju navedeno, uspešnost konstrukcije hedonističkog modela zavisi od kvaliteta podataka. Neophodno je da podaci budu detaljni i da dobro opisuju posmatranu nepokretnost. Ovakva baza podataka nije javno dostupna za teritoriju Republike Srbije. Shodno tome, podaci koji su korišćeni za formiranje hedonističkog modela za procenu cena nepokretnosti su preuzeti sa Kaggle platforme⁶. Podaci su objavljeni u okviru organizovanog takmičenja na platformi. Podaci obuhvataju 545 opservacija i uključuju podatke o ceni kuća, površini, broju spavaćih soba, broju kupatila, broj spratova, postojanje podruma, tople vode, klimatizacije, gostinjske sobe, da li je na glavnoj ulici, da li kuća nameštena, nenameštena ili polunameštena te da li je na preferiranoj lokaciji[23]. Formiranje modela je izvršeno upotrebom programa *R Studio*. Modeliranje je rađeno po uzoru na više radova koji istražuju ovu tematiku[24][25][11].

Kvalitativne karakteristike su predstavljene kao kategorijalne promenljive, što znači da su opisane u obliku kategorija ili grupa. Ove promenljive uključuju prisustvo podruma, grejanja i tople vode, klimatizacije, gostinjske sobe, i poziciju kuće u odnosu na glavnu ulicu. Takođe, opremljenost kuće je opisana kroz dva različita slučaja: polunameštena i nenameštena, što omogućava analitičarima da uzmu u obzir različite nivoe opremljenosti prilikom modeliranja i analize cena nepokretnosti.

Prvi korak u formiranju modela je analiza i grafički prikaz podataka.

Zavisna promenljiva p_i predstavlja cenu $i - te$ kuće

Kvantitativne nezavisne promenljive $i - te$ kuće:

x_{i1} : površina kuće, izražena u kvadratnim stopama (originalni podaci)

x_{i2} : broj spavaćih soba

x_{i3} : broj kupatila

x_{i4} : broj spratova

x_{i5} : broj parking mesta

Kvalitativne karakteristike $i - te$ kuće su posmatrane kao dummy varijable i definišu se na sledeći način:

y_{i1} : da li kuća ima podrum

$$y_{i1} = \begin{cases} 1, & \text{ako kuća ima podrum} \\ 0, & \text{ako kuća nema podrum} \end{cases} \quad (7)$$

y_{i2} : da li kuća ima toplu vodu i grejanje

$$y_{i2} = \begin{cases} 1, & \text{ako kuća ima toplu vodu i grejanje} \\ 0, & \text{ako kuća nema toplu vodu i grejanje} \end{cases} \quad (8)$$

⁶Kaggle je online platforma za učenje mašinskog učenja i analizu podataka. Osnovana je 2010. godine i postala je popularna destinacija za data naučnike, istraživače i profesionalce u oblasti podataka. Platforma pruža niz resursa i alatki koji pomažu korisnicima da uče o mašinskom učenju, razmenjuju ideje, takmiče se u takmičenjima i radu na stvarnim projektima analize podataka.

y_{i3} : da li je kuća klimatizovana

$$y_{i3} = \begin{cases} 1, & \text{ako je kuća klimatizovana} \\ 0, & \text{ako kuća nije klimatizovana} \end{cases} \quad (9)$$

y_{i4} : da li kuća ima gostinsku sobu

$$y_{i4} = \begin{cases} 1, & \text{ako kuća ima gostinsku sobu} \\ 0, & \text{ako kuća nema gostinsku sobu} \end{cases} \quad (10)$$

y_{i5} : da li se kuća nalazi na glavnoj ulici

$$y_{i5} = \begin{cases} 1, & \text{ako se kuća nalazi na glavnoj ulici} \\ 0, & \text{ako se kuća ne nalazi na glavnoj ulici} \end{cases} \quad (11)$$

Kvalitativna karakteristika koja se odnosi na opremljenost i -te kuće je posmatrana pomoću dve dummy varijable:

y_{i6} : da li je kuća nenameštena

$$y_{i6} = \begin{cases} 1, & \text{ako je kuća nenameštena} \\ 0, & \text{ako je kuća nameštena} \end{cases} \quad (12)$$

y_{i7} : da li je kuća polunameštena

$$y_{i7} = \begin{cases} 1, & \text{ako je kuća polunameštena} \\ 0, & \text{ako kuća nije polunameštena} \end{cases} \quad (13)$$

Posmatrano je ukupno 5 kvantitativnih obeležja i ukupno 7 kvalitativnih obeležja.

Na osnovu tabele 10 koja sadrži deskriptivni opis podataka možemo uvideti da medijalna kuća ima 3 spavaće sobe, 1 kupatilo, 2 sprata i nema parking mesto. Površina ovakve kuće iznosi 4600 sqft i njena cena je 4340000\$.

Metric	Cena	Površina	Broj spavaćih soba	Broj kupatila	Broj spratova	Parking
Min.	1750000	1650	1	1	1	0
1st Qu.	3430000	3600	2	1	1	0
Median	4340000	4600	3	1	2	0
Mean	4766729	5151	2,965	1,286	1,806	0,6936
3rd Qu.	5740000	6360	3	2	2	1
Max.	13300000	716200	6	4	4	3

Tabela 10: Deskriptivni opis kvantitativnih podataka

Na sledećem prikazu (tabela 11) možemo da vidimo koliko su u procentualnom iznosu zastupljene promenljive glavna ulica, gostinska soba, podrum, topla voda i grejanje, klimatizacija u odnosu na nivo opremljenosti kuća. Većina kuća iz posmatranog skupa se nalazi na glavnoj ulici. Manje od 50% kuća ima podrum i klimatizaciju. Takođe gostinsku sobu poseduje samo 18% kuća.

	Udeo	Glavna ulica	Gostinska soba	Podrum	Topla voda i grejanje	Klimatizacija
Polunameštena	42%	36%	8%	16%	3%	12%
Nenameštena	33%	26%	4%	9%	1%	8%
Nameštena	26%	24%	6%	10%	1%	11%
Ukupno	100%	86%	18%	35%	5%	32%

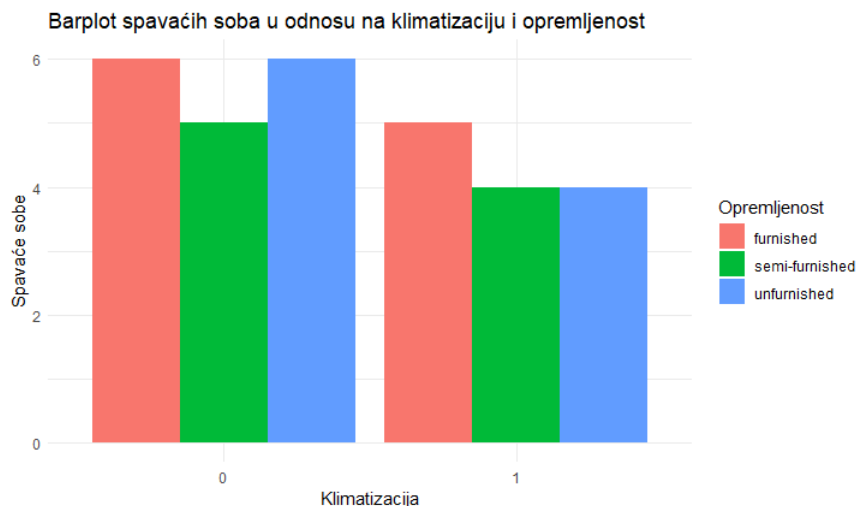
Tabela 11: Procentualne zastupljenosti kvalitativnih osobina u odnosu na nivo opremljenosti

Bar plot, koji je prikazan na frafiku 6, nam ukazuje kako se kreće broj kupatila u odnosu na opremljenost kuće i da li kuća ima toplu vodu i grejanje. Na primer, možemo uvesti da kuće koje imaju četiri kupatila su opremljene, ali nemaju toplu vodu i grejanje.



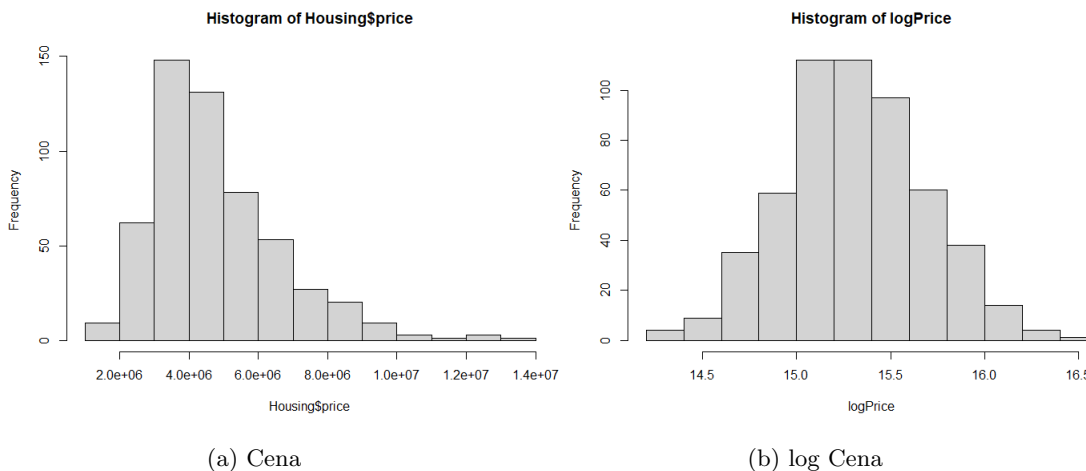
Grafik 6: Bar plot kupatila u odnosu na toplu vodi i grejanje i opremljenost

Bar plot, koji je prikazan na grafiku 7, nam ukazuje kako se kreće broj spavaćih soba u odnosu na opremljenost kuće i da li je kuća klimatizovana. Tako na primer, kuće koje imaju šest soba i koje su nameštene ili nenameštene nisu klimatizovane.



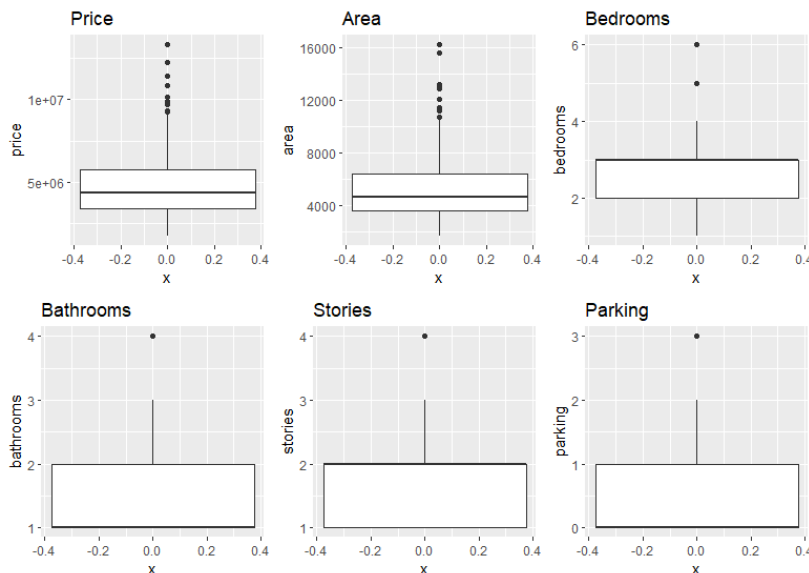
Grafik 7: Bar plot spavaćih soba u odnosu na klimatizaciju i opremljenost

Prikazom cene preko histograma (grafik 8) uočena je asimetrija varijable u desno, što ukazuje da zavisna promenljiva Cena nema normalnu raspodelu. Kako bi se poboljšala raspodela promenljive, posmatrana je logaritamska cena. Vizuelnim posmatranjem histograma raspodela ima normalan zvonasti oblik što ukazuje na normalnu raspodelu. Ova zapažanja su potvrđena primenom Dickey-Fuller. Za $p < 0.05$ odbacujemo hipotezu o normalnosti, odnosno za $p > 0.05$ nemamo razloga da odbacimo nultu hipotezu.



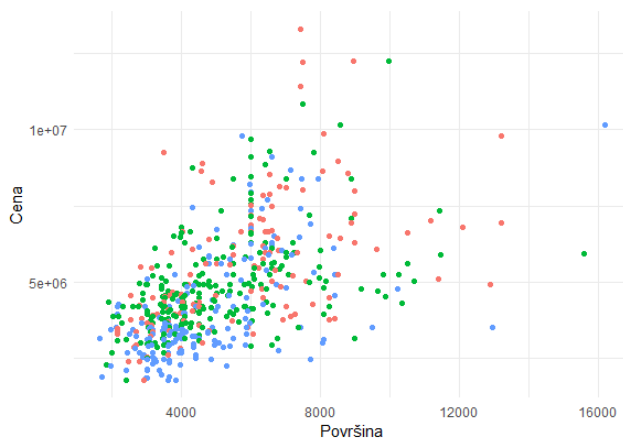
Grafik 8: Histogram

Na Box plot prikazu (grafik 9) možemo da vidimo distribuciju podataka. Dno i vrh kutije predstavljaju prvi i treći kvartil skupa podataka, kao i vrednost mediane za svaku promenljivu. Brkovi su linije koje se protežu od vrha i dna kutije ukazujući na varijabilnost izvan gornjeg i donjeg kvartila. Individualne tačke udaljene od ostalih podataka jesu outlier-i i oni su na Box plot prikazu izvan kutije i udaljeni od brkova. Najveći broj kuća iz seta podataka ima 3 spavaće sobe. Broj kupatila je najčešće 1. Većina kuća nema parking, dok je broj spratova 2. Možemo uvideti da na posmatranom skupu, kuće sa 4 i 5 soba, kuće sa tri parkinga, kuće sa 4 sprata i 4 kupatila predstavljaju outlier-e.



Grafik 9: Box plot

Na tačkastom grafikonu (grafik 10) možemo videti da kuće površine preko 12.000 sqft i kuće sa cenom većom od 10.000.000 su udaljene od ostalih podataka. Narandžastom bojom su prikazane nameštene kuće, zelenom bojom su prikazane polunameštene kuće, a plavom bojom su prikazane nenameštene kuće.



Grafik 10: Tačkasti grafikon cene i površine

Korelaciona matrica se može predstaviti na sledeći način (tabela 12) što olakšava vizuelnu identifikaciju korelacije između nezavisnih promenljivih. Posmatranjem korelacione matrice utvrđujemo koliko iznose koeficijenti korelacije između nezavisnih promenljivih i kolika je jačina njihove povezanosti.

Površina, broj spavaćih soba, broj kupatila, broj spratova i broj parking mesta predstavljaju promenljive koje imaju pozitivan uticaj na cenu. Najveći koeficijent korelacije je između cene i površine i iznosi 0.54. Zatim koeficijent korelacije između cene i broja kupatila je 0.49. Broj spratova i cena su povezane koeficijentom 0.42, dok najmanji uticaj na cenu imaju broj spavaćih soba i broj parking mesta.

Možemo uočiti da je povećan nivo korelacije između nekih nezavisnih promenljivih: Spavaće sobe i Broj kupatila koeficijent iznosi 0.37, Spavaće sobe i Broj spratova koeficijent je 0.41, dok koeficijent korelacije za Broj kupatila i Broj spratova iznosi 0.33. Povezanost ovih varijabli je logična. Više spratova će uključivati više soba. Sa druge strane, posmatranjem kupatila, ima smisla ako računamo da se na svakom spratu nalazi po jedno ili da svaka soba ima svoje kupatilo. Takođe korelacija je povišena i između Površina Parking i iznosi 0.35.

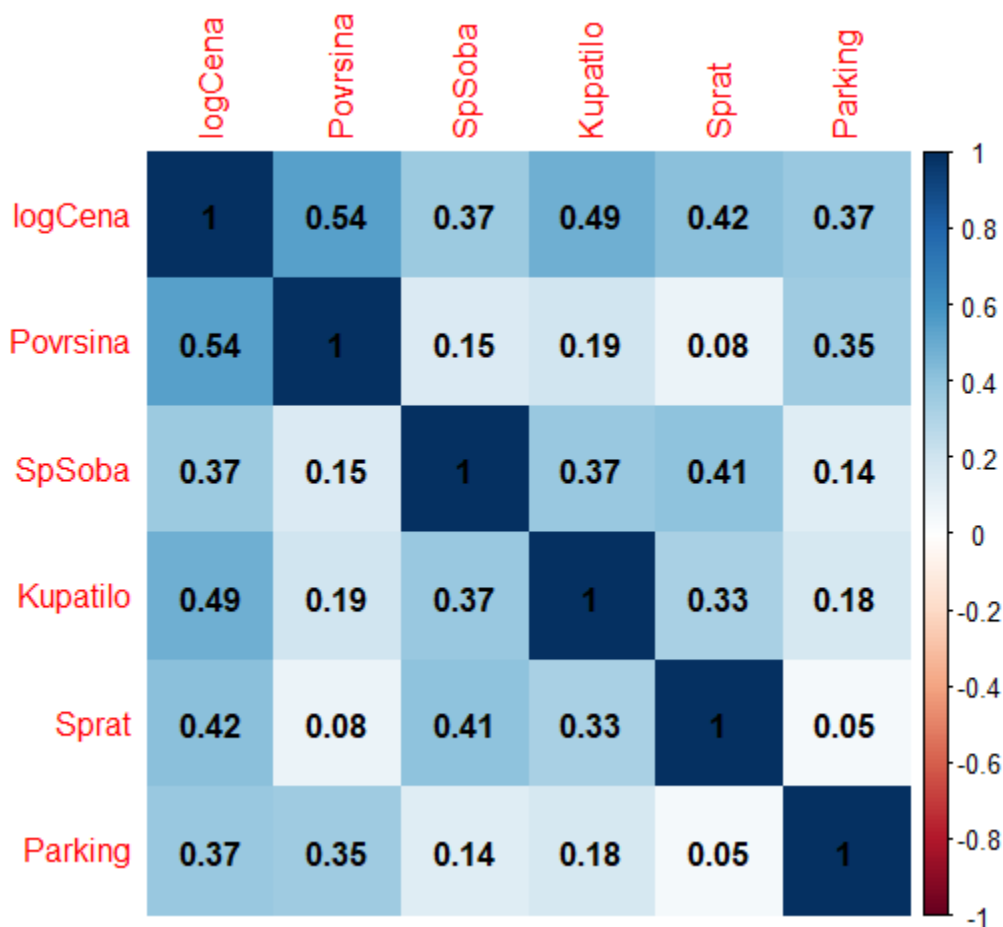


Tabela 12: Korelaciona matrica

Modeli su formirani na istom broju opservacija. Ukupan broj opservacija u posmatranom skupu podataka je bio 545. Nakon uklanjanja outliera, posmatrani skup je obuhvatio 531 opservaciju. Podaci su podeljeni na dva podskupa: trening skup i testing skup u odnosu 80:20. Trening skup čini 80% slučajno odabranih podataka i on je korišćen za formiranje, odnosno treniranje modela. Preostalih 20% opservacija čini testing skup na kojem je izvršena evaluacija modela. Formirano je 8 različitih modela koji uključuju različite nezavisne varijable. U prikazanoj tabeli 13 Možemo da vidimo da je kod svakog modela navedeno koliko iznosi procenjeni koeficijent i koliko iznosi p vrednost test statistike. Ova vrednost nam govori da li je određeni procenjeni koeficijent za posmatranu promenljivu statistički značajno različit od nule, odnosno da li je potrebno odrađenu promenljivu uključiti u model.

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
Odsečak	14370 < 2e-16 ***	14420 < 2e-16 ***	1220466 < 2e-16 ***	14390 < 2e-16 ***	1226124 < 2e-16 ***	1220466 < 2e-16 ***	1208687 < 2e-16 ***	12235316 < 2e-16 ***
Površina	0,05 < 2e-16 ***	0,05 < 2e-16 ***	0.29394 < 2e-16 ***	0,05 < 2e-16 ***				
Log Površina					0.28752 < 2e-16 ***	0.29394 < 2e-16 ***	0.30409 < 2e-16 ***	0.308320 < 2e-16 ***
Broj spavaćih soba	15,62 0.308163			53,60 0.000381 ***			0.05653 0.000144 ***	
Broj kupatila	179,80 8.08e-15 ***	185,60 2.31e-16 ***	0.18437 < 2e-16 ***	208,40 < 2e-16 ***	0.18326 < 2e-16 ***	0.18437 < 2e-16 ***	0.20719 < 2e-16 ***	
Broj spratova	87,38 2.98e-10 ***	92,40 1.07e-12 ***	0.09346 2.47e-13 ***		0.09180 6.49e-13 ***	0.09346 2.47e-13 ***		
Parking	45,65 0.000270 ***	46,01 0.000219 ***	0.04111 0.000789 ***	34,73 0.007707 **	0.04270 0.000498 ***	0.04111 0.000789 ***		
Glavna ulica	128,30 1.24e-05 ***	122,60 2.31e-05 ***	0.10566 0.000240 ***	159,90 1.85e-07 ***	0.10357 0.000311 ***	0.10566 0.000240 ***	0.14658 1.43e-06 ***	0.138076 3.84e-06 ***
Gostinska soba	63,10 0.025384 *	61,86 0.028340 *			0.04669 0.093974 .			
Podrum	96,44 2.92e-05 ***	99,01 1.61e-05 ***	0.11519 9.37e-08 ***	77,46 0.000512 ***	0.10314 5.19e-06 ***	0.11519 9.37e-08 ***	0.07801 0.000395 ***	0.106336 1.78e-06 ***
Topla voda i grejanje	192,40 0.000239 ***	194,70 0.000201 ***	0.19817 0.000124 ***	198,60 0.000316 ***	0.20046 0.000101 ***	0.19817 0.000124 ***	0.21086 0.000107 ***	0.188679 0.000425 ***
Klima	151,80 2.72e-11 ***	150,10 3.98e-11 ***	0.14035 4.01e-10 ***	186,70 2.11e-15 ***	0.13823 7.01e-10 ***	0.14035 4.01e-10 ***	0.17694 2.83e-14 ***	0.158428 8.20e-12 ***
Preferirana lokacija	125,80 3.83e-07 ***	126,60 3.13e-07 ***	0.13438 3.21e-08 ***	133,60 2.91e-07 ***	0.13268 4.54e-08 ***	0.13438 3.21e-08 ***	0.13773 7.72e-08 ***	0.133639 1.08e-07 ***
Nenamešten	-112,10 2.66e-05 ***	-129,00 2.12e-09 ***	-0.12933 1.14e-09 ***	-138,40 1.02e-09 ***	-0.12839 1.39e-09 ***	-0.12933 1.14e-09 ***	-0.14287 1.47e-10 ***	-0.141846 1.00e-10 ***
Polu-nameštena	26,38 0.281097							
Spavaća soba: Sprat								0.017615 2.06e-06 ***
Spavaća soba: Kupatilo								0.039229 1.64e-11 ***
					Nivo značajnosti: $p < 0.001$ *** $p < 0.01$ ** $p < 0.05$ * $p < 0.1$.			
R ²	72,04%	70,08%	70,85%	66,72%	71,04%	70,85%	67,49%	68,56%
Adj R ²	69,32%	69,30%	70,16%	65,94%	70,29%	70,16%	66,81%	67,90%
AIC	-150,0542	-151,6689	-165,0533	-107,3521	-165,9411	-165,0533	-119,5594	-134,1369
BIC	-88,88954	-98,65957	-116,1216	-58,42034	-112,9317	-116,1216	-74,70533	-89,2828

Tabela 13: Modeli

Model 1 je inicijalno formiran na osnovu svih nezavisnih promenljivih. Možemo uvideti da koeficijenti koji stoje uz promenljive broj spavaćih soba i polunameštena nisu statistički značajno različiti od 0 što implicira da ove promenljive možemo da izbacimo iz modela. Koeficijent koji stoji uz promenljivu gostinska soba je statistički značajan na nivou 5%.

Primenom Anderson-Darling testa za testiranje normalnosti zaključeno je da reziduali modela 1 ne prate normalnu raspodelu, vrednost p iznosi 0.007 što je ispod praga značajnosti od 5%. Takođe sprovedeni su testovi za proveru homoskedastičnosti (Goldfeld-Quandt) i test za proveru autokorelacije (Durbin-Watson). Ovim testovima je potvrđeno da model ne pati od heteroskedastičnosti i da nema autokorelacije među rezidualima. Vrednost VIF iznosi 3.36 što je manje od granice 5 pa time zaključujemo da nema multikolinearnosti. Outlier testom koji se zasniva na Studentovoj t -raspodeli potvrđeno je odsustvo outlier-a.

Model 2 se u odnosu na model 1 razlikuje time što ne uključuje promenljivu broj spavaćih soba i promenljivu polunameštena jer koeficijenti uz ove promenljive u modelu 1 nisu doprinosili formiranju cene. Koeficijent koji stoji uz promenljivu gostinska soba je statistički značajan na nivou 5%. Primena testova je sprovedena na modelu 2 i dala je jednake rezultate kao za model 1, reziduali ne prate normalnu raspodelu, ali nema prisustva heteroskedastičnosti i multikolinearnosti. Kod ovog modela VIF iznosi 3.34.

Model 3 je formiran tako što je u odnosu na model 2 izbačena promenljiva gostinska soba. U ovom modelu svi koeficijenti su statistički značajni na nivou od 0.1%. Sprovedenjem testa o normalnosti reziduala utvrđeno je da reziduali ne prate normalnu raspodelu, ali su homoskedastični, te je utvrđeno da nisu autokorelisani. VIF kod ovog modela iznosi 3.30.

Model 4 je formiran tako da u odnosu na model 3 uključuje nezavisnu promenljivu broj spavaćih soba, ali ne uključuje nezavisnu promenljivu broj spratova. Koeficijent uz promenljivu parking je statistički značajan na nivou od 5%. Primenom testa o normalnosti, utvrđeno je da reziduali prate normalnu raspodelu jer je $p = 0.1134$ što je veće od praga značajnosti od 5%. Takođe, $p = 0.1247 > 0.05$ prilikom testiranja homoskedastičnosti, što znači da reziduali imaju konstantnu varijansu. Pored toga $p = 0.5814 > 0.05$ prilikom sprovođenja testa o autokorelaciji. Ovim potvrđujemo da su sve pretpostavke modela ispunjene. Kod modela 4, vrednost VIF=3, što je manje od granice 5. Takođe, potvrđeno je odsustvo outlier-a.

Zatim je posmatrana logaritamska površina kao nezavisna promenljiva. Formirana su četiri modela sa ovom nezavisnom promenljivom.

Model 5 je formiran sa nezavisnom promenljivom logaritamska površina. Koeficijent uz promenljivu gostinska soba je statistički značajan na nivou od 10%. Ostali koeficijenti su statistički značajni na nivou od 0.1%. Kod ovog modela reziduali ne prate normalnu raspodelu ($p = 0.01184 < 0.05$) i nisu homoskedastični ($p = 0.03366 < 0.05$). Reziduali nisu autokorelisani. VIF=3.45 kod modela 5.

Kako je u modelu 5 koeficijent uz promenljivu gostinska soba statistički značajan na nivou od 10%, u model 6 ova promenljiva nije uključena. Reziduali kod ovog modela su homoskedastični ($p = 0.06759 > 0.05$) i nisu autokorelisani ($p = 0.5257 > 0.05$). Reziduali ne prate normalnu raspodelu ($p = 0.0182 < 0.05$). Kod modela 6, VIF=3.43.

U model 7 je dodata promenljiva spavaća soba, ali je izbačena promenljiva broj spratova i parking, za razliku od modela 6. Svi koeficijenti su statistički značajno različiti od 0 na nivou značajnosti 0.1%. U modelu 7, reziduali prate normalnu raspodelu ($p = 0.2478 > 0.05$), reziduali su homoskedastični ($p = 0.08138 > 0.05$) i nisu autokorelisani ($p = 0.2656 > 0.05$). Vrednost VIF iznosi 3.08 kod modela 7.

Model 8 pored logaritamske površine uključuje promenljivu formiranu kao kombinacija broja spavaćih soba i broja spratova, kao i promenljivu koja predstavlja kombinaciju promenljivih broj spavaćih soba i broj kupatila. Svi koeficijenti u ovom modelu su statistički značajno različiti od 0 na nivou značajnosti 0.1%. Model 8 ima normalno distribuirane rezidualne, reziduali su homoskedastični i ne postoji autokorelacija između njih. Takođe nema potvrde o prisustvu outlier-a. Vrednost VIF za model 8 iznosi 2.86.

Ako posmatramo koliko iznosi vrednost koeficijenta R^2 , najviši je kod modela 1 i 5. Međutim, model 1 uključuju i one nezavisne promenljive koje ne doprinose formiranju cene i reziduali ovih modela ne prate normalnu raspodelu. Vrednost kriterijuma AIC je najmanja kod modela 5, dok je vrednost BIC najmanja kod modela 3 i 6.

Nakon sprovedenih testova, zaključujemo da jedino modeli 4, 7 i 8 ispunjavaju sve pretpostavke o rezidualima. Imajući to u vidu za formiranje predikcija su korišćeni upravo ti modeli.

Predikcije modela su sprovedene na testing skupu. Rezultati predikcija su prikazani u tabeli 14 ispod. Možemo videti da su vrednosti R^2 i $AdjR^2$ najviše kod modela 8. Takođe prikazane su srednja apsolutna greška, srednja kvadratna greška i koren srednje kvadratne greške. Iznosi grešaka su najmanji kod modela 8.

Model	R-squared	Adj R-squared	MAE	MSE	RMSE
Model 4	64,28%	63,95%	0,1945479	0,06178581	0,2485675
Model 7	64,57%	64,24%	0,1921719	0,06059674	0,246164
Model 8	68,07%	67,77%	0,1816375	0,05549878	0,2355818

Tabela 14: Evaluacija modela

U tabeli 15 su prikazani parametri modela 8 koji predstavljaju nezavisne promenljive pri određivanju logaritamske cene:

Nezavisna promenljiva	Parametri
Odsečak	12235316
Log Površina	0.308320
Glavna ulica	0.138076
Podrum	0.106336
Topla voda i grejanje	0.188679
Klima	0.158428
Preferirana lokacija	0.133639
Nenamešten	-0.141846
Spavaća soba: Sprat	0.017615
Spavaća soba: Kupatilo	0.039229

Tabela 15: Parametri modela 8

Možemo da vidimo da elastičnost između površine i cene nije jednaka 1 što ukazuje da kuća koja je na primer duplo veća neće koštati duplo više. Ukoliko se kuća nalazi na glavnoj ulici, odnosno ako je vrednost indikator varijable, ona ima pozitivan uticaj na cenu, što znači da su kuće na glavnim ulicama skuplje. Pozitivan uticaj imaju i druge indikator promenljive. Tako da postojanje podruma u kući je plus, postojanje tople vode i grejanja takođe pozitivno utiče na nivo cene. Kuće koje su klimatizovane koštaju više jer tada ne iziskuju ulaganje u klimatizaciju. Ukoliko je vrednost indikator promenljive preferirana lokacija 1, ona će imati pozitivan uticaj na cenu. Koeficijent uz indikator varijablu nenameštena je negativan što znači da ukoliko kuća nije nameštena cena će biti manja. Kombinovana promenljiva broj spavaćih soba po spratu ima pozitivan uticaj, što ukazuje da ukoliko broj soba po spratu veći, cena kuće će biti viša. Nasuprot tome, kuće koje imaju više spratova biće jeftinije u odnosu na one koje imaju isti broj soba na manje spratova. Kombinovana promenljiva broj spavaćih soba po kupatilu ima pozitivan uticaj. Istu logiku možemo primeniti i na odnos između broja soba i kupatila. Ovo bi ukazivalo da je broj soba važniji faktor za povećanje cene od broja kupatila.

Zaključak

Tržište nekretnina je složena i dinamična sfera koja odražava mnoge aspekte ekonomije. Uz stalne promene u ekonomskom okruženju, faktori poput kamatnih stopa, inflacije, urbanizacije i čak globalnih trendova mogu imati značajan uticaj na tržišne trendove i vrednost nekretnina. Upravo iz tog razloga, primena sofisticiranih modela, poput ARIMA i hedonističkog modela, postaje od suštinske važnosti. U okviru ovog master rada, cilj je bio analizirati tržište nekretnina primenjujući dve različite metode - ARIMA i hedonistički model. Dok je hedonistički model bio fokusiran na karakteristike kuća kao što su površina, broj soba, postojanje podruma i druge, sa ciljem da precizno odredi njihovu tržišnu vrednost, ARIMA je pružila analitički okvir za procenu i predviđanje tržišnih trendova baziranih na vremenskim serijama prosečnih cena stanova po kvadratnom metru. Modeliranje je izvršeno korišćenjem R Studio programa.

Rezultati ARIMA modela omogućavaju dublje razumevanje dinamike tržišta nekretnina u posmatranom periodu. Analiza, koja je rađena na podacima prikupljenim od strane Republičkog zavoda za statistiku[22], je pokazala da tržište nekretnina ima sezonske oscilacije, što je ključno za investitore i agente za nekretnine, jer razumevanje sezonskih varijacija može biti od presudnog značaja za planiranje i donošenje odluka. Na primer, dekompozicija vremenske serije cena otkrila je sezonski karakter, što može uticati na strategije ulaganja i prodaje. ARIMA model, posebno ARIMA(3,2,0)(1,1,0), pokazao se kao najefikasniji za analizu dostupnih podataka. Diferenciranje vremenske serije tri puta omogućilo je postizanje stacionarnosti, a značajne autokorelacione funkcije pružile su uvide o povezanosti različitih vremenskih tačaka. Ovo je od suštinske važnosti jer prepoznavanje ovih veza može pružiti ključne informacije za predviđanje budućih tržišnih kretanja. Ovo može biti od značajne vrednosti za razvojne strategije, planiranje budžeta i čak za makroekonomske analize koje se bave projekcijama rasta.

U okviru hedonističkog modeliranja, konstruisani su različiti modeli čija je primarna svrha procena vrednosti kuća. Izvorni podaci koji su poslužili kao temelj za izradu ovih modela su dobijeni sa Kaggle platforme[23]. Za modeliranje, koje je takođe vršeno u R Studio programu, korišćene su različite kombinacije nezavisnih promenljivih. Najrelevantnijim se pokazao model 8. Za njegovu konstrukciju su korišćene sledeće nezavisne promenljive: logaritamska površina, podatak o tome da li se kuća nalazi na glavnoj ulici, da li kuća poseduje podrum, da li ima toplu vodu i grejanje, da li je klimatizovana, da li se nalazi na preferiranoj lokaciji, podatak o tome da li je kuća nameštena, odnosno nenameštena kao i odnos promenljivih broja spavaćih soba i broja spratova i odnos promenljivih broja spavaćih soba i broja kupatila. Analiza parametara ovog modela otkriva da logaritamska površina ima dominantan uticaj na određivanje cene kuće. Interesantno je da ovaj parametar nije jednak 1, što upućuje na zaključak da elastičnost između površine i cene nije jedinična. U praktičnom smislu, to implicira da kuća dvostruko veće površine neće imati dvostruko veću cenu. Ovaj fenomen može se dalje istražiti kroz detaljniju analizu, koja može uključivati i faktore kao što su kvalitet materijala, starost objekta, i druge specifične karakteristike koje mogu imati uticaj na tržišnu vrednost kuće.

Kada se kombinuju nalazi iz oba pristupa, jasno je da hedonistički model pruža detaljan uvid u to kako specifične karakteristike nekretnine utiču na njenu cenu, dok ARIMA model omogućava razumevanje i predviđanje širih tržišnih trendova. U svetlu ovih rezultata, može se zaključiti da su ARIMA i hedonistički modeli neophodni alati za sveobuhvatno razumevanje tržišta nekretnina. Ove analize i modeli ne samo da omogućavaju bolje razumevanje prošlih trendova, već su i ključni za anticipiranje budućih promena, što može biti značajno za investitore, analitičare i sve one koji su aktivno uključeni u tržište nekretnina.

Literatura

- [1] Abraham H Maslow. “A theory of human motivation”. **in:** *Psychological Review* 50.4 (1943), **pages** 370–396.
- [2] *ZAKON O PROMETU NEPOKRETNOSTI*. Službeni glasnik RS, br. 93/2014, 12/2014 i 6/2015. 2015. URL: https://www.paragraf.rs/propisi/zakon_o_prometu_nepokretnosti.html (**urlseen** 15/10/2023).
- [3] Nicholas G Pirounakis. *Real Estate Economics: A Point-to-Point Handbook*. Routledge, 2012.
- [4] Donald Rapp. “Bubbles, Booms, and Busts”. **in:** Springer New York, 2009. **chapter** A Short History Of Booms, Bubbles, And Busts.
- [5] Crowe Christopher **and** others. “How to deal with real estate booms: Lessons from country experiences”. **in:** *Journal of Financial Stability* 9.3 (2013), **pages** 300–319.
- [6] B. Marjanović **and** R. Pervan. “POSTUPAK I ZNAČAJ SEGMENTACIJE TRŽIŠTA NA PRIMJERU TRŽIŠTA NEKRETNINA”. **in:** *Zbornik radova Veleučilišta u Šibeniku* 14.1-2 (2020), **pages** 35–43. URL: <https://hrcak.srce.hr/241609> (**urlseen** 15/10/2023).
- [7] Odeljenje za procenu i vođenje nepokretnosti. *Izveštaj o stanju na tržištu nepokretnosti za prvo polugodište 2022. godine*. techreport. Preuzeto sa <https://www.rgz.gov.rs/>. Republički geodetski zavod, 2022.
- [8] Zlatko Kovačić. *Analiza Vremenskih Serija*. Belgrade: Ekonomski fakultet, Univerzitet u Beogradu, 1995.
- [9] Nataša Krklec Jerinkić. “Vremenske serije”. Beleške sa predavanja. 2018.
- [10] Nathalie Picard, Constantinos Antoniou **and** André de Palma. “Econometric Models”. **in:** *SustainCity Working Paper* 2.4 (2010).
- [11] Ben J. Sopranzetti. “Hedonic Regression Models”. **in:** *Handbook of Financial Econometrics and Statistics*. Springer New York, 2015, **pages** 2119–2134.
- [12] Allen Goodman. “Andrew Court and the Invention of Hedonic Price Analysis”. **in:** *Journal of Urban Economics* 44 (1998), **pages** 291–298.
- [13] Hendler Reuven. “Lancaster’s New Approach to Consumer Demand and Its Limitations”. **in:** *The American Economic Review* 65 (1975), **pages** 194–199. URL: <http://www.jstor.org/stable/1806408> (**urlseen** 15/10/2023).
- [14] S. Rosen. “Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition”. **in:** *Journal of Political Economy* 82 (1974), **pages** 34–55.
- [15] Patrick Bajari **and** Matthew E Kahn. *Estimating Housing Demand with an Application to Explaining Racial Segregation in Cities*. Working Paper 9891. National Bureau of Economic Research, 2003.
- [16] A. M. Freeman. “Hedonic Prices, Property Values, and Measuring Environmental Benefits: A Survey of the Issues”. **in:** *Scandinavian Journal of Economics* 81 (1979), **pages** 154–171.
- [17] OECD **and** others. “Hedonic Regression Methods”. **in:** *Handbook on Residential Property Price Indices*. Luxembourg: Publications Office of the European Union, 2013, **pages** 50–64.
- [18] Zorana Lužanin. “Ekonometrija”. Beleške sa predavanja. 2018.
- [19] Dorota Dejniał **and** Janusz Dabrowski. “Application of ARIMA models in real estate market forecasting”. **in:** *GIS ODYSSEY*. 2017.
- [20] Arvydas Jadevicius **and** Simon Huston. “Property market modelling and forecasting: Simple vs complex models”. **in:** *Journal of Property Investment Finance* 33 (2015), **pages** 337–361.
- [21] Simon Stevenson. “A comparison of the forecasting ability of ARIMA models”. **in:** *Journal of Property Investment & Finance* 25.3 (2007), **pages** 223–240.

- [22] Republički zavod za statistiku Srbije. *Statistički podaci Republičkog zavoda za statistiku Srbije*. Pristupljeno 15. oktobra 2023. 2023. URL: <https://data.stat.gov.rs/?caller=SDDB&languageCode=sr-Latn>.
- [23] Yasser Hamed. *Housing Prices Dataset*. Pristupljeno 15. oktobra 2023. 2021. URL: <https://www.kaggle.com/datasets/yasserh/housing-prices-dataset/data>.
- [24] Kunovac Davor **and others**. “Primjena hedonističke metode za izračunavanje indeksa cijena nekretnina u Hrvatskoj”. **in:** *Hrvatska narodna banka* (2008). ISSN: 1332-1900. URL: <https://www.hnb.hr/documents/20182/121648/i-020.pdf/67a39d10-1fee-4447-9121-69a1f4a21f13>.
- [25] Zakaria Firano **and** Fatine Filali A. “Towards the hedonic modelling and determinants of real estates price in Morocco”. **in:** *Social Sciences Humanities Open* 4 (2021). URL: <https://www.sciencedirect.com/science/article/pii/S2590291121000723>.

Biografija

Silvija Sokol, rođena 25. maja 1995. godine u Beogradu. Osnovnu školu "Šremski front" u Šidu je završila 2009. godine. Srednju tehničku školu "Nikola Tesla", smer poljoprivredni tehničar je završila u Šidu 2013. godine. Po završetku srednje škole započela je studije na Prirodno - matematičkom fakultetu u Novom Sadu, na Departmanu za matematiku i informatiku, smer Primenjena matematika - matematika finansija i uspešno ih završila 2017. godine sa prosečnom ocenom 8.23. Iste godine upisuje master studije Primenjene matematike - matematika finansija. Zaključno sa julskim rokom 2020. godine položila je sve ispite predviđene planom i programom sa prosečnom ocenom 8.67. Karijeru je započela kao procenitelj vrednosti nepokretnosti u kompaniji NAI WMG gde je stekla bolji uvid u dinamiku tržišta nekretnina. Karijerni put je dalje vodi na trenutnu poziciju, Analitičar za investiranje u kompaniji DDOR osiguranje.



UNIVERZITET U NOVOM SADU
PRIRODNO - MATEMATIČKI FAKULTET
KLJUČNA DOKUMENTACIJSKA INFORMACIJA

Redni broj:

RBR

Identifikacioni broj:

IBR

Tip dokumentacije: Monografska dokumentacija

TD

Tip zapisa: Tekstualni štampani materijal

TZ

Vrsta rada: Master rad

VR

Autor: Silvija Sokol

AU

Mentor: dr Zorana Lužanin

MN

Naslov rada: Matematički modeli za analizu tržišta nepokretnosti

NR

Jezik publikacije: srpski(latinica)

JP

Jezik izvoda: srpski/engleski

JI

Zemlja publikovanja: Srbija

ZP

Uže geografsko područje: Vojvodina

UGP

Godina: 2023

GO

Izdavač: Autorski reprint

IZ

Mesto i adresa: Novi Sad, Departman za matematiku i informatiku, Prirodno-matematički fakultet, Trg
Dositeja Obradovića 4

MA

Fizički opis rada: (4/54/0/15/1/10/0) (broj poglavlja/strana/lit. citata/tabela/slika/grafika/priloga)

FO

Naučna oblast: Matematika

NO

Naučna disciplina: Primenjena matematika

ND

Predmetna odrednica/Ključne reči: tržište nekretnina, cene nekretnina, vremenske serije,
hedonistička regresija

PO

UDK:

Čuva se: Biblioteka Departmana za matematiku i informatiku Prirodno-matematičkog fakulteta u
Novom Sadu

ČU

Važna napomena:**VN**

Izvod: Cilj rada je evaluacija i poređenje efikasnosti ARIMA i hedonističkog modela u analizi i predviđanju dinamike tržišta nekretnina. Kroz empirijsko istraživanje, rad će pokušati da odgovori na pitanja o preciznosti, pouzdanosti i praktičnosti ovih modela u realnim tržišnim uslovima. Osim toga, cilj je i demonstracija primene i adaptacije ovih modela u R Studio okruženju, istraživanje njihove fleksibilnosti i mogućnosti za doprinos boljem razumevanju i analizi tržišta nekretnina.

IZ

Datum prihvatanja teme od strane NN veća: 05.10.2023.

DP

Datum odbrane: 27.10.2023

DO**Članovi komisije:**

Predsednik: dr Andreja Tepavčević, redovni profesor, Prirodno - matematički fakultet, Univerzitet u Novom Sadu

Mentor: dr Zorana Lužanin, redovni profesor, Prirodno - matematički fakultet, Univerzitet u Novom Sadu

Član: dr Sanja Rapajić, redovni profesor, Prirodno - matematički fakultet, Univerzitet u Novom Sadu

KO

UNIVERSITY OF NOVI SAD
FACULTY OF SCIENCE AND MATHEMATICS
KEY WORD DOCUMENTATION

Accession number:

ANO

Identification number:

INO

Document type: Monograph type

DT

Type of record: Printed text

TR

Contents code: Master's thesis

CC

Author: Silvija Sokol

AU

Mentor: Zorana Lužanin, PhD

MN

Title: Mathematical models for real estate market analysis

TI

Language of text: Serbian(latin)

LT

Language of abstract: Serbian/English

LA

Country of publication: Serbia

CP

Locality of publication: Vojvodina

LP

Publication year: 2023

PY

Publisher: Author's reprint

PU

Publication place: Novi Sad, Department of Mathematics and Informatics, Faculty of Science and Mathematics, University of Novi Sad, Trg Dositeja Obradovica 4

PP

Physical description: (4/54/0/15/1/10/0)

(chapters/pages/literature/tables/pictures/graphics/appendices)

PD

Scientific field: Mathematics

SF

Scientific discipline: Applied mathematics

SD

Subject/Key words: real estate market, real estate price, time series, hedonic regression

SKW

UC

Holding data: The Library of the Department of Mathematics and Informatics, Faculty of Science and Mathematics, University of Novi Sad

HD**Note:****N**

Abstract: The aim of this study is to evaluate and compare the effectiveness of the ARIMA and hedonic models in analyzing and predicting the dynamics of the real estate market. Through empirical research, the study will seek to address questions about the accuracy, reliability, and practicality of these models in real market conditions. Additionally, the goal is to demonstrate the application and adaptation of these models in the R Studio environment, explore their flexibility, and their potential contribution to a better understanding and analysis of the real estate market.

AB**Accepted by Scientific Board on:** October 5th 2023**ASB****Defended:** October 27th 2023**DE****Thesis defend board:**

President: Andreja Tepavcevic, PhD, full professor, Faculty of Science and Mathematics, University of Novi Sad

Mentor: Zorana Luzanin, PhD, full professor, Faculty of Science and Mathematics, University of Novi Sad

Member: Sanja Rapajic, PhD, full professor, Faculty of Science and Mathematics, University of Novi Sad

DB