

Analiza čimbenika rizika za srčane bolesti

Motivacija i opis problema

Podaci o pacijentima prikupljeni su s ciljem istraživanja čimbenika koji utječu na pojavu srčanih bolesti te unapređenja prevencije i liječenja. Razumijevanje ovih čimbenika omogućava bolje donošenje odluka u medicinskoj praksi, s ciljem smanjenja rizika od obolijevanja i poboljšanja zdravstvenih ishoda. Ova analiza pruža uvid u ključne medicinske parametre koji su povezani s pojavom srčanih bolesti.

Opis skupa podataka i smjernice za izradu projekta

Skup podataka obuhvaća informacije o demografskim karakteristikama pacijenata, njihovom zdravstvenom stanju te rezultatima ključnih medicinskih pretraga. Podaci uključuju varijable kao što su dob, spol, krvni tlak, kolesterol, vrste boli u prsima te prisutnost ili odsutnost srčane bolesti.

Pri izradi projekta vodite se ovim istraživačkim pitanjima:

- Postoji li razlika u maksimalnom broju otkucaja srca između pacijenata s anginom i onih bez angine?
- Postoji li značajna razlika u starosti pacijenata s obzirom na različite vrste boli u prsima?
- Možemo li predvidjeti prisutnost srčane bolesti na temelju maksimalnog broja otkucaja srca i starosti?
- Postoji li razlika u razini kolesterola između muškaraca i žena?

Također razmislite o vlastitim istraživačkim pitanjima i hipotezama koje želite analizirati koristeći dostupne podatke.

Napomena: Ako imate dodatnih pitanja vezanih uz ovaj konkretan projekt, javite se na e-mail adresu tomislav.kurtovic@fer.unizg.hr.

Analiza filmova na IMDb-u

Motivacija i opis problema

IMDB je najveća javno dostupna baza podataka o filmovima, serijama, glumcima, redateljima, scenaristima i mnoštvu drugih informacija vezanih uz kinematografiju. Osim samog prikaza podataka, IMDB dopušta korisnicima i ocjenjivanje te komentiranje filmova i serija. Analizom takvog skupa podataka moguće je doći do raznih zanimljivih zaključaka kao što su preferencije korisnika ili koje značajke čine film popularnim.

Opis skupa podataka i smjernice za izradu projekta

Podatci se sastoje od opisa 5044 filmova preuzetih s internetskih stranica IMDb-a. Uz svaki film dane su i općenite informacije poput žanra filma, godine izdanja, zemlje podrijetla ili imena redatelja. Dodatno, dani su i metrički podatci kao što su budžet filma, zarada filma, IMDB ocjena i slično.

Pri izradi projekta vodite se ovim istraživačkim pitanjima:

- Postoji li razlika u zaradi filmova namijenjenih za opću publiku (PG) i filmovima namijenjenih za osobe određene dobi (R)?
- Postoji li razlika u budžetima filmova s obzirom na njihov žanr?
- Možemo li temeljem danih varijabli predvidjeti IMDB ocjenu nekog filma?
- Razlikuju li se IMDB ocjene filmova s obzirom na vrijeme premijere filma?

Također razmislite o vlastitim istraživačkim pitanjima i hipotezama koje želite analizirati koristeći dostupne podatke.

Napomena: Ako imate dodatnih pitanja vezanih uz ovaj konkretan projekt, javite se na e-mail adresu krunoslav.jurcic@fer.unizg.hr.

Analiza glazbenih preferencija i njihovog utjecaja na mentalno zdravlje

Motivacija i opis problema

Muzikoterapija (MT) je metoda koja koristi glazbu i njezine elemente kako bi smanjila stres, poboljšala raspoloženje i unaprijedila opće mentalno zdravlje pojedinaca koristeći glazbu kao katalizator za hormone sreće poput oksitocina. Muzikoterapija obuhvaća širok spektar različitih glazbenih žanrova, pri čemu se izbor žanrova može značajno razlikovati među različitim organizacijama. Istraživanja, edukacija i klinička praksa u muzikoterapiji oslanjaju se na stručne standarde, uzimajući u obzir kulturološke, socijalne i političke kontekste.

Opis skupa podataka i smjernice za izradu projekta

Skup podataka obuhvaća 736 odgovora iz ankete koja se bavi slušateljskim navikama, glazbenim preferencijama i mentalnim zdravljem. Sadrži 33 varijable, uključujući dob ispitanika, podatke o tome jesu li instrumentalisti i/ili skladatelji, koji streaming servis koriste, slušaju li glazbu dok rade, slušaju li glazbu na stranom jeziku, istražuju li novu glazbu te prosječan broj sati dnevno provedenih slušajući glazbu. Također se bilježe učestalosti slušanja različitih žanrova, omiljeni žanr zajedno s njegovim BPM-om, samoprocijenjene razine mentalnih poremećaja i utjecaj glazbe na njih.

Pri izradi projekta vodite se ovim istraživačkim pitanjima:

- Slušaju li skladatelji više glazbe dnevno od drugih?
- Prikažite korelacije između frekvencija slušanja različitih glazbenih žanrova i samoprocijenjenih razina mentalnih poremećaja. Također, testirajte je li korelacija između frekvencije slušanja metal žanra i samoprocijenjene razine depresije statistički značajna.
- Postoji li povezanost između korištenog streaming servisa i sklonosti istraživanju?
- Razlikuje li se prosječni broj sati slušanja glazbe značajno među korisnicima ovisno o njihovim omiljenim žanrovima?
- Možete li se iz zadanih podataka predvidjeti dob ispitanika?

Također razmislite o vlastitim istraživačkim pitanjima i hipotezama koje želite analizirati koristeći dostupne podatke.

Napomena: Ako imate dodatnih pitanja vezanih uz ovaj konkretan projekt, javite se na e-mail adresu jana.juros@fer.unizg.hr.

Analiza podataka ljudi oboljelih od dijabetesa

Motivacija i opis problema

Dijabetes ili šećerna bolest se naziva globalnom epidemijom jer broj oboljelih osoba u svijetu raste na razine koje su bile nezamislive prije samo nekoliko desetljeća. Postoji više razloga za ovu pojavu, a najznačajniji je promjena načina života koju obilježava veći unos visokokalorične hrane i smanjena tjelesna aktivnost. Svjetske procjene govore da 537 milijuna osoba u dobi od 20 do 79 godina živi sa šećernom bolešću. Čak polovica oboljelih osoba nema postavljenu dijagnozu bolesti te ne prima liječenje. Statistička analiza raznih biokemijskih te zdravstvenih parametara može pomoći istraživačima i zdravstvenim djelatnicima u analizi i razumijevanju složenih interakcija i čimbenika rizika povezanih s različitim vrstama dijabetesa.

Opis skupa podataka i smjernice za izradu projekta

Skup sadrži zbirku medicinskih podataka vezanih uz različite oblike dijabetesa (steroidni dijabetes, predijabetes, dijabetes tipa 1 itd.). Sadrži brojne značajke koje obuhvaćaju genetske markere, okolišne čimbenike, životne navike i druge relevantne medicinske procjene.

Pri izradi projekta vodite se ovim istraživačkim pitanjima:

- Može li se na temelju zadanih parametara procijeniti dob pacijenta?
- Postoji li razlike u oblicima bolesti među pušačima i nepušačima?
- Postoje li razlike u razini zdravlja gušterače među oblicima dijabetesa pacijenata?
- Utječe li razina fizičke aktivnosti na razinu inzulina u krvi pacijenta?

Također razmislite o vlastitim istraživačkim pitanjima i hipotezama koje želite analizirati koristeći dostupne podatke.

Napomena: Ako imate dodatnih pitanja vezanih uz ovaj konkretan projekt, javite se na e-mail adresu krunoslav.jurcic@fer.unizg.hr.

Analiza podataka zdravstvenog pregleda

Motivacija i opis problema

Praćenje zdravstvenih pokazatelja i identifikacija faktora rizika ključni su za prevenciju kroničnih bolesti i unapređenje kvalitete života. Za medicinske stručnjake i zdravstvene organizacije važno je imati točne i pouzdane informacije o zdravlju populacije kako bi se mogli donijeti pravovremeni i informirani medicinski savjeti ili strategije prevencije.

Opis skupa podataka i smjernice za izradu projekta

Prikupljeni podatci predstavljaju informacije o zdravstvenim pregledima ispitanika. Podaci uključuju 17 značajki za svakog pojedinca, uključujući parametre poput visine, težine, krvnog tlaka, razine kolesterola, glukoze, te navika poput pušenja i konzumacije alkohola. Skup podataka također sadrži informacije o tjelesnoj aktivnosti i prisutnosti kardiovaskularnih bolesti. Cilj analize ovih podataka je istražiti povezanost zdravstvenih navika i rizika za kardiovaskularne bolesti, kao i identificirati ključne faktore koji utječu na zdravlje populacije.

Pri izradi projekta vodite se ovim istraživačkim pitanjima:

- Kakva je distribucija razina kolesterola među različitim dobnim skupinama i spolovima?
- Postoji li značajna razlika u prosječnom krvnom tlaku između pušača i nepušača?
- Razlikuje li se prosječni krvni tlak značajno među skupinama s različitom učestalošću tjelesne aktivnosti?
- Kakav je odnos između BMI-a i krvnog tlaka te možemo li predvidjeti krvni tlak na temelju BMI-a, dobi i učestalosti tjelesne aktivnosti?

Također razmislite o vlastitim istraživačkim pitanjima i hipotezama koje želite analizirati koristeći dostupne podatke.

Napomena: Ako imate dodatnih pitanja vezanih uz ovaj konkretan projekt, javite se na e-mail adresu tomislav.kovacevic@fer.unizg.hr.

Analiza potrošačkih navika kupaca

Motivacija i opis problema

Analiza potrošačkih navika kupaca omogućuje poduzećima prilagodbu ponude proizvoda različitim segmentima potrošača, s ciljem zadržavanja što većeg broja kupaca i poboljšanja prodajnih rezultata. Razumijevanje potrošačkih preferencija povećava učinkovitost marketinških kampanja kroz ciljano oglašavanje, optimizira asortiman te pospješuje upravljanje zalihama proizvoda.

Opis skupa podataka i smjernice za izradu projekta

Skup podataka sadrži osnovne demografske varijable poput dobi kupca, obrazovanja, bračnog statusa i prihoda kućanstva. Potrošnja kupaca razvrstana je prema različitim kategorijama proizvoda i prodajnim kanalima. Dodatno je na raspolaganju informacija je li pojedini potrošač ostvario kupnju u sklopu posebnih ponuda ili promotivnih kampanja.

Pri izradi projekta vodite se ovim istraživačkim pitanjima:

- Troše li kupci s djecom više na slatkiše od drugih klijenata?
- Mogu li dostupne varijable predvidjeti ukupnu potrošnju kupca?
- Postoji li razlika u prihodima kućanstva među klijentima različitog stupnja obrazovanja?
- Postoji li zavisnost između dobi klijenta i najčešće korištenih prodajnih kanala?

Također razmislite o vlastitim istraživačkim pitanjima i hipotezama koje želite analizirati koristeći dostupne podatke.

Napomena: Ako imate dodatnih pitanja vezanih uz ovaj konkretan projekt, javite se na e-mail adresu andro.mercep@fer.unizg.hr.

Analiza tržišta nekretnina

Motivacija i opis problema

Kupci nekretnina često žele zadovoljiti različite potrebe (npr. broj soba zbog djece ili potreba za većom garažom zbog dva auta), no takve “luksuze” ne žele preplatiti. Također, često su cijene nekretnina zbog različitih razloga napuhane, dok je bankama u interesu objektivno procijeniti vrijednost nekretnine za potrebe kreditiranja klijenata. Zbog toga se javlja potreba za prikupljanjem podataka o prodanim nekretninama.

Opis skupa podataka i smjernice za izradu projekta

Prikupljeni podatci predstavljaju informacije o prodanim nekretninama u gradu Ames (SAD). Svaka je nekretnina opisana sa 79 značajki, a ukupni broj prikupljenih podataka je 1460. Nekretnine su opisane značajkama poput kvadrature, blizine javnog prijevoza, brojem spavaćih soba, posjedovanja podruma, orijentacije i slično. Cilj analize ovih podataka je istražiti povezanost oblika, veličine i lokacije zemljišne čestice i nekretnine s njihovom cijenom.

Pri izradi projekta vodite se ovim istraživačkim pitanjima:

- Postoji li veza između zemljišne čestice i posjedovanja garaže?
- Kakva je razlika u kvadraturi podruma između dva proizvoljno odabrana kvarta?
- Razlikuje li se prosječna cijena po kvadratnom metru značajno između različitih tipova zgrada?
- Možemo li procijeniti konačnu prodajnu cijenu nekretnine na temelju ključnih karakteristika poput površine iznad razine zemlje, broja kupaonica itd.?

Također razmislite o vlastitim istraživačkim pitanjima i hipotezama koje želite analizirati koristeći dostupne podatke.

Napomena: Ako imate dodatnih pitanja vezanih uz ovaj konkretan projekt, javite se na e-mail adresu tomislav.kovacevic@fer.unizg.hr.

Analiza uspjeha učenika

Motivacija i opis problema

Podaci o učenicima prikupljeni su kako bi se istražili čimbenici koji utječu na njihov školski uspjeh. Cilj je razumjeti kako različiti aspekti, poput vremena provedenog u učenju, obiteljske podrške i izostanaka, doprinose postignućima učenika. Ova analiza omogućuje bolji uvid u područja koja su ključna za poboljšanje obrazovnih rezultata te za donošenje preporuka u svrhu optimizacije obrazovnog procesa.

Opis skupa podataka i smjernice za izradu projekta

Skup podataka obuhvaća informacije o demografskim karakteristikama učenika, njihovim obrazovnim navikama i ocjenama iz ključnih predmeta poput matematike i portugalskog jezika. Podaci uključuju varijable kao što su dob, spol, obiteljski status, vrijeme učenja, izostanci, te dodatna podrška u nastavi.

Pri izradi projekta vodite se ovim istraživačkim pitanjima:

- Jesu li prosječne konačne ocjene iz matematike različite između spolova?
- Postoji li razlika u prvoj ocjeni iz matematike s obzirom na mjesto stanovanja učenika?
- Možemo li predvidjeti prolaz iz završnog ispita iz jezika na temelju sociodemografskih varijabli poput spola, obrazovanja roditelja i veličine obitelji?
- Postoji li razlika u broju izostanaka iz matematike između učenika koji dolaze iz manjih i većih obitelji?

Također razmislite o vlastitim istraživačkim pitanjima i hipotezama koje želite analizirati koristeći dostupne podatke.

Napomena: Ako imate dodatnih pitanja vezanih uz ovaj konkretan projekt, javite se na e-mail adresu tomislav.kurtovic@fer.unizg.hr.

Analiza uspješnosti marketinške kampanje

Motivacija i opis problema

Poduzeća za potrebe promocije vlastitih proizvoda i usluga odabiru jedan od dva glavna pristupa: masovne marketinške kampanje, usmjerene na široku publiku, ili ciljani marketing, koji je usmjeren na specifičnu skupinu potrošača. Tradicionalne masovne kampanje, poput reklamnih panoa, tipično imaju vrlo nisku uspješnost koja prema provedenim studijama često iznosi i ispod 1%. S druge strane, ciljani marketing često pokazuje značajno veću učinkovitost zbog usmjerenosti na kupce koji su skloniji kupnji određenih proizvoda i usluga.

Opis skupa podataka i smjernice za izradu projekta

Skup podataka odnosi se na ciljanu telefonsku marketinšku kampanju koju je provela jedna bankarska institucija. Dostupne varijable uključuju osnovne sociodemografske podatke o klijentu (dob, obrazovanje, bračni status i slično), podatke o odnosu klijenta s bankom (poput stanja računa i iskorištenja kredita), informacije o prethodnim kontaktima s klijentom te podatke o uspješnosti održane marketinške kampanje.

Pri izradi projekta vodite se ovim istraživačkim pitanjima:

- Postoji li zavisnost između zanimanja i bračnog statusa klijenta?
- Imaju li klijenti s otvorenim kreditom više novca na računu od ostalih klijenata?
- Postoji li razlika trajanja poziva marketinške kampanje među klijentima različitog stupnja obrazovanja?
- Mogu li dostupne varijable predvidjeti uspješnost marketinške kampanje?

Također razmislite o vlastitim istraživačkim pitanjima i hipotezama koje želite analizirati koristeći dostupne podatke.

Napomena: Ako imate dodatnih pitanja vezanih uz ovaj konkretan projekt, javite se na e-mail adresu andro.mercep@fer.unizg.hr.

Analiza Vivino podataka

Motivacija i opis problema

Vivino je popularna aplikacija koja omogućava korisnicima da prodaju i ocjenjuju vina. Korisnici mogu pronaći sveobuhvatne informacije, uključujući cijene i recenzije, koje im olakšavaju donošenje informiranih odluka prilikom kupnje vina. Pored toga, platforma potiče dijeljenje iskustava i preporuka među ljubiteljima vina.

Opis skupa podataka i smjernice za izradu projekta

Skup podataka sadrži 12.205 dokumentiranih vina, a svako vino je opisano s 13 varijabli. Te varijable uključuju ime vina, regiju proizvodnje (kao što su Bordeaux, Napa Valley), državu proizvodnje i kod države, godinu berbe, jedinstveni identifikacijski broj (ID), ime vinarije, ocjenu vina (rating) koja se temelji na recenzijama korisnika, broj recenzija koje vino ima, cijenu vina, informaciju je li vino označeno kao prirodno ili nije (prirodno vino podrazumijeva proizvodnju uz minimalno intervenciju, tj. bez kemikalija i pesticida u vinogradu i dodavanja raznih aditiva samom vinu), primarnu sortu grožđa (npr. Cabernet Sauvignon, Chardonnay), te vrstu vina (crno, bijelo, rose, pjenušavo...).

Pri izradi projekta vodite se ovim istraživačkim pitanjima:

- Postoji li razlika u ocjenama vina među različitim vrstama vina?
- Postoji li povezanost između vrste vina i oznake je li vino prirodno?
- Jesu li vina iz Francuske popularnija (imaju više recenzija) od onih iz Italije?
- Može li se na temelju dostupnih podataka predvidjeti cijena vina?
- Koja vinska regija nudi najbolji omjer ocjene i cijene vina? Koju vinariju biste posjetili iz te regije?

Također razmislite o vlastitim istraživačkim pitanjima i hipotezama koje želite analizirati koristeći dostupne podatke.

Napomena: Ako imate dodatnih pitanja vezanih uz ovaj konkretan projekt, javite se na e-mail adresu jana.juros@fer.unizg.hr.