

SVEUČILIŠTE U ZAGREBU  
FAKULTET ORGANIZACIJE I INFORMATIKE  
VARAŽDIN

Ada Tikvan

**ANALIZA SPORTSKIH USPJEHA NA  
OLIMPIJSKIM IGRAMA**

**SEMINAR**

Varaždin, 2025.

SVEUČILIŠTE U ZAGREBU

FAKULTET ORGANIZACIJE I INFORMATIKE

V A R A Ž D I N

**Ada Tikvan**

**Matični broj: 0016158840**

**Studij: Informacijski i poslovni sustavi**

## **ANALIZA SPORTSKIH USPJEHA NA OLIMPIJSKIM IGRAMA**

**SEMINAR**

**Mentor:**

prof. dr. sc. Markus Schatten

**Varaždin, siječanj 2025.**

*Ada Tikvan*

**Izjava o izvornosti**

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

*Autorica potvrdila prihvaćanjem odredbi u sustavu FOI Radovi*

---

## Sažetak

Ovaj seminarski rad bavi se analizom sportskih uspjeha na Olimpijskim igrama. Podaci su sakupljeni iz dva izvora preuzeta sa stranice Kaggle. Prvi izvor je 'athletes.xlsx' datoteka u kojoj se nalaze imena sportaša, iz koje zemlje dolaze i u kojoj su se disciplini natjecali. Drugi izvor podataka je 'medals.csv', .csv datoteka sadržava podatke o broju osvojenih medalja po zemljama i kontinentima te plasman na olimpijskim igrama. Spajanjem ova dva izvora zadržani su svi redove iz .csv, a dodane informacije iz .xlsx samo za podudarne zemlje. Nadalje su obrađeni podaci spremjeni u SQLite bazu podataka, relevantni podaci su spremjeni u nove .xlsx datoteke te je napravljen grafikon koji prikazuje koliko je kojih medalja osvojeno po kontinentu.

**Ključne riječi:** heterogeni skup podataka, obrada podataka, vizualizacija podataka, spremanje u bazu podataka

# **Sadržaj**

<b>1.</b>	<b>Uvod</b>	.	.	.	.	.	.	1
<b>2.</b>	<b>Opis implementacije</b>	.	.	.	.	.	.	2
<b>3.</b>	<b>Prikaz rada aplikacije</b>	.	.	.	.	.	.	3
<b>4.</b>	<b>Kritički osvrt</b>	.	.	.	.	.	.	4
<b>5.</b>	<b>Zaključak</b>	.	.	.	.	.	.	5
<b>Popis literature</b>								6
<b>Popis slika</b>								7

# 1. Uvod

Tema analize sportskih uspjeha na Olimijskim igrama proizašla je iz potrage za skupovima podataka koji zahtjevaju dobru obradu, mogućnost spremanja u bazu i mogućnost dobre vizualne obrade. Baze podataka obrade sportskih uspjeha mogu biti od koristi istraživačima, medijima i svim pratiteljima sporta koji prate povijesne trendove, a jasnoća obrađenih podataka omogućava lako razumijevanje svima.

U ovom radu bit će prikazano spajanje relevantnih stupaca iz oba korištena izvora preuzetih sa stranice Kaggle [1], [2] kako bismo dobili konkrente i smislene podatke za daljnju analizu. Obrađeni podaci spremljeni su u SQLite bazu podataka, relevantni podaci su spremljeni u nove tablice te je napravljen grafikon koji prikazuje koliko je kojih medalja osvojeno po kontinentu **oraictic2011AkademskoPismoStrategije**.

## 2. Opis implementacije

Za implementaciju aplikacije korišteno je VS Code razvojno okruženje dok je sami kod napisan u Pythonu.

Na početku su uvezeni svi potrebni moduli kako bi se podaci mogli obraditi, spremiti u stvorenu bazu podataka, vizualizirati te izraditi REST API.

**Obrada podataka** Podaci su učitani uz pomoć pandas biblioteke (pd.read\_csv i pd.read\_excel) Učitani podaci obrađeni su odabirom ključnih stupaca iz oba skupa podataka. Spajanje obrađenih podataka (pd.merge) napravljeno je na temelju stupca 'Country' koji se nalazi u obje tablice

**SQLite baza podataka** Baza podataka olympics\_data.db stvorena je u SQLite alatu, a objektom veze 'conn' otvaramo vezu prema toj bazi. Korištenjem pokazivača 'cursor' na bazu podataka omogućeno je izvršavanje naredbi poput stvaranja tablica i upita. Kreira se tablica 'olympics' u koju se spremaju podaci iz spojenog skupa podataka, naziv države i koliko svaka država ima kojih medalja. Uz svaku državu je pridružen sportaš iz te države i u kojoj se disciplini natjecao. Nakon spremanja u bazu (conn.commit()), veza prema bazi se zatvara pomoću conn.close()

**Flask API** Za ovaj dio koda pomoć sam zatražila od umjetne inteligencije chatbota ChatGPT.[3] Na upit da mi u Pythonu kreira aplikaciju koja će omogućiti korisnicima pregled podataka o državama, osvojenim medaljama i sportašima preporučeno je korištenje Flask framework-a za izradu REST-APIja. [4]. Prvo se definira funkcija 'get\_data\_from\_db(query)' za dohvaćanje podataka iz baze. Nakon uspostavljanja veze s SQLite bazom podataka kreiran je kurzor koji izvršava zadani upit kojim se dohvaća lista svih redaka koji su rezultat upita. Nakon upita veza s bazom se zatvara. Kreiranjem rute 'medals' dohvaćeni su svi podaci prilikom GET zahtjeva prema adresi. SQL upit dohvaća sve podatke iz tablice *olympics*. Tako dohvaćeni podaci predstavljaju listu redaka, a petljom for se iterira kroz svaki redak kako bi se podaci pretvorili u rječnike koji se dodaju u listu result koju Flask koristi uz funkciju jsonify kako bi pretvorio podatke u JSON format.

**Vizualizacija podataka:** SQL upitom podatke grupiramo po tome koliko svaki kontinent ima koliko medalja. Rezultat je prikazan kao stupčasti grafikon.

**Spremanje obrađenih podataka:** obrađeni podaci o osvojenim medaljama prema kontinentima spremjeni su u medals\_by\_continent.xlsx, a podaci o državi, sportašima koji su barem jednu medalju, disciplinama i broju osvojenih medalja spremjeni su u athletes\_with\_medals.xlsx.

Aplikacija se pokreće naredbom app.run(debug=True)

### **3. Prikaz rada aplikacije**

Pokretanjem aplikacije prvo se čitaju učitani skupovi podataka koji se potom obrađeni spremaju u SQLite bazu podataka u tablicu 'olympics'. Da vidimo stanje u tablici možemo pokrenuti VS Code terminal ili upisati naredbe u Windows PowerShell.

```
Windows PowerShell
Copyright (C) Microsoft Corporation. All rights reserved.

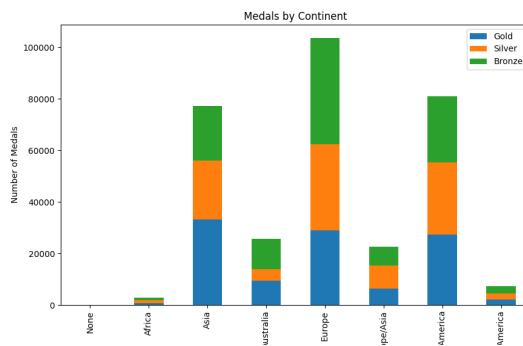
Try the new cross-platform PowerShell https://aka.ms/powershell

PS C:\Users\Ada> cd C:\Users\Ada\Desktop
PS C:\Users\Ada\Desktop> sqlite3 olympics data.db
SQLite version 3.30.2 2020-05-14 09:05:00
Enter ".help" for usage hints.
sqlite> .tables
olympics
athletes
headers on
sqlite> .mode column
sqlite> SELECT * FROM olympics LIMIT 10;
Country          Gold Medal    Silver Medal   Bronze Medal Total Continent      Name           Discipline
-----          -----       -----       -----       -----
United States of America 39.0        41.0        33.0        113.0 North America   ABBOTT Monica  Baseball/Softball
United States of America 39.0        41.0        33.0        113.0 North America   ABDIRAWAN Abdi   Athletics
United States of America 39.0        41.0        33.0        113.0 North America   ADEBARAifice Femi  Basketball
United States of America 39.0        41.0        33.0        113.0 North America   ALI Saeed Ali    Basketball/Softball
United States of America 39.0        41.0        33.0        113.0 North America   ASHINGER Nicole Trampoline Gymnastics
United States of America 39.0        41.0        33.0        113.0 North America   AKINADEWO Foluke Volleyball
United States of America 39.0        41.0        33.0        113.0 North America   ALLEN Jennifer   Athletics
United States of America 39.0        41.0        33.0        113.0 North America   ALLEN Niles     Basketball/Softball
United States of America 39.0        41.0        33.0        113.0 North America   ALLMAN Valarie  Athletics
United States of America 39.0        41.0        33.0        113.0 North America   ALVAREZ Anita   Artistic Swimming
sqlite> .exit
```

Slika 1: Prikaz baze podataka

Stvaranjem rute /medals aplikacija dohvaća sve podatke iz tablice `olympics` i vraća ih u JSON formatu.

Analiza i vizualizacija podataka omogućena je bibliotekama matplotlib i seaborn. Stupičasti grafikon prikazuje koliko je kojih medalja osvojeno po kontinentu.



Slika 2: Stupičasti grafikon

Relevantni obradeni podaci spremaju se u medals\_by\_continent.xlsx i athletes\_with\_medals.xlsx.

## **4. Kritički osvrt**

S obzirom da sam u zadnji tren morala promijeniti temu ovog rada zbog nedosljednosti podataka iz prijavljene teme, zadovoljna sam implementacijom ovog rješenja. Korišteni izvori podataka nisu komplikirani, ali ipak zahtjevaju smislenu obradu, analizu i povezivanje te spremanje u bazu podataka i kasniju vizualizaciju. Ovako obrađene podatke mogu pratiti statističari, mediji i pratitelji sporta

## **5. Zaključak**

Ovaj seminarski rad uspješno prikazuje obradu, analizu i vizualizaciju skupova podataka. Za implementaciju koda je korišten VS Code i jezik Python zbog jednostavnosti, a obrada podataka omogućena je uvozom raznih biblioteka. Kreirana aplikacija omogućuje pristup podacima i njihovu pohranu u bazi podataka SQLite.

# Popis literatúre

- [1] A. P. Sarkhel. „2021 Olympics in Tokyo.” Dostupno na: <https://www.kaggle.com/datasets/arjunprasadsarkhel/2021-olympics-in-tokyo>. (2022.), adresa: <https://www.kaggle.com/datasets/arjunprasadsarkhel/2021-olympics-in-tokyo>.
- [2] Ž. S. „Olympic Games 2021 Medals.” Dostupno na: <https://www.kaggle.com/datasets/stefanzivanov/olympic-games-2021-medals>. (2022.), adresa: <https://www.kaggle.com/datasets/stefanzivanov/olympic-games-2021-medals>.
- [3] „ChatGPT.” (), adresa: <https://chatgpt.com/>.
- [4] Flask. „Quickstart.” Dostupno na: <https://flask.palletsprojects.com/en/stable/quickstart/>. (), adresa: <https://flask.palletsprojects.com/en/stable/quickstart/>.

# **Popis slika**

1. Prikaz baze podataka . . . . .	3
2. Stupičasti grafikon . . . . .	3