

1.KV1 - Definiranje projektnog zadatka

1.1. Projektni zadatak

[Opisati glavni cilj projekta i svrhu vizualizacije podataka.

Ovaj zadatak zahtijeva jasno definiranje problema koji se pokušava riješiti. Potrebno je postaviti jasne ciljeve i granice projekta, kao i odrediti koji će se podaci prikupljati i analizirati kako bi se riješio problem.]

Naziv zadatka:

Vizualizacija podataka pokémona od 1. do 10. generacije

Opis problema:

Podaci o Pokémonima sadrže velik broj atributa poput osnovnih statistika (HP, Attack, Defense, Sp. Atk, Sp. Def, Speed), tipova (Type 1 i Type 2), fizičkih osobina (visina, težina), te raznih varijanti (forme, regionalne razlike, mega evolucije). Zbog velikog broja zapisa i višedimenzionalnosti podataka, korisnicima je otežano brzo snalaženje, usporedba i donošenje zaključaka o odnosima među Pokémonima.

Također, kombinacije tipova i njihova distribucija nisu odmah vidljive, kao ni razlike između Single-Type i Dual-Type Pokémona.

Opis zadatka:

Cilj zadatka je izraditi interaktivnu web-aplikaciju pomoću D3.js koja će omogućiti korisniku dinamički uvid u podatke o Pokémonima kroz vizualne prikaze. Implementirat će se više različitih tipova grafova.

Cilj projekta:

Krajnji cilj projekta je omogućiti korisniku brz, jasan i interaktivan uvid u osnovne karakteristike Pokémona, uz mogućnost dublje analize odnosa između tipova i statistika.

Poveznica na git repozitorij projekta:

<https://github.com/mbenkovic02/Pokemon-Data-Visualization>

1.2. Podatci

[Pronaći izvore podataka i opisati podatke koji će se koristiti za vizualizaciju.]

- Z-1.2.1. *U ovom zadatku potrebno je pronaći odgovarajuće izvore podataka koji će se koristiti za rješavanje problema definiranog u prvom zadatku. Važno je osigurati da su podaci kvalitetni i relevantni za problem koji se rješava te da su dostupni za upotrebu.*
- Z-1.2.2. *Potrebno je opisati odabrane podatke kao i navesti pripadajuće izvore.*

Za potrebe vizualizacije korišten je skup podataka *Pokemon Database.csv* preuzet s platforme [Kaggle](https://www.kaggle.com/datasets/mrdew25/pokemon-database/data), (<https://www.kaggle.com/datasets/mrdew25/pokemon-database/data>). Dataset obuhvaća Pokémone od 1. do 10. generacije i sadrži velik broj atributa koji omogućuju sveobuhvatnu analizu njihovih karakteristika.

Podaci uključuju:

- **Osnovne identifikatore:** ID, ime Pokémona, generacija, regionalna forma
- **Statistike:** HP, Attack, Defense, Sp. Atk, Sp. Def, Speed, Total
- **Tipove:** Type 1 i Type 2
- **Fizičke karakteristike:** visina, težina
- **Napredni atributi:** legendarni status, mitološki status, Baby Pokémon, Mega forma
- **Effort Value (EV) stupci:** broj EV bodova koje Pokémon daje kada bude poražen (Health EV, Attack EV, Defense EV, Special Attack EV, Special Defense EV, Speed EV, EV Yield Total)
- **Pomoćni opisi:** klasifikacija, spolni omjeri, sreća, grupa za uzgoj (Egg Group), evolucijske poveznice itd.

Podaci su kvalitetni i detaljno strukturirani, što omogućuje složene analize poput usporedbe po tipu, statistici, distribuciji EV-a, razlika između generacija.

1.3. Obrada podataka

[Opisati obavljeni postupak obrade i povezanja podataka.]

- Z-1.3.1. *Obraditi prikupljene podatke i povezati ih kako bi se stvorio cjelovit skup podataka. Ovo uključuje čišćenje i obradu podataka, kao i provjeru njihove konzistentnosti, aktualnosti, cjelovitosti, tj. kvalitete i ispravnosti.*

Nakon preuzimanja dataset-a obavljena je obrada podataka kako bi se osigurao kvalitetan i pregledan skup za vizualizaciju. U početnoj fazi provedena je selekcija relevantnih stupaca, dok su izbačeni oni koji se ne koriste u analizi ili ne doprinose korisničkom iskustvu.

Iz skupa podataka uklonjeni su svi stupci koji nisu doprinosili analitičkoj vrijednosti u kontekstu vizualizacije. Tako je izbačen Pokemon Id jer je Pokedex Number u kombinaciji s Alternate Form Name dovoljan za identifikaciju varijanti. Stupci poput Original Pokemon ID, Special Event Ability, Special Event Ability Description, Male Ratio, Female Ratio i Base Happiness također su uklonjeni zbog rijetke prisutnosti ili irelevantnosti za prikaz. Dodatno su uklonjeni gameplay-mehanički atributi poput Catch Rate, Experience Growth, Egg Group i Egg Cycle Count, koji nemaju značajnu vrijednost za komparativnu analizu statova i tipova.

1.4. Relevantne vrste prikaza za korištene podatke

[Predložiti nekoliko različitih načina prikaza podataka koji bi bili prikladni za ovaj projekt]

- Z-1.4.1. *Predložiti moguće načine prikaza podataka koji će pomoći u razumijevanju podataka i rješavanju problema koji je postavljen u prvom zadatku. Ovo može uključivati odabir najprikladnijeg načina vizualizacije podataka, ali to je zadatak iduće vježbe.*

1. Radar (spider) graf – Usporedba statistika Pokémona

Radar graf omogućuje usporedbu više Pokémona na temelju njihovih osnovnih statistika (HP, Attack, Defense, Sp. Atk, Sp. Def, Speed). Svaka os predstavlja jednu statistiku, a svaki

Pokémon je prikazan kao višekutnik.

Omogućena je selekcija 2–3 Pokémona, koji se prikazuju na istom grafikonu. Promjena odabira koristi D3.js *Enter–Update–Exit* obrazac s glatkim prijelazima i tooltipovima. Sa strane se prikazuju dodatni detalji o svakom Pokémonu koji nije vizualno prikazan unutar grafa (npr. tip, slika, ukupna snaga).

2. Boxplot – Usporedba distribucije statova po tipu

Boxplot prikazuje raspodjelu jedne od osnovnih statistika unutar odabranih `Type 1` tipova.

Korisnik može odabrati jednu metriku (HP, Attack, Defense, Sp. Atk, Sp. Def, Speed, Total) i do šest tipova Pokémona za usporedbu.

Svaki tip prikazan je kao box s prikazom medijana, kvartila, raspona i outliera. Promjene se ažuriraju korištenjem D3.js tranzicija i *Update* logike, čime se osigurava glatko vizualno iskustvo bez brisanja i ponovnog crtanja grafa.

3. Stacked bar chart – Broj Pokémona po tipu i tipu kombinacije (Single vs Dual-Type)

Ovaj prikaz prikazuje distribuciju `Type 1` tipova ovisno o tome imaju li Pokémoni i `Type 2` (Dual-Type) ili ne (Single-Type). Za svaki tip prikazuje se segmentirana traka s brojem Single-Type i Dual-Type Pokémona.

Prikaz koristi stacked layout, sortiranje po ukupnom broju Pokémona, tooltipove s brojem i postotkom, te ikonice tipova za veću preglednost.

4. Histogrami – Raspodjela baznih statova među svim Pokémonima

Za svaki osnovni stat (HP, Attack, Defense, Sp. Atk, Sp. Def, Speed) izrađen je histogram koji prikazuje distribuciju vrijednosti statova u populaciji Pokémona. Os X prikazuje raspon vrijednosti, dok os Y prikazuje broj (ili postotak) Pokémona unutar svakog intervala (bin-a). Grafikon omogućuje korisniku pregled gdje se većina Pokémona nalazi (npr. 50–100) te prepoznavanje rijetkih ekstremnih slučajeva.

5. Rang-liste – Top 10 Pokémona po pojedinim statistikama

Za svaku od osnovnih statistika prikazana je lista Pokémona s najvišim vrijednostima. Prikaz uključuje redni broj, ime Pokémona, vrijednost stat-a i sliku.

Korisnik može selektirati metriku koja se rangira (npr. Defense) te brzo pregledati koji Pokémoni dominiraju u toj kategoriji. Ovaj prikaz služi kao dopuna distribucijskim prikazima te omogućuje fokus na najjače pojedinačne entitete u dataset-u.

6. Stacked bar chart – Raspodjela EV statova po Pokémonima s najvećim EV Yieldom

Ova vizualizacija prikazuje koje i koliko EV bodova pojedini Pokémon daje kada bude poražen.

Svaki Pokémon s ukupnim EV Yieldom većim od nule prikazan je kao vodoravna traka, segmentirana po statovima (HP, Attack, Defense, Sp. Atk, Sp. Def, Speed) koje doprinosi.

Dužina i boja svakog segmenta predstavljaju broj EV bodova po stat-u, dok je ukupna dužina trake jednaka vrijednosti EV Yield Total. Pokémoni su sortirani po ukupnoj vrijednosti EV-a koje daju, čime se korisniku omogućuje brza identifikacija najefikasnijih ciljeva za EV treniranje (grindanje).

Tooltip prikazuje ime Pokémona, točnu raspodjelu EV-a po statovima

