Scenarijsko prognoziranje i ublažavanje okolišnih rizika

Ishodi učenja

Nakon završetka ove cjeline studenti će moći:

- Objasniti koncept scenarijskog prognoziranja u analizi okolišnih i klimatskih podataka.
- Opisati postupak agregiranja vremenskih nizova okolišnih podataka u godišnje klimatske pokazatelje i objasniti njegovu ulogu u analizi dugoročnih rizika.
- Primijeniti metodu nenadziranog strojnog učenja (K-means) za grupiranje klimatskih godina prema njihovim ukupnim obilježjima.
- Interpretirati dobivene klastere kao različite klimatske scenarije
- Raspraviti kako rezultati scenarijske analize mogu poduprijeti donošenje odluka

Uvod

Suvremeni sustavi praćenja okoliša sve se više oslanjaju na kontinuirano prikupljanje podataka iz različitih izvora, uključujući senzorske mreže, meteorološke postaje, satelitska opažanja i otvorene ekološke repozitorije. Ti sustavi redovito generiraju velike količine podataka i čine osnovu za praćenje stanja okoliša tijekom dugih razdoblja.

Za razliku od tradicionalnih analitičkih metoda, koje se često oslanjaju na statične ili ograničene skupove podataka, praćenje okoliša dinamičan je proces u kojem se podaci kontinuirano prikupljaju i ažuriraju. U takvom okruženju strojno učenje ne uključuje samo razvoj pojedinačnih modela, nego i održavanje trajnog analitičkog pristupa za obradu vremenski strukturiranih podataka i provedbu dugoročnih analiza.

Najvažnije okolišne varijable, poput razina onečišćujućih tvari u zraku, temperature, oborina ili razine mora, prirodno su organizirane kao vremenski nizovi. Ti podaci pokazuju sezonalnost, kratkoročne fluktuacije i dugoročne trendove, zbog čega su potrebne metode koje analiziraju kako se uvjeti mijenjaju tijekom vremena, a ne samo metode koje promatraju trenutačna stanja.

U tom kontekstu strojno učenje nudi metode za praćenje trendova i otkrivanje anomalija u okolišnim podacima. Da bi bile praktično korisne u operativnim sustavima, te metode moraju biti integrirane u širi okvir upravljanja ekološkim podacima, umjesto da se koriste kao izolirani analitički alati, čime podupiru informirane i odgovorne okolišne odluke kao što su rano otkrivanje problema s kvalitetom zraka, procjena rizika od ekstremnih vremenskih događaja ili praćenje dugoročnih promjena razine mora.

Od kontinuiranog praćenja okoliša do klimatskih scenarija

U dugoročnom praćenju okoliša oslanjanje samo na pojedinačna mjerenja ili kratka vremenska razdoblja često nije dovoljno. Kako bi se otkrile smislene promjene povezane s klimatskim rizicima, podaci se obično sažimaju tijekom duljih razdoblja, primjerice na godišnjoj razini. Takav pristup omogućuje da pojedine godine predstavljaju sveobuhvatne presjeke okolišnih uvjeta, a ne samo izolirana mjerenja.

Osim prosječnih vrijednosti, ekstremni događaji poput toplinskih valova ili sušnih razdoblja ključni su u analizi okolišnog rizika. Njihova važnost proizlazi ne samo iz intenziteta nego i iz učestalosti pojavljivanja i trajanja. Stoga se u analizi ne razmatraju samo prosječne temperature i oborine, nego i učestalost, trajanje i kombinacije nepovoljnih uvjeta.

Kada se ti pokazatelji promatraju zajedno, svaka se godina može opisati skupom značajki koje zajednički definiraju njezin klimatski profil. U tom se kontekstu zatim mogu primijeniti tehnike strojnog učenja za prepoznavanje sličnosti među godinama i njihovo grupiranje na temelju ukupnih obilježja.

Algoritam K-means klasteriranja u kontekstu okolišnih podataka

K-means [1], [2], [3] jednostavan je i široko korišten algoritam klasteriranja za prepoznavanje sličnosti u velikim skupovima podataka. Njegova je osnovna ideja podijeliti podatke u unaprijed zadan broj skupina (klastera), pri čemu svaka skupina sadržava zapise koji su međusobno najsličniji na temelju promatranih značajki. U ovoj metodi svaka je godina predstavljena skupom numeričkih značajki, a algoritam istodobno ispituje sve te značajke, tražeći ukupnu sličnost među zapisima umjesto podudaranja samo u jednoj varijabli.

U praksi algoritam započinje početnim procjenama središta skupina. Zatim ih višekratno prilagođava kako bi smanjio udaljenost između svakog zapisa i centroida njegove skupine. Taj se postupak nastavlja dok se raspodjela zapisa po skupinama ne stabilizira. Rezultat su skupine koje predstavljaju uobičajene obrasce u podacima.

U okolišnim analizama K-means grupira različita vremenska razdoblja, primjerice pojedine godine, na temelju njihova ukupnog klimatskog profila, a ne samo jedne varijable. Ova metoda pomaže prepoznati stabilne, prijelazne i ekstremne klimatske uvjete bez oslanjanja na unaprijed definirane pragove ili pravila.

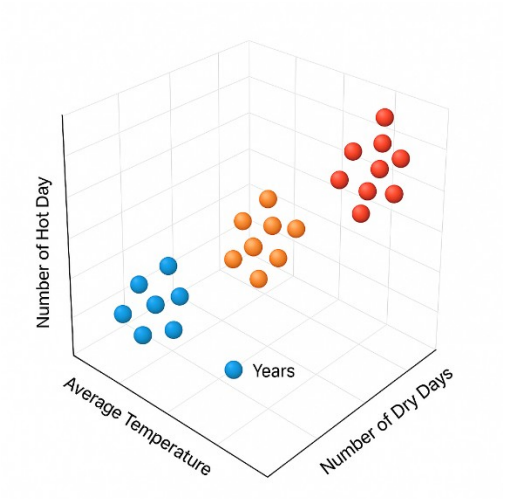
Zamislimo da svake godine pratimo nekoliko jednostavnih pokazatelja, poput prosječne temperature, broja vrlo vrućih dana i broja dana bez kiše. Promatraju li se ti pokazatelji pojedinačno, svaka se godina čini jedinstvenom, što otežava njihovo jasno opisivanje.

Međutim, kada se svi pokazatelji razmatraju zajedno, neke godine pokazuju vrlo slične obrasce, kao što je prikazano na slici 5.1. Primjerice, neke godine mogu imati umjerene temperature, rijetke toplinske valove i dovoljno oborina. U drugoj polovici godine obično se javljaju više temperature, češći toplinski valovi i dulja sušna razdoblja.

U ovom slučaju algoritam K-means ne određuje unaprijed broj skupina; umjesto toga automatski otkriva skupine godina koje su najsličnije prema svim promatranim pokazateljima. Svaka dobivena skupina može se interpretirati kao tipičan klimatski obrazac.

Jedna skupina može predstavljati stabilnije, umjerene godine; druga može odgovarati izraženom toplinskom stresu; a treća može biti povezana s ekstremnom vrućinom i sušom. Pri analizi novog skupa podataka svaka se godina može usporediti s tim obrascima i svrstati u skupinu kojoj je, prema svojim obilježjima, najsličnija.

Studija slučaja prikazana u nastavku temelji se na ovom konceptu te koristi dugoročne meteorološke vremenske nizove za klasifikaciju stvarnih klimatskih godina i njihovu analizu u kontekstu klimatskih rizika.



Slika 5.1 Ilustrativni primjer klasteriranja godišnjih klimatskih profila

Studija slučaja: analiza klimatskog rizika na temelju meteoroloških vremenskih nizova

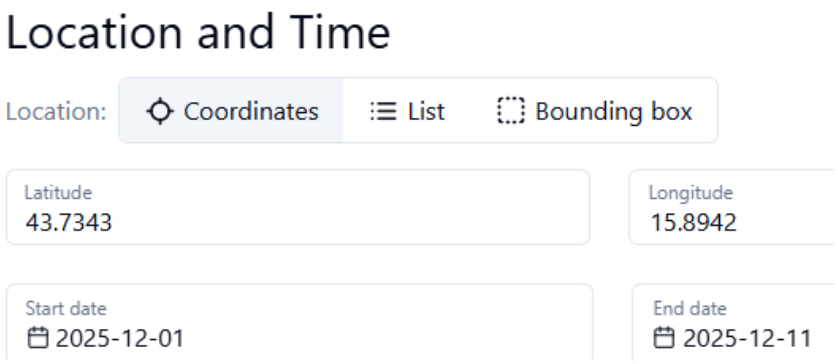
Izvor podataka Open-Meteo

Za ovu se studiju slučaja kao referentni izvor meteoroloških podataka koristi platforma Open-Meteo [4]. Open-Meteo omogućuje besplatan pristup povijesnim vremenskim podacima za bilo koju lokaciju u svijetu putem javno dostupnog API-ja, pri čemu korisnici mogu odabrati vremenski raspon i skup meteoroloških varijabli. Povijesni podaci dostupni su od 1940. godine, uz malo kašnjenje (do 1 dan) u pristupu najnovijim zapisima.

Iako je osnovni pristup podacima besplatan, Open-Meteo ograničava otvoreni model primjene ograničenjem broja API poziva i naprednih značajki. Za profesionalne i poslovne korisnike dostupni su prošireni API planovi koji nude veća ograničenja, više skupova podataka i pouzdaniju infrastrukturu.

Podaci iz Open-Meteo potječu iz javnih izvora, uključujući nacionalne meteorološke agencije i klimatske reanalize. Dostupni su pod licencijom Creative Commons Attribution 4.0 (CC BY 4.0), što znači da korisnici moraju navesti izvor. Ova licencija dopušta njihovu uporabu u obrazovnom i istraživačkom radu te potiče transparentnost i ponovljivost analize.

Preuzimanje se treba provesti putem jednostavnog mrežnog sučelja, uz određivanje geografske lokacije, vremenskog raspona (slika 5.2), vremenske rezolucije (satna ili dnevna) i odabranih meteoroloških varijabli. Takav pristup osigurava točno i dosljedno prikupljanje podataka za različite lokacije i analitičke scenarije.



Slika 5.2 Odabir lokacije podataka i vremenskog niza na platformi Open-Meteo

Primjer klasteriranja klimatskih godina

Za analizu klimatskog rizika važno je odabrati samo varijable koje su izravno povezane s dugoročnim promjenama i okolišnim stresom te čije je značenje jasno razumljivo. U tom će se kontekstu naglasak staviti na varijable koje opisuju toplinske uvjete, hidrološku bilancu i ekstremne vremenske događaje.

Osnovni pokazatelji toplinskog stresa uključuju prosječnu dnevnu temperaturu zraka na 2 metra (Mean Temperature (2 m) – temperature_2m_mean) i najvišu dnevnu temperaturu zraka (Maximum Temperature (2 m) – temperature_2m_max). Ta će mjerenja pomoći u praćenju dugoročnih trendova zagrijavanja i prepoznavanju razdoblja ekstremne vrućine. Osim toga, koristit će se Mean Apparent Temperature (2 m) (apparent_temperature_mean), koja uzima u obzir učinke vlage i vjetra, kako bi se dobila točnija procjena toplinskog opterećenja ljudi i okoliša.

Za procjenu hidrološkog i sušnog stresa odabrat će se varijable koje opisuju dostupnost i gubitak vode u okolišu. Ukupna dnevna oborina (Precipitation Sum – precipitation_sum) koristit će se za procjenu ulaza vode u sustav, dok će referentna evapotranspiracija (Reference Evapotranspiration – et0_fao_evapotranspiration) predstavljati potencijalni gubitak vode zbog isparavanja i transpiracije. Prosječna vlažnost tla u sloju 0-100 cm (Mean Soil Moisture (0-100 cm) – soil_moisture_0_100cm_mean) služit će kao pokazatelj stvarne dostupnosti vode u tlu i produljenoga sušnog stresa.

Kao dopunske varijable koje pružaju širi meteorološki kontekst i pomažu u interpretaciji rezultata koristit će se najveća brzina vjetra (Maximum Wind Speed (10 m) – wind_speed_10m_max), srednja relativna vlažnost (Mean Relative Humidity (2 m) – relative_humidity_2m_mean) i srednji tlak zraka reduciran na razinu mora (Mean Sea Level Pressure – pressure_msl_mean). Iako te varijable nisu središnje za analizu, poboljšat će razumijevanje uvjeta u kojima se klimatski rizici razvijaju i mijenjaju.

Razdoblje opažanja i vremenski obuhvat analize

Za analizu klimatskog rizika koristit će se povijesni meteorološki podaci od 1. siječnja 1950. do 31. prosinca 2024. Odabrano razdoblje pruža dovoljno dug vremenski niz za prepoznavanje dugoročnih promjena toplinskih i hidroloških uvjeta, uz zadržavanje relativno visoke i stabilne kvalitete podataka, što je ključno za točnu analizu i interpretaciju.

Podaci će se preuzimati na dnevnoj razini, čime se postiže ravnoteža između vremenske rezolucije i stabilnosti vremenskog niza. Dnevna rezolucija pomaže u praćenju sezonskih

obrazaca i promjena učestalosti nepovoljnih uvjeta, uz smanjenje šuma i kratkoročnih fluktuacija prisutnih u podacima veće frekvencije.

Analitički tijek rada i uporaba metoda strojnog učenja

Pri analizi toplinskih i hidroloških pokazatelja naglasak neće biti samo na bilježenju prisutnosti sezonskih obrazaca, koji su očekivani i dobro poznati (npr. ljetna sušna razdoblja), nego na tome kako se njihova obilježja mijenjaju tijekom duljih razdoblja. Konkretno, analiza će ispitati pomake u učestalosti, trajanju i intenzitetu nepovoljnih uvjeta, kao i moguće promjene njihova vremena pojavljivanja unutar godine.

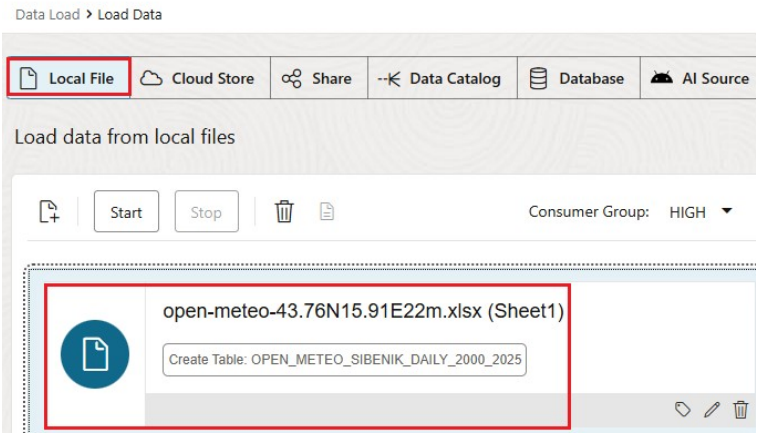
U kontekstu hidrološkog stresa analiza će se usmjeriti na razdoblja u kojima se dostupnost vode u tlu smanjuje tijekom više uzastopnih godina, u kombinaciji s povećanom evapotranspiracijom. Također će se procijeniti pojavljuju li se ta razdoblja ranije u sezoni, traju li dulje ili pokazuju veći intenzitet nego u prethodnim godinama. Te promjene upućuju na povećan rizik od suše, iako sezonski obrazac pojave ostaje nepromijenjen.

Slično tome, toplinski pokazatelji neće pratiti samo ljetne maksimume, nego i promjene u broju vrućih dana, trajanju toplinskih razdoblja i učestalosti visokih temperatura izvan tipičnih sezonskih raspona. Time se prirodni sezonski obrasci razlikuju od dugoročnih pomaka, što ima izravne implikacije za klimatske rizike.

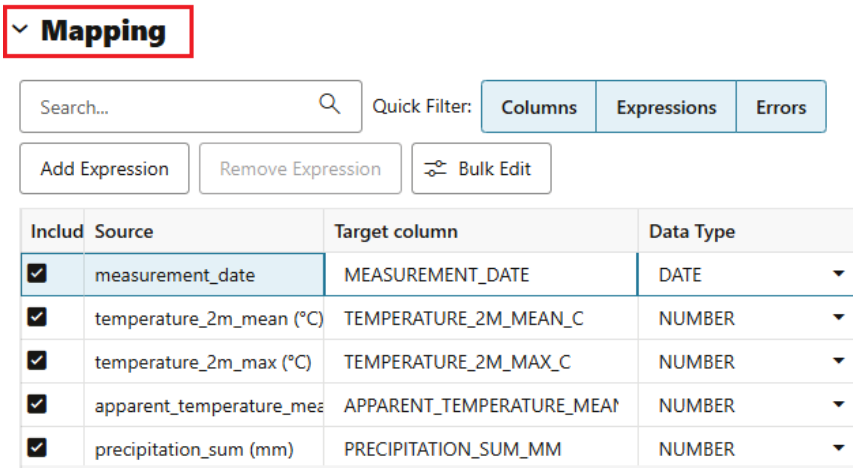
Uvoz podataka u Oracleovo skladište podataka

Preuzeti meteorološki podaci uvoze se putem ugrađenog sučelja Data Load u Oracle Autonomous Data Warehouse [5]. Sučelje omogućuje učitavanje podataka iz lokalnih .xlsx datoteka i njihovo automatsko pretvaranje u relacijske tablice unutar baze podataka, kao što je prikazano na slici 5.3.

Tijekom učitavanja navodi se naziv ciljne tablice, a stupci Excel datoteke preslikavaju se na odgovarajuće tipove podataka u bazi podataka. Datumski zapisi pohranjuju se kao vremenski atributi, dok se numeričke meteorološke varijable učitavaju kao numeričke vrijednosti prikladne za daljnju analizu, kao što je prikazano na slici 5.4.

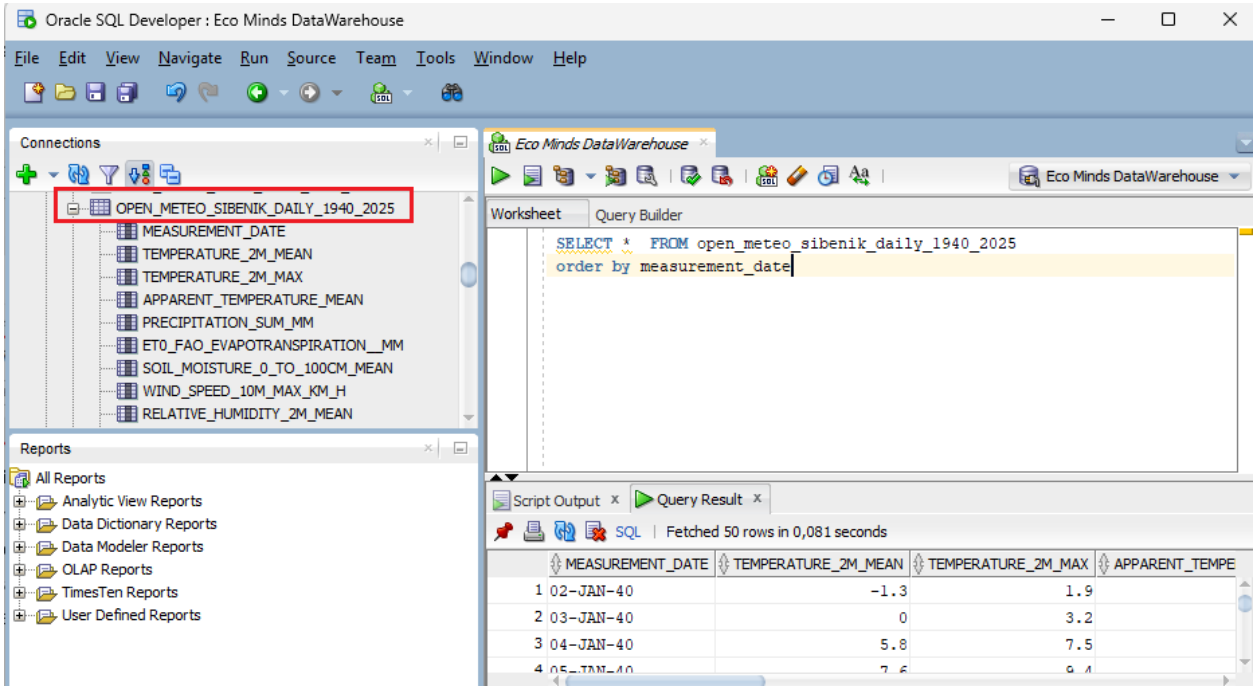


Slika 5.3 Sučelje Data Load za uvoz meteoroloških podataka u XLSX formatu



Slika 5.4 Preslikavanje stupaca iz Excel datoteke na tipove podataka u bazi tijekom uvoza

Ovaj postupak rezultira strukturiranom tablicom s dnevnim vremenskim nizovima odabranih meteoroloških varijabli za promatranu lokaciju (Šibenik) i zadano razdoblje (od 01.01.1940. do 31.12.2024.), kao što je prikazano na slici 5.5.



Slika 5.5 Struktura tablice OPEN_METEO_SIBENIK_DAILY_1940_2025 s dnevnim meteorološkim podacima

Agregiranje podataka i inženjering značajki

Nakon uvoza dnevnih meteoroloških mjerenja u relacijsku bazu podataka, sljedeći je korak agregirati podatke i stvoriti godišnje ulazne značajke (skup značajki). Budući da je cilj analize identificirati dugoročne obrasce i tipove klimatskih godina, dnevni se zapisi pretvaraju u godišnje sažetke koji opisuju toplinske, hidrološke i ekstremne uvjete svake godine kao jedan višedimenzijski zapis.

Agregiranje se provodi SQL upitom koji koristi više CTE-ova (common table expressions). Svaki dio izračunava određeni skup pokazatelja, a rezultati se spajaju u jedinstvenu godišnju tablicu značajki.

U bloku yearly_temps godišnja toplinska slika generira se pomoću statističkih mjera temperature zraka na 2 metra i osjetne temperature. Za svaku se godinu izračunavaju:

- prosječna i medijalna srednja dnevna temperatura (AVG, PERCENTILE_CONT(0.5)), čime se istodobno obuhvaćaju „tipična" temperatura i robusna mjera manje osjetljiva na ekstreme,
- minimalne i maksimalne vrijednosti (MIN, MAX), pri čemu se maksimum uzima iz temperature_2m_max kako bi se bolje obuhvatilo ekstremno toplinsko opterećenje,

- analogne mjere osjetne temperature (apparent_temperature_mean) kao pokazatelja stvarnog toplinskog stresa, koji uključuje utjecaj vlage i vjetra,
- osim toga, izračunava se prosječna relativna vlažnost kao kontekstna varijabla koja pomaže interpretirati razlike između „suhe vrućine" i razdoblja s povećanim toplinskim opterećenjem zbog vlage.

Ekstremni toplinski događaji procjenjuju se pomoću kriterija toplinskog vala. U odjeljku heatwave najprije se identificiraju dani u kojima temperature_2m_max prelazi 30 °C, pri čemu su ti dani uzastopni. Kontinuitet se uspostavlja grupiranjem uzastopnih datuma na temelju razlike između datuma i rednog broja u nizu. Samo epizode koje traju najmanje 5 dana (HAVING COUNT(*) >= 5) klasificiraju se kao toplinski valovi. Na temelju tih epizoda heatwave_stats izračunava godišnje ukupne vrijednosti.

- broj toplinskih valova u godini,
- prosječno, minimalno i maksimalno trajanje toplinskog vala.

Osim same razine temperature, obuhvaća se i struktura ekstremnih događaja (primjerice koliko se često pojavljuju i koliko dugo traju), što je ključno za razumijevanje klimatskog rizika.

Kako bi se sušni rizik opisao integrirano (a ne isključivo kroz oborine), u skupu podataka yearly_drought uveden je kompozitni pokazatelj „dan rizika od suše". Takav je dan definiran istodobnim ispunjenjem triju uvjeta:

- vrlo niska oborina (precipitation_sum_mm<1),
- visoka referentna evapotranspiracija (et0_fao_evapotranspiration__mm>4),
- niska vlažnost tla (soil_moisture_0_to_100cm_mean <0,20).

Osim toga, identificiraju se posebni dani s visokom evapotranspiracijom i niskom vlažnošću tla, što omogućuje razlikovanje godina u kojima dominira „atmosferska potražnja za vodom" od godina u kojima je primarni problem dugoročan deficit vode u tlu. Na taj se način sušni stres ne pojednostavnjuje u jedan signal. Umjesto toga modelira se kao funkcija ulaza (oborine), gubitaka (ET0) i stvarne dostupnosti vode (vlažnost tla).

Blok yearly_wind dodaje komponentu ekstremnih uvjeta vjetra kroz:

- najveću zabilježenu brzinu vjetra u godini,
- broj dana s jakim vjetrom (prag 50 km/h).

Ti pokazatelji nisu glavni fokus toplinsko-hidrološkog rizika, ali pomažu u razumijevanju ekstremnih događaja relevantnih za operativno praćenje i upravljanje rizicima.

Završni SELECT spaja sve izračunane godišnje skupove značajki u jedan rezultat, koristeći LEFT JOIN za pokazatelje koji ne moraju postojati u svakoj godini (npr. toplinski valovi), pri čemu se nedostajuće vrijednosti popunjavaju nulom (COALESCE). Izlaz je godišnji skup ulaznih značajki, u kojem je svaka godina jedan zapis s više čimbenika okolišnog stresa: vrućina, oborine, suša i vjetar. Ta matrica značajki predstavlja ulazni podatkovni prostor za naknadnu primjenu metoda nenadziranog strojnog učenja (npr. K-means).

```
WITH yearly_temps AS (
  SELECT
    EXTRACT(YEAR FROM measurement_date) AS year,

    -- Air temperature statistics
    ROUND(AVG(temperature_2m_mean), 1) AS avg_year_temp,
    ROUND(
      PERCENTILE_CONT(0.5)
      WITHIN GROUP (ORDER BY temperature_2m_mean),
      1
    ) AS median_year_temp,
    MIN(temperature_2m_mean) AS min_year_temp,
    MAX(temperature_2m_max) AS max_year_temp,

    -- Apparent temperature (thermal stress)
    ROUND(AVG(apparent_temperature_mean), 1) AS avg_apparent_temp,
    ROUND(
      PERCENTILE_CONT(0.5)
      WITHIN GROUP (ORDER BY apparent_temperature_mean),
      1
    ) AS median_apparent_temp,
    MIN(apparent_temperature_mean) AS min_apparent_temp,
    MAX(apparent_temperature_mean) AS max_apparent_temp,

    -- Relative humidity
    ROUND(AVG(relative_humidity_2m_mean), 1) AS avg_relative_humidity

  FROM open_meteo_sibenik_daily_1940_2025
  WHERE measurement_date BETWEEN
        DATE '1950-01-01' AND DATE '2024-12-31'
  GROUP BY EXTRACT(YEAR FROM measurement_date)
),
yearly_hydrology AS (
  SELECT
    EXTRACT(YEAR FROM measurement_date) AS year,

    -- Hydrological indicators
    COUNT(CASE WHEN precipitation_sum_mm < 1 THEN 1 END) AS dry_days,
    COUNT(CASE WHEN precipitation_sum_mm >= 1 THEN 1 END) AS rainy_days,
    MAX(precipitation_sum_mm) AS max_daily_precipitation_mm

  FROM open_meteo_sibenik_daily_1940_2025
```

```
  WHERE measurement_date BETWEEN
        DATE '1950-01-01' AND DATE '2024-12-31'
  GROUP BY EXTRACT(YEAR FROM measurement_date)
),
heatwaves AS (
  SELECT
    grp,
    MIN(measurement_date) AS start_date,
    COUNT(*) AS duration_days
  FROM (
    SELECT
      measurement_date,
      measurement_date
        - ROW_NUMBER() OVER (ORDER BY measurement_date) AS grp
    FROM open_meteo_sibenik_daily_1940_2025
    WHERE measurement_date BETWEEN
          DATE '1950-01-01' AND DATE '2024-12-31'
      AND temperature_2m_max > 30
  )
  GROUP BY grp
  HAVING COUNT(*) >= 5
),
heatwave_stats AS (
  SELECT
    EXTRACT(YEAR FROM start_date) AS year,

    -- Heatwave indicators
    COUNT(*) AS heatwave_count,
    ROUND(AVG(duration_days), 1) AS heatwave_avg_duration_days,
    MIN(duration_days) AS heatwave_min_duration_days,
    MAX(duration_days) AS heatwave_max_duration_days

  FROM heatwaves
  GROUP BY EXTRACT(YEAR FROM start_date)
),
yearly_drought AS (
  SELECT
    EXTRACT(YEAR FROM measurement_date) AS year,

    -- Combined drought risk (all conditions met)
    COUNT(
      CASE
        WHEN precipitation_sum_mm < 1
         AND et0_fao_evapotranspiration__mm > 4
         AND soil_moisture_0_to_100cm_mean < 0.20
        THEN 1
      END
    ) AS drought_risk_days,

    -- Individual drought-related conditions
    COUNT(
      CASE
        WHEN et0_fao_evapotranspiration__mm > 4
```

```
        THEN 1
      END
    ) AS high_et0_days,

    COUNT(
      CASE
        WHEN soil_moisture_0_to_100cm_mean < 0.20
        THEN 1
      END
    ) AS low_soil_moisture_days

  FROM open_meteo_sibenik_daily_1940_2025
  WHERE measurement_date BETWEEN
        DATE '1950-01-01' AND DATE '2024-12-31'
  GROUP BY EXTRACT(YEAR FROM measurement_date)
),
yearly_wind AS (
  SELECT
    EXTRACT(YEAR FROM measurement_date) AS year,

    -- Wind extremes
    MAX(wind_speed_10m_max_km_h) AS max_wind_speed_km_h,
    COUNT(
      CASE
        WHEN wind_speed_10m_max_km_h > 50
        THEN 1
      END
    ) AS strong_wind_days_50kmh

  FROM open_meteo_sibenik_daily_1940_2025
  WHERE measurement_date BETWEEN
        DATE '1950-01-01' AND DATE '2024-12-31'
  GROUP BY EXTRACT(YEAR FROM measurement_date)
)
SELECT
  t.year,

  -- Heatwave indicators
  COALESCE(h.heatwave_count, 0) AS heatwave_count,
  COALESCE(h.heatwave_avg_duration_days, 0) AS heatwave_avg_duration_days,
  COALESCE(h.heatwave_min_duration_days, 0) AS heatwave_min_duration_days,
  COALESCE(h.heatwave_max_duration_days, 0) AS heatwave_max_duration_days,

  -- Air temperature
  t.avg_year_temp,
  t.median_year_temp,
  t.min_year_temp,
  t.max_year_temp,

  -- Apparent temperature
  t.avg_apparent_temp,
  t.median_apparent_temp,
  t.min_apparent_temp,
```

```
  t.max_apparent_temp,

  -- Relative humidity
  t.avg_relative_humidity,

  -- Hydrology
  y.dry_days,
  y.rainy_days,
  y.max_daily_precipitation_mm,

  -- Drought stress
  COALESCE(d.drought_risk_days, 0) AS drought_risk_days,
  COALESCE(d.high_et0_days, 0) AS high_et0_days,
  COALESCE(d.low_soil_moisture_days, 0) AS low_soil_moisture_days,

  -- Wind
  w.max_wind_speed_km_h,
  w.strong_wind_days_50kmh

FROM yearly_temps t
LEFT JOIN heatwave_stats h ON t.year = h.year
JOIN yearly_hydrology y ON t.year = y.year
LEFT JOIN yearly_drought d ON t.year = d.year
LEFT JOIN yearly_wind w ON t.year = w.year
ORDER BY t.year;
```

Analiza i vizualizacija agregiranih godišnjih pokazatelja

Na temelju godišnjih ukupnih vrijednosti izračunanih SQL upitom opisanim u prethodnom potpoglavlju provedena je vizualna analiza ključnih toplinskih i hidroloških pokazatelja. Ona služi kao dopunski analitički alat za razumijevanje dugoročnih obrazaca i praćenje promjena klimatskih uvjeta tijekom vremena.

Slika 5.6 prikazuje prosječnu godišnju temperaturu zraka u Šibeniku od 1950. do 2024., izračunanu kao godišnji prosjek dnevne prosječne temperature zraka na 2 m (temperature_2m_mean).

Vizualizacija potvrđuje dugoročni trend porasta temperatura, uz izraženu varijabilnost iz godine u godinu. Osobito od početka 2000-ih prosječne su temperature više od prethodnih višegodišnjih prosjeka, a i posljednjih godina zabilježena su razdoblja znatno povišenih temperatura.

Taj trend ne upućuje samo na postupno zagrijavanje, nego i na pomak u raspodjeli godišnjih toplinskih uvjeta, što je ključno za naknadnu analizu toplinskih valova i za klasteriranje klimatskih godina metodama strojnog učenja.

Slika 5.7 prikazuje godišnji broj toplinskih valova, definiranih kao razdoblja od najmanje pet uzastopnih dana s maksimalnom dnevnom temperaturom iznad 30 °C, zajedno s njihovim prosječnim trajanjem. Na početku niza toplinski su valovi rijetki i općenito kratki. Međutim, posljednjih godina njihova se učestalost povećala, a trajanje pojedinačnih događaja znatno je naraslo.

Ovaj obrazac pokazuje promjenu načina na koji toplinski stres utječe na sustave: umjesto povremenih ekstremnih događaja, posljednjih su godina sve češća dulja razdoblja visokih temperatura, što znatno utječe na ljudsko zdravlje, ekosustave i uporabu resursa.

Slika 5.8 prikazuje godišnji broj sušnih i kišnih dana, pri čemu su sušni dani definirani kao dani s manje od 1 mm oborine. Grafikon otkriva znatnu varijabilnost iz godine u godinu, uz noviji trend prema većem broju sušnih i manjem broju kišnih dana.

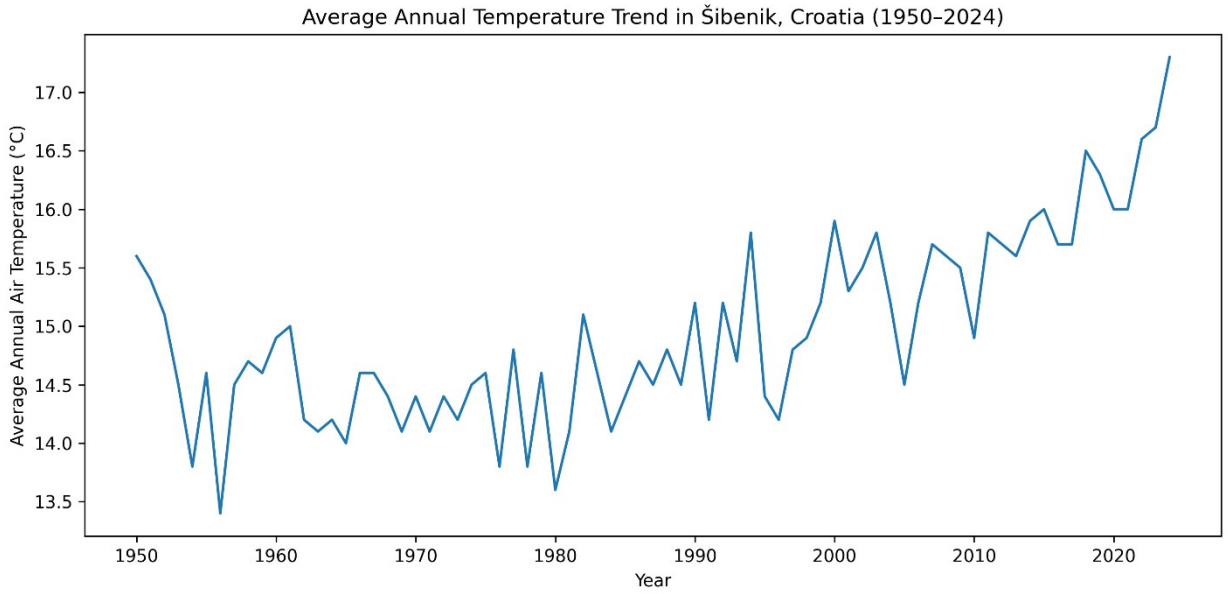
Ta promjena u raspodjeli dana s oborinama ne znači nužno smanjenje ukupne godišnje količine oborina. Umjesto toga pokazuje veću koncentraciju kišnih događaja i dulja sušna razdoblja, što je važno za razvoj sušnog stresa.

Slika 5.9 prikazuje godišnji broj dana koji zadovoljavaju kriterije izraženog sušnog rizika, obilježenog vrlo niskom oborinom, povećanom referentnom evapotranspiracijom i smanjenom vlažnošću tla. Tijekom većeg dijela promatranog razdoblja takvi uvjeti nisu istodobno ispunjeni; međutim, posljednjih godina zabilježen je znatno veći broj sušnih dana.

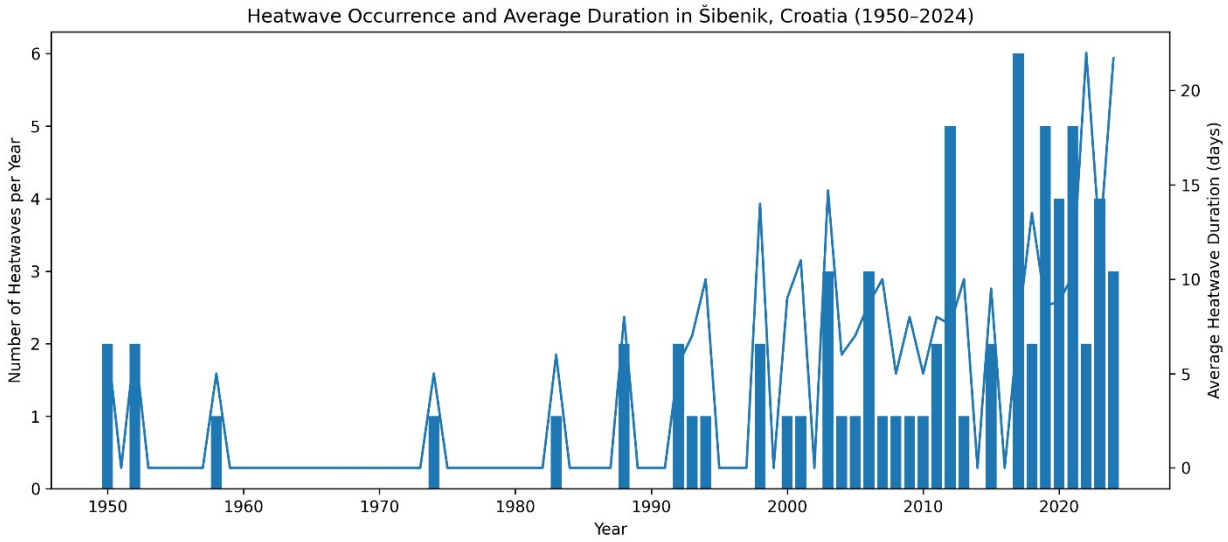
Pojava višedesetljetnih razdoblja s kombiniranim aridnim stresom upućuje na promjenu načina interakcije atmosferskih i hidroloških čimbenika, pri čemu suša više nije samo posljedica manjka oborina, nego i povećanog gubitka vode te smanjene vlažnosti tla.

Važno je prepoznati da se nepovoljna obilježja ne moraju nužno pojaviti svake godine; njihove kombinacije i razine izraženosti razlikuju se od godine do godine. Višedimenzijska priroda godišnjih klimatskih značajki naglašava potrebu za metodom koja može istodobno analizirati više pokazatelja i otkriti prirodne obrasce njihova pojavljivanja.

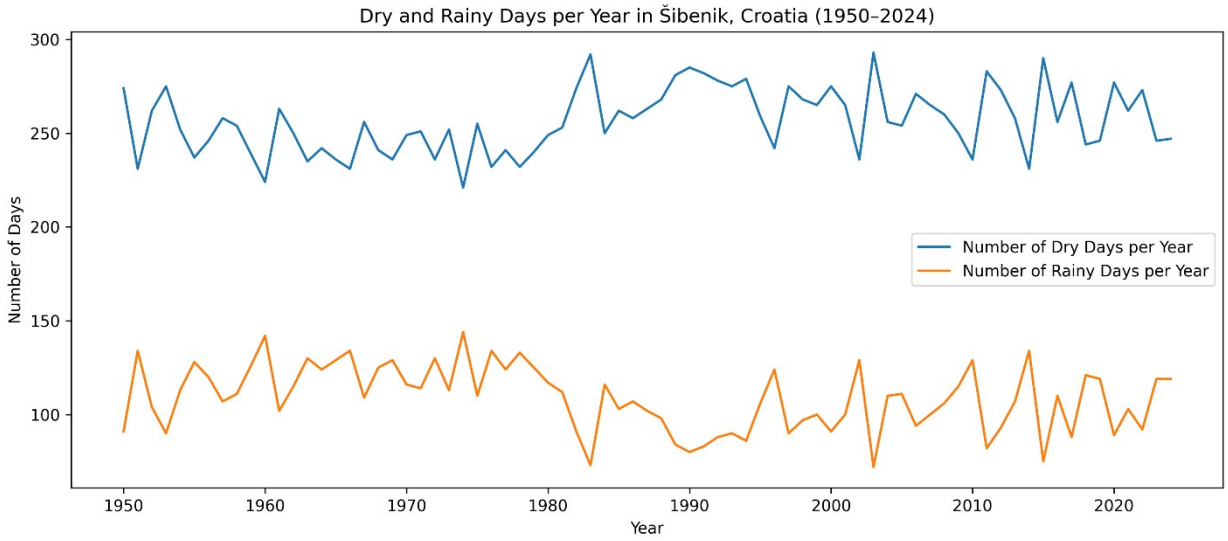
Sljedeći je korak u analizi primjena metoda nenadziranog strojnog učenja koje godišnje zapise tretiraju kao cjelovite klimatske profile. Takav pristup omogućuje grupiranje godina prema njihovim ukupnim sličnostima, što pomaže identificirati različite tipove klimatskih godina i interpretirati ih u smislu klimatskih rizika.



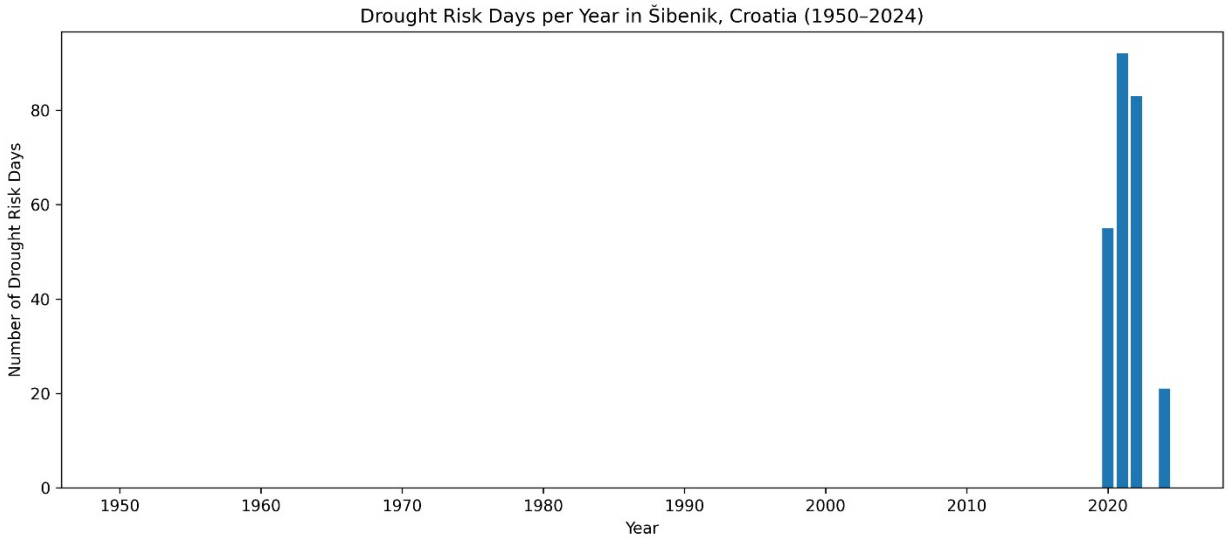
Slika 5.6 Prosječna godišnja temperatura zraka u Šibeniku (1950.–2024.)



Slika 5.7 Učestalost i prosječno trajanje toplinskih valova u Šibeniku (1950.–2024.)



Slika 5.8 Broj sušnih i kišnih dana po godini u Šibeniku (1950.–2024.)



Slika 5.9 Broj dana koji zadovoljavaju kriterij izraženog rizika od suše u Šibeniku (1950.–2024.)

Nenadzirano strojno učenje za grupiranje godišnjih klimatskih profila

Nakon agregiranja dnevnih meteoroloških mjerenja u godišnje skupove značajki, na podatkovni je prostor primijenjena metoda K-means klasteriranja. Analiza je provedena izravno u Oracleovu skladištu podataka pomoću ugrađenog paketa DBMS_DATA_MINING [6], koji omogućuje integraciju pripreme podataka, modeliranja i interpretacije rezultata unutar jedinstvenog relacijskog okvira.

Model grupiranja definiran je kao model klasteriranja (CLUSTERING) primijenjen na tablicu CLIMATE_YEARLY_FEATURES_SIBENIK, koja je izravan rezultat SQL agregacijskog upita opisanog u prethodnom potpoglavlju. Sadržava po jedan zapis za svaku godinu s pripadajućim skupom godišnjih agregiranih značajki, uključujući toplinske, hidrološke, sušne i vjetrovne pokazatelje.

Broj klastera postavljen je na 3 u tablici parametara modela (CLUS_NUM_CLUSTERS = 3). Takav odabir omogućuje uravnoteženu segmentaciju godišnjih klimatskih profila u manji broj razumljivijih tipova, bez pretjeranog usitnjavanja rezultata.

Model K-means istreniran je na svim dostupnim godišnjim značajkama, pri čemu su ulazne varijable automatski normalizirane i standardizirane unutar okvira Oracle Data Mining. Time se osigurava da varijable s većim numeričkim rasponima (poput broja dana ili ekstremnih vrijednosti) ne dominiraju postupkom klasteriranja.

Nakon izgradnje modela svakoj se godini dodjeljuje određeni klaster pomoću ugrađene funkcije CLUSTER_ID. Rezultati pokazuju da su godine grupirane u tri različita klastera, od kojih svaki ima različite vremenske nizove i klimatske značajke, kao što je prikazano u tablici 1.

Jedan klaster (ID = 4), koji se može smatrati stabilnim klimatskim scenarijem (označen zelenom bojom), uključuje najveći broj godina i obilježen je najnižim prosječnim temperaturama, gotovo potpunim izostankom toplinskih valova ili sušnog stresa te najvećim brojem kišnih dana. Taj se klaster sastoji od godina s relativno stabilnom, umjerenom klimom. Slika 5.10 pokazuje da taj klaster dominira tijekom većeg dijela promatranog razdoblja, osobito u drugoj polovici 20. stoljeća.

Drugi klaster (ID = 5), koji predstavlja prijelazni klimatski scenarij (označen plavom bojom), obilježen je višim prosječnim temperaturama, češćim toplinskim valovima te većim brojem sušnih dana uz manje kišnih dana. Ta kombinacija upućuje na povećan toplinski i hidrološki stres, iako ne doseže ekstremne vrijednosti sušnog pokazatelja. Slika 5.10 pokazuje da se godine u ovom klasteru sporadično pojavljuju u ranijim razdobljima, ali postaju znatno češće nakon 2000. godine.

Treći klaster (ID = 3), koji predstavlja nepovoljan (ekstremni) klimatski scenarij označen narančastom bojom, iako obuhvaća manji broj godina, jasno je obilježen najvišim prosječnim temperaturama, najvećim brojem toplinskih valova i iznimno velikim brojem dana s kombiniranim aridnim stresom. Kao što je prikazano na slici 5.10, taj se klaster pojavljuje gotovo isključivo u najnovijem promatranom razdoblju, što upućuje na kvalitativni pomak u klimatskim obilježjima, a ne samo na postupnu promjenu pojedinačnih varijabli.

Tablica 1. Rezultati K-means klasteriranja klimatskih pokazatelja za Šibenik

ID klastera	Broj godina u klasteru	Prosječna godišnja temperatura (°C)	Prosječna osjetna temperatura (°C)	Prosječan broj toplinskih valova godišnje	Prosječan broj dana s rizikom od suše	Prosječan broj sušnih dana godišnje	Prosječan broj kišnih dana godišnje	Prosječna najveća brzina vjetra (km/h)
3	**4**	**16,5**	**15**	**3,5**	**62,8**	**264,8**	**100,8**	**44,5**
4	46	14,5	12,7	0,1	0	249,4	115,8	53,7
5	25	15,5	13,9	2,1	0	268,1	97,2	51,8

Slika 5.10 Dodjela godina klimatskim klasterima u Šibeniku (1950.–2024.)

```
-- ======================================================
-- K-means clustering of yearly climate profiles
-- ------------------------------------------------------
-- This script implements unsupervised machine learning
-- directly inside Oracle Autonomous Data Warehouse.
-- The input table CLIMATE_YEARLY_FEATURES_SIBENIK
-- contains one row per year with aggregated climatic
-- indicators derived from daily meteorological data.
--
-- The goal of the analysis is to group years with
-- similar climatic characteristics using the K-means
-- clustering algorithm.
-- ======================================================

-- ------------------------------------------------------
-- Step 1: Define K-means model settings
-- ------------------------------------------------------
-- Create a settings table used by Oracle Data Mining
-- to configure the clustering algorithm.

CREATE TABLE KMEANS_SETTINGS (
  setting_name  VARCHAR2(30),
  setting_value VARCHAR2(30)
);

-- Specify the number of clusters (K = 3)
INSERT INTO KMEANS_SETTINGS
VALUES ('CLUS_NUM_CLUSTERS', '3');


-- ------------------------------------------------------
-- Step 2: Create the K-means clustering model
-- ------------------------------------------------------
-- The model is trained directly inside the database
```

```
-- using the DBMS_DATA_MINING package.
-- Each row represents one year (case),
-- and all remaining columns are used as input features.

BEGIN
  DBMS_DATA_MINING.CREATE_MODEL(
    model_name          => 'CLIMATE_KMEANS_SIBENIK',
    mining_function     => DBMS_DATA_MINING.CLUSTERING,
    data_table_name     => 'CLIMATE_YEARLY_FEATURES_SIBENIK',
    case_id_column_name => 'YEAR',
    settings_table_name => 'KMEANS_SETTINGS'
  );
END;


-- ------------------------------------------------------
-- Step 3: Verify that the model was successfully created
-- ------------------------------------------------------

SELECT model_name, mining_function
FROM user_mining_models
WHERE model_name = 'CLIMATE_KMEANS_SIBENIK';


-- ------------------------------------------------------
-- Step 4: Assign each year to a cluster
-- ------------------------------------------------------
-- The CLUSTER_ID function returns the cluster
-- to which each yearly climate profile belongs.

SELECT
  year,
  CLUSTER_ID(CLIMATE_KMEANS_SIBENIK USING *) AS cluster_id
FROM CLIMATE_YEARLY_FEATURES_SIBENIK
ORDER BY year;


-- ------------------------------------------------------
-- Step 5: Profile clusters using average feature values
-- ------------------------------------------------------
-- This query summarizes the climatic characteristics
-- of each cluster by calculating average values
-- of selected indicators.

SELECT
  CLUSTER_ID(CLIMATE_KMEANS_SIBENIK USING *) AS cluster_id,
  COUNT(*) AS years_in_cluster,

  -- Temperature characteristics
  ROUND(AVG(avg_year_temp), 1) AS avg_temp,
  ROUND(AVG(avg_apparent_temp), 1) AS avg_apparent_temp,

  -- Heat stress indicators
```

```
  ROUND(AVG(heatwave_count), 1) AS avg_heatwaves,

  -- Drought-related indicators
  ROUND(AVG(drought_risk_days), 1) AS avg_drought_days,
  ROUND(AVG(dry_days), 1) AS avg_dry_days,

  -- Precipitation characteristics
  ROUND(AVG(rainy_days), 1) AS avg_rainy_days,

  -- Wind extremes
  ROUND(AVG(max_wind_speed_km_h), 1) AS avg_max_wind

FROM CLIMATE_YEARLY_FEATURES_SIBENIK
GROUP BY CLUSTER_ID(CLIMATE_KMEANS_SIBENIK USING *)
ORDER BY cluster_id;
```

Potvrda identificiranih klimatskih obrazaca pomoću novijih podataka

Rezultati studije slučaja na području Šibenika (Hrvatska) pokazuju da se klimatske promjene očituju ne samo kroz postupno povećanje temperatura, nego i kroz promjene u strukturi klimatskih godina te češće nepovoljne kombinacije toplinskih i hidroloških uvjeta. Vizualna analiza godišnjih pokazatelja otkriva povećanje učestalosti i trajanja toplinskih valova, kao i pojavu godina sa značajnim kombiniranim sušnim stresom.

Rezultati klasterske analize primjenom metode K-means pokazali su da se od 2020. u promatranom području pojavio poseban tip klimatske godine, obilježen najizraženijim toplinskim i sušnim stresom. Taj se klaster razlikuje po visokim prosječnim i maksimalnim temperaturama, češćim i duljim toplinskim valovima te produljenim razdobljima kombiniranog sušnog stresa; tijekom vremena pojavljuje se gotovo isključivo u najnovijem razdoblju analize.

Uvid u gotovo potpune meteorološke podatke za 2025., prikazane u tablici 2 (zaključno s 12. prosinca 2025.), pokazuje da se obilježja te godine vrlo dobro podudaraju s prethodno identificiranim obrascem. Zabilježena su četiri toplinska vala prosječnog trajanja duljeg od 13 dana, iznimno visoka maksimalna dnevna temperatura od 38,4 °C i više od 100 dana s kombiniranim sušnim stresom, dok se broj kišnih dana smanjio. Ta kombinacija toplinskih i hidroloških pokazatelja odgovara profilu najekstremnijeg klastera identificiranog analizom povijesnih podataka.

To potvrđuje da primijenjena K-means analiza nije samo retrospektivan opis prošlih uvjeta, nego i alat za otkrivanje strukturnih promjena u klimatskom ponašanju sustava. Stoga 2025. godina nije izoliran ekstrem, nego dio obrasca koji je započeo početkom 2020-ih, što naglašava važnost ovog pristupa za rano prepoznavanje nepovoljnih scenarija i procjenu okolišnog rizika.

Tablica 2. Ključni klimatski pokazatelji za 2025. (Šibenik)

Pokazatelj	Vrijednost
Broj toplinskih valova	4
Prosječno trajanje toplinskog vala (dani)	13,5
Maksimalna dnevna temperatura (°C)	38,4
Prosječna godišnja temperatura zraka (°C)	17
Prosječna osjetna temperatura (°C)	15,3
Sušni dani	252
Broj kišnih dana	94
Broj dana s kombiniranim sušnim stresom	101
Broj dana s povećanom evapotranspiracijom	128
Najveća brzina vjetra (km/h)	48,4

Implikacije za scenarijsko ublažavanje okolišnih rizika

Na području Šibenika identificirani klimatski klasteri služe kao scenariji utemeljeni na dokazima koji prikazuju različite razine okolišnog rizika. Godine s relativno stabilnim klimatskim uvjetima predstavljaju osnovni scenarij. Međutim, jedan je klaster obilježen značajnim toplinskim i sušnim stresom te predstavlja nepovoljan najgori scenarij. Sve veća učestalost takvih godina u posljednjim desetljećima, zajedno s gotovo potpunim podacima za 2025., upućuje na to da ti teški uvjeti postaju redovito obilježje lokalne klime, a ne rijedak događaj.

U praksi povećana učestalost godina s nepovoljnim klimatskim obilježjima naglašava potrebu za boljim upravljanjem vodnim resursima, osobito tijekom dugih sušnih razdoblja i turističke sezone, kada potražnja za vodom raste. Sve veća učestalost i trajanje toplinskih valova ističu

važnost preventivnih mjera u javnom zdravstvu, urbanoj prilagodbi i potrošnji energije, osobito u gusto naseljenim područjima i ljetnim mjesecima. Osim toga, razumijevanje pomaka od stabilnijih prema rizičnijim klimatskim tipovima može pomoći u razvoju lokalnih sustava ranog upozoravanja i planova prilagodbe klimatskim promjenama.

Proširenje analize na dodatne gradove

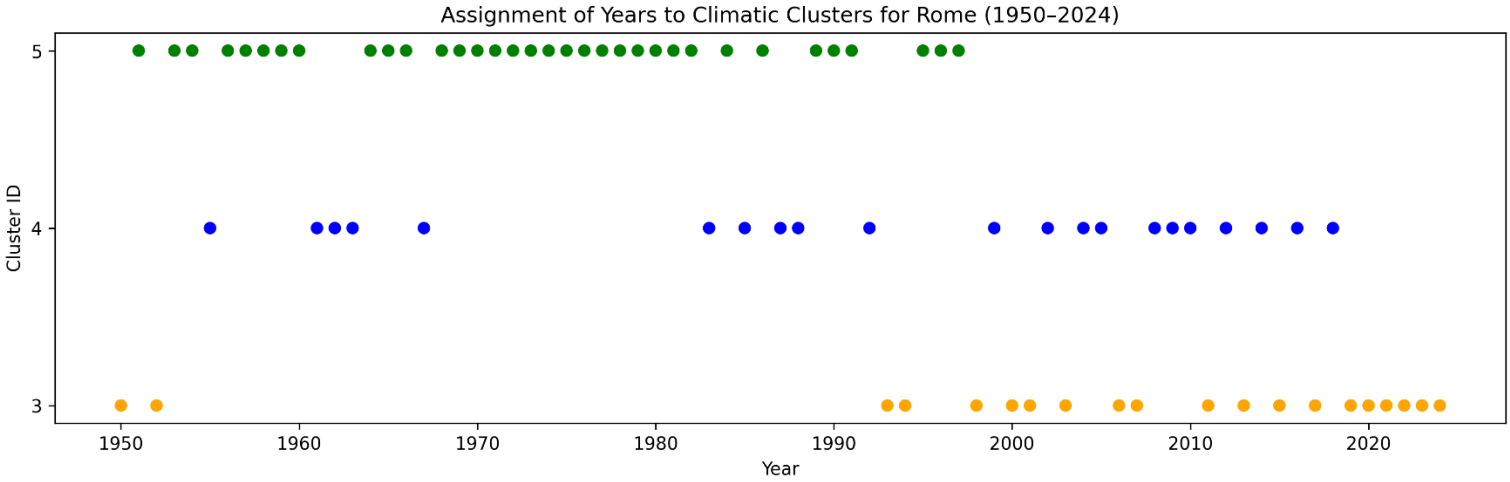
Kako bi se provjerila primjenjivost predloženog pristupa, provedene su dodatne analize za tri grada u različitim geografskim i klimatskim regijama: Rim (Italija), Skopje (Sjeverna Makedonija) i Kijev (Ukrajina).

Za svaki je grad primijenjen isti analitički postupak kao za Šibenik, uključujući odabir meteoroloških varijabli, agregiranje dnevnih mjerenja na godišnju razinu i primjenu algoritma K-means. Međutim, model je treniran zasebno na podacima specifičnima za pojedinu lokaciju. Time se osigurava da identificirani klasteri predstavljaju lokalne klimatske obrasce i odnose među pokazateljima unutar svakog grada.

Rezultati klasterske analize za Rim, Skopje i Kijev prikazani su u pripadajućim tablicama, koje prikazuju prosjeke ključnih klimatskih pokazatelja po klasteru (tablice 3-5), kao i na slikama koje prikazuju raspodjelu klimatskih godina po klasterima za razdoblje 1950.-2024. (slike 5.11-5.13). Iako je metodologija dosljedna u svim slučajevima (kao prethodno za Šibenik), dobiveni se klasteri razlikuju po svojim obilježjima i vremenskim trendovima, što potvrđuje da se klimatski obrasci i rizici razlikuju među regionalnim kontekstima.

Tablica 3. Rezultati K-means klasteriranja klimatskih pokazatelja za Rim

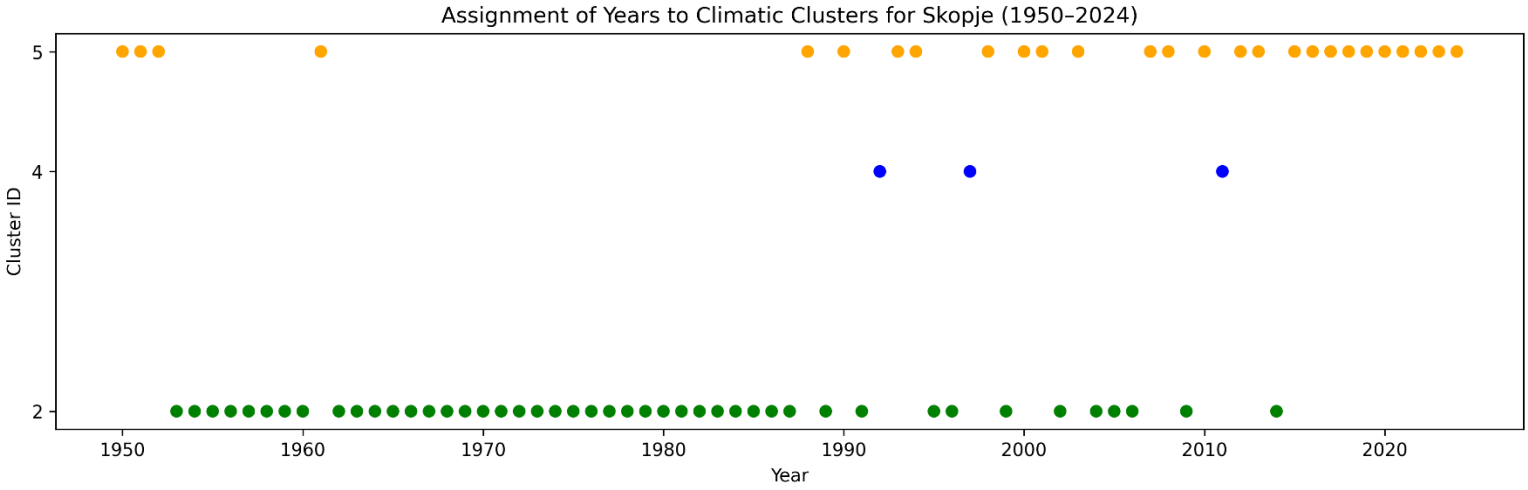
ID klastera	Broj godina u klasteru	Prosječna godišnja temperatura (°C)	Prosječna osjetna temperatura (°C)	Prosječan broj toplinskih valova godišnje	Prosječan broj dana s rizikom od suše	Prosječan broj sušnih dana godišnje	Prosječan broj kišnih dana godišnje	Prosječna najveća brzina vjetra (km/h)
3	20	16	15,3	3	40,3	268,7	96,6	35,3
4	21	15,4	14,8	2,8	15,1	255,5	109,8	34,4
5	34	14,7	14	1,2	5	250,1	115,1	34,4



Slika 5.10. Dodjela godina klimatskim klasterima u Rimu (1950.–2024.)

Tablica 4. Rezultati K-means klasteriranja klimatskih pokazatelja za Skopje

ID klastera	Broj godina u klasteru	Prosječna godišnja temperatura (°C)	Prosječna osjetna temperatura (°C)	Prosječan broj toplinskih valova godišnje	Prosječan broj dana s rizikom od suše	Prosječan broj sušnih dana godišnje	Prosječan broj kišnih dana godišnje	Prosječna najveća brzina vjetra (km/h)
2	45	12,6	11,1	2,6	36,6	269,8	95,4	26,7
4	3	13,1	11,4	2,3	66,7	301,3	64	31
5	27	14,1	12,6	4,6	78,2	285,1	80,1	26,2



Slika 5.11 Dodjela godina klimatskim klasterima u Skopju (1950.–2024.)

Tablica 5. Rezultati K-means klasteriranja klimatskih pokazatelja za Kijev

ID klastera	Broj godina u klasteru	Prosječna godišnja temperatura (°C)	Prosječna osjetna temperatura (°C)	Prosječan broj toplinskih valova godišnje	Prosječan broj dana s rizikom od suše	Prosječan broj sušnih dana godišnje	Prosječan broj kišnih dana godišnje	Prosječna najveća brzina vjetra (km/h)
3	2	7	4,1	0	0	248	117,5	52,5
4	63	7,7	4,8	0	0	243,2	122	39,6
5	10	8,9	6,1	1,2	0	255,6	109,6	37,1

Slika 5.11 Dodjela godina klimatskim klasterima u Kijevu (1950.–2024.)

Analiza dodatnih gradova pokazuje da je u Rimu, Skopju i Kijevu posljednjih godina došlo do pomaka prema godinama s izraženijim ekstremnim klimatskim obilježjima. Međutim, ti se obrasci razlikuju. U sva tri grada povećala se učestalost godina s višim temperaturama i većim toplinskim stresom. Suprotno tome, stabilniji klimatski obrasci češće se pojavljuju u starijem dijelu podatkovnog niza.

Osobito je uočljivo da su od početka 2000-ih, a još više nakon 2010., ekstremniji klasteri postali češći u kontinentalnim gradovima poput Skopja i Kijeva, što upućuje na širenje toplinskog stresa izvan tipičnih mediteranskih regija. Iako se intenzitet i kombinacije ekstremnih uvjeta razlikuju od onih u obalnim gradovima, nalazi upućuju i na strukturni pomak klimatskih obrazaca u tim područjima.

Ovi nalazi potvrđuju da identificirani obrasci nisu samo lokalne anomalije, nego dio šireg trenda klimatskih promjena, pri čemu se ekstremni scenariji postupno pojavljuju u različitim klimatskim zonama. To naglašava potrebu za lokalno prilagođenim, ali dosljedno primijenjenim analitičkim metodama.

Zadaci za vježbu

- Preuzmite povijesne dnevne vremenske podatke za odabranu lokaciju koristeći se npr. Open-Meteo API-jem.
- Agregirajte dnevne meteorološke podatke u godišnje klimatske pokazatelje (npr. prosječnu temperaturu, broj vrlo vrućih dana, broj sušnih dana).
- Uspostavite godišnje klimatske profile na temelju odabranih pokazatelja.
- Primijenite algoritam K-means za grupiranje klimatskih godina prema njihovim obilježjima.
- Dodijelite svaku godinu njezinu klasteru i prikažite rezultate u tabličnom ili grafičkom obliku.
- Interpretirajte dobivene klastere kao različite klimatske scenarije (stabilan, prijelazan, nepovoljan).
- Usporedite novije podatke s prethodno identificiranim klasterima i procijenite kojem tipu klimatske godine pripadaju.

Provjera znanja

Koji je primarni cilj uporabe algoritma K-means u prikazanoj studiji slučaja?

A. Numeričko prognoziranje buduće temperature
B. Grupiranje klimatskih godina prema njihovim sličnostima

C. Izračun sezonskih prosjeka meteoroloških varijabli
D. Otkrivanje pojedinačnih ekstremnih dana

Koja je svrha agregiranja dnevnih meteoroloških podataka na godišnju razinu?

A. Smanjenje količine podataka bez gubitka informacija
B. Lakše uočavanje dugoročnih obrazaca i klimatskih tipova
C. Uklanjanje sezonalnosti iz podataka
D. Priprema podataka za nadzirano učenje

Koji skup pokazatelja najbolje opisuje nepovoljan (ekstremni) klimatski scenarij identificiran u studiji slučaja?

A. Visoke temperature, mnogo kišnih dana, slab vjetar
B. Umjerene temperature i rijetki toplinski valovi
C. Visoke temperature, česti toplinski valovi i velik broj dana s kombiniranim sušnim stresom
D. Velika količina oborina i kratkotrajni ekstremi

Zašto se klaster identificiran nakon 2020. u slučaju grada Šibenika smatra pokazateljem strukturne promjene klimatskih uvjeta?

A. Zato što se temelji samo na jednoj ekstremnoj godini
B. Zato što uključuje nove kombinacije toplinskog i hidrološkog stresa
C. Zato što se pojavljuje gotovo isključivo u novijem razdoblju
D. Zato što je unaprijed definiran pragovima

Koja tvrdnja najbolje opisuje ulogu algoritma K-means u kontekstu scenarijskog prognoziranja okolišnih rizika?

A. K-means izravno predviđa buduće klimatske uvjete kao scenarije
B. K-means identificira tipične obrasce koji se mogu interpretirati kao scenariji
C. K-means koristi klasičnu meteorološku analizu kao scenarije
D. K-means se koristi za kratkoročne prognoze

Dodatni izvori

[1] G. James, D. Witten, T. Hastie i R. Tibshirani, An Introduction to Statistical Learning: With Applications in R, 2. izd. New York, NY, SAD: Springer, 2021. [Online]. Dostupno: https://hastie.su.domains/ISLR2/ISLRv2_website.pdf, pristupljeno: 15. prosinca 2025.

[2] Lindholm, N. Wahlström, F. Lindsten i T. B. Schön, Machine Learning – A First Course for Engineers and Scientists. Cambridge, U.K.: Cambridge University Press, 2022. [Online]. Dostupno: http://smlbook.org/book/sml-book-draft-latest.pdf, pristupljeno: 15. prosinca 2025.

[3] O. Simeone, A brief introduction to machine learning for engineers, arXiv preprint, arXiv:1709.02840, 2017. [Online]. Dostupno: https://arxiv.org/pdf/1709.02840.pdf, pristupljeno: 15. prosinca 2025.

[4] Historical Weather API Documentation, 2024. Dostupno: https://open-meteo.com/en/docs/historical-weather-api, pristupljeno: 15. prosinca 2025.

[5] Oracle, Autonomous AI Lakehouse. [Online]. Dostupno: https://www.oracle.com/autonomous-database/autonomous-ai-lakehouse, pristupljeno: 15. prosinca 2025.

[6] Oracle, DBMS_DATA_MINING Package Reference, Oracle Database 19c Documentation. [Online]. Dostupno: https://docs.oracle.com/en/database/oracle/oracle-database/19/arpls/DBMS_DATA_MINING.html, pristupljeno: 15. prosinca 2025.