

Jure Arsovič, Matic Klopčič in Karla Gliha

EKSTRAKCIJA PODATKOV

Druga naloga pri predmetu Iskanje in ekstrakcija podatkov s spleta

MENTOR: prof. dr. Marko Bajec in asistent Timotej Knez

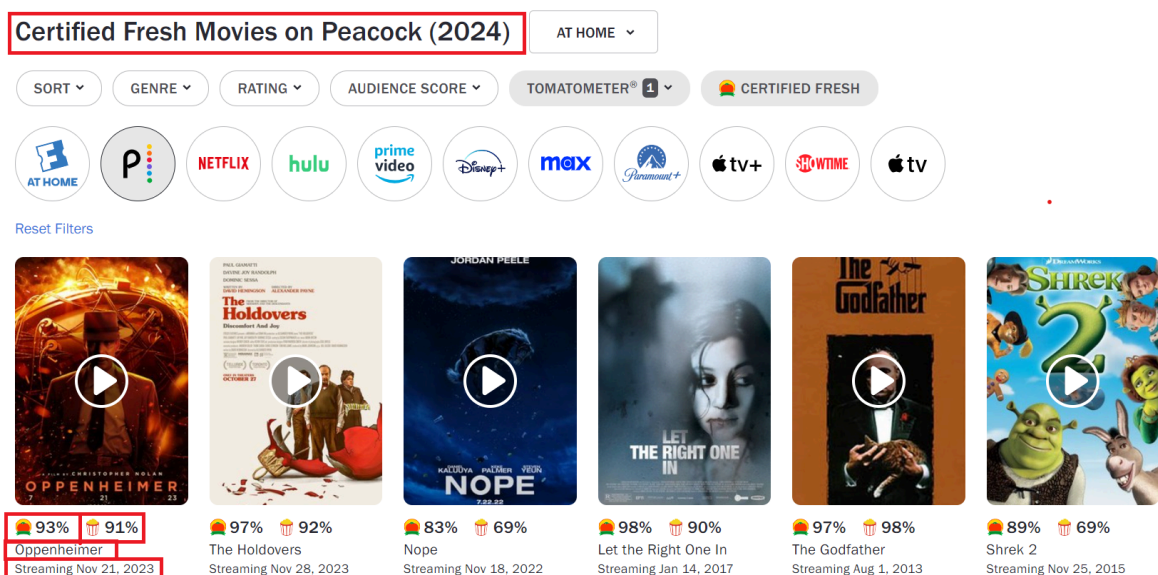
1 UVOD

Druga domača naloga pri predmetu Iskanje in ekstrakcija podatkov s spleta je predstavljala implementacijo metod za pridobivanje podatkov s spleta. Potrebno je bilo razumevanje strukture spletnih strani in implementacija treh različnih načinov ekstrakcije podatkov, s katerimi smo iz spletnih strani izluščili pomembne podatke. Prvi dve metodi podatke pridobivata neposredno, zadnja pa avtomatsko.

2 IZBRANI SPLETNI STRANI

Za tretji tip spletne strani smo si izbrali strani iz domene rottentomatoes.com, ki prikazujejo seznam najbolj ocenjenih filmov na specifičnih streaming storitvah, v našem primeru Amazon Prime in Peacock. Strani smo izbrali zato, ker obe vsebujeta strukturiran seznam, a se popolnoma razlikujeta v vsebini seznama. Podatki ki jih je bilo treba pridobiti iz njih so:

- Naslova obeh strani
- Naslov vsakega filma v seznamu
- Datum izdaje oziroma začetka predvajanja vsakega filma iz seznama
- Ocena publike za vsak film v seznamu
- Ocena kritikov za vsak film v seznamu



Slika 1: Z rdečo so označeni podatki, ki smo jih iz strani pridobili

3 IMPLEMENTACIJA

3.1 REGEX

Pri implementaciji regexa smo bili pozorni, da smo implementirali take regularne izraze, da bo isti izraz pridobil podatke iz obeh strani, ne glede na razlike v njuni strukturi. Za vsak podatek smo uporabili izključno en regularni izraz, nikoli nismo uporabili “nested” izrazov, da bi izrazili katerikoli podatek.

OVERSTOCK	
Title	([^\<]+)
List price	<s>([^\<]+)</s>
Price	<td align="right" nowrap="nowrap">Price:</td><td align="left" nowrap="nowrap">([^\<]+)</td>
Saving *	<td align="right" nowrap="nowrap">You Save:</td><td align="left" nowrap="nowrap">([^\<]+)</td>
Content	<td valign="top">([^\<]+) <a href
RTV	
Title	<h1[^\>]*>(.*?)</h1>
Author/ published time **	([^\<]+)\s*([d]{1,2}\s*[a-zA-Z]+\s*[d]{4}\s*ob\s*[d]{1,2}:[d]{2})
Subtitle	<div class="subtitle">([^\<]+)</div>
Lead	<p class="lead">([^\<]+)</p>
Content	^(?..*)<div class="article-body"> <div class="gallery">(?!.*\$)<script[^\>]*>.*?</script> <[^\>]+>
ROTTEN TOMATOES	
Title	<title>([^\<]+)</title>
Film title	\s*(.*?)\s*\s*
Scores ***	<score-pairs-deprecated.*?audiencescore="(\d+)".*?criticsscore="(\d+)".*?</score-pairs-deprecated>\s*
Streaming from	\s*(.*?)\s*

* Opomba pri saving (overstock): podatke pridobljene z regularnim izrazom in XPathom prikazanom pod saved smo potem naknadno razdelili na “saved price” in “saved percentage” s .split(“ “)[0] in .split(“ “)[1].

** Opomba pri author/published time (rtv): podatke pridobljene z regularnim izrazom prikazanom pod author/published time smo potem naknadno razdelili na “author” in “published time” z .group(1) in .group(2) saj smo z re.search te dva podatka razdelili v dve grupi.

*** Opomba pri scores (rotten tomatoes): podatke pridobljene z regularnim izrazom prikazanom pod scores smo potem naknadno razdelili na “critics score” in “audience score”.

3.2 XPATH

Pri XPath smo najprej vsako stran sparsali z ukazom `tree = html.fromstring(html_page)`, nato pa posamezne elemente ekstrahirali z uporabo spodnjih XPathov kot argument funkcije `tree.xpath(“XPath”)`.

OVERSTOCK	
Title	"/html/body/table[2]/tbody/tr[1]/td[5]/table/tbody/tr[2]/td/table/tbody/tr/td/table/tbody//a/b/text()
List price	'//td[b[text()="List Price:"]]/following-sibling::td/s/text()'
Price	'//td[b[text()="Price:"]]/following-sibling::td/span/b/text()'
Saving *	'//td[b[text()="You Save:"]]/following-sibling::td/span/text()'
Content	'//span[@class="normal"]/text()'
RTV	
Title	'//header[@class="article-header"]/h1/text()'
Author	'//div/div[3]/div/header/div[3]/div[1]/strong/text()'
Published time	'normalize-space(substring-after(//div[@class="author-timestamp"]/strong/following-sibling::text()[1], " '))'
Subtitle	'//div[@class="subtitle"]/text()'
Lead	'//p[@class="lead"]/text()'
Content	'//article[@class="article"]//p/text()'
ROTTEN TOMATOES	
Page title	'/html/body/div[3]/main/div[1]/div[1]/div[1]/discovery-title/h1/text()'
Film title	'//span[@data-qa="discovery-media-list-item-title"]/text()'

Critics score	'//score-pairs-deprecated/@criticsscore'
Audience score	'//score-pairs-deprecated/@audiencescore'
Streaming from	'//span[@data-qa="discovery-media-list-item-start-date"]/text()'

3.3 ROADRUNNER

Za tretjo metodo smo izbrali implementacijo algoritma RoadRunner, sledili smo korakom opisane implementacije [1]. V implementaciji metode RoadRunner smo najprej pripravili HTML vsebine z obdelavo v BeautifulSoup objekt in odstranili neželene elemente za poenostavitev analize. Nato smo HTML dokumente tokenizirali za lažjo primerjanje struktur. Glavna funkcija `road_runner` je primerjala te žetone, iskala skupne vzorce in obdelovala razlike, da identificira ujemajoče se strukture med dokumenti. Končni rezultati so bili shranjeni in oblikovani za jasen izpis, kar nam omogoča robustno pridobivanje podatkov iz kompleksnih in raznolikih spletnih strani.

3.3.1 PSEVDOKODA

Class RoadRunner:

Method `extract_data(html1, html2)`:

- Prepare soup objects from `html1` and `html2`
- Call `road_runner` on the soup objects to find common patterns
- Store the result in a dictionary '`extracted_data`' with key '`wrapper`'
- Write the result to a file using `pretty_format` method

Method `pretty_format(wrapper, indent)`:

- Initialize variables for formatting the HTML content for readability
- Split the wrapper content into tokens
- Process each token to handle indentation and formatting based on tag types and text content

Method `prepare_soup(html_content)`:

- Create a BeautifulSoup object with `html_content`
- Remove unwanted elements from the soup by calling `decompose_unwanted`

Method `decompose_unwanted(soup)`:

- Remove tags specified in `ignore_tags` and other specific tags like '`script`', '`style`', '`meta`'
- Remove comments from the soup

Method `tokenize_and_compare(html1, html2)`:

- Tokenize `html1` and `html2`
- Call `road_runner` with the tokens to compare and find common patterns

Method `tokenize(html)`:

- Parse HTML to create a list of tokens representing the structure of the document
- Iterate over elements in HTML and generate tokens for each element, including tags and text

Method `road_runner(tokens1, tokens2)`:

Compare the HTML structures represented by tokens1 and tokens2 to identify common patterns

Handle mismatches, direct text comparisons, and nested tag structures recursively

4 ZAKLJUČEK

Pri drugi domači nalogi smo se naučili kako ter na kakšne načine lahko pridobivamo podatke s spleta. Z implementacijo prvih dveh metod, regex in XPath, smo bolje spoznali samo strukturo in obliko spletnih strani, kar nam je pomagalo pri oblikovanju izrazov za ekstrakcijo željenih podatkov. Pri zadnji metodi, implementaciji algoritma RoadRunner, smo se poglobili v dinamično primerjavo in analizo struktur HTML, kar nas je naučilo, kako obvladovati bolj kompleksne in manj strukturirane podatkovne vire. Ta domača naloga nam je tako ponudila celovit vpogled v različne tehnike in orodja za pridobivanje podatkov ter izboljšala naše tehnične sposobnosti in razumevanje podatkovnih struktur na spletu.

5 LITERATURA

[1] V. Crescenzi, G. Mecca, and P. Merialdo, "RoadRunner: Towards Automatic Data Extraction from Large Web Sites," Proceedings of the 27th VLDB Conference, Roma, Italy, 2001.

6 IZHODI PROGRAMA

6.1 REGEX

6.1.1 RTV

```
{
  "RTV1": [
    {
      "TITLE": "Audi A6 50 TDI quattro: nemir v premijskem razredu",
      "AUTHOR": "Miha Merljak",
      "PUBLISHED_TIME": "28. december 2018 ob 08:51",
      "SUBTITLE": "Test nove generacije",
      "LEAD": "To je novi audi A6. V razred najdražjih in najbolj premijskih žrebcev je vnesel nemir, še preden je sploh zapeljal na parkirni prostor, rezerviran za izvršnega direktorja. ",
      "CONTENT": "Audi A6 je avtomobil za direktorje, ki se mu na avtocesti spoštljivo umikajo, kot bi šlo za Pahorjev ali Šarčev avtomobil. Foto: David Šavli Oglas Samo pogledjte njegovo masko – to ogromno satovje z radarji na takem položaju, da se ti na avtocesti tudi pri 120 km/h vsi spoštljivo umikajo, saj so prepričani, da gre za Pahorjev ali Šarčev avto. Seveda, novi A6 lahko cesto in promet skenira s kar petimi radarji, petimi kamerami, infrardečo kamero za nočni vid, dvanajstimi ultrazvočnimi senzorji in laserskim čitalnikom – lidarjem. V glavnem vojaška tehnologija v službi varnosti za fante, ki smo radi gledali Top Gun, Bonda in druge možakarja s finimi igračkami. Novo poglavje Vozniški delovni prostor je novo poglavje digitalne dobe, z dvema ogromnima zaslonoma, ki tako kot naprednejši telefoni dregnejo blazinice vaših prstov, kot se sprehajate po steklu. A še bolj se nam zdi
```

pomembno, da so osnovna stikala tam, kjer jih pričakujete. Najprej so torej zagotovili enostavno osnovo, tisti bolj \"advanced\" vozniki pa si lahko nato vse skupaj še veliko bolj prilagodijo. Velik korak naprej pri kabinskem udobju zaznavajo tudi na zadnji klopi, tam je prostora v vseh smereh precej več. Če vam pogled na Audijev spisek dodatne opreme ne odvzame volje do življenja, potem vsekakor toplo priporočamo nakup zračnega vzmetenja, saj dobi z njim A6 več različnih in vozniško zelo uporabnih karakterjev. Enako velja za seksi luči z inteligentno matrično osvetlitvijo, pa za športno podvozje in vsekakor za štirikolesno krmiljenje. S tem postane A6 med ovinki v občutku na volanu še veliko krajši in bolj agilen. Vse naštetu smo preskušali v družbi agregata 50 TDI, ki je v resnici klasični trilitrski dizel, podkrepjen z elektromotorjem. Ja, ta audi je mehki hibrid z izjemnim navorom in dovolj moči kadar koli in kjer koli. Si pa mislimo, da bo največji del trga zadovoljil že učinkovit dvolitrski mehki hibrid z močjo 150 kilovatov. Ključni tehnični podatki:- na testu Audi A6 50 TDI quattro tiptronic

Mere:- dolžina: 4,9 m- medosna razdalja: 2,9 m- obračalni krog: 12,1 m- prtljažnik: 530 l- masa: 1.900 kg

Pogon:- trilitrski šestvaljni dizelski motor- moč: 210 kW- navor: 620 Nm- 8-stopenjski samodejni menjalnik- pogon na vsa štiri kolesa- pnevmatike: 225/60 R17- poraba: 6,6 l/100 km = 8,9 EUR/100 km- posoda za gorivo: 73 l- doseg: 1.106 km- izpusti CO2: 147 g/km

Stroški pri 15.000 km in 5-letni uporabi:- nakupna cena: 69.080 EUR- stroški finančnega lizinga: 4.463 EUR/5 let- stroški registracije: 10.829 EUR/5 let- stroški vzdrževanja: 1.926 EUR/5 let- stroški goriva: 6.702 EUR/75.000 km- strošek 1 kompleta pnevmatik: 716 EUR- vrednosti po 5 letih po Eurotaxu: 33.964 EUR- stroški skupaj: 1.001 EUR/mesec"

```

    }
  ],
  "RTV2": [
    {
      "TITLE": "Volvo XC 40 D4 AWD momentum: suvereno med najboljše v razredu",
      "AUTHOR": "Miha Merljak",
      "PUBLISHED_TIME": "25. januar 2019 ob 15:23",
      "SUBTITLE": "Test novega modela",
      "LEAD": "XC 40 je najmanjši Volvov SUV, ki se oblikovno skoraj v celotni naslanja na oba večja predhodnika. Že samo s tem so mu vrata do denarnic tistih kupcev, ki iščejo izstopajočo, a hkrati visoko kultivirano in prečiščeno dizajnersko govorico, na pol odprta.",
      "CONTENT": "XC40 se v mestu odlično zlije z okolico. Foto: David Šavli Volvo se je nižjih srednjih razredov v preteklosti izogibal ali pa je vanje vstopal z zelo nišnimi produkti, ki niso pustili večjega tržnega pečata. V primeru XC 40 ni težko napovedati, da bo ta tradicija prekinjena. Ponuja namreč visoko kakovost končne izdelave in v kabini odlično premišljeno funkcionalnost ter na dotik prijetne materiale. Še posebej hvalimo število, iznajdljivost in velikost različnih odlagalnih prostorov ter široke, čvrste in zelo udobne sedeže. Intuitivno in enostavno logično je upravljanje z velikim vmesnikom, ki z večfunkcijskim zaslonom na dotik kraljuje na z roko lahko dostopnem mestu na sredinski armaturi. Razočaranj ne bo niti v velikosti in uporabnosti prtljažnega prostora, ki s 460 litri prostornine sicer ni med večjimi v razredu, a se v uporabniškem smislu odkupi z dobro urejenostjo ter domiselnimi rešitvami pregrajevanja. Foto: David Šavli XC 40 je od tal odmaknjen konkretnih 21 cm, a sta vzmetenje in krmilni mehanizem tako nastavljena, da ponuja tudi v hitro odpeljanih ovinkih zelo dolgo nevtrarno in predvidljivo lego. V premeru preskušane modela, ki je imel v paketu R design vzmetenje še nekoliko bolj trdo, se je to samo še bolj potrdilo, a je v tem primeru treba računati na manj udobno vožnjo čez različne asfaltne grbine. Podoben razmislek velja opraviti tudi pri izbiri motorja. Preskušani 2-litrski dizel s 190 KM predstavlja vrh ponudbe, ki z močjo, udobjem in tudi povprečno porabo navduši predvsem pri avtocestnih dolgoprogaških izzivih, v počasni mestni vožnji ter pri
    }
  ]
}

```

pogostih postankih in speljevanjih pa deluje preveč robusten. XC 40 je s čvrsto gradnjo, funkcionalno in udobno kabino ter številnimi asistenčnimi sistemi in izstopajočim skandinavskim dizajnom v premišljenem trenutku vstopil na trg modnih mestnih terencev, v katerem se brez ene same sence dvoma suvereno postavi med najdražje in najbolj premijske v mestu. Ključni tehnični podatki:- na testu Volvo XC40 2.0 TD avt awd momentum Mere:- dolžina: 4,4 m- medosna razdalja: 2,7 m- obračalni krog: 11,4 m- oddaljenost od tal: 21 cm- prtljažnik: 432 l- masa: 2.250 kg Pogon:- 2-litrski 4-valjni bencinski motor- moč: 140 kW- navor: 400 Nm- 8-stopenjski samodejni menjalnik- pogon na vsa štiri kolesa- pnevmatike: 235/50 R19 - poraba: 6,3 l/100 km = 8,2 EUR/100km- posoda za gorivo: 54 l- doseg: 857 km- izpusti CO2: 133 g/km Stroški pri 15.000 km in 5-letni uporabi:- nakupna cena: 43.619 EUR- stroški finančnega leasinga: 3.268 EUR/5 let- stroški registracije: 8.701 EUR/5 let- stroški vzdrževanja: 2.320 EUR/5 let- stroški goriva: 6.190 EUR/75.000 km- strošek 1 kompleta pnevmatik: 923 EUR- vrednosti po 5 letih po Eurotaxu: 18.886 EUR- stroški skupaj: 774 EUR/mesec"

```

    }
  ]
}

```

6.1.2 OVERSTOCK

```

{
  "OVS1": [
    {
      "TITLE": "10-kt. Seven Diamond Ladies Heart Ring (0.08 TW)",
      "LIST_PRICE": "$149.00",
      "PRICE": "$69.99",
      "SAVED_AMOUNT": "$79.01",
      "SAVED_PERCENTAGE": "(53%)",
      "CONTENT": "This ladies fashion ring dazzles with hearts and diamonds. The gold band is crafted into delicate, open hearts. Seven brilliant-cut diamonds add a bit of sparkle. "
    },
    {
      "TITLE": "10-Kt. Diamond Ring (.25 TW)",
      "LIST_PRICE": "$250.00",
      "PRICE": "$74.90",
      "SAVED_AMOUNT": "$175.10",
      "SAVED_PERCENTAGE": "(70%)",
      "CONTENT": "Nineteen round diamonds accent this 10-karat yellow gold ring with filigree accents."
    },
    {
      "TITLE": "10-kt. Pearl and Diamond Butterfly Earrings",
      "LIST_PRICE": "$149.00",
      "PRICE": "$42.99",
      "SAVED_AMOUNT": "$106.01",
      "SAVED_PERCENTAGE": "(71%)",
      "CONTENT": "Perfectly proportioned 5.5- to 6-mm cultured pearls on 10-karat yellow gold settings highlight these petite earrings. A dainty rhodium-plated gold butterfly studded with a diamond (0.02 total carat weight, J-K color, I-2 clarity) rests atop each pearl."
    }
  ]
}

```

```

    },
    {
      "TITLE": "14-kt. Diamond 'S' Tennis Bracelet (2.00 TW)",
      "LIST_PRICE": "$1,539.99",
      "PRICE": "$499.99",
      "SAVED_AMOUNT": "$1,040.00",
      "SAVED_PERCENTAGE": "(67%)",
      "CONTENT": "Invest in a swirl of light with this diamond 'S' tennis bracelet. Crafted
in 14-karat gold, the piece features 49 diamonds for two full carats. The 7.25-inch bracelet
closes with a pressure clasp."
    },
    {
      "TITLE": "10-kt. Diamond Band Fashion Ring (.11 TW)",
      "LIST_PRICE": "$179.99",
      "PRICE": "$79.99",
      "SAVED_AMOUNT": "$100.00",
      "SAVED_PERCENTAGE": "(55%)",
      "CONTENT": "Crafted in white and yellow gold, this ring displays a band of seven
round diamonds. Order your new gold and diamond fashion ring today at our low, online
price."
    },
    {
      "TITLE": "14-kt. White Gold, Pearl and Diamond Ring",
      "LIST_PRICE": "$419.99",
      "PRICE": "$149.99",
      "SAVED_AMOUNT": "$270.00",
      "SAVED_PERCENTAGE": "(64%)",
      "CONTENT": "Show your romantic side with this 14-karat diamond and pearl ring.
Set in a domed band of 14-karat white gold, the ring features a 7-mm cultured pearl. Curved
rows of diamonds flank the pearl."
    },
    {
      "TITLE": "14-kt. Gold Diamond Present Future Pendant (.25TW)",
      "LIST_PRICE": "$299.00",
      "PRICE": "$149.99",
      "SAVED_AMOUNT": "$149.01",
      "SAVED_PERCENTAGE": "(49%)",
      "CONTENT": "Designed with three large, sparkling diamonds to represent past,
present, and future, this stunning pendant is set in gleaming 14-karat gold. It incorporates a
total of nine diamonds (0.25 total carat weight, K color, I-2 to I-3 clarity). "
    },
    {
      "TITLE": "14-kt. Diamond Solitaire Pendant (.33 TW)",
      "LIST_PRICE": "$1,019.99",
      "PRICE": "$319.99",
      "SAVED_AMOUNT": "$700.00",
      "SAVED_PERCENTAGE": "(68%)",
      "CONTENT": "In this simple, yet elegant pendant, a round brilliant diamond (0.33
total carat weight, H-J color, I-1 to I-2 clarity) is prong-set in 14-karat white gold."
    },
  },

```



```

{
  "TITLE": "14-kt. Diamond Solitaire Earrings (0.33 TW)",
  "LIST_PRICE": "$639.99",
  "PRICE": "$199.99",
  "SAVED_AMOUNT": "$440.00",
  "SAVED_PERCENTAGE": "(68%)",
  "CONTENT": "Dazzle your way into her heart, with these classic diamond solitaire earrings. Two brilliant-cut diamonds (0.33 total carat weight, G-H color, I-1 to I-2 clarity) are set in four prongs of 14-karat white gold."
},
{
  "TITLE": "14-kt. Diamond Cross Pendant (.06 TW)",
  "LIST_PRICE": "$305.00",
  "PRICE": "$119.99",
  "SAVED_AMOUNT": "$185.01",
  "SAVED_PERCENTAGE": "(60%)",
  "CONTENT": "Over a cleanly sculpted Roman cross of 14-karat white gold drapes a slender banner containing three bright prong-set round diamonds (0.06 total carat weight, H-I color, I clarity)."
},
{
  "TITLE": "14-kt. Diamond Solitaire Stud Earrings (.50 TW)",
  "LIST_PRICE": "$999.99",
  "PRICE": "$359.99",
  "SAVED_AMOUNT": "$640.00",
  "SAVED_PERCENTAGE": "(64%)",
  "CONTENT": "Every jewelry collection needs a classic pair of diamond solitaire earrings. Set in 14-karat gold, these diamond stud earrings (0.50 total carat weight) have post backs with butterfly clasps."
},
{
  "TITLE": "14-kt. Cultured Pearl Diamond Earrings",
  "LIST_PRICE": "$508.99",
  "PRICE": "$179.99",
  "SAVED_AMOUNT": "$329.00",
  "SAVED_PERCENTAGE": "(64%)",
  "CONTENT": "Create an elegant appearance with these pearl and diamond stud earrings. Set in 14-karat yellow gold, each earring features an 8 to 8.5-mm cultured white pearl. Prong-set round diamonds accent the pearls. Posts with butterfly clasps secure the earrings."
},
{
  "TITLE": "14-kt. Diamond 7.5-8 mm Pearl Pendant",
  "LIST_PRICE": "$196.99",
  "PRICE": "$69.99",
  "SAVED_AMOUNT": "$127.00",
  "SAVED_PERCENTAGE": "(64%)",
  "CONTENT": "Add a classic to your jewelry collection with this 14-karat gold, diamond, and pearl necklace. The 7.5-8 mm cultured white pearl creates the focal point of the pendant, while a diamond (0.10 TW) adds sparkle."
}

```

```

    },
    {
      "TITLE": "14-kt. Diamond Solitaire Earrings (.50 TW)",
      "LIST_PRICE": "$1,369.99",
      "PRICE": "$409.99",
      "SAVED_AMOUNT": "$960.00",
      "SAVED_PERCENTAGE": "(70%)",
      "CONTENT": "This earring set has two brilliant-cut diamonds (0.50 total carat weight, G-H color, I-1 to I-2 clarity) set in four prongs of 14-karat white gold."
    },
    {
      "TITLE": "14-kt White Gold Diamond Band (0.50 TW)",
      "LIST_PRICE": "$1,635.00",
      "PRICE": "$609.99",
      "SAVED_AMOUNT": "$1,025.01",
      "SAVED_PERCENTAGE": "(62%)",
      "CONTENT": "Crafted of 14-karat white gold, this stylish ring features a bright row of 20 channel-set, princess-cut baguette diamonds. Treat her like royalty and save when you buy jewelry treasures at Overstock.com."
    }
  ],
  "OVS2": [
    {
      "TITLE": "14-kt. Green Jade Hoops",
      "LIST_PRICE": "$90.00",
      "PRICE": "$46.99",
      "SAVED_AMOUNT": "$43.01",
      "SAVED_PERCENTAGE": "(47%)",
      "CONTENT": "Hoops of cool green jade rest between 14-karat yellow gold endpieces. The hoops graduate in thickness from 3 mm at the ends to 6 mm in the center, with approximately 29 mm overall diameter."
    },
    {
      "TITLE": "14-kt. Jade Doughnut Pendant",
      "LIST_PRICE": "$150.00",
      "PRICE": "$48.99",
      "SAVED_AMOUNT": "$101.01",
      "SAVED_PERCENTAGE": "(67%)",
      "CONTENT": "The 25-mm disk hangs delicately from a 14-karat gold chain. The disk features a dramatic gold Chinese character in the center, accompanied by four stylized gold bees."
    },
    {
      "TITLE": "14-kt. Charcoal Jade and Ruby Elephant Pendant",
      "LIST_PRICE": "$100.00",
      "PRICE": "$28.99",
      "SAVED_AMOUNT": "$71.01",
      "SAVED_PERCENTAGE": "(71%)",

```

"CONTENT": "Carved of rich dark grey jade, this elephant pendant has 14-karat yellow gold applied to mark the feet, tusk, tail, and blanket. A 2-mm round faceted ruby in a gold bezel setting forms the eye. The pendant hangs from an 18-inch chain."

},

{

"TITLE": "14-kt. Carved Lavender Jade Earrings",

"LIST_PRICE": "\$80.00",

"PRICE": "\$39.99",

"SAVED_AMOUNT": "\$40.01",

"SAVED_PERCENTAGE": "(50%)",

"CONTENT": "Luscious 8-mm lavender jade balls, carved with intricate Asian style, dangle from a 14-karat yellow gold French hook."

},

{

"TITLE": "14-kt. Jade Cross Pendant",

"LIST_PRICE": "\$150.00",

"PRICE": "\$49.99",

"SAVED_AMOUNT": "\$100.01",

"SAVED_PERCENTAGE": "(66%)",

"CONTENT": "Green jade and gold create this beautiful cross pendant. Cylindrical bars of green jade feature caps and center of 14-karat yellow gold."

},

{

"TITLE": "14-kt. Multicolored Jade Earrings",

"LIST_PRICE": "\$375.00",

"PRICE": "\$99.99",

"SAVED_AMOUNT": "\$275.01",

"SAVED_PERCENTAGE": "(73%)",

"CONTENT": "A delicate wrapping of 14-karat yellow gold wire holds six 6 x 4 pear shapes of jade in various shades: brilliant green, orange, lavender, black, pale yellow, and white. The post earrings have butterfly backs."

},

{

"TITLE": "14-kt. Multicolored Jade Ring",

"LIST_PRICE": "\$250.00",

"PRICE": "\$56.99",

"SAVED_AMOUNT": "\$193.01",

"SAVED_PERCENTAGE": "(77%)",

"CONTENT": "A delicate wrapping of 14-karat yellow gold wire holds six 6 x 4 ovals of jade in various shades: brilliant green, orange, lavender, black, pale yellow, and white. A narrow gold band divides to support the setting."

},

{

"TITLE": "14-kt. Onyx and Ruby Elephant Pendant",

"LIST_PRICE": "\$100.00",

"PRICE": "\$35.99",

"SAVED_AMOUNT": "\$64.01",

"SAVED_PERCENTAGE": "(64%)",

```

    "CONTENT": "Carved of rich black onyx, this elephant pendant has 14-karat yellow
gold applied to mark the feet, tusk, tail, and blanket. A 2-mm round faceted ruby in a gold
bezel setting forms the eye. The pendant hangs from an 18-inch chain."
  }
]
}

```

6.1.3 ROTTEN TOMATOES

Title1: Certified Fresh Movies on Peacock (2024) | Rotten Tomatoes

```

[
  {
    "title": "Oppenheimer",
    "streaming_date": "Streaming Nov 21, 2023",
    "audience_score": "91%",
    "critic_score": "93%"
  },
  {
    "title": "The Holdovers",
    "streaming_date": "Streaming Nov 28, 2023",
    "audience_score": "92%",
    "critic_score": "97%"
  },
  {
    "title": "Nope",
    "streaming_date": "Streaming Nov 18, 2022",
    "audience_score": "69%",
    "critic_score": "83%"
  },
  {
    "title": "Let the Right One In",
    "streaming_date": "Streaming Jan 14, 2017",
    "audience_score": "90%",
    "critic_score": "98%"
  },
  {
    "title": "The Godfather",
    "streaming_date": "Streaming Aug 1, 2013",
    "audience_score": "98%",
    "critic_score": "97%"
  },
  {
    "title": "Dredd",
    "streaming_date": "Streaming Sep 2, 2014",
    "audience_score": "72%",
    "critic_score": "80%"
  },
  {

```

```
"title": "Memento",
"streaming_date": "Streaming Mar 1, 2017",
"audience_score": "94%",
"critic_score": "94%"
},
{
  "title": "Shrek 2",
  "streaming_date": "Streaming Nov 25, 2015",
  "audience_score": "69%",
  "critic_score": "89%"
},
{
  "title": "Shrek",
  "streaming_date": "Streaming Nov 25, 2015",
  "audience_score": "90%",
  "critic_score": "88%"
},
{
  "title": "John Wick",
  "streaming_date": "Streaming Jun 7, 2016",
  "audience_score": "81%",
  "critic_score": "86%"
},
{
  "title": "Jackie Brown",
  "streaming_date": "Streaming Dec 17, 2015",
  "audience_score": "85%",
  "critic_score": "88%"
},
{
  "title": "Spider-Man",
  "streaming_date": "Streaming Apr 25, 2013",
  "audience_score": "67%",
  "critic_score": "90%"
},
{
  "title": "Noah",
  "streaming_date": "Streaming Apr 9, 2016",
  "audience_score": "41%",
  "critic_score": "75%"
},
{
  "title": "Promising Young Woman",
  "streaming_date": "Streaming Jan 15, 2021",
  "audience_score": "87%",
  "critic_score": "90%"
},
{
  "title": "The Godfather, Part II",
  "streaming_date": "Streaming Sep 15, 2012",
```

```
"audience_score": "97%",
"critic_score": "96%"
},
{
  "title": "The Big Lebowski",
  "streaming_date": "Streaming Dec 1, 2015",
  "audience_score": "93%",
  "critic_score": "80%"
},
{
  "title": "Tropic Thunder",
  "streaming_date": "Streaming Dec 29, 2012",
  "audience_score": "71%",
  "critic_score": "82%"
},
{
  "title": "Tár",
  "streaming_date": "Streaming Dec 20, 2022",
  "audience_score": "74%",
  "critic_score": "91%"
},
{
  "title": "Superbad",
  "streaming_date": "Streaming Apr 16, 2012",
  "audience_score": "87%",
  "critic_score": "88%"
},
{
  "title": "Harry Potter and the Sorcerer's Stone",
  "streaming_date": "Streaming Dec 16, 2012",
  "audience_score": "82%",
  "critic_score": "81%"
},
{
  "title": "Melancholia",
  "streaming_date": "Streaming Jan 13, 2015",
  "audience_score": "67%",
  "critic_score": "80%"
},
{
  "title": "Casino",
  "streaming_date": "Streaming May 18, 2016",
  "audience_score": "93%",
  "critic_score": "78%"
},
{
  "title": "Scarface",
  "streaming_date": "Streaming Jun 15, 2012",
  "audience_score": "93%",
  "critic_score": "79%"
}
```

```

    },
    {
      "title": "Spider-Man 2",
      "streaming_date": "Streaming May 8, 2015",
      "audience_score": "82%",
      "critic_score": "93%"
    },
    {
      "title": "The Endless",
      "streaming_date": "Streaming Jun 26, 2018",
      "audience_score": "68%",
      "critic_score": "92%"
    },
    {
      "title": "Coherence",
      "streaming_date": "Streaming Nov 16, 2015",
      "audience_score": "81%",
      "critic_score": "88%"
    },
    {
      "title": "Sick",
      "streaming_date": "Streaming Jan 13, 2023",
      "audience_score": "73%",
      "critic_score": "85%"
    },
    {
      "title": "John Wick: Chapter 3 -- Parabellum",
      "streaming_date": "Streaming Aug 23, 2019",
      "audience_score": "86%",
      "critic_score": "89%"
    }
  ]
}
Title2: Certified Fresh Movies on Prime Video (2024) | Rotten Tomatoes
[
  {
    "title": "Hundreds of Beavers",
    "streaming_date": "Streaming Apr 15, 2024",
    "audience_score": "94%",
    "critic_score": "95%"
  },
  {
    "title": "Dune: Part Two",
    "streaming_date": "Streaming Apr 16, 2024",
    "audience_score": "95%",
    "critic_score": "93%"
  },
  {
    "title": "One Life",
    "streaming_date": "Streaming Apr 9, 2024",
    "audience_score": "95%",

```

```
"critic_score": "90%"
},
{
  "title": "The Zone of Interest",
  "streaming_date": "Streaming Feb 20, 2024",
  "audience_score": "78%",
  "critic_score": "93%"
},
{
  "title": "Poor Things",
  "streaming_date": "Streaming Feb 27, 2024",
  "audience_score": "79%",
  "critic_score": "92%"
},
{
  "title": "All of Us Strangers",
  "streaming_date": "Streaming Feb 22, 2024",
  "audience_score": "91%",
  "critic_score": "96%"
},
{
  "title": "The Gentlemen",
  "streaming_date": "Streaming Mar 24, 2020",
  "audience_score": "84%",
  "critic_score": "75%"
},
{
  "title": "Coup de Chance",
  "streaming_date": "Streaming Apr 12, 2024",
  "audience_score": "81%",
  "critic_score": "83%"
},
{
  "title": "Dream Scenario",
  "streaming_date": "Streaming Dec 22, 2023",
  "audience_score": "68%",
  "critic_score": "91%"
},
{
  "title": "Wonka",
  "streaming_date": "Streaming Jan 30, 2024",
  "audience_score": "91%",
  "critic_score": "82%"
},
{
  "title": "Megan Leavey",
  "streaming_date": "Streaming Jul 10, 2017",
  "audience_score": "82%",
  "critic_score": "86%"
},
},
```



```
{
  "title": "Oppenheimer",
  "streaming_date": "Streaming Nov 21, 2023",
  "audience_score": "91%",
  "critic_score": "93%"
},
{
  "title": "Dune",
  "streaming_date": "Streaming Oct 22, 2021",
  "audience_score": "90%",
  "critic_score": "83%"
},
{
  "title": "Talk to Me",
  "streaming_date": "Streaming Sep 12, 2023",
  "audience_score": "82%",
  "critic_score": "95%"
},
{
  "title": "Godzilla vs. Kong",
  "streaming_date": "Streaming Mar 31, 2021",
  "audience_score": "91%",
  "critic_score": "76%"
},
{
  "title": "The Fallout",
  "streaming_date": "Streaming Jan 27, 2022",
  "audience_score": "84%",
  "critic_score": "93%"
},
{
  "title": "Perfect Days",
  "streaming_date": "Streaming Mar 5, 2024",
  "audience_score": "89%",
  "critic_score": "96%"
},
{
  "title": "The Iron Claw",
  "streaming_date": "Streaming Feb 13, 2024",
  "audience_score": "94%",
  "critic_score": "89%"
},
{
  "title": "Stopmotion",
  "streaming_date": "Streaming Mar 15, 2024",
  "audience_score": "50%",
  "critic_score": "90%"
},
{
  "title": "American Fiction",
```

```
"streaming_date": "Streaming Feb 6, 2024",
"audience_score": "96%",
"critic_score": "93%"
},
{
  "title": "Split",
  "streaming_date": "Streaming Apr 5, 2017",
  "audience_score": "79%",
  "critic_score": "78%"
},
{
  "title": "Ordinary Angels",
  "streaming_date": "Streaming Mar 26, 2024",
  "audience_score": "99%",
  "critic_score": "84%"
},
{
  "title": "Killers of the Flower Moon",
  "streaming_date": "Streaming Dec 5, 2023",
  "audience_score": "84%",
  "critic_score": "93%"
},
{
  "title": "X",
  "streaming_date": "Streaming Apr 14, 2022",
  "audience_score": "75%",
  "critic_score": "94%"
},
{
  "title": "Baby Driver",
  "streaming_date": "Streaming Jul 10, 2017",
  "audience_score": "86%",
  "critic_score": "92%"
},
{
  "title": "The Holdovers",
  "streaming_date": "Streaming Nov 28, 2023",
  "audience_score": "92%",
  "critic_score": "97%"
},
{
  "title": "The Invisible Man",
  "streaming_date": "Streaming Mar 20, 2020",
  "audience_score": "88%",
  "critic_score": "92%"
},
{
  "title": "War for the Planet of the Apes",
  "streaming_date": "Streaming Apr 1, 2019",
  "audience_score": "84%",
```

```

    "critic_score": "94%"
  }
]

```

6.2 XPATH

6.2.1 RTV

```

{
  "TITLE": "Audi A6 50 TDI quattro: nemir v premijskem razredu",
  "AUTHOR": "Miha Merljak",
  "PUBLISHED_TIME": "28. december 2018 ob 08:51",
  "SUBTITLE": "Test nove generacije",
  "LEAD": "To je novi audi A6. V razred najdražjih in najbolj premijskih žrebcev je vnesel nemir, še preden je sploh zapeljal na parkirni prostor, rezerviran za izvršnega direktorja. ",
  "CONTENT": [
    "Samo pogledjte njegovo masko – to ogromno satovje z radarji na takem položaju, da se ti na avtocesti tudi pri 120 km/h vsi spoštljivo umikajo, saj so prepričani, da gre za Pahorjev ali Šarčev avto. Seveda, novi A6 lahko cesto in promet skenira s kar petimi radarji, petimi kamerami, infrardečo kamero za nočni vid, dvanajstimi ultrazvočnimi senzorji in laserskim čitalnikom – lidarjem. V glavnem vojaška tehnologija v službi varnosti za fante, ki smo radi gledali Top Gun, Bonda in druge možakarja s finimi igračami.",
    "Vozniški delovni prostor je novo poglavje digitalne dobe, z dvema ogromnima zaslonoma, ki tako kot naprednejši telefoni dregnejo blazinice vaših prstov, kot se sprehajate po steklu. A še bolj se nam zdi pomembno, da so osnovna stikala tam, kjer jih pričakujete. Najprej so torej zagotovili enostavno osnovo, tisti bolj \"advanced\" vozniki pa si lahko nato vse skupaj še veliko bolj prilagodijo. Velik korak naprej pri kabinskem udobju zaznavajo tudi na zadnji klopi, tam je prostora v vseh smereh precej več.",
    "Če vam pogled na Audijev spisek dodatne opreme ne odvzame volje do življenja, potem vsekakor toplo priporočamo nakup zračnega vzmetenja, saj dobi z njim A6 več različnih in vozniško zelo uporabnih karakterjev.",
    "Enako velja za seksi luči z inteligentno matrično osvetlitvijo, pa za športno podvozje in vsekakor za štirikolesno krmiljenje. S tem postane A6 med ovinki v občutku na volanu še veliko krajši in bolj agilen. Vse naštetu smo preskušali v družbi agregata 50 TDI, ki je v resnici klasični trilitrski dizel, podkrepjen z elektromotorjem. Ja, ta audi je mehki hibrid z izjemnim navorom in dovolj moči kadar koli in kjer koli. Si pa mislimo, da bo največji del trga zadovoljil že učinkovit dvolitrski mehki hibrid z močjo 150 kilovatov.",
    "- na testu Audi A6 50 TDI quattro tiptronic",
    "- dolžina: 4,9 m",
    "- medosna razdalja: 2,9 m",
    "- obračalni krog: 12,1 m",
    "- prtljažnik: 530 l",
    "- masa: 1.900 kg",
    "- trilitrski šestvaljni dizelski motor",
    "- moč: 210 kW",
    "- navor: 620 Nm",
    "- 8-stopenjski samodejni menjalnik",
    "- pogon na vsa štiri kolesa",
    "- pnevmatike: 225/60 R17",
  ]
}

```

"- poraba: 6,6 l/100 km = 8,9 EUR/100 km",
"- posoda za gorivo: 73 l",
"- doseg: 1.106 km",
"- izpusti CO",
": 147 g/km",
"- nakupna cena: 69.080 EUR",
"- stroški finančnega lizinga: 4.463 EUR/5 let",
"- stroški registracije: 10.829 EUR/5 let",
"- stroški vzdrževanja: 1.926 EUR/5 let",
"- stroški goriva: 6.702 EUR/75.000 km",
"- strošek 1 kompleta pnevmatik: 716 EUR",
"- vrednosti po 5 letih po Eurotaxu: 33.964 EUR",
"- stroški skupaj: 1.001 EUR/mesec"
]

}
{

"TITLE": "Volvo XC 40 D4 AWD momentum: suvereno med najboljše v razredu",
"AUTHOR": "Miha Merljak",
"PUBLISHED_TIME": "25. januar 2019 ob 15:23",
"SUBTITLE": "Test novega modela",

"LEAD": "XC 40 je najmanjši Volvov SUV, ki se oblikovno skoraj v celotni naslanja na oba večja predhodnika. Že samo s tem so mu vrata do denarnic tistih kupcev, ki iščejo izstopajočo, a hkrati visoko kultivirano in prečiščeno dizajnersko govorico, na pol odprta.",

"CONTENT": [

"Volvo se je nižjih srednjih razredov v preteklosti izogibal ali pa je vanje vstopal z zelo nišnimi produkti, ki niso pustili večjega tržnega pečata. V primeru XC 40 ni težko napovedati, da bo ta tradicija prekinjena. Ponuja namreč visoko kakovost končne izdelave in v kabini odlično premišljeno funkcionalnost ter na dotik prijetne materiale. ",

"Še posebej hvalimo število, iznajdljivost in velikost različnih odlagalnih prostorov ter široke, čvrste in zelo udobne sedeže. Intuitivno in enostavno logično je upravljanje z velikim vmesnikom, ki z večfunkcijskim zaslonom na dotik kraljuje na z roko lahko dostopnem mestu na sredinski armaturi. Razočaranj ne bo niti v velikosti in uporabnosti prtljažnega prostora, ki s 460 litri prostornine sicer ni med večjimi v razredu, a se v uporabniškem smislu odkupi z dobro urejenostjo ter domiselnimi rešitvami pregrajevanja.",

"XC 40 je od tal odmaknjen konkretnih 21 cm, a sta vzmetenje in krmilni mehanizem tako nastavljena, da ponuja tudi v hitro odpeljanih ovinkih zelo dolgo nevtrarno in predvidljivo lego. V premeru preskušane modela, ki je imel v paketu R design vzmetenje še nekoliko bolj trdo, se je to samo še bolj potrdilo, a je v tem primeru treba računati na manj udobno vožnjo čez različne asfaltne grbine. Podoben razmislek velja opraviti tudi pri izbiri motorja.",

"Preskušani 2-litrski dizel s 190 KM predstavlja vrh ponudbe, ki z močjo, udobjem in tudi povprečno porabo navduši predvsem pri avtocestnih dolgoprogaških izzivih, v počasni mestni vožnji ter pri pogostih postankih in speljevanjih pa deluje preveč robusten.",

"XC 40 je s čvrsto gradnjo, funkcionalno in udobno kabino ter številnimi asistenčnimi sistemi in izstopajočim skandinavskim dizajnom v premišljenem trenutku vstopil na trg modnih mestnih terencev, v katerem se brez ene same sence dvoma suvereno postavi med najdražje in najbolj premijske v mestu.",

"- na testu Volvo XC40 2.0 TD avt awd momentum",

"Mere:",

"- dolžina: 4,4 m",

```

"- medosna razdalja: 2,7 m",
"- obračalni krog: 11,4 m",
"- oddaljenost od tal: 21 cm",
"- prtljažnik: 432 l",
"- masa: 2.250 kg",
"Pogon:",
"- 2-litrski 4-valjni bencinski motor",
"- moč: 140 kW",
"- navor: 400 Nm",
"- 8-stopenjski samodejni menjalnik",
"- pogon na vsa štiri kolesa",
"- pnevmatike: 235/50 R19 ",
"- poraba: 6,3 l/100 km = 8,2 EUR/100km",
"- posoda za gorivo: 54 l",
"- doseg: 857 km",
"- izpusti CO2: 133 g/km",
"Stroški pri 15.000 km in 5-letni uporabi:",
"- nakupna cena: 43.619 EUR",
"- stroški finančnega leasinga: 3.268 EUR/5 let",
"- stroški registracije: 8.701 EUR/5 let",
"- stroški vzdrževanja: 2.320 EUR/5 let",
"- stroški goriva: 6.190 EUR/75.000 km",
"- strošek 1 kompleta pnevmatik: 923 EUR",
"- vrednosti po 5 letih po Eurotaxu: 18.886 EUR",
"- stroški skupaj: 774 EUR/mesec"
]
}

```

6.2.2 OVERSTOCK

```

{
  "TITLE": "10-kt. Seven Diamond Ladies Heart Ring (0.08 TW)",
  "LIST_PRICE": "$149.00",
  "PRICE": "$69.99",
  "SAVED_AMOUNT": "$79.01",
  "SAVED_PERCENTAGE": "(53%)",
  "CONTENT": "This ladies fashion ring dazzles\nwith hearts and diamonds. The gold
band is crafted into delicate, open hearts.\nSeven brilliant-cut diamonds add a bit of sparkle. "
}
{
  "TITLE": "10-Kt. Diamond Ring (.25 TW)",
  "LIST_PRICE": "$250.00",
  "PRICE": "$74.90",
  "SAVED_AMOUNT": "$175.10",
  "SAVED_PERCENTAGE": "(70%)",
  "CONTENT": "Nineteen round diamonds accent this 10-karat yellow gold ring with
filigree accents."
}
{
  "TITLE": "10-kt. Pearl and Diamond Butterfly Earrings",

```

```

    "LIST_PRICE": "$149.00",
    "PRICE": "$42.99",
    "SAVED_AMOUNT": "$106.01",
    "SAVED_PERCENTAGE": "(71%)",
    "CONTENT": "Perfectly proportioned 5.5- to 6-mm cultured pearls on 10-karat
yellow gold settings highlight these petite earrings. A dainty rhodium-plated gold butterfly
studded with a diamond (0.02 total carat weight, J-K color, I-2 clarity) rests atop each
pearl."
  }
  {
    "TITLE": "14-kt. Diamond 'S' Tennis Bracelet (2.00 TW)",
    "LIST_PRICE": "$1,539.99",
    "PRICE": "$499.99",
    "SAVED_AMOUNT": "$1,040.00",
    "SAVED_PERCENTAGE": "(67%)",
    "CONTENT": "Invest in a swirl of light with this diamond 'S' tennis bracelet.
Crafted in 14-karat gold, the piece features 49 diamonds for two full carats. The 7.25-inch
bracelet closes with a pressure clasp."
  }
  {
    "TITLE": "10-kt. Diamond Band Fashion Ring (.11 TW)",
    "LIST_PRICE": "$179.99",
    "PRICE": "$79.99",
    "SAVED_AMOUNT": "$100.00",
    "SAVED_PERCENTAGE": "(55%)",
    "CONTENT": "Crafted in white and yellow gold, this ring displays a band of seven
round diamonds. Order your new gold and diamond fashion ring today at our low, online
price."
  }
  {
    "TITLE": "14-kt. White Gold, Pearl and Diamond Ring",
    "LIST_PRICE": "$419.99",
    "PRICE": "$149.99",
    "SAVED_AMOUNT": "$270.00",
    "SAVED_PERCENTAGE": "(64%)",
    "CONTENT": "Show your romantic side with this 14-karat diamond and pearl ring.
Set in a domed band of 14-karat white gold, the ring features a 7-mm cultured pearl. Curved
rows of diamonds flank the pearl."
  }
  {
    "TITLE": "14-kt. Gold Diamond Present Future Pendant (.25TW)",
    "LIST_PRICE": "$299.00",
    "PRICE": "$149.99",
    "SAVED_AMOUNT": "$149.01",
    "SAVED_PERCENTAGE": "(49%)",
    "CONTENT": "Designed with three large, sparkling diamonds to represent past,
present, and future, this stunning pendant is set in gleaming 14-karat gold. It incorporates a
total of nine diamonds (0.25 total carat weight, K color, I-2 to I-3 clarity). "
  }
  {

```

```

    "TITLE": "14-kt. Diamond Solitaire Pendant (.33 TW)",
    "LIST_PRICE": "$1,019.99",
    "PRICE": "$319.99",
    "SAVED_AMOUNT": "$700.00",
    "SAVED_PERCENTAGE": "(68%)",
    "CONTENT": "In this simple, yet elegant pendant,\na round brilliant diamond (0.33
total carat weight, H-J color, I-1 to I-2\nclearity) is prong-set in 14-karat white gold."
}
{
    "TITLE": "14-kt. Diamond Solitaire Earrings (0.33 TW)",
    "LIST_PRICE": "$639.99",
    "PRICE": "$199.99",
    "SAVED_AMOUNT": "$440.00",
    "SAVED_PERCENTAGE": "(68%)",
    "CONTENT": "Dazzle your way into her heart,\nwith these classic diamond solitaire
earrings. Two brilliant-cut diamonds\n(0.33 total carat weight, G-H color, I-1 to I-2 clarity)
are set in four\nprongs of 14-karat white gold."
}
{
    "TITLE": "14-kt. Diamond Cross Pendant (.06 TW)",
    "LIST_PRICE": "$305.00",
    "PRICE": "$119.99",
    "SAVED_AMOUNT": "$185.01",
    "SAVED_PERCENTAGE": "(60%)",
    "CONTENT": "Over a cleanly sculpted Roman\ncross of 14-karat white gold drapes a
slender banner containing three bright\nprong-set round diamonds (0.06 total carat weight,
H-I color, I clarity)."
}
{
    "TITLE": "14-kt. Diamond Solitaire Stud Earrings (.50 TW)",
    "LIST_PRICE": "$999.99",
    "PRICE": "$359.99",
    "SAVED_AMOUNT": "$640.00",
    "SAVED_PERCENTAGE": "(64%)",
    "CONTENT": "Every jewelry collection needs\na classic pair of diamond solitaire
earrings. Set in 14-karat gold, these\ndiamond stud earrings (0.50 total carat weight) have
post backs with butterfly\nclasps."
}
{
    "TITLE": "14-kt. Cultured Pearl Diamond Earrings",
    "LIST_PRICE": "$508.99",
    "PRICE": "$179.99",
    "SAVED_AMOUNT": "$329.00",
    "SAVED_PERCENTAGE": "(64%)",
    "CONTENT": "Create an elegant appearance with\nthese pearl and diamond stud
earrings. Set in 14-karat yellow gold, each\nearring features an 8 to 8.5-mm cultured white
pearl. Prong-set round diamonds\naccent the pearls. Posts with butterfly clasps secure the
earrings."
}
{

```

```

    "TITLE": "14-kt. Diamond 7.5-8 mm Pearl Pendant",
    "LIST_PRICE": "$196.99",
    "PRICE": "$69.99",
    "SAVED_AMOUNT": "$127.00",
    "SAVED_PERCENTAGE": "(64%)",
    "CONTENT": "Add a classic to your jewelry\ncollection with this 14-karat gold,
diamond, and pearl necklace. The 7.5-8\mmm cultured white pearl creates the focal point of
the pendant, while a diamond\n(0.10 TW) adds sparkle."
}
{
    "TITLE": "14-kt. Diamond Solitaire Earrings (.50 TW)",
    "LIST_PRICE": "$1,369.99",
    "PRICE": "$409.99",
    "SAVED_AMOUNT": "$960.00",
    "SAVED_PERCENTAGE": "(70%)",
    "CONTENT": "This earring set has two brilliant-cut\ndiamonds (0.50 total carat
weight, G-H color, I-1 to I-2 clarity) set in\nfour prongs of 14-karat white gold."
}
{
    "TITLE": "14-kt White Gold Diamond Band (0.50 TW)",
    "LIST_PRICE": "$1,635.00",
    "PRICE": "$609.99",
    "SAVED_AMOUNT": "$1,025.01",
    "SAVED_PERCENTAGE": "(62%)",
    "CONTENT": "Crafted of 14-karat white gold,\nthis stylish ring features a bright row
of 20 channel-set, princess-cut baguette\ndiamonds. Treat her like royalty and save when you
buy jewelry treasures\nat Overstock.com."
}
{
    "TITLE": "14-kt. Green Jade Hoops",
    "LIST_PRICE": "$90.00",
    "PRICE": "$46.99",
    "SAVED_AMOUNT": "$43.01",
    "SAVED_PERCENTAGE": "(47%)",
    "CONTENT": "Hoops of cool green jade rest\nbetween 14-karat yellow gold
endpieces. The hoops graduate in thickness from\n3 mm at the ends to 6 mm in the center,
with approximately 29 mm overall\ndiameter."
}
{
    "TITLE": "14-kt. Jade Doughnut Pendant",
    "LIST_PRICE": "$150.00",
    "PRICE": "$48.99",
    "SAVED_AMOUNT": "$101.01",
    "SAVED_PERCENTAGE": "(67%)",
    "CONTENT": "The 25-mm disk hangs delicately\nfrom a 14-karat gold chain. The
disk features a dramatic gold Chinese character\nin the center, accompanied by four stylized
gold bees."
}
{
    "TITLE": "14-kt. Charcoal Jade and Ruby Elephant Pendant",

```



```

    "LIST_PRICE": "$100.00",
    "PRICE": "$28.99",
    "SAVED_AMOUNT": "$71.01",
    "SAVED_PERCENTAGE": "(71%)",
    "CONTENT": "Carved of rich dark grey jade,\nthis elephant pendant has 14-karat
yellow gold applied to mark the feet,\ntusk, tail, and blanket. A 2-mm round faceted ruby in a
gold bezel setting\nforms the eye. The pendant hangs from an 18-inch chain."
}
{
    "TITLE": "14-kt. Carved Lavender Jade Earrings",
    "LIST_PRICE": "$80.00",
    "PRICE": "$39.99",
    "SAVED_AMOUNT": "$40.01",
    "SAVED_PERCENTAGE": "(50%)",
    "CONTENT": "Luscious 8-mm lavender jade balls, carved with intricate Asian style,
dangle from a 14-karat yellow gold French hook."
}
{
    "TITLE": "14-kt. Jade Cross Pendant",
    "LIST_PRICE": "$150.00",
    "PRICE": "$49.99",
    "SAVED_AMOUNT": "$100.01",
    "SAVED_PERCENTAGE": "(66%)",
    "CONTENT": "Green jade and gold create this\nbeautiful cross pendant. Cylindrical
bars of green jade feature caps and\ncenter of 14-karat yellow gold."
}
{
    "TITLE": "14-kt. Multicolored Jade Earrings",
    "LIST_PRICE": "$375.00",
    "PRICE": "$99.99",
    "SAVED_AMOUNT": "$275.01",
    "SAVED_PERCENTAGE": "(73%)",
    "CONTENT": "A delicate wrapping of 14-karat\nyellow gold wire holds six 6 x 4
pear shapes of jade in various shades: brilliant\ngreen, orange, lavender, black, pale yellow,
and white. The post earrings\nhave butterfly backs."
}
{
    "TITLE": "14-kt. Multicolored Jade Ring",
    "LIST_PRICE": "$250.00",
    "PRICE": "$56.99",
    "SAVED_AMOUNT": "$193.01",
    "SAVED_PERCENTAGE": "(77%)",
    "CONTENT": "A delicate wrapping of 14-karat\nyellow gold wire holds six 6 x 4
ovals of jade in various shades: brilliant\ngreen, orange, lavender, black, pale yellow, and
white. A narrow gold band\ndivides to support the setting."
}
{
    "TITLE": "14-kt. Onyx and Ruby Elephant Pendant",
    "LIST_PRICE": "$100.00",
    "PRICE": "$35.99",

```

```

    "SAVED_AMOUNT": "$64.01",
    "SAVED_PERCENTAGE": "(64%)",
    "CONTENT": "Carved of rich black onyx, this\nelephant pendant has 14-karat yellow
gold applied to mark the feet, tusk,\ntail, and blanket. A 2-mm round faceted ruby in a gold
bezel setting forms\nthe eye. The pendant hangs from an 18-inch chain."
}

```

6.2.3 ROTTEN TOMATOES

Page title: Certified Fresh Movies on Peacock

```

{
  "title": "Oppenheimer",
  "streaming_date": "Streaming Nov 21, 2023",
  "audience_score": "91%",
  "critic_score": "93%"
}
{
  "title": "The Holdovers",
  "streaming_date": "Streaming Nov 28, 2023",
  "audience_score": "92%",
  "critic_score": "97%"
}
{
  "title": "Nope",
  "streaming_date": "Streaming Nov 18, 2022",
  "audience_score": "69%",
  "critic_score": "83%"
}
{
  "title": "Let the Right One In",
  "streaming_date": "Streaming Jan 14, 2017",
  "audience_score": "90%",
  "critic_score": "98%"
}
{
  "title": "The Godfather",
  "streaming_date": "Streaming Aug 1, 2013",
  "audience_score": "98%",
  "critic_score": "97%"
}
{
  "title": "Dredd",
  "streaming_date": "Streaming Sep 2, 2014",
  "audience_score": "72%",
  "critic_score": "80%"
}
{
  "title": "Memento",
  "streaming_date": "Streaming Mar 1, 2017",

```

```

    "audience_score": "94%",
    "critic_score": "94%"
  }
  {
    "title": "Shrek 2",
    "streaming_date": "Streaming Nov 25, 2015",
    "audience_score": "69%",
    "critic_score": "89%"
  }
  {
    "title": "Shrek",
    "streaming_date": "Streaming Nov 25, 2015",
    "audience_score": "90%",
    "critic_score": "88%"
  }
  {
    "title": "John Wick",
    "streaming_date": "Streaming Jun 7, 2016",
    "audience_score": "81%",
    "critic_score": "86%"
  }
  {
    "title": "Jackie Brown",
    "streaming_date": "Streaming Dec 17, 2015",
    "audience_score": "85%",
    "critic_score": "88%"
  }
  {
    "title": "Spider-Man",
    "streaming_date": "Streaming Apr 25, 2013",
    "audience_score": "67%",
    "critic_score": "90%"
  }
  {
    "title": "Noah",
    "streaming_date": "Streaming Apr 9, 2016",
    "audience_score": "41%",
    "critic_score": "75%"
  }
  {
    "title": "Promising Young Woman",
    "streaming_date": "Streaming Jan 15, 2021",
    "audience_score": "87%",
    "critic_score": "90%"
  }
  {
    "title": "The Godfather, Part II",
    "streaming_date": "Streaming Sep 15, 2012",
    "audience_score": "97%",
    "critic_score": "96%"
  }

```

```

}
{
  "title": "The Big Lebowski",
  "streaming_date": "Streaming Dec 1, 2015",
  "audience_score": "93%",
  "critic_score": "80%"
}
{
  "title": "Tropic Thunder",
  "streaming_date": "Streaming Dec 29, 2012",
  "audience_score": "71%",
  "critic_score": "82%"
}
{
  "title": "Tár",
  "streaming_date": "Streaming Dec 20, 2022",
  "audience_score": "74%",
  "critic_score": "91%"
}
{
  "title": "Superbad",
  "streaming_date": "Streaming Apr 16, 2012",
  "audience_score": "87%",
  "critic_score": "88%"
}
{
  "title": "Harry Potter and the Sorcerer's Stone",
  "streaming_date": "Streaming Dec 16, 2012",
  "audience_score": "82%",
  "critic_score": "81%"
}
{
  "title": "Melancholia",
  "streaming_date": "Streaming Jan 13, 2015",
  "audience_score": "67%",
  "critic_score": "80%"
}
{
  "title": "Casino",
  "streaming_date": "Streaming May 18, 2016",
  "audience_score": "93%",
  "critic_score": "78%"
}
{
  "title": "Scarface",
  "streaming_date": "Streaming Jun 15, 2012",
  "audience_score": "93%",
  "critic_score": "79%"
}
{

```

```

        "title": "Spider-Man 2",
        "streaming_date": "Streaming May 8, 2015",
        "audience_score": "82%",
        "critic_score": "93%"
    }
    {
        "title": "The Endless",
        "streaming_date": "Streaming Jun 26, 2018",
        "audience_score": "68%",
        "critic_score": "92%"
    }
    {
        "title": "Coherence",
        "streaming_date": "Streaming Nov 16, 2015",
        "audience_score": "81%",
        "critic_score": "88%"
    }
    {
        "title": "Sick",
        "streaming_date": "Streaming Jan 13, 2023",
        "audience_score": "73%",
        "critic_score": "85%"
    }
    {
        "title": "John Wick: Chapter 3 -- Parabellum",
        "streaming_date": "Streaming Aug 23, 2019",
        "audience_score": "86%",
        "critic_score": "89%"
    }

```

Page title: Certified Fresh Movies on Prime Video

```

    {
        "title": "Hundreds of Beavers",
        "streaming_date": "Streaming Apr 15, 2024",
        "audience_score": "94%",
        "critic_score": "95%"
    }
    {
        "title": "Dune: Part Two",
        "streaming_date": "Streaming Apr 16, 2024",
        "audience_score": "95%",
        "critic_score": "93%"
    }
    {
        "title": "One Life",
        "streaming_date": "Streaming Apr 9, 2024",
        "audience_score": "95%",
        "critic_score": "90%"
    }
    {
        "title": "The Zone of Interest",

```

```

    "streaming_date": "Streaming Feb 20, 2024",
    "audience_score": "78%",
    "critic_score": "93%"
  }
  {
    "title": "Poor Things",
    "streaming_date": "Streaming Feb 27, 2024",
    "audience_score": "79%",
    "critic_score": "92%"
  }
  {
    "title": "All of Us Strangers",
    "streaming_date": "Streaming Feb 22, 2024",
    "audience_score": "91%",
    "critic_score": "96%"
  }
  {
    "title": "The Gentlemen",
    "streaming_date": "Streaming Mar 24, 2020",
    "audience_score": "84%",
    "critic_score": "75%"
  }
  {
    "title": "Coup de Chance",
    "streaming_date": "Streaming Apr 12, 2024",
    "audience_score": "81%",
    "critic_score": "83%"
  }
  {
    "title": "Dream Scenario",
    "streaming_date": "Streaming Dec 22, 2023",
    "audience_score": "68%",
    "critic_score": "91%"
  }
  {
    "title": "Wonka",
    "streaming_date": "Streaming Jan 30, 2024",
    "audience_score": "91%",
    "critic_score": "82%"
  }
  {
    "title": "Megan Leavey",
    "streaming_date": "Streaming Jul 10, 2017",
    "audience_score": "82%",
    "critic_score": "86%"
  }
  {
    "title": "Oppenheimer",
    "streaming_date": "Streaming Nov 21, 2023",
    "audience_score": "91%",

```

```

    "critic_score": "93%"
  }
  {
    "title": "Dune",
    "streaming_date": "Streaming Oct 22, 2021",
    "audience_score": "90%",
    "critic_score": "83%"
  }
  {
    "title": "Talk to Me",
    "streaming_date": "Streaming Sep 12, 2023",
    "audience_score": "82%",
    "critic_score": "95%"
  }
  {
    "title": "Godzilla vs. Kong",
    "streaming_date": "Streaming Mar 31, 2021",
    "audience_score": "91%",
    "critic_score": "76%"
  }
  {
    "title": "The Fallout",
    "streaming_date": "Streaming Jan 27, 2022",
    "audience_score": "84%",
    "critic_score": "93%"
  }
  {
    "title": "Perfect Days",
    "streaming_date": "Streaming Mar 5, 2024",
    "audience_score": "89%",
    "critic_score": "96%"
  }
  {
    "title": "The Iron Claw",
    "streaming_date": "Streaming Feb 13, 2024",
    "audience_score": "94%",
    "critic_score": "89%"
  }
  {
    "title": "Stopmotion",
    "streaming_date": "Streaming Mar 15, 2024",
    "audience_score": "50%",
    "critic_score": "90%"
  }
  {
    "title": "American Fiction",
    "streaming_date": "Streaming Feb 6, 2024",
    "audience_score": "96%",
    "critic_score": "93%"
  }
}

```

```
{
  "title": "Split",
  "streaming_date": "Streaming Apr 5, 2017",
  "audience_score": "79%",
  "critic_score": "78%"
}
{
  "title": "Ordinary Angels",
  "streaming_date": "Streaming Mar 26, 2024",
  "audience_score": "99%",
  "critic_score": "84%"
}
{
  "title": "Killers of the Flower Moon",
  "streaming_date": "Streaming Dec 5, 2023",
  "audience_score": "84%",
  "critic_score": "93%"
}
{
  "title": "X",
  "streaming_date": "Streaming Apr 14, 2022",
  "audience_score": "75%",
  "critic_score": "94%"
}
{
  "title": "Baby Driver",
  "streaming_date": "Streaming Jul 10, 2017",
  "audience_score": "86%",
  "critic_score": "92%"
}
{
  "title": "The Holdovers",
  "streaming_date": "Streaming Nov 28, 2023",
  "audience_score": "92%",
  "critic_score": "97%"
}
{
  "title": "The Invisible Man",
  "streaming_date": "Streaming Mar 20, 2020",
  "audience_score": "88%",
  "critic_score": "92%"
}
{
  "title": "War for the Planet of the Apes",
  "streaming_date": "Streaming Apr 1, 2019",
  "audience_score": "84%",
  "critic_score": "94%"
}
```


6.3 ROAD RUNNER

Izpis metode Road Runner smo priložili v obliki tekstovnih datotek.