

Verificação de Tendências em Popularidade Regional de Jogos Digitais



Julien Hervot de Mattos Vaz (19202276)
Bruno Pamplona Huebes (19102917)

Introdução

- Objetivos: Utilizar métricas de popularidade de jogos digitais populares em diferentes países do planeta para verificar o efeito de características e funções disponíveis nos hábitos de consumo regionais.
- Coleta dos seguintes dados: lista ordenada de jogos mais populares na plataforma escolhida, para cada um dos países do escopo. Para cada um dos jogos, encontrar seus gêneros, idiomas disponíveis, e preços em cada região.
- Objetiva-se realizar uma análise de dados para encontrar fatores de grande impacto para o sucesso dos produtos em regiões específicas.
- Uma pesquisa desta forma poderia ser interessante para distribuidoras interessadas em atender mercados regionais menos representados, e em encontrar potenciais nichos não atendidos.

Metodologia

- Utilização da Steam, como plataforma mais popular para distribuição de jogos digitais, para encontrar as classificações de popularidade dos produtos, com seus idiomas disponíveis listados no próprio site.
- Lista de popularidade é específica para cada país, requerendo uso de proxies para sua obtenção.
- Para obtenção dos marcadores de gêneros (ex: ação, suspense, casual), utilização de uma Spider sobre o site RAWG.io, que lista dados sobre diversos jogos.
- Uso da API IsThereAnyDeal para obter informações quanto aos preços regionais.

Objetivo

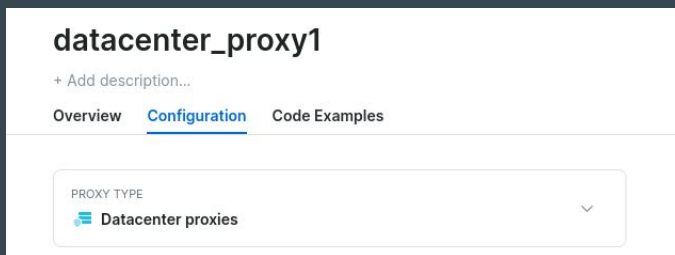
- Arquivos iniciais: Listas de Países (com códigos/siglas e moedas) a serem trabalhados.
- Páginas de entrada: Páginas Web com conteúdo principal em HTML (em alguns casos contido no conteúdo JavaScript das páginas):
 - Página de Melhores Vendidos por Região, da Steam.
 - Páginas dos produtos, com tabela de idiomas no conteúdo HTML.
 - Páginas de jogos no RAWG.io, com referência ao ID na Steam dentro de conteúdo JavaScript, e título do jogo em Header HTML.
- Conteúdo adicional de entrada: Conteúdo Retornado via API do IsThereAnyDeal.

Objetivo

- Conteúdo de Suporte: Lista de Proxies disponíveis no serviço utilizado (Bright Data) para decidir os países a serem tratados.
- Conteúdo de Saída: Para cada script, gerar um JSON relacionando o jogo tratado (marcado com seu ID de aplicativo) com os dados obtidos nesse script (lista de idiomas, lista de preços, lista de gêneros,etc.) e a URL fonte.
- Ao final, gerar um JSON que combina os dados de todos os outros arquivos JSON gerados.

Uso de Proxies para obtenção de conteúdo regional

Lista de jogos mais populares depende da região de acesso, portanto, proxies providenciadas pelo Bright Data, foram utilizadas para obter as listas.

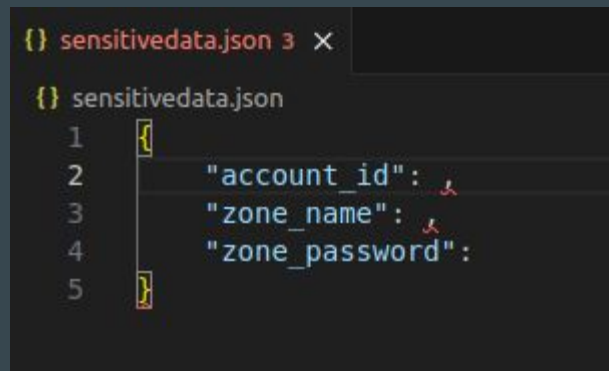


| | | |
|------------------|--------------|--------------------------------|
| Dota 2 | 9/jul./2013 | Gratuito |
| Counter-Strike 2 | 21/ago./2012 | Gratuito |
| Path of Exile 2 | 6/dez./2024 | R\$79,80 |
| ELDEN RING | 24/fev./2022 | -40% R\$229,90 R\$137,94 |

Página de jogos mais populares (19 de Dez. de 2024), acessado do Brasil

| | | |
|---------------------|--------------|------|
| NARAKA: BLADEPOINT | 11 Aug. 2021 | Free |
| Dota 2 | 9 Jul. 2013 | Free |
| Counter-Strike 2 | 21 Aug. 2012 | Free |
| PUBG: BATTLEGROUNDS | 21 Dec. 2017 | Free |

Página de jogos mais populares (19 de Dez. de 2024), acessado com VPN para Hong Kong




























Arquivo sensitivedata.json usado para inserir dados de login e de acesso a proxies.

Dados de Entrada: Lista de Popularidade

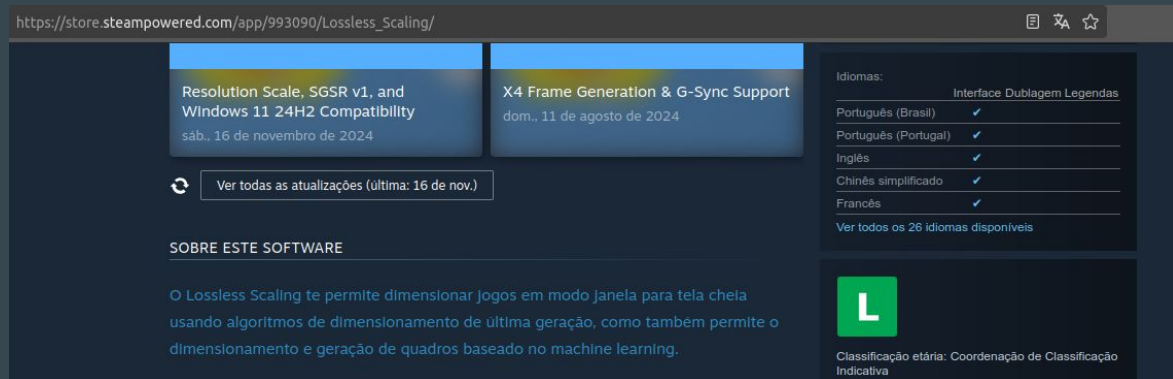
- Jogos ordenados em lista de popularidade regional recente , com separação em páginas.
- Página pode ser alterada em um simples valor no URL de acesso, facilmente alterável (para todos os países, são lidas no máximo 150 páginas)
- Código HTML contém os IDs dos jogos claramente marcados em ‘data-ds-appid’, título também é extraído para poder ler e validar os arquivos JSON gerados.

```
<a class="search result row ds_collapse flag ds_flagged ds_owned app_impression tracked" href="https://store.steampowered.com/app/105600/Terraria/?snr=1 7 7 7000 150 3" data-ds-appid="105600" data-ds-itemkey="App 105600" data-ds-tagids="[1100689,3810,1662,3871,3859,21,3964]" data-ds-crtrids="[32940202]" onmouseover="GameHover( this, event, 'global hover', { 'type': 'app', 'id': 105600, 'v6': 1 } );" onmouseout="HideGameHover( this, event, 'global hover' )" data-search-page="3" data-gpnav="item" data-ds-steam-deck-compat-handled="true">
<a class="search result row ds_collapse flag app_impression tracked" href="https://store.steampowered.com/app/1304930/The_Outlast_Trials/?snr=1 7 7 7000 150 3" data-ds-appid="1304930" data-ds-itemkey="App 1304930" data-ds-tagids="[1667,3859,1685,3978,1721,3839,6730]" data-ds-descids="[1,2,5]" data-ds-crtrids="[33520780]" onmouseover="GameHover( this, event, 'global hover', { 'type': 'app', 'id': 1304930, 'v6': 1 } );" onmouseout="HideGameHover( this, event, 'global hover' )" data-search-page="3" data-gpnav="item" data-ds-steam-deck-compat-handled="true">
<a class="search result row ds_collapse flag app_impression tracked" href="https://store.steampowered.com/app/1774580/STAR_WARS_Jedi_Survivor/?snr=1 7 7 7000 150 3" data-ds-appid="1774580" data-ds-itemkey="App 1774580" data-ds-tagids="[19,21,29482,4182,1697,1695,3942]" data-ds-crtrids="[37913704,36135791,37385824]" onmouseover="GameHover( this, event, 'global hover', { 'type': 'app', 'id': 1774580, 'v6': 1 } );" onmouseout="HideGameHover( this, event, 'global hover' )" data-search-page="3" data-gpnav="item" data-ds-steam-deck-compat-handled="true">
<a class="search result row ds_collapse flag app_impression tracked" href="https://store.steampowered.com/app/524220/NieRAutomata/?snr=1 7 7 7000 150 3" data-ds-appid="524220" data-ds-itemkey="App 524220" data-ds-tagids="[1756,1742,7208,1646,19,1695,122]" data-ds-crtrids="[1012195,42407224]" onmouseover="GameHover( this, event, 'global hover', { 'type': 'app', 'id': 524220, 'v6': 1 } );" onmouseout="HideGameHover( this, event, 'global hover' )" data-search-page="3" data-gpnav="item" data-ds-steam-deck-compat-handled="true">
```

| Inserir termo de busca ou marcador | | Buscar | Ordenar por | Relevância | |
|--|--------------------------------------|-------------|-------------|-------------------|-----------|
| 4.906 resultados correspondem à sua busca. 412 títulos foram excluídos de acordo com as suas preferências. No entanto, esses títulos não fazem parte da primeira página de resultados. | | | | | |
| 4.906 resultados correspondem à sua busca. | | | | | |
|  | Balatro | 26/fev/2024 | -10% | Gratuito | R\$40,49 |
|  | Stardew Valley | 26/fev/2016 | -40% | Gratuito | R\$14,99 |
|  | Call of Duty®: Warzone™ | 16/mar/2022 | | Gratuito | |
|  | Black Myth: Wukong | 19/ago/2024 | | | R\$229,99 |
|  | Battlefield™ 2042 | 19/mar/2021 | -92% | | R\$19,92 |
|  | iRacing | 12/jun/2015 | -66% | | R\$11,21 |
|  | Street Fighter™ 6 | 29/jun/2023 | -30% | | R\$99,50 |
|  | The Sims™ 4 | 18/jun/2020 | | Gratuito | |
|  | Ready or Not | 13/ago/2023 | -30% | | R\$25,99 |
|  | Red Dead Redemption | 29/mar/2014 | -20% | | R\$199,00 |
|  | God of War | 14/jun/2022 | -60% | | R\$79,96 |
|  | Battlefield™ V | 23/mar/2020 | -93% | | R\$16,03 |
|  | Resident Evil Village | 7/mar/2021 | -80% | | R\$55,00 |
|  | Resident Evil 7 Biohazard | 24/jun/2017 | -60% | | R\$20,76 |
|  | State of Decay 2: Juggernaut Edition | 13/mar/2020 | -90% | | R\$7,99 |
|  | Ghost of Tsushima VERSÃO DO DIRETOR | 16/mar/2024 | -20% | | R\$199,92 |
|  | Mass Effect™ Legendary Edition | 14/mar/2021 | -92% | | R\$19,92 |
|  | Valheim | 2/fev/2021 | -30% | | R\$18,99 |
|  | Squad | 23/mar/2020 | -60% | | R\$53,99 |
|  | Automobilista 2 | 30/jun/2020 | -70% | Compatível com RV | R\$39,99 |
|  | Terraria | 16/mar/2011 | -30% | | R\$10,49 |
|  | The Outlast Trials | 5/mar/2024 | -50% | | R\$53,99 |
|  | STAR WARS Jedi: Survivor™ | 28/mar/2023 | -75% | | R\$74,75 |
|  | NieR:Automata™ | 17/mar/2017 | -80% | | R\$42,80 |
|  | Cuphead | 29/mar/2017 | -30% | | R\$25,89 |
| exibindo resultados 31-75 de 4906 | | | | | |

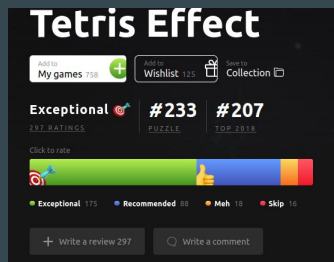
Dados de Entrada: Lista de Idiomas Disponíveis

- Página do jogo contém lista de idiomas no canto, são extraídos os idiomas com marcação na coluna de interface



```
<tr class="" style="">
  <td class="ellipsis" style="width: 94px; text-align: left">Português (Brasil)</td>
  <td class="checkcol">
    <span></span>
  </td>
  <td class="checkcol">
  </td>
  <td class="checkcol">
  </td>
</tr>
<tr class="" style="">
  <td class="ellipsis" style="width: 94px; text-align: left">Português (Portugal)</td>
  <td class="checkcol">
    <span></span>
  </td>
  <td class="checkcol">
  </td>
  <td class="checkcol">
  </td>
</tr>
<tr class="" style="">
  <td class="ellipsis" style="width: 94px; text-align: left">Inglês</td>
  <td class="checkcol">
    <span></span>
  </td>
  <td class="checkcol">
  </td>
  <td class="checkcol">
  </td>
</tr>
```


Dados de Entrada: Lista de Marcadores de Gênero (via Crawler)



```
<div class="game_head"> flex
  <div class="game_head-meta"> ... </div> flex
  <h1 class="heading heading_1 game_title" itemprop="name">Tetrtris Effect</h1>
</div>
```

About

From Tetsuya Mizuguchi and the same people who brought you the award-winning Rez Infinite and Lumines comes the new PS4 game TETRIS® EFFECT. Named after a real-world phenomenon where players' brains are so engrossed that images of the iconic falling Tetrimino blocks (i.e. the Tetris playing pieces) linger in their vision, thoughts, and even dreams, TETRIS EFFECT amplifies this magical feeling of total immersion by surrounding you with fantastic, fully three-dimensional worlds that react and evolve based on how you play. Music, backgrounds, sounds, speci... [Read more](#)

Platforms

PlayStation 4, PC

Metascore

90

Genre

Puzzle

Release date

Nov 9, 2018

Developer

Enhance Games

Age rating

10+ Everyone 10+

Other games in the series

Tetrtris Effect: Connected

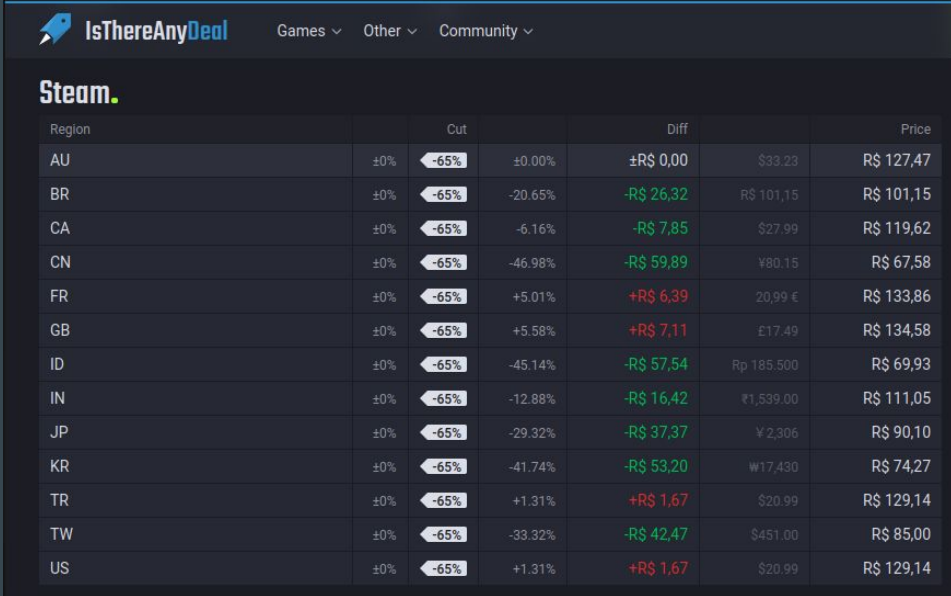
Tags

Atmospheric, Great Soundtrack, Classic, exclusive, puzzles, Tetris, vivid

```
<div class="game_meta"> flex
  <div class="game_meta-block"> ... </div>
  <div class="game_meta-block"> ... </div>
  <div class="game_meta-block">
    <div class="game_meta-title">Genre</div>
    <div class="game_meta-text">
      <meta itemprop="genre" content="Puzzle">
      <a class="game_meta-filter-link" href="/games/puzzle">Puzzle</a> event
    </div>
  </div>
```

- No rawg.io, a Spider procurou o site por páginas de jogos cujos nomes ou link para compra correspondesse aos jogos previamente encontrados, para listar seus gêneros.

Dados de Entrada: Extração de Preços via API



The screenshot shows the 'IsThereAnyDeal' website interface. At the top, there's a navigation bar with the logo and links for 'Games', 'Other', and 'Community'. Below this, the 'Steam.' section is highlighted. The main content is a table comparing prices across different regions. The table has columns for Region, Cut, Diff, and Price. The 'Cut' column shows a consistent -65% discount across all regions. The 'Diff' column shows the price difference relative to the AU region, with values like ±R\$ 0,00 for AU, -R\$ 26,32 for BR, and -R\$ 59,89 for CN. The 'Price' column shows the final price in Brazilian Reals (R\$) for each region.

| Region | | Cut | | Diff | | Price |
|--------|-----|------|---------|------------|------------|------------|
| AU | ±0% | -65% | ±0.00% | ±R\$ 0,00 | \$33.23 | R\$ 127,47 |
| BR | ±0% | -65% | -20.65% | -R\$ 26,32 | R\$ 101,15 | R\$ 101,15 |
| CA | ±0% | -65% | -6.16% | -R\$ 7,85 | \$27.99 | R\$ 119,62 |
| CN | ±0% | -65% | -46.98% | -R\$ 59,89 | ¥80.15 | R\$ 67,58 |
| FR | ±0% | -65% | +5.01% | +R\$ 6,39 | 20,99 € | R\$ 133,86 |
| GB | ±0% | -65% | +5.58% | +R\$ 7,11 | £17.49 | R\$ 134,58 |
| ID | ±0% | -65% | -45.14% | -R\$ 57,54 | Rp 185.500 | R\$ 69,93 |
| IN | ±0% | -65% | -12.88% | -R\$ 16,42 | ₹1,539.00 | R\$ 111,05 |
| JP | ±0% | -65% | -29.32% | -R\$ 37,37 | ¥ 2,306 | R\$ 90,10 |
| KR | ±0% | -65% | -41.74% | -R\$ 53,20 | ₩17,430 | R\$ 74,27 |
| TR | ±0% | -65% | +1.31% | +R\$ 1,67 | \$20.99 | R\$ 129,14 |
| TW | ±0% | -65% | -33.32% | -R\$ 42,47 | \$451.00 | R\$ 85,00 |
| US | ±0% | -65% | +1.31% | +R\$ 1,67 | \$20.99 | R\$ 129,14 |

Tabela de preços do jogo na Steam em diferentes regiões.

Por ser conteúdo dinâmico, tivemos de usar a API do site para obter os dados.

Exemplo dos dados de saída: Popularidade

`{}` list_BR.json ×

src > regional_lists > `{}` list_BR.json > `{}` 0 > title

```
1  [
2    {
3      "appid": "2694490",
4      "title": "Path of Exile 2"
5    },
6    {
7      "appid": "730",
8      "title": "Counter-Strike 2"
9    },
10   {
11     "appid": "2767030",
12     "title": "Marvel Rivals"
13   },
14 ]
```

Exemplo de lista de popularidade regional, nesse caso, para o Brasil. Existe um arquivo para cada país trabalhado.

`{}` list_total.json > ...

```
1  {
2    "2694490": {
3      "title": "Path of Exile 2",
4      "ranks": {
5        "rank_LK": 1,
6        "rank_JP": 1,
7        "rank_US": 1,
8        "rank_TH": 1,
9        "rank_SG": 1,
10       "rank_PH": 1,
11       "rank_PE": 1,
12       "rank_PA": 1,
13       "rank_MX": 1,
14       "rank_MY": 1,
15       "rank_LA": 1,
16       "rank_JM": 1,
17       "rank_ID": 1,
```

Lista com todas as popularidades regionais trabalhadas, gerada a partir dos arquivos intermediários. Nota-se que para jogos muito populares, é comum que estejam no topo em múltiplas regiões. Caso um jogo não aparece nos no placar de algum país, seu rank é listado como -1

Exemplo dos dados de saída: Preços

```
src > {} game_references.json > ...
```

```
1  {  
2    "2694490": "018d937f-798a-718a-8a1b-dc62e236c3bb",  
3    "730": "018d937f-7851-7004-b780-3f657a301f9a",  
4    "2767030": "018e807a-f53d-73bf-b813-23ce161f577f",  
5    "2677660": "018d937f-7b13-7325-845c-6f7e319f1a5a",  
6    "578080": "018d937f-57ce-723f-b7e7-7c7b5df93471",  
7    "2933620": "018fc559-910a-7290-82bc-55ef6bc1bc59",  
8    "1174180": "018d937f-3a3b-7210-bd2d-0d1dfb1d84c0",  
9    "570": "018d937f-19a5-7057-bb6d-314d586e6dbc",  
10   "1293830": "018d937f-0686-7187-9bf1-c96d4b6cdc4a",  
11   "2669320": "0190c17a-16d1-72d2-8ab6-5d5f3853e03a",  
12   "227300": "018d937f-20cc-703a-963c-8dd3f93cb88b",  
13   "1962663": "018d937f-3fb1-73ea-becb-da9c7b2d716c",  
14   "1643320": "018f4953-3ed5-70b5-a582-94a4de019373",  
15   "252490": "018d937f-060f-705e-b364-8c5bba4d62f6",  
16   "2507950": "01917195-2e81-724c-90dc-aef4dd164330",  
17   "2074920": "018d937f-6eea-7164-b49f-cc638d28f549",
```

Lista relacionando os IDs na loja trabalhada (Steam) e os IDs das entradas equivalentes na API consultada.

```
r_currency > {} prices_AED.json > ...
```

```
"018d937e-ea24-70bd-91e9-b1a8c6782f58": 34.99,  
"018d937e-f022-7110-a507-755f61a998cc": 19.99,  
"018d937e-f06b-7284-bf6c-c4e4381bb40a": 0,  
"018d937e-f06b-7284-bf6c-c4e440623880": 49.99,  
"018d937e-f4a8-70de-8c52-07048217de4c": 29.99,  
"018d937e-f4e1-72d7-80a8-3946c6237baf": 59.99,  
"018d937e-f4fb-70ee-b1c7-ff968a297b4c": 19.99,  
"018d937e-f518-7166-9649-affe7f72a110": 35,  
"018d937e-f519-71e5-b1d0-b00cbb24c647": 49.99,  
"018d937e-f803-7329-b166-49740f47e000": 19.99,  
"018d937e-f81a-7288-868c-49236397885a": 59.99,  
"018d937e-f835-710c-b95b-928e277b187e": 14.99,  
"018d937e-f85d-703b-aa5d-a38a1c1f5bc2": 29.99,  
"018d937e-f886-70e0-a872-4970e8bd9503": 39.99,  
"018d937e-f8b9-702f-ab65-2b3ba5da16e6": 0,  
"018d937e-f8e0-72c9-8f3c-0c3f92129010": 0,  
"018d937e-f98c-71cb-9624-eb9231e0b0ea": 29.99,
```

Como as popularidades, os preços são salvos em múltiplos arquivos diferentes, um para cada moeda. Exemplo para dirham.

Exemplo dos dados de saída: Gêneros

```
src > gamecrawl > {} output.json > ...
```

```
1  [
2  {"801800": {"title": "Atomfall", "genres": ["RPG"], "url": "https://rawg.io/games/atomfall"}},
3  {"1845910": {"title": "Dragon Age™: The Veilguard", "genres": ["Action", "Adventure", "RPG", "e
4  {"2429640": {"title": "THRONE AND LIBERTY", "genres": ["Action", "Adventure", "RPG", "Massively
5  {"1771300": {"title": "Kingdom Come: Deliverance II", "genres": ["Action", "Adventure", "RPG"],
6  {"2933620": {"title": "Call of Duty®: Black Ops 6", "genres": ["Action", "Shooter"], "url": "ht
7  {"2679460": {"title": "Metaphor: ReFantazio", "genres": ["Action", "Adventure", "RPG"], "url":
8  {"1085660": {"title": "Destiny 2", "genres": ["Action", "Shooter"], "url": "https://rawg.io/ga
9  {"42670": {"title": "Singularity™", "genres": ["Shooter"], "url": "https://rawg.io/games/singu
10 {"295250": {"title": "Stranded", "genres": ["Action", "Adventure", "Strategy", "Indie"], "url":
11 {"314790": {"title": "Silence", "genres": ["Action", "Adventure", "Indie"], "url": "https://rav
12 {"206420": {"title": "Saints Row IV: Re-Elected", "genres": ["Action", "Adventure"], "url": "ht
13 {"536220": {"title": "The Walking Dead: A New Frontier", "genres": ["Action", "Adventure"], "ur
14 {"282900": {"title": "Hyperdimension Neptunia Re;Birth1", "genres": ["RPG", "Strategy"], "url":
15 {"47790": {"title": "Medal of Honor™", "genres": ["Shooter"], "url": "https://rawg.io/games/mec
16 {"256190": {"title": "Enemy Front", "genres": ["Action", "Shooter"], "url": "https://rawg.io/ga
17 {"214490": {"title": "Alien: Isolation", "genres": ["Action", "Adventure"], "url": "https://rav
18 {"333950": {"title": "Medieval Engineers", "genres": ["Action", "Strategy", "Simulation", "Indi
19 {"12830": {"title": "Operation Flashpoint: Dragon Rising", "genres": ["Action", "Shooter"], "ur
20 {"578330": {"title": "LEGO® City Undercover", "genres": ["Action", "Adventure"], "url": "https:
21 {"9010": {"title": "Return to Castle Wolfenstein", "genres": ["Shooter"], "url": "https://rawg.
22 {"1030830": {"title": "Mafia II: Definitive Edition", "genres": ["Action", "Adventure"], "url":
23 {"555570": {"title": "Infestation: The New Z", "genres": ["Action", "RPG", "Massively Multiplay
```

Arquivo gerado ao rodar a Spider com output em um arquivo “output.json”, contendo os gêneros dos jogos extraídos do site rawg.io.

Exemplo dos dados de saída: Idiomas

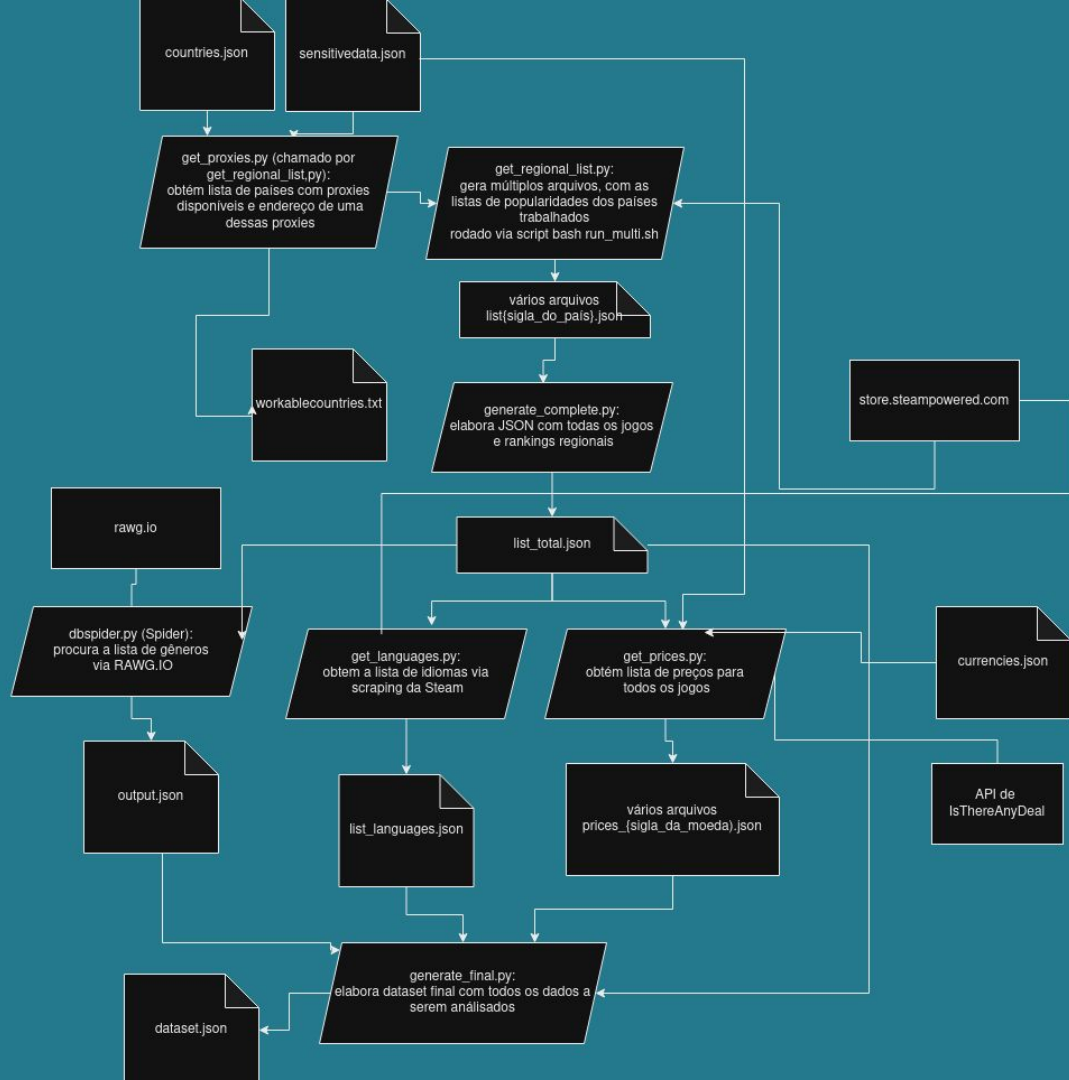
```
{ } list_languages.json > ...  
1  {  
2    "2694490": {  
3      "languages": [  
4        "English",  
5        "French",  
6        "German",  
7        "Spanish - Spain",  
8        "Japanese",  
9        "Korean",  
10       "Portuguese - Brazil",  
11       "Russian",  
12       "Thai"  
13     ],  
14     "url": "https://store.steampowered.com/app/2694490"  
15   },
```

Arquivo contendo os idiomas de todos os jogos trabalhados.

Dataset Final

| dataset.json M | dataset.json M |
|--|--|
| <pre>{ "22380": { "title": "Fallout: New Vegas", "ranks": { "rank_LK": 466, "rank_VN": 461, "rank_CN": 1636, "rank_JP": 2251, "rank_KR": 933, "rank_PH": 469, "rank_PE": 480, "rank_PA": 480, "rank_MX": 480, "rank_MY": 475, "rank_LA": 469, "rank_JM": 475, "rank_ID": 475, "rank_EC": 480, "rank_DO": 475, "rank_CR": 449, "rank_CO": 449, "rank_CL": 448, "rank_CA": 597, "rank_KH": 439, "rank_BO": 449, "rank_AR": 449, "rank_BR": 435, "rank_AE": 437, "rank_TR": 1011, "rank_CH": 397, "rank_ZA": 446, "rank_SA": 436, "rank_PK": 437, "rank_NO": 1071, "rank_IN": 443, "rank_AU": 545, "rank_GB": 455, "rank_US": -1, "rank_TH": -1, "rank_TW": -1, "rank_SG": -1, "rank_HK": -1, "rank_UA": -1 } } }</pre> | <pre>{ "rank_IL": -1, "rank_AZ": -1 }, "languages": ["English", "French", "German", "Italian", "Spanish - Spain"], "genres": ["Action", "Shooter", "RPG"], "url": ["https://store.steampowered.com/app/22380", "https://rawg.io/games/fallout-new-vegas"], "prices": { "price_LKR": 9.99, "price_VND": 9.99, "price_CNY": 41, "price_JPY": 1100, "price_USD": 9.99, "price_THB": 9.99, "price_TWD": 299, "price_KRW": 11000, "price_SGD": 9.99, "price_PHP": 9.99, "price_PEN": 9.99, "price_PAB": 9.99, "price_MXN": 9.99, "price_MYR": 9.99, "price_LAK": 9.99, "price_JMD": 9.99, "price_IDR": 89999, "price_HKD": 9.99, "price_DOP": 9.99, "price_CRC": 9.99, "price_COP": 9.99, "price_CLP": 9.99, "price_CAD": 13.49 }</pre> |

Descrição sobre o trabalho



Descrição sobre o trabalho

- Qual o nome do crawler utilizado?
 - GameCrawl (dbspider.py)
- Infraestrutura tecnológica usada:
 - Python, JSON, Scrapy (CrawlSpider e Rules), Regex, ItemAdapter, BeautifulSoup
- Qual foi a abordagem e métodos usados?
 - Extração de dados via requisição HTTP e parsing utilizando BeautifulSoup.
 - Extração de dados dinâmicos via API do IsThereAnyDeal.

Descrição sobre o trabalho

- Onde os dados foram coletados/extraídos?
 - <https://rawg.io>
 - <http://store.steampowered.com>
 - <https://api.isthereanydeal.com>
- Quem estaria interessado em comprar os dados coletados/extraídos?
 - Jogadores, donos de lojas de jogos e produtores
- De quais locais ou áreas geográficas os dados foram coletados/extraídos?
 - Os dados foram coletados de 45 países diferentes (lista no arquivo `workable_countries.txt`).

Descrição sobre o trabalho

- Qual o idioma considerado?
 - Inglês.
- Qual foi o formato de dados coletado/extraídos?
 - JSON.
- Quais foram os atributos coletados/extraídos? Identificá-los como dados qualitativos, quantitativos ou ambos.
 - Nome (qualitativo), ID na Steam (qualitativo), popularidade no país (qualitativo), idioma (qualitativo), URL (qualitativo), UUID na API IsThereAnyDeal (qualitativo) e preço (quantitativo).

Descrição sobre o trabalho

- Quais fontes de dados devem ser usadas para coletar/extrair informações? Listas as seeds – identificá-las e mostrar as que não eram seeds e que foram alcançadas.
 - <https://rawg.io> (seed)
 - <https://rawg.io/games/<game-name>>
 - <http://store.steampowered.com> (seed)
 - http://store.steampowered.com/search/?sort_by=&sort_order=0&category1=998&filter=topsellers&page=<iterator>
 - <http://store.steampowered.com/app/<game-id>>
 - <https://api.isthereanydeal.com> (seed)
 - <https://api.isthereanydeal.com/games/<API>>

Descrição sobre o trabalho

- Por quanto tempo os dados foram coletados (o escopo da coleta/extração de dados)?
 - Por mais de duas horas de cada fonte.
- Datas da coleta:
 - <https://rawg.io> (16/12/2024)
 - <http://store.steampowered.com> (04/12/2024, 09/12/2024 e 17/12/2024)
 - <https://api.isthereanydeal.com> (19/12/2024)

Análise dos resultados finais: Metodologia para lidar com Ranking

- Apenas o ranking dos jogos, por país, foi extraída. Foi necessária alguma forma para transformar essa métrica em algo tratável na análise de dados.
- Solução: inverter a ordem dos rankings (se em um país foram coletados 500 jogos, o jogo #1 se torna o #500), e aplicar operação de logaritmo sobre esse novo valor somado a um. A operação de logaritmo reduz a diferença de peso entre jogos com posições similares.
- Para dados qualitativos (gênero e idiomas), foi gerada uma soma destas pontuações para avaliar o efeito deste na popularidade regional.
- Para dados quantitativos (preço e método de conversão para preços regionais), foi analisada a correlação entre esta e a pontuação de popularidade.

Análise de Gêneros de Jogos

| | |
|-----------------------|-------------|
| genre | |
| Action | 8683.159092 |
| Indie | 8411.409599 |
| Adventure | 7026.791256 |
| Simulation | 5175.633547 |
| RPG | 4672.415709 |
| Strategy | 4530.446147 |
| Casual | 3494.702496 |
| Shooter | 1308.751753 |
| Massively Multiplayer | 1117.254897 |
| Sports | 703.567680 |
| Puzzle | 697.146174 |
| Racing | 651.762865 |
| Platformer | 494.265797 |
| Arcade | 300.855414 |
| Fighting | 275.212406 |
| Family | 200.304631 |
| Board Games | 118.499621 |
| Card | 114.160003 |
| Educational | 103.729076 |

| | |
|-----------------------|-------------|
| genre | |
| Indie | 7873.653038 |
| Action | 6924.773733 |
| Adventure | 6010.625233 |
| Simulation | 4131.071663 |
| Casual | 3845.524453 |
| RPG | 3427.435948 |
| Strategy | 3234.611312 |
| Shooter | 967.295233 |
| Puzzle | 651.593383 |
| Massively Multiplayer | 633.332430 |
| Sports | 486.099284 |
| Racing | 378.595185 |
| Platformer | 378.011091 |
| Arcade | 331.225515 |
| Fighting | 232.520870 |
| Family | 114.755667 |
| Card | 107.727424 |
| Educational | 68.866768 |
| Board Games | 48.433976 |

Rankings de popularidade de gêneros para Estados Unidos e Coreia do Sul

Análise de Idiomas de Jogos

| language | | language | | language | |
|-------------------------|--------------|-------------------------|--------------|-------------------------|-------------|
| English | 20463.237634 | English | 20612.562504 | English | 9519.849572 |
| German | 15766.284541 | Simplified Chinese | 17057.242897 | German | 7316.645305 |
| French | 15724.235484 | Japanese | 14941.688133 | French | 7306.622955 |
| Spanish - Spain | 14283.683865 | German | 14619.960857 | Spanish - Spain | 6674.990063 |
| Simplified Chinese | 13768.075434 | French | 14512.353917 | Simplified Chinese | 6229.755810 |
| Japanese | 13259.062116 | Spanish - Spain | 13407.374857 | Russian | 6029.897466 |
| Russian | 12571.177430 | Russian | 12458.030672 | Japanese | 5896.010163 |
| Italian | 11938.345807 | Korean | 11786.285286 | Italian | 5635.878920 |
| Korean | 10881.135459 | Traditional Chinese | 11426.534915 | Portuguese - Brazil | 5179.305806 |
| Portuguese - Brazil | 10822.373931 | Italian | 10683.868886 | Korean | 4931.473911 |
| Traditional Chinese | 9606.205100 | Portuguese - Brazil | 10423.023895 | Polish | 4429.675758 |
| Polish | 8928.921261 | Polish | 8331.840252 | Traditional Chinese | 4181.907221 |
| Turkish | 4999.351957 | Turkish | 5134.573156 | Turkish | 2678.374262 |
| Spanish - Latin America | 4795.298334 | Spanish - Latin America | 4537.981140 | Spanish - Latin America | 2236.144944 |
| Portuguese - Portugal | 3166.541905 | Portuguese - Portugal | 3304.027135 | Portuguese - Portugal | 1753.217359 |
| Czech | 2870.406705 | Czech | 2671.126756 | Dutch | 1520.094118 |
| Dutch | 2714.565206 | Dutch | 2594.209077 | Czech | 1487.935669 |
| Ukrainian | 2264.075448 | Ukrainian | 2326.410710 | Ukrainian | 1218.459060 |
| Thai | 1847.247250 | Thai | 2014.314262 | Swedish | 1062.557596 |
| Arabic | 1804.855358 | Arabic | 1682.500109 | Thai | 944.699109 |

Rankings de popularidade de jogos com certos idiomas para Ucrânia, China e Noruega.
Nota-se que o efeito da disponibilidade do idioma local tem efeitos diferentes entre países

Correlação entre preço e popularidade

- Correlação entre o preço dos jogos e sua popularidade
- Não inclui jogos gratuitos
- Números positivos indicam maior popularidade de produtos mais caros, e negativos, mais baratos.
- Proximidade de zero indica pouca correlação entre preços e hábitos de compra.
- Exemplos para Brasil, Austrália e Grã-Bretanha.

| | price | rank |
|-------|----------|----------|
| price | 1.000000 | 0.360432 |
| rank | 0.360432 | 1.000000 |

| | price | rank |
|-------|----------|----------|
| price | 1.000000 | 0.193093 |
| rank | 0.193093 | 1.000000 |

| | price | rank |
|-------|----------|---------|
| price | 1.000000 | 0.31225 |
| rank | 0.31225 | 1.00000 |

Correlação entre taxa de conversão e popularidade

- Como jogos na Steam tem seu preço padrão em dólar, analisamos a correlação entre a popularidade e a taxa de conversão utilizada
- Por exemplo, para certos países com economias mais fracas, é relativamente comum que os preços na região sejam colocados abaixo da conversão direta em dólar. Queremos avaliar se há uma correlação perceptível neste quesito.
- Exemplos da Índia, México e Ucrânia.

| | conversion | rank |
|------------|------------|----------|
| conversion | 1.000000 | 0.007383 |
| rank | 0.007383 | 1.000000 |

| | conversion | rank |
|------------|------------|----------|
| conversion | 1.000000 | 0.298501 |
| rank | 0.298501 | 1.000000 |

| | conversion | rank |
|------------|------------|-----------|
| conversion | 1.000000 | -0.224975 |
| rank | -0.224975 | 1.000000 |

Casos de falha

- Falhas do processo de extração:
 - No site IsThereAnyDeal, como o conteúdo do web app era carregado por um script JS dinamicamente, a div onde ele se encontrava estava vazia na resposta da requisição.
 - No site rawg.io, houve uma tentativa de dinamicamente encontrar os mesmos jogos que foram encontrados pela Steam. Contudo, em 4% dos jogos, não foi encontrado uma página representativa pela Spider.
 - Para alguns jogos não foi possível coletar os preços, devido a respostas sem conteúdo no corpo.

Conclusão

- O que teriam feito diferente?
 - Sentimos que poderíamos ter feito uma análise dos dados mais visual e mais completa para retratar melhor a realidade.
- O trabalho tem potencial para ser algo visto com bons olhos pelo mercado? Por quê? O que faltaria desenvolver para ser um produto?
 - Sim, principalmente para o público gamer. Se tivéssemos mais pessoas na equipe responsáveis pela análise dos dados, poderia ser utilizado para recomendar jogos aos usuários.