

Task 2

Done by jana faour

1.A brief description of the website you chose and what data you scraped?

About website:

The website name is Arcom group its an e-commerce platform specializing in tiles and other related products. It features a catalog of products that users can browse, with each product having specific details like size, price, and images.

Purpose of Scraping:

The goal of scraping this data is to gather essential information about the products for purposes like displaying them on another platform, analyzing prices, or inventory management.

2.The Java code you wrote for the `CommandLineRunner`?

```
WebScrapingApplication v | ▶ | 🔍 | ⚙️ | -
WebScrapingApplication.java pom.xml (web-scraping) Tiles.java ScrapingService.java x ScrapingApplication.java
6 import org.jsoup.select.Elements;
7
8 import java.io.IOException;
9 import java.util.ArrayList;
10 import java.util.List;
11
12 public class ScrapingService { 2 usages
13
14 @ public static List<Tiles> scrapeTiles(String url) throws IOException { 1 usage
15     List<Tiles> tilesList = new ArrayList<>();
16     Document doc = Jsoup.connect(url).get();
17
18     // Select all elements that represent a tile product
19     Elements productElements = doc.select(cssQuery: "div.product-index");
20
21     for (Element productElement : productElements) {
22         // Extract name
23         String name = productElement.select(cssQuery: "h3.product-title a").text();
24
25         // Extract link
26         String link = "https://arcomgroup.com" + productElement.select(cssQuery: "h3.product-title a").attr(attributeKey: "href");
27
28         // Extract image URL
29         String imageUrl = "https:" + productElement.select(cssQuery: ".prod-image img").attr(attributeKey: "src");
30
31         // Extract price (check for sale price first, then regular price)
32         String price = productElement.select(cssQuery: ".price__sale .price-item--sale").text();
33         if (price.isEmpty()) {
34             price = productElement.select(cssQuery: ".price__regular .price-item--regular").text();
35         }
36
37         // Extract size
38         String size = productElement.select(cssQuery: ".size-swatch .options li").attr(attributeKey: "data-option-title");
39
40         // Create Tiles object with extracted data
41         Tiles tile = new Tiles(name, link, imageUrl, price, size);
42
43         // Add the tile to the list
44         tilesList.add(tile);
45     }
46
47     return tilesList;
}
```

3.A screenshot of the console output showing the scraped data.

Enter link to scrape or 'exit' to stop:

<https://arcomgroup.com/collections/tiles?page=5>

Tiles{name='Baltra Pearl Pulido', link='https://arcomgroup.com/collections/tiles/products/t-por-s-tn-fa-24803', imageUrl='https://arcomgroup.com/cdn/shop/files/24803-BALTRAPEARLPULIDO0X120.jpg?v=1714986940&width=600', price='4

Tiles{name='Camile Gold Pulido', link='https://arcomgroup.com/collections/tiles/products/t-por-s-tn-fa-24790', imageUrl='https://arcomgroup.com/cdn/shop/files/24790-CAMILLEGOLDPULIDO0X120.jpg?v=1714987191&width=600', price='\$36

Tiles{name='Sakhir Natural', link='https://arcomgroup.com/collections/tiles/products/tiles-23-t-por-s-tn-fa-23929', imageUrl='https://arcomgroup.com/cdn/shop/files/23929-SAKHIRNATURAL60X120ANTIDTRECT_1.jpg?v=1714741958&wid

Tiles{name='Sakhir Grey', link='https://arcomgroup.com/collections/tiles/products/tiles-23-t-por-s-tn-fa-23928', imageUrl='https://arcomgroup.com/cdn/shop/files/23928-SAKHIRGREY60X120ANTIDTRECT.jpg?v=1714987428&width=600',

Tiles{name='Symphony White Pulido', link='https://arcomgroup.com/collections/tiles/products/tiles-23-t-por-s-tn-fa-21332', imageUrl='https://arcomgroup.com/cdn/shop/files/21332-SYMPHONYWHITE60X120PUL.jpg?v=1714987688&widht=

Tiles{name='Symphony Noir Pulido', link='https://arcomgroup.com/collections/tiles/products/tiles-23-t-por-s-tn-fa-21331', imageUrl='https://arcomgroup.com/cdn/shop/files/21331-SYMPHONYNOIRPULIDO60X120.jpg?v=1714987686&widht=600

Tiles{name='Fiorland White Pulido', link='https://arcomgroup.com/collections/tiles/products/tiles-23-t-por-s-tn-fa-21328', imageUrl='https://arcomgroup.com/cdn/shop/files/21328-FIORLANDWHITEPULIDO60X120.png?v=1714987968&widht=

Tiles{name='Even White Pulido', link='https://arcomgroup.com/collections/tiles/products/tiles-23-t-por-s-tn-fa-21327', imageUrl='https://arcomgroup.com/cdn/shop/files/21327-EVENWHITEPULIDO60X120.jpg?v=1714984226&widht=600', pr

Tiles{name='Baltra Ivory Pulido', link='https://arcomgroup.com/collections/tiles/products/tiles-23-t-por-s-tn-fa-21323', imageUrl='https://arcomgroup.com/cdn/shop/files/21323-BALTRAIVORYPULIDO60X120.jpg?v=1714988576&widht=600',

Tiles{name='Roma Pietra Grey', link='https://arcomgroup.com/collections/tiles/products/tiles-23-t-por-i-fa-fa-24595', imageUrl='https://arcomgroup.com/cdn/shop/files/roma-stone-pietra-grey-matt.jpg?v=1714992629&widht=600', f

Tiles{name='Onyx Turquoise Pulido', link='https://arcomgroup.com/collections/tiles/products/tiles-23-t-por-s-ba-fa-22549', imageUrl='https://arcomgroup.com/cdn/shop/files/22549-ONYXTURQUOISEPULIDO60X120.jpg?v=1714814991&wid

Tiles{name='Onyx Bleu Pulido', link='https://arcomgroup.com/collections/tiles/products/tiles-23-t-por-s-ba-fa-22546', imageUrl='https://arcomgroup.com/cdn/shop/files/22546-ONYXBLEUPULIDO60X120.jpg?v=1714815858&widht=600', pr

Tiles{name='MoonLight Pulido', link='https://arcomgroup.com/collections/tiles/products/tiles-23-t-por-s-ba-fa-22545', imageUrl='https://arcomgroup.com/cdn/shop/files/22545-MOONLIGHTPULIDO60X120.jpg?v=1714815037&widht=600', f

Tiles{name='Lemurian Pulido', link='https://arcomgroup.com/collections/tiles/products/tiles-23-t-por-s-ba-fa-22543', imageUrl='https://arcomgroup.com/cdn/shop/files/22543-LEMURIANPULIDO60X120.1.jpg?v=1714813379&widht=600', f

Tiles{name='Manaos Green Pulido', link='https://arcomgroup.com/collections/tiles/products/tiles-23-t-por-s-ba-fa-21197', imageUrl='https://arcomgroup.com/cdn/shop/files/21197-MANAOGREENPULIDO60X120.jpg?v=1714813385&widht=60

Tiles{name='Atlantida Pulido', link='https://arcomgroup.com/collections/tiles/products/tiles-23-t-por-s-ba-fa-21196', imageUrl='https://arcomgroup.com/cdn/shop/files/21196-ATLANTIDAPULIDO60X120.jpg?v=1714813384&widht=600', f

Tiles{name='Nexside Blue Pulido', link='https://arcomgroup.com/collections/tiles/products/tiles-23-t-por-s-ba-fa-19931', imageUrl='https://arcomgroup.com/cdn/shop/files/19931-NEXSIDEBLUEPULIDO60X120.jpg?v=1714813040&widht=60

Tiles{name='Adaggio Gold Pulido', link='https://arcomgroup.com/collections/tiles/products/t-por-s-ba-fa-19905', imageUrl='https://arcomgroup.com/cdn/shop/files/19905-ADAAGGIOLDLPULIDO60X120.jpg?v=1714813042&widht=600', price

Tiles{name='Eternal Pearl Pulido', link='https://arcomgroup.com/collections/tiles/products/tiles-23-t-por-s-ba-fa-19320', imageUrl='https://arcomgroup.com/cdn/shop/files/19320-ETERNALPEARLPULIDO60X120.jpg?v=1714813041&widht=

Tiles{name='Patmos Pulido', link='https://arcomgroup.com/collections/tiles/products/tiles-23-t-por-s-ba-fa-16363', imageUrl='https://arcomgroup.com/cdn/shop/files/16363-PATMOSPULIDO60X120.jpg?v=1714812322&widht=600', price='

Tiles{name='Lilac Grigio Siena Lux', link='https://arcomgroup.com/collections/tiles/products/t-por-i-ak-a11-22606', imageUrl='https://arcomgroup.com/cdn/shop/products/22606-LILAC_GRIGIO_SIENA_LUX_120X280_e86c753a-055a-4fa3-l

Tiles{name='Oyster Gold Lux', link='https://arcomgroup.com/collections/tiles/products/t-por-i-ak-a11-24123', imageUrl='https://arcomgroup.com/cdn/shop/products/24123-OYSTER_GOLD_LUX_120X280_2774d99b-a5e4-49bc-8f2e-b5b7b48006

Tiles{name='Oyster White Lux', link='https://arcomgroup.com/collections/tiles/products/t-por-i-ak-a11-24124', imageUrl='https://arcomgroup.com/cdn/shop/products/24124-OYSTER_WHITE_LUX_120X280_050f2f03-7ea1-465e-9c06-0955b096f

Tiles{name='Calacatta Vena Oro', link='https://arcomgroup.com/collections/tiles/products/t-por-i-ak-a11-22607', imageUrl='https://arcomgroup.com/cdn/shop/products/22607-CALACATTA_VENA_ORO_LUX_120X280_975cf1d-b6f8-49b6-a0d0-

Tiles{name='Gems Crystal', link='https://arcomgroup.com/collections/tiles/products/t-por-i-ak-a11-24719', imageUrl='https://arcomgroup.com/cdn/shop/products/24719-GEMS_CRYSTAL_LUX_120X280_da5f8e13-255c-4da7-aaf4-e8c881c81d8f

Tiles{name='Calacatta Michelangelo', link='https://arcomgroup.com/collections/tiles/products/t-por-i-ak-a11-24126', imageUrl='https://arcomgroup.com/cdn/shop/products/24126-CALACATTA_MICHELANGELO_LUX_120X280_def3aed9-1858-4d

Tiles{name='Statuario Light', link='https://arcomgroup.com/collections/tiles/products/t-por-i-ak-a11-24132', imageUrl='https://arcomgroup.com/cdn/shop/products/24132-STATUARIO_LIGHT_LUX_120X280_c922d46b-980f-4d55-a360-815bbd

Tiles{name='Sensi Carrara', link='https://arcomgroup.com/collections/tiles/products/t-por-i-ak-a11-24720', imageUrl='https://arcomgroup.com/cdn/shop/products/24720-CARRARA_LUX_120X280_7697d9ee-0a90-43cb-9f08-25075814832e.jpg

Tiles{name='Calacatta Bright Lux', link='https://arcomgroup.com/collections/tiles/products/t-por-i-ak-a11-24724', imageUrl='https://arcomgroup.com/cdn/shop/files/24724-CalacattaBrightLux120x280_1.jpg?v=1714483707&widht=800',

Tiles{name='Calacatta Bright Megabarocco Lux', link='https://arcomgroup.com/collections/tiles/products/t-por-i-ak-a11-24726', imageUrl='https://arcomgroup.com/cdn/shop/files/24726-CALACATTABRIGHTMEGAROCOCCOLUX120X280_1.jpg?v=

Tiles{name='Statuario White Megabarocco Lux', link='https://arcomgroup.com/collections/tiles/products/t-por-i-ak-a11-24725', imageUrl='https://arcomgroup.com/cdn/shop/files/24725-STATUARIOWHITEMEGAROCOCCOLUX120X280_1.jpg?v=

Tiles{name='Statuario White Lux Bookmatch', link='https://arcomgroup.com/collections/tiles/products/t-por-i-ak-a11-24723', imageUrl='https://arcomgroup.com/cdn/shop/files/24723-STATUARIOWHITELUX_A_120X280_1.jpg?v=1714483798f

Tiles{name='Calacatta Michelangelo Antique 3D', link='https://arcomgroup.com/collections/tiles/products/tiles-23-t-por-i-ak-fa-24130', imageUrl='https://arcomgroup.com/cdn/shop/products/24130-CALACATTA_MICHELANGELO_ANTIQUE_3

Tiles{name='Statuario Light Antique 3D', link='https://arcomgroup.com/collections/tiles/products/tiles-23-t-por-i-ak-fa-24129', imageUrl='https://arcomgroup.com/cdn/shop/products/24129-STATUARIO_LIGHT_ANTIQUE_3D_60X120_a9d8c

Tiles{name='Carrara Antique 3D', link='https://arcomgroup.com/collections/tiles/products/tiles-23-t-por-i-ak-fa-24115', imageUrl='https://arcomgroup.com/cdn/shop/products/24115-Carrara_Antique_3D_60X120_dfc53f93-d9cb-4484-9f