

Тестовое задание

Задача

Вам предстоит собрать информацию с сайта <https://www.petsonic.com>

Нас интересует та часть сайта, который занимается продажей товаров для животных - <https://www.petsonic.com/dermatitis-y-problemas-piel-para-perros/>

Нужно собрать **информацию о всех товарах** из категории и записать в **csv** файл.

По каждому товару нужно собрать три параметра

1. Название товара
2. Цену
3. Изображение

Важно! На случай проблем с доступностью категории, обратитесь к (HR) менеджер

Пример разбора

Вот пример страницы товара

<https://www.petsonic.com/pienso-para-perros-affinity-advance-adulto-razas-medianas.html>

🏠 > Perros > Comida > Piensos para perros > Advance Adulto Razas Medianas pienso para perros Medium breed



Название → **Advance Adulto Razas Medianas Pienso Para Perros Medium Breed**

Affinity Advance Adulto de Razas Medianas es un pienso para perros entre 1 y 7 años de edad. Su fórmula incorpora una adecuada combinación entre proteínas y grasas, para óptimo desarrollo y refuerzo de la salud intestinal y cardíaca



Referencia: 8410650150352

PESO	Весовка
3 Kg. 14.50 € 4.83 €/Kg	14 Kg. 44.95 € 3.21 €/Kg

Цена

Ahorra comprando 2 o más unidades...

x2 28.42 € 14.21 €	- 0.58 €	x4 56.24 € 14.06 €	- 1.74 €
--------------------------	----------	--------------------------	----------

Total: 14.50 € - 1 + Añadir al carrito

Такую страницу мы называем страницей мультипродукта. Это означает, что на одной странице находится информация о нескольких разновидностях продукта. В данном случае разная весовка: 3 kg, 14 kg.

С этой страницы в выходной файл должно быть добавлено 2 записи:

Name	Price	Image
Advance Adulto Razas Medianas Pienso Para Perros Medium Breed - 3 Kg.	14.50	https://www.petsonic.com/17833-thickbox_default/pienso-para-perros-affinity-advance-adulto-razas-medianas.jpg
Advance Adulto Razas Medianas Pienso Para Perros Medium Breed - 14 Kg	44.95	https://www.petsonic.com/17833-thickbox_default/pienso-para-perros-affinity-advance-adulto-razas-medianas.jpg

* Ссылки на картинки могут быть разными у разных вариаций

На что обратить внимание

1. Название сформировано, как общее название продукта **Advance Adulto Razas Medianas Pienso Para Perros Medium Breed** плюс весовка **3 Kg**.
2. Если у продуктов нет вариаций, то и весовка будет указана только одна. В таком случае с одной страницы продукта получится одна запись в файле
3. Список продуктов категории состоит из нескольких страниц

Требования к оформлению кода

1. Должна **быть написана программа/скрипт на Ruby**. Программа получает на вход:
 - a. ссылка на страницу категории (может передаваться любая категория сайта)
 - b. имя файла в который будет записан результат
2. После отработки скрипта результаты записывает в выходной файл с заданными именем
3. В тестовом задании нужно использовать **XPATH** (не путать с CSS) для получения содержимого html элементов, таких как цена, название и т.п. Возможно, понадобится использовать regex для того чтобы достать отдельные части данных, если xpath недостаточно
4. Для скачивания страниц нужно использовать библиотеку **curb** (*Могут быть проблемы под Windows, лучше использовать Linux. Если не получится использовать curb, используйте библиотеки open-uri или mechanize и подробно опишите, почему не использовалась библиотека curb*), для парсинга страниц **nokogiri**, для записи в csv файл – модуль **csv**
5. Вместе с решением нужно прислать пример запуска скрипта
6. Если в скрипте используются сторонние gem-ы, то приложить Gemfile
7. Скрипт должен работать достаточно быстро и при этом не привлекать к себе внимание администраторов сайта.
8. Во время выполнения скрипта должно быть понятно, что он сейчас делает

Теоретические вопросы

Если что-то не делали или не знаете - укажите, не оставляйте пустых ответов

Блок Q1

1. С какими ОС Вы работали? Работали ли с консолью? Какие консольные утилиты Linux Вы знаете и использовали?
2. Как Вы повышаете свою профессиональную квалификацию (названия книг, курсов, сайтов, скринкастов и т.д.)?
3. Какие сторонние библиотеки Вы используете чаще всего для разработки? Какие плюсы в них Вы выделяете для себя? (Ruby или тот язык, на котором пишете)

Блок Q2

1. Какие инструменты для профайлинга и дебага Вы используете? Какие у них минусы?
2. Объясните почему происходит следующее (**в контексте языка Ruby**):
 - a. $1660 / 100 \neq 16.6$
 - b. $24.0 * 0.1 \neq 2.4$
3. С какими СУБД вы работали? Проектировали ли вы свою собственную БД? Сколько было в ней таблиц? С какой самой большой таблицей по количеству записей Вы работали?
4. Оптимизировали ли Вы запросы в SQL? Как Вы это делали? Как Вы понимаете что запрос выполняется оптимально?

Как все отправить

Перед тем, как отправить ваши результаты, проверьте по следующему чеклисту:

1. Ваше резюме
2. Файл(ы) с кодом решения задачи
3. Команду на запуск скрипта
4. Скрин тестового запуска
5. Подробные ответы на все вопросы. Если что-то не делали или не знаете - укажите, не оставляйте пустых ответов.
6. Заархивируйте всё это в zip архив - не все почтовые клиенты пропускают скрипты

Данное задание является собственностью компании Profitero
Просьба не публиковать задание и решение в открытых источниках!

До встречи на собеседовании!