

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/snacks-huesos-para-perros/>

Нужно собрать **все товары** из категории и записать в **csv** файл.

По каждому товару нужно собрать три параметра

1. Полное название товара
2. Цену
3. Изображение

Важно! На случай проблем с доступностью категории, воспользуйтесь альтернативной ссылкой:

<https://www.petsonic.com/hobbit-half/>

Пример разбора

Вот пример страницы товара <https://www.petsonic.com/hobbit-alf-galletas-granja-para-perros.html>



Название

HOBBIT ALF
Galletas Granja para Perro

★★★★★ 0 opiniones

Las galletas de granja para perro tienen divertidas formas de animales y un delicioso aroma a vainilla. No contienen gluten.

Весовка

Peso	Precio IVA INCLUIDO	Cantidad
<input checked="" type="radio"/> 200 Gr.	1.50 € 1.35 €	- 1 +
<input type="radio"/> 430 gr	2.60 € 2.34 €	
<input type="radio"/> 4 Kg.	18.70 € 16.83 €	

Цена

Precio total: **1.35 €**

 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: 200 gr, 430 gr, 4 kg.

С этой страницы в выходной файл должно быть добавлено 3 записи:

Name	Price	Image
Galletas Granja para Perro - 200 gr	1.35	http://www.petsonic.com/5830-large_default/galletas-granja-para-perro.jpg
Galletas Granja para Perro - 430 gr	2.34	http://www.petsonic.com/5830-large_default/galletas-granja-para-perro.jpg
Galletas Granja para Perro - 4 kg	16.83	http://www.petsonic.com/5830-large_default/galletas-granja-para-perro.jpg

* Ссылки на картинки могут быть разными у разных вариаций

На что обратить внимание

1. Название сформировано, как общее название продукта *Galletas Granja para Perro* плюс весовка из таблицы *200 gr*
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** для получения содержимого html элементов, таких как цена, название и т.п. Возможно, понадобится использовать `gegexr` 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*), 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?

*на след. страничке очень важная информация!

!!! В каком виде нужно отправлять готовое ТЗ:

1. Ваше резюме (файл, а не ссылка)
2. Ссылка на репозиторий с решением задачи (либо файл с кодом)
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 `tar` или `zip` архив - не все почтовые клиенты пропускают скрипты и назовите данную папку «Фамилия Имя.ТЗ»

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

До встречи на собеседовании! ;)