CST3130 Advanced Web Development with Big Data

Coursework 1 Project Report

Jan Ciepiela M00779169

Project Overview

For the first piece of coursework I have created a project that shows the flow of data from being scraped from the website, to being put through to a database, to being displayed in a web application. The project consist of a few components working together. Java program is used to scrape multiple websites using JSoup to extract content from them, working on five threads, and place it into a SQL database with the help of Hibernate. The Java project is built using Maven, which also allows for executing JUnit tests created for the program. Another part is implemented using Javascript. Web service is used to retrieve data from the database and parse it to JSON, which is then retrieved by Axios and displayed on the web application using Vue.js. The dependencies and web service are managed with the help of Node.js and Express.js.

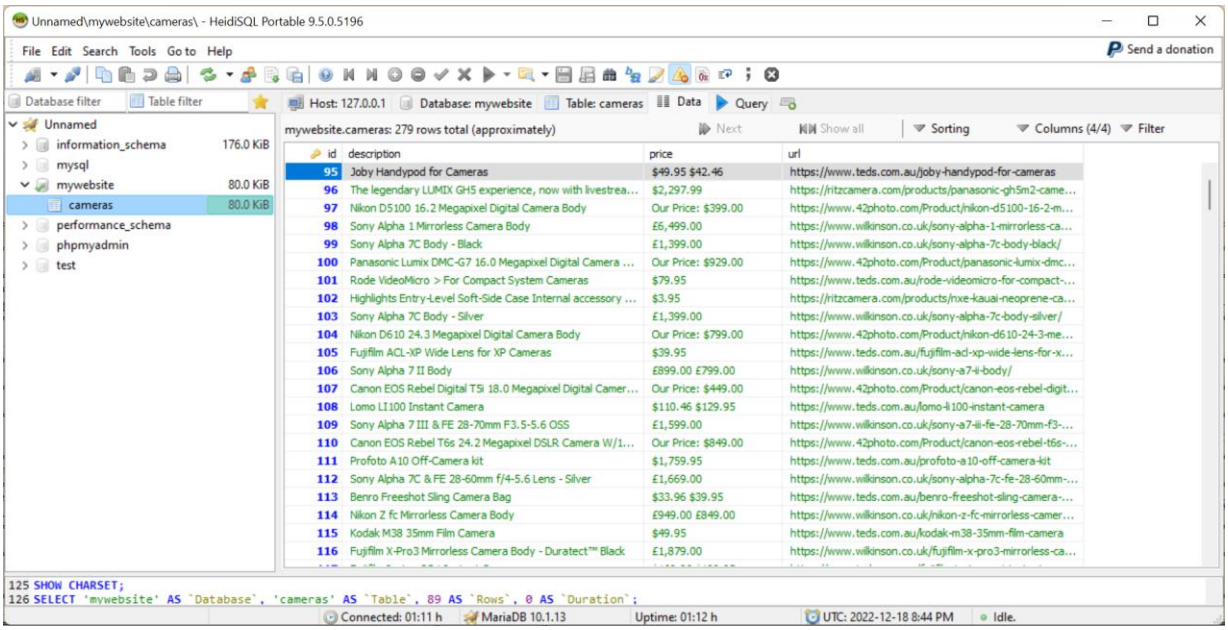
Websites scraped

The list of websites used for the purpose of the project:

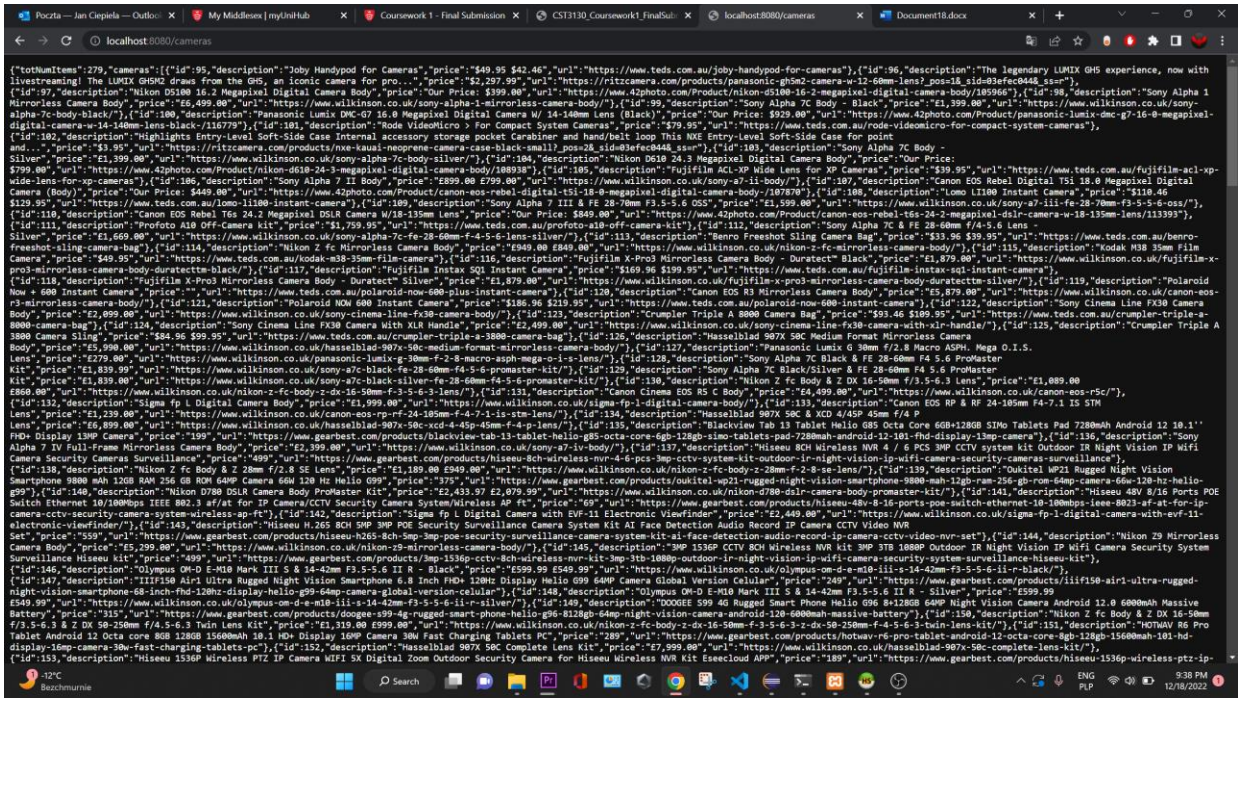
- Gearbest: https://www.gearbest.com/
- Ritzcamera: https://ritzcamera.com/
- Teds: https://www.teds.com.au/
- Wilkinson: https://www.wilkinson.co.uk/
- 42photo: https://www.42photo.com/

Data flow

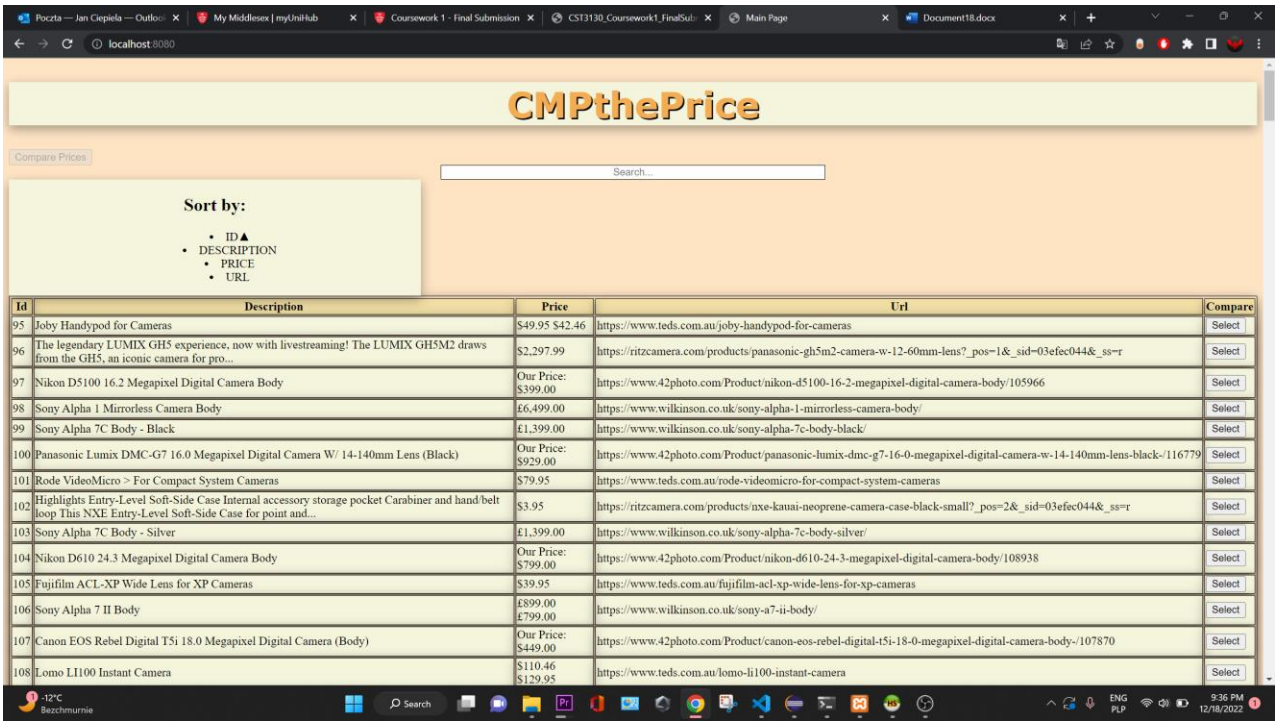
Once the data has been scraped, Hibernate is used to place it into a database. The view from HeidiSQL with the data placed in a “cameras” table:



The data is then retrieved by the web service and parsed into JSON. An example of the contents parsed after using the path “/cameras”:

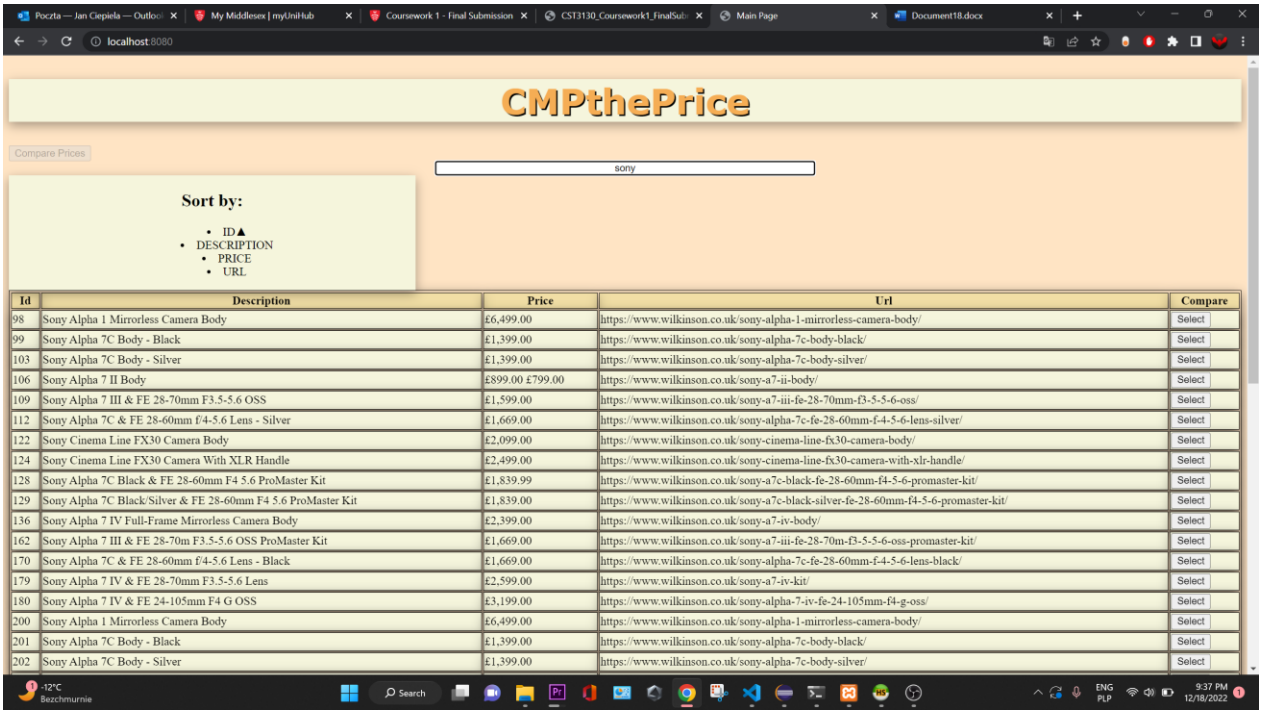


Data is then loaded into an array and displayed with the help of Vue.js:



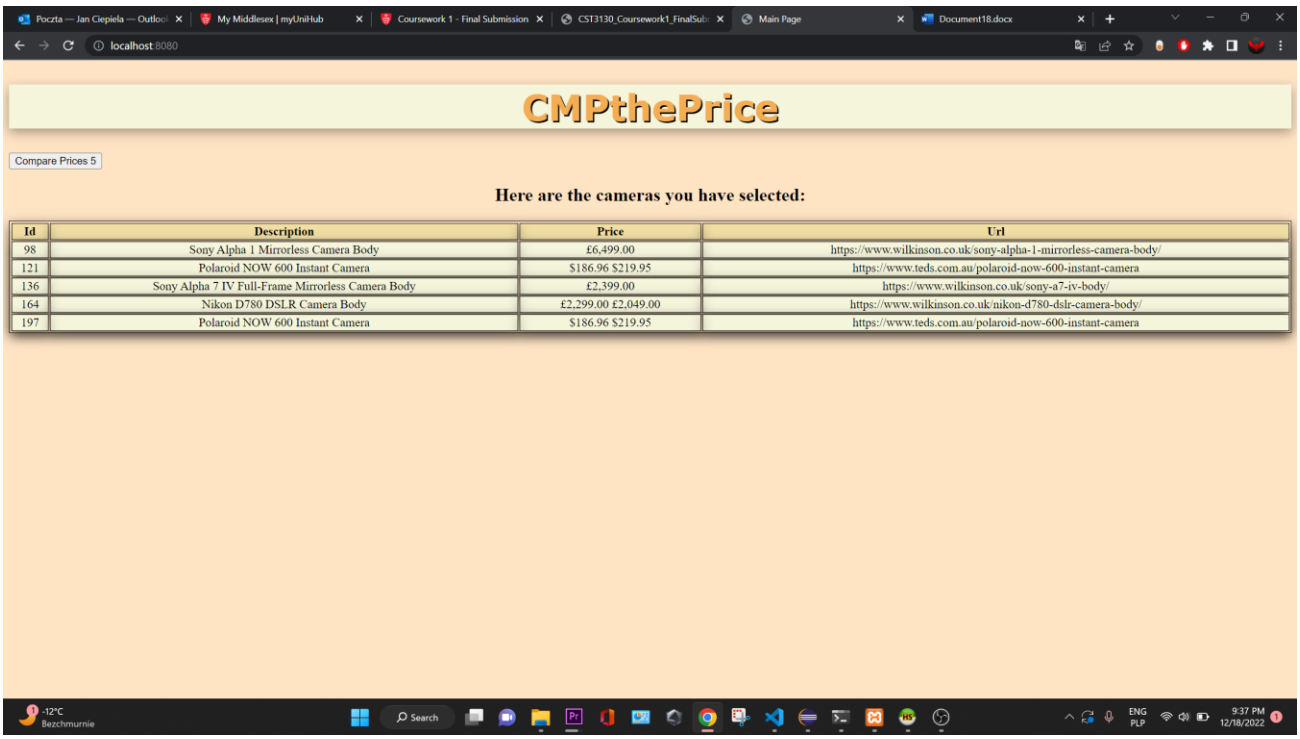
Id	Description	Price	Url	Compare
95	Joby Handypod for Cameras	$49.95 $42.46	https://www.teds.com.au/joby-handypod-for-cameras	Select
96	The legendary LUMIX GH5 experience, now with livestreaming! The LUMIX GH5M2 draws from the GH5, an iconic camera for pro...	$2,297.99	https://ritzcamera.com/products/panasonic-gh5m2-camera-w-12-60mm-lens?_pos=1&_sid=03efec044&_ss=r	Select
97	Nikon D5100 16.2 Megapixel Digital Camera Body	Our Price: $399.00	https://www.42photo.com/Product/nikon-d5100-16-2-megapixel-digital-camera-body/105966	Select
98	Sony Alpha 1 Mirrorless Camera Body	£6,499.00	https://www.wilkinson.co.uk/sony-alpha-1-mirrorless-camera-body/	Select
99	Sony Alpha 7C Body - Black	£1,399.00	https://www.wilkinson.co.uk/sony-alpha-7c-body-black/	Select
100	Panasonic Lumix DMC-G7 16.0 Megapixel Digital Camera W/ 14-140mm Lens (Black)	Our Price: $929.00	https://www.42photo.com/Product/panasonic-lumix-dmc-g7-16-0-megapixel-digital-camera-w-14-140mm-lens-black-/116779	Select
101	Rode VideoMicro > For Compact System Cameras	$79.95	https://www.teds.com.au/rode-videomicro-for-compact-system-cameras	Select
102	Highlights Entry-Level Soft-Side Case Internal accessory storage pocket Carabiner and hand/belt loop This NXE Entry-Level Soft-Side Case for point and...	$3.95	https://ritzcamera.com/products/nxe-kauai-neoprene-camera-case-black-small?_pos=2&_sid=03efec044&_ss=r	Select
103	Sony Alpha 7C Body - Silver	£1,399.00	https://www.wilkinson.co.uk/sony-alpha-7c-body-silver/	Select
104	Nikon D610 24.3 Megapixel Digital Camera Body	Our Price: $799.00	https://www.42photo.com/Product/nikon-d610-24-3-megapixel-digital-camera-body/108938	Select
105	Fujifilm ACL-XP Wide Lens for XP Cameras	$39.95	https://www.teds.com.au/fujifilm-acl-xp-wide-lens-for-xp-cameras	Select
106	Sony Alpha 7 II Body	£899.00 £799.00	https://www.wilkinson.co.uk/sony-a7-ii-body/	Select
107	Canon EOS Rebel Digital T5i 18.0 Megapixel Digital Camera (Body)	Our Price: $449.00	https://www.42photo.com/Product/canon-eos-rebel-digital-t5i-18-0-megapixel-digital-camera-body-/107870	Select
108	Lomo LI100 Instant Camera	$110.46 $129.95	https://www.teds.com.au/lomo-li100-instant-camera	Select

User can search the products, sort them by id, description, price and url as well as increasingly or descresingly:

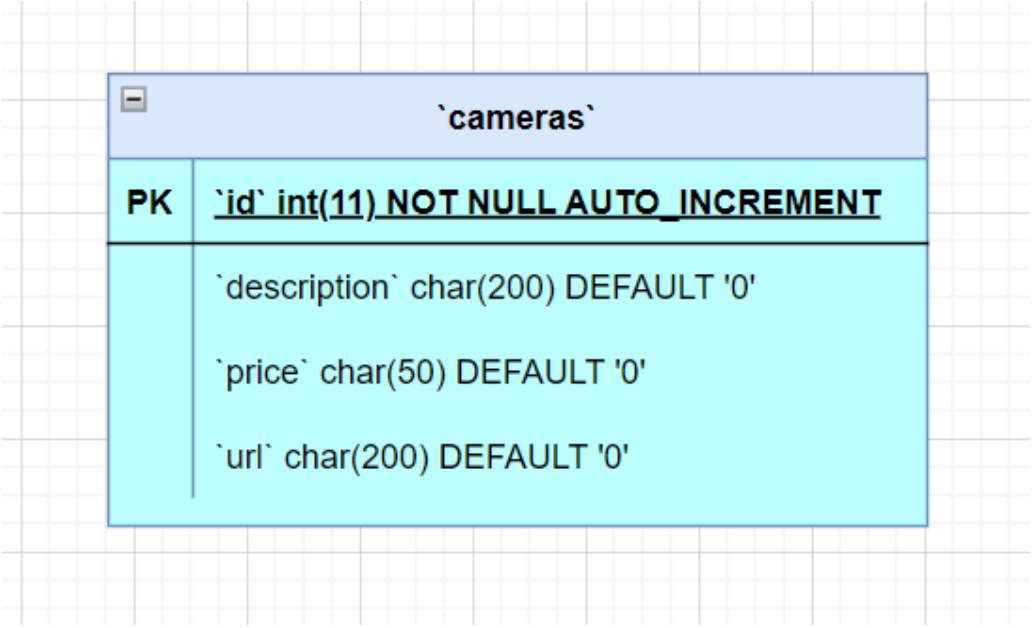


Id	Description	Price	Url	Compare
98	Sony Alpha 1 Mirrorless Camera Body	£6,499.00	https://www.wilkinson.co.uk/sony-alpha-1-mirrorless-camera-body/	Select
99	Sony Alpha 7C Body - Black	£1,399.00	https://www.wilkinson.co.uk/sony-alpha-7c-body-black/	Select
103	Sony Alpha 7C Body - Silver	£1,399.00	https://www.wilkinson.co.uk/sony-alpha-7c-body-silver/	Select
106	Sony Alpha 7 II Body	£899.00 £799.00	https://www.wilkinson.co.uk/sony-a7-ii-body/	Select
109	Sony Alpha 7 III & FE 28-70mm F3.5-5.6 OSS	£1,599.00	https://www.wilkinson.co.uk/sony-a7-iii-fe-28-70mm-f3-5-5-6-oss/	Select
112	Sony Alpha 7C & FE 28-60mm f/4-5.6 Lens - Silver	£1,669.00	https://www.wilkinson.co.uk/sony-alpha-7c-fe-28-60mm-f-4-5-6-lens-silver/	Select
122	Sony Cinema Line FX30 Camera Body	£2,099.00	https://www.wilkinson.co.uk/sony-cinema-line-fx30-camera-body/	Select
124	Sony Cinema Line FX30 Camera With XLR Handle	£2,499.00	https://www.wilkinson.co.uk/sony-cinema-line-fx30-camera-with-xlr-handle/	Select
128	Sony Alpha 7C Black & FE 28-60mm F4 5.6 ProMaster Kit	£1,839.99	https://www.wilkinson.co.uk/sony-a7c-black-fe-28-60mm-f4-5-6-promaster-kit/	Select
129	Sony Alpha 7C Black/Silver & FE 28-60mm F4 5.6 ProMaster Kit	£1,839.00	https://www.wilkinson.co.uk/sony-a7c-black-silver-fe-28-60mm-f4-5-6-promaster-kit/	Select
136	Sony Alpha 7 IV Full-Frame Mirrorless Camera Body	£2,399.00	https://www.wilkinson.co.uk/sony-a7-iv-body/	Select
162	Sony Alpha 7 III & FE 28-70m F3.5-5.6 OSS ProMaster Kit	£1,669.00	https://www.wilkinson.co.uk/sony-a7-iii-fe-28-70m-f3-5-5-6-oss-promaster-kit/	Select
170	Sony Alpha 7C & FE 28-60mm f/4-5.6 Lens - Black	£1,669.00	https://www.wilkinson.co.uk/sony-alpha-7c-fe-28-60mm-f-4-5-6-lens-black/	Select
179	Sony Alpha 7 IV & FE 28-70mm F3.5-5.6 Lens	£2,599.00	https://www.wilkinson.co.uk/sony-a7-iv-kit/	Select
180	Sony Alpha 7 IV & FE 24-105mm F4 G OSS	£3,199.00	https://www.wilkinson.co.uk/sony-alpha-7-iv-fe-24-105mm-f4-g-oss/	Select
200	Sony Alpha 1 Mirrorless Camera Body	£6,499.00	https://www.wilkinson.co.uk/sony-alpha-1-mirrorless-camera-body/	Select
201	Sony Alpha 7C Body - Black	£1,399.00	https://www.wilkinson.co.uk/sony-alpha-7c-body-black/	Select
202	Sony Alpha 7C Body - Silver	£1,399.00	https://www.wilkinson.co.uk/sony-alpha-7c-body-silver/	Select

Upon selection of at least two products, user can then compare them on the comparison page, to which one may switch using the button in the top left corner. The view from the comparison page:



Database Diagram



Unit testing

To verify if the project works correctly there have been designed five JUnit tests to examine the Java program, specifically the scrapers:

```
[INFO] Tests run: 5, Failures: 0, Errors: 0, Skipped: 0, Time elapsed: 39.068 s - in janciepiela.ScraperTest
[INFO]
[INFO] Results:
[INFO]
[INFO] Tests run: 5, Failures: 0, Errors: 0, Skipped: 0
[INFO]
[INFO] ------------------------------------------------------------------------
[INFO] BUILD SUCCESS
[INFO] ------------------------------------------------------------------------
[INFO] Total time:  42.318 s
[INFO] Finished at: 2022-12-18T18:41:02+01:00
[INFO] ------------------------------------------------------------------------
PS C:\Users\johnn\Desktop\java_website\maven_project>
```

As well as two Javascript tests using Mocha and Chai to examine the connection to the database and the default API route:

```
PS D:\www> npm test

> www@1.0.0 test
> mocha --timeout 10000

App listening on port 8080


  First Test Collection
    ✓ should test default API route
    ✓ Should connect to database without an error


  2 passing (44ms)
```