

Parmanand Menghwar

A Product Pricing Comparison Model and Data Visualization for Online Retailers

Helsinki Metropolia University of Applied Sciences

Master of Engineering

Information Technology

Master's Thesis

02 June 2019

PREFACE

When I started my first job as an eCommerce Consultant in 2016, I got a chance to work on different ecommerce projects and got an opportunity to create a Magneto based online store from scratch.

Thanks to ArvoPartners Oy for allowing me to use six working hours on my studies every week for one semester. As a result, I completed my two major courses “Big Data” and “Cloud Computing”, and I was able to help two of our clients who were struggling in deciding product prices as they had lots of data but didn’t exactly know what to do with it.

Thanks to my parents, it would have been impossible without their support and prayers.

Thanks to my wife for taking care of my son when I was gone whole day on every Tuesday from 8 am to 9 pm and facing all the side effects of this journey.

The last whole year was tough for me as I switched my job, started writing this thesis at the same time and having a family with a kid that taught me patience, time management and hair loss.

Thanks to Sir Ville for trusting in me and supporting through my master studies and in thesis work.

Thanks to my friend Dr Muhammad Zeeshan Asghar for being so supportive from last seven years and mentoring in whole thesis writing, I still remember your advice when I was applying for master studies and looking for fulltime job at the same time. Thanks to my colleague and friend Tomi Nares for supporting me in this writing process.

Helsinki, 02 June 2019

Parmanand Menghwar

Author(s) Title	Parmanand Menghwar A Product Pricing Comparison Model and Data Visualization for Online Retailers
Number of Pages Date	37 pages 02 June 2019
Degree	Master of Engineering
Degree Programme	Information Technology
Instructor(s)	Ville Jääskeläinen, Head of master's Program in IT
<p>As the internet continues to progress and online shopping is growing rapidly, more and more people are using internet to buy products and services. In online shopping consumers and retailers are always looking for a competitive price for a certain product. A consumer can search the best prices using online search tools or apps, but it is challenging for a retailer to manage regularly a good price level for a product. Retailers need to consider many factors such as competition, prices of thousands of products, ongoing campaigns and steadily changing market situations.</p> <p>The purpose of this study is to help retailers to decide a suitable product price by comparing it to an average market price. Normally, retailers decide the product price based on their production and logistics cost. This study explains the design and implementation of a product comparison tool for retailers with regards/reference to their top five competitors. Based on the preliminary feedback from multiple retailers this study can help them to adjust their product prices, get a better conversion rate, increase sales and revenue.</p>	
Keywords	Prize Analysis, Prize Optimization, Web Scraping, Online Store, Data Visualization, eCommerce.

Table of Contents

Preface	
Abstract	
List of Figures	
List of Abbreviations	
1 Introduction	1
2 Current State	5
3 Theoretical and Technical Framework	6
3.1 Online shopping and online store	6
3.2 Magento 9.1	6
3.3 Php-MySQL, CSS and HTML	8
3.4 Sonassi	8
3.5 Web Scrapping	9
3.6 Open Web Scraper	10
3.7 Commercial Tools for Web Scrapping	10
3.7.1 Webhooks	10
3.7.2 Scheduling	11
3.8 MySQL database	11
3.9 Comparison shopping engines	11
3.9.1 Vertaa	12
3.9.2 Hintaseuranta	13
3.9.3 Shopzilla	15
3.10 Pricing strategies	16
3.11 Normalize Data	17
3.12 Data Visualization	17
3.12.1 Power-Bi	17
3.12.2 Data Studio by Google	18
4 Implementation	20
4.1 Web scraper setup	20
4.2 Import CSV data into MySQL database	22
4.3 Searching and filtering of data	23
4.4 Visualization of data	24
4.5 Creating Dashboard in Data Studio	25
5 Discussions and Conclusions	29

References

List of Figures

- Figure 1. The system architecture of this study.
- Figure 2. File structure of Magento CE 1.9.3.
- Figure 3. Web Scraped data example.
- Figure 4. Screenshot of Open Web Scraper.
- Figure 5. Screenshot of MySQL table with columns and rows.
- Figure 6. Vertaa search result page.
- Figure 7. Vertaa product detail page.
- Figure 8. Hintaseuranta search result page.
- Figure 9. Hintaseuranta product detail page.
- Figure 10. Shopzilla search result page.
- Figure 11. Power BI Dashboards for Desktop, Mobile & Tablet.
- Figure 12. Google Data Studio Search Console.
- Figure 13. Web Scraper sitemap.
- Figure 14. Data preview before scraping is completed.
- Figure 15. Web scraped data extracted/download file in CSV format.
- Figure 16. All imported CSV files in a single table.
- Figure 17. Products filtered according to the product title.
- Figure 18. Products filtered according to the product title.
- Figure 19. The visual form of the data as a chart.
- Figure 20. Supported data source options for creating a dashboard.
- Figure 21. Columns with names and types which user need to update.
- Figure 22. Available chart options.
- Figure 23. Options set for dashboard main chart.

List of Abbreviations

AJAX	Asynchronous JavaScript & XML
API	Application Programming Interface
CRUD	Create, read, update and delete
CSS	Cascading Style Sheet
CSS3	Cascading Style Sheet level 3
CSV	Comma Separated Values
DSS	Data Security Standards
GUI	Graphical User Interface
HTML5	Hypertext Markup Language version 5
HTTP	Hypertext Transfer Protocol
JSON	JavaScript Object Notation
Magento CE	Magento Community Version
Magento EE	Magento Enterprise Version
MAMP	Macintosh, Apache, MySQL & PHP
MySQL	My S-Q-L Open Source Relational Database
NDA	Non-Disclosure Agreement
PHP	Hypertext Pre-Processor (Initially stood as Personal Home Page)
REST	Representational State Transfer
URL	Universal Resource Locator
XML	Extensible Markup Language

1 Introduction

As the internet continues to progress it has become an essential part of people's everyday life. It has also become an important channel for consumers to buy regularly everything from shoes to medicine online. An online shopping is a form of e-commerce that enables consumers and business owners to combine such commercial activities as buying and selling products and services online, thus enhancing consumers buying intention. All the products in online stores are described by their name, description with photos and multimedia files, and price.

Many factors influence consumers to buy a certain product online. Saad Akbar and Paul James has identified 9 factors that can influence a consumer to buy a product and is very helpful for the retailers. Among the 9 factors, the strongest predictors from highest to lowest were: Price, Refund, Convenience, Auction websites, Security, Brand, Search engines, Promotion and Online shopping malls.

According to PWC research these are the main reasons why customers visited an e-commerce website:

- 61% to compare pricing
- 23% to participate in promotions
- 41% to look for coupons

A typical consumer follows multiple ways to find the best price for a product. This means that consumers value highly the ability to research prices of the products they are interested in. There are many product comparison shopping engines and apps available that assist the growing need of consumers.

On the other hand, a product pricing for a seller on e-commerce platform is highly challenging. Sellers typically consider several factors such as:

- competition,
- prices for thousands of products, and
- steadily changing market situations.

A pricing strategy refers to the method a company uses to price their products or services. A successful pricing strategy can significantly increase sales, result in better cooperation with suppliers, and boost revenue. If one does not base the pricing on any strategy, chances are that one might set product prices too high or too low. One can lose customers if the prices are too high and if the prices are set too low one will get small margins.

Generally, pricing strategies include the following five strategies [1]:

- Cost-Plus Pricing — simply calculating the costs and adding a markup.
- Competitive Pricing — setting a price based on what the competition charges.
- Value-Based Pricing — setting a price based on how much the customer believes the product is worth.
- Price Skimming — setting a high price and lowering it as the market evolves.
- Penetration Pricing — setting a low price to enter a competitive market and raising it later.

This study is focusing on a new tool that can be used for a competitive pricing strategy because it helps the retailer to set the price in accordance with what the competition is charging and retailers are also able to attract more customers due to their analytical approach for online shopping.

In order to successfully apply and get the benefits of this strategy, the retailers should have a solid data collection system to get the fresh product prices from the competition.

Furthermore, to make the competitive analysis efficient, the retailers need to automate it because comparing a big number of product prices with a competitor is a cumbersome and time-consuming process. Finally, a presentation of collected data in a pictorial or graphical form enables the retailers to see the results of the analysis visually and helps them to decide the product prices.

Currently retailers are using e-commerce technologies such as a competitor price tracking software to make the pricing more competitive. But there is a lack of simple and customizable tool that can help to decide a right product pricing.

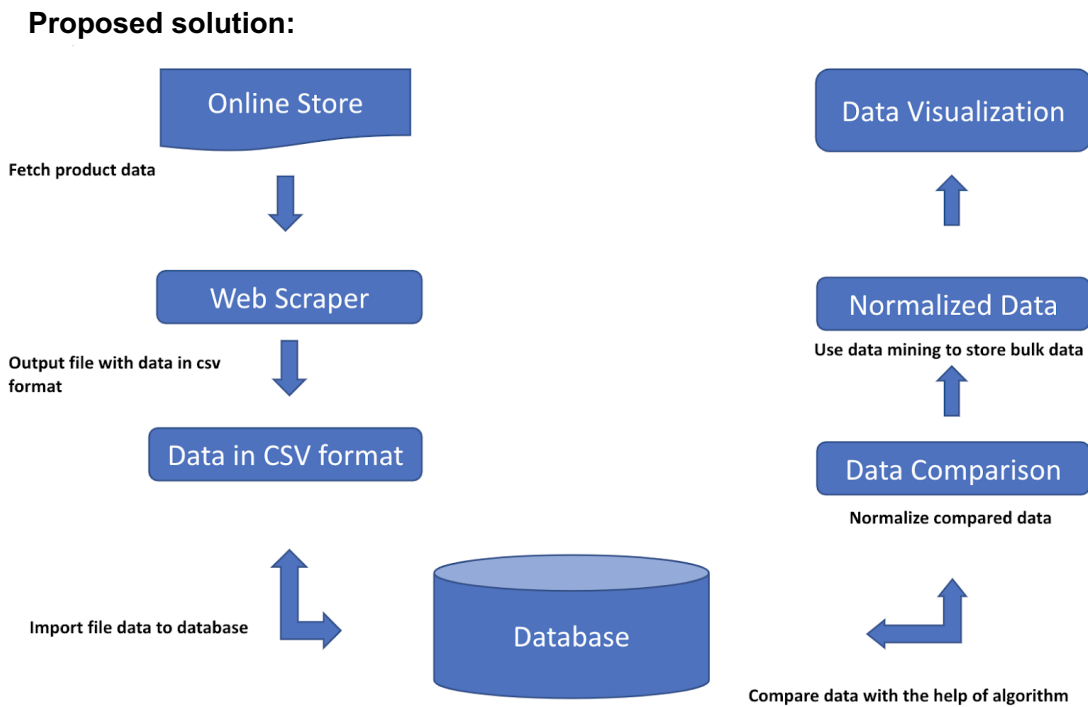


Figure 1. The system architecture of this study.

This study explains the design and implementation of a product comparison tool for the retailers with reference to their top 5 competitors. Moreover, this tool helps the retailer to decide the product price compared to average market price. Normally, retailers decide the product price based on their production and logistics cost.

This tool chooses top 5 competitors product prices and apply web scraping to retrieve the product data such as a product title, a category and a price. There are many ways to store the retrieved data. In this study open source browser-based extension was used to retrieve the data and stored in an excel sheet. Then the data was imported into a MySQL database to normalize it. In normalization process, only relevant or needed data is stored in separate table and remove redundant data from that table.

The data was normalized by matching the product titles and categories and then they were stored into a database table. This stored data can be used to compare prices of the available products. By applying a comparison algorithm, the product prices can be decided.

This algorithm will take retailer's product titles and matches them with competitor's product titles and puts all 5 results in a row which can generate the dashboard. In order to understand this data clearly, a data visualization process/tool can be used.

With the help of Data Studio by Google or Power-BI by Microsoft, the retailer can create personalized dashboards with a unique, 360-degree view of their business. The dashboard will help the retailers to decide the prices of each product based on an average market price and a certain competitor product pricing.

The thesis has been divided into five sections. The first section introduces the problem, scope or big picture and proposed solution of the study. The second chapter explains the current state of the system, issues found out in the current system and discussed the solution of each problem. In the third chapter, tools and technologies were explained which are used in the old system and in the current solution. The fourth chapter described the implementation of the proposed solution. The proposed solution consists of various other tools and required various steps to be done properly. Finally, the fifth chapter discussed the overall thesis and provides conclusions.

2 Current State

This study is based on a real time problem faced by an e-commerce store with more than ten thousand products in each store. The problem was, how to decide the price of a product? Currently, product price was decided based on the following formula.

$$\text{Retail Price} = [(\text{Cost of item}) \div (100 - \text{markup percentage})] \times 100$$

For example, a user wants to price a product that costs user \$15 at a 45% markup instead of the usual 50 [2]. Here's how you would calculate your retail price:

$$\text{Retail Price} = [(15) \div (100 - 45)] \times 100$$

$$\text{Retail Price} = [(15 \div 55)] \times 100 = \$27$$

While this is a relatively simply markup formula, this pricing strategy doesn't work for every product in every retail business.

This formula created two problem for the retailers. The first problem was that some products were sold a lot, and some were sold less.

The reason of the above problem that retailers are not familiar with the product prices of their competitors. They can check product prices using available product comparison tool and apps mostly build for end user or customers. But checking prices of thousands of products online is a time consuming and cumbersome process which needs to be done multiple times depends on season or market, that leads to the second problem. The second problem was, how to easily find the product price of competitors and compare them with own products?

The solution for this problem is to check each product price using a customized tool and compare them with own products. Moreover, a visual representation of these compared product can help the retailer in deciding product prices, increase sales and profits. The online store was hosted on Sonassi, Magento was used as an ecommerce platform which is based on PHP, MySQL, CSS and HTML.

3 Theoretical and Technical Framework

This chapter explain the tools and technical Framework used in the existing system. It also explains the details of tools and technologies used in the proposed solution. The existing retailer online store was developed using Magento framework which is based on PHP-MySQL, html and CSS. Sonassi was used for hosting and storage purpose.

The proposed solution is designed and developed using various tools. Firstly, web scraper was used for scrapping the online stores data. MySQL database was used to save the data. Finally, Google Studio was used for visualizing the data and for dashboard creation etc.

3.1 Online shopping and online store

Online shopping is a convenient and time saving way for consumers to directly buy goods and services from a retailer over the internet. Consumers do not need to travel or wait in lines and online stores are open all the time. They are accessible from anywhere and anytime. Modern online stores provide rich information to consumer about a product such as price, availability and quantity. Moreover, consumer can use various online tools and apps to compare and make purchase decisions among various products and service [3].

3.2 Magento 9.1

Magento known as open source ecommerce platform which helps creating online stores and can be managed by merchants, it comes with free Community version (CE) and paid Enterprise version (Strong support, many built-in feature). More than 50 payment gateways, many plugin and extensions make Magento development easier and its strong developers community contribution [4]. Figure 2 shows the folder structure of Magento community version 1.9.3.

In this study own store data such as product prices, product title and product purchase total count which is taken from Magento database, Google analytics and physical store system is used to compared with three or four competitors store data.

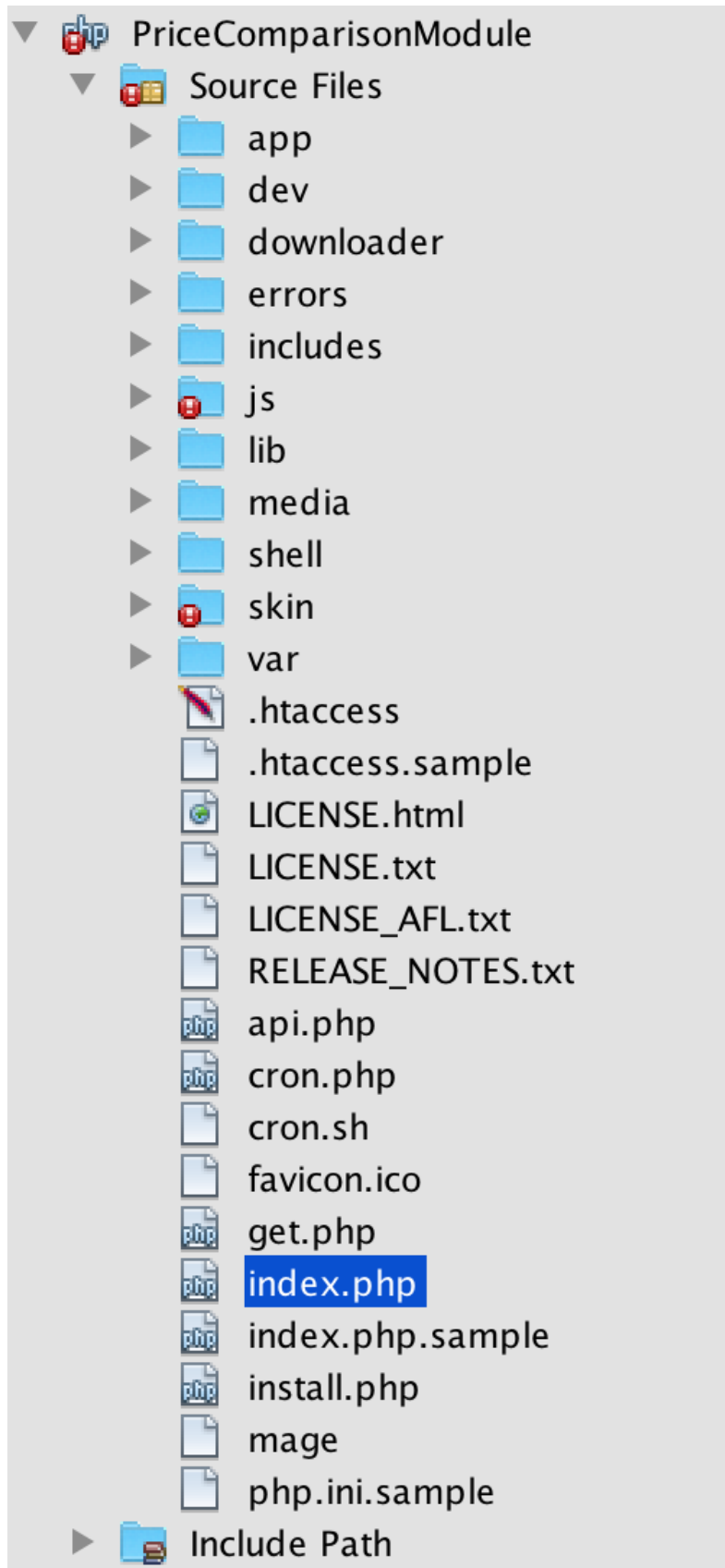


Figure 2. File structure of Magento CE 1.9.3.

3.3 Php-MySQL, CSS and HTML

PHP is a server-side scripting language that was introduced for developing dynamic web applications. PHP code is embedded into HTML source file with PHP tags and interpreted by web server.

“The original PHP release was created by Rasmus Lerdorf in June 1995, to make various common web programming tasks easier and less repetitive. The name originally stood for “Personal Home Page,” but has since become a recursive acronym, standing for “PHP: Hypertext Preprocessor.” The goal of that release was to minimize the amount of code required to achieve results, which led to PHP being HTML-centric—that is, PHP code was embedded inside HTML [5].

Its mixture of different languages e.g. Perl, C and Java programming, due to its simple structure, easy connectivity with MySQL database and ability to do all kinds of things became famous in the developer’s community.

“MySQL, the most popular Open Source SQL database management system, is developed, distributed, and supported by Oracle Corporation.

MySQL is a database management system. A database is a structured collection of data. It may be anything from a simple shopping list to a picture gallery or the vast amounts of information in a corporate network. To add, access, and process data stored in a computer database, you need a database management system such as MySQL Server. Since computers are very good at handling large amounts of data, database management systems play a central role in computing, as standalone utilities, or as parts of other applications [5].

Cascading Style Sheets (CSS) is a style sheet language used for describing the presentation of a document written in a markup language like HTML. [5] CSS is a cornerstone technology of the World Wide Web, alongside HTML and JavaScript.

3.4 Sonassi

Sonassi provides specially modified hosting for Magento with reference to MageStack, it provides all components, documentation to help Magento Developers and retailers with full ROOT access to server, PCI compliance.

Sonassi offers free migration, no setup fee, premium support if needed (Charged in minutes used), really good in security point of view, Sonassi internal scan system once a day give administrator or developer full scan report which includes check for all Magento and third-party extension security check, difference report for JavaScript, PHP and CSS, difference report play vital role in debugging when some change happens, sometime developers set some files in .gitignore (means changes in these file won't show in Version Control) so those changes can only be tracked by MageScan file difference logs. Automatic scheduled backup, cache clearance facility, system health check and sending warning or updates make it different from normal hosting providers [7].

3.5 Web Scrapping

Web Scrapping known as “web data extraction”, “data harvesting”, “web crawling” or “web spidering” to extract data from websites with http (hypertext transfer protocol). The first step is extract data from the online websites [8].

With Web Scrapping it's possible to get all visible data on a html page in excel format or CSV format, then it can be checked and imported to database or Big Query for next phase of analysis.

Figure 3 shows Web Scraped data from competitor 1 and competitor 2 with mean price, price difference (pre-calculated price difference which can be used later for dashboard), availability and brand name etc.

Product Name	Our Store	Competitor 1	Competitor 2	Mean Price	Price Difference	Availability	Product Brand
Estelle & Thild All-in-one Tinted Moisturizer Light	299	299 NA	299	299	0	0	Estelle
Estelle & Thild All-in-one Tinted Moisturizer Medium	299	299 NA	299	299	0	0	Estelle
Estelle & Thild BioCalm Anti-Redness Rescue Serum	299	299 NA	299	299	0	0	Estelle
Estelle & Thild BioCalm Extra Nourishing Night Cream	279	279 NA	279	279	0	0	Estelle
Estelle & Thild BioCalm Optimal Comfort Rescue Oil	329	299 NA	299	299	30	0	Estelle
Estelle & Thild BioCalm Soothing Eye Balm	229	229 NA	229	229	0	0	Estelle
Estelle & Thild BioCalm Soothing Moisture Day Cream	259	259 NA	259	259	0	0	Estelle
Estelle & Thild BioCleanse 3-in-1 Cleansing Foam	199	199 NA	199	199	0	0	Estelle
Estelle & Thild BioCleanse Deep Cleansing Detox Mask	299	299 NA	299	299	0	0	Estelle
Estelle & Thild BioCleanse Eye Make Up Remover	189	189 NA	189	189	0	0	Estelle
Estelle & Thild BioCleanse Fragrance Free Micellar Cleansing Water	199	199 NA	199	199	0	0	Estelle
Estelle & Thild BioCleanse Fragrance Free Cleansing Gel	189	189 NA	189	189	0	0	Estelle
Estelle & Thild BioCleanse Fragrance Free Cleansing Milk	199 NA	NA	NA	NA	NA	0	Estelle
Estelle & Thild BioCleanse Multi-Action Cleansing Gel	189	189 NA	189	189	0	0	Estelle
Estelle & Thild BioCleanse Multi-Action Facial Toner	189	189 NA	189	189	0	0	Estelle
Estelle & Thild BioCleanse Radiance Micro Polish	249	245 NA	245	245	4	0	Estelle
Estelle & Thild BioCleanse Silky Soft Cleansing Milk	199	199 NA	199	199	0	0	Estelle
Estelle & Thild BioDefense Antioxidant Eye Cream	269	269 NA	269	269	0	0	Estelle
Estelle & Thild BioDefense Antioxidant Face Cream	399	379 NA	379	379	20	0	Estelle
Estelle & Thild BioDefense Instant Recovery Night Cream	429	399 NA	399	399	30	0	Estelle
Estelle & Thild BioDefense Multi-Action Youth Serum	425	425 NA	425	425	0	0	Estelle
Estelle & Thild BioDefense Multi-Nutrient Youth Oil	399	379 NA	379	379	20	0	Estelle
Estelle & Thild BioHydrate Intense Moisture Night Cream	279	279 NA	279	279	0	0	Estelle
Estelle & Thild BioHydrate Refreshing Eye Gel	229	229 NA	229	229	0	0	Estelle
Estelle & Thild BioHydrate Thirst Relief Vitamin Serum	299	299 NA	299	299	0	0	Estelle
Estelle & Thild BioHydrate Total Moisture Day Lotion	259	259 NA	259	259	0	0	Estelle

Figure 3. Web Scraped data example.


3.6 Open Web Scraper

This study has used the Open Web Scraper tool. These tools are now part of famous browsers and available as an extension such as Google Chrome Web Scraper Browser extension. Anybody can install the Web Scraper extension from the chrome web store to make it an easy to use data scraping tool and its free to use. The best part is, you can stay in the comfort zone of your browser while the scraping happens. This tool does not demand any technical skills, and anybody can use it for quick data scraping [9].

Once browser extension is installed user need to go to Chrome Browser menu -> View -> Developers -> Inspect Element -> Web Scraper.

Test Sites

- Home
- Computers >
- Phones >



Asus VivoBook X441NA-GA190 \$295.99

Asus VivoBook X441NA-GA190 Chocolate Black, 14", Celeron N3450, 4GB, 128GB SSD, Endless OS, ENG kbd

14 reviews ★★ ★

Elements
Console
Sources
Network
Performance
Memory
Application
Security
Audits
Web Scraper

Sitemaps
Sitemap demo-scraping-stie
Create new sitemap

._root
category-links
subcategory-link
product-link

ID	Selector	type	Multiple	Parent selectors	Actions
name	h4:nth-of-type(2)	SelectorText	no	product-link	Element preview Data preview Edit Delete
price	h4.pull-right	SelectorText	no	product-link	Element preview Data preview Edit Delete

Add new selector

Figure 4. Screenshot of Open Web Scraper.

3.7 Commercial Tools for Web Scrapping

In Web Scrapping there are many challenges to do repetitive task or it depends how often you scrape data from a website, you IP can get blocked etc. to avoid these kinds of situation and some commercial solutions solved some problems and gave freedom in form of Webhooks and Scheduling [10].

3.7.1 Webhooks

In Web scraping web hooks are used to fetch selected data, it can be one element or multiple object and can be configured, when to run and what to fetch.

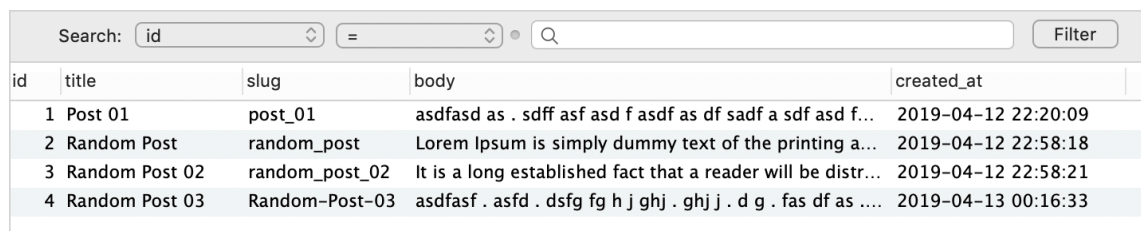
Choosing right webhook with right fetching frequency can make big difference, plus it also depends on how much data you are planning to fetch.

3.7.2 Scheduling

In Web scraping scheduling used to schedule Webhook, you can define at what time and how frequently a specific Webhook event take place. Commonly its used to fetch full follow.

3.8 MySQL database

This study has used MySQL an open source relational database management system. The web scraping tool retrieved the data in a CSV format which needs to be imported to a database system. A database system like MySQL provides many features such as filtering the data, normalize the data and searching the required data. Furthermore, it can be integrated to any data visualization software such as Data Studio by Google as a data-source [6].



id	title	slug	body	created_at
1	Post 01	post_01	asdfasd as . sdff asf asd f asdf as df sadf a sdf asd f...	2019-04-12 22:20:09
2	Random Post	random_post	Lorem Ipsum is simply dummy text of the printing a...	2019-04-12 22:58:18
3	Random Post 02	random_post_02	It is a long established fact that a reader will be distr...	2019-04-12 22:58:21
4	Random Post 03	Random-Post-03	asdfasf . asfd . dsfg fg h j ghj . ghj j . d g . fas df as	2019-04-13 00:16:33

Figure 5. Screenshot of MySQL table with columns and rows.

3.9 Comparison shopping engines

Comparison shopping engines helps the consumer to find the best price of a product. These engines respond the results to the consumer's search query into single results page. This page contains collective product information such as pricing and retailer information. These engines also give ecommerce retailers a glance to see what types of prices are attracting their customers into a purchase, attract new customer and increase sales. They will also see which competitors are being shown to your customers. Some of the popular pricing comparison websites are listed below [11] [12].

3.9.1 Vertaa

Vertaa.fi is Finland's large comparison site, where consumers can compare products based on the price, delivery time and reviews. The site provides consumer to search products by category and it also lists the products by category.

A user can search a product directly into the search field. Another option is to browse the desired categories and select the desired product. To start the search the user needs to enter the required product and press enter. After that the search engine will start searching the products and shows all the relevant products and its price as shown in Figure 6 [13].

The screenshot shows the search results for 'samsung Puhelimet (91)'. The page includes a sidebar with filters for 'Quick filters', 'Malli', and 'Merkki'. The main content area displays a grid of 9 smartphone listings, each with a product image, name, specifications, and price.

Product Name	Price (Alk.)	Link
Samsung Galaxy S9 64GB, Musta	500,90	Katsota halvin hinta
Samsung Galaxy A50 128GB, Musta	293,90	Katsota halvin hinta
Samsung Galaxy A40 64GB, Musta	237,24	Katsota halvin hinta
Samsung Galaxy A7 (2018) 64GB Musta	202,90	Katsota halvin hinta
Samsung Galaxy S10 128GB, Prism White	700,90	Katsota halvin hinta
Samsung Galaxy S10e 128GB, Prism Black	637,00	Katsota halvin hinta
Samsung Galaxy A70 128GB, Musta	403,35	Katsota halvin hinta
Samsung Galaxy J6+ 32GB Kulta	138,90	Katsota halvin hinta
Samsung Galaxy A70 128GB, Valkoinen	419,00	Katsota halvin hinta

Figure 6. Vertaa search result page.

VERTAA.FI
...ennen kuin ostat!

Etsi tuotteita... Kategoriat

Vertaa.fi / Puhelimet ja internet / Laitteet / Puhelimet / Samsung / Galaxy S9 64GB, Musta

Samsung Galaxy S9 64GB, Musta
Viimeisin päivitys 02 kesäkuuta, 2019 klo 17:32
[Kirjoita arvostelu](#) 👍 paras ostos [Hinnat Samsung Galaxy S9 64GB, Musta](#)

Järjestä perusteella Halvin hinta

Myyjä	Toimitusaika	Hinta	Toiminto
GIGANTTI Kauppaan 1 arvostelu	Toimitusaika: 3 pv	504,90 Sis. toimituskulut	KATSO TUOTETTA
POWER Kauppaan 10 arvostelua	Toimitusaika: ?	504,90 Sis. toimituskulut	KATSO TUOTETTA
RDE.FI Kauppaan Ei arvostelua	Toimitusaika: 1 vk	506,86 Sis. toimituskulut	KATSO TUOTETTA
KNAITEK Kauppaan Ei arvostelua <i>alennus</i>	Toimitusaika: 5 pv <i>21% alennus</i>	657,00 522,00 Sis. toimituskulut	KATSO TUOTETTA
PROSHOP Kauppaan 120 arvostelua	Toimitusaika: 3 pv	579,00 Sis. toimituskulut	KATSO TUOTETTA
HUVILAITE Kauppaan	Toimitusaika: 5 pv	583,90	KATSO TUOTETTA

Figure 7. Vertaa product detail page.

When user selects the required product and click more. On the next page the user can see more detail of the product. The detail consists of its price, availability in other stores, delivery time as shown in Figure 7.

3.9.2 Hintaseuranta

This website fetches the prices for one product group and then provides the user with price comparisons from different websites. It also contains different search and filtering options. It can see where the desired product is sold at the lowest price. You can also see recent product price trends on the site.

Hintaseuranta.fi
Löydä nopeimmat toimitusajat

samsung **HAE**

Viihde ja elektroniikka Tietotekniikka Koti ja perhe Työ ja hyvinvointi Liikkuminen ja ulkoilu Matkailu Suosituimmat

Haku > Haku > samsung

Haku: "samsung"

Hinta (€)
0 — 40953

Osastot

- Tulostinten mustekasetit (575)
- Puhelinkotelot ja suojakalvot (358)
- Akut puhelimille (184)
- Televisiot (180)
- Näytöt (173)
- Muut urheiluvälineet (105)
- Jääkaappipakastimet (101)
- Näytönsuojat, suojakuoret ja -laukut (79)
- Muut puhelintarvikkeet (70)
- Tulostimet (68)
- Kirjat (67)
- Sisäiset kovalevyt (64)

Näytön tarkkuus

- 4K (167)
- Full HD (124)
- HD ready (7)

Valmistaja

- Samsung (2084)
- HP (222)
- Insmat (182)
- Celly (169)
- MTP Products (163)

30 / 2655 tuotetta Näytä: Järjestä: Suosituimmat





	UUSIN Samsung Galaxy S9 64GB, puhelin Alkaen 399,00 €	Katso halvin Katso kaikki
	Samsung Galaxy S10 128GB, puhelin Alkaen 699 €	Katso halvin Katso kaikki
	Samsung Galaxy A7 (2018), puhelin Alkaen 199,00 €	Katso halvin Katso kaikki
	Samsung Galaxy S10e 128GB 6GB RAM, puhelin Alkaen 630 €	Katso halvin Katso kaikki


Figure 8. Hintaseuranta search result page.

This website works on the same mechanism as vertaa.fi but it provides more search and filtering options. User can search the product directly from the search field. Figure 8 shows the results of the product search and shows all the products with the given name and prices in the main right view. In the left view it shows different options for the user to select such as price range, departments, manufacturer, rating of the product etc.

User can select the desired product and its detail. The detail page can be seen in Figure 9. This page contains various information such as the user can see the description about the product, actual price of the product, product price including delivery charges, availability in different stores, price in different stores and estimated delivery time [14].

Samsung Galaxy S9 64GB, puhelin

★★★★★ Jätä arvostelu ensimmäisenä ja osallistut arvontaan.



Halvin hinta
ilman toimituskuluja

399,00 €

Katso tuotetta ▶

Halvin hinta toimitettuna **401,90 €**

▶

[Katso kaikki kuvat \(10\)](#)

Upeilla ääni- ja kuvausominaisuuksilla varustettu Samsung Galaxy S9 on rohkea ja edistyksellinen high-end älypuhelin, jonka vahva työteho ja monipuolinen kamera varmistavat ilahduttavan käyttökokemuksen. Samsung Galaxy S9:n kameran aukko mukautuu olosuhteisiin automaattisesti valaistuksen vaihtuessa ja poimii kuviin hämärässäkin kaikki yksityiskohdat. Super Slow Motion -kuvaustilassa tallennat jopa 960 kuvaa sekunnissa ja luot upeita hidastuskuvia, joten Samsung Galaxy S9 näkee paljon sellaista, mitä paljas silmä ei huomaa. [Näytä lisää ▶](#)

Katso myös muut [Samsung-puhelimet](#).






Kaupat ⁵⁶	Tuotetiedot ⁷⁹	Hintakehitys			
Kauppa	Vaihtoehdot	Toimitusaika	Hinta ▼	Toimitettuna	
 ★★★★★	Samsung Galaxy S9 violetti Samsung Galaxy S9 kulta Samsung Galaxy S9 musta Samsung Galaxy S9 sininen	1-2 päivää	399,00 € HALVIN	401,90 € HALVIN TOIMITETTUNA	Katso tuotetta ▶
 ★★★★★	SAMSUNG GALAXY S9 LILAC PURPLE	3 - 7 päivää	495 €	502 €	Katso tuotetta ▶
 ★★★★★	SAMSUNG GALAXY S9 LILAC PURPLE	3 - 7 päivää	495 €	505 €	Katso tuotetta ▶
 ★★★★★	SAMSUNG GALAXY S9 MIDNIGHT BLACK SAMSUNG GALAXY S9 CORAL BLUE	ei tiedossa	495 €	505 €	Katso tuotetta ▶
 ★★★★★	SAMSUNG GALAXY S9 MIDNIGHT BLACK SAMSUNG GALAXY S9 CORAL BLUE	ei tiedossa	495 €	502 €	Katso tuotetta ▶

Figure 9. Hintaseuranta product detail page.

3.9.3 Shopzilla

This website comes in Google's top ten comparison websites, but it provides less information about a product and its prices. Figure 10 shows the result of a searched product. It shows the list of products in the main view and its prices. Left view of this page shows different filtering options such as price range, brands, different stores details etc. If a user clicks on any product, site will redirect user to that specific site which sells that product [15].

The screenshot shows the Shopzilla search results page for 'samsung galaxy s9'. The page features a search bar at the top with the query 'samsung galaxy s9' and a 'go' button. Below the search bar, there are related search suggestions and a 'See all departments' link. The main content area displays a list of search results, each with a product image, title, description, price, and a 'go to store' button. The results are sorted by 'Best Match' and include a 'Tax & Shipping: enter zip code' field. On the left side, there are navigation menus for 'You're in Cell Phone', 'Accessories', 'Narrow Search Results', 'Price Range', and 'By Brand'.

Product Title	Description	Price	Store
Samsung Galaxy S9 Unlocked Smartphone - 128GB - Coral Blue - US ...	Super Speed Dual Pixel Camera Infinity Display; edge-to-edge immersive screen, enhancing your entertainment experience High-performance for 4K UHD video recording, high resolution pictures, gaming, and music, for use in Smartphones, action ... more	\$689.45	Amazon.com (10,000+)
Samsung Galaxy S9 Unlocked (US Version) 256GB - Midnight Black with ...	Super Speed Dual Pixel Camera Infinity Display; edge-to-edge immersive screen, enhancing your entertainment experience High-performance for 4K UHD video recording, high resolution pictures, gaming, and music, for use in Smartphones, action ... more	\$759.45	Amazon.com (10,000+)
Total Wireless Samsung Galaxy S9 4G LTE Prepaid Smartphone	Capture memories with the 12 MP Camera/8 MP Front-Facing Camera/Video Recorder and bring them to life on the 5.8-inch Quad HD Super AMOLED Screen display. This prepaid smartphone's wide dimension ensures it fits well in your palms for ... more	\$599.99	Amazon.com (10,000+)
Samsung Accessories Samsun Galaxy Note8 Color: Gray Size: Os	Unlocked. Excellent Condition. more	\$600.00	Poshmark, Inc.
Samsung Other Galaxy Note 8 Orchid Color: black Size: Os	Great Condition. Verizon Service. Few Scratches But Barely Noticeable. Wanting Firm Price Or Trade For An Iphone 8 Plus. more	\$790.00	Poshmark, Inc.
Samsung Galaxy S9 Plus - Midnight Black - 64GB with qualifying plan	The Samsung Galaxy S9+ Is Here With The Camera Re-Imagined! Take Professional Quality Photos With The Dual Aperture, Dual 1mp Rear Facing Cameras, And Capture Stunning Pictures Whether You're In Bright Daylight, Moonlight, Or Super Low ... more	\$840.00	T-Mobile (100+)
Samsung Galaxy S9 - Midnight Black - 64GB with qualifying plan	The Samsung Galaxy S9 Is Here With The Camera Re-Imagined! Take Professional Quality Photos With The Dual Aperture Camera, And Capture Stunning Pictures Whether You're In Bright Daylight, Moonlight, Or Super Low Light. Slow Down Reality In ... more	\$720.00	T-Mobile (100+)

Figure 10. Shopzilla search result page.

3.10 Pricing strategies

The greatest challenge for a retailer is to accurately price their product and service. A good pricing strategy helps the retailer to decide the price of their products. As a retailer various questions popped up into the mind while thinking about the pricing strategies for ecommerce. Questions such as How do you really determine what the fair price is? How do you know the real worth of a product you are trying to sell? Five common approaches are listed below [16].

- Cost-plus pricing
- Competitive pricing
- Value-based pricing
- Price skimming
- Penetration pricing

3.11 Normalize Data

The web scraping tools provides the data in a CSV format that data needs normalization further. Normalization is the process of efficiently organizing data in a database. As web scraped data from different online stores which needs to be normalized by filtering or taking only needed data.

3.12 Data Visualization

Data visualization tools provide an accessible way to see and understand trends, outliers, and patterns in data. It involves the creation and study of the visual representation of data. To communicate information clearly and efficiently, data visualization uses statistical graphics, plots, information graphics and other tools. Numerical data may be encoded using dots, lines, or bars, to visually communicate a quantitative message. Effective visualization helps users analyze and reason about data and evidence. It makes complex data more accessible, understandable and usable. Users may have particular analytical tasks, such as making comparisons or understanding causality, and the design principle of the graphic (i.e., showing comparisons or showing causality) follows the task. Tables are generally used where users will look up a specific measurement, while charts of various types are used to show patterns or relationships in the data for one or more variable [17].

3.12.1 Power-Bi

Power BI Desktop designed for desktop or local computers, it's easy to use, connect to any data source or connect one then one data source at a time by data modeling, after connection, data can be filtered and visualized in dashboard and reports, its easily share-able by Power BI service. Initial plan was to use Power BI as Google's Data Studio offered was only to premiums customers [18].

It supports hundreds of data connections and all Power Bi reports can be view on most of latest mobiles and tablets with the help of Power BI apps.

Figure 11 shows demo example of Power BI report dashboard, for desktop, mobile and tablet.

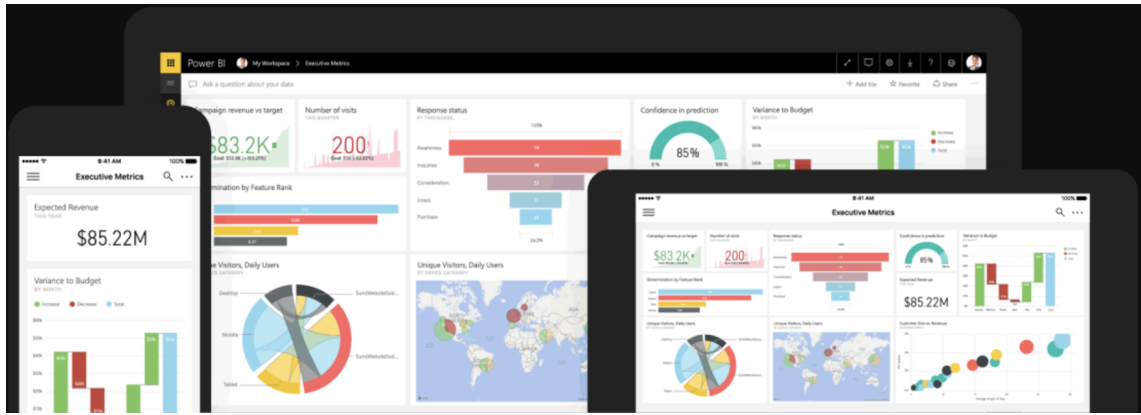


Figure 11. Power BI Dashboards for Desktop, Mobile & Tablet.

3.12.2 Data Studio by Google

Data Studio is Google's reporting solution for power users who want to go beyond the data and dashboards of Google Analytics. The data widgets in Data Studio are notable for their variety, customization options, live data and interactive controls (such as column sorting and table pagination). Data sources include Google products (Analytics, AdWords, Search Console, Sheets, YouTube, etc.), database connectors, file upload, and "community" connectors to popular marketing services [18].

Google Data Studio started out as part of the enterprise Google Analytics 360 suite. In May 2016, Google announced a free version of Data Studio for individuals and smaller teams. At the time, there were differences between the paid version and the free version, such as the number of reports that could be created per account. By February 2017, Google announced free, unlimited Data Studio reports. Since then, the free version of Google Data Studio has continued to add powerful features and usability enhancements.

Reason to switch from Power-BI to Data Studio?

Power-Bi tool is an expensive tool to use as it has subscription-based cost. It was feasible to use for one user but for multiple users it costs extra money. At the same time Google was providing a free to use tool for dashboarding and data visualization that is called Data Studio. It was the main reason to switch to Google Data Studio.

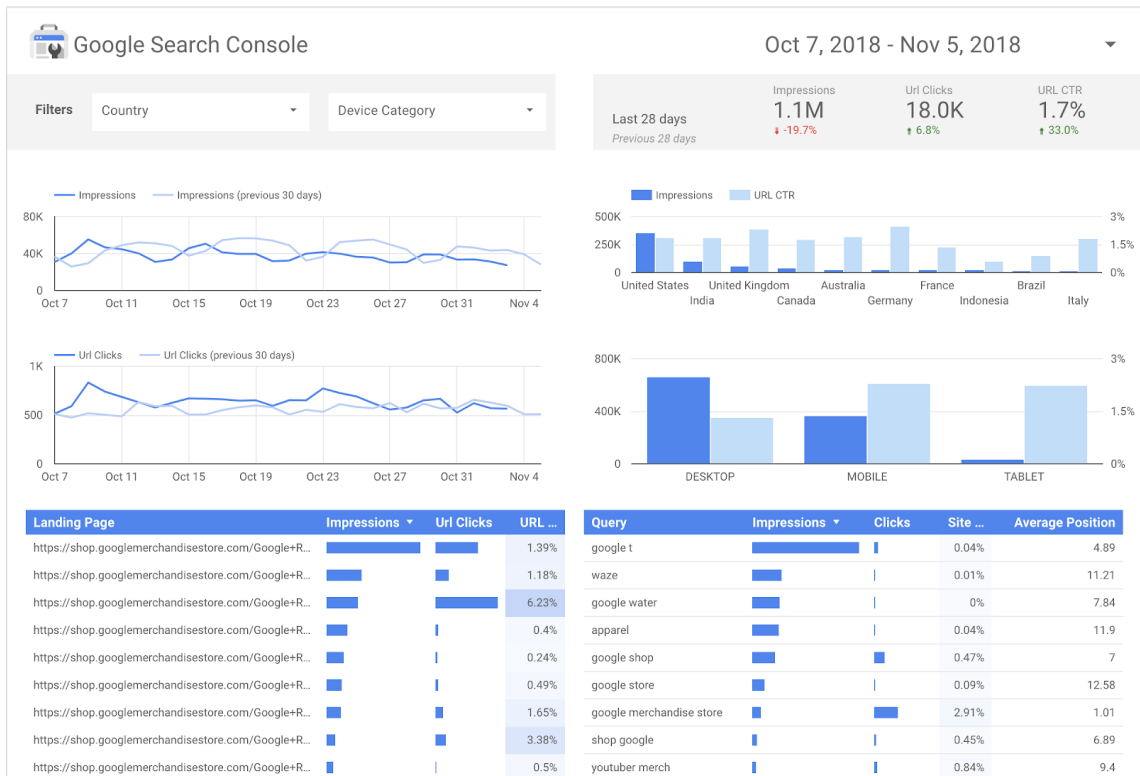


Figure 12. Google Data Studio Search Console.

4 Implementation

The implementation of this tool consists of various other tools and required various steps to work properly. Firstly, products data were required from three different competitors online stores.

4.1 Web scrapper setup

For the data collection, an Open Web Scraper browser extension was used in Chrome browser. This browser extension needs a sitemap to extract the relevant data, which can be defined once and imported to scrape data multiple times. A user needs to create the sitemap based on the keywords such as a store URL or root access location, main category links, sub-category links, product title, product price and product stock. A script is generated based on the sitemap and it can be visualized in graph format as a sitemap shown in Figure 13.

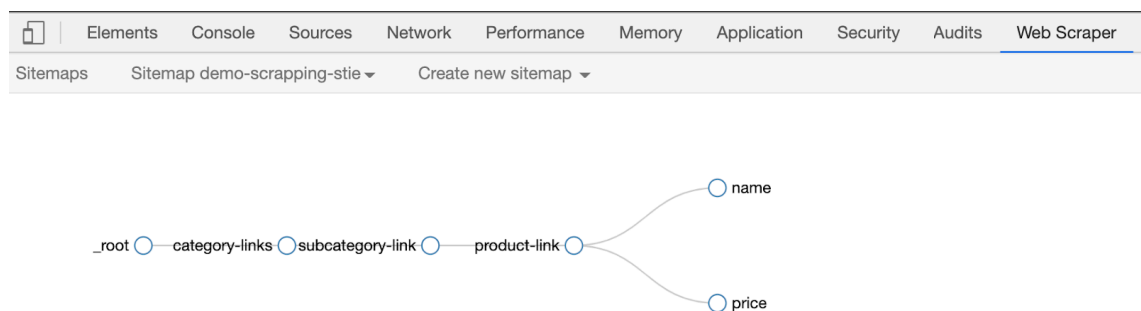


Figure 13. Web Scraper sitemap.

This sitemap can be saved and re-used in future without repeating the same process. After this, a data preview is generated as shown in Figure 14. The preview shows the data coming from the online stores. The preview contains different information in tabular format. First column tells the order number, second column shows URL of the online store from where the data is coming, third and fourth column shows the product category and its corresponding web link such as phone, fifth column and sixth columns shows the subcategory and its corresponding link of the product such as Touch phones, seventh columns depicts the actual product and its corresponding link. The links can help the user to go directly to the product store. The eighth column shows the actual product name and ninth column shows the price of the product [9].

web-scraping-order	web-scraping-start-url	category-links	category-links-href	subcategory-link	subcategory-link-href	product-link	product-link-href	name	price
1558635468-3	https://www.webscraper.io/test-sites/e-commerce/allinone	Phones	https://www.webscraper.io/test-sites/e-commerce/allinone/phones	Touch	https://www.webscraper.io/test-sites/e-commerce/allinone/phones/touch	Iphone	https://www.webscraper.io/test-sites/e-commerce/allinone/product/198	Iphone	\$899.99
1558635472-5	https://www.webscraper.io/test-sites/e-commerce/allinone	Phones	https://www.webscraper.io/test-sites/e-commerce/allinone/phones	Touch	https://www.webscraper.io/test-sites/e-commerce/allinone/phones/touch	Sony Xperia	https://www.webscraper.io/test-sites/e-commerce/allinone/product/196	Sony Xperia	\$118.99
1558635478-8	https://www.webscraper.io/test-sites/e-commerce/allinone	Phones	https://www.webscraper.io/test-sites/e-commerce/allinone/phones	Touch	https://www.webscraper.io/test-sites/e-commerce/allinone/phones/touch	LG Optimus	https://www.webscraper.io/test-sites/e-commerce/allinone/product/193	LG Optimus	\$57.99
1558635480-9	https://www.webscraper.io/test-sites/e-commerce/allinone	Phones	https://www.webscraper.io/test-sites/e-commerce/allinone/phones	Touch	https://www.webscraper.io/test-sites/e-commerce/allinone/phones/touch	Nokia 123	https://www.webscraper.io/test-sites/e-commerce/allinone/product/192	Nokia 123	\$24.99
1558635464-1	https://www.webscraper.io/test-sites/e-commerce/allinone	Phones	https://www.webscraper.io/test-sites/e-commerce/allinone/phones	Touch	https://www.webscraper.io/test-sites/e-commerce/allinone/phones/touch	Iphone	https://www.webscraper.io/test-sites/e-commerce/allinone/product/200	Iphone	\$899.99
1558635476-7	https://www.webscraper.io/test-sites/e-commerce/allinone	Phones	https://www.webscraper.io/test-sites/e-commerce/allinone/phones	Touch	https://www.webscraper.io/test-sites/e-commerce/allinone/phones/touch	Samsung Galaxy	https://www.webscraper.io/test-sites/e-commerce/allinone/product/194	Samsung Galaxy	\$93.99
1558635466-2	https://www.webscraper.io/test-sites/e-commerce/allinone	Phones	https://www.webscraper.io/test-sites/e-commerce/allinone/phones	Touch	https://www.webscraper.io/test-sites/e-commerce/allinone/phones/touch	Iphone	https://www.webscraper.io/test-sites/e-commerce/allinone/product/199	Iphone	\$899.99
1558635470-	https://www.webscraper.io/test-	Phones	https://www.webscraper.io/test-	Touch	https://www.webscraper.io/test-sites/e-	Ubuntu	https://www.webscraper.io/test-sites/e-	Ubuntu	\$499.99

Figure 14. Data preview before scraping is completed [9].

Later the data was exported as CSV (comma-separated values) format as shown in Figure 15. CSV is a simple file format used to store tabular data, such as a spreadsheet or database. The same procedure is used to get other three stores product data and saved into three different CSV files. Total four CSV files were generated using this process.

1	Product Name,Date,Price		
2	Samsung 4* N**** Smart 4K UHD TV Led-Televisio,1/3/2019,1014.62		
3	Philips 5*****/12 55" 4K UHD Smart-TV,1/27/2019,989.00		
4	Philips 3*****/12-LED-tv 32",1/28/2019,989.00		
5	Philips 3*" HD 3*****/12 LED TV,1/29/2019,989.00		
6	LG Ultra HD 4K TV - 55U***,1/30/2019,1000.66		
7	LG Ultra HD 4K TV - 55UK***,2/12/2019,893.55		
8	Philips 65****/12 65" 4K UHD Smart-TV,2/13/2019,893.55		
9	Philips 65****/12 65" 4K UHD OLED Android-TV,2/14/2019,852.05		
10	TX-49FX**** 4K HDR TV,2/26/2019,958.19		
11	43PUS****/12 4K UHD Smart-TV 43",2/27/2019,958.19		
12	LG Ultra HD 4K TV - 43,Ç"-ù 43U*****,2/28/2019,949.57		
13	LG Ultra HD 4K TV - 43,Ç"-ù 43UK*****,3/15/2019,961.81		
14	LG Super UHD 4K TV - 49,Ç"-ù 49S*****,3/17/2019,922.23		
15	Philips 65P*****/12 65" 4K UHD Android-TV,3/22/2019,841.89		
16	49P*****/12 4K UHD Android-TV 49",3/28/2019,641.00		
17	55P*****/12 4K UHD Android TV 55",3/29/2019,671.85		

Figure 15. Web scraped data extracted/download file in CSV format.

4.2 Import CSV data into MySQL database

The tabular form data in CSV files can easily imported to any database. In this study, the data is imported to MySQL database. The competitor's data were stored in three separate CSV files. To get a single product information user needs to search all these files separately. In order to solve this problem, the CSV files were imported to a single database [6].

While importing web scraped data CSV files to database there is need to differentiate each file data from one another, so an extra column added as 'store_name' which represents store or competitor as shown in Figure 16.

id	store_name	product_name	date	price
1	Store1	Samsung 4*" N**** Smart 4K UHD TV Led-Televisio	1/3/2019	1014.62
2	Store1	Philips 5*****/12 55" 4K UHD Smart-TV	1/27/2019	989
3	Store1	Philips 3*****/12-LED-tv 32"	1/28/2019	989
4	Store1	Philips 3*" HD 3*****/12 LED TV	1/29/2019	989
5	Store1	LG Ultra HD 4K TV - 55U***	1/30/2019	1000.66
6	Store1	LG Ultra HD 4K TV - 55UK***	2/12/2019	893.55
7	Store1	Philips 65****/12 65" 4K UHD Smart-TV	2/13/2019	893.55
8	Store1	Philips 65****/12 65" 4K UHD OLED Android-TV	2/14/2019	852.05
9	Store2	Samsung 4*" N**** Smart 4K UHD TV Led-Televisio	1/3/2019	696.16
10	Store2	Philips 5*****/12 55" 4K UHD Smart-TV	1/27/2019	890.81
11	Store2	Philips 3*****/12-LED-tv 32"	1/28/2019	890.81
12	Store2	Philips 3*" HD 3*****/12 LED TV	1/29/2019	890.81
13	Store2	LG Ultra HD 4K TV - 55U***	1/30/2019	890.81
14	Store2	LG Ultra HD 4K TV - 55UK***	2/12/2019	744.33
15	Store2	Philips 65****/12 65" 4K UHD Smart-TV	2/13/2019	744.33
16	Store2	Philips 65****/12 65" 4K UHD OLED Android-TV	2/14/2019	622.24
17	Store3	Samsung 4*" N**** Smart 4K UHD TV Led-Televisio	1/3/2019	1024
18	Store3	Philips 5*****/12 55" 4K UHD Smart-TV	1/27/2019	1079.83
19	Store3	Philips 3*****/12-LED-tv 32"	1/28/2019	1079.83
20	Store3	Philips 3*" HD 3*****/12 LED TV	1/29/2019	1079.83
21	Store3	LG Ultra HD 4K TV - 55U***	1/30/2019	1079.83
22	Store3	LG Ultra HD 4K TV - 55UK***	2/12/2019	1054
23	Store3	Philips 65****/12 65" 4K UHD Smart-TV	2/13/2019	1054
24	Store3	Philips 65****/12 65" 4K UHD OLED Android-TV	2/14/2019	1054
25	Store4	Samsung 4*" N**** Smart 4K UHD TV Led-Televisio	1/3/2019	1649
26	Store4	Philips 5*****/12 55" 4K UHD Smart-TV	1/27/2019	2099
27	Store4	Philips 3*****/12-LED-tv 32"	1/28/2019	2099
28	Store4	Philips 3*" HD 3*****/12 LED TV	1/29/2019	2099
29	Store4	LG Ultra HD 4K TV - 55U***	1/30/2019	1732.33
30	Store4	LG Ultra HD 4K TV - 55UK***	2/12/2019	1765.66
31	Store4	Philips 65****/12 65" 4K UHD Smart-TV	2/13/2019	1765.66
32	Store4	Philips 65****/12 65" 4K UHD OLED Android-TV	2/14/2019	1765.66

Figure 16. All imported CSV files in a single table.

After getting the data into the database a single table was generated as shown in Figure 16. The figure shows all the store names as store1, store 2, store 3, store 4, different product names, date when the product was retrieved and price of each product.

4.3 Searching and filtering of data

Next step is to filter the matching products using the product title. For example, a user may want to check if Samsung 4K UHD TV is available in other competitors stores or not and if a product stock level is zero that that product price needs to be ignored as out of stock product can have old price which can mislead compared prices.

Then the next step is to find the prices of these matching products as explained in the Figure 17.

id	product_name	store1	store2	store3	store4
1	Samsung 4*" N**** Smart 4K UHD TV Led-Televisio	1014.62	696.16	1024	1649
2	Philips 5*****/12 55" 4K UHD Smart-TV	989	890.81	1079.83	2099
3	Philips 3*****/12-LED-tv 32"	989	890.81	1079.83	2099
4	Philips 3*" HD 3*****/12 LED TV	989	890.81	1079.83	2099
5	LG Ultra HD 4K TV - 55U***	1000.66	890.81	1079.83	1732.33
6	LG Ultra HD 4K TV - 55UK***	893.55	744.33	1054	1765.66
7	Philips 65****/12 65" 4K UHD Smart-TV	893.55	744.33	1054	1765.66
8	Philips 65****/12 65" 4K UHD OLED Android-TV	852.05	622.24	1054	1765.66
9	TX-49FX**** 4K HDR TV	958.19	700.51	1008.09	1949
10	43PUS****/12 4K UHD Smart-TV 43"	958.19	700.51	1008.09	1949
11	LG Ultra HD 4K TV - 43? 43U*****	949.57	700.51	989.9	2265.66
12	LG Ultra HD 4K TV - 43? 43UK*****	961.81	852.85	996.33	2265.66
13	LG Super UHD 4K TV - 49? 49S*****	922.23	961.94	996.33	1949
14	Philips 65P*****/12 65" 4K UHD Android-TV	841.89	769.94	1005.18	1949
15	49P*****/12 4K UHD Android-TV 49"	641	461.51	754.33	1999
16	55P****/12 4K UHD Android TV 55"	671.85	479.54	754.33	1999

Figure 17. Products filtered according to the product title.

After matching the products by title and price the next step was to get most sold and least sold products as shown in Figure 18. This process shows the user which products have good price as compare to their competitors and which product price needs to be improved. Next the average price of each product was collected and can be seen in the

same figure. The average price is another way to evaluate the product price and guide the user finalize the product price.

id	date	store1	store2	store3	store4	store1_sold	average
1	1/3/2019	1014.62	696.16	1024	1649	1	1095.94
2	1/27/2019	989	890.81	1079.83	2099	4	1264.66
3	1/28/2019	989	890.81	1079.83	2099	2	1264.66
4	1/29/2019	989	890.81	1079.83	2099	1	1264.66
5	1/30/2019	888.66	890.81	1079.83	1732.33	5	1147.91
6	2/12/2019	893.55	744.33	1054	1765.66	2	1114.39
7	2/13/2019	693.55	744.33	1054	1765.66	6	1064.39
8	2/14/2019	852.05	622.24	1054	1765.66	3	1073.49
9	2/26/2019	958.19	700.51	1008.09	1949	1	1153.95
10	2/27/2019	958.19	700.51	1008.09	1949	1	1153.95
11	2/28/2019	949.57	700.51	989.9	2265.66	0	1226.41
12	3/15/2019	961.81	852.85	996.33	2265.66	4	1269.16
13	3/17/2019	922.23	961.94	996.33	1949	2	1207.38
14	3/22/2019	841.89	769.94	1005.18	1949	0	1141.5
15	3/28/2019	641	461.51	754.33	1999	1	963.96
16	3/29/2019	671.85	479.54	754.33	1999	0	976.18
17	3/30/2019	750.35	479.54	762.11	1999	1	997.75
18	3/31/2019	629	461.51	762.11	1999	2	962.905
19	4/5/2019	566.73	461.51	657.75	1999	5	921.247
20	4/7/2019	571.57	455.11	657.75	1999	3	920.857

Figure 18. Products filtered according to the product title.

4.4 Visualization of data

So far, data extracted from 3 different competitors using web scrapping method. Then the data was imported to a single database table from 3 different CSV files. After importing the data, data was filtered and compared with competitors' products. All this data can be seen in a tabular form, but this data looks really simple and less interactive. To make the output data interactive and colorful, it needs to be view in the form a chart [17]. Firstly, the data from excel sheet imported to a excel chart as shown in Figure 1.

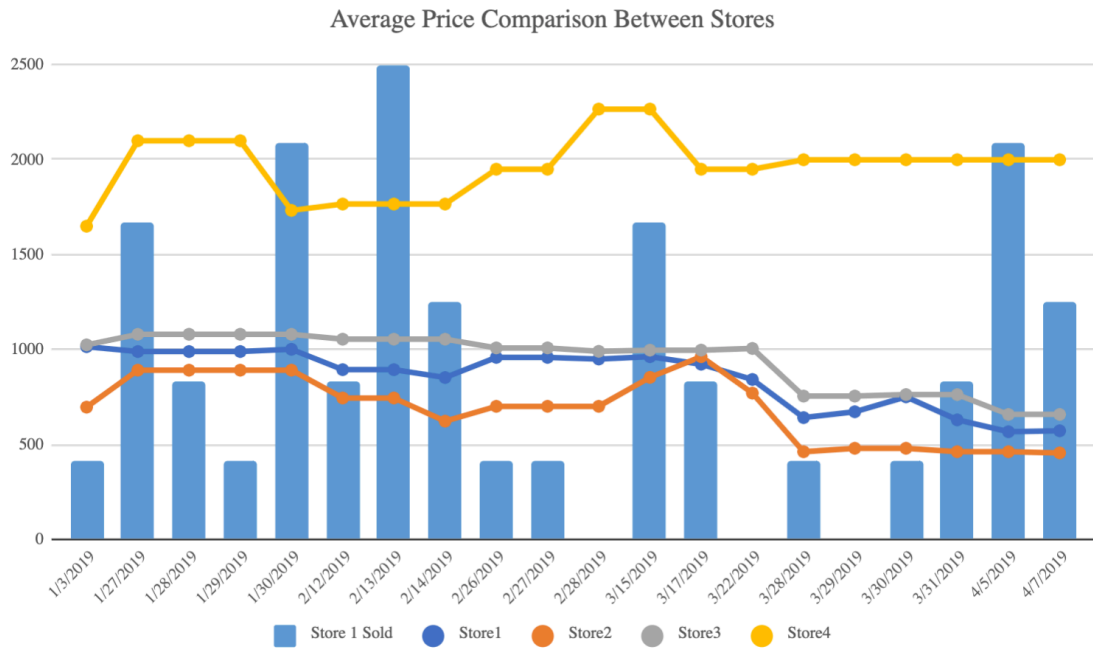


Figure 19. The visual form of the data as a chart.

The above chart shows the average price comparison between stores. The chart takes store1 sold product price as a main parameter and shows the product price of competitors. Moreover, it shows the comparison of prices and average price. The chart shows the product sold date at the x-axis; the y-axis shows the quantity of each sold product.

The product data of each competitor stores shows in different color lines. As dark blue line shows the product price of store 1, orange line shows the product price of store 2, grey line shows the product data of store 3 and yellow line shows the product data of store 4. These lines clearly indicate when store 1 is selling product expensive and when it is selling products cheaper as compared to their competitors.

Finally, the data visualization tool such as Data Studio by Google was used to generate a dashboard.

4.5 Creating Dashboard in Data Studio

To create a dashboard in Data Studio a user needs to create an account or use an existing one and needs to define a data resource. Data Studio offers various options as a data source and a user can select multiple data sources for a dashboard. Figure 20 shows supported data source options for creating a dashboard [19 pp. 26, 27, 28].

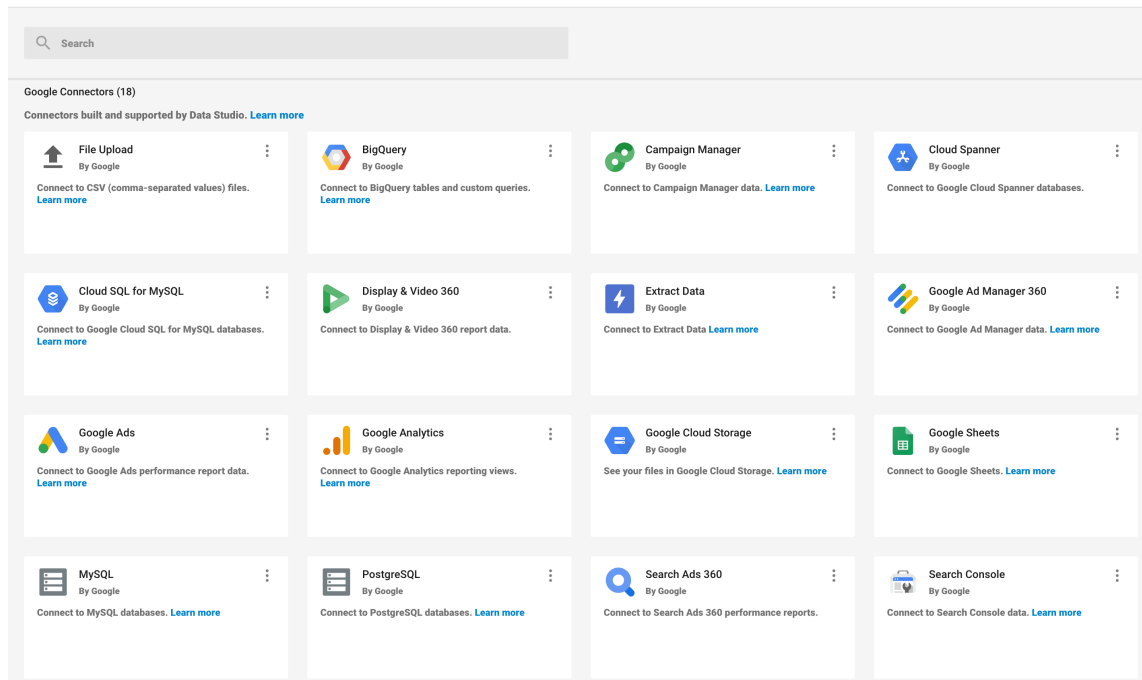


Figure 20. Supported data source options for creating a dashboard.

In this study, the above filtered data file discussed in section 4.1 in a CSV format was used as a data source. The MySQL database cannot be used as a data source at this time because MySQL database should be hosted on a server. After selecting the file, user needs to confirm column fields types or aggregation which will help Data Studio in dashboard creation.

Index	Field	Type	Aggregation	Description
1	id	123 Number	None	
2	date	Date (YYYYMMDD)	None	
3	store1	123 Number	None	
4	store2	123 Number	None	
5	store3	123 Number	None	
6	store4	123 Number	None	
7	store1_sold	123 Number	None	
8	average	123 Number	None	

Figure 21. Columns with names and types which user need to update.

After successfully setting up the data source, a user can create the dashboard. Data Studio offers various chart options, in this study a table chart was used. Figure 22 shows the different chart options and different layout and theme options that can be easily selected.

One dashboard can have multiple pages and a user can add multiple charts or graphs in one page.

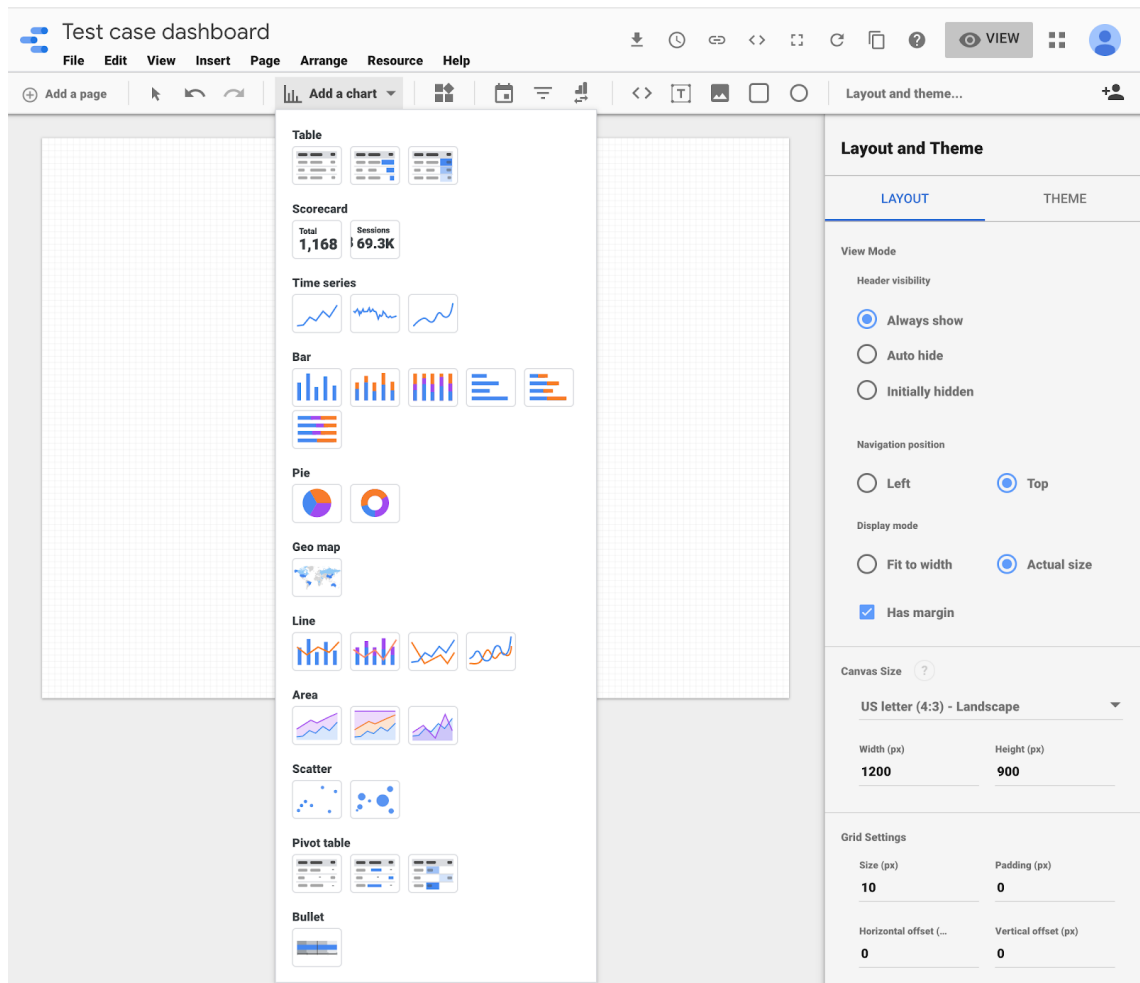


Figure 22. Available chart options.

The next step is to select the required fields available in the selected data source and give this data to the selected chart. For example, one could select product prices and average prices from all stores. Figure 23 illustrates the selected interactive chart and selected data. A user can view the selected data in different colors and a user can see extra information by hovering the mouse on any chart bar.

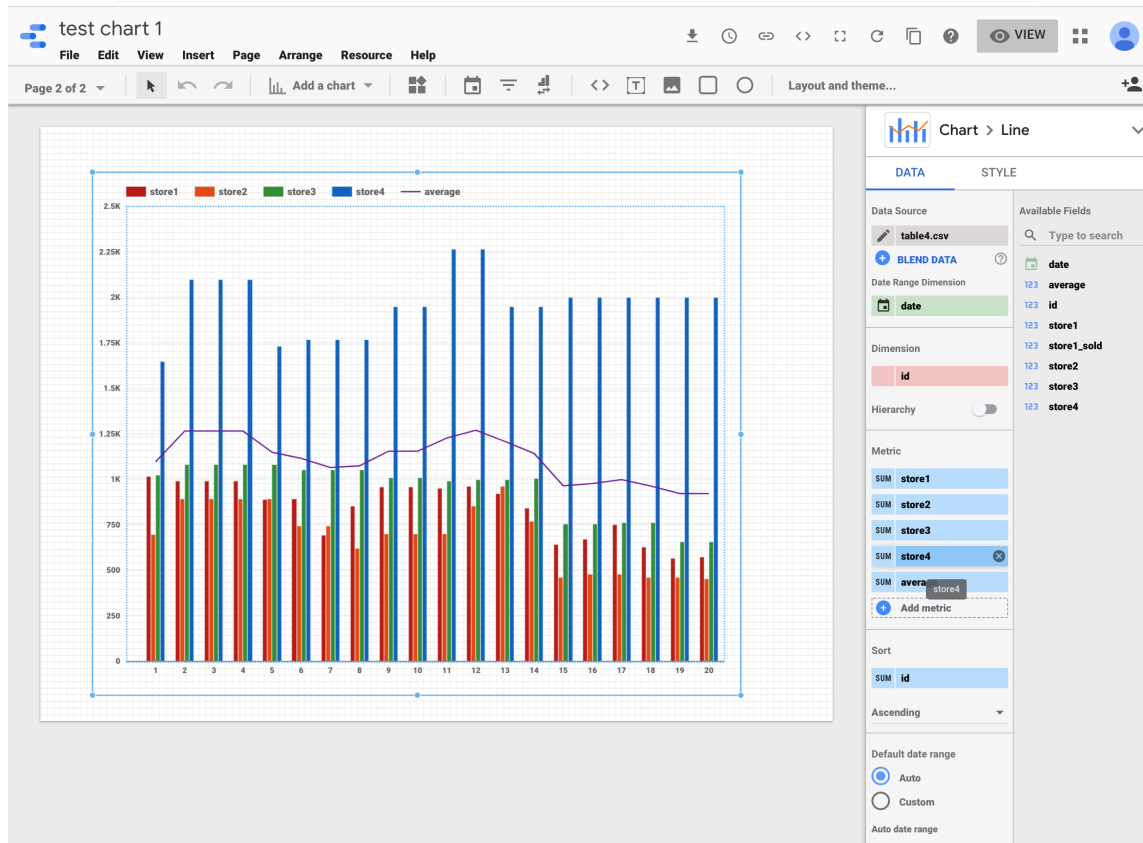


Figure 23. Options set for dashboard main chart.

In Figure 23, next to Data tab there is a STYLE tab, where user can define color or type for each store, for example store1 has red color with series as bar and average has purple color and series are line. Data and STYLE tabs gives many customization options which could be easily used for future improvement.

5 Discussions and Conclusions

This study was based on a real time problem faced by an online store retailer. The problem was how to decide a right price for a product? The retailer was using a basic formula for product pricing. This formula had two main problems

- Some products were sold a lot, and some were sold less.
- Finding product prices of competitors and to compare them with own products prices.

This study explained the implementation of a simple tool that can solve the above problems. The used tools performed the following steps:

- gathering the competitor's data,
- filtering the data,
- comparing the data and
- visualizing the data with the help of dashboard.

In the process of developing the tool the required knowledge was gained in several aspects:

- new tools.
- Web Scraping technologies,
- dashboard design for data visualization in Data Studio,
- server setup and maintenance,
- documentation, and
- NDA rules.

Based on the preliminary feedback from four retailers this tool study can help them to adjust their product prices. Moreover, it provides the following advantages:

- increased sales, average order values,
- reduced the logistics and supply chain costs, and
- encourage customers to buy more the one product.

References

1. <https://www.accountingtools.com/articles/2017/5/16/cost-plus-pricing>
2. <https://fitsmallbusiness.com/product-pricing/>
3. <https://medium.com/@nyxonedigital/importance-of-e-commerce-and-online-shopping-and-why-to-sell-online-5a3fd8e6f416>
4. <https://devdocs.magento.com/guides/m1x/>
5. Hudson, P (2006). PHP in a Nutshell, Sebastopol: O'Reilly Media. Access 2.05.2018
6. MySQL available at <https://dev.mysql.com/doc/refman/8.0/en/what-is-mysql.html> Accessed 2.05.2018
7. Sonassi available at <https://www.sonassi.com/magento-hosting> Accessed 11.05.2018
8. <https://www.scrapehero.com/open-source-web-scraping-frameworks-and-tools/> Accessed 30.05.2019
9. <https://webscraper.io/documentation> Accessed 30.05.2019
10. https://linuxhint.com/top_20_webscraping_tools/ Accessed 30.05.2019
11. <https://www.ppchero.com/comparison-shopping-engines-a-deep-dive-into-comparison-shopping-engines-strategies-part-1/> Accessed 30.05.2019
12. <https://www.shopify.com/blog/7068398-10-best-comparison-shopping-engines-to-increase-ecommerce-sales> Accessed 30.05.2019
13. Vertaa <https://www.vertaa.fi/> Accessed 30.05.2019
14. Hintaseuranta <https://hintaseuranta.fi/> Accessed 30.05.2019
15. Shopzilla <http://www.shopzilla.com/> Accessed 30.05.2019
16. <https://www.thebalancesmb.com/retail-pricing-strategies-2890279> Accessed 30.05.2019
17. <https://infogram.com/page/data-visualization> Accessed 30.05.2019
18. Power-BI available at <https://powerbi.microsoft.com/en-us/> Accessed 10.05.2018
19. <https://datastudio.google.com> Accessed 05.11.2018