

UNDERGRADUATE PROJECT PROPOSAL

|  |  |
| --- | --- |
| **Project Title:** | **Design of a web-based price comparison system for an online shopping platform** |
| **Surname:** | **Zhu** |
| **First Name:** | **Xunran** |
| **Student Number:** | **201918010408** |
| **Supervisor Name:** | **Albert Xu** |
| **Module Code:** | **CHC 6096** |
| **Module Name:** | **Project** |
| **Date Submitted:** | **2022/11/15** |

**Table of Contents**

[1 Introduction 3](#_Toc119407001)

[1.1 Background 3](#_Toc119407002)

[1.2 Aim 3](#_Toc119407003)

[1.3 Objectives 3](#_Toc119407004)

[1.4 Project Overview 4](#_Toc119407005)

[1.4.1 Scope 4](#_Toc119407006)

[1.4.2 Audience 4](#_Toc119407007)

[2 Background Review 4](#_Toc119407008)

[3 Methodology 5](#_Toc119407009)

[3.1 Approach 5](#_Toc119407010)

[3.2 Technology 7](#_Toc119407011)

[3.3 Version management plan 7](#_Toc119407012)

[4 Project Management 7](#_Toc119407013)

[4.1 Activities 7](#_Toc119407014)

[4.2 Schedule 8](#_Toc119407015)

[4.3 Data management plan 8](#_Toc119407016)

[4.4 Project Deliverables 8](#_Toc119407017)

[5 References 9](#_Toc119407018)

# Introduction

## Background

Online shopping is becoming more and more popular today. Many consumers regard online shopping as another shopping channel or mode [1]. Take China as an example, the sales volume of the Tmall Double Eleven shopping carnival in 2020 reached 498.2 billion yuan [2]. Because of the different operation mode, for the same goods, these shopping networks at the same time for the same goods are often different prices. For customers, when shopping online, they all want to get the goods they want at the lowest price. In order to achieve this goal, we need to go to each shopping website to check the price of the goods we want, and then we can know the lowest price after comparison. Obviously, this is a lot of hassle, requiring users to open many Web pages, search for products, record prices, and then compare. If a commodity price comparison system can automatically collect commodity price information on shopping websites and show the results of price comparison, it will provide users with great convenience. The purpose of this paper is to study how to realize this kind of price comparison system.

## Aim

This paper uses the method of data extraction based on Web. By integrating the search function of the homepage of various shopping websites, it searches the goods that users want to buy, compares the queried results, and finally presents the comparison results to users.

## Objectives

(1) Consult and collect literature related to the subject of the project, and conduct background investigation for the project. On this basis, the project is studied and studied, and the project conceptual model is proposed.

(2) Design the price comparison system, and use crawler technology for data extraction and data collection.

(3)Design UI, that is, design web user interface, so that users can operate hardware conveniently and efficiently to achieve two-way interaction.

(4) Design the database, including requirements analysis, conceptual structure design, logic design (ER modeling), physical design, successful implementation of the database and database maintenance.

(5) Test all the function.

(6) To achieve the web page to show users the results of commodity price comparison of different platforms.

## Project Overview

### Scope

The price comparison system of online shopping platform is based on the web, which aims to directly compare the price of the same commodity on different platforms for users, without the need to open multiple web interfaces, which greatly provides convenience for users. The price comparison system first requires the user to register an account, and after logging in the account, a homepage will be displayed, which contains a search box. When the user enters the same type of commodity that he wants to compare the price in the search box, the corresponding price comparison results and the store platform information of the commodity will be displayed on the interface. In addition, the results of commodity price comparison can be arranged in ascending or descending order.

### Audience

The project's main audience is those who like to shop online and compare prices.

# Background Review

Sharma and Gupta[3] report that Web crawler traverses the Web by following the ever-changing, dense and distributed hyperlink structure. In Shi, Shi and Lin's article [4], text-based crawling and web link-based crawling are traditional crawling strategies. Text content-based methods mainly analyze text and URL in web pages, represented by fish search strategy, shark search strategy [10] and optimal priority strategy [11]. Vadivel et al. [5] explained in their article that Web crawler obtains Web pages and other multimedia files from the Web, preprocesses the pages, and extracts information [12].

HTML5 technology is widely used in the development of Web service platform. As mentioned in Zhao's article [6], analyzing interactions is the first step in the design process. By collecting users' ideas, a good foundation is laid for the design of the system. The second step is to identify the view interface. The third step is to list the elements. Four is grouping. The fifth is to create a paper prototype. The sixth is the lo-fi prototype. Seventh, feasibility analysis. In the article Yi,Feng and Li[7], the basic front-end technique is described. These are HTML, CSS, and JavaScript. HTML is its Structured Standard Language. CSS as the standard language. JavaScript is the standard language for behavior.

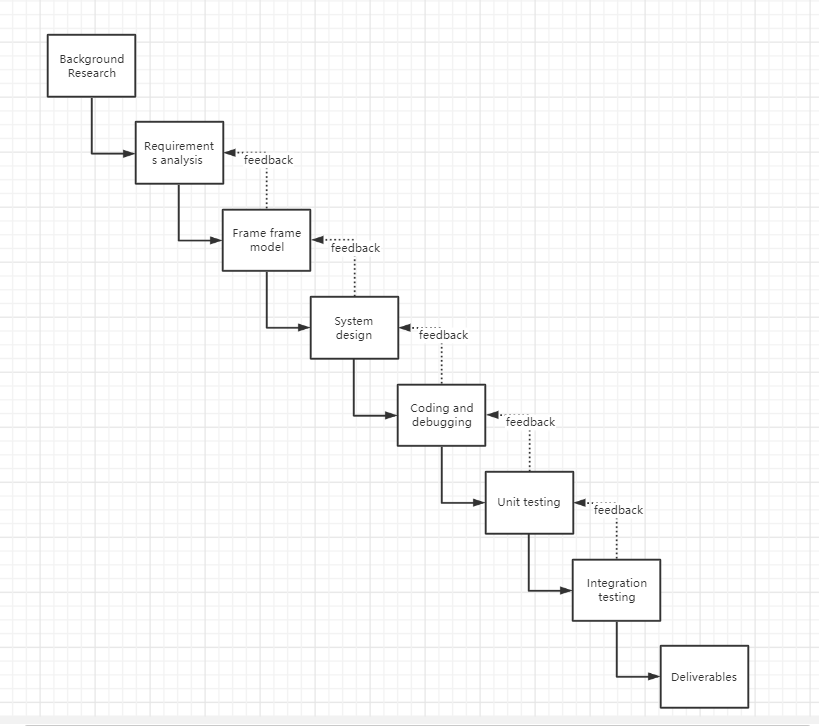
As can be seen from Wu's article [8], database system is a system composed of database system and its corresponding database management software. DBS is an ideal data processing core developed to meet the needs of data processing. For most users, the only way the user interacts with the database is through the database management system. Attaran and Far[9] pointed out the design methods of different types of databases. Generally, three schema concepts should be followed: conceptual schema, logically structured data, and physical database design [13]. Edgar. F. Codd, inventor of the relational model, proposed the concept of normalization in the 1970s, including the first, second and third normal forms, namely the currently known INF, 2MF and 3NF[14].

# Methodology

## Approach

This part is mainly divided into four parts, development model diagram, requirements analysis, technical framework and testing.

3.1.1

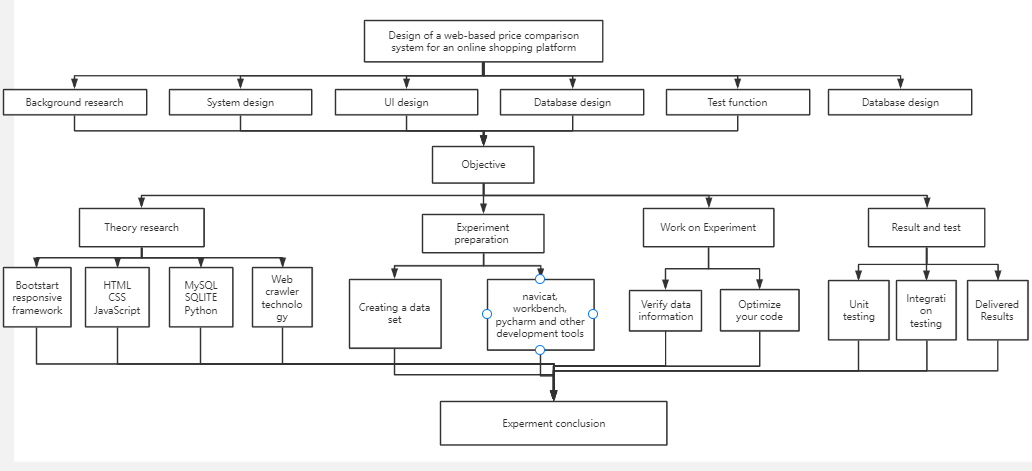


(Figure 1: Development model diagram)

3.1.2 Requirements analysis

The price comparison system must realize the registration and login of the user, after login, the data can be obtained through the search box of the home page, and the extracted data information can be displayed on the user interface, such as: "commodity price", "commodity name" and so on. To achieve this goal, it is necessary to use the crawler technology, build the user interface and build the database. The back-end development software uses pycharm, the framework is django, and the database uses mysql or sqlite. The user interface can use the bootstart responsive framework.

3.1.3



(Figure 2: Technical circuit diagram)

3.1.4 Test

Black box testing, also known as functional testing, is mainly aimed at finding the requirements of software design or the defects in the software design specification. The methods adopted are boundary value analysis and false speculation. There are three methods for white box testing. One is program structure analysis, which analyzes the structure of the program according to the flowchart drawn. The second is the logical coverage square test, which tests all the paths according to the internal structure of the program. The third is the basic path test, which analyzes the paths in the program according to the logic judgment of the program and then designs the use cases.

## Technology

Software: PyCharm, Navicat

Programming language: Python

Browser: Microsoft Edge

Database: MySQL5.7

Frame: Django

## Version management plan

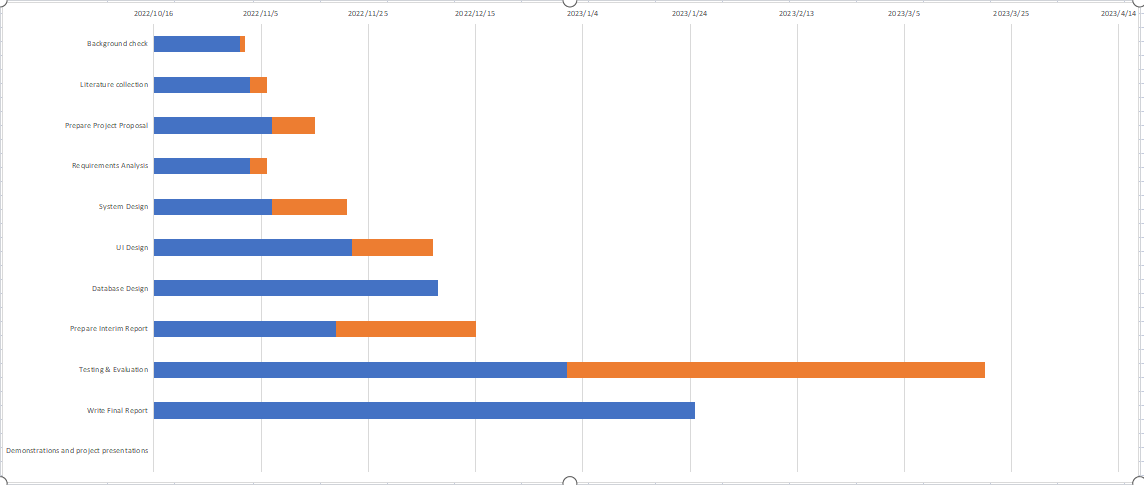
The project will use GitHub as a version management tool to create a separate folder called Project for adding and storing code and files, and categorizing each part of it.

# Project Management

## Activities

|  |  |
| --- | --- |
| Objective 1: Study existing system models | Ac1: Background check |
| Ac2: Collection of literature |
| Objective 2: Design of price comparison system | Ac1: Understand web crawler technology |
| Ac2: Using web crawler technology to develop the system |
| Objective 3: User interface design | Ac1: Design the user interface, mainly including login and registration homepage, as well as each interface contains pictures and interactive buttons |
| Ac2: Code |
| Objective 4: Design of database | Ac1: Create a data set according to the model, collect and store the data |
| Ac2: Simple processing and analysis of data |
| Ac3: Will be presented to the user through the user interface |
| Objective 5: Test all function | Ac1: Unit test each part of the system randomly for several times and observe the results |
| Ac2: Test the whole system randomly for several times and observe the results |
| Objective 6: Show the comparison results | Ac: Successfully display the price comparison results of the same commodity on different platforms |

## Schedule



(Figure 3: Gantt Chart)

## Data management plan

Data related to the project will be stored on GitHub in a folder called Project and updated as the developer progresses until the project is introduced.

## Project Deliverables

4.4.1 Project Proposal

4.4.2 Project Weekly Report

4.4.3 Project Progress report

4.4.4 Project Final Report

4.4.5 Project Code

4.4.6 Project Presentation

4.4.7 Price comparison system of online shopping based on web

# References

[1] Han’guk Chŏnja T’ongsin Yŏn’guwŏn, Han’guk Chŏngbo Sahoe Chinhŭngwŏn, Global IT Research lnstitute (Korea), & IEEE Communications Society. (n.d.). *The 11th International Conference on Advanced Communication Technology : ubiquitous ICT convergence makes life better! : ICACT 2009 : Phoenix Park, Korea, Feb. 15-18, 2009 : proceedings*.

[2] Guo, Q., Yu, W., Ma, Z., & Kong, J. (2021). User Behavior Monitoring Mechanism of Online Shopping Processes based on Hierarchical Colored Petri Net. *Proceeding - 2021 China Automation Congress, CAC 2021*, 6118–6122.<https://doi.org/10.1109/CAC53003.2021.9728033>

[3] Sharma, S., & Gupta, P. (2015). The anatomy of web crawlers. *International Conference on Computing, Communication and Automation, ICCCA 2015*, 849–853. https://doi.org/10.1109/CCAA.2015.7148493

[4] Shi, Z., Shi, M., & Lin, W. (2016). *The Implementation of Crawling News Page Based on Incremental Web Crawler; The Implementation of Crawling News Page Based on Incremental Web Crawler*. https://doi.org/10.1109/ACIT-CSII-BCD.2016.74

[5] Ieee. (2012). *2012 International Conference on Advances in Engineering, Science and Management.* IEEE.

[6] Haitao, Z. (2021). Research on Mobile Web Front-end Design Based on HTML5 Technology. *Proceedings - 2021 6th International Symposium on Computer and Information Processing Technology, ISCIPT 2021*, 290–293. https://doi.org/10.1109/ISCIPT53667.2021.00065

[7] Yi, N., Feng, X., & Li, C. (n.d.). *2017 7 th International Conference on Communication Systems and Network Technologies The Design and Implementation of the Front End of the Art Play Library System*. https://doi.org/10.1109/CSNT.2017.54

[8] Wu, Y. (2017). Design of user database resource management system based on web. *Proceedings - 2017 International Conference on Computer Technology, Electronics and Communication, ICCTEC 2017*, 364–367. https://doi.org/10.1109/ICCTEC.2017.00084

[9] Attaran, H., & Far, A. H. (2011). A novel technique for object oriented relational database design. *Proceedings of 2011, 10th IEEE International Conference on Cybernetic Intelligent Systems, CIS 2011*, 128–132. https://doi.org/10.1109/CIS.2011.6169147

[10] M. Hersovici, M. Jacovi, Y. S. Maarek, et al, “The shark-search algorithm. An application: tailored Web site mapping,” Computer Networks and ISDN Systems, 1998, pp. 317-326.

[11] F. Menczer, G. Pant, P. Srinivasan, “Topical web crawlers: Evaluating adaptive algorithms,” ACM Transactions on Internet Technology (TOTIT), 2004, pp. 378-419.

[12] V. Shkapenyuk , T. Suel, "Design And implementation of a High Performance Distributed WebCrawler", fn fCD£. 2002.

[13] R. Kaula, and J. Chin “A database approach towards flexible manufacturing: A conceptual framework, “ in Computers and Industrial Engieering, Vol 24, No 2, pp. 131-141, Apr 1993

[14] P. K. Janert, “Practical Database Design, Part 2; Normalization, History, and Event Logging, “ From http://www.devx.com/ibm/Article/20859/4047, Accessed 2011