Clothes Shopping Website

**GROUP 5 of BigData3:**

Leader:

闫冰洁(Yan Bingjie)

Members:

牛聪(Niu Cong)

滕畅(Teng Chang)

王楠(Wang Nan)

2019.12.18

CERTIFICATE

This clothes shopping website project&data analysis report reflects the original project independently completed by 4 students from the fifth group of class 3, big data, according to the requirements of niit .

Coordinator: Rupali、Abhijeet、Delucia

# ACKNOWLEDGEMENT

**1.What inspired us to write this project**

First of all, in terms of technology, there are many websites similar to online clothing sales. From the perspective of type and data, the requirements and related analysis of such websites are very comprehensive.If our production hits a bottleneck, there will be a lot of information for us to refer to.Secondly, in terms of demand, such websites are more in line with the requirements of the public, and the data will be more careful and large. The implementation of relevant demand analysis can provide more accurate data support. We can effectively integrate the quality and quantity of user feedback in limited time, providing us with more opportunities to maintain and upgrade.

**2.Who helped us**

In the process of making the difficult project, we have received a lot of help from many people, including Mr. Liu decai, Ms. Rupali, Mr. Abhijeet and all the classmates and relatives who have put forward ideas, ideas and provided help for our project, all of whom have provided us with great technical and emotional help.

**3. Show appreciation**

We would like to express our sincere thanks to all the people who helped us overcome the difficulties and stayed with us through this boring and interesting day.

**4. Briefly introduce the process of developing this project**

At the center of the panel discussion, we had four major discussions on the development of the project.In our first discussion, we determined the division of labor.According to the weight of each task, the team leader arranged different Numbers of team members to complete the task, which pointed out the direction for our future efforts.The second time, nearly a week after the first discussion, we completed and integrated all our project contents, and made comments and Suggestions for improvement within the group.Then we had the next round of rectification.The third time, two days after the second discussion, we consolidated and discussed all the contents again and confirmed the completion of the project.In the fourth discussion, we determined the content and presentation of PPT.

## Description about Technologies used

Operating system: Windows、Linux

Database: HBase

Servlet: Tomcat

Database.Connection.driver: Java web integrated development environment.

Data Analysis: Hive MapReduce

# System Analysis(Requirements)

## 1. Functional requirement

|  |
| --- |
| **User login and registration** |
| **View the main category of goods** |
| **Search for goods** |
| **Add to shopping cart** |
| **Payment** |
| **Data Analysis** |

1.1 Function block description

(1)Registration function.

Customers first have to register as onlinemall users. When registering,youonly need to fill in the login user name, password and E-mail address. After registration, the user can continue to truthfully fill in the detailed personal information and consignee information, and can change the password, query and modify the order.  
(2)Select product features.

Customers browse online shopping malls, their own needs of products into the shopping cart, can continue to add goods.

(3) Manage the shopping cart.

After selecting the goods, the customer can go to the shopping cart page, view the goods he wantsto buy, modify the quantity of a certain item, cancel the purchase of a certain item and empty the whole shopping cart.

(4) Order function.

The customer submits the order after identifying the items in the shopping cart. If the customer has filled in the consignee information, the page displays the information and the customer confirms it. If it is not filled in, the corresponding form will be displayed for it to fill in. The system will record the consignee information submitted by the customer for use in the next shopping. After submitting the order, customers can check the order inthe online mall and cancel or modify the order that has not been processed.

## Analysis requirement for own log

* 1. Top 5 most popular items

The number of clicks on a product can reflect the popularity of a product, so our definition of the most popular product is Top5 of the most clicked products of all products. Make block statistics for each item in the table through "Group by", then sort all statistics in reverse with "Order by DESc", select the top 5 statistics in "Limit 5", and finally get the 5 items with the most clicks.

* 1. Top 5 best sellers

The sales volume of a product can reflect the best-selling degree of a product. We use "Group by" to block each "actid=1" in the table, then reverse order all statistics using "Order by DESc" and filter the first five statistics using "Limit 5" to get the top five items sold.

* 1. Top 3 purchase rate items

The product purchase rate can reflect the user group's loyalty to the product and the degree of acceptance of the product description. The number of actid==1 was summed and then compared to the total number of actid==1 to determine the top five items with the highest purchase rate.

* 1. Top 5 purchase combinations

By analyzing the combination of purchases, you can use bundled sales or special promotions to increase sales. Use Map to count the individual purchases of all users, count all product combinations during the Reduce phase, and finally calculate Top5 during the cleanup phase.

## 3.Analysis requirement for internet log

3.1 Top 5 most popular items

The number of clicks on a product can reflect the popularity of a product, so our definition of the most popular product is Top5 of the most clicked products of all products. Make block statistics for each item in the table through "Group by", then sort all statistics in reverse with "Order by DESc", select the top 5 statistics in "Limit 5", and finally get the 5 items with the most clicks.

3.2 Top 5 best sellers

The sales volume of a product can reflect the best-selling degree of a product. We use "Group by" to block each "actid=1" in the table, then reverse order all statistics using "Order by DESc" and filter the first five statistics using "Limit 5" to get the top five items sold.

3.3 Top 3 purchase rate items

The product purchase rate can reflect the user group's loyalty to the product and the degree of acceptance of the product description. The number of actid==1 was summed and then compared to the total number of actid==1 to determine the top five items with the highest purchase rate.

3.4 Top 5 purchase combinations

By analyzing the combination of purchases, you can use bundled sales or special promotions to increase sales. Use Map to count the individual purchases of all users, count all product combinations during the Reduce phase, and finally calculate Top5 during the cleanup phase.

# Detailed introduction

## 1. Software positioning

The positioning of this website is to provide users with a more convenient and fast way to buy clothes.That is, it breaks through the limitation of region in space, and gives users more power to choose clothes.In terms of time, it saves the time of going to different shopping malls and stores to choose and try on clothes, adding more fun to people's life.

## 2. Basic functions of the software

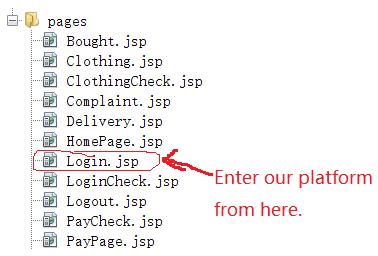
The website is a user-oriented platform, including: online registration, online login, browse the product page, shopping, order submission, payment, view orders, complaints and other operations.

## 3. How to run your website

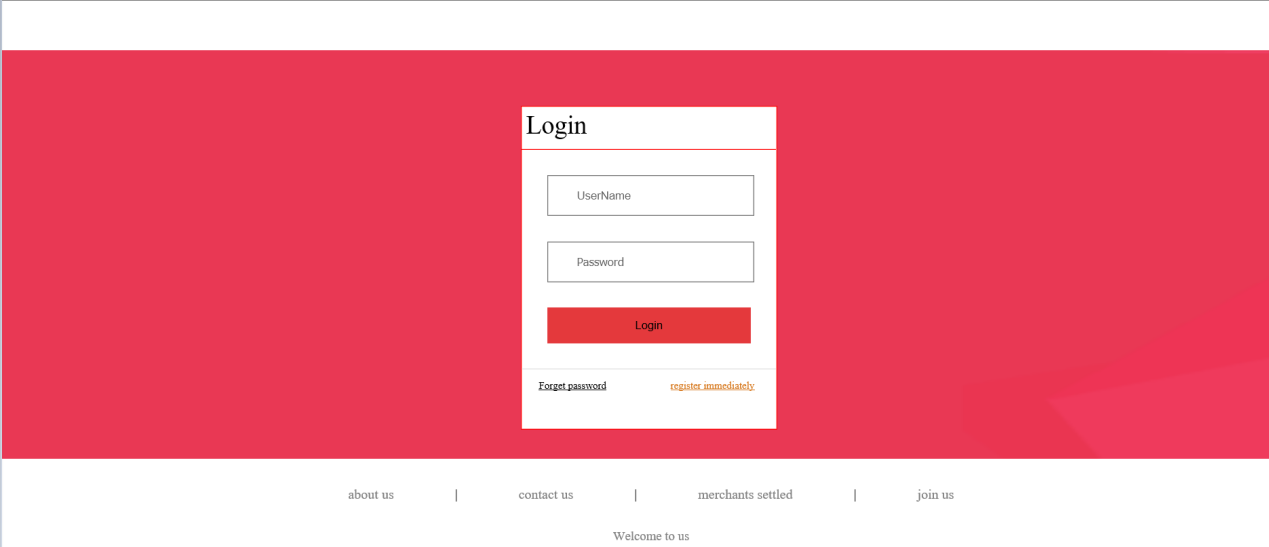
Use netbeans IDE software to run the program.

## Head-end instructions

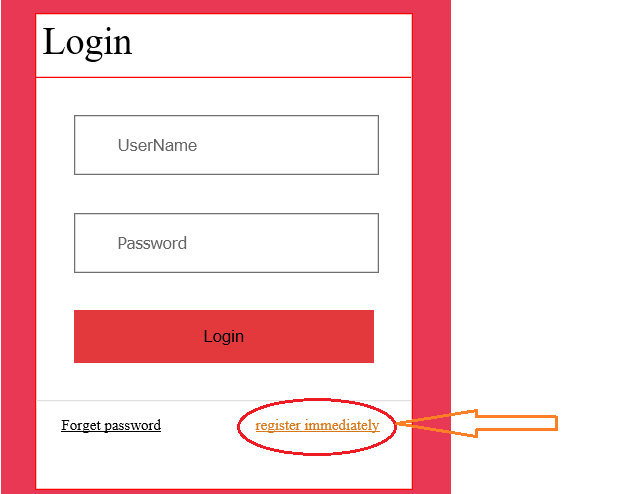
(1) From/Project/webpage/first/pages for the Login. jsp, the right to run the. The JSP file



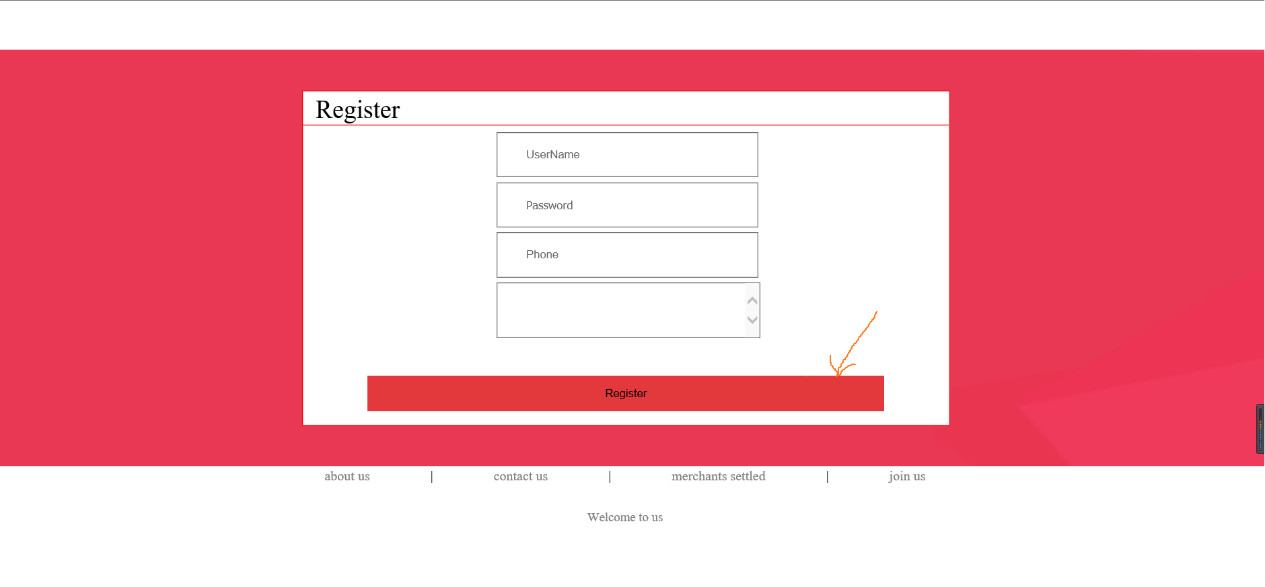
1. After running, the following interface appears



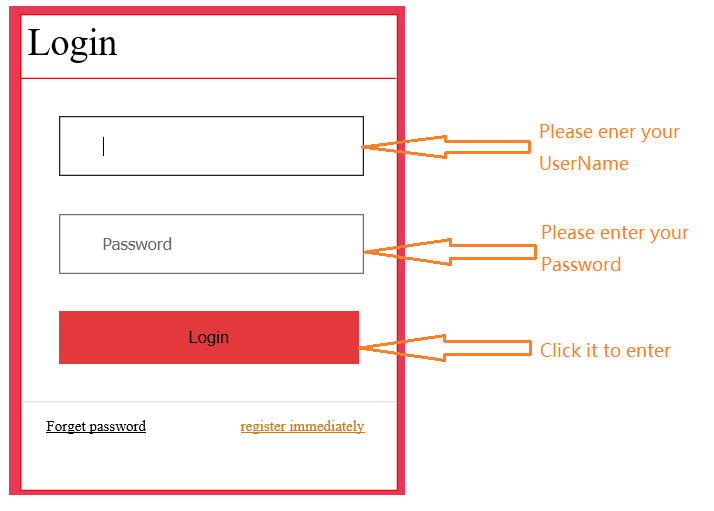
1. If this is the first time for you to log into our website, please click "Register immediately" in the bottom right corner for user registration (if you have already registered, ignore this step)



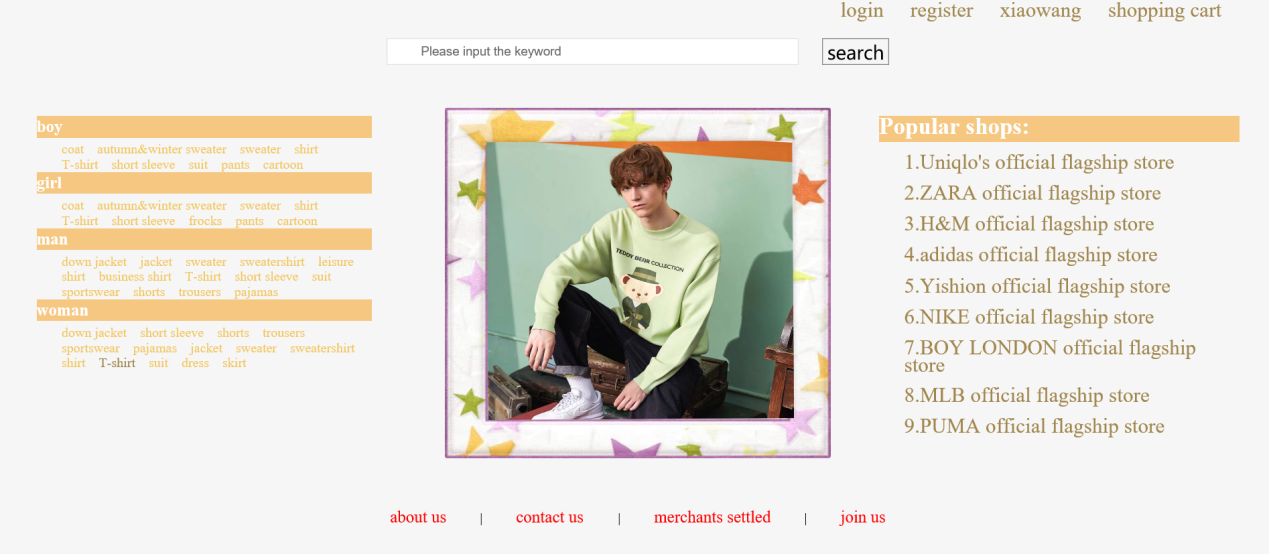
(4) After filling in the user registration interface, click the "Register" button to go back to the Login interface to Login (if you have already registered, ignore this step)



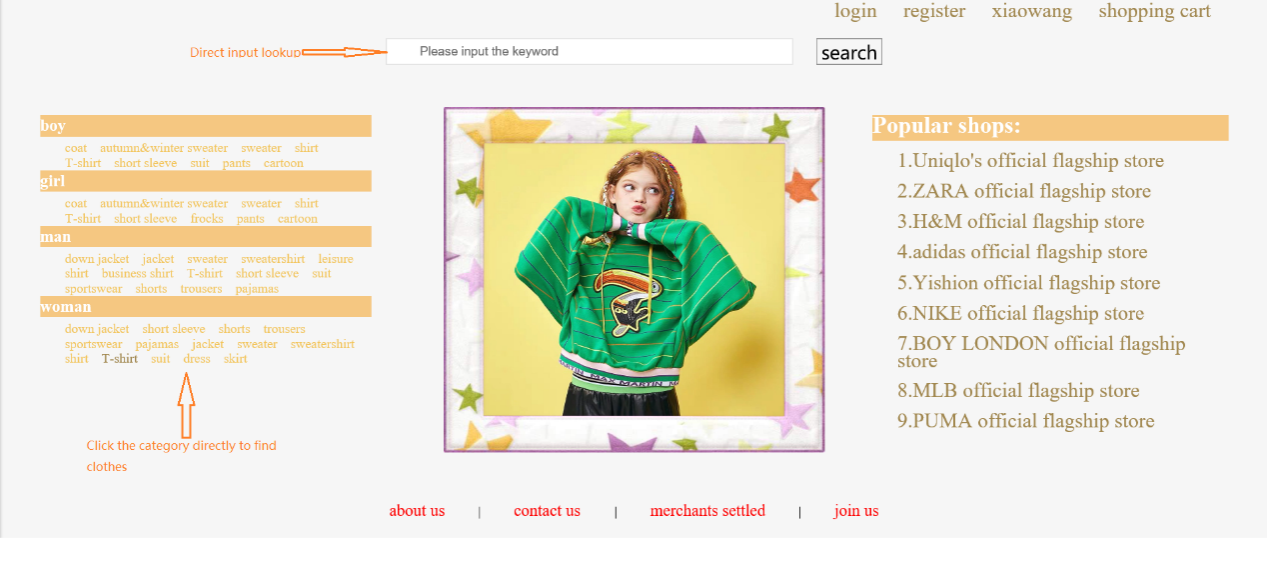
(5) Enter the account and password, then click "Login" button to enter the website.



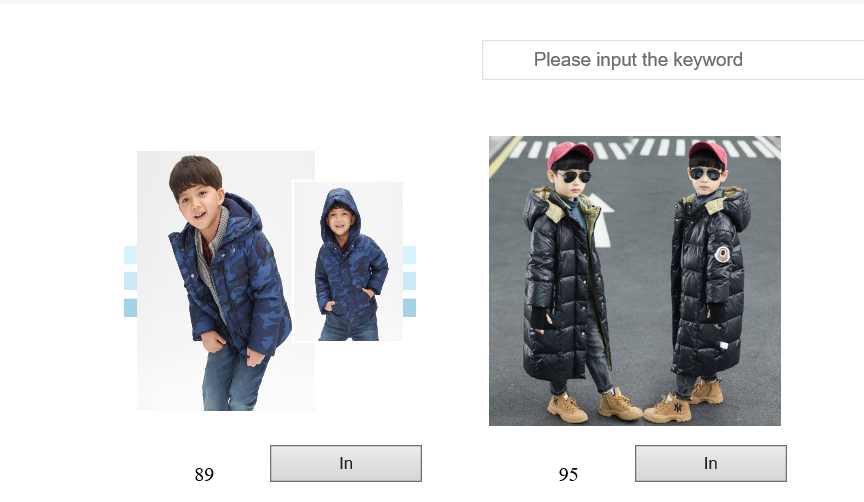
1. The following page appears after entering this website:



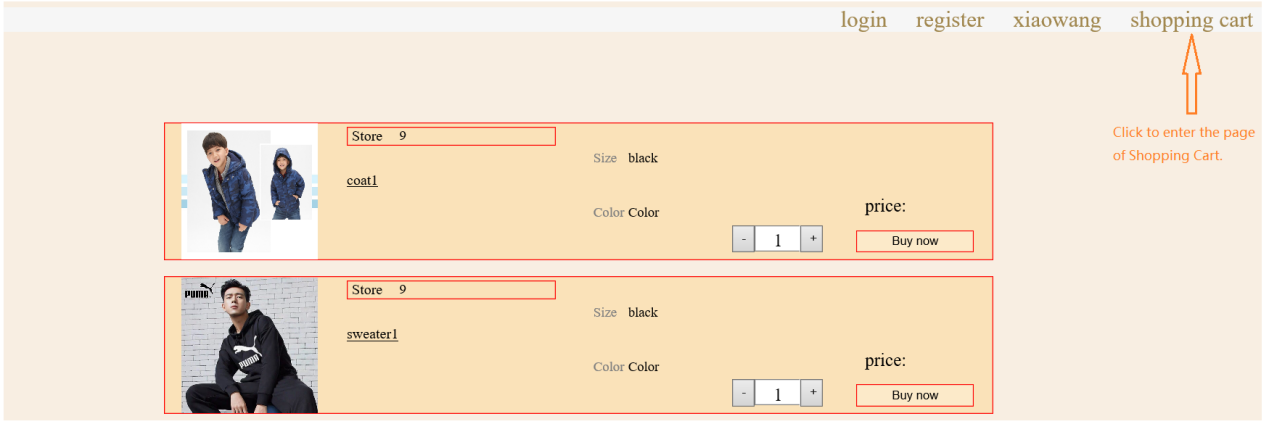
1. There are two ways to find products in this website:



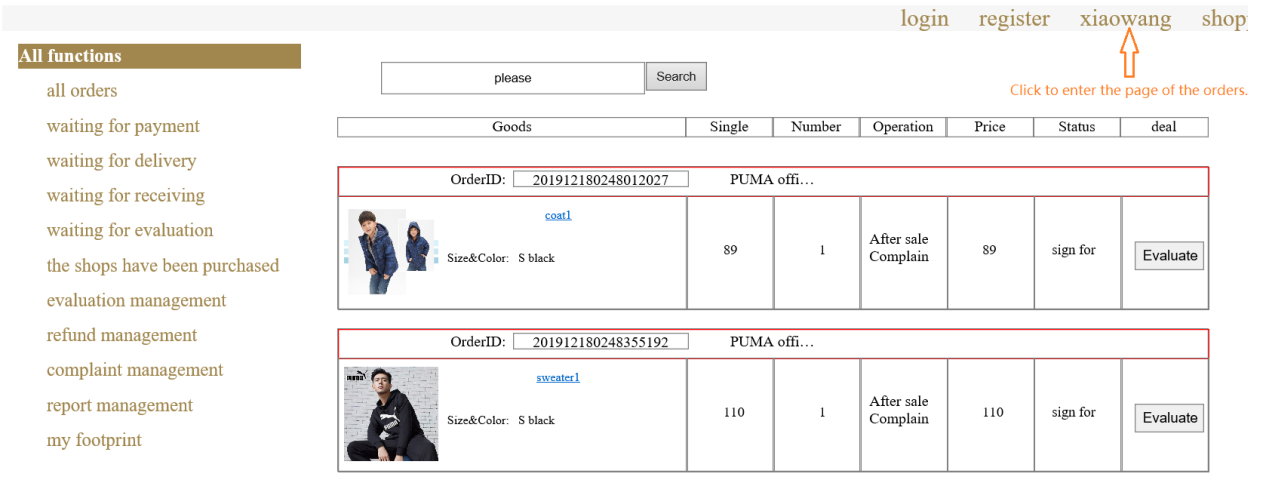
1. Enter the product interface to browse the product



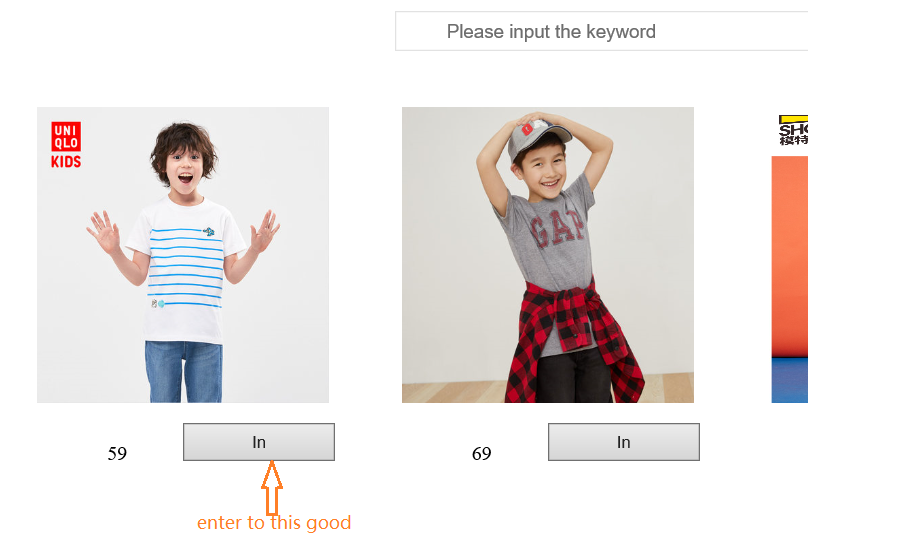
1. Go to the shopping cart to view the items that have been added to the shopping cart



1. View purchased items (unreceived and received)



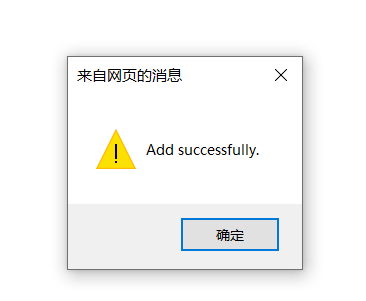
1. Enter the product page from the browse product page



(12) Click add to cart



(13) Add to cart successfully, click ok to return



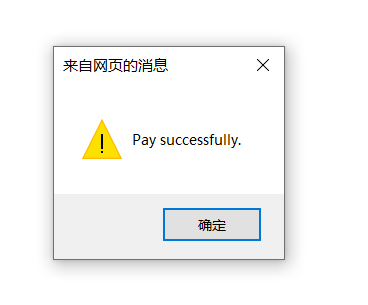
1. Click to pay



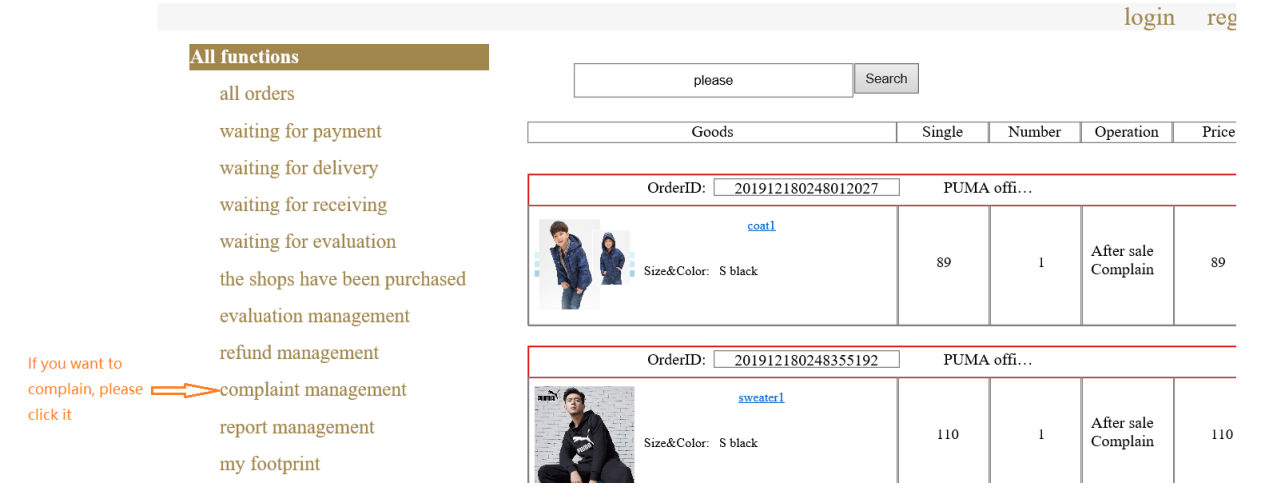
1. Pay for



(16) Purchase successful, click "ok" to return

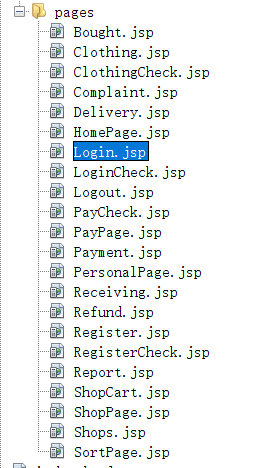


(17) If you want to complain, please click the button in the page



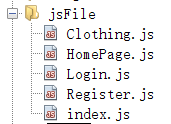
## 5. Description of the code directory structure of the Head-end

(1) The directory used to display the page



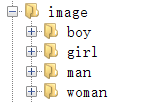
The above is the directory used to display the page, where the Bought. JSP is the order interface, which displays the information of each order Bought by the user in detail.Clothing. JSP is the product interface that displays the details of an item;Clothingcheck.jsp carried on clothing.jsp, logging the actions of the user on the item, such as adding to the shopping cart, buying directly, etc.;Complaint.jsp displays the user's historical comments or Suggestions about a product or the site;The delivery.jsp displays the user's order status;HomePage. JSP as the user's HomePage;Login.jsp is the Login interface;Logincheck.jsp for login.jsp to undertake, through the log of the user Login status, including the user Login success, Login failure details;PayCheck. JSP is the payment confirmation interface, which logs the information of the user's payment success and payment failure.Paypage.jsp is the output interface of the payment password, which displays the necessary information such as the amount of payment;The payment.jsp displays the status of each of our expenditures;Receiving. JSP is the logistics information page, showing the delivery progress of the goods for the user;Refoundation.jsp is the product return interface, showing the user the detailed history of the product return and exchange.Register.jsp for the registration interface, the user must enter the user name, password and other necessary information can be logged in successfully;Registercheck.jsp for register.jsp to do the following, through the log record the user registration success of the information;Report.jsp is the user feedback interface, showing each feedback of the user to this website;Shopcart.jsp for the shopping cart interface, detailed display of the user added to the shopping cart information;Shoppage.jsp for the store page, showing all the store's treasures;Sortpage.jsp for browsing the product page, through the log record users for each product browsing.

(2) The directory used to decorate the display page with events



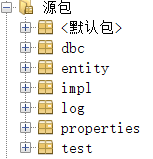
Above is the catalog used to decorate the display page with events, where clothes.js is the collection of click events for clothes.jsp, including the increase or decrease of the number of items, the selection of size and color, and the addition to the shopping cart and direct purchase operations;HomePage. Js is the collection of click events for HomePage. JSP;Login.js is the collection of click events for login. JSP, which includes immediate registration, forgetting the password and successfully logging into the next page;Register.js is the collection of click events of register.jsp, including the blank of user name and password, wrong mobile phone number, and successfully jump to the login page;Index.js is the set of the operation of the first page picture rotation.

(3) The catalog used to store the pictures of the products



The above figure is a collection of commodity pictures, all concentrated under the image package.

(4) The directory used to beautify the display page



The above figure shows the directory used to connect to the database and call the collection of class packages displayed on the website of the database data.Entity is used for adding, deleting, modifying and checking the data of a single commodity, order, shopping cart and store.The impl package is used to search for the overall merchandise, order, shopping cart, store, and input the overall data;The log package is used to hold log file information;The properties package is used to hold information about database users, passwords, and so on.The test package contains the code used to test the site.

## Analysis instructions

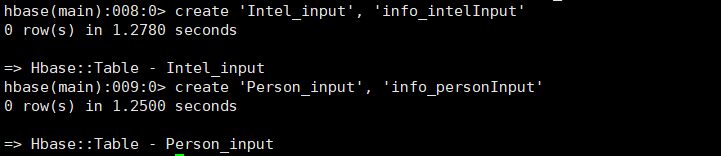
We analyze the data in the following three steps.

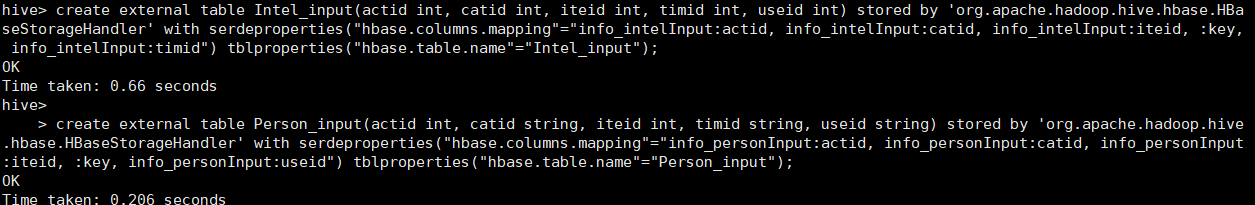
* 1. Create tables

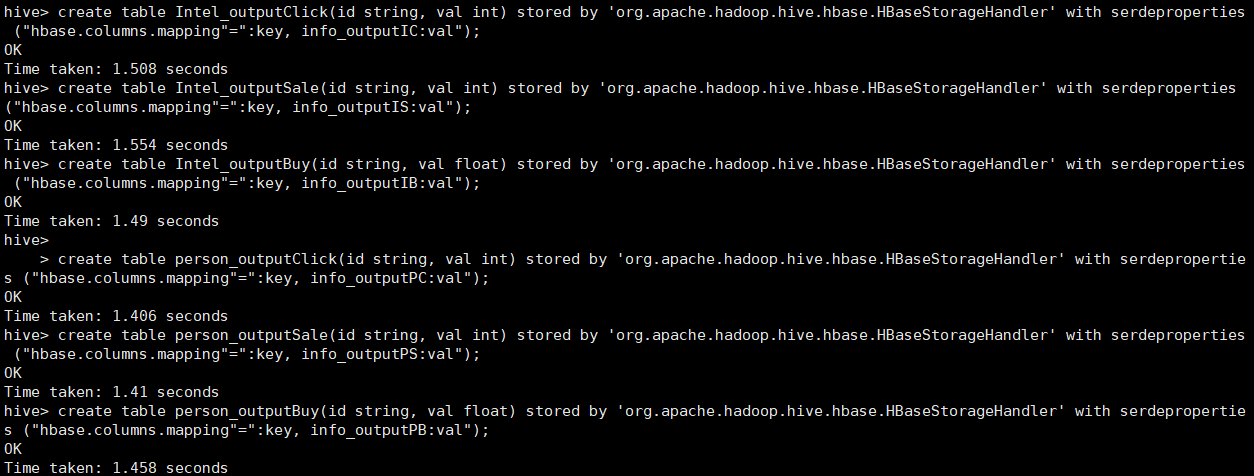
First we created two external hBase-based tables on Hive so that we could analyze the data directly on Hive after importing it to HBase.We then created six internal tables of Hive and connected them to HBase for subsequent export of the analysis results.

We used NetBeans to import our generated log files into HBase, and Hive statements to import the web log files directly into HBase.

The results are shown in the figure below:



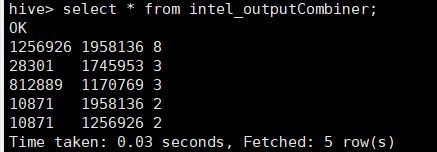
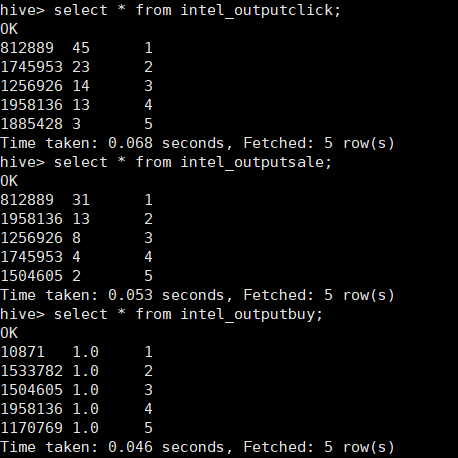


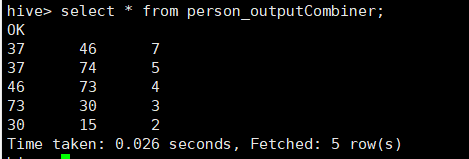
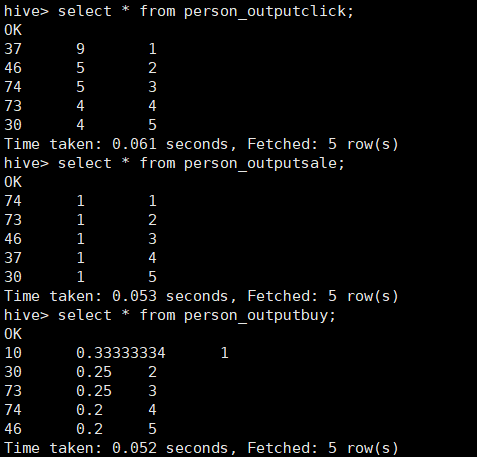


6.2 Data Analysis

We use Hive statements to analyze the data and import the results directly into the HBase table.

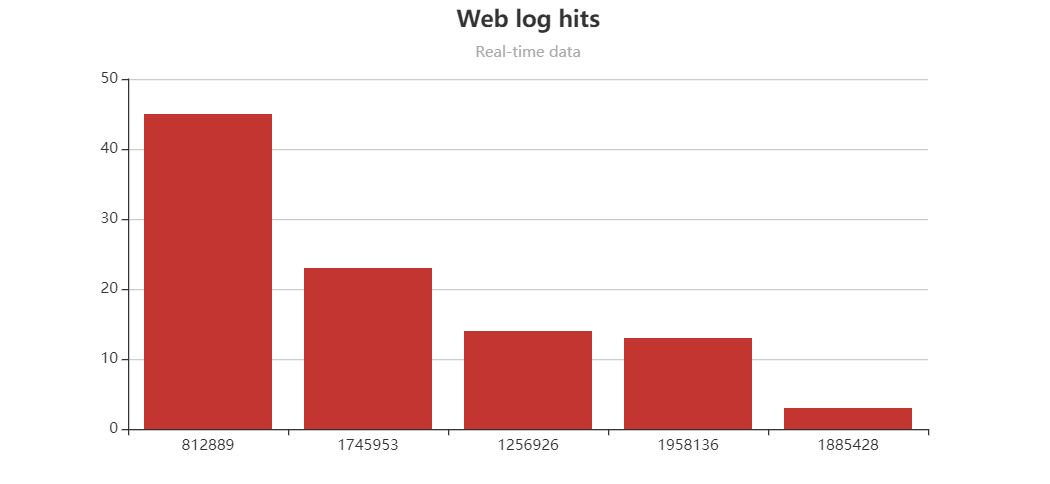
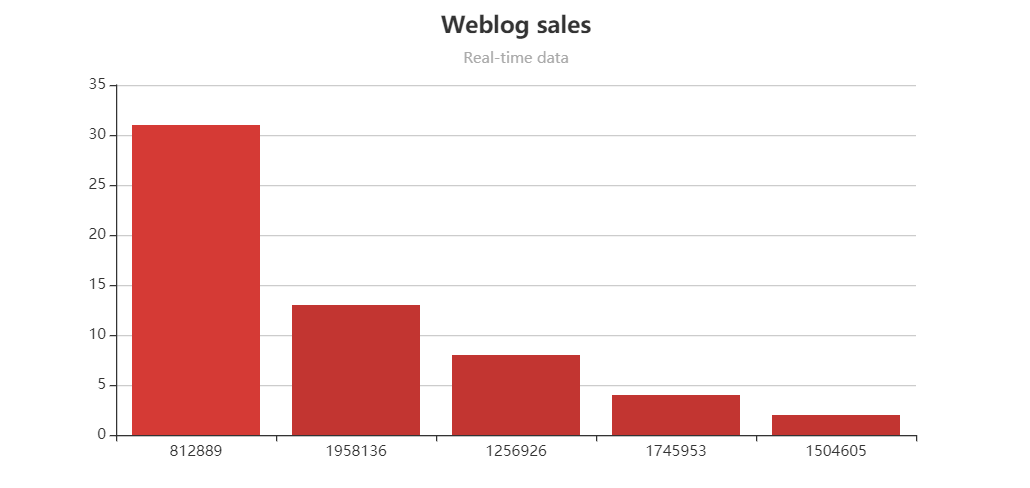
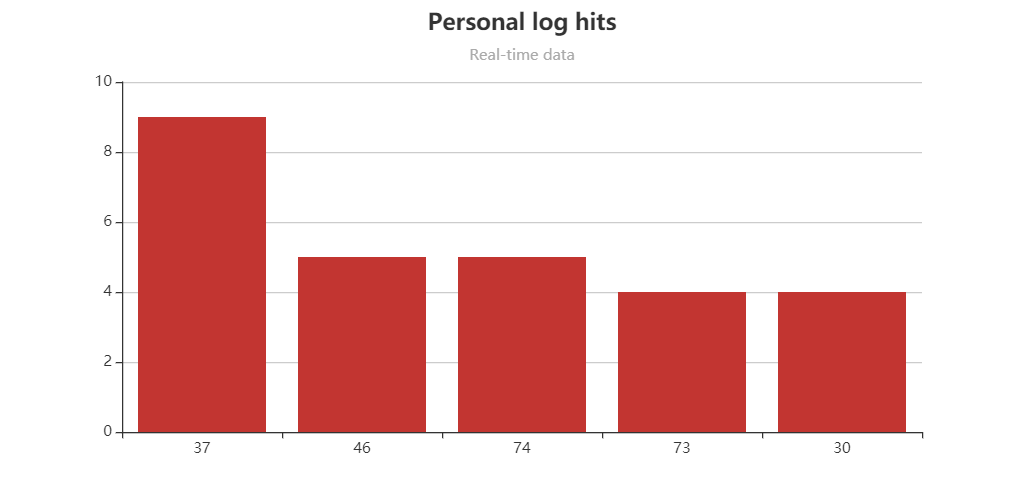
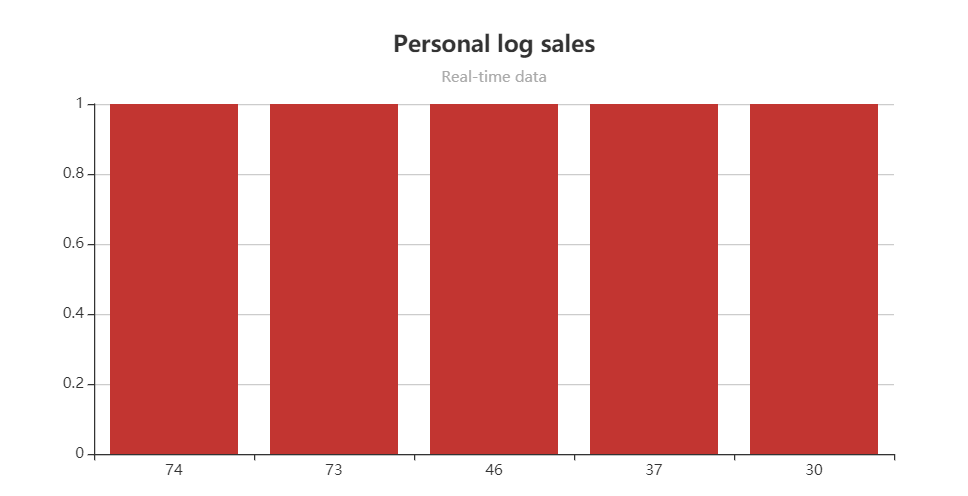
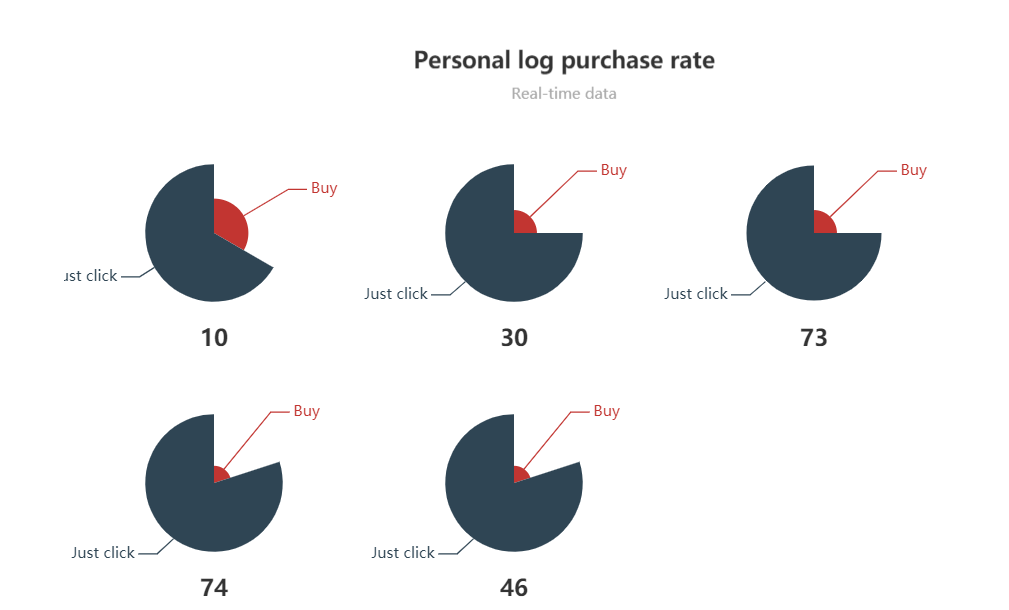
Some results are shown in the figure below():

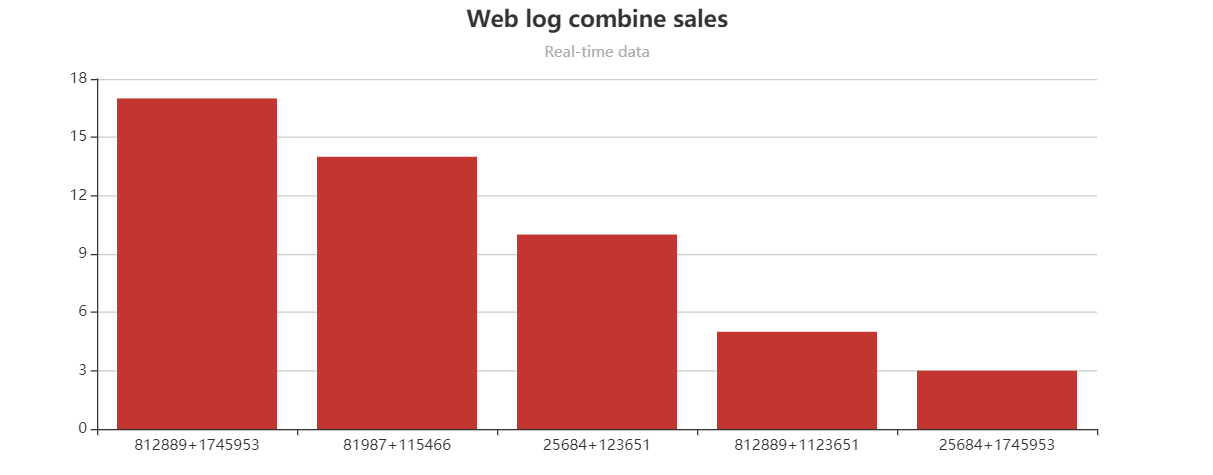


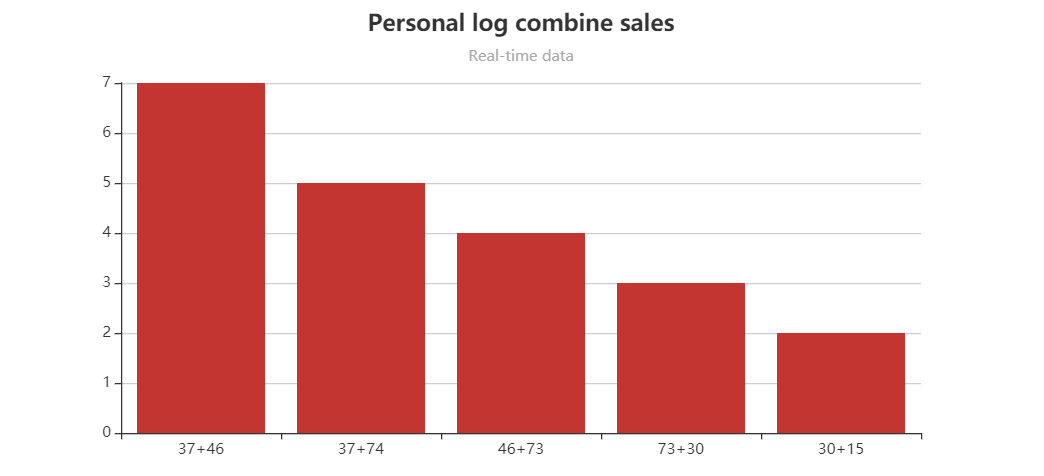


6.3 Obtain data from the front end

We use NetBeans to connect to HBase, turn the result into JSON data, take it out and put it into a FRONT-END JSP file as a parameter to construct the dynamic diagram. The final result is shown in the figure below:

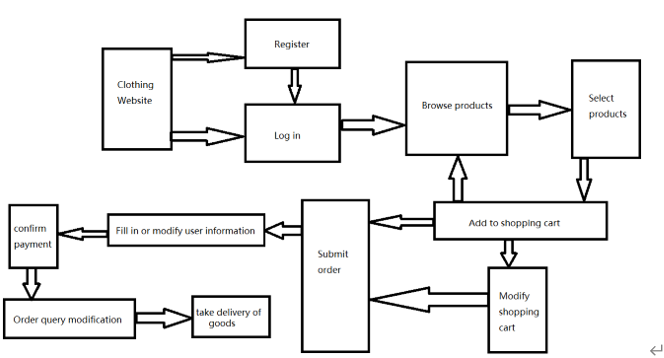
     



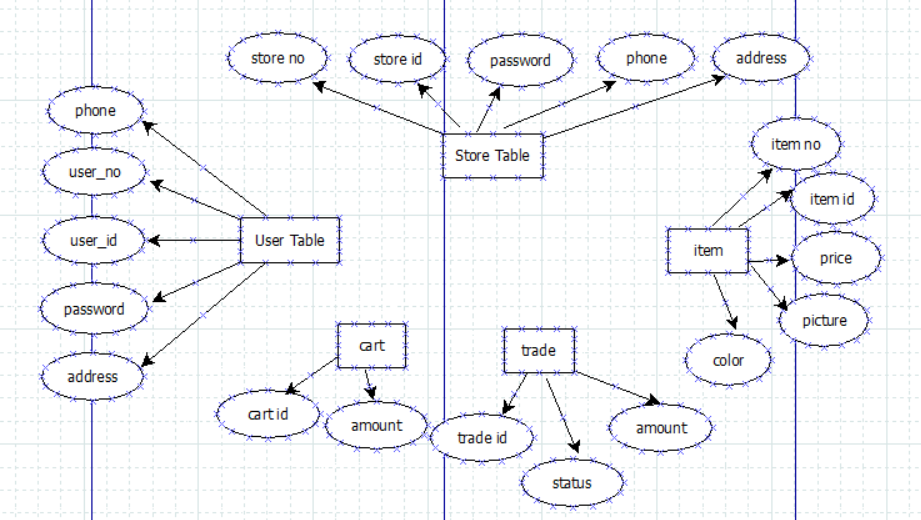


# Flow Chart

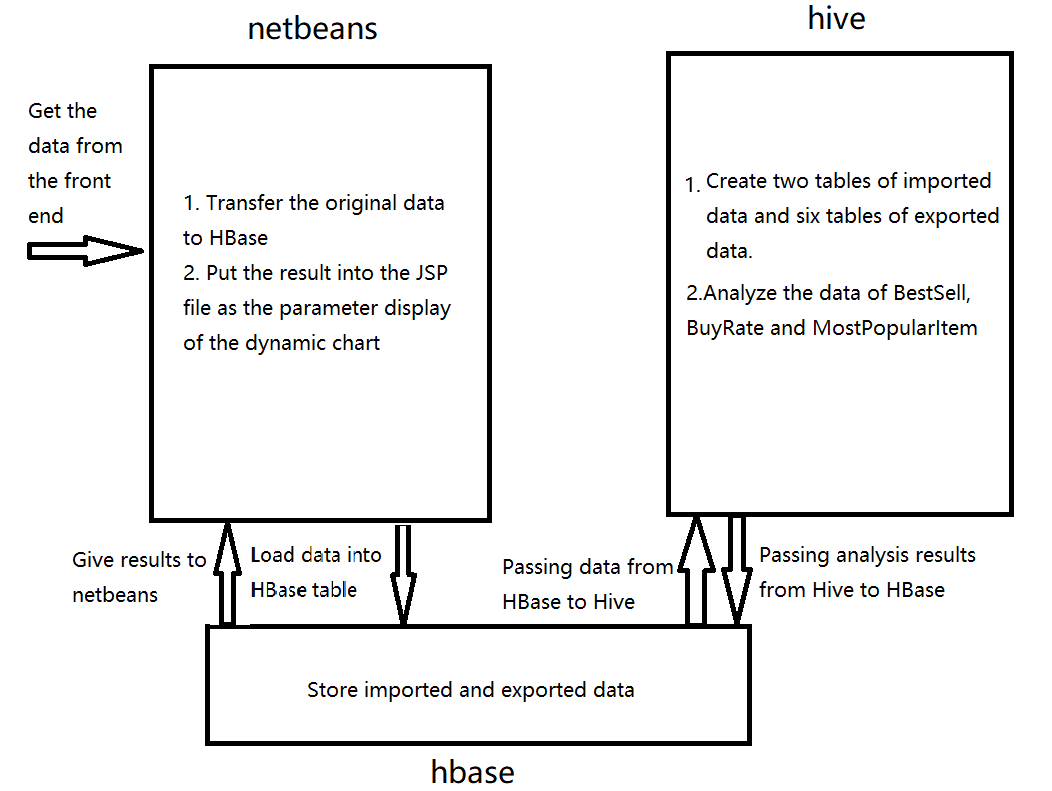
## Client function



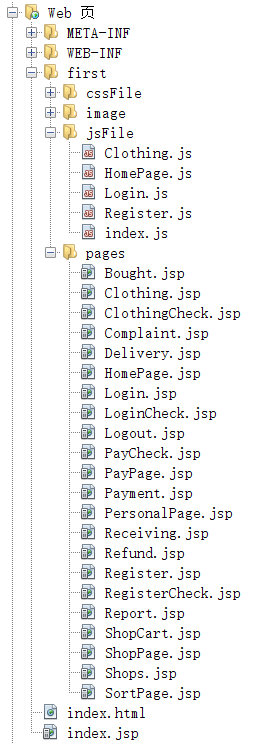
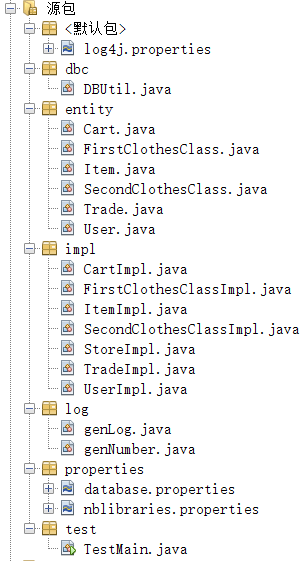
## ER diagram for database Design



## Data analysis flow chart



# Class Diagram



# Individual Milestone Schedule

|  |  |  |  |
| --- | --- | --- | --- |
| **Task Name** | **Start Time** | **Finish Time** | **Developer Name** |
| SRS | 2019/11/02 | 2019/11/04 | TengChang、WangNan |
| Find web logs | 2019/11/06 | 2019/11/08 | YanBingjie |
| Web Design | 2019/11/06 | 2019/11/25 | TengChang、WangNan |
| DataBase | 2019/11/10 | 2019/11/25 | NiuCong |
| J-Mysql-DBC | 2019/11/12 | 2019/11/25 | YanBingjie |
| Hadoop Log Analysis | 2019/12/01 | 2019/12/17 | YanBingjie |
| Project Report | 2019/12/17 | 2019/12/18 | TengChang、WangNan |
| PPT | 2019/12/18 | 2019/12/18 | NiuCong |
| J-HBase-DBC | 2020/06/20 | 2020/06/22 | YanBingjie |
| Hive Log Analysis | 2020/06/15 | 2020/06/22 | TengChang WangNan |
| Dynamic graph | 2020/06/23 | 2020/06/23 | NiuCong |
| Project Report | 2020/06/23 | 2020/06/24 | WangNan |
| PPT | 2020/06/24 | 2020/06/25 | TengChang |

# Important Code

/\*Log file\*/

public static void Log(String user\_id, String ip\_address, String action\_id, String category\_id1, String category\_id2,

String item\_id, String seller\_id) { // timestamp, user\_id, ip\_address, action\_id, first\_category\_id, second\_category\_id, item\_id, seller\_id

Logger log = Logger.getLogger(genLog.class.getName());

SimpleDateFormat df = new SimpleDateFormat("yyyy-MM-dd HH:mm:ss");

String time = df.format(new Date());

log.info(time + "," + user\_id + "," + ip\_address + "," + action\_id + "," + category\_id1 + "," + category\_id2 + "," + item\_id + "," + seller\_id);

}

/\*User operation file\*/

public class UserImpl {

public static int add(User user) throws Exception {

String sql = "insert into user(user\_no,user\_id,user\_passwd,user\_phone, user\_addr) values(?,?,?,?,?)";

if(user.getUser\_no() == null || user.getUser\_no().length() == 0) {

user.setUser\_no(genNumber.getno());

}

try (Connection conn = getConnection();) {

return exceuteUpdate(conn, sql, new Object[]{user.getUser\_no(), user.getUser\_id(), user.getUser\_passwd(), user.getUser\_phone(), user.getUser\_addr()});

}

}

public static User findIdUser(String name) throws Exception {

String sql = "select \* from user where user\_id=?";

User use = null;

try (Connection conn = getConnection();

PreparedStatement pst = conn.prepareStatement(sql)) {

pst.setString(1, name);

try (ResultSet rs = pst.executeQuery()) {

if (rs.next()) {

use = new User(rs.getString("user\_no"), rs.getString("user\_id"), rs.getString("user\_passwd"), rs.getString("user\_phone"), rs.getString("user\_addr"));

}

return use;

}

}

}

public static User findNoUser(String name) throws Exception {

String sql = "select \* from user where user\_no=?";

User use = null;

try (Connection conn = getConnection();

PreparedStatement pst = conn.prepareStatement(sql)) {

pst.setString(1, name);

try (ResultSet rs = pst.executeQuery()) {

if (rs.next()) {

use = new User(rs.getString("user\_no"), rs.getString("user\_id"), rs.getString("user\_passwd"), rs.getString("user\_phone"), rs.getString("user\_addr"));

}

return use;

}

}

}

}

/\*PayCheck file\*/

<body>

<%

String pass = request.getParameter("pass");

User u = (User) session.getAttribute("user");

genLog log = new genLog();

// genIp ip = new genIp();

String item\_id = (String)session.getAttribute("item\_id");

String store = StoreImpl.findById(item\_id);

Cookie[] co = request.getCookies();

if (!u.getUser\_passwd().equals(pass)) {

log.Log(u.getUser\_id(), InetAddress.getLocalHost().getHostAddress(), "08", co[0].getName(), co[1].getName(), item\_id, store);

out.println("<script>alert(\"Password is incorrect.\")</script>");

out.println("<script> window.history.back(-1); </script>");

} else {

log.Log(u.getUser\_id(), InetAddress.getLocalHost().getHostAddress(), "09", co[0].getName(), co[1].getName(), item\_id, store);

Item item = ItemImpl.findIdItem(item\_id);

String item\_no = item.getItem\_no();

List<Cart> c = CartImpl.findCartUser(u.getUser\_no()); //通过item\_no找到该cart

int i = 0;

while (i < c.size()) {

if (c.get(i).getItem\_no().equals(item\_no)) {

break;

} else {

i++;

}

}

Trade trade = new Trade(null, c.get(i).getItem\_amount(), item\_no, c.get(i).getUser\_no(), c.get(i).getItem\_size(), c.get(i).getItem\_color());

int t = TradeImpl.add(trade);

if(c.get(i).getCart\_no()!=null){

CartImpl.deleteCartId(c.get(i).getCart\_no());

}

out.println("<script>alert(\"Pay successfully.\")</script>");

out.println("<script> window.location.href=\"HomePage.jsp\"; </script>");

}

%>

</body>

/\*connect database\*/

public static void init() throws IOException {

System.out.println("initing");

conf = HBaseConfiguration.create();

conf.set("hbase.rootdir", "hdfs://192.168.233.128:9000/hbase");

conf.set("hbase.zookeeper.quorum", "192.168.233.128:2181");

connection = ConnectionFactory.createConnection(conf);

System.out.println("initsuccess");

}

/\*Update database\*/

public static int exceuteUpdate(Connection conn, String preparedSql, Object[] param) throws SQLException {

int num;

try (Connection connection = conn;

PreparedStatement pstmt = connection.prepareStatement(preparedSql);) {

if (param != null) {

for (int i = 0; i < param.length; i++) {

pstmt.setObject(i + 1, param[i]);

}

}

num = pstmt.executeUpdate();

return num;

}

}

/\*BestSellMapReduce\*/

1. on HBase: create 'Intel\_input', 'info\_intelInput'

on Hive: create external table Intel\_input(actid int, catid int, iteid int, timid int, useid int)

stored by 'org.apache.hadoop.hive.hbase.HBaseStorageHandler' with serdeproperties

("hbase.columns.mapping"="info\_intelInput:actid, info\_intelInput:catid, info\_intelInput:iteid, :key, info\_intelInput:timid")

tblproperties("hbase.table.name"="Intel\_input");

2. on HBase: create 'Person\_input', 'info\_personInput'

on Hive: create external table Person\_input(actid int, catid string, iteid int, timid string, useid string)

stored by 'org.apache.hadoop.hive.hbase.HBaseStorageHandler' with serdeproperties

("hbase.columns.mapping"="info\_personInput:actid, info\_personInput:catid, info\_personInput:iteid, :key, info\_personInput:useid")

tblproperties("hbase.table.name"="Person\_input");

/\*BuyRateMapReduce\*/

create table Intel\_outputSale(id string, val int) stored by 'org.apache.hadoop.hive.hbase.HBaseStorageHandler' with serdeproperties ("hbase.columns.mapping"=":key, info\_outputIS:val");

insert into table Intel\_outputSale select t.item,count(\*) from (select explode(split(iteid,'\n')) as item from Intel\_input where actid=1) as t group by t.item;

/\*MostPopularItem\*/

create table person\_outputClick(id string, val int) stored by 'org.apache.hadoop.hive.hbase.HBaseStorageHandler' with serdeproperties ("hbase.columns.mapping"=":key, info\_outputPC:val");

insert into table Intel\_outputClick select t.item,count(\*) from (select explode(split(iteid,'\n')) as item from Intel\_input) as t group by t.item;

/\*getAnalRes\*/

public static List<Result> getRecords(String name) throws IOException{

HBUtil.init();

List<Result> res = new ArrayList<>();

Table hTable = HBUtil.getTable(name);

Scan scan = new Scan();

// Scan scan = new Scan("1".getBytes(), "4".getBytes()); //从1到4

scan.setCacheBlocks(true);

scan.setCaching(1000);

try(ResultScanner scanner = hTable.getScanner(scan)){

for(Result r: scanner){

res.add(r);

}

}

return res;

}

public static List<Product> exportList(List<String> list) throws IOException{

List<Product> res = new ArrayList<Product>();

for(int k=0; k<list.size(); k++){

String tableName = list.get(k);

//取数据

List<Result> ress = getRecords(tableName);

if(ress.isEmpty()){

System.out.println(tableName+" has no data");

continue;

}

System.out.println("---------------------------------------------");

for(int i=0; i<5; i++){

Result r = ress.get(i);

String rowkey = new String(r.getRow());

List<Cell> cellList = r.listCells();

for(int j=0; j<cellList.size(); j++) {

System.out.println(new String(cellList.get(j).getValue()));

}

// String name = new String(cellList.get(0).getValue());

float num = new Float(new String(cellList.get(0).getValue()));

res.add(new Product(rowkey, num));

}

}

return res;

}

/\*getHBaselog\*/

public static void put(String key, String actid, String catid, String iteid, String timid, String useid) {

try {

HBUtil.init();

} catch (IOException ex) {

Logger.getLogger(TextHBase.class.getName()).log(Level.SEVERE, null, ex);

}

String tableName = "Person\_input";

try {

Put put = HBUtil.createPut(key);

// System.out.println("get put object");

HBUtil.addValueOnPut(put, "info\_personInput", "actid", actid);

HBUtil.addValueOnPut(put, "info\_personInput", "catid", catid);

HBUtil.addValueOnPut(put, "info\_personInput", "iteid", iteid);

HBUtil.addValueOnPut(put, "info\_personInput", "timid", timid);

HBUtil.addValueOnPut(put, "info\_personInput", "useid", useid);

// System.out.println("???????????????????????????");

HBUtil.put(tableName, put);

// System.out.println("put successfully");

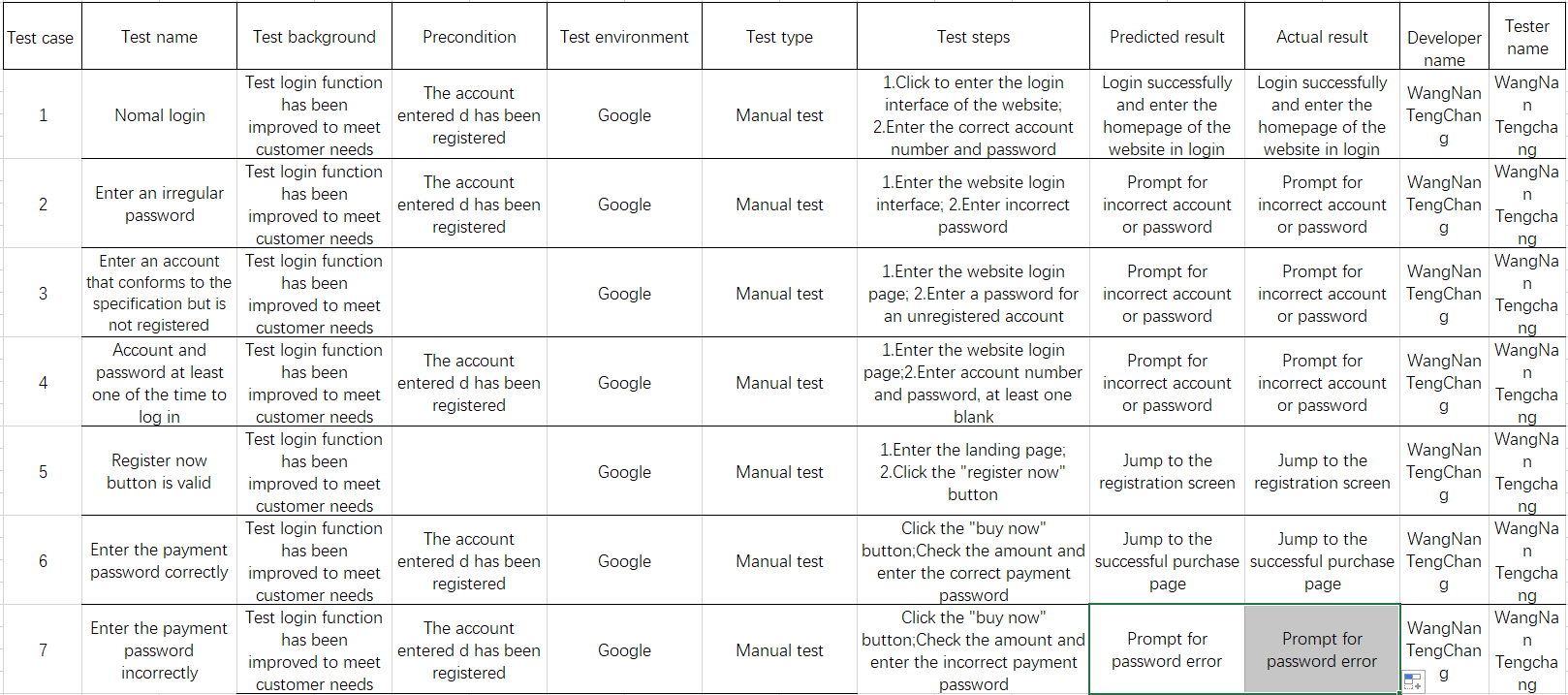
} catch (Exception e) {

e.printStackTrace();

}

}

# Test

****

# Potential Improvements

## 1. The security

In the current system, even if the user USES the appropriate user name and password to enter the bank Intranet site, it is not completely secure, we can use session to prevent this problem.

## 2. Data validation

In the current system, all the number and date fields of the added page have data validation, but the update page does not.If the user makes an error during the update process, you should add data validation to all update pages.

## 3. Limited historical memory

Users sometimes log off their accounts by mistake. However, the accounts that have been logged off by mistake contain some data that the users have not been able to move out. At this time, there should be a period of logout.

## 4. Page index

Page index function if the table view page has a page index function, the behavior of searching from the drop-down menu would be inappropriate.For example, when the user views all the products and goes to the second page, the user cannot select any other options from the drop-down box.With this problem, the page index function is removed from the current system.