

Math 113 - Grocery Store Project

This is a semester long project involving many of the aspects of Statistics. It can also serve as an example for the project that students will have to perform.

Group Composition

The class will be divided into groups of approximately 4 students each. Each group will be assigned one grocery store. If this project is ran in conjunction with the other statistics class, it may be necessary to add additional stores.

Communication

Since this project will require out-side of class communication between 4-6 people and it can be difficult to get 4-6 people together at a common time, a conferencing system has been set up using [WebBoard](http://webboard.richland.cc.il.us/~james/). The system is available over the Internet at <http://webboard.richland.cc.il.us/~james/>

WebBoard allows students / faculty to communicate by leaving and responding to messages. There is no requirement that people be online at the same time to use the system. You can leave a message and the other person can respond when they get the chance.

WebBoard can also be configured to send the messages to your email account so that you do not have to log into the system to check messages. Responses can also be sent by email. It also allows news group reading.

WebBoard also allows interactive chat if more than one user is online at the same time. WebBoard is also available from any of the computer labs on campus, including the one in S137.

Steering Committee

One member of each group will serve on a steering committee that serves to make decisions affecting the entire class.

When the steering committee has a task to perform and the group has a task to perform, the members of the steering committee are exempt from performing their groups task.

The members of the steering committee will need to have Internet access so that they can communicate with each other using WebBoard.

While actively involved in the project, members of the steering committee should check WebBoard at least every other day. An alternative would be to configure WebBoard to send emails when a message is posted and then check your email.

Assignments

Periodically, information will be turned in for a grade. All assignments submitted for a grade should have the name of the group and the name of all the students in the group at the top of the document. The name of the groups will be "The Cub Foods Group", "The Eagles Group", etc., and "The Steering Committee".

All information submitted to the instructor is to be printed.

In addition, certain assignments will be posted to WebBoard so that the results are available to the class.

Parts of the project will be due at different times. Due dates and point values will be posted.

Phase 1 - Organization into Groups

This will take place very early in the semester. Either the first or second class period.

Each group should elect one person to serve as a liaison officer that is a member of a steering committee to coordinate with other groups. The liaison officer should have Internet access.

The assignment for this phase is log into WebBoard and post a message introducing yourself to the rest of the class in the "Math 113 - Introductions" conference. Your message should contain your name and why you're taking this class. Also include how many hours you're taking this semester and how many hours you work a week. Give the names of the other people in your group. Be sure to identify your group's liaison officer to the steering committee.

Phase 2 - Planning

Decide upon categories of food. As a starting point, here are the food groups recognized by the US Department of Agriculture and the Supermarket Business magazine.

USDA Food Groups (1)

1. Dairy and Egg Products
2. Spices and Herbs
3. Baby Foods
4. Fats and Oils
5. Poultry Products
6. Soups, Sauces, and Gravies
7. Sausages and Luncheon Meats
8. Breakfast Cereals
9. Fruits and Fruit Juices
10. Pork Products
11. Vegetables and Vegetable Products
12. Nut and Seed Products
13. Beef Products
14. Beverages
15. Finfish and Shellfish Products
16. Legumes and Legume Products
17. Lamb, Veal, and Game Products
18. Baked Products
19. Snacks and Sweets
20. Cereal Grains and Pasta
21. Fast Foods
22. Meals, Entrees, and Sidedishes

Supermarket Business Groups (2)

1. Dry Grocery (Food)
2. Dry Grocery (Non Food)
3. GM/HBC (includes prescriptions)
4. Perishables
5. Meat/Fish/Poultry
6. Service Deli
7. Produce
8. Baked Goods
9. In-Store Bakery
10. Dairy
11. Frozen Foods

You will probably want to eliminate some categories and add other, non-food, categories that are common to all stores.

Decide which foods you want to include in each group. Each category that your groups decides upon must have at least five (5) items. At this stage in the process, list brands / items, but not sizes. Here's an example:

Breakfast Cereals

Lucky Charms
Froot Loops
CrunchBerries
Post Raisin Bran
Corn Chex

Each group should get together - or use WebBoard to create their list. It should be typed and posted on WebBoard for the steering committee to use. The liaison should not be the person to type up the list from each group (unless they really want to).

The steering committee will combine the lists into one master list. The master list should contain at least eight (8) but no more than twelve (12) items for each food category.

Each category should be assigned a number. The numbers should be multiples of 100. Each item should be assigned a number. The item number should be the category number plus a count. For example, if Breakfast Cereals is the first group, then call it group 100 and then number the cereals as 101, 102, 103, 104, etc..

The master list will be provided to each group for data collection.

Phase 3 - Data Collection

Once the master list of items is compiled and distributed, the stores will be visited and data collected. One problem is that there are no sizes for the items and most items have several sizes.

There are two ways of solving that problem.

1. Visit the stores, record the available sizes for all of the items, come back together, agree on sizes, go back to the stores and collect prices for those sizes.
2. Visit the stores, record the sizes and prices for all available sizes, come back together, agree on the sizes, and discard the extra information.

I think that most reasonable people will agree that the second option is the preferred one since they only have to go the store once to collect data.

The amount of data collection can be reduced by agreeing to the following guidelines ahead of time.

1. If there is only one size, use it.
2. If there are two sizes, choose the (smaller / larger) size.
3. If there are three sizes, choose the (smallest / middle / largest) size.
4. If there are four or more sizes, choose the _____ size.

A list of items will be prepared and distributed. When collecting the data, record the size, regular price, and, if applicable, sale price for each item.

Stores normally hold their prices for a week, from Sunday to Saturday. In order to ensure a fair comparison, it is important that all stores be visited during the same week.

Sample Data Collection Sheet

Store : _____			
Date : _____			
Collected By : _____			
Breakfast Cereals	Size	Regular Price	Sale Price
101	Kellogg's Rice Crispies		
102	Post Raisin Bran		
103	Lucky Charm's		
104	CrunchBerries		
105	Fruit Loops		

The steering committee is responsible for the design of the structure of the SPSS file. They should come up with variable names, labels, and value labels so that the data is consistent throughout all five stores. The steering committee is not responsible for the data entry, only the design.

Each group should have an SPSS file for your store. The individual SPSS files will be combined into one master file for use.

If you are unable to collect prices for an item because the store does not carry it, please report it to your liaison officer as soon as possible so that they can consult the steering committee. The data will either have to be deleted or a different size chosen. This will be the decision of the steering committee. It may involve another trip to the store to collect more data.

Each category must have at least five (5) items in it for the final report. Since prices were collected for at least eight (8) items, the steering committee has three (3) items that they can eliminate per category before sending people back for more information.

Once all of the data is combined, compare data to make sure the product sizes are all the same. If they are not the same, bring it to the attention of the steering committee so that they take the appropriate action.

Phase 4 - Descriptive Statistics

Each group is to describe the data for their store (apply a filter to the master file). The steering committee is to describe the combined data for all of the stores.

Describe the following for both the regular price and the sale price.

- Five number summary for each category.
- Simple Boxplots for each category.
- Histogram for all of the data.
- Mean, standard deviation, and total for each category as well as the entire dataset for your store.

The output from SPSS should be cleaned up and annotated. After the entire group has checked the results, post the output file (.SPO extension) to WebBoard.

Phase 5 - Inferential Statistics

The number in parentheses after each question is the section number the material covers.

Q1: Does the mean price of your store differ significantly from the mean of all the stores? (7.4)

In the data description phase, the steering committee found the mean price for all of the data. Compare the mean for your store (set a filter) to that single value. Give results for both regular and sale price. Give the results for both the categories and the combined store data.

Example. If the mean regular price for all of the data is \$1.24, then compare the regular price data for your store to the value 1.24. If the sale price of all items is \$1.22, then compare the sale price data for your store to the value 1.22.

The steering committee should work with their individual groups for this part.

Q2: Is the sale price for your store significantly cheaper than the regular price? (8.2)

Set a filter for your store and then compare regular price and sale price. The steering committee should use all available data.

Compare the data for the entire store and for each category (split the file).

Q3: Is the mean price of your store significantly different from each of the other stores? (8.5)

There will be four comparisons here. For example, if you are in the Cub Foods group, then you should make the following comparisons: 1) Cub-Eagle's, 2) Cub-Kroger's, 3) Cub-Schnuck's, and 4) Cub-WalMart.

Give the results for both the regular and sale prices. Compare the data for the entire store and each category (split the file).

The steering committee should work with their individual groups for this question.

Q4: Is there a significant difference in the mean prices of the stores? (11.2)

This will be the same report by all of the groups, so the steering committee should work with their individual groups.

Give the results for both the regular and sale prices. Give the results for each category (split the file) and the overall prices.

If there are significant differences in the mean prices, which stores are different?

Which is the cheapest store? Which is the most expensive store?

Spring 2000 Deadlines

Date	Description
1/26	Login into WebBoard and leave introductory message
1/28	Each groups list of products due to steering committee
2/4	Compiled list of products due to instructor
2/7	Data sheets for data collection handed out to class
2/7-2/12	Collect data from stores
2/18	Data entry due
2/25	Data combined and consistency checks performed
3/10	Descriptive Statistics (Part 4) due
4/7	Inferential Statistics (Part 5), questions 1-3, due
5/5	Inferential Statistics (Part 5), question 4, due

Top 25 US Food Retailers 1999 (By 1998 Grocery Sales) (3)

Rank	Retailer	Headquarters	1998 Sales (Millions)	Stores
1	The Kroger Co.	Cincinnati, OH	\$28,203	2207
2	Safeway Inc.	Pleasanton, CA	\$24,484	1493
3	Ahold USA, Inc.	Atlanta, GA	\$16,200	880
4	Albertson's, Inc.	Boise, ID	\$16,005	985
5	Costco Companies, Inc.	Issaquah, WA	\$15,251	280
6	Fred Meyer, Inc.	Portland, OR	\$14,025	853
7	American Stores Company	Salt Lake City, UT	\$13,756	807
8	Winn-Dixie Stores, Inc.	Jacksonville, FL	\$13,617	1180
9	Wal-Mart Supercenters	Bentonville, AR	\$12,800	564
10	Publix Super Markets, Inc.	Lakeland, FL	\$12,100	586
11	Food Lion, Inc.	Salisbury, NC	\$10,219	1207
12	Great Atlantic & Pacific Tea Co.	Montvale, NJ	\$10,179	913
13	Sam's Club	Bentonville, AR	\$8,267	451
14	H.E. Butt Grocery Co.	San Antonio, TX	\$7,000	258
15	The Stop & Shop Supermarket Co.	Quincy, MA	\$5,500	190
16	SUPERVALU, Inc.	Minneapolis, MN	\$4,984	330
17	Giant Food, Inc.	Washington, DC	\$4,400	173
18	Meijer, Inc.	Grand Rapids, MI	\$4,064	117
19	Pathmark Stores Inc.	Carteret, NJ	\$3,711	133
20	Giant Food Stores, Inc.	Carlisle, PA	\$3,700	164
21	Fleming Companies Inc.	Oklahoma City, OK	\$3,558	381
22	Hannaford Bros. Co.	Portland, ME	\$3,324	150
23	Top's Markets, Inc.	Buffalo, NY	\$3,238	240
24	Hy-Vee, Inc.	West Des Moines, IA	\$2,946	221
25	Bi-Lo, Inc.	Greenville, SC 29607	\$2,900	253

Of the top twenty five supermarkets, only Kroger's, Wal-Mart, and Sam's Club are in the Decatur area. Sam's Club is a warehouse club does not carry the same diversity of products or sizes as the other stores.

References

1. U.S. Department of Agriculture, Agricultural Research Service. 1999. *USDA Nutrient Database for Standard Reference, Release 13*. Nutrient Data Laboratory Home Page, <http://www.nal.usda.gov/fnic/foodcomp>
2. *51st Annual Consumer Expenditures Study*, Supermarket Business. September 1998, pg. 39-40. The data is available online at <http://www.fmi.org/keyfacts/grocerydept.html>
3. *Directory of Supermarket, Grocery & Convenience Store Chains '99*. Chain Store Guide Information Services. This information is available online from the Food Marketing Institute at http://www.fmi.org/food/top_retailers.html.