**EK 125 Final Project: (Minseok\_Sakong.zip)**

**Minseok Sakong \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**Yerin Hur \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**HyoJoo Kim \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**Purpose and Solution:**

In recent years, the constant increase in the manufacturing industry of food and cosmetics is offering consumers a lot more choices to purchase products than before; there are similar but slightly different kinds of products from a variety of brands. Though it is beneficial for the consumers to have a wide spectrum of products to buy, consumers also have high chances of being misled into some of the products that contain hazardous chemicals. Even though it is compulsory for all the companies to indicate the ingredients that are used in their products on the label, consumers easily get confused by its numerous, incomprehensible jargons. Therefore, we assure this application will be useful to the society since consumers can easily find and avoid potentially harmful products based on the diseases or symptoms. By providing a way to filter the proper products they need, we believe this application can offer a guide to be smart consumers in the mass market.

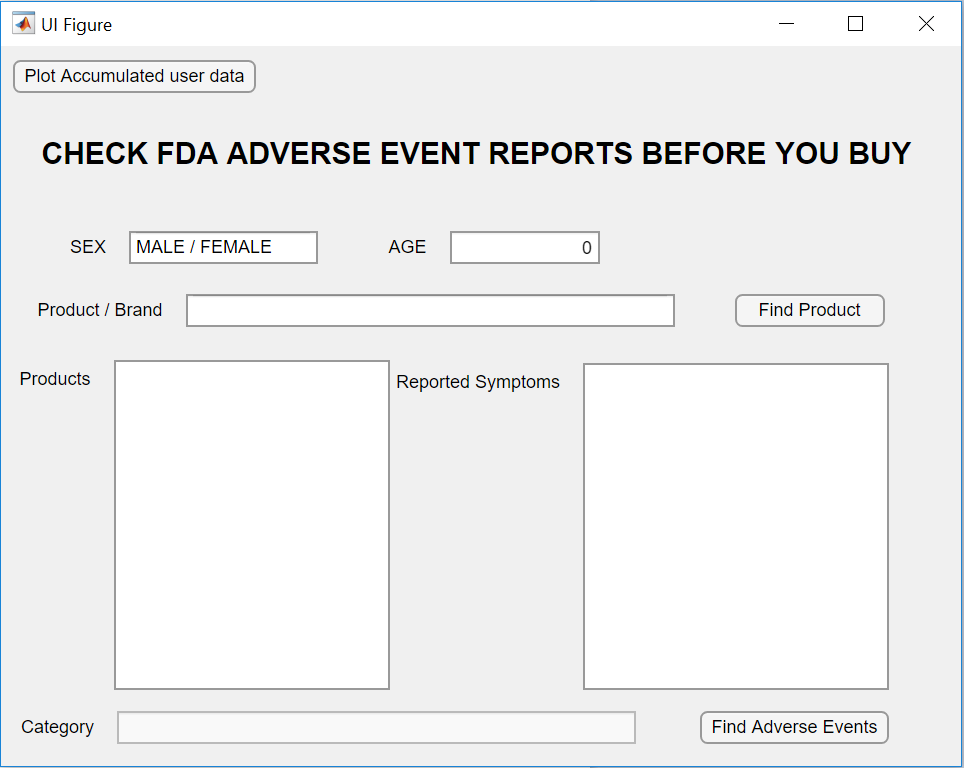
**Project Database:** https://www.fda.gov/Food/ComplianceEnforcement/ucm494015.htm#file

Our project is based on the data set of the adverse event and product complaints reports that were submitted to the FDA, Food and Drug Administration of the United States. We chose this data set to find the brand and its product that triggers the allergic reactions to customers. As consumers search by brand or product names, diseases or symptoms will be displayed accordingly in order to give a warning before their purchase.

**Project-Derived Analyzed Data:**

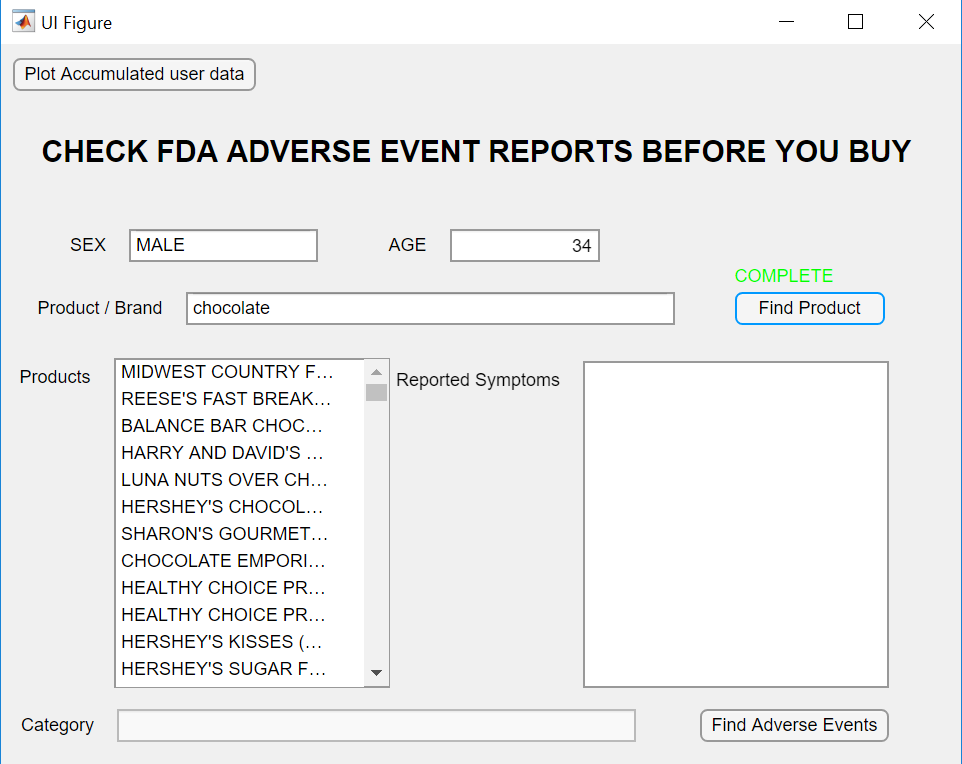
Every time users search products in the app, the app analyzes and saves user-searched products data, categorized by gender and age, in its own database. Since users search products, which they are interested in to purchase, the accumulated analytics can be used as real-world marketing data and statistics resources for trending products in the mass market. The app not only helps consumers to purchase right product, but also provides practical big data sets for various studies and researches.

1. **draft.mlapp**: The front-end(app-designer) of our program. It is name as a draft, but it is not a draft. It is enthusiastically-written and fully-tested final product.
2. **getthedata.m**: The function uses xlsread function in order to **gather all of the data from the excel file to arrays of structures** **(line 3~18).** It saves the data from the back in order to pre-allocate arrays of structures. Also, it returns the arrays of structures as output.
3. **extractbyname.m**: The function extracts some of the data from the whole chunk of big data, which was gathered from “getthedata” function. The extracted data is used as a cache file for searching, which allows the program to save much time and memory.
4. **searchbyname.m**: The function is the main search function of this modular program. It searches 1. product items that has string “name (input variable)” in their product names or brand name, 2. reported symptoms of the item, and 3. the category of the item from the arrays of structures (cache structure).
5. **write2file.m**: The function, “write2file” **tracks what products user searched with our app. It writes user-searched product name, the category of the item, user’s gender, and user’s age to the “data.txt” file (line 4~8).**
6. **saveplotdatam.m**: The function analyzes and **writes numerical user data to the “.dat” file, categorized by age and gender (line 5~46).** The numerical data is used for plotting.
7. **plotuserdata.m**: The function **plots stacked-bar graph with the numerical data, accumulated by the “saveplotdatam” function (line 6~14).** The plot is designed to easily recognize user preferences of products, categorized by age and gender.
8. **category.mat**: This contains string arrays of product categories.
9. **maledat.dat & femaledat.dat**: This contains numerical data of analyzed user data for plotting.
10. **data.txt**: This contains analyzed data of user-searched products.
11. **CAERS-ASCII-2004-2018Q2.csv**: The big data sets from the FDA report portal.

**The starting screen**

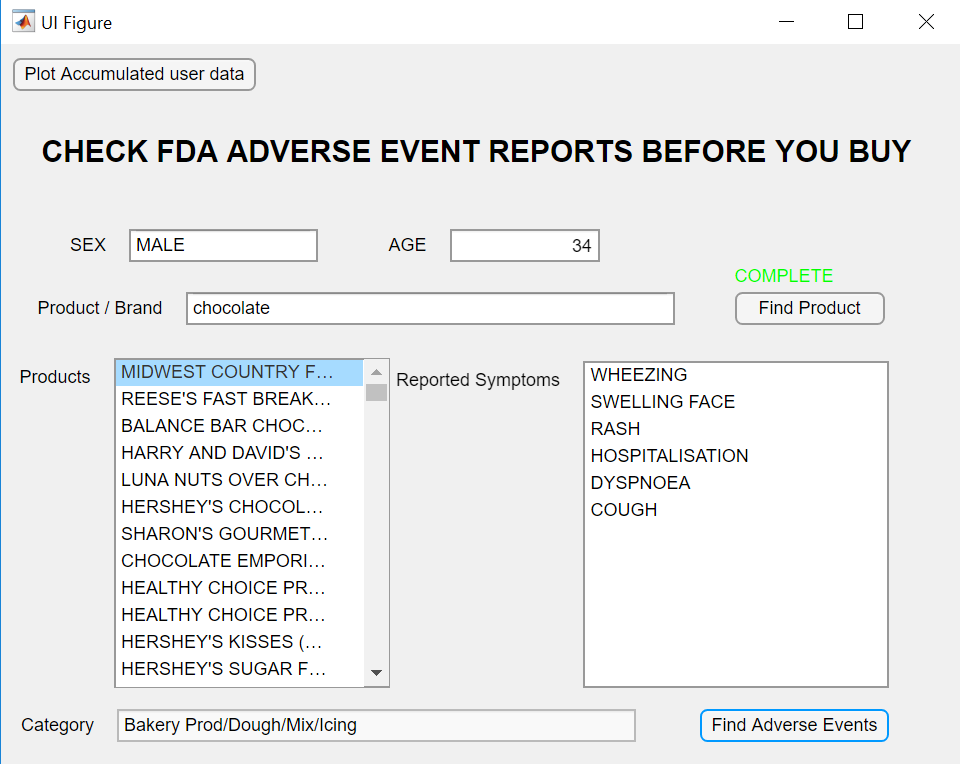
Front-end: Nothing yet

Back-end: All the data from the excel file is already loaded in to arrays of structures (in the global property: productlist)

**User typed-in gender, age, product name, and pressed Find Product Button**

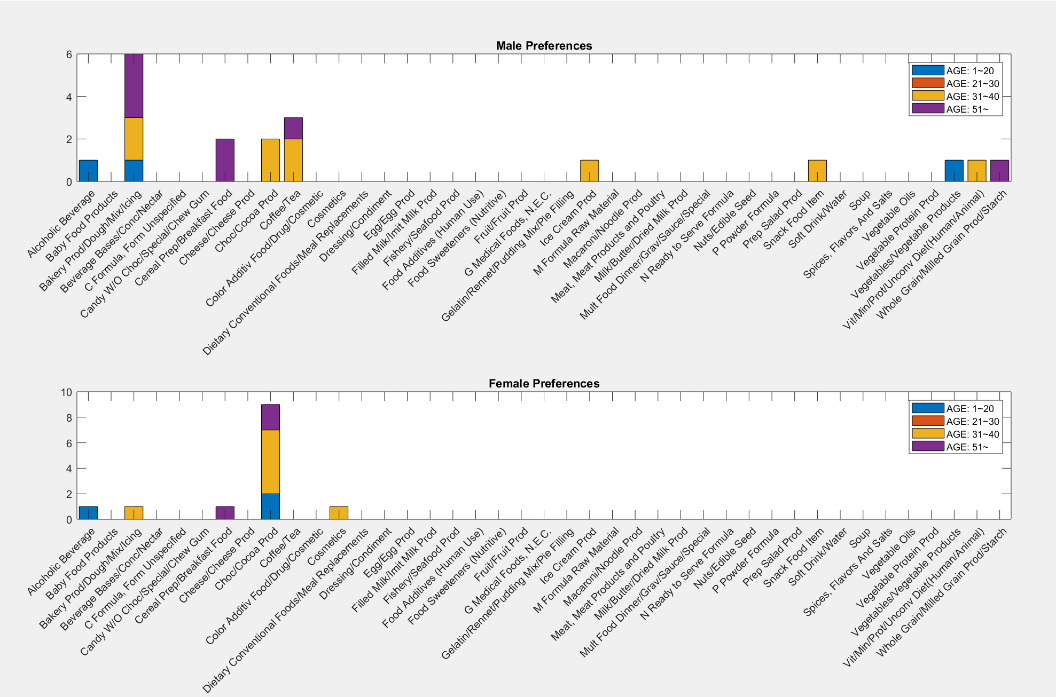
Front-end: Products that have “chocolate” in their names or brands pop out in the products list box

Back-end: When Find Product Button is pressed, structures that contains “chocolate” in their name fields are extracted to cachestruct property and searchbyname function extracts product name data to products list box as arrays of character arrays. (searchbyname function searches relevant data from cachestruct property, not productlist property in order to save time.)

**User selected one product and pressed Find Adverse Events Button**

Front-end: Reported symptoms lists are displayed in organized list and the category of the product is displayed.

Back-end: searchbyname function extracts symptoms and the category of selected product to those list boxes. Also, write2file and saveplotdatam functions are executed to accumulate, analyze, and save user-searched data.

**BU Questrom student pressed the Plot Accumulated user data Button**

Front-end: Stacked bar graphs are displayed. It is well organized by gender and age of users.

Back-end: plotuserdata function is executed in order to plot analyzed user-searched data from “maledat.dat” and “femaledat.dat”. It can be used to analyze product preferences by gender and age.