Baby Names Trends Application

Yasmine Fung

CONTENTS

Letter of Transmittal

Phone: (555)555-5555
Email: yasminef@customtech.com
Website: www.customtech.com

Custom Tech
1973 DuPont St. NE
Pueblo, CO 81749

7/15/2020

Matthew Stewart
Chief Product Officer
Baby Registry Inc.
461 Sunset Ave. S.
Pueblo, CO 810256

Dear Mr. Stewart,

Within this proposal you will find information on the baby name popularity application and our plan of action. The proposal explains in detail the cost, benefits, and methodology for development. A general timeline is given so you know what to expect each week along with a list of all deliverables.

Thank you for reaching out to us at Custom Tech and considering us for your application. If you have any questions or need further assistance, feel free to contact us at any time. We look forward to working with you in the future.

Sincerely,

Yasmine Fung

A. PROJECT RECOMMENDATION (SECTION A)

A.1. PROBLEM SUMMARY

Baby Registry Inc. (BRI) has decided to incorporate a few changes into its company to gain an advantage over its competitors. Baby Registry Inc. has begun producing embroidered products with baby names to provide a personalized touch for its consumers. To cut down on manufacturing costs, BRI would like a program that will show any trends in names to help the company decide which to embroider on goods. Baby Registry Inc. would also like the program to provide an additional service to assist parents with picking a baby name.

A.2. APPLICATION BENEFITS

Creation of a new application would benefit Baby Registry Inc. in manufacturing costs as well as sales. The program will provide a way to analyze baby name popularity among the years as well as view the most searched names among its customers. This would ensure BRI is producing product that will be popular among its consumers.

The new program will also provide a new service for expecting parents to help decide on a baby name. Some parents may want a unique name and others might want one that is more old-fashioned. With this new service, parents will have the ability to look up the popularity of a name they may have in mind or look up the top baby names for a given year.

Both of these benefits will set Baby Registry Inc. apart from other companies in the same field.

A.3. DATA PRODUCTS

The program will consist of a stand-alone program with a user-friendly interface that can be installed on all Baby Registry Inc. computers. The following features will be included:

- Search dataset for name view times used to apply for Social Security card each year
- Search for top 5 boy and girl names for a given year
- View most names most searched by customers

The program will be written in Java and utilize SQL to search the database for necessary information. The database will be stored on a local server that can be remotely accessed through the program. No login will be required as all data is public information.

A.4. DATA DESCRIPTION

The dataset that will be used for this program is provided by a data science company named Kaggle who has gathered data from the Social Security Administration. The data consists of a count of each name used to apply for a Social Security card. Records are provided from 1880 to 2017. Only names given to 5 or more babies are included in the dataset for privacy reasons. (Social Security Number Holders, n.d.)

The dataset can be accessed through the link below:
https://www.kaggle.com/hassenmorad/us-national-baby-names-18802017

A.5. OBJECTIVE AND HYPOTHESES

The objective of this program is to create a program that will allow for easy analysis of baby names by Baby Registry Inc. to increase product sales as well as provide an additional service for parents. The creation and incorporation of this program into BRI will help the company create more products that will be in high demand.

Our hypothesis is that recording the most searched names along with the ability to view popular names for each year will assist the company in deciding on which products to produce and at what volume. With a higher production of products with popular names and a lower production of others, the company will be able to decrease overhead costs for storage.

A.6. METHODOLOGY

The best methodology for this project would be the waterfall method.

With a static scope in mind, the waterfall method will allow for the creation of clear milestones. For the waterfall method to work, the project scope must be static. (Dunleavy, 2019) If there is already a clear goal in sight, all of the milestones are able to be laid out with confidence.

Before moving onto the next phase, the current phase must be completed. This would provide a more organized approach to developing and documenting the process. The waterfall method requires team members to keep detailed documents which will be useful in providing updates on progress and tracking any issues that may arise. Scope, cost, and time are all taken into account when developing a plan of action. With these three constraints, we would come up with a solution without compromising on any of these factors.

A.7. FUNDING REQUIREMENTS

Service	Cost per Hour	Required Hours	Total Cost
Planning and Design	$85	40	$3400
Development	$150	120	$18000
Testing	$100	50	$5000
Review and Analysis	$75	40	$3000
		Total	$29400

A.8. STAKEHOLDER IMPACT

Implementation of this new technology will allow for BRI to make financially sound decisions that will appeal to a greater consumer base. Analysis of trends will allow the company to

better predict product demand. With a better understanding of what consumers want, less merchandise will be taking up storage space. Popular products can also be mass produced to cut down on manufacturing costs.

Additionally, consumers will have the option look up information on the popularity of baby names. This unique service will likely help to attract more business.

A.9. PRECAUTIONS

All data used is public information. Names given to less than 5 babies a year are not included in the dataset for privacy reasons. This data only contains information on names given throughout the United States of America. Any international branches will require a separate data set.

For the most accuracy, data for subsequent years will have to be added to the dataset once released to the public.

Although this program will be custom built for Baby Registry Inc., Custom Tech holds the rights to the application. We do not permit the resale or duplication of our application for unintended uses. If your party is found to be in violation, legal action will be taken.

A.10. EXPERTISE

Our company consists of developers with expertise and experience in many areas. For the best fit for our customers, each team is hand selected for a project based on experience, knowledge, and work ethic. If any additional expertise is required, our company will not hesitate to seek out any additional contract developers for the project.

The main team working on the Baby Registry Inc. application will consist of three of our developers. All of our employees are required to have at the minimum:

- Bachelor's degree in computer science or a related field
- 3 years of programming experience
- Knowledge of different development lifecycle processes

The application team will be led by Yasmine Fung, one of our lead developers. She has experience creating Java applications as well as working with databases.

B. TECHNICAL PROPOSAL (SECTION B)

B.1. PROBLEM STATEMENT

Design and develop a stand-alone application that will allow Baby Registry Inc. employees to analyze baby name trends. Program should be able to display top names for year searched and number of babies given name if name is searched. A record of searches conducted by customers should be kept for analysis purposes.

B.2. CUSTOMER NEED

Baby Registry Inc. has been supplying parents with the newest baby technology since it was founded in 2015. To stay ahead of the competition, BRI wants to stand out from competitors by offering embroidered products. The company would like a program to help with the analysis of baby names to determine which names to embroider on goods. Embroidered baby supplies will appeal to a greater number of consumers as well as cut overhead costs associated with storage and manufacturing. BRI would also like to include a new service available to parents that will provide the most popular names for each year and how many babies were given a certain name.

B.3. EXISTING SYSTEM ANALYSIS

Baby Registry Inc. has an existing system that keeps track of product orders, customer information, and registry lists. BRI would like to create a separate application to help with name trend analysis for both employers and customers.

B.4. DATA REQUIREMENTS

Data used is provided by Kaggle, a subsidiary of Google LLC. The dataset contains information on the names used on applications for a Social Security card. The dataset represents names used on five or more applications for security purposes. Data has been gathered from the Social Security Administration and put into a CSV file by Kaggle.

All data used is public information and the latest dataset should be used when released to the public for accuracy.

B.5. METHODOLOGY

The waterfall method will be used to develop the new program. Development will be divided into 6 stages:

- Requirement gathering
- Analysis
- Design
- Coding
- Testing

- Deployment

Our team will ensure requirements are met for a stage before moving onto the next one. Thorough documentation will be provided to keep stakeholders up to date and serve as a guideline incase backtracking is needed. All of the testing will be conducted by developers not assigned to the project thus requiring minimum involvement from the client. With a clear scope established milestones can be laid out before development begins.

B.6. DELIVERABLES AND OUTCOMES

Project deliverables include:

- Project timeline
- Weekly status reports
- Testing methods
- Testing results
- Training manual

The finished product will include:

- Data and program housed in current BRI server
- Compatibility with Windows
- User-friendly interface
- Ability to search dataset for name to display times name was used to apply for Social Security card per year
- Ability to search for top 5 boy and girl names for a given year
- View of most popular searches conducted by customers

B.7. IMPLEMENTATION PLAN AND OUTCOMES

Implementation:

- Product added to Baby Registry Inc. server: application and data will be housed on BRI's main server
- Remote network set up: network will be set up for remote access from all branches
- Training/demonstration: employees will be given demonstration on how to use new application
- Maintenance: IT will be given demonstration on how to update server with new data for subsequent years

Outcome:

- Search dataset for name view times used to apply for Social Security card each year
- Search for top 5 boy and girl names for a given year
- View most names most searched by customers

B.8. EVALUATION PLAN

To measure customer satisfaction:
Baby Registry Inc. currently sends out a customer satisfaction survey once parents close their BRI account. The survey contains questions on all services available. An additional section will be added for the new popular baby name application. Employees are also given monthly satisfaction surveys regarding work conditions. A section will be added to the employee survey as well to determine usability. The program will be considered successful if there is a 85% satisfaction rate from both employees and consumers.

To measure financial requirements:
Successful analysis of names should provide the company with information on which products will be most popular among parents. Manufacturing should be increased to lower the cost per unit. Less storage should be needed with inventory being made up mainly of in demand merchandise. The program will be considered successful financially if products produced as a factor of the program can increase profits by 10% within the first year.

B.9. RESOURCES AND COSTS

Human resource requirements:

Requirement Gathering			
Service	**Cost per Hour**	**Required Hours**	**Total Cost**
Planning and Design	$85	20	$1700

Analysis			
Service	**Cost per Hour**	**Required Hours**	**Total Cost**
Current system analysis	$85	5	$425

Design			
Service	**Cost per Hour**	**Required Hours**	**Total Cost**
Application design	$85	15	$1275

Coding			
Service	**Cost per Hour**	**Required Hours**	**Total Cost**
Development	$150	120	$18000

Testing			
Service	**Cost per Hour**	**Required Hours**	**Total Cost**
Application testing	$100	50	$5000

Deployment			
Service	**Cost per Hour**	**Required Hours**	**Total Cost**
Review and Analysis	$75	40	$3000

Total: $29400

The following are required to run the program:

- Microsoft Windows 7, 8, 10
- In-house server
- MySQL database

These requirements will not affect the cost of the product as Baby Registry Inc. possesses all of the necessary resources. The program will utilize the existing remote server held at the main office. All of Baby Registry Inc. computers operate on Windows 10. The program is designed to run on Windows to prevent any additional costs such as software updates or new computers. The only cost would be for the human resources needed to develop and create the application.

B.10. TIMELINE AND MILESTONES

Milestone	Start Date	Completion Date	Duration	Dependencies	Resources
Requirements Meeting	09/04/2020	09/04/2020	20	N/A	• BRI server analysis
UI Design	09/08/2020	09/10/2020	10	• Requirements meeting	• Requirements list
Language and Framework Finalization	09/14/2020	09/15/2020	10	• UI design	• UI Design
Architectural Design	09/16/2020	09/25/2020	110	• Framework/ • language finalization	• Framework
Design Review Meeting	09/28/2020	09/29/2020	10	• Architectural design	N/A
Testing	09/30/2020	10/02/2020	50	• Architectural design	•
Set Up Development Environment	09/30/2020	10/01/2020	15	• Review meeting • Testing	• Server access
Data Analysis	10/06/2020	10/07/2020	15	• Setup environment	N/A
Program Deployment	10/08/2020	10/09/2020	8	• Data analysis	N/A
Product Demo Meeting	10/12/2020	10/12/2020	2	• Deployment	N/A

C. POST IMPLEMENTATION REPORT (SECTION D)

C.1.PURPOSE STATEMENT

For Baby Registry Inc. to be able to provide embroidered products with baby names, being aware of popular baby names and trends is the best way to accurately predict the demand for goods. We at Custom Tech understand there is a consistent need to set your company apart from others in the same industry. Baby Registry Inc. has reached out to our company to come up with a solution that will show any trends in baby names to help with predicting the demand of embroidered products with baby names. To provide an extra service not offered by other companies, this program will also assist parents with naming their child by providing information on the popularity of a name through the years as well as displaying popular names for any given year.

Vision	**What is our vision for goal for creating this product?**		
Target Group	**Needs**	**Product**	**Value**
The program will be utilized by both the Baby Registry Inc. company as well as consumers.	For BRI, the product will provide a method for analyzing name trends to increase profits. For consumers, an additional service will be provided to assist with any questions related to baby names.	The product will be a standalone application that will display available necessary data in a way that will allow for easy analysis.	With the ability to analyze trends, BRI will be able to better predict the demand of each item. Popular products can be mass produced to lower manufacturing costs as well as storage overhead costs.

C.2. DATASETS

The data used is public information made available by the Social Security Association and Kaggle. Data is updated periodically to add statistics for subsequent years. Information provided for each data item includes the name, gender, count, and year. The dataset accounts for different spellings of the same name if 5 or more people have applied for a Social Security card with that spelling. Due to security reasons, names given to fewer than 5 people are not included in the dataset.

Data used was converted to a CSV file by Kaggle.

b_names	name_sex	name_count	name_year
Lorena	F	510	2000
Ryder	M	203	2000
Adrienne	F	496	2000
Deion	M	175	2000
Jet	M	56	2000
Yogi	M	5	2002
Adama	F	13	2000
Demarcus	M	302	2000
Sheily	F	9	2000

Figure 1: Raw ataset

The original data contained information on the name, sex, count, and year. To better format the data for use, an ID column was added to act as the primary key and a column to keep track of the number of times a name is searched.

id	b_names	name_sex	name_count	name_year	search
1332922	Lorena	F	510	2000	9016
1350843	Ryder	M	203	2000	8356
1332935	Adrienne	F	496	2000	8276
1350939	Deion	M	175	2000	7822
1351930	Jet	M	56	2000	7382
1422905	Yogi	M	5	2002	7281
1339981	Adama	F	13	2000	7202
1350668	Demarcus	M	302	2000	6935
1343019	Sheily	F	9	2000	6890

Figure 2: Dataset with added attributes

C.3.CODE FOR ANALYSIS

The accuracy of the program is dependent on the number and group of people using the application. With this program created specifically for Baby Registry Inc., only those registered with or working for BRI will have access to the service. A different user group would yield different results, however BRI is most interested in what its consumers are searching.

The program works by gathering the user's input and searching the database for information to display. A SQL statement is sent to the server to search for relevant data. If the user's search cannot be found in the database, an error message is displayed.

The application also records each search on the Popularity Through the Years screen using the search attribute.

	id	b_names	name_sex	name_count	name_year	search
▶	1332922	Lorena	F	510	2000	9016
	1350843	Ryder	M	203	2000	8356
	1332935	Adrienne	F	496	2000	8276
	1350939	Deion	M	175	2000	7822
	1351930	Jet	M	56	2000	7382
	1422905	Yogi	M	5	2002	7281
	1339981	Adama	F	13	2000	7202
	1350668	Demarcus	M	302	2000	6935
	1343019	Sheily	F	9	2000	6890
	1340023	Arynn	F	13	2000	6829
	1340691	Kelyn	F	12	2000	6723

Figure 3: All attributes of the dataset

Each time a name is searched, the program increases the current attribute count by 1 and updates the record.

```
public static void updateCount(CountUpdate updateCount) {
    try {
        statement = connect.prepareStatement("UPDATE baby_names SET search = ?  WHERE id = " + updateCount.getId());
        statement.setInt(1, updateCount.getSearch() + 1);
        statement.execute();
    } catch (SQLException e) {
        e.printStackTrace();
    }
}
```

Figure 4: Method to record number of searches

To retrieve the most searched names, the application sends a SQL statement to the server to process and order the information. All records are sorted by the search attribute in ascending order. The top 10 records are returned to the application.

```
public static ObservableList<CommonSearch> topSearches () {
    ObservableList<CommonSearch> searches = FXCollections.observableArrayList();
    int rank = 0;
    try {
        statement = connect.prepareStatement("SELECT * FROM baby_names ORDER BY search DESC");
        ResultSet result = statement.executeQuery();
        while(result.next()) {
            if (rank < 10) {
                rank++;
                CommonSearch list = new CommonSearch();
                list.setRank((Integer.toString(rank)));
                list.setName(result.getString("b_names"));
                list.setGender(result.getString("name_sex"));
                list.setCount(result.getInt("name_count"));
                list.setYear(result.getString("name_year"));
                list.setSearch(result.getInt("search") + 1);
                searches.add(list);
            }
        }
         return searches;
    } catch (SQLException e) {
        e.printStackTrace();
    }
    return searches;
}
```

Figure 5: Method for retrieving the most searched names

C.4. HYPOTHESIS VERIFICATION

The hypothesis that will be looked at is:

- Does the program contribute to an increase in profits?

Within the first year, the program should lead to a 10% increase in profits either from increased revenue or from lower manufacturing and storage costs. Although this cannot be verified until the program has been in use for a full year, we can test to see if all of the necessary tools are present to help with increasing profits. The program should allow for easy analysis of name trends by displaying the number of times a name has been used to apply for a Social Security card through the years. A counter should also keep track of how many times each name is searched to provide a more accurate prediction of popular names among Baby Registry Inc. consumers. With both of these features present, BRI should be on track to increase profit margins if the results are applied during product design and manufacturing.

C.5.VISUALIZATION AND REPORTING

Each feature includes a table to display all the necessary data. Information from these tables are also included in various graphs for visual data comparison.

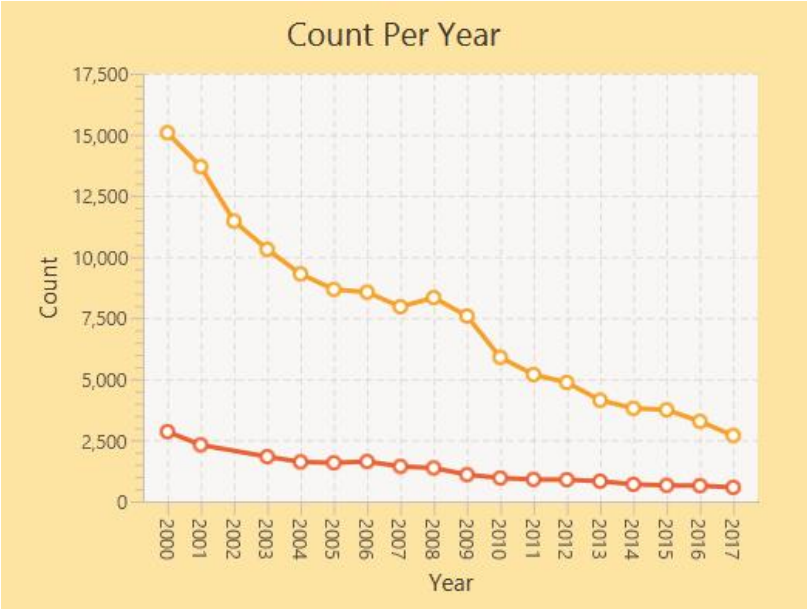

Figure 6: Number of times the name Taylor was given each year to boys (red) and girls (orange)

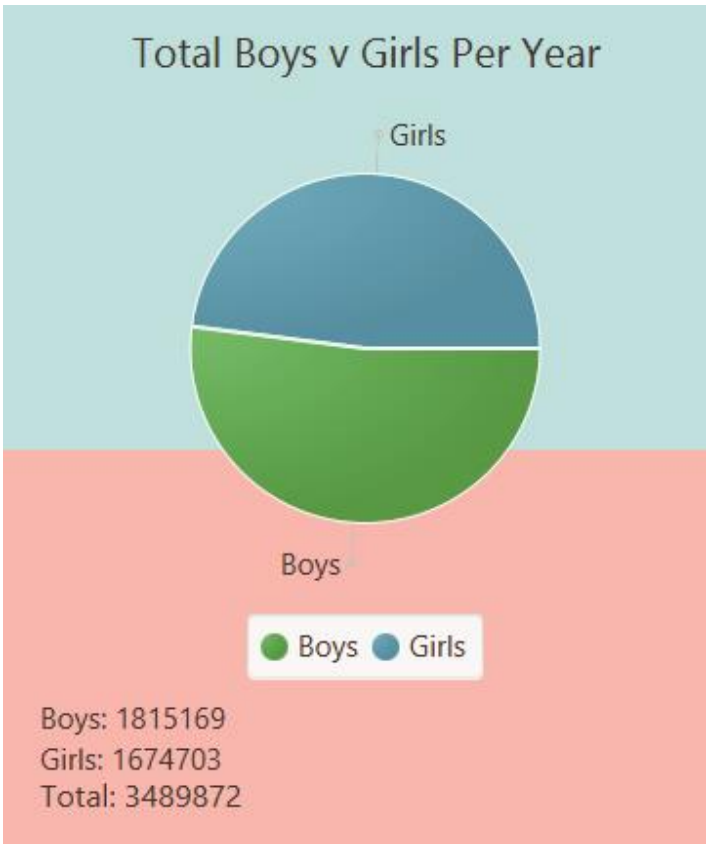

Figure 7: Total number of boys vs. girls that applied for a Social Security card in 2017

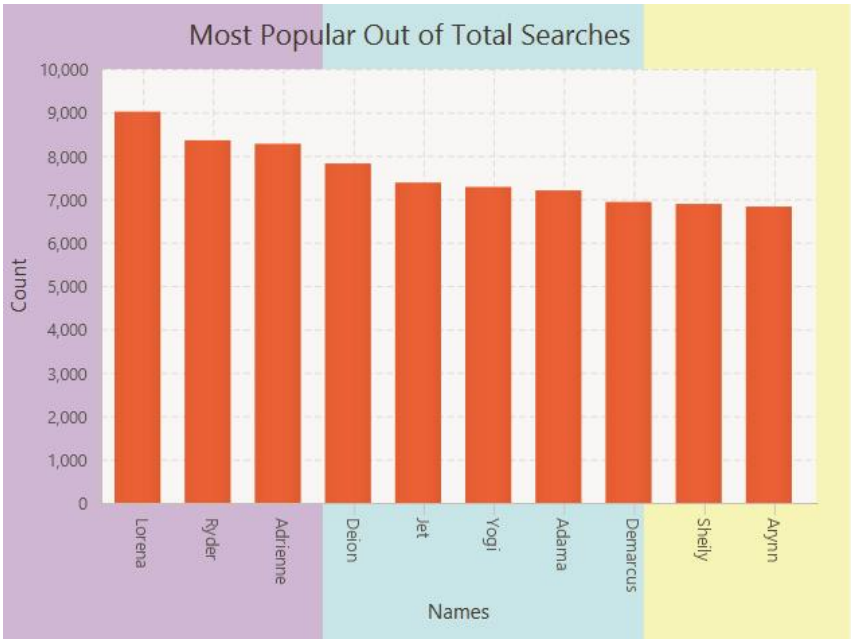

Figure 8: Most searched names by users

C.6.ACCURACY ANALYSIS

The program will be put through a number of tests to ensure accuracy:

- The same search will be entered multiple times to confirm the program is producing consistent results.
- Program results will be compared to the expected results manually pulled from the dataset.
- Tests will be conducted with multiple datasets.

C.7.APPLICATION TESTING

Black box testing was conducted, and outputs were compared to the database for accuracy. Valid inputs were tested as well as invalid inputs to ensure the application did not produce any unwanted results. Multiple datasets were used for testing.

C.8.APPLICATION FILES

The following table provides a list of files needed to execute the application:

Filename	Description
MainFile.java	Calls the program to display the begin screen
BeginScreen.fxml	Properties for the beginning screen
BeginScreenController. java	Contains code to instruct the program on what to do when buttons are clicked on the beginning screen
CommonSearchNames.fxml	Properties for the Most Commonly Searched Names screen
CommonSearchNamesController. java	Contains code to instruct the program on what to do when buttons are clicked and data to display on the Most Commonly Searched Name screen
MainMenu.fxml	Properties for the main menu
MainMenuController. java	Contains code to instruct the program on what to do when buttons are clicked on the main menu
NamePopularity.fxml	Properties for the Popularity Through the Years screen
NamePopularityController. java	Contains code to instruct the program on what to do when buttons are clicked and data to display on the Popularity Through the Years screen
YearList.fxml	Properties for the Popularity by Year screen

YearListController. java	Contains code to instruct the program on what to do when buttons are clicked and data to display on the Popularity by Year screen
BoyRankings. java	Class object for boy names used on Popularity by Year screen
CommonSearch. java	Class object for names used on Most Commonly Searched Name screen
CountUpdate. java	Increases count by 1 when name is searched
GirlRankings. java	Class object for girl names used on Popularity by Year screen
Year. java	Class object for names used on Popularity Through the Years screen
DatabaseConnectionString. java	Contains information needed to connect to the MySQL server
NamesDatabase. java	Utilizes SQL to search database for information to display on Popularity Through the Years screen
SearchDatabase. java	Utilizes SQL to search database for information to display on Most Commonly Searched Name screen
YearsDatabase. java	Utilizes SQL to search database for information to display on Popularity by Year screen
Baby_Names.jar	Executable file that utilizes all files above to run application

C.9.USER'S GUIDE

Installation guide:

1) Download the zipped file submitted on the Task Submission page.
2) Unzip the file.
3) Navigate to the dist folder inside the Performance Assessment folder.

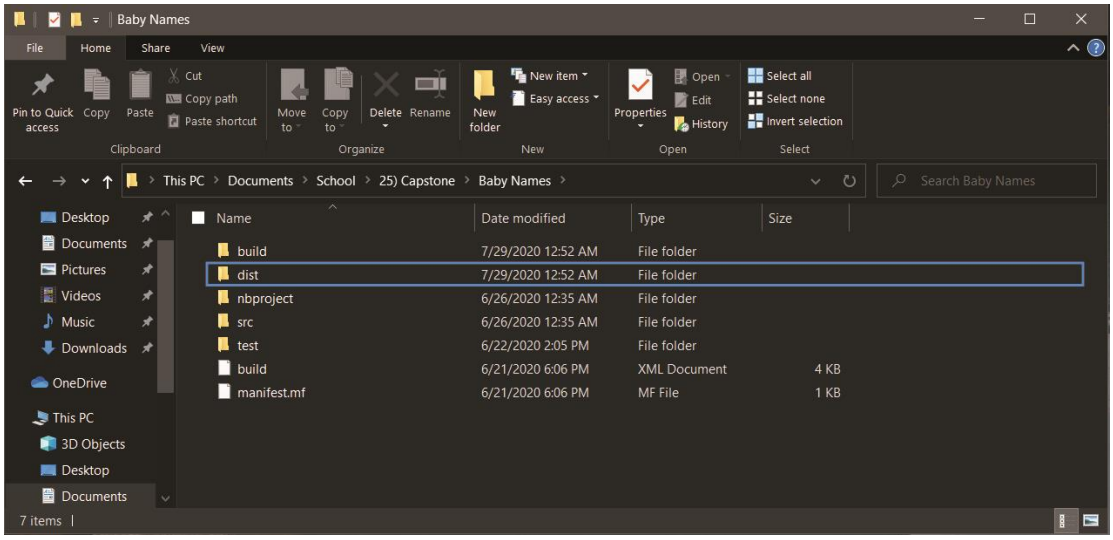

Figure 9: File directory

4) Double clicking on the Baby_Names Executable Jar File will run the application.

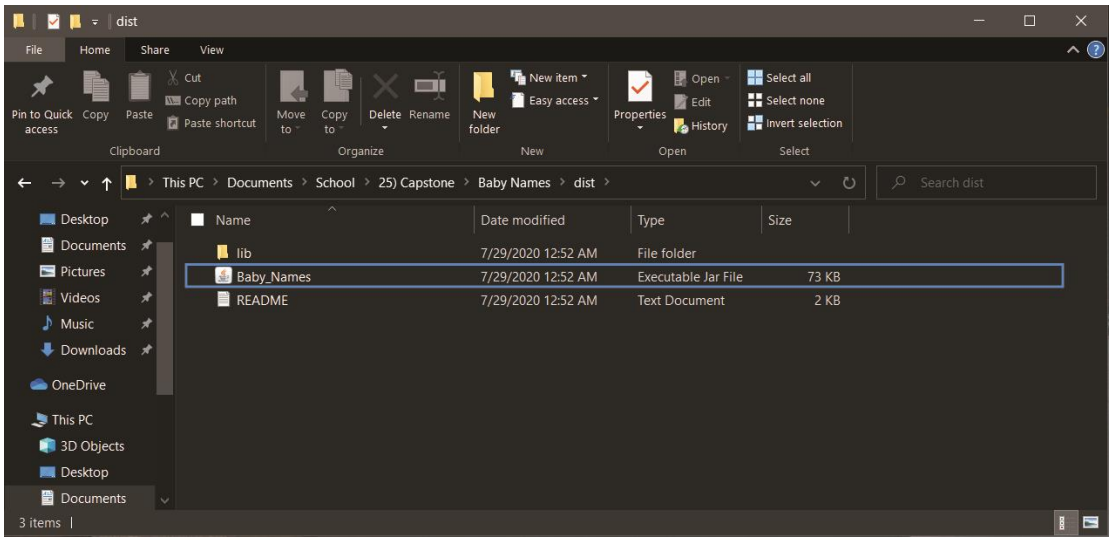

Figure 10: File directory

5) For easy access, a shortcut to the Baby Names file from the desktop can be created.

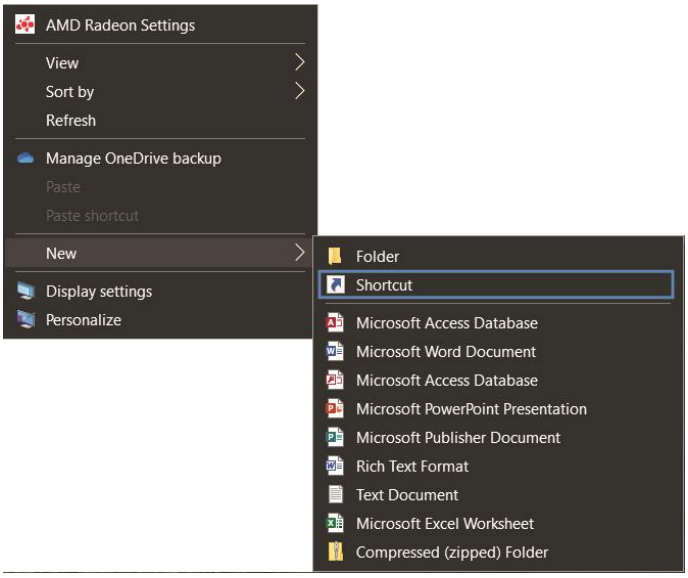

Figure 11: Creating a shortcut

6) Navigate to where the Baby_Names Executable Jar File is located.

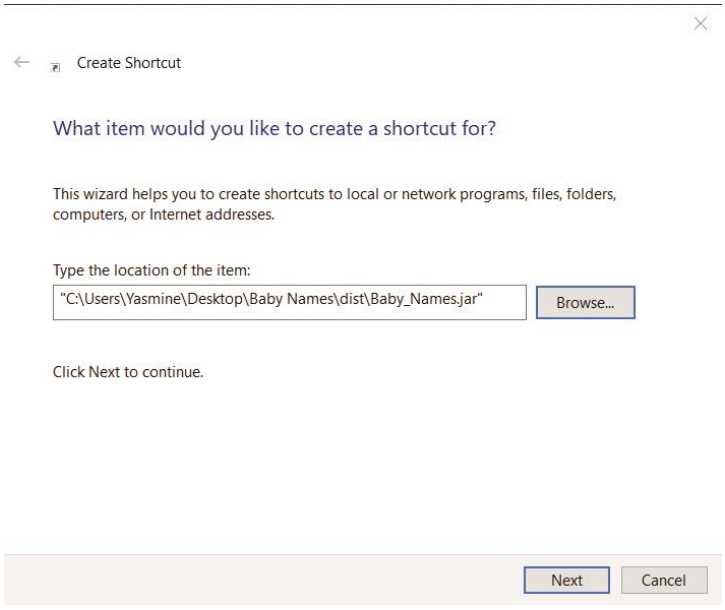

Figure 12: Creating a shortcut

Application Guide:

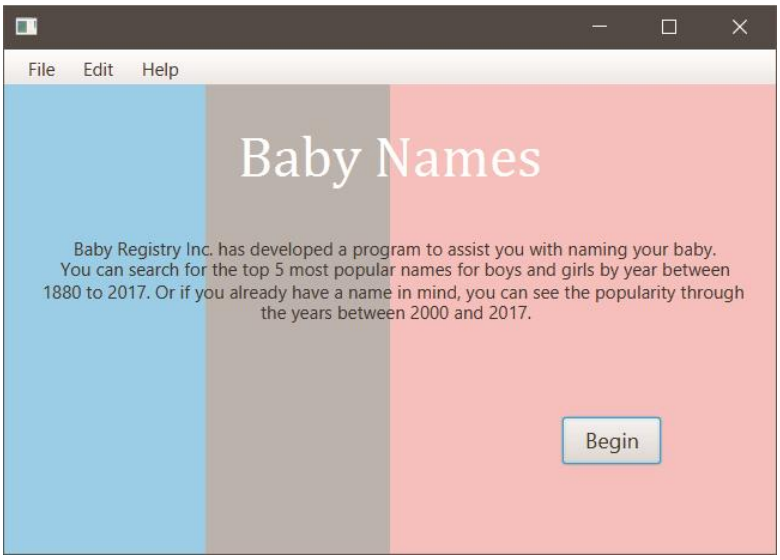

Figure 13: Opening screen

Figure 14: Main menu

1) Search for a name to display information about how many people of each gender applied for a Social Security with that name each year between 2000 – 2017.

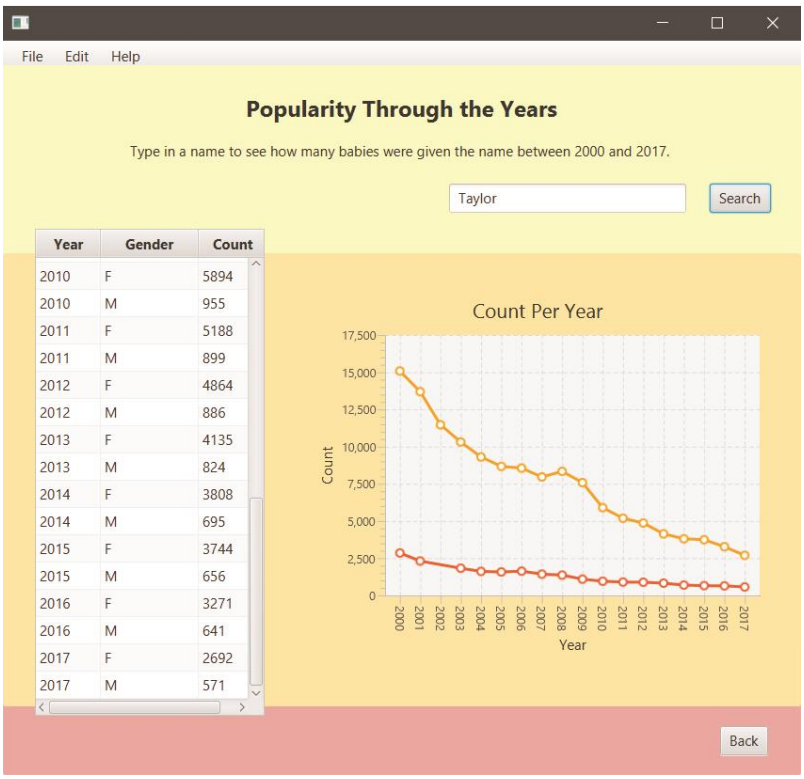

Figure 15: Popularity through the years screen

2) Search for any year between 1880 – 2017 to display the top 5 boy and girl names. A pie chart displays the ratio of boys to girls that applied for a Social Security card that year.

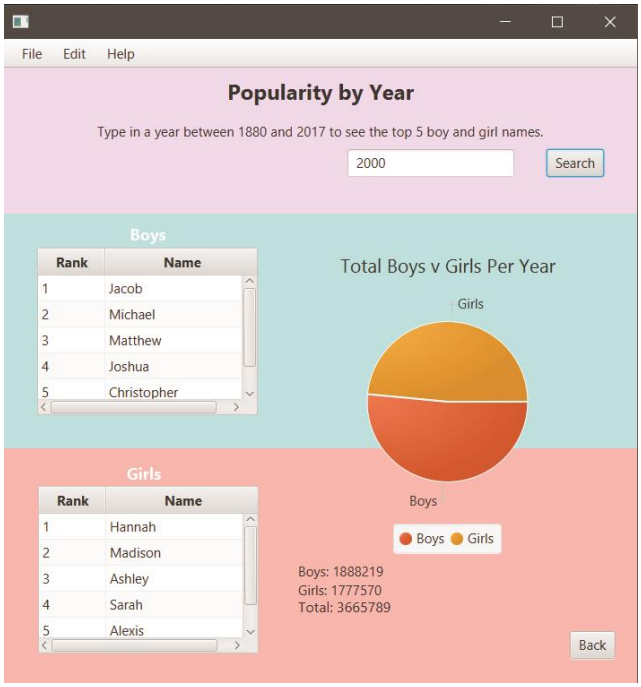

Figure 16: Popularity by year screen

3) The top 10 most searched names by Baby Registry Inc. customers are displayed.

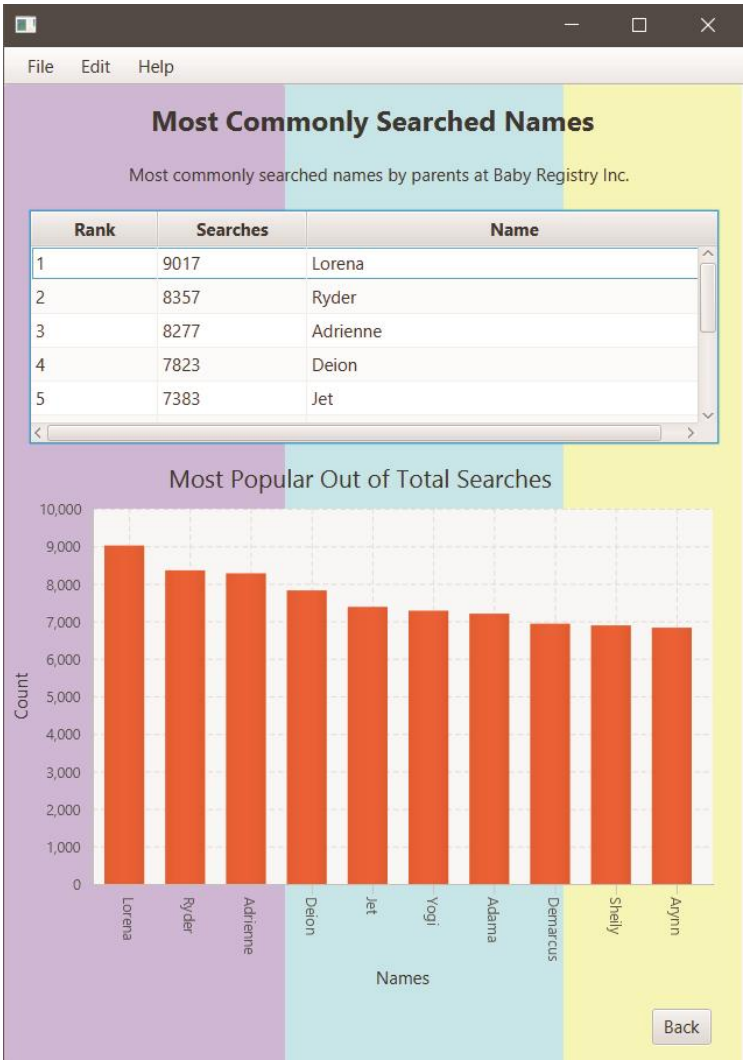

Figure 17: Most commonly searched names screen

D. SOURCES

Dunleavy, J., Gallagher, A., & Reeves, P. (2019, April 23b). The Waterfall Methodology: Why You Should Use It. Retrieved from https://developer.ibm.com/technologies/devops/articles/waterfall-model-advantages-disadvantages/

Morad, H. (2018, September 14). US National Baby Names (1880-2017). Retrieved July 6, 2020, from https://www.kaggle.com/hassenmorad/us-national-baby-names-18802017

Social Security Number Holders. (n.d.). Retrieved July 6, 2020, from https://www.ssa.gov/OACT/babynames/numberUSbirths.html