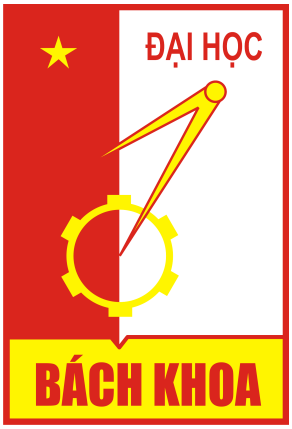
HANOI UNIVERSITY OF SCIENCE AND TECHNOLOGY



Project 1

Program for Facebook information reconnaissance and management with AirTable

Vũ Hoàng Ngọc - 20210646
Nguyễn Duy An - 20214943
ngoc.vh210646@sis.hust.edu.vn
an.nd214943@sis.hust.edu.vn

Major: Cyber Security
Specialization: K66

Supervisor: Associate Professor Tống Văn Vạn _____

Signature

Department: Cyber Security

School: Information and Communications Technology

HANOI, 07/2023

TABLE OF CONTENTS

CHAPTER 1. INTRODUCTION

1.1 Objectives and scope of the project

The purpose of this project is to develop a Java-based program for social media intelligence gathering, specifically focusing on Facebook, and synchronize the gathered information with AirTable, a cloud-based collaboration platform, while ensuring the security and privacy of the users involved. By leveraging the power of OSINT techniques, the program will enable users to gather information from Facebook users and groups, perform data analysis, generate visual charts, and export results in various formats.

The project's scope includes designing and implementing a Java-based program that integrates with Facebook's API and AirTable's platform. We choose Facebook which is one of the most popular Social Media platforms in the world and get data from Facebook Api Tool explorer. The permissions we get in our Meta App includes public_profile and groups_access_member_info through feature "Get User Acess Token" of the Facebook Api Tool explorer.

1.2 Approach

The approach involves practicing and applying security programming rules in Java, studying security and privacy concerns, utilizing OSINT techniques, implementing data analysis algorithms, and providing export functionalities.

1.3 Work Contribution

Name	Task Performed	Contribution Level
Nguyen Duy An - 20214943	• Focus on Facebook API, create a Facebook virtual application to get access tokens for testing. • Build the Use case list and diagram for the tool. • Scan to check source code with Jenkins & SonarQube. • Analyze the data. • Write the report & Present.	50%
Vu Hoang Ngoc - 20210646	• Focus on Airtable API, create Airtable base and do the tasks which work with Airtable. • Build the model and the structure of source code. • Check source code again following the SEI CERT Oracle Coding Standard for Java and fix it. • Write the report & Present. • Check the source code of the cross-check group.	50%

Table 1.1: Contributor Information

CHAPTER 2. REQUIREMENTS ANALYSIS

Project Description: *"Programming to build a program for social network information reconnaissance and synchronization to AirTable."* To deal with this problem, we are suggested the following approach. After that, we will make some adjustments to comply with the limited API of the social network we have chosen.

- **The program should be written in Java OOP**

  We utilized the OOP skills and secure coding skills we learned in class to program it in Java using Eclipse IDE.

- **Research Security and Privacy Issues on Social Networks**

  We have conducted research on Facebook APIs using the following public references:

  1. References for Meta Developers: `https://developers.facebook.com/docs/`
  2. Meta Platform, Facebook Graph API Terms & Policy for Developers: `https://developers.facebook.com/terms/`
  3. Privacy Policy for Users: `https://www.facebook.com/privacy/policy/`

- **Research AirTable and its APIs**

  We learned how to use Airtable and its API to GET or POST data to an Airtable Base. We completed all of these tasks by following the references provided by Airtable's documentation.

  1. Airtable general guides for all Users: `https://www.airtable.com/guides`
  2. API tutorial for Developers: `https://airtable.com/developers/web/api/`

- **Learn about OSINT - Open Source Intelligence**

  Learn about OSINT - the technique of collecting identity information on social networks. Allow for the collection of information about members, groups, hashtags, etc., on Facebook.

- **Create a Virtual Application on Facebook**

  We have created a virtual application on Facebook. To access the application, users need to read and agree to our requirements (including accessing the user's profile and data related to the groups that the user participates in). When the user agrees to grant the virtual application access to their personal information, we will receive an access token from the customer and use it to collect the necessary data. After thoroughly researching Facebook's terms and conditions for applications that collect user data, we obtained the following information about the user:

  - User ID
  - User name
  - The groups that the user is a member of (including the group name, group ID, number of members, and whether the user is an administrator of the group or not)

- **Analyze and Summarize Data. Generate Charts and Extract Results in Excel**

  We collected data and created data records in Airtable. The data was then analyzed and summarized from the Airtable database, and we created a few statistical charts and backed up all the data into an Excel file.

CHAPTER 3. SYSTEM DESIGN

3.1 Usecases Diagram

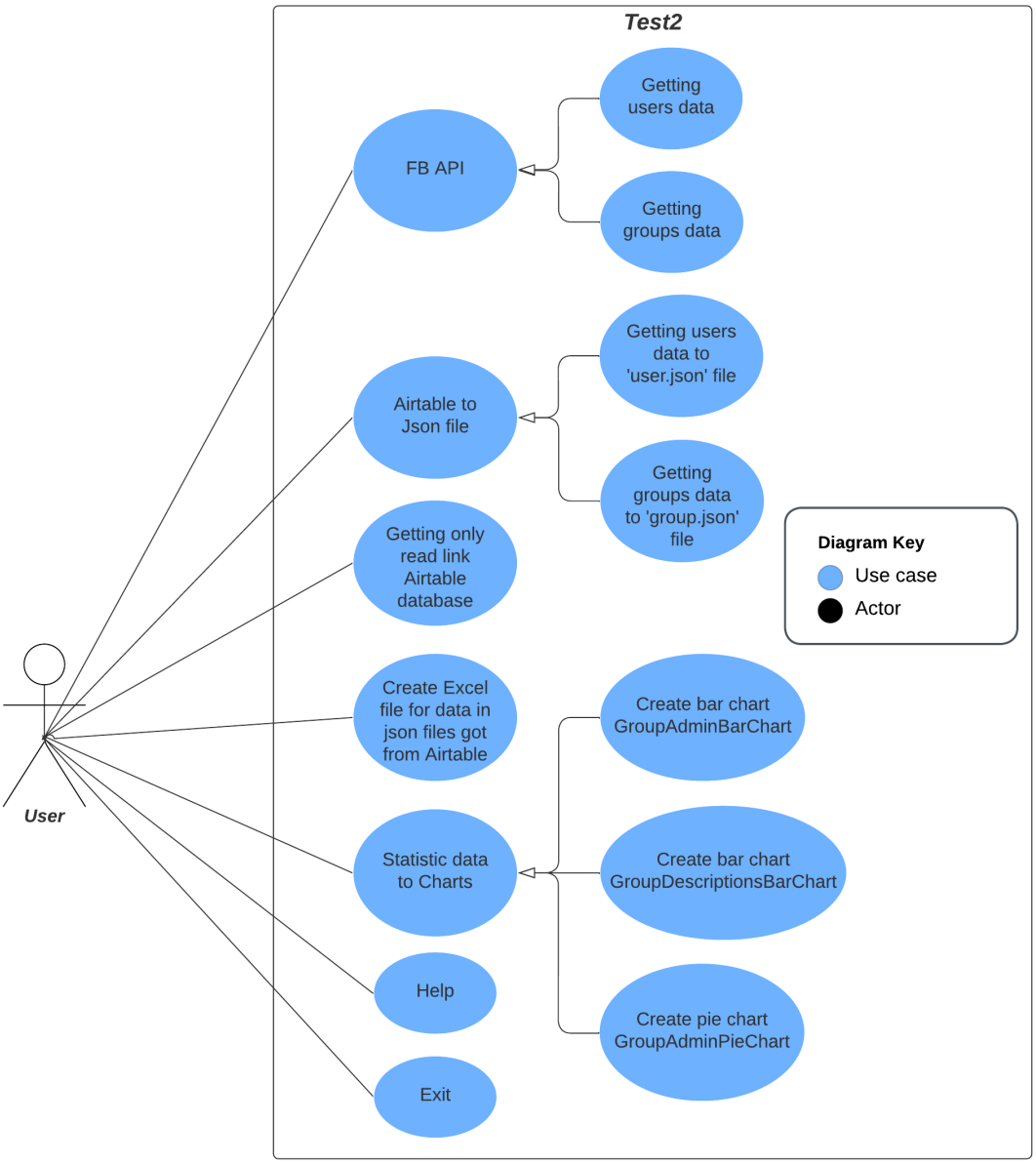


Figure 3.1: Use cases diagram

3.2 Package (UML diagram)

Abstract PostRequest class: To post data get from FB APi on Airtable, 2 'detailed' classes RecordGroup and RecordUser to get their table case. GetRequest Class: To get data from table we want from Airtable

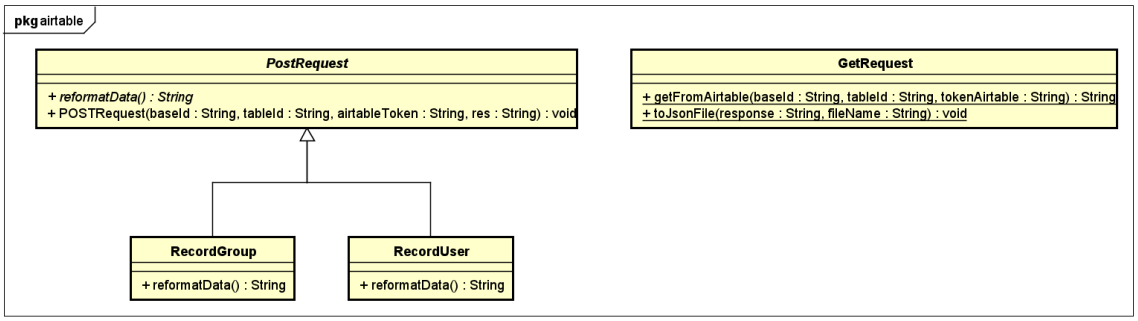


Figure 3.2: package airtable

GetData class: has 4 subclass for get data through perform HTTP request detailed cases (User, Group, GroupFeed, GroupAdmin)

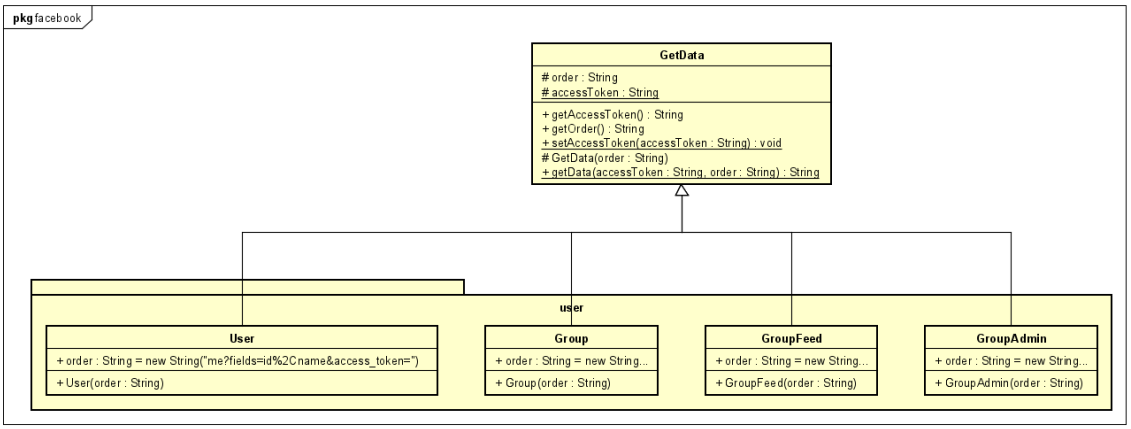


Figure 3.3: package facebook

Package stat: have 4 classes to generate excel, groupAdminBarchart(number of group user is admin), groupAdminPieChart(portion of group user is admin) and groupDescriptionBarChart(number of group has description).

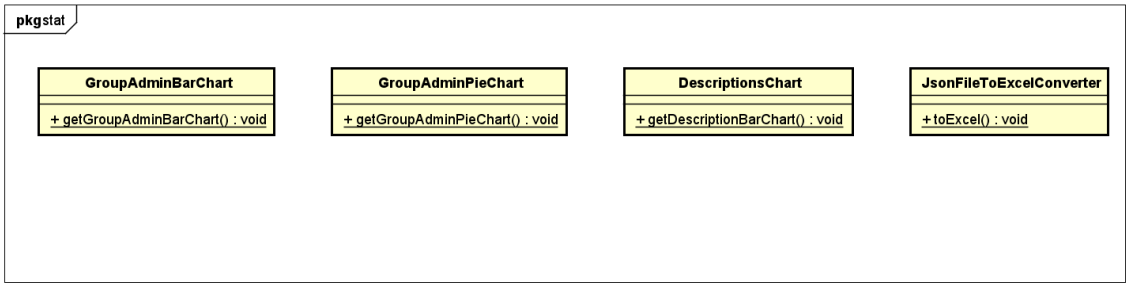


Figure 3.4: package statistic

4.1 Testing

Checking by ourselves

- By seeing the warning in the Markdown of Eclipse IDE, somewhere we forgot to close the scanner, workbook. This is violated with FIO04-J which can lead to "Resources leaks". So, we closed it after using it.
- Some sensitive information like the API key, token of Airtable is in our source code, it also can lead to resource leaks. Therefore, we moved this information to another JSON file.
- If the tool is given a Facebook access token that is either invalid or unauthorized, it will cause the program to halt abruptly when the token is requested. So, we have added a satisfied exception for that error case.

Checking by Jenkins & SonarQube

1. When we scanned at the first time, we had seen 1 bug and 1 kind of security hotspot:
    - All 8 security hotspots have the same problem. Because we print out the StackTrace of the error when we catch it. It violated the rule ERR00-J.

Rule	Severity	Likelihood	Remediation Cost	Priority	Level
ERR00-J	Low	Probable	Medium	P4	L3

Figure 4.1: ERR00-J: Do not suppress or ignore checked exceptions

2. The second time, after we updated the CLI of the tool, 2 bugs were found. But no security point happens.
    - The first is the closing of FileWriter when we create a JSON file. So we need to handle it also in the catch block.

Figure 4.2: FIO04-J: Close resources when they are no longer needed

- The second is the end condition for the CLI. So we added a break method to end it.

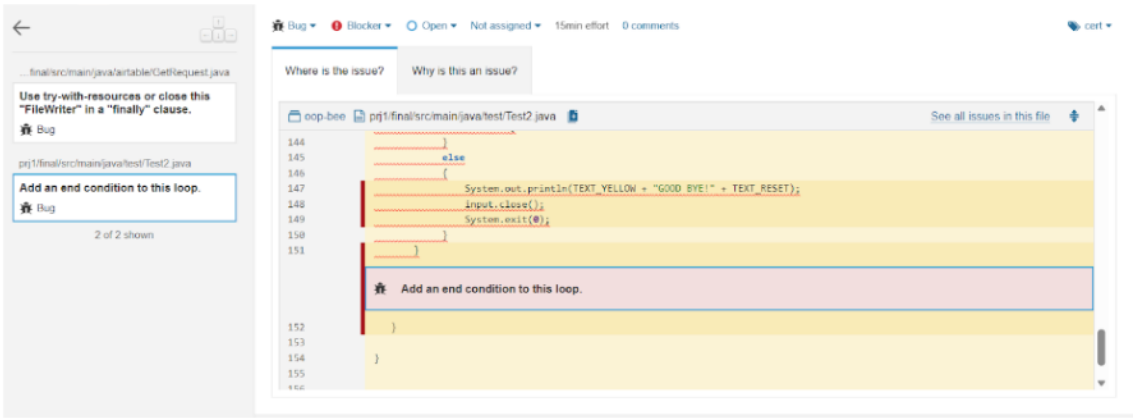
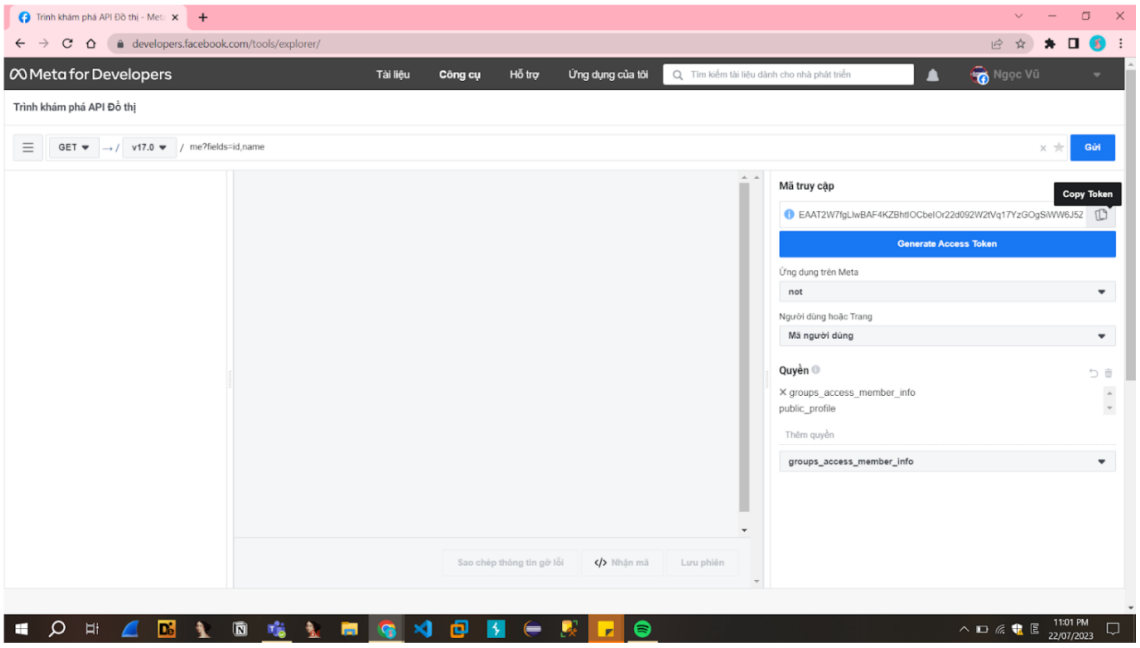


Figure 4.3: Close loop bug

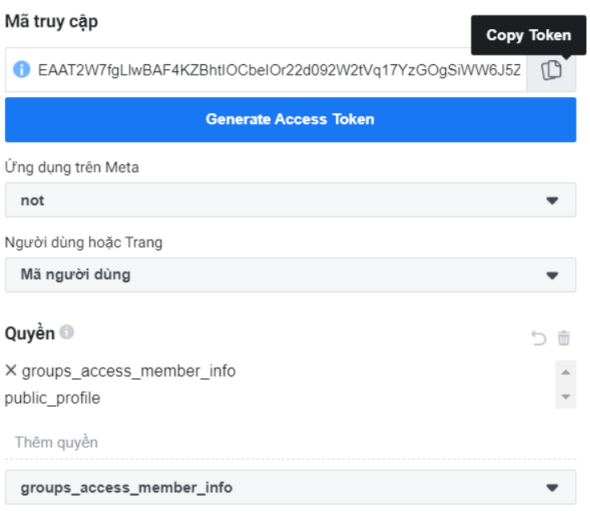
4.2 Implementaion

Get Access Token To get an access token, follow these steps:

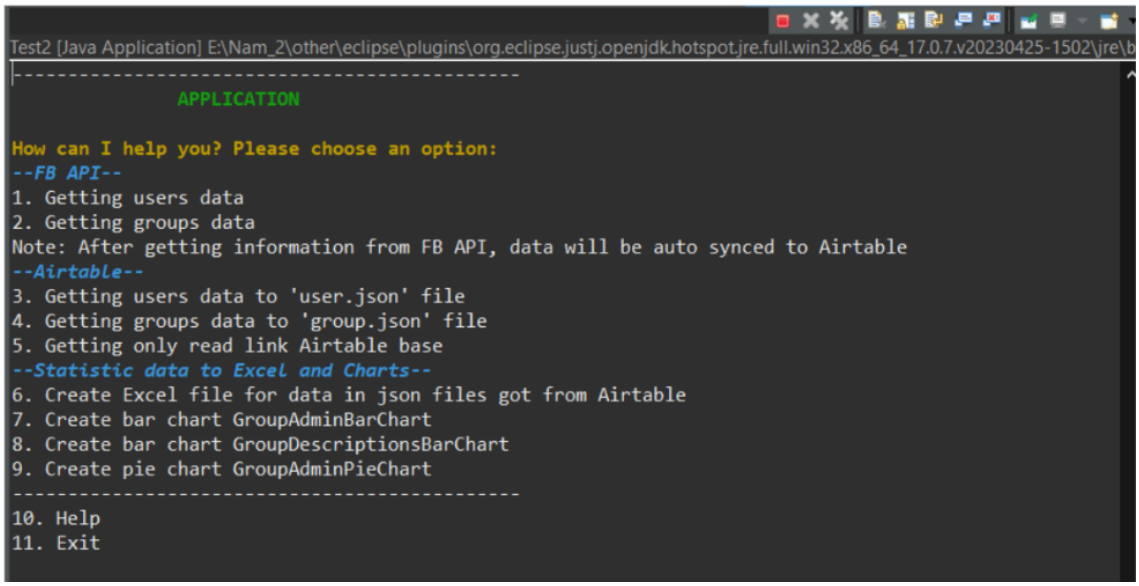
- Go to the link: `https://developers.facebook.com/tools/explorer/` and choose your virtual application. Generate the access token and accept the terms.



- Copy the generated access token for use in your tool.



Getting Started with our Tools: When running the tool, an instruction menu with 11 choices will appear:



Main Methods There are 3 main methods with this tool:

1. Get Facebook data to post into Airtable base (Option 1, 2)
2. Get data from Airtable (Option 3, 4, 5)
3. Statistic data with Excel file and generate charts (Option 6, 7, 8, 9)

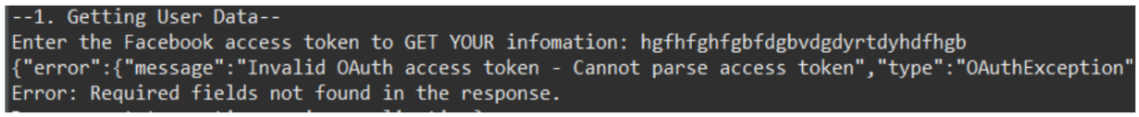
4.2.1 Get Facebook data to post into Airtable base (Option 1, 2)

- Option 1: The tool will get data from a user (Name, ID, etc.).
- Option 2: The tool will get data from a Group (Name, ID, Description, user who we investigate is admin or not).

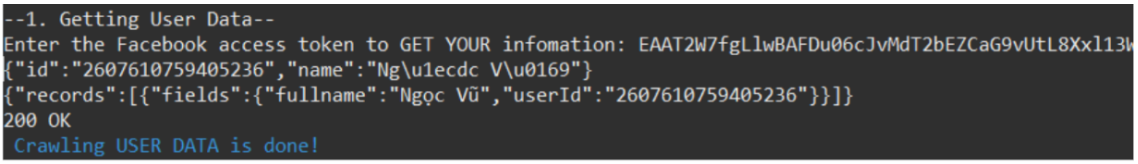
Both options require the following step:

- Declare a valid access token from Facebook.

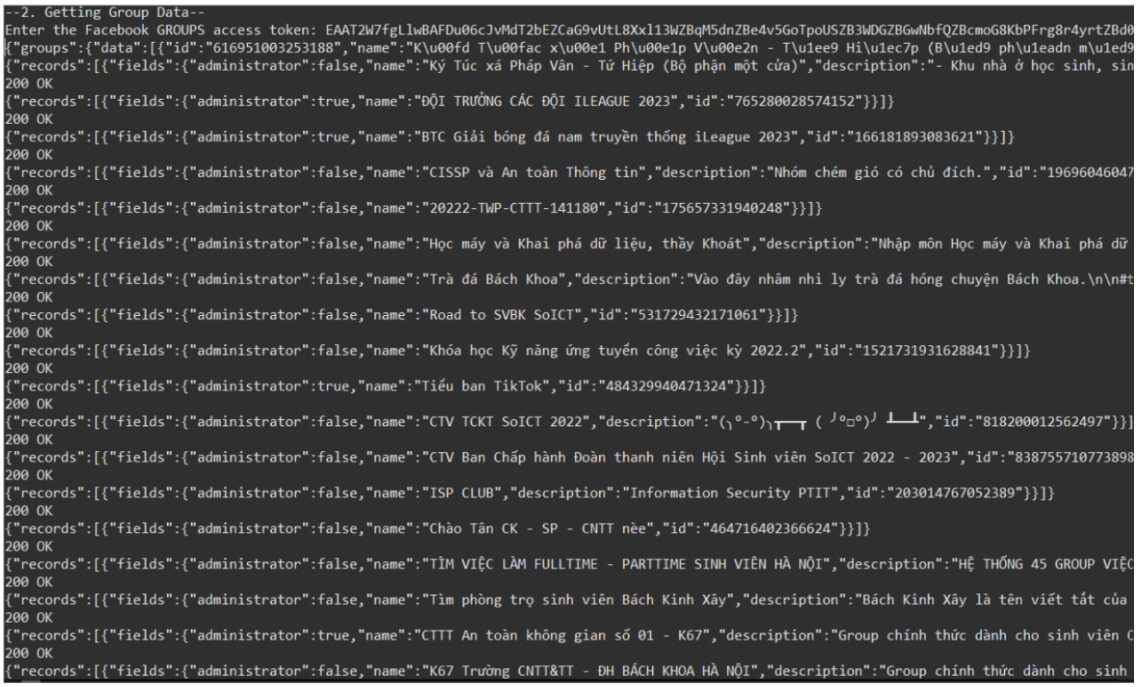
- If the access token entered is invalid, an error will occur, and the program will raise an exception.



- If the access token is valid, the program will receive data and attempt to post it to Airtable, resulting in a ’200 OK’ status if successful. If posting to Airtable fails, the program may raise another status code, such as ’422’.



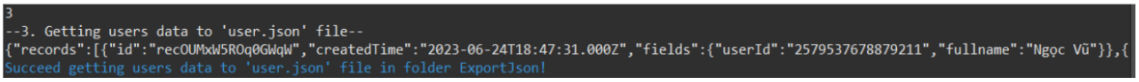
- Similar to option 1, after entering a valid access token to retrieve group data, the program will prompt the user to perform tasks similar to those required to retrieve profile data.



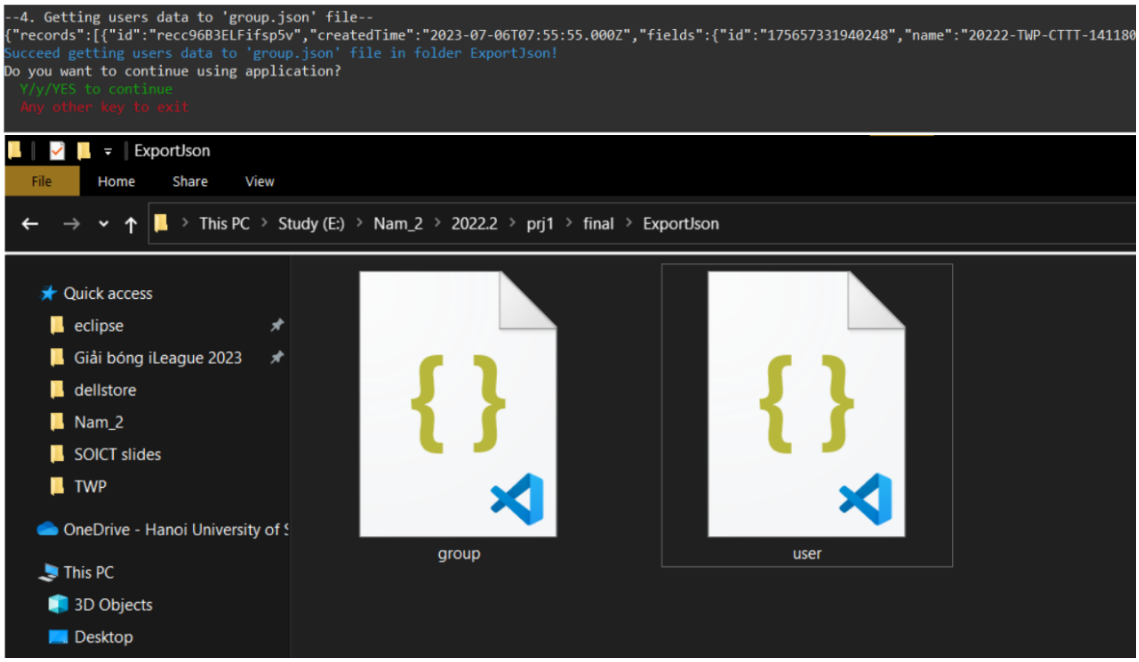
4.2.2 Get data from Airtable (Option 3, 4, 5)

All data recorded will be shown if you choose one of these options:

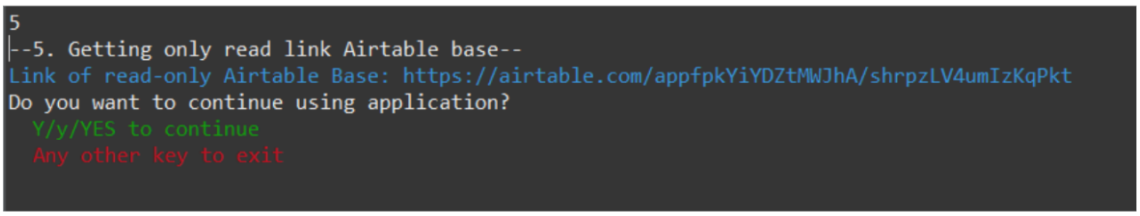
1. Data of every user obtained by the tool will be exported to a JSON file.



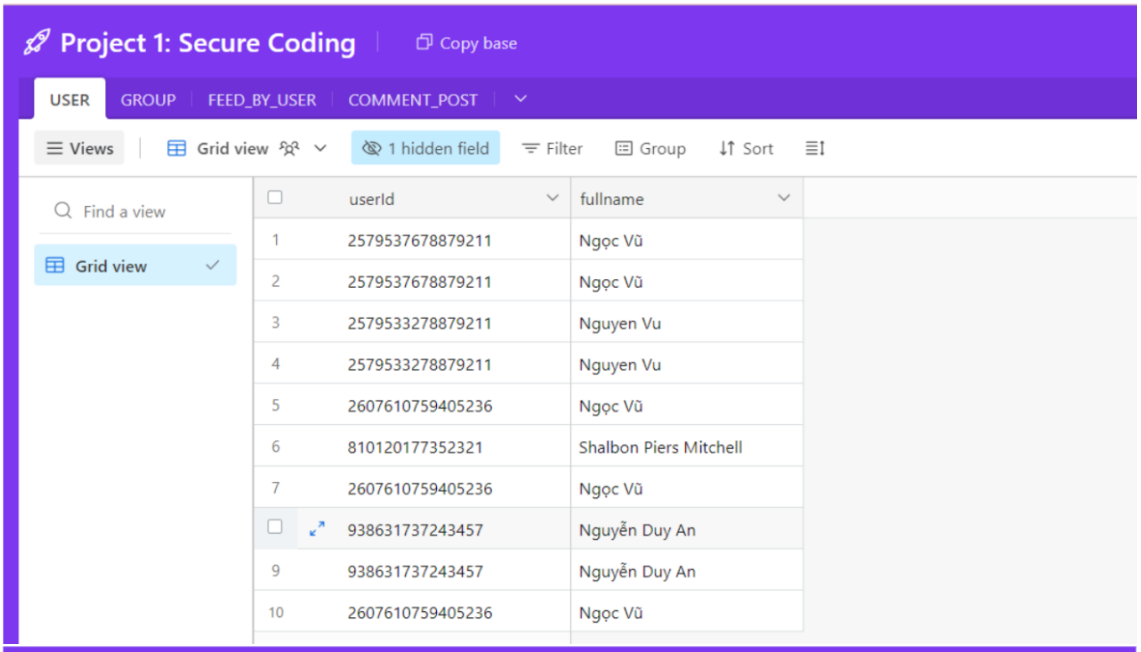
2. Data of every group obtained by the tool will be exported to a JSON file.

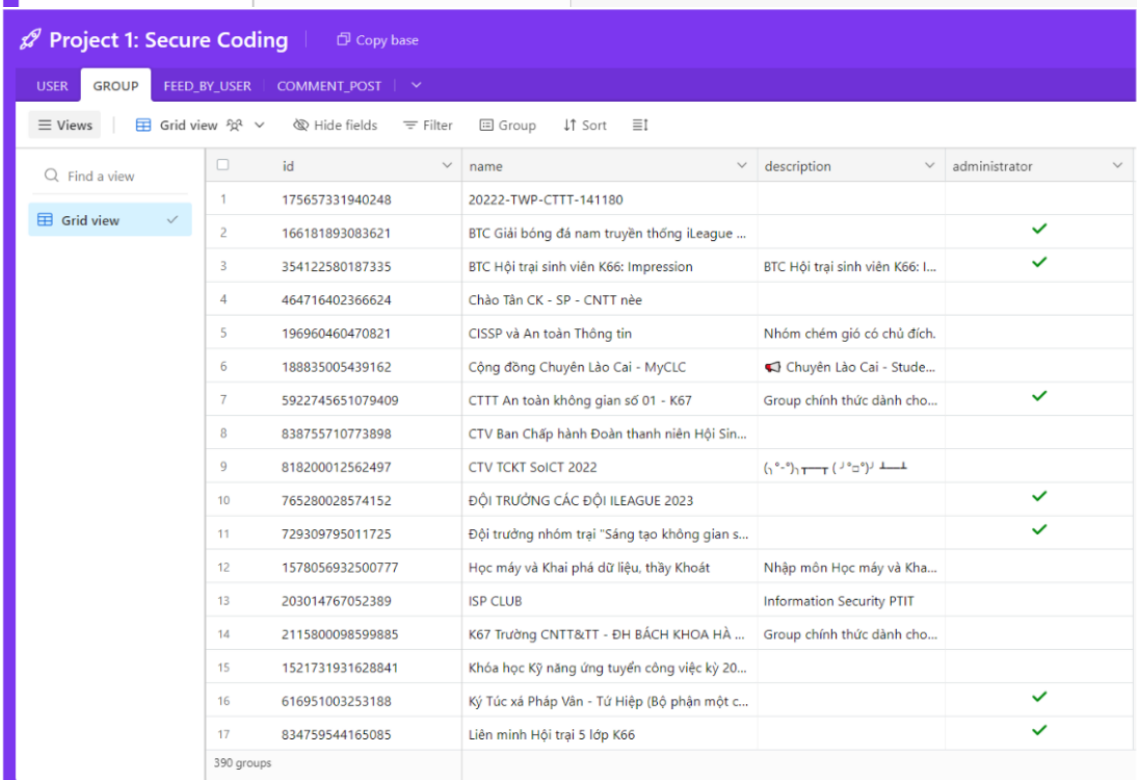


3. You can get a link to view the Airtable base when choosing this option.



When go to the link, result that we can see here:

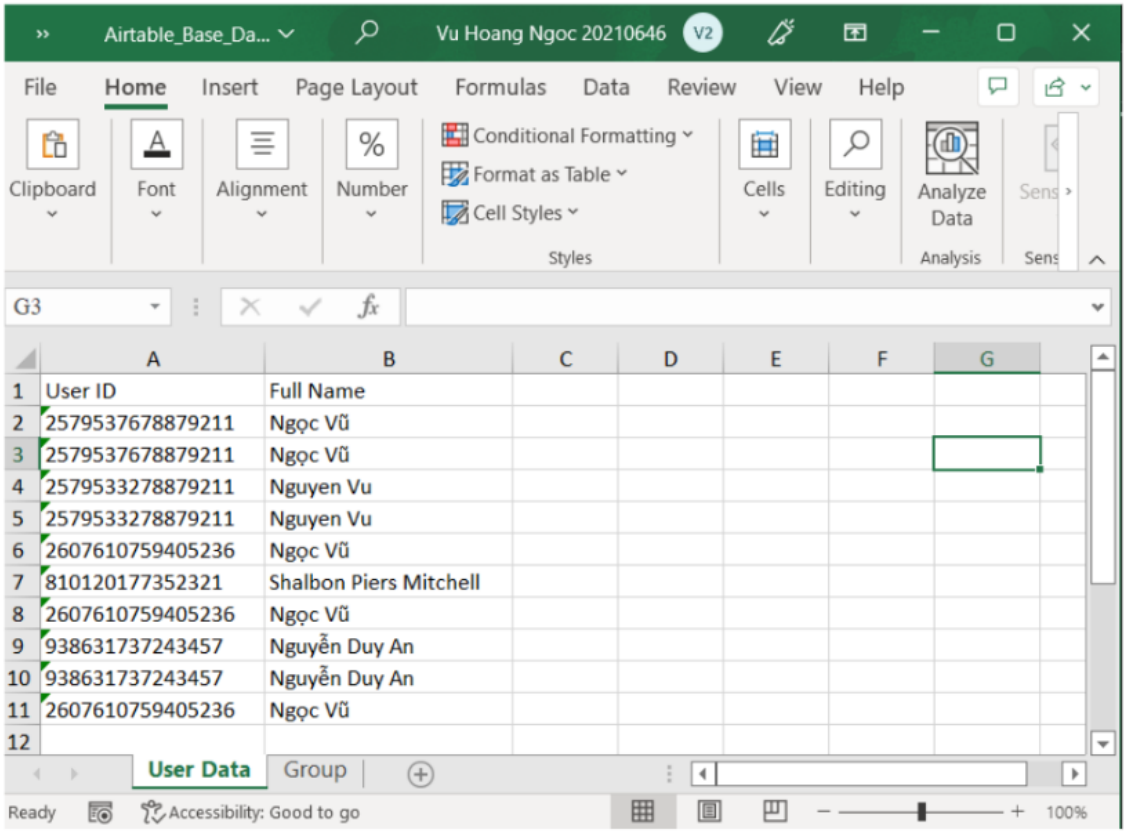


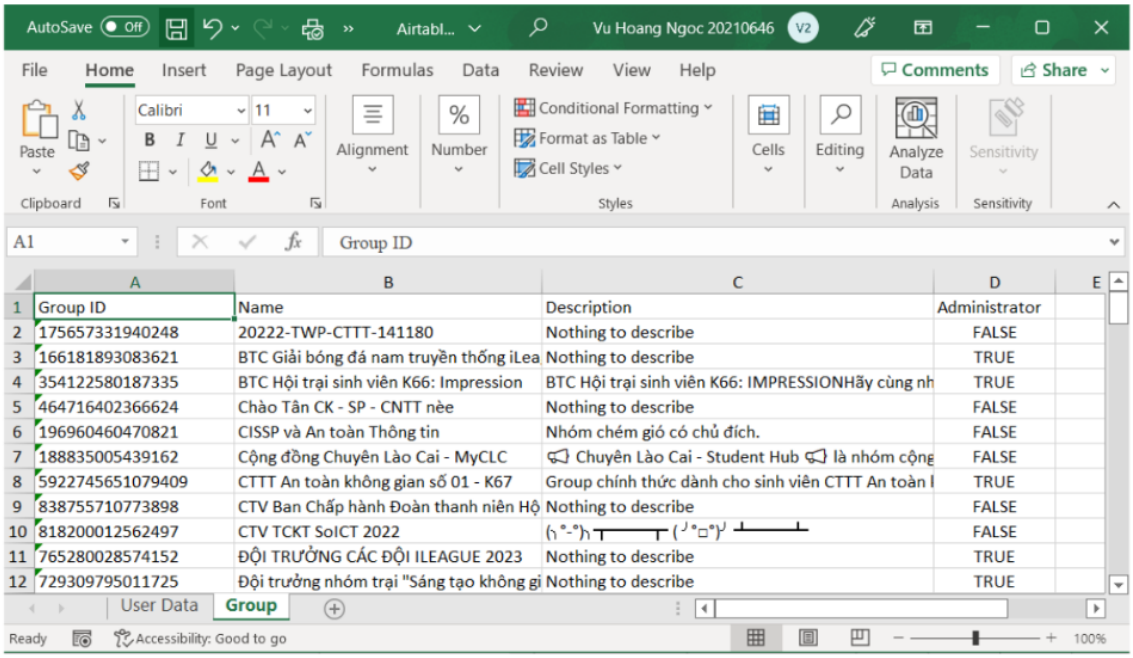


4.2.3 Statistic data with excel file and generate charts(Option 6, 7, 8, 9)

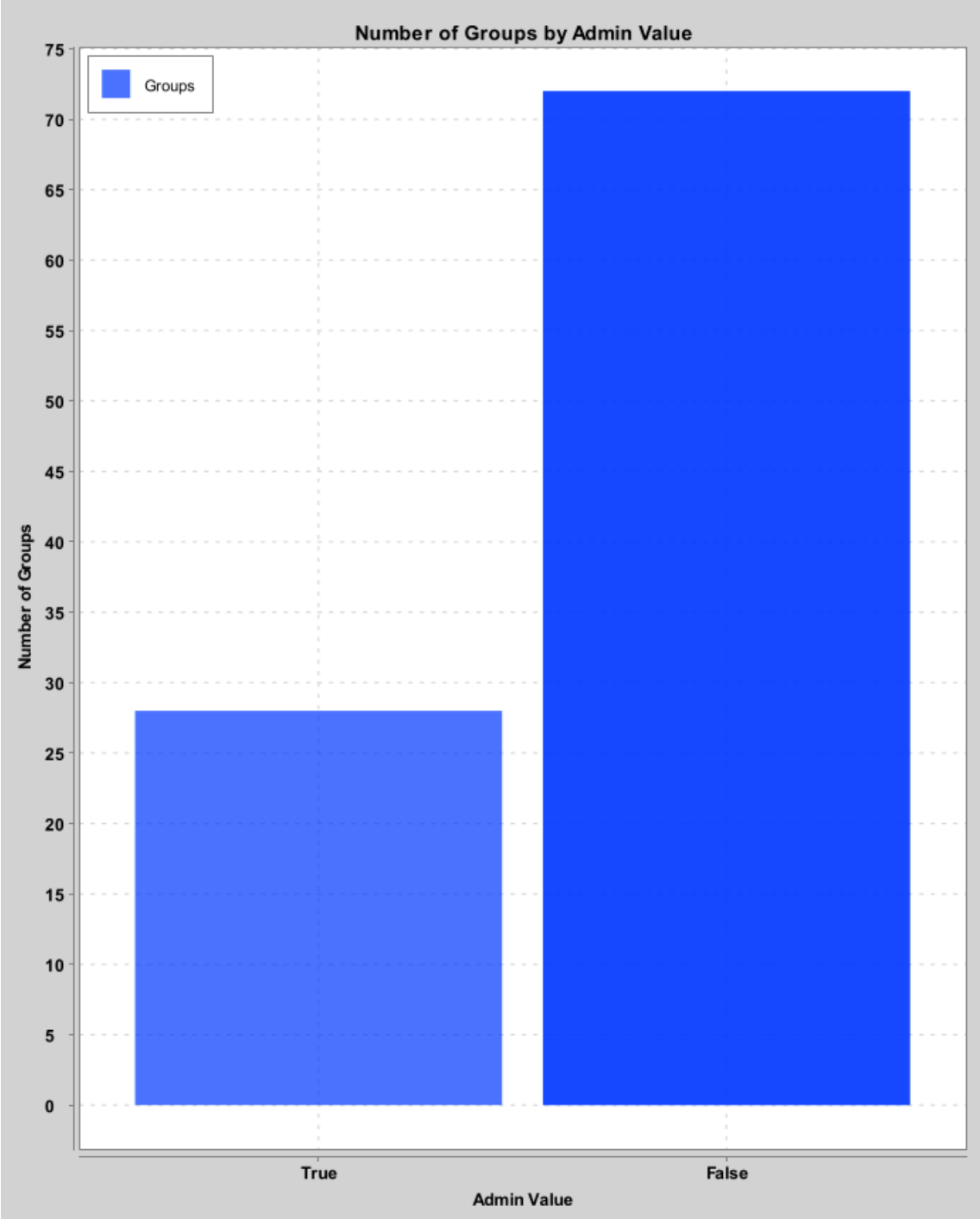
1. Option 6: Export the Airtable base from Option 5 to an Excel file.

```
--6. Create Excel file for data got from Airtable--
Check file Airtable_Base_Data.xlsx in the folder RESULT
```

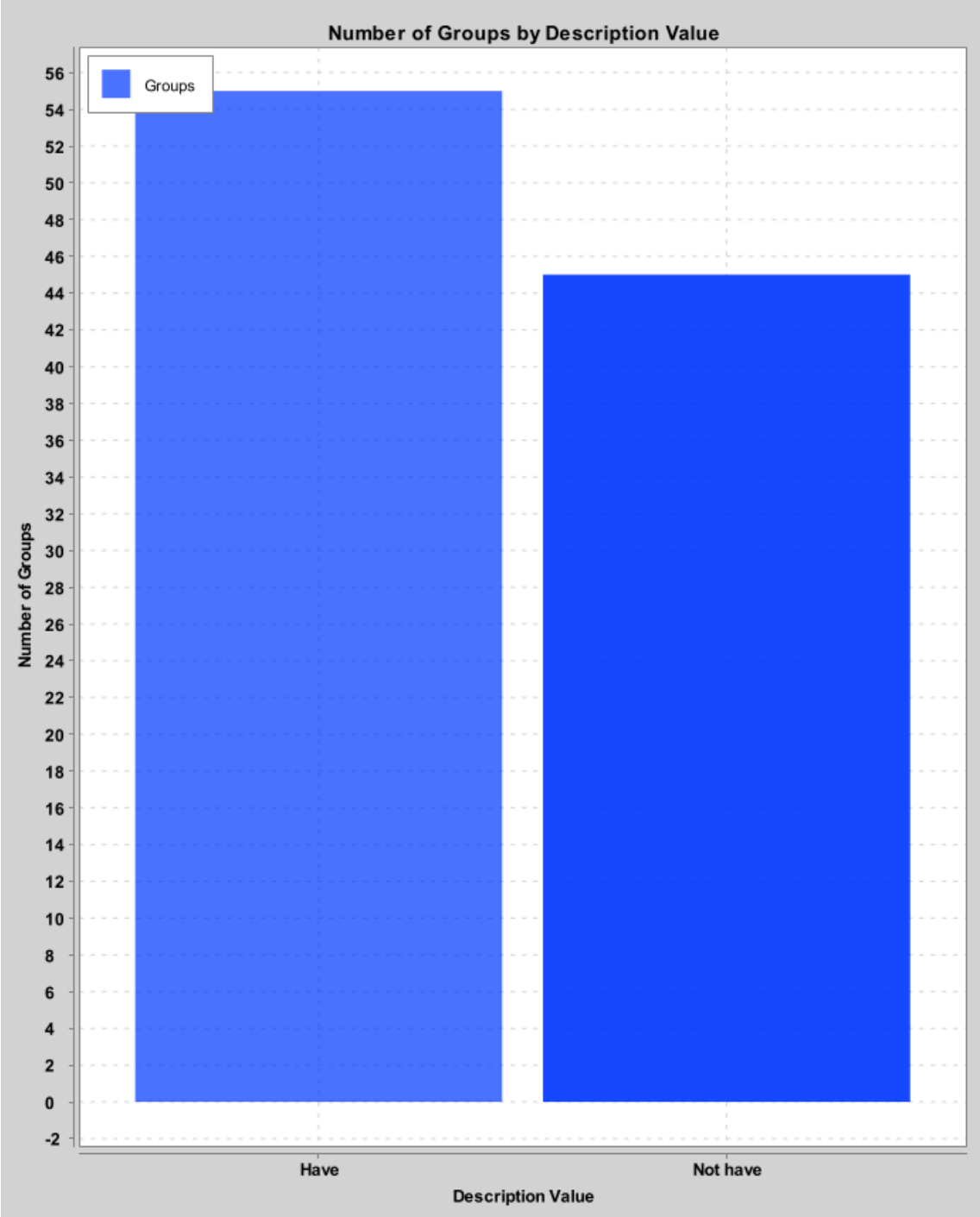




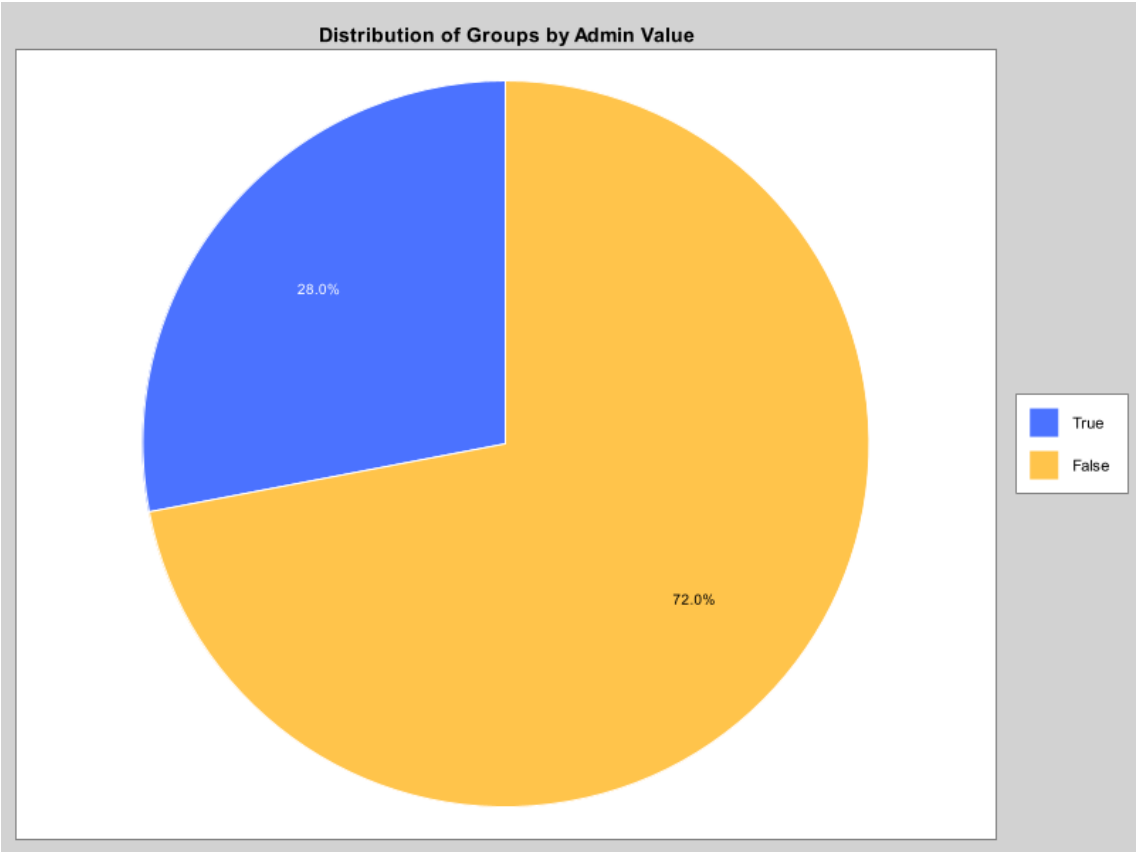
2. Option 7: Get the Group Administrator bar chart.



3. Option 8: Get the Group description count bar chart.



4. Option 9: Get the Group Administrator pie chart.



The result will include an Excel file and 3 charts that will be saved in the folder named "RESULT."

CHAPTER 5. CONCLUSION AND FUTURE WORK

5.1 Conclusion

After completing the Project 1, we have learned a lot about object-oriented programming, secure programming with Java, how to collect data on social networks, data analysis, and how to use Airtable to create an automatic database. In reality, we can use this software for the following purposes:

Collect information: The program can collect information from social networks such as Facebook, Twitter, LinkedIn, Instagram, and YouTube to help users obtain necessary information for research, market analysis, or managing interactions with customers.

Analyze data: The program can analyze the collected data to provide results and predictions, helping users understand trends or behaviors of customers.

Manage interactions: The program can help users manage interactions with customers on social networks, for example by responding to comments, interacting with shared content, or following interested users.

5.2 Future work

In the future, we can develop this software to collect data from even more social networks. From there, we can accurately analyze the psychology and trends of customers to keep up with the latest trends and achieve the most optimal communication effectiveness.