

User behaviour analysis: Banking mobile app

Yinhao Yang



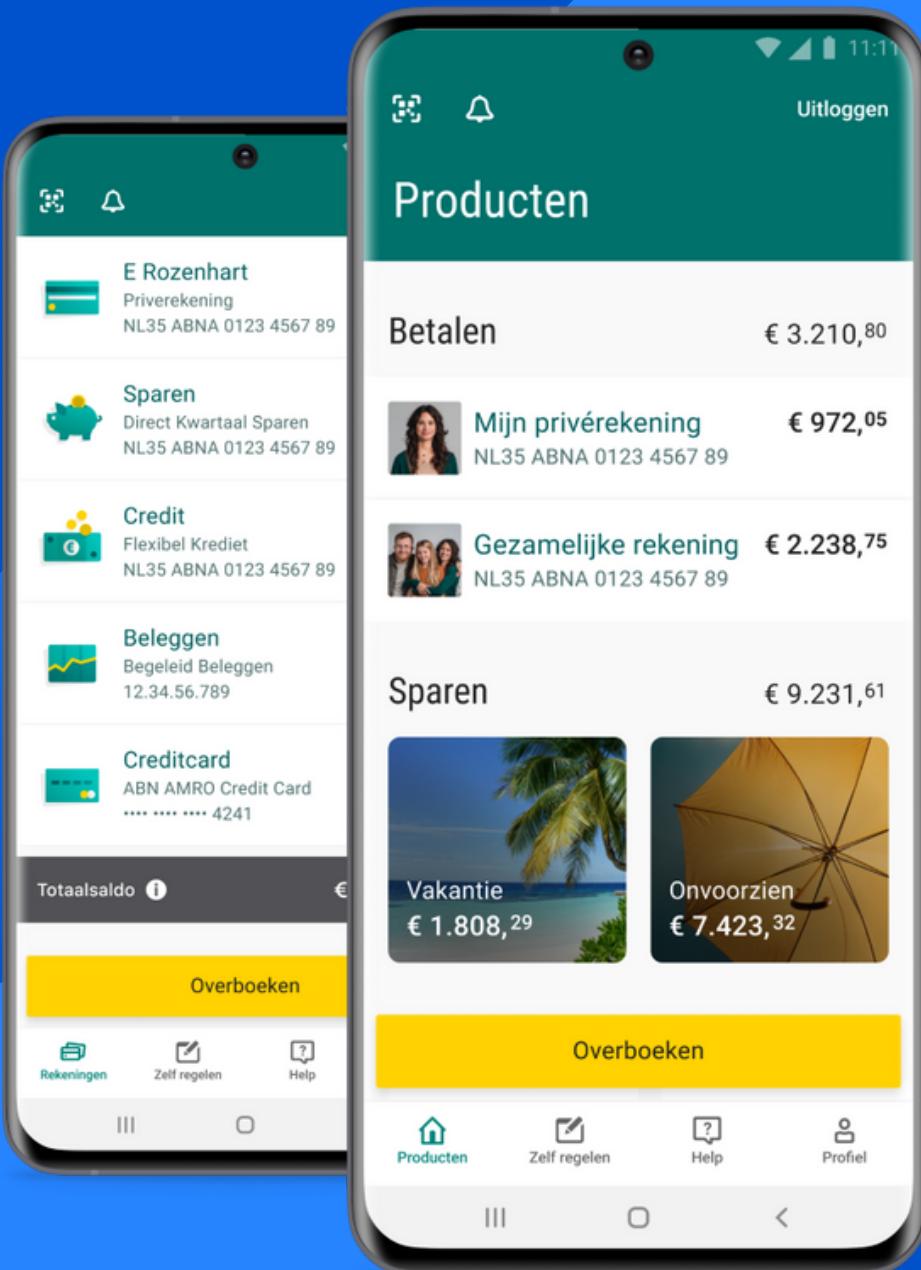
[HTTPS://GITHUB.COM/VERIONYANG/PROJECT-USER_USAGE_BANKING_APP](https://github.com/verionyang/project-user_usage_banking_app)

BASIC INFO

About the dataset:

24-h free trial before enrollment in paid services:

- 50000 rows * 14 columns
- The dataset captures the user behaviours of a banking mobile app.
- App behaviour is characterised:
 - Number of touch points
 - Whether they used premium services
 - Whether they played mini gams
- Their registration dates and enrollment dates are recorded as well.



OUR PURPOSE

We hope to understand the following questions:



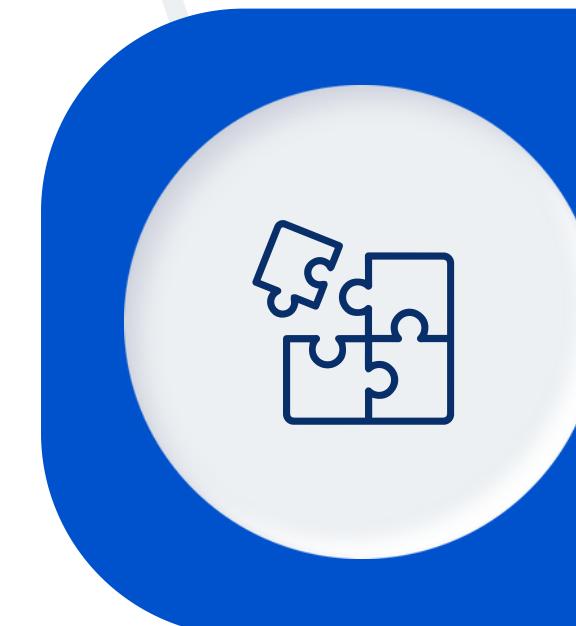
User patterns

- When do users use the app the most?
- Does age play an important role in user's app usage, mini games usage, and enrollment?



Decision-making

- What are the main factors that drive a user to sign up for the paid service?
- Does the implementation of the mini games help increase the likelihood of enrollment?



Functionality

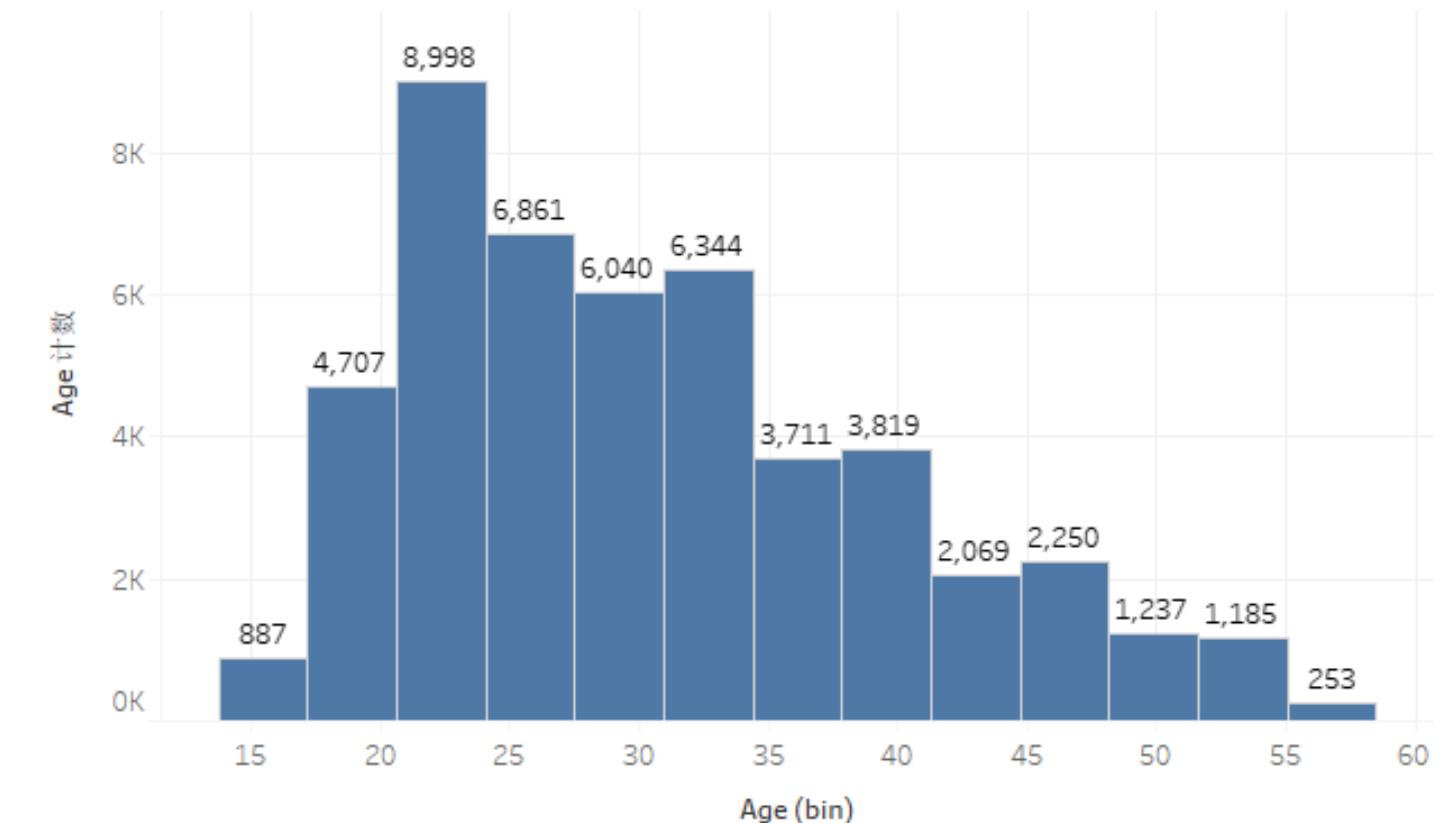
- What kinds of functions are appreciated by users and what are not?
- Among all the functions, what types of functions have positive impact on the likelihood of enrollment?

CONTINUOUS DATA

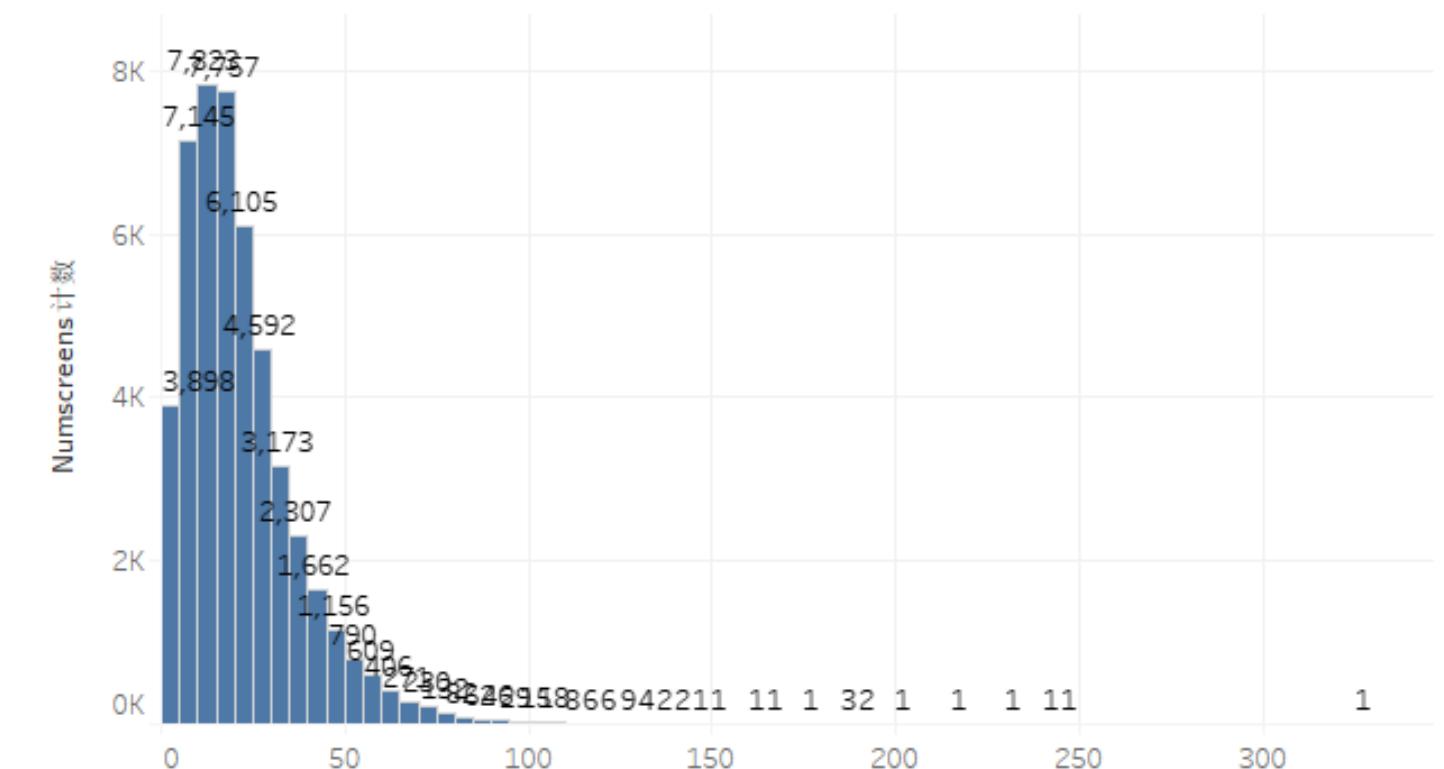
Age & number of touch points:

The dataset consists of 2 continuous numerical columns: age and number of touch points. They both follow a positively-skewed normal distribution.

Age Distribution



Num screens distribution



APP USEAGE PER MONTH,PER DAY, PER HOUR

User patterns: Enrolled VS Not

Split the date to 4 new columns: date, hour, month, year.

These plots are separated by whether the user has enrolled in the paid services.

As we can see, **Q2 of year 2, Thursday-Sunday, from late afternoon till midnight** are the peak moments in terms of app usage and user activities.

Lastly, one can easily tell that enrolled users are more active than those who have not.



AGE VS APP USGAGE

How age affects user patterns?

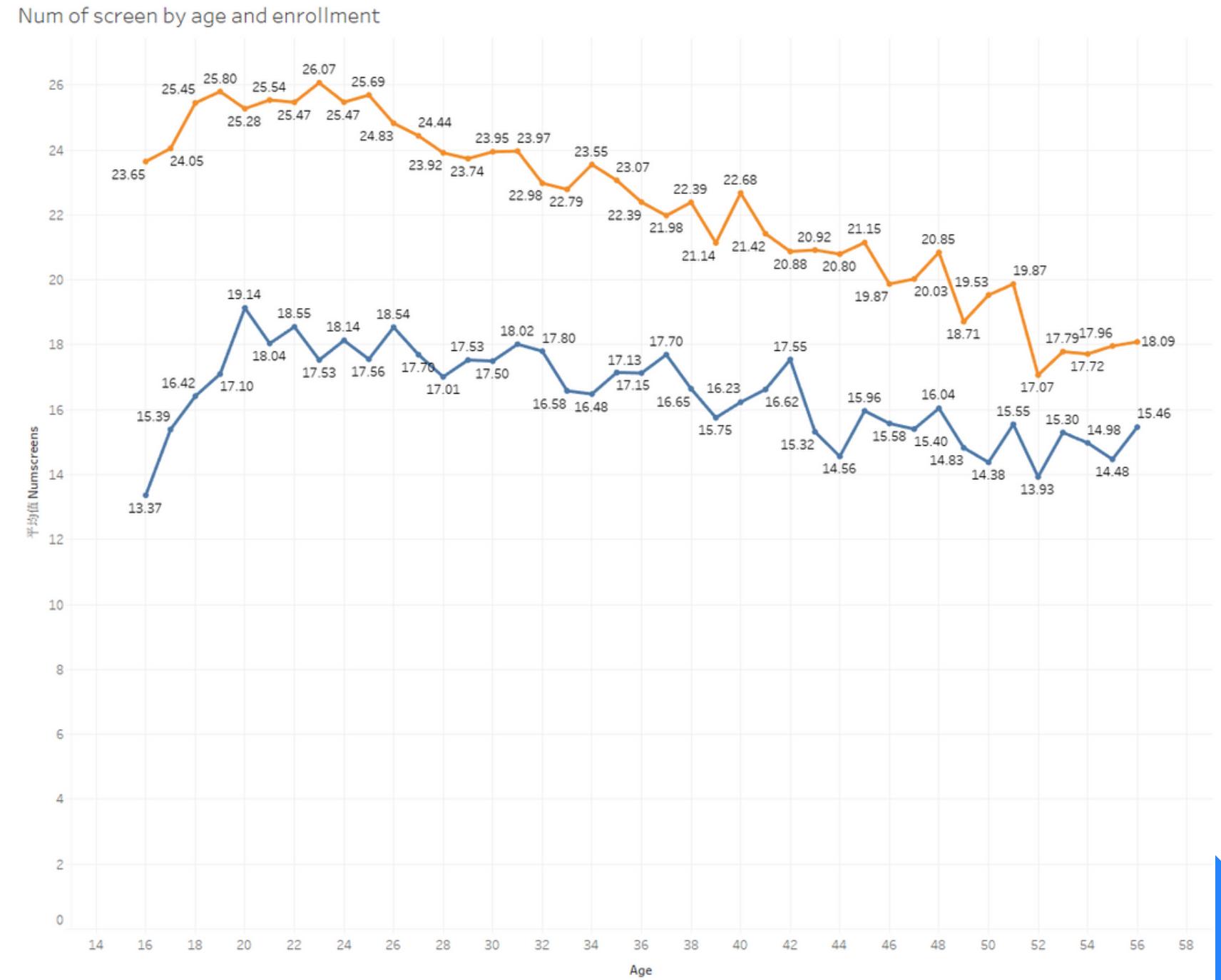
We are interested if age plays a significant role in user behaviour and user decision-making. So we will run 2 hypotheses testings.

The first hypothesis test is to test if age plays a role in number of screens a user has used.

- H0: People in different age have used same number of screens.
- H1: People in different age have significant differences in the number of screens used.

The second hypothesis is to test whether age has an impact on whether users have enrolled.

- H0: Age and enrollment are NOT correlated.
- H1: Age and enrollment are correlated.

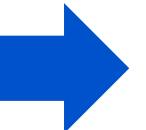


TOUCH POINTS

Clean up the screens

The original dataset records users' touch points as a list. We need to process the column for further analysis.

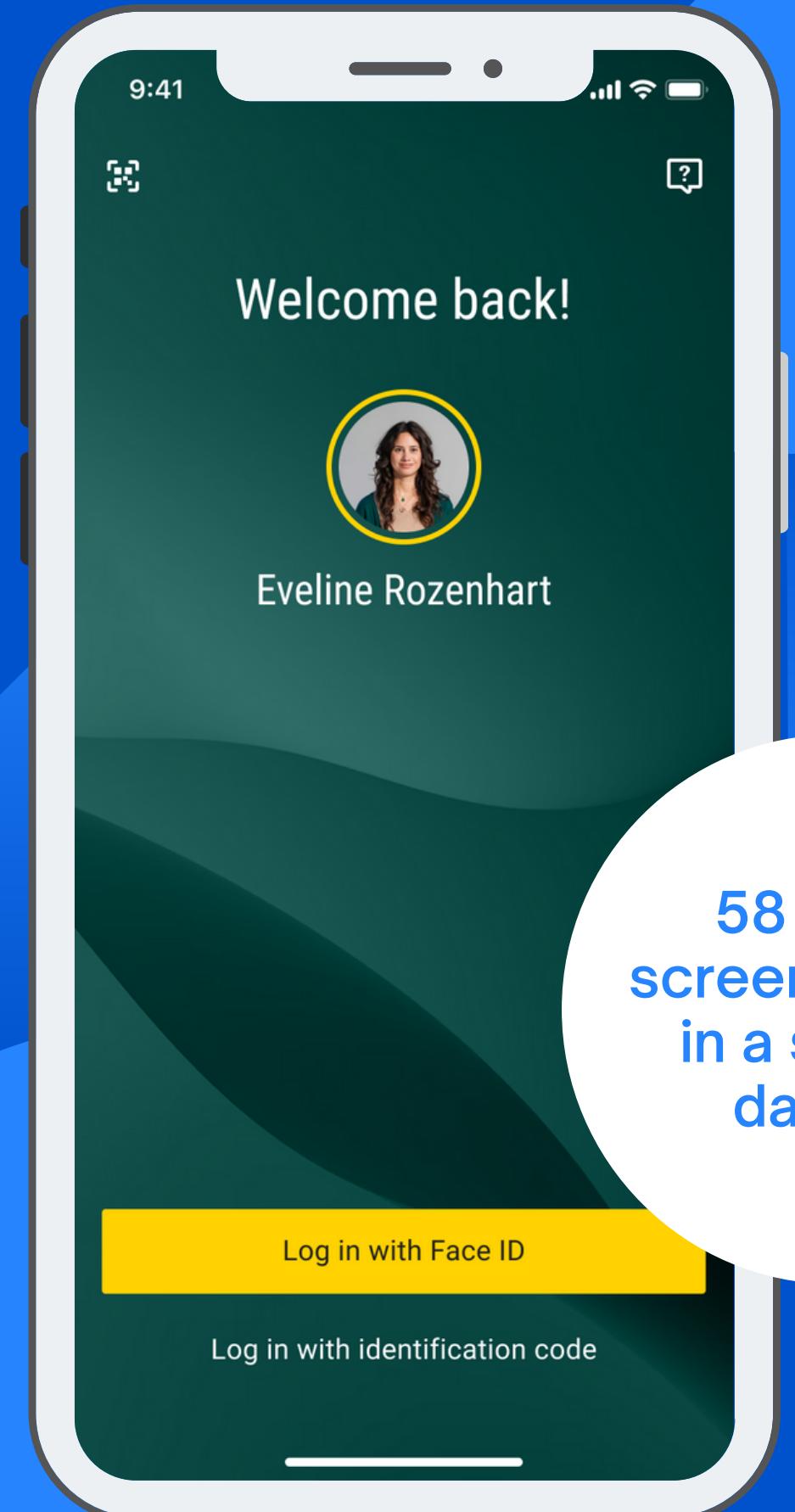
screen_list
idscreen,joinscreen,Cycle,product_review,ScanP...
joinscreen,product_review,product_review2,Scan...
Splash,Cycle,Loan
product_review,Home,product_review,Loan3,Finan...
idscreen,joinscreen,Cycle,Credit3Container,Sca...



- Group them by function:
1. Credit
 2. Loan
 3. Saving
 4. Basic
 5. Game
 6. Verification
 7. Profile

1 categorical column

7 dummy columns



**58 touch
screens stored
in a second
dataset**

DATA PROCESS

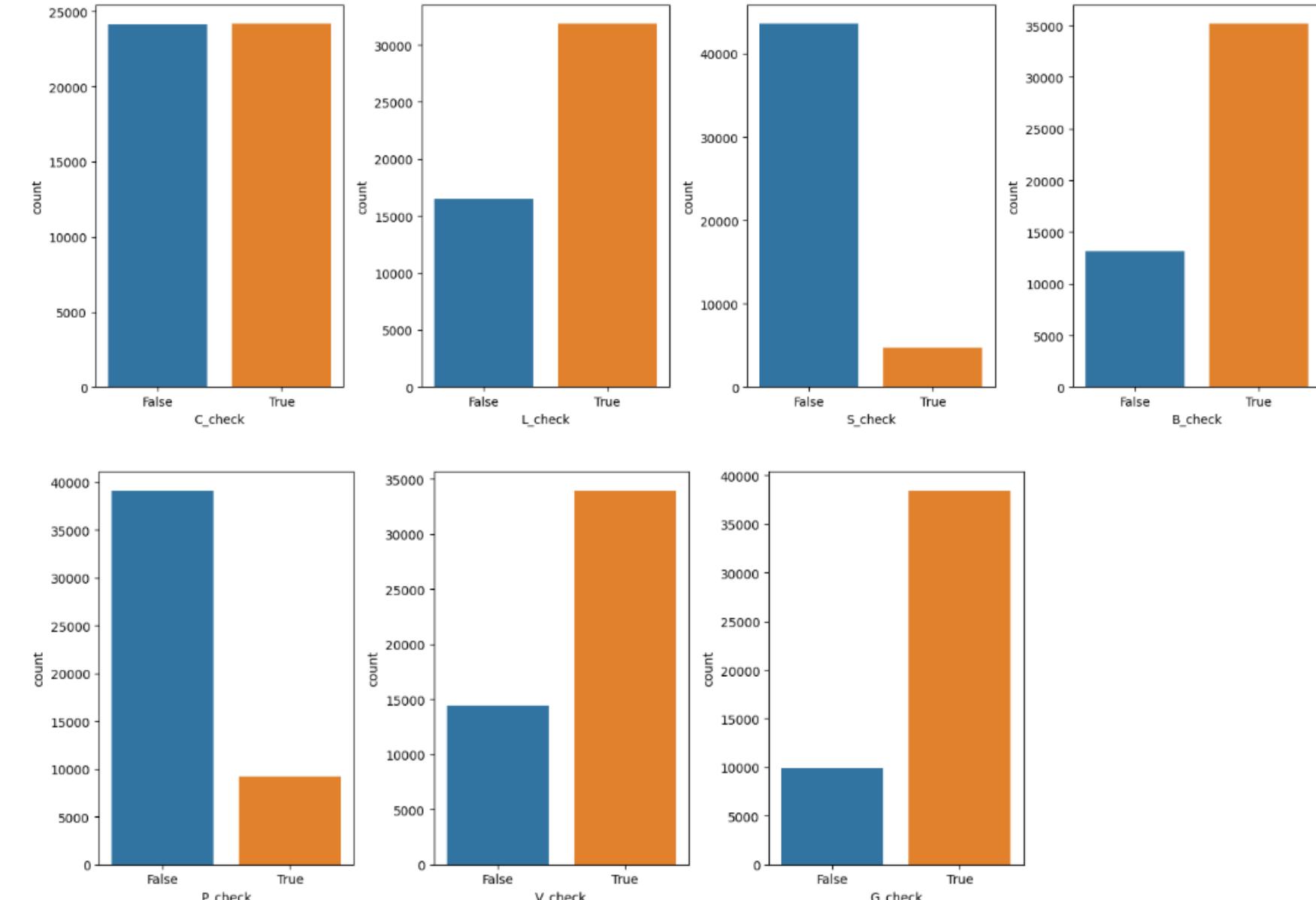
Touch points & enrolled hours

Plot the countplots of all 7 columns and drop 2 of the least used.

Subtract enrolled hours to registration hours.

75% of the users signed up within the first 3 hours.

The enrolled dates and hours columns are hence dropped.



MODELLING

Train,test,split

Process:

- Power transformer because of the positive skews in the continuous columns.
- No OneHotEncoder because the cleaned dataset doesn't carry any categorical columns
- MinMaxScaler

Results:

- Compared to other models, Logistic model performs the best.
- Downsampling due to data imbalance.
- The Cohen kappa score increases from 0.34 to 0.38.

Coefficient table	
first_open_hours	-0.295157
first_open_month	-0.298704
first_open_year	0.812956
dayofweek	0.078982
age	-0.595227
numscreens	2.989859
minigame	0.847566
used_premium_feature	-0.301069
liked	0.024321
C_check	0.344929
L_check	-0.679621
S_check	-0.463568
V_check	1.064628
G_check	-0.013839

CONCLUSION

Observations

From the coefficient table, we can draw the following conclusions:

- The app is more **favoured by younger people**.
- Number of touch points has a **direct impact on** whether or not a user will enrol in the paid version.
- The implementation of minigames does **increase the likelihood** of enrollment.
- Among all grouped functions, **Credit-related functions** and **Verification-related functions** are the biggest driving factors for enrollment.
- Loan functions and Saving functions both have a **negative impact** on the enrollments.
- The model suggests us that in **the middle of the week or weekends** are more positively related to the final enrollments.



CONCLUSION

Suggestion

Some managerial implications can be:

- Target the correct audience.
- Iterate and optimise the user interfaces, making it **more interactive** so the number of touch points per user increases.
- Given that this is a banking mobile app, the core functions such as **saving, loaning, and etc need immediate attention** for improvement. Conduct a more in-depth (qualitative) research on why users seem to be unsatisfied with these functions and update them as soon as possible.
- Pay more attention to the **timing of your running campaigns**.



CONCLUSION

Weaknesses

The analysis comes with the following limitations:

- The dataset doesn't come with more information on demographic info other than ages. Further studies can collect other info such as gender, races, and etc to further explore the differences.
- The list of top screens **are not self-explanatory**.
- Some are obvious but not all are clear. So the UI team should work on **a better classification** of different functions and conduct more research on how each grouped function affects users' decision-making.
- Other dimensions of the usage such as user's average time spent is not recorded.



Thank You



[HTTPS://GITHUB.COM/VERIONYANG/PROJECT_USER_USAGE_BANKING_APP](https://github.com/VERIONYANG/PROJECT_USER_USAGE_BANKING_APP)



Yiniao Yang