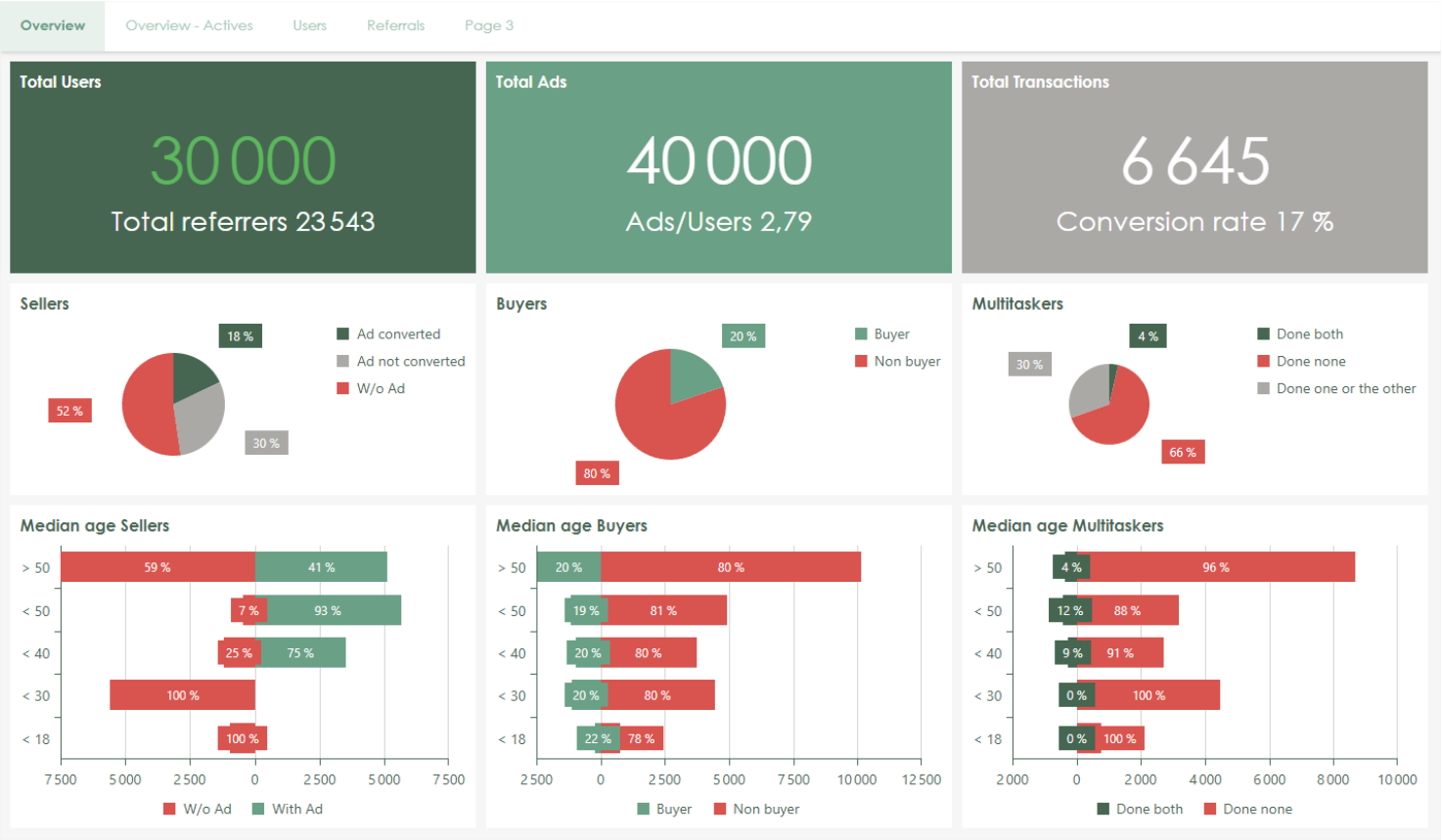
Data Analysis

As you read my observations, you will find **text in bold:** This is what I would have looked into further had I had more time.

Overview

The « overview » page gives a general idea of how the product is used and by whom.



A good proportion of users are referrers. That seems to work well.

In 4 years, only 2.8 ads per user who actually posted an ad is quite low. **Can it be boosted?**

Maybe also look into **how to appeal to the half of the users with no ads whatsoever**? Half is a lot.

The conversion rates of the ads is extremely low : less than 1 on 5. **How come? It's an important topic to look into, for the attractivity of the product.**

I looked into the ages relative to type of activity, and we can see that

- for selling it's more the "over 30 yo"

- for buying it's more evenly distributed, with a peak over 50 yo.

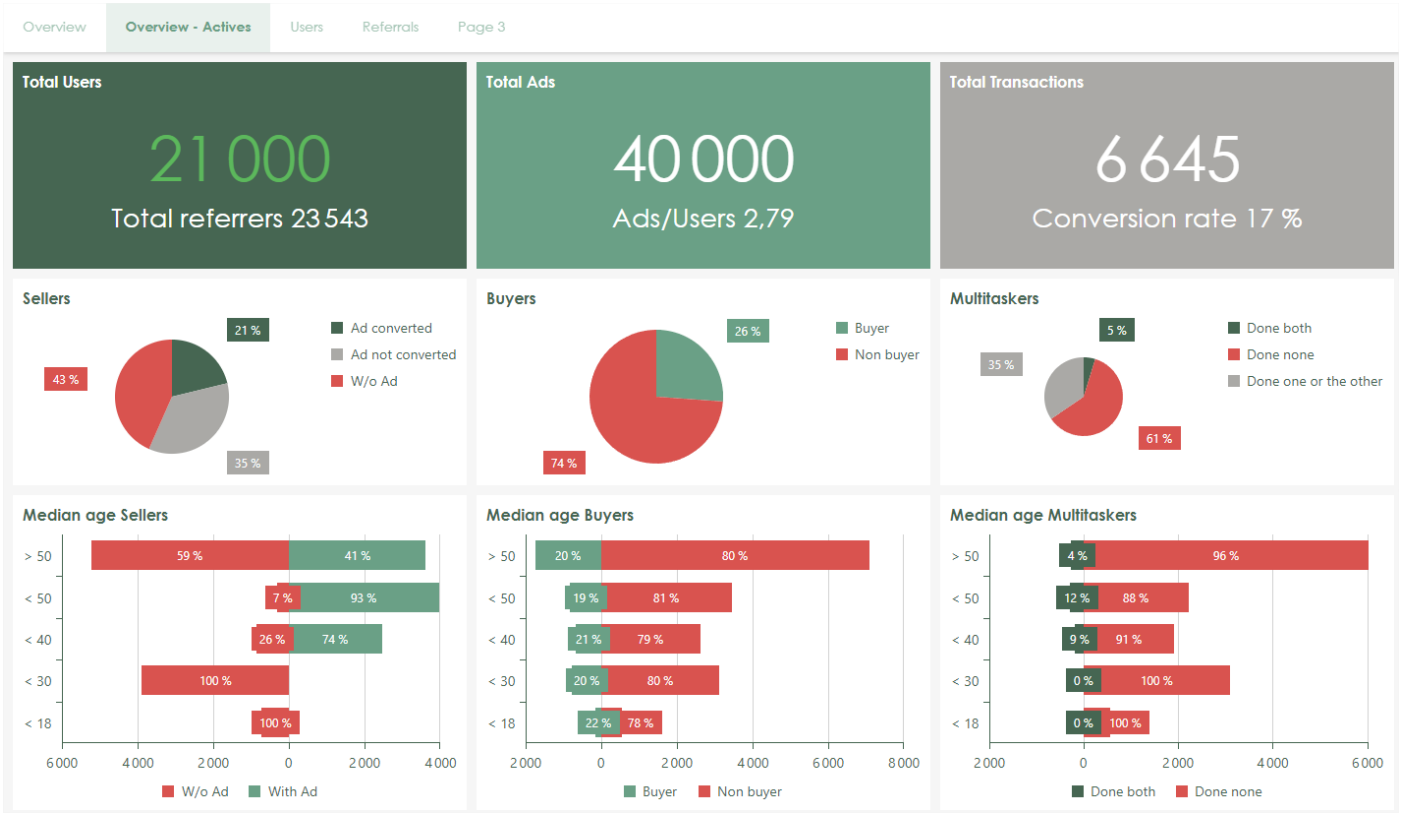
- and those most active in both at once are between 40 and 50 yo.

How to appeal to the younger generations?

Is referral an issue?

Overview – Actifs

Here is when I realized I had assumed, wrongly, that I was working only with active users. However a lot of the users did not have enough referrals to be active.
At least if we still consider for this dataset ad theoretical product that we need 3 referrals to be able to interact.
So I looked at the same data, but this time without those inactive users.
I kept both pages in this document because I find interesting to see both sets of results. More conclusions can be drawn that way.



A third of the users are not active. We can see that you can vouch for people without being active.

Users without ad go from 52% to 43%. Still a lot and worth looking into.

Users who never did a purchase go from 80% to 74%.

So for both 8-9% of total.

The distribution by age does not change sensibly.

Users

This page is a first analysis of the users.
At first I looked overall with age and sex, then I noticed the peak in women over 50 and looked at it a little closer.

No marked evolution as to ages at time of subscription or ages of users over time, nor of the proportion of women relative to men.

I would like to look into what appeals to younger generations, as they are fewer in numbers. Then again, with only 2 categories it might stay a little too basic.

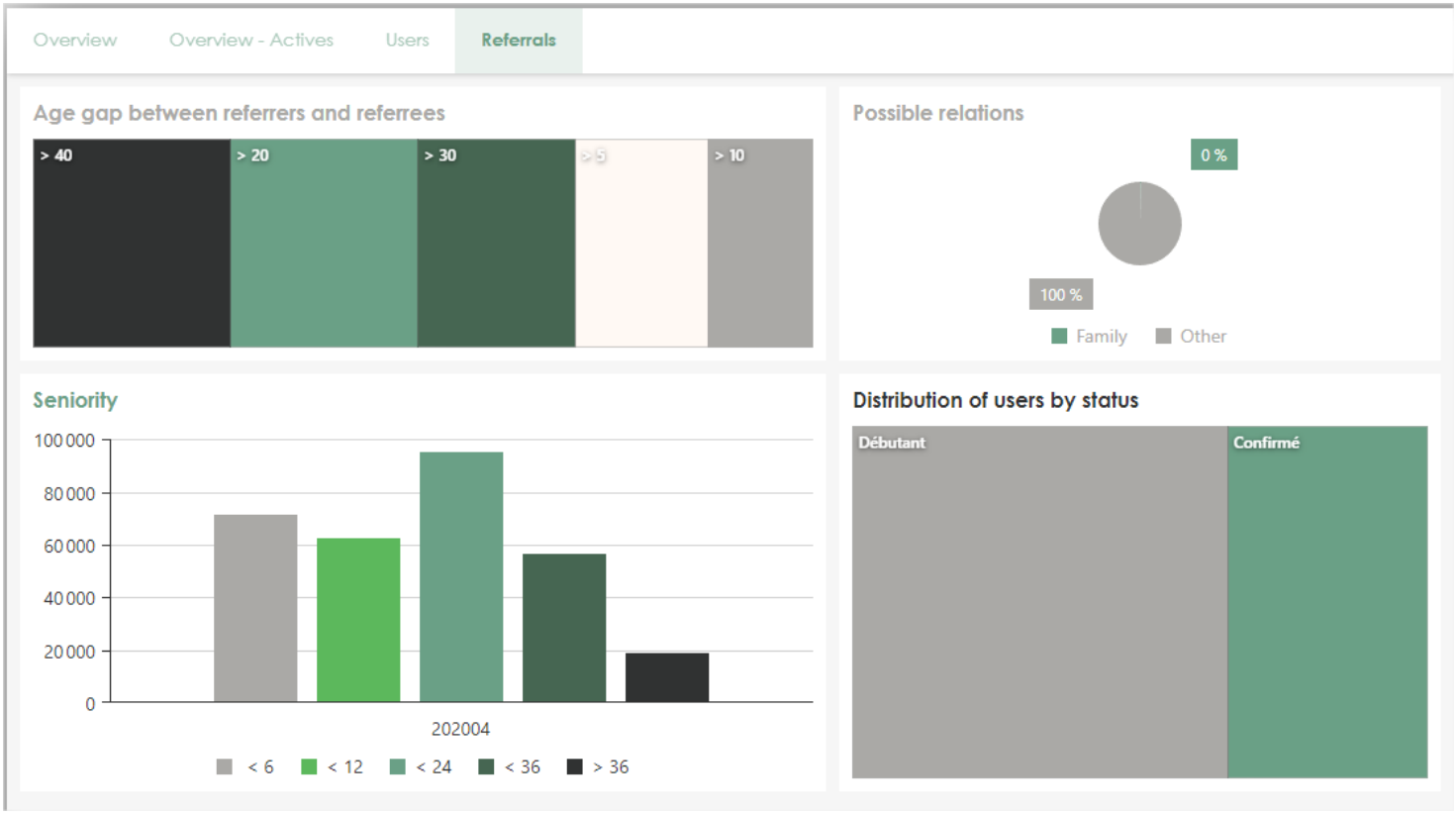
Good to note : Twice as many women as men.
No big surprise in woman wanting to feel more safe in who they deal with, so turning to a vouching solution in bigger numbers. **But maybe other reasons? Maybe what attracts them could be used in ad campaign to attract a wider public?**
There is a little loss of data with the N/A (again, how it is possible?) but 1 it's negligeable, and 2 it's probably proportional anyway.

Another point of note : most women are over 50 yo.
Worth looking into : cities or countryside?
What type of item? : fifty-fifty⁉ Fake data galore.
Why no more ads after 2019?
Evolution of subscription and their origins? Meh. I don't believe it. Too clean. I'd like to see the real data on those 2 points.

I would also add a map of the activity, overall and by category, and by year, and maybe by year/category if it feels relevant after a first enquiry.

Referrals

This page is a first analysis of the referrals.



Interesting how very progressive the number of referrals is with age difference. Does not seem very realistic.

The “Possible relations” graph is very rudimentary and does not bring any clarity.

The seniority shows that all referrals have been made during the same year and month. **Bug?**

The “Status” graph shines a light on how difficult it is to climb that ladder.

I would look into the average seniority of each status. And for each sub-category (see here) too.

Has it been set randomly? If not, how?