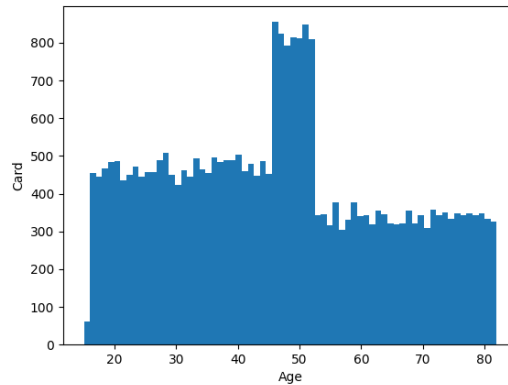# Gens de confiance

Data Product Analysis

1. AUDIENCE

a. Age is fairly distributed with a spike for people aged around 50.

The spike may be explained by the fact that people around the age of 50 are more prone to be interested in real estate. In any case this category could be a source of growth for real estate transactions.

b. Gender: some data missing

Sex	Total
M.	1490
Mr	1509
M	5908
F	18041
Missing	3052

We can note that there women account for more than half the population on the platform, even after repartition of the missing data. Does that mean that women are more trustworthy than men and find referrers more easily? We can also note an issue with data quality. We could improve data by transforming 'M.', 'M' & 'Mr' to 'Man'. We should also fix the site in this way and avoid missing data as this is valuable information to retarget ads for instance.

c. Location: 3 main cities

City	Total_users
Marseille	4901
Nantes	4900
Laval	4900

d. Devices

browser_family	#
Opera	6087
IE	5973
Chrome	4804
Safari	3980
Firefo	3572
Mobile Safari	2022
Firefox Mobile	1208
Chrome Mobile iOS	1198
Firefox iOS	1156

os_family	#
Windows	13402
Linux	5482
iOS	4376
Mac OS X	4360
Android	2380

device_family	#
Other	17699
Mac	4349
Generic Smartphone	3183
iPod	2022
iPad	1189
iPhone	1165
Generic Tablet	345
LG UG	48

This data can be communicated to the front end developer team to make sure UX and responsiveness are optimized on these main device configurations.

2. ACQUISITION

utm_source	COUNT(*)
facebook	14991
mailing	9024
(missing or organic)	5985

utm_medium	COUNT(*)
ad	6823
emails	11497
social	4714
(missing or organic)	6966

utm_campaign	COUNT(*)
referral.referBack	6079
ad.recap	6013
aide-coronavirus	5990
message.reply	5976
(missing or organic)	5942

The data encourages us to pursue campaigns as 80% of users arrived on the platform from campaigns. E-mail is the most efficient medium for these campaigns.

3. CONVERSION / VALUE CREATION

a. Vehicles, an untapped source of revenues?

Category	Number_of_ads
realestate	14087
REAL_ESTATE	3945
relestate	1968
vehicle	20000

Category	Turnover
real estate	861 060 533 €
vehicle	24 958 237 €

Category	Average_sold_price
real estate	265 000 €
vehicle	7 401 €

We can note 2 main categories in ads: 50% are for real estate and 50% for vehicles (in number of ads). Today users only have to pay for real estate ads. Vehicles is maybe an untapped source of revenue for GDC as it accounts for 50% of its ads (and users pay per ad) and users may be willing to pay for them because they create value from them (7400€ on average).

WARNING: As for 'Sex', we can note an issue with data quality. We should fix the different spelling both in production and in data cleaning to improve queries. There are also some strangely low prices for real estate transactions (less than 100 euros) that should be investigated to improve the precision of above figures.

b. GDC creating value sellers

category	on_price_or_below_price	above_price
real estate	1669	1604
vehicle	1700	1672

We can see that half of the time sellers get a better price than asked for ! This could be advertised to entice more people to sell on GDC.

4. Further analysis

Other analysis with available data could tackle:

- Links between referrers and referees. Are there influencers ? How to use new users to bring more people to the platform. Which users are more involved in referrals ? Is there a main profile ?
- Profile of sellers and buyers: mainly men, women, age …?
- Time series data to display trends.