Background

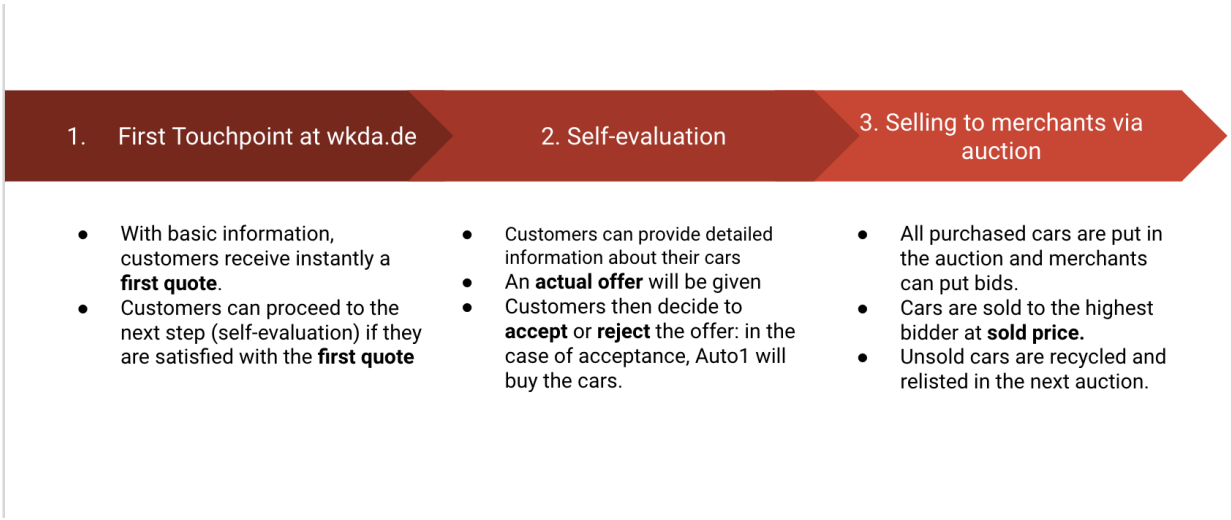
There are 3 main business pillars at the Auto1 group

- Autohero (consumer-to-consumer)
- Remarketing (merchant-to-merchant)
- Auto1.com (consumer-to-merchant)

Each of the pillars has its own business opportunities, characteristics and challenges. The role of a data scientist is to support the company to take advantage of the vast amount of data and make data-driven business decisions.

Challenge Description

Auto1 purchases used cars from consumers (under https://www.wirkaufendeinauto.de/ or wkda.de) and sells them to merchants. The following flow chart provides a simplified overview of the funnel.



Please complete:

- Task 1
- ***Either*** Task 2a ***or*** Task 2b (see Instructions for details)

using the data attached with the email.

Task 1

(20-60 mins)

We have attached 3 separate csv files, each of them containing data from each of the respective steps in the above flow chart. There is a common column `id` in each of the files to relate/link the data in the files.

First quote: it is an indication price but ***non-binding***
Actual offer: it is an actual offer and is ***binding***

Here we would like to ask you to identify one key performance indicator (KPI) at ***each*** step in the funnel. For *each KPI*, please ***briefly*** describe and show the KPI. Also explain why the chosen KPI is important for measuring the success of the Auto1 business model.

Task 2a

(45-90 mins)
This morning you received a call from the local MD: it turns out, since early 2021, there are very few customers *starting* the self-evaluation (see step 2 in the flow chart above).

You will need to

i) investigate the root cause of this observation and provide the explanation for why so few customers are starting self-evaluation
ii) prototype a solution to ***fix*** this issue (few customers starting self-evaluation) that can be implemented ***as quickly as possible*** (**a simple** model would be sufficient).

Task 2b

(45-90 mins)
During a regular meeting, one of your team members has the following proposal:

> `We should boost the conversion rate of self-evaluation by increasing the "first quote". The rationale is since "first quote" is non-binding, it will only benefit conversion without any drawbacks.`

Do you agree with this proposal? Please

- explain your answer and use the provided data to support your argument
- build a **simple** prediction model to pinpoint the main driver for conversion. What strategies, in terms of pricing (both first quote and actual offer), would improve conversion?

Instruction for tasks

- The whole challenge should approximately take between 1 - 5 hours
  (time for cosmetic changes or prettifying your solution is not taken into account and is purely optional)

First quote: it is an indication price but ***non-binding***
Actual offer: it is an actual offer and is ***binding***

- Please use the ***second*** character of your ***last/family name*** to decide whether you should complete Task 2a or Task 2b:
    - [A-M]: Task 2a
    - [N-Z]: Task 2b
    - Non English alphabet: Task 2b
    - If your family/last name has only 1 character: Task 2a
- Please keep your answer brief. Feel free to use a combination of
    - Text document (Microsoft Word/powerpoint/google doc/google presentation)
    - Visual illustrations
    - Jupyter notebook (please export the notebook to html) / python script if necessary
- Please put all materials in a folder named with your first and last name, zip/tar the folder and send it back.

Suggestions

- Answer *concisely* so it is easy to follow your arguments and thought processes
- Illustrate/support your answers with visual aids
- Write clean and easy-to-understand codes:
    - Use meaningful variables names
    - Wrap your codes into functions with meaningful names
    - Supplement your codes with comments when necessary

Data attachment

Data for challenge
(Links only work internally, candidates please use the attached file in your email)