boberdoo.com Data Science Developer

*Abc LLC is one of our clients providing health insurance services. They bid on **pings** through the boberdoo system. Pings are partial information about potential customers that usually consist of state, gender, age and estimated household income. If their offered bid is the highest, they receive the full **post** information: phone number, email and address. In such cases Abc LLC contacts the lead and if the lead accepts Abc LLC's offer, they sign a contract with them to receive health insurance services for a monthly premium. After this they become **closed leads**.*

Sometimes Abc LLC buys leads that they cannot close if their offer is not competitive enough.

Please find two tables attached: **leads_hw** and **bids_hw**.

Leads_hw is an exported table from Abc LLC's data warehouse that contains lead information:

- leadID: unique ID of the lead
- state_18: the state where the lead lives in
- gender_18: lead's gender, which is 1 if male, and 2 if female
- estimated_household_income_18: this is the lead's estimation of their own household income per year in dollars
- cpa_status_18: 1 for closed leads, 0 for lost leads
- premium_amount_18: value of the contract signed, to be paid monthly in dollars

Bids_hw contains the bids offered to each incoming ping.

- created: time of the bid
- leadID: unique ID of lead
- max_bid: bid offered by Abc LLC for that particular lead
- won: 1 if their bid was the highest and they won the bid, 0 otherwise

Questions:

1. Considering the information on closed leads, how would you describe Abc LLC's current most typical client?

2. Considering the information on closed leads, what customer segments should Abc LLC target to maximize their income?

3. How would you describe the current market potential for the segments from Question 2?

4. How much more should Abc LLC bid on average on the segments from Question 2 to increase their income by 30%? Would the investment pay off?

Note:

Please save your queries / functions / visuals to be able to show them later. Present the code for your queries to show how you got to your answers. Do not hesitate to contact us should you have any question on the given data or definitions.