What identifies high-income people?

An analysis of Census data

Langyi Tian

Jan 2020

Executive Summary

We want to target our marketing/service to higher income people.

Taking the data, what do we know about them? what can we do to identify them?

Summary of findings:

- Exploratory analysis at group level

Characteristics such as white skin, highly educated, prestigous occupation (manager, professional, etc.) and being married are identified to be linked with higher income group.

- Individual-level modeling

Male and nation of origin also emerge as foundamental determinants contributing to higher income

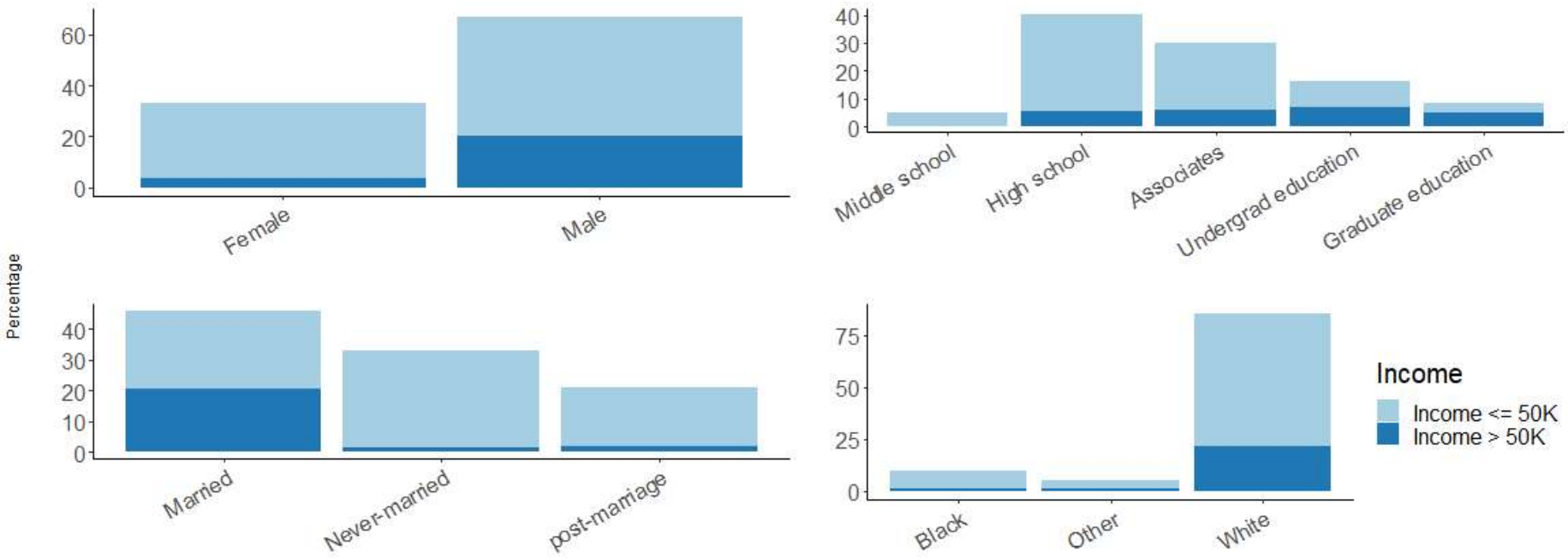
- Identify them in the future

To address future need to identify high-income people, we tested a machine learning algorithm and find it more challenging to predict the higher-income than the lower-income.

- Know more about the targeted segment

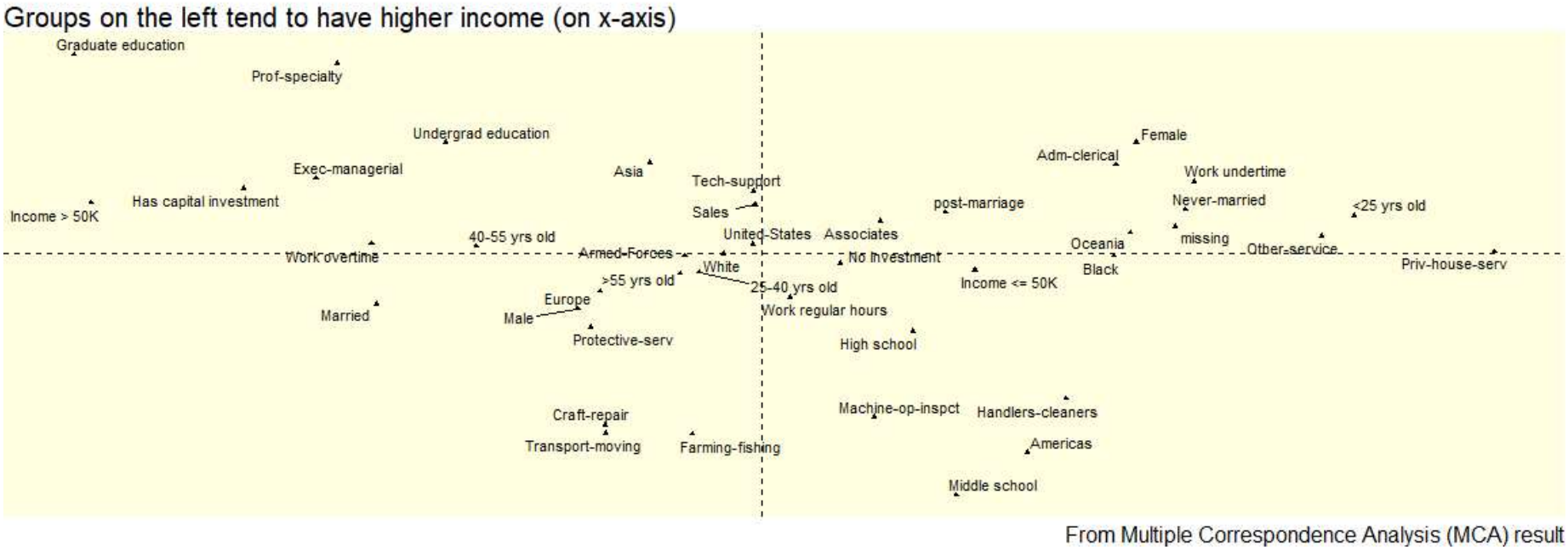
To understand specific consumer tribes, we mined 7 tribes with differnt likelihood to have high income.

Exploratory analysis: specific groups to target



- There are more high-income people among male, the educated, married families, and white people.
- At a rough level, targeting at these groups will be more effecitve than reaching everyone
- However, only observing by one dimension is too rough. e.g., how to compare the advantage of white people and college graduates?
- The next page gives an overview from multiple dimensions.

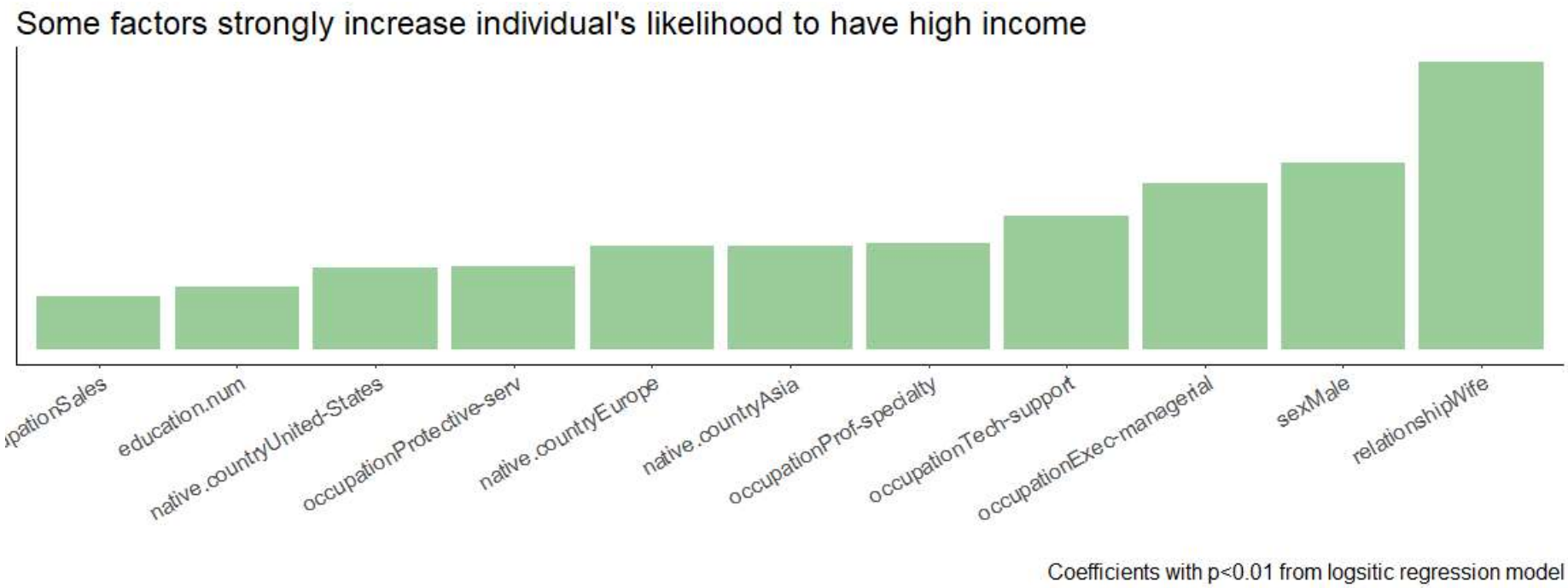
Exploratory analysis: a more comprehensive overview



From Multiple Correspondence Analysis (MCA) result

- A number of groups consist of married families, college graduates, senior people (40-55 yrs old), and people with good occupations (managers, professionals). People there are likely to have high income--much more likely than male or white people as a whole.
- But some people are both male and managers. Is it being male or being a manager (or both) that makes his money? How do these two factors compare?

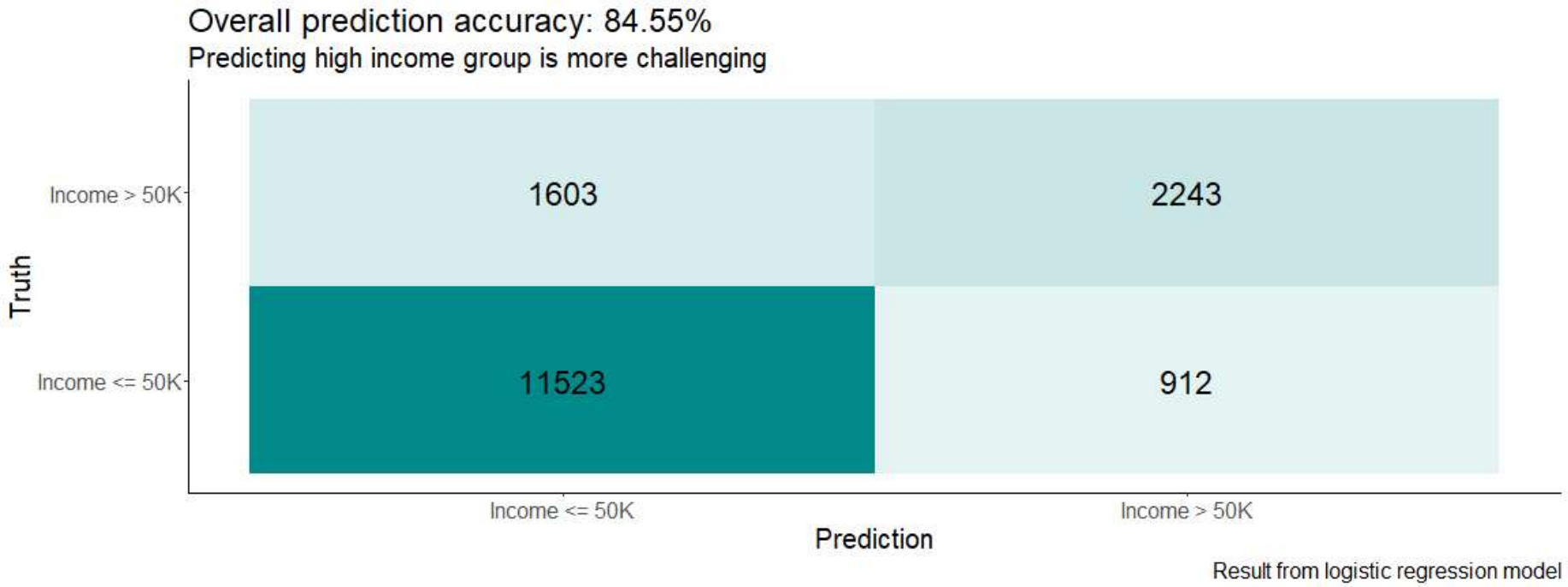
Individual-level modeling: identify influential factors



- Consistent with findings at the group level, being in marriage, having a good occupation and education will create more chance to have high income
- Being a male and originating from Asia or Europe actually help a lot more than what seems at group level. Probably, these origin help people get better education, occupation, etc.

Identify them in the future: machine learning prediction

- What if we want to identify high income people from our future data without knowing their income?
- Machine learning will help us to predict. This is how well we can do at a preliminary stage.



- Looks like it's more challenging to find people with higher income than lower income, but surely they are of higher consumer value. Keep improving!

Know more about them

- Now we know by what attributes/group membership can we find high income people, and how well we can find them.
- Finally, what if we want to design a detailed marketing message to some specific consumer segments/tribes?
- Let' say, we'd like to not only find those with higher income but also know other group-level characteristics such as demographics and occupation, so that we can tailor our strategy?

So, we mined 7 segments in the population that help us understand who we can target and what are they like.

Know more about them: 7 consumer segments

Show 10 entries Search:

Consumer segment profiles: Cluster medoids with predicted probability of high income

Sex	Age	Marriage	Race	Education level	Job	% Probabiliy of >50k income	Segment % in sample
Male	46	Married	Other	13	Adm-clerical	53	6
Male	38	Married	White	10	Protective-serv	36	20
Male	36	Married	White	9	Transport-moving	21	18
Male	30	Never-married	Black	10	Protective-serv	6	22
Male	48	Never-married	White	10	missing	3	14
Female	59	post-marriage	White	9	Other-service	2	3

Going forward: applications

- Identification of high-income individual with machine learning
- Design targeted marketing messages to specific segments
- ...

THANK YOU!