5. Sampling: Explain possible sampling bias through Weibo, such as gender bias, etc.

The most straightforward bias from weibo is the nationality, because only Asian people use weibo whereas the users in other countries seldom use this platform. So if we only sample the data from weibo, this will lead to nationality bias to the brand.

The other one is age bias. Even this kind of social network is easy for young people. However, it is a little bit complicated for our parents or even grandmother or grandfather. So we are not able to extract the information from them through weibo. And usually the people in this generation have more capacity to buy the product as compared with young people. This will lead to a serious bias for the brand.

6. What are some possible algorithms to identify users who showed interest in Michael Kors over Kate Spade? What are some data points that can be used to illustrate the algorithm's utility? Give a couple examples and discuss pros and cons.