

Final Report: Statistical Modeling and Analysis Results for Ticketmaster

(Team: Black Group)

Introduction: We generated large useful sample data from the given datasets and worked on three main databases. The aim of the analysis includes identifying the scalpers, discovering one-time buyer and “true fans”, using probabilistic selling to implement price discrimination to high-price-elasticity buyers for profit maximization.

Languages: Subtracting the sample data with Python and data analysis with R language.

User segments:

Finding scalpers: From the user behavior data, we analyzed totals_pageviews and totals_timeonsites and found the tendency for scalpers to spend more time on the sites and clicks for certain pages. From the purchase data, we chose the distance to venue, tickets purchased quantity, the time period between the purchase and the event date with the user id to distinguish the one-time user and “true fans” (or scalpers). Our hypothesis was that reseller datas are most likely to be abnormal. We implemented the analysis by fitting these two variables into a ***multivariate Gaussian distribution and calculated Mahalanobis distance*** from the points to the Gaussian distribution. After plotting a 2D graph and labelled those scattered dots, we can identify those outliers in the data, who are more likely to be resellers.

Remote buyer: We fit the dist_to_ven data into Gaussian distribution and identify those with large z-score as remote buyers based on dist_to_ven.

One time buyer: We discovered one-time buyer based on uniqueness of purch_party_lkup_id in purchase table. We have also discovered one-time buyers with short distance from the event.

Area: we categorize users using geonetwork_region and the polt bar chart. The y-axis is total number of hits

and this can be useful for the future pricing optimization because this user group have higher-elasticity and price promotions work better on them. For true fans with multi-times consumption, they have lower elasticity and we can charge more on them.

Marketing Plan:

It's beneficial that we do something to reduce the number of scalpers. As we find out a way to detect scalpers, we can raise the price of the ticket for scalpers to reduce their profit and achieve this goal. And we can provide a reasonable discount for remote buyers because they are more likely to be sensitive to price. Furthermore, as we find out that over 90 percent of the buyers are one-time buyers, we concluded that most of the buyers in the market are probably not “crazy” fans who have particular favor in a certain field. We referred to several papers and find out that we can add probabilistic selling to our marketing plan. “Probabilistic good” is an offer involving a probability of getting any one of a set of multiple distinct items. We separate out those one-time buyers who are closed to the events (ranked 0.1% in distance). We can provide probabilistic choice. For example, a ticket for either Saturday concert or Sunday concert with lower price (especially for less popular events) but the buyers should accept whichever the seller assigns to them. For those one-time buyers who prefer to buy in advance and are insensitive to price, we can provide another probabilistic choice with a set of multiple events in advance but the price will be higher so that buyers can choose to redeem whichever ticket several days before the event.