PROPOSAL BY GROUP 5

Make Full Use of 12306 and Guarantee You a Convenient Life

Na Su, Weijia Mi, Qiutong Dong

- January 27, 2018

Background Information

China Railway Customer Service Center (www.12306.cn) is an important platform for railway to serve customers. It will integrate the entire passengers as well as cargo transportation information and freight transport business and public information inquiry service for social and railway customers.

By logging in this website (www.12306.cn), customers can check the train schedules, fares, whether the train is late, remaining tickets, ticketing outlets, cargo tariffs, vehicle technical parameters and relevant passenger and freight regulations. Railway freight customers can handle business through this site as well.

Customers and Needs

Our project focuses on providing a new web page in which we will scrape data from travel and transportation websites, process and visualize these data and enable them to offer valuable information to our customers. Our customers would be Chinese government, economic researchers, train and high-speed rail passengers and investors who are interested.

With the aim of improving citizens' life quality, Chinese government have done a lot to maintain and optimize Chinese transportation system. Specifically, high-speed rail plays an integral role in public transportation, no matter for people who are travelers or businessmen. These data, related to rail transportation, to some extent, can offer government, economic research and investors insights of passenger flow in different cities, population migration during some festival and even, the economy development.

As for passengers, what they need are comfortable and enjoyable tours. An informative travel data website would bring efficiency, comfort and convenience to their tours and hectic lives. Our project is going to process the data into some indexes and suggestions including crowd rate, simultaneous arrival number, adult to child ratio, refund rate and optimal trip suggestion. Also, some data will be visualized to density traffic map and real-time map of the remaining tickets.

Problems and Necessities

Problems:

1. What the distribution of population flow during holiday seasons within a year is.
2. What changes in urban population structure in China through population migration during the Spring Festival in last 10 years are getting through.
3. How to figure out the most optimal trip with lower price and shorter time for passengers, especially during holiday seasons.
4. What the picture of remaining ticket information in real-time looks like.
5. How to estimate the demand for taxi at specific moment in order to inform the passengers in advance that how long they will wait for a taxi after arrival.

Necessities :

For Problem #1:What the distribution of population flow during holiday seasons within a year is.

By comparing the data of holiday trips horizontally, we can figure out the popularity of different tourist cities in the specific year. Likewise, Through longitudinal comparison, we can also see the trend of popularity of the tourist cities among consumers . The analysis of these data can provide a useful reference to the Chinese Government in terms of tourism city planning, so that more national capital can be invested in the tourist cities with a rising trend, the efficiency of capital utilization can be improved, and the most promising tourist cities can gain better developing opportunities. In the end, all people will be benefited.

For Problem #2:What changes in urban population structure in China through population migration during the Spring Festival in last 10 years are getting through.

Through the analysis of population migration during the Spring Festival in the past decade, we can have of better general picture of how the distribution of urban population is changed in China in recent years. Based on this data, the Chinese government can adjust the process of urban planning appropriately, such as adding public facilities in urban areas with a larger population or diverting urban population, etc., so that the cities can achieve a healthier and steadier development.

For Problem #3:How to figure out the most optimal trip with lower price and shorter time for passengers, especially during holiday seasons.

When buying high-speed rail tickets, we often struggle to find the optimal trip with shorter time and lower price. Only by constantly changing and entering the place of departure and arrival to compare the different trips one by one can we find out ‘optimal’ trips, which results in a waste of time and often misses a truly optimal trip. With the data scraped, we can help ticket-bookers quickly fetch the best real-time itineraries, including transfer information, which will save them money and improve efficiency.

For Problem #4:What the picture of remaining ticket information in real-time looks like and how comfortable and suitable it is to take a train.

Using the data, we can make a real-time map of the remaining tickets, which gives the ticket-bookers the newest information in the fastest way, so that they can make the purchase decision as quickly as they can, which is of great significance during rush seasons.
-Our data also shows real-time population distribution information in different carriages since the ticket prices for children and adults are different. Therefore, we can provide ticket-bookers with information about the number of children in a particular carriage. Passengers who are noise-sensitive or want to work in the train can choose to avoid the carriage with many children. On the contrary, passengers with children can also choose the carriage with more kids so that their children will not be alone on the road. In this way, the needs of different customer segments can be met perfectly.

For Problem #5:How to estimate the demand for taxi at specific moment in order to inform the passengers in advance that how long they will wait for a taxi after arrival.

Long queues of waiting for taxis often appear at high-speed rail arrivals. Passengers usually wait for an hour to get in. We can scrape and use the real-time data to analyze the arrivals and departures within half an hour before and after the passengers’ arrival time, that is the congestion index. Passengers who do not want to wait for a long time can find out the train with lower congestion index to have a more enjoyable trip.

Data Sources

1. www.12306.com, the website of China Railway Customer Service Center
2. The official website of Ministry of Transport of People's Republic of China
3.Weibo, Wechat
4.Other websites selling Chinese mainland trains and high-speed rail tickets： Qunar, Xiecheng, Feizhu

*Weibo and Weichat are Chinese versions of Facebook and Twitter.
*Qunar is a Chinese-language online travel information provider and mainland search engine for web-based and mobile users. Xiecheng and Feizhu are similar to Qunar.