



Adventura Data Analytics Project

**Analytical Approach to Consumer Market Data and
Algorithmic Design**



Project Deliverables

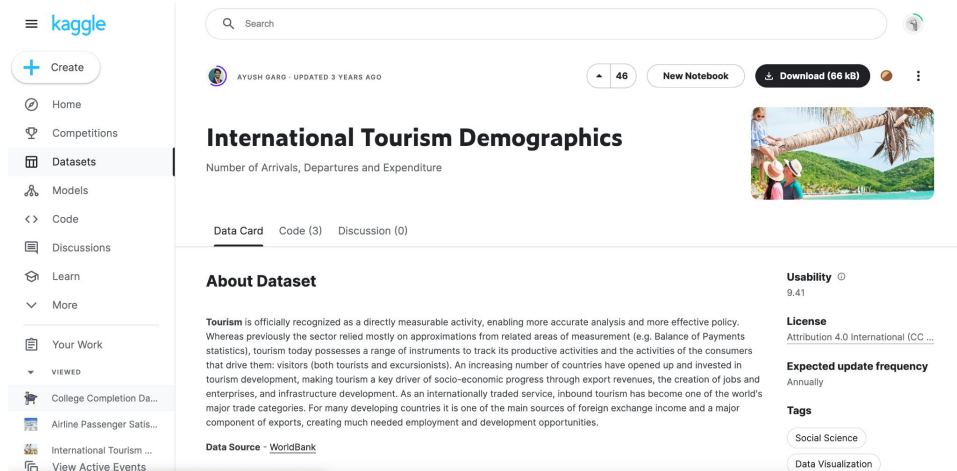
- I used the dataset to help inform my travel management platform about suggestions they could make to improve the recommendation algorithm and increase the conversion rate for users on the site.
- A higher conversion rate translates to greater revenue that our partners will provide us commission for. The datasets will allow us to conduct descriptive and diagnostic analysis on the data set, and then generate rules that can be applied to predict unknown outcomes.
- The final end product will be a report that outlines key metrics and relationships of preferences of travelers and tourists that correlate with higher levels of satisfaction.



Methodology

Exploratory Data Analysis

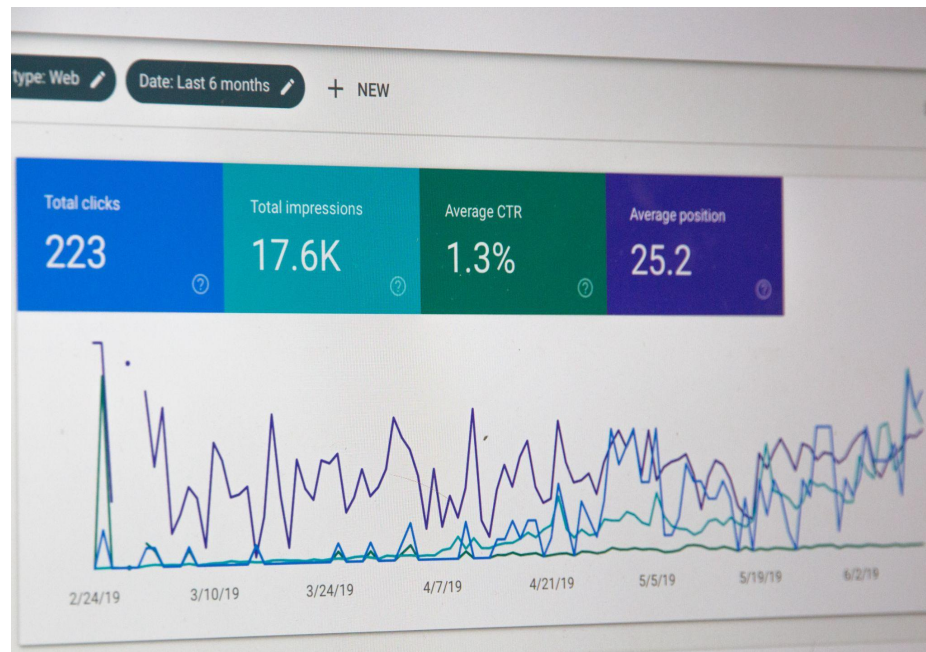
- Each dataset was collected from the given sources and then analyzed through data cleaning, and analysis methods.
- Using a combination of pivot tables and systematic categorization I created visualizations categorizing the various relationships and showing how certain attributes contribute more to users leaning towards a particular element.
- The specific features we will look for will be relationships that showcase trends relating to the higher ratings from tourists.



The screenshot shows the Kaggle interface for the dataset 'International Tourism Demographics' by Ayush Garg. The left sidebar contains navigation links: Home, Competitions, Datasets (selected), Models, Code, Discussions, Learn, and More. Below these are 'Your Work' and a 'VIEWED' list including 'College Completion Da...', 'Airline Passenger Satis...', and 'International Tourism ...'. The main content area displays the dataset title, a description of arrivals, departures, and expenditure, and tabs for 'Data Card', 'Code (3)', and 'Discussion (0)'. The 'About Dataset' section explains that tourism is a directly measurable activity and provides context on its economic importance. On the right, a 'Usability' score of 9.41 is shown, along with a 'License' of Attribution 4.0 International (CC BY) and an 'Expected update frequency' of Annually. At the bottom right, there are 'Tags' for 'Social Science' and 'Data Visualization'.

The Chosen KPI

- ***The chosen KPI:*** Reviews
- ***Indicators of value in KPI:*** 4-5 Star Reviews/Ratings and numerical scores of above average on a numerical scale will indicate a highly rated area and a total amount of 500+ Reviews from customers. Reviews such as *Satisfied* and *Superior* will be translated into numbers

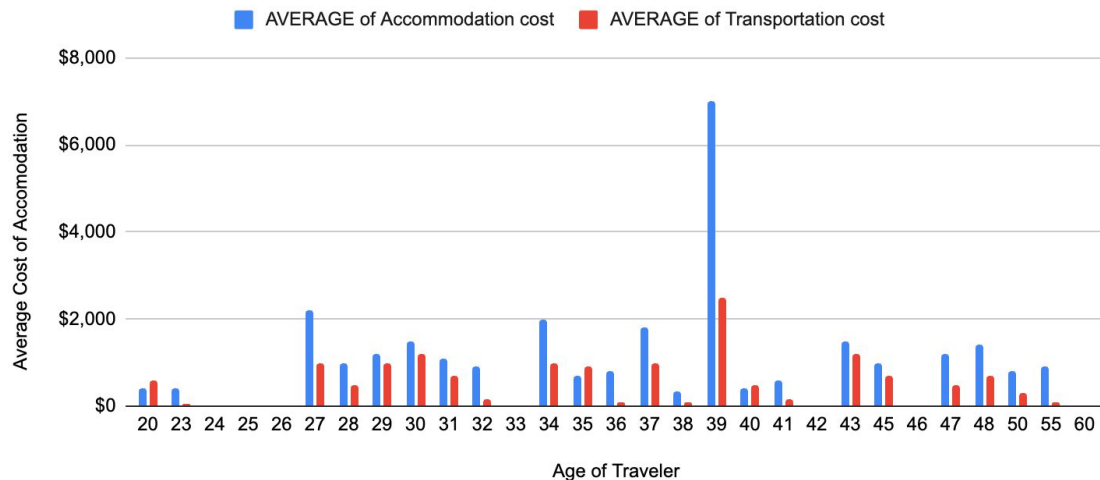


Average Costs for Accommodations and Transportation by Age of Traveler

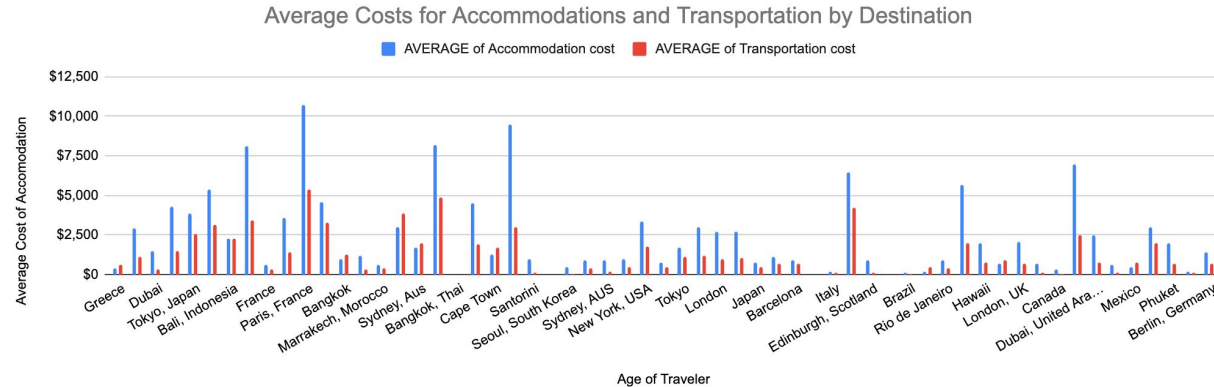


- Distribution
 - Shape - roughly even with a sharp spike around the 39 age mark
- **Insight:** Age 39 travelers appear to spend the most on average across all datasets. Targeting this age group with higher priced and more luxurious packages may be advantageous for both us and our partners.

Average Costs for Accommodations and Transportation by Age of Traveler



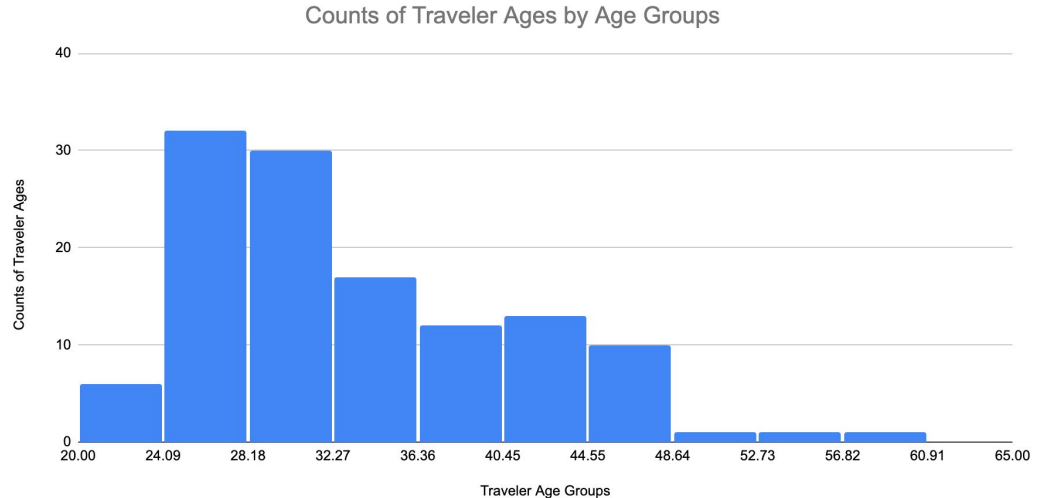
Specific Destinations can be Ranked Based on Cost Values



- The chart above shows the range of costs and spending associated with each travel destination. Aggregating the location from lowest to highest price may be desirable by users of the platform looking to budget. It may be useful to rank the destinations based on cost to help address the user needs.
- **A ranking system can be implemented for users on a budget that can filter destinations by price this ensures that**

Ages 24-32 are the Most Active Age Group of Travelers

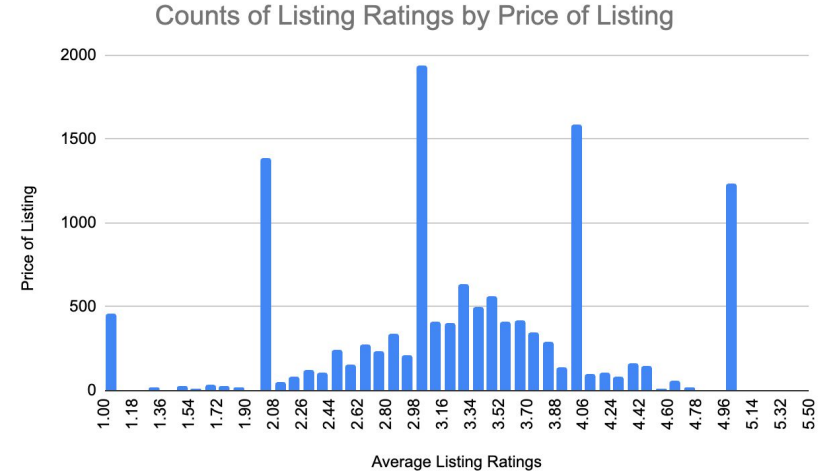
- The above graph showcases the counts of travelers by age group, showing trends in the number of travelers based on the age.
- While age may be a main factor, it should be noted that external factors such as income level and country of origin may play a role in these data trends as well.
- **Tailored targeting toward the age group of 24-32 may produce the greatest ROI.**



Higher and Lower Pricing results in Both Lower and Higher Average Listing Ratings



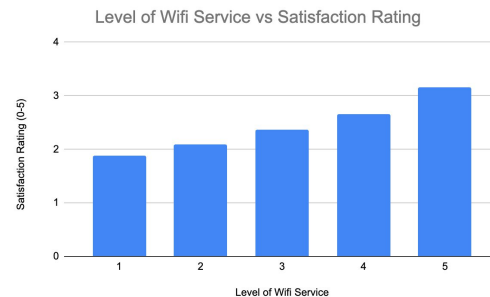
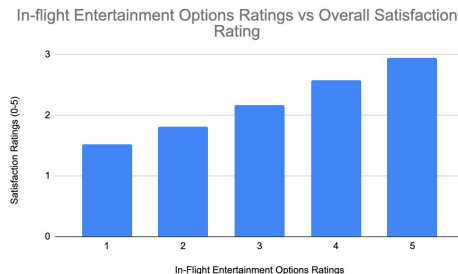
- The graph shows the count of listing ratings compared to the price of the listing.
- This relationship reveals clear trends around the price where higher prices actually resulted in mid-level ratings around 3 and lower prices correlated with both lower and higher ratings.
- It should be considered that there are probably other considerations such as location and the condition of the listing however this is an interesting datapoint to include in the data algorithms nonetheless.



Transportation Services that Prioritize These Key Factors Should be Ranked Higher in our Algorithm



- According to the listed relationships there appears to be a clear correlation between the level of satisfaction for the experience overall and the rated satisfaction rate associated with the chosen categories and variables.
- **This shows clear factors that we should consider when choosing airline transportation services to recommend to users on the platform.**

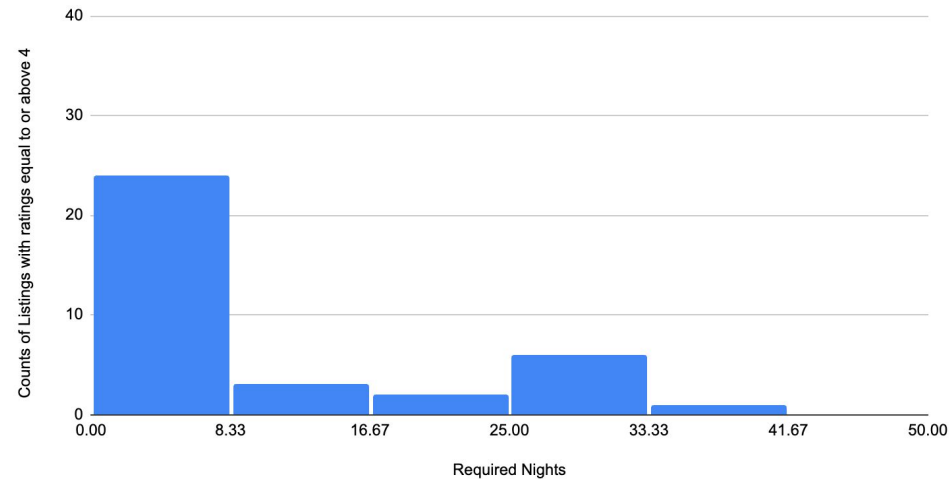


Listings with a More Accommodating Scheduling are Ranked Higher by Travelers



- The chart shows the relationship between high level ratings and the minimum number of required nights at the listing.
- As revealed by the chart, the higher level listings are correlated with lower minimum required nights (less than 8 nights).
- This may reveal a relationship between listings that are more accommodating of the schedule flexibility and higher user satisfaction which allows us to aggregate those listings higher up in our own rankings and algorithms.

Counts of Listing Ratings Equal to or Greater than 4 by Minimum Required Nights



Next Steps

- Reviewing the recommendations
 - Key Metrics related to higher user satisfaction rates have been identified
 - Key factors have been identified that draw connections in relationships to users preferences

We can determine that a targeted campaign to the users, partners and research of additional datasets will allow us to build out a comprehensive algorithm for the travel management service based on the revealed insights of relationships. These insights will inform the AI data model of how to rank and sort listings and categories for users that have specialized preferences, improving the overall user experience.



Additional Resources Needed

A full comprehensive data report that is able to identify ample evidence connecting the variables to the the KPIs will allow us to finally create a fully fleshed out algorithm that is able to improve over time as users begin utilizing the platform.

Additionally, some in-depth market research into the specific listing platforms will inform on the preferences of the target market.

