

Predicting Skytrax's Official World Airline Star ratings from customer reviews

Jose D. PÉREZGONZÁLEZ* and Andrew GILBEY
School of Aviation, Massey University, New Zealand

Abstract. *Skytrax audits and ranks airlines internationally in its yearly 'Official World Airline Star Ranking'. Unfortunately, its activity is severely restricted, at most covering just 30% of those airlines listed as open for review by passengers in its website in 2010. This research article explored the possibility of using a readily available variable as predictor, as an alternative way of ranking the remaining 70% of airlines in a simpler and more straightforward manner. The regression formula retained correlated highly with the criterion variable, accounting for 47% of its variance, thus supporting the viability of using customer reviews as a possible way of predicting ranking scores for airlines not officially audited by Skytrax.*

Introduction

Skytrax is a consultancy firm based in London (UK), which does research and advisory consultancy mostly with the aviation sector (Skytrax, 2011a). It is probably better known for its yearly airline and airport ratings, and customer-choice-based 'World Airline Awards' and 'World Airport Awards' (Wikipedia, 2011a). The former rating ranks airlines and airports according to quality, after auditing done by Skytrax itself (2010). The second recognises the best airlines and airports as chosen by passengers by means of an international survey (Skytrax, 2011b).

Skytrax (2010) claims to be the world's largest airline review site, with over 670 airlines [reviewed or open for review] and "millions of airline and airport reviews online". Yet, only 641 airlines were listed as opened for customer review in 2010 (or about 16% of passenger airlines in the world, as per Wikipedia, 2011b), only 244 of these airlines had been in the "star ranking programme" (or 30% of the airlines opened for review), and only 192 of these airlines had been audited and obtained a ranking in 2010 (or 79% of the 244 airlines in the star ranking programme) (Pérezgonzález & Gilbey, 2011a).

In previous research we have reported regression models for predicting Skytrax ratings from customers' reviews, both for airports (Pérezgonzález & Gilbey, 2011b) and for airlines (Pérezgonzález & Gilbey, 2011a). However, those models tend to include several variables, require the compilation of a database of individual reviews, and require the calculation of subsequent averages for further analysis. For example, the best prediction model for Skytrax airlines ratings included three variables: 'Customer rating', 'Recommended [as valuable airline]', and 'Value for money [evaluation]', resulting in a relatively high multiple correlation index between predictors and criterion ($R = .689$).

Still so, the compilation of those variables is cumbersome and time consuming. In a recent article, however (see Pérezgonzález & Gilbey, 2011c), we reported a regression model for predicting Skytrax ratings for airports from a single, readily available, averaged variable provided by Skytrax in its website. The resulting model was slightly

* Correspondence author: Dr. Jose D. PÉREZGONZÁLEZ, School of Aviation, Massey University, Turitea Campus, Private Bag 11222, Palmerston North 4442, New Zealand. <http://aviation.massey.ac.nz/>

less powerful than the one reported in Pérezgonzález and Gilbey (2011b), yet simpler and more straightforward to use.

The aim of this research now is to ascertain how well Skytrax airline rankings can be predicted from a similar single variable: the average customers' reviews offered by Skytrax as 'Customer review scoring'.

Methods

For this research, we investigated the population of 115 airlines which obtained a Skytrax rating for 2010 and had a minimum of 10 customer reviews during the year. From the data available at Skytrax's website, we selected two variables, readily available in the website, per airline.

The first variable acted as our criterion (or dependent) variable. This was the 'Skytrax official ranking', a score on a five-anchor ordinal scale running from "1, Very poor", to "5, Excellent").

The second variable acted as our predictor (or independent) variable. This was the 'Customer review scoring', an average score on an ordinal scale running from 0 to 10 – although there is no information of how this average was obtained, it plausibly represents the average 'Customer review scoring' for all reviews, including those of previous years).

SPSS-v16 was used for all analyses. A pre-screening data analysis showed that both variables were normally distributed and suitable for analysis using a linear model. Therefore, we used parametric tests for all subsequent analysis.

Results

Results show that it is possible to retain a model for predicting Skytrax airline rankings from the 'Customer review scoring' supplied by Skytrax. This model is statistically significant ($F = 87.302$, $p < 0.01$), and its correlation with the criterion is relatively high ($R = .688$).

The regression model was the following:

$\text{Predicted Skytrax Ranking} = 1.675 + (.291 * \text{Customer review scoring})$
--

Conclusion

Skytrax rates airlines according to quality. It also aims to be the leading institution in doing so thanks to its "Official World Airline Star Ranking". Unfortunately, its activity is severely restricted, at most covering just 30% of those airlines which customers have reviewed in its website. The regression model obtained in this research may help in covering a portion of the remaining 70% of airlines which are not audited by Skytrax. The model presented here is potentially less useful than another predictive model presented elsewhere (see Pérezgonzález & Gilbey, 2011a), as the predictor variable includes all customer reviews, not just reviews for 2010, and, therefore, it reflects less contemporary opinions. However, it is much simpler (one predictor variable instead of three) and straightforward (the score is readily available in Skytrax's website, instead of needing the compilation of a database of individual scores and the computation of averages), and what it loses in representativeness is gained in convenience without necessarily affecting the underlying prediction.

References

PÉREZGONZÁLEZ Jose D & Andrew GILBEY (2011a). *Predicting Skytrax airline ratings from customer reviews*. Article sent for peer-review to the Journal of Air Transport Management (pending)

PÉREZGONZÁLEZ Jose D & Andrew GILBEY (2011b). *Predicting Skytrax Airport Awards from customer reviews*. Journal of Airport Management, volume 5, number 4, (in press).

PÉREZGONZÁLEZ Jose D & Andrew GILBEY (2011c). *A convenient regression formula for predicting Skytrax's Official World Airport Star ratings*. Aviation Education and Research Proceedings (ISSN 1176-0729), vol2011, pp 45-47.

SKYTRAX (2010). *The A-Z index to official world airline star ranking*. Retrieved from "<http://www.airlinequality.com/StarRanking/summary.htm>" on 7 December 2010.

SKYTRAX (2011a). Retrieved from "<http://www.airlinequality.com/>" on 23 February 2011.

SKYTRAX (2011b). *World airline awards*. Retrieved from "<http://www.worldairlineawards.com/>" on 13 May 2011.

WIKIPEDIA (2011a). *Skytrax*. Retrieved from "<http://en.wikipedia.org/wiki/Skytrax>" on 23 February 2011.

WIKIPEDIA (2011b). *List of passenger airlines*. Retrieved from "http://en.wikipedia.org/wiki/List_of_passenger_airlines" on 18 April 2011.