OLLSCOIL NA hÉIREANN
NATIONAL UNIVERSITY OF IRELAND, GALWAY

MS5107 BUSINESS MODELLING & ANALYTICS
_____

ASSIGNMENT 2

Predicting Airfares on New Routes

Task Description:

Several new airports have opened in major cities, opening the market for new routes
(a route refers to a pair of airports), and Southwest Airlines has not announced
whether it will cover routes to/from these cities. In order to price flights on these
routes, a major airline collected information on 638 air routes in the United States.
Some factors are known about these new routes: the distance travelled, demographics
of the city where the new airport is located, and whether this city is a vacation
destination. Other factors are yet unknown (e.g., the number of passengers that will
travel this route). A major unknown factor is whether Southwest Airlines or another
discount airline will travel on these new routes. Southwest's strategy (point-to-point
routes covering only major cities, use of secondary airports, standardized fleet, low
fares) has been very different from the model followed by the older and bigger
airlines (hub-and-spoke model extending to even smaller cities, presence in primary
airports, variety in fleet, pursuit of high-end business travellers). The presence of
discount airlines is therefore believed to reduce the fares greatly.
The file Airfares.xlsx contains real data for existing routes that were collected
recently:

S_CODE	Starting airport's code
S_CITY	Starting city
E_CODE	Ending airport's code
E_CITY	Ending city
COUPON	Average number of coupons (a one-coupon flight is a non-stop flight, A two-coupon flight is a one stop flight, etc.) for that route
NEW	Number of new carriers entering that route
VACATION	Whether a vacation route (Yes) or not (No).
SW	Whether Southwest Airlines serves that route (Yes) or not (No)
HI	Herfindel Index - measure of market concentration
S_INCOME	Starting city's average personal income
E_INCOME	Ending city's average personal income
S_POP	Starting city's population
E_POP	Ending city's population
SLOT	Whether either endpoint airport is slot controlled or not; This is a

	measure of airport congestion
GATE	Whether either endpoint airport has gate constraints or not; This is another measure of airport congestion
DISTANCE	Distance between two endpoint airports in miles
PAX	Number of passengers on that route during period of data collection
FARE (the response)	Average fare on that route

Note that some cities are served by more than one airport, and in those cases the airports are distinguished by their 3-letter code.

Objective:

A. Using your knowledge in Business Modelling and Analytics, build a model that predicts average fare on a new route.
1. Prepare a summary report to describe the model building process and why you believe that your model is good.
2. Using the model, predict the average fare on a route with the following characteristics: COUPON=1.202, NEW=3, VACATION=No, SW=No, HI=4442.141, S_INCOME = $28760, E_INCOME=$27664, S_POP=4557004, E_POP=3195503, SLOT=Free, GATE=Free, PAX=12782, DISTANCE=1976 miles.
3. Predict the reduction in average fare on the above route if Southwest Airlines decides to cover this route.

B. In reality, which of the factors (predictor variables) will not be available for predicting the average fare from a **new airport**? (i.e. before flights start operating on those new routes)
1. Briefly comment on your assumptions.
2. Based on the settings and findings of the model from item A, build another model using the available (in your opinion) variables only. Comment on your solution.
3. Use this model to predict the average fare using only the available (in your opinion) data from the record in item A.2.
4. Compare performance of this model with performance of the model from item A. Is this model good enough, or is it worthwhile re-evaluating the model once flights commence on the new route?

Requirements:

- Work in **groups** of 4±1. Email to the lecturer the group names as soon as you can.
- **Submit** an Excel file with essential worksheets (only) of your solution to the Blackboard electronic drop box provided. Separately, submit a hard copy (printout) of your report to the MS5107 drop box in Cairnes building. The report is limited to about 4000 words, printed in Times New Roman or similar font, 12-point, single spaced (which is about 15 pages incl. figures).
- Submit a signed group assignment submission form stapled to the assignment. **Individually, confidentially, and separately** from the report, each student

has to submit both peer evaluation form and individual contribution form, all available on the Blackboard. You can drop the forms in the MS5107 box in sealed envelops or hand them to the lecturer.

- The assignment is due on or before **Fri, 29th Nov 2019,** 23:59 and carries 25% of the total marks.