

STAT3500/7500 Consulting Competition

Deliverables

2.5%	half-time check	due 6pm 1 Oct
7.5%	seminar	during a lecture in the week of 22 Oct
20%	report	due 6pm 26 Oct

Task. Your task is to develop model(s) using given Translink™ passenger flow data, so as to answer queries of the form

*Predict the traffic **volume** going from **Location A** to **Location B** between **Time 1** and **Time 2** on this **Date***

Data. The training data is for the period from 23 Feb to 24 Mar, 2016. It contains the following attributes

- **Location A** and **Location B** are two regions/suburbs.
- **Time 1** and **Time 2** are two time points.
- **Date** is a day.
- **volume** is the number of travel trajectories.

The test data is sampled for the period from 25 to 31 Mar, 2016, with all the above attributes given except that volume is to be predicted.

Additionally, a small set of validation data sampled from the same period as the test data is given.

The training, validation and test data can be downloaded from Blackboard > Assessments > Consulting Competition. Jiwon's guest lecture slides will be available from Blackboard.

Score. Your predictions will be scored using the relative total absolute error

$$\text{RTAE} = \frac{\sum_{i=1}^N |Y_i - \hat{Y}_i|}{\sum_{i=1}^N Y_i}$$

- Y_i 's are the true counts of trajectories, and
- \hat{Y}_i 's are the counts predicted by your model.

Half-time check (2.5%, due 6pm 1 Oct). Each team will email a file containing your predictions (one line, space separated) to nan.ye@uq.edu.au. The scores for the teams will be announced in the next week.

Seminar (7.5%, week of 22 Oct). Each team will make a presentation describing the problem, modelling methods tried, results obtained, and conclusions and limitations.

Report (20%, due 6pm 26 Oct). Each team will submit a written Consultation Report containing the following sections:

- Introduction and background to the consulting project, including purpose of the project.
- Description of the data, including where the data come from, how the data are obtained, and present relevant summary statistic(s) and/or plot(s) of the data.
- Outline the modelling procedure undertaken by the team, including why and how the model(s) were chosen and/or constructed.
- Report final fitted model(s), and use your model(s) make predictions on the test set.
- Discuss any limitation(s) of your model(s) and identify ways in which the consulting task can be improved.

Although there is no page limit on the written reports, each will be assessed on presentation (be clear and concise, and include relevant diagrams/charts), correctness, and model quality.

Group members also need to rate each other's contributions to the project. Marks awarded to each student may differ according to their peer-rating scores.

Email your report and peer-rating scores, together with your *final predictions* and your *code*, to nan.ye@uq.edu.au.

Prize. The team with the smallest final RTAE will win a prize (TBD).