AAVAIL

REVENUE PROJECTION PROJECT

MODELING

INITIAL ATTEMPTS

In order to accelerate the building of models, I reused the solution guidance material

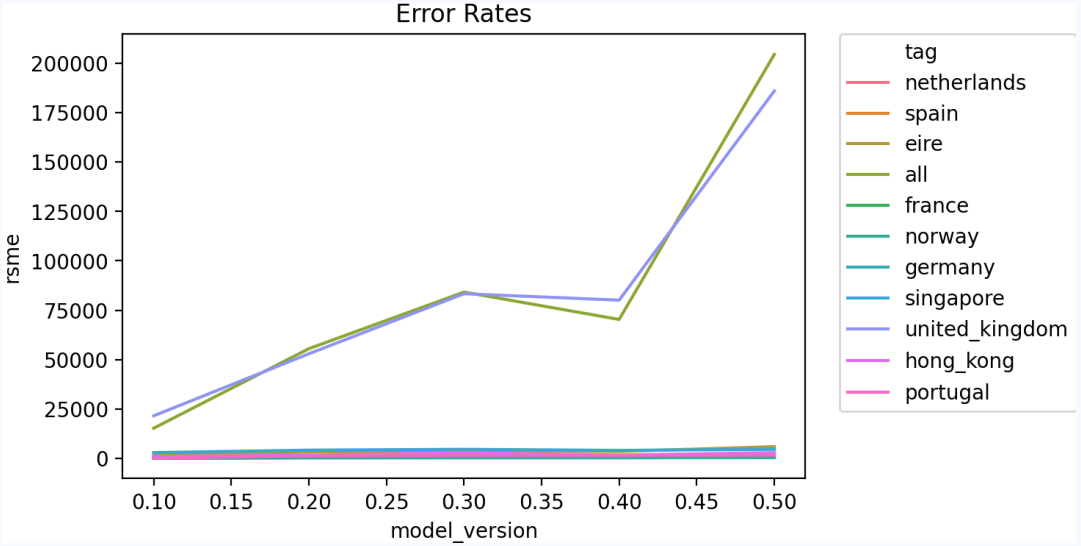
This used an initial RandomForest model. In order to compare this against other models, I selected both simpler and more complex model (each versioned)
V0.1 Random Forest
V0.2 Linear Regression
V0.3 Suppport Vector
V0.4 Bagging Regression
V0.5 Switched to Tensor Flow Neural Network (and altered code to handle loading either)

Code available in Github to look at configuration of each model

TRAINING ANALYSIS (PT. 1)

Initial analysis of the RSME indicates that the model version in the solution guidance (Random Forest) should be selected as it has the lowest error rate.

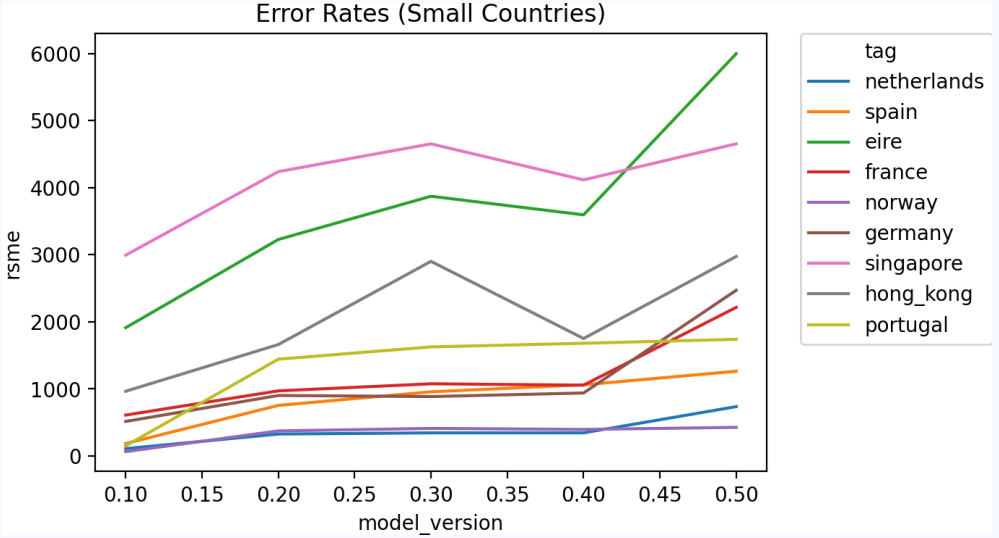
This means it most accurately predicts revenue.

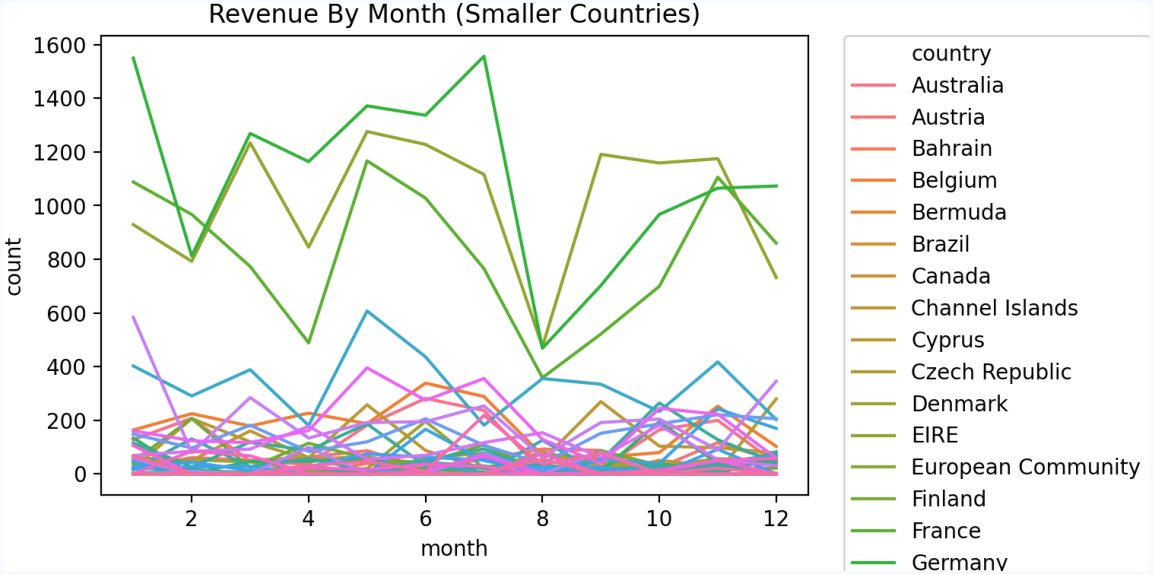


TRAINING ANALYSIS (PT. 2)

Excluding the larger datasets of "All" and "United Kingdom" we still see better error results on the Random Forest, though the error is less pronounced.

The performance of Radom Forest appears acceptable against revenue figures.





CONCLUSION

The Random Forest gives the best error rate and should be selected as the model of choice.