Investigating the Education Consultancy Business

Shaun Chua

DSI 13

Agenda

Agenda

- **Problem Statement**

- **Data**

- **Modelling**

- **Discussion**

- **Limitations**

- **Questions**

Problem Statement

"How much do I sell for?"

Every business owner ever

Problem Statement

The real one.

GeBiz

- **One-stop G2B Procurement**

- **Transparency**

- **Fair and Open Competition**

- **Value for Money**

Education Consultancies
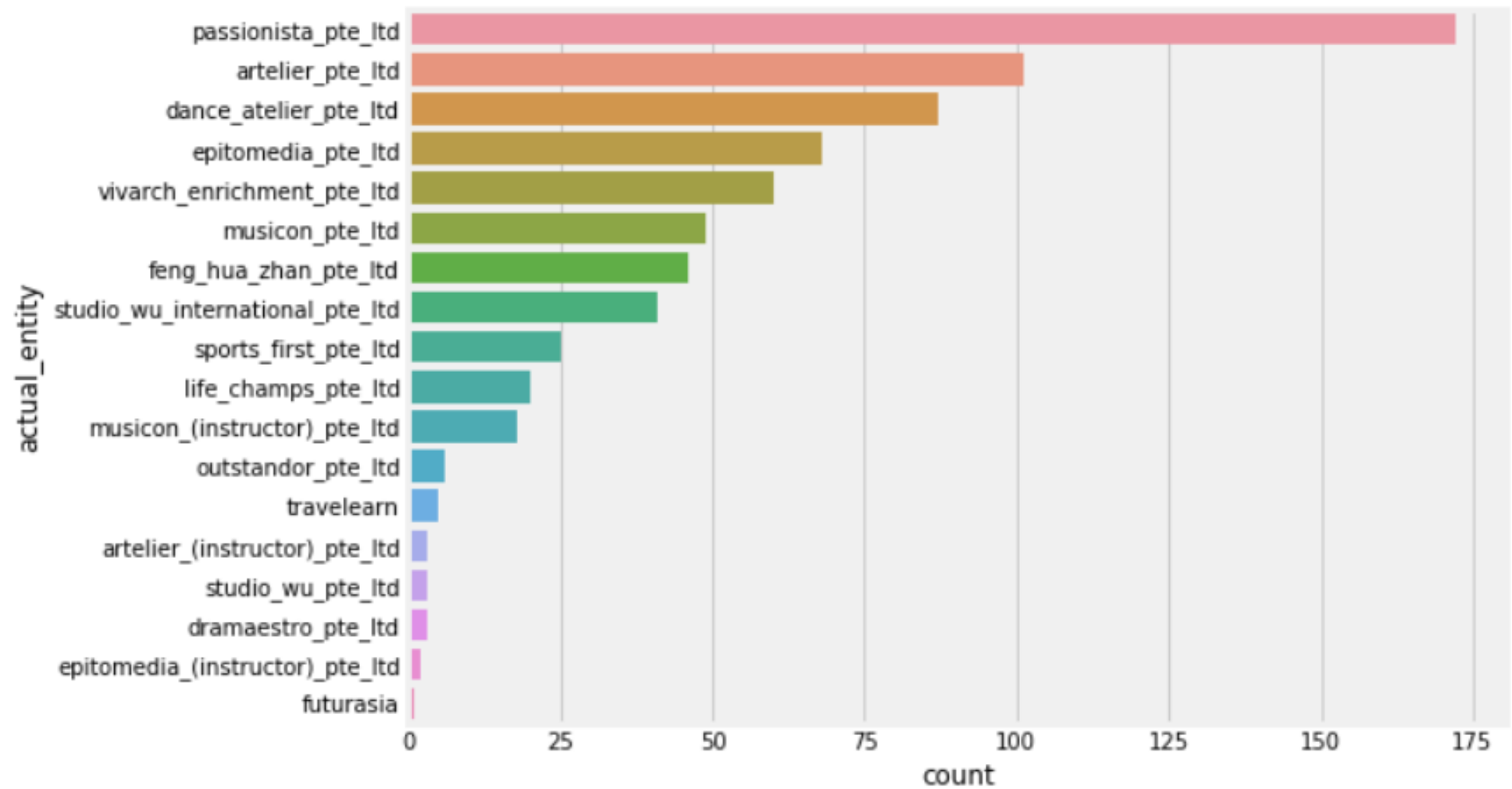
- **Pricing**

- **Business Development**

Data

Data

2018 Sales Data

- **Business Entities**

- **Consultants**

- **Zones**

- **UOM**

2019 Sales Data

- **Business Entities**

- **Consultants**

- **Zones**

- **UOM**

EDA
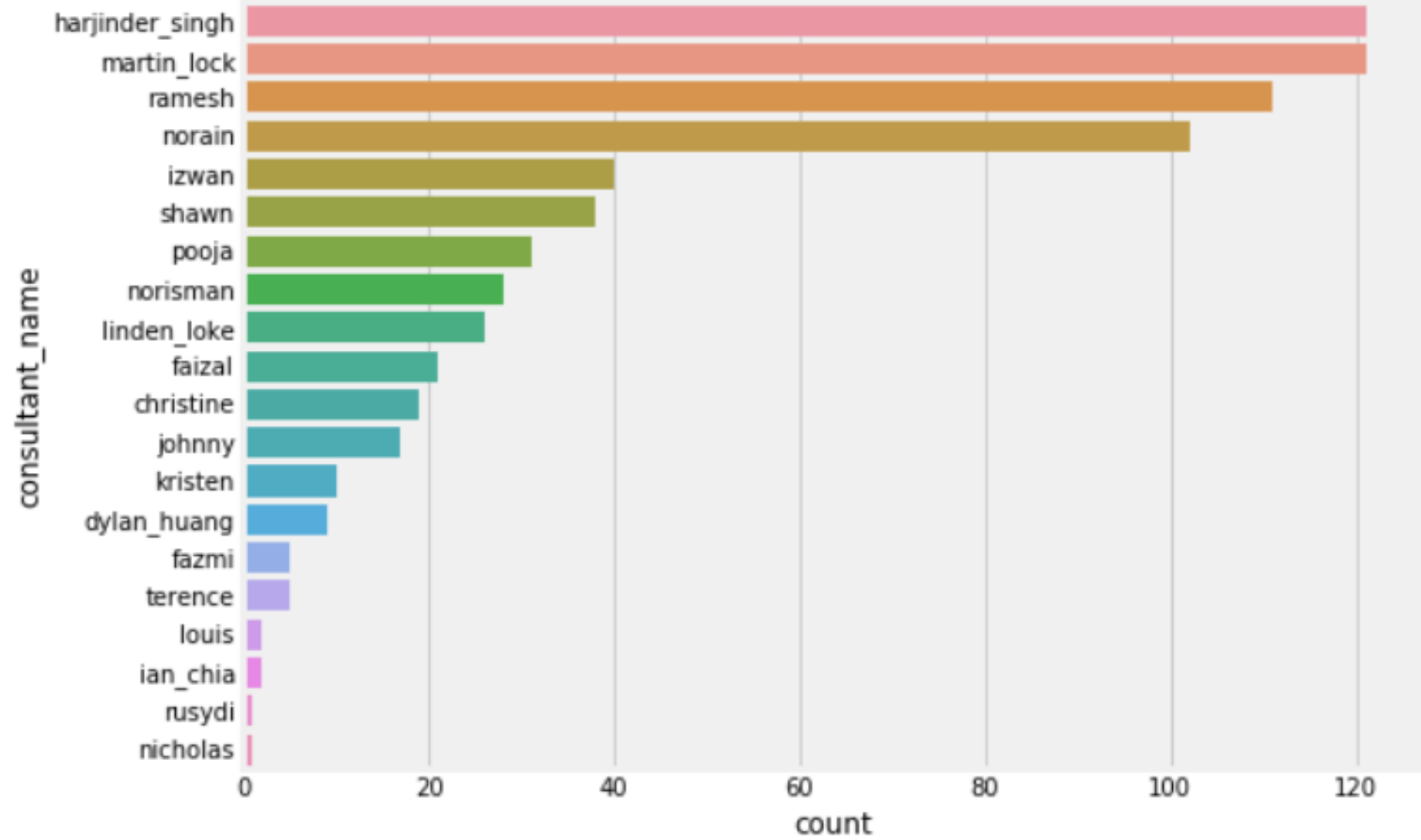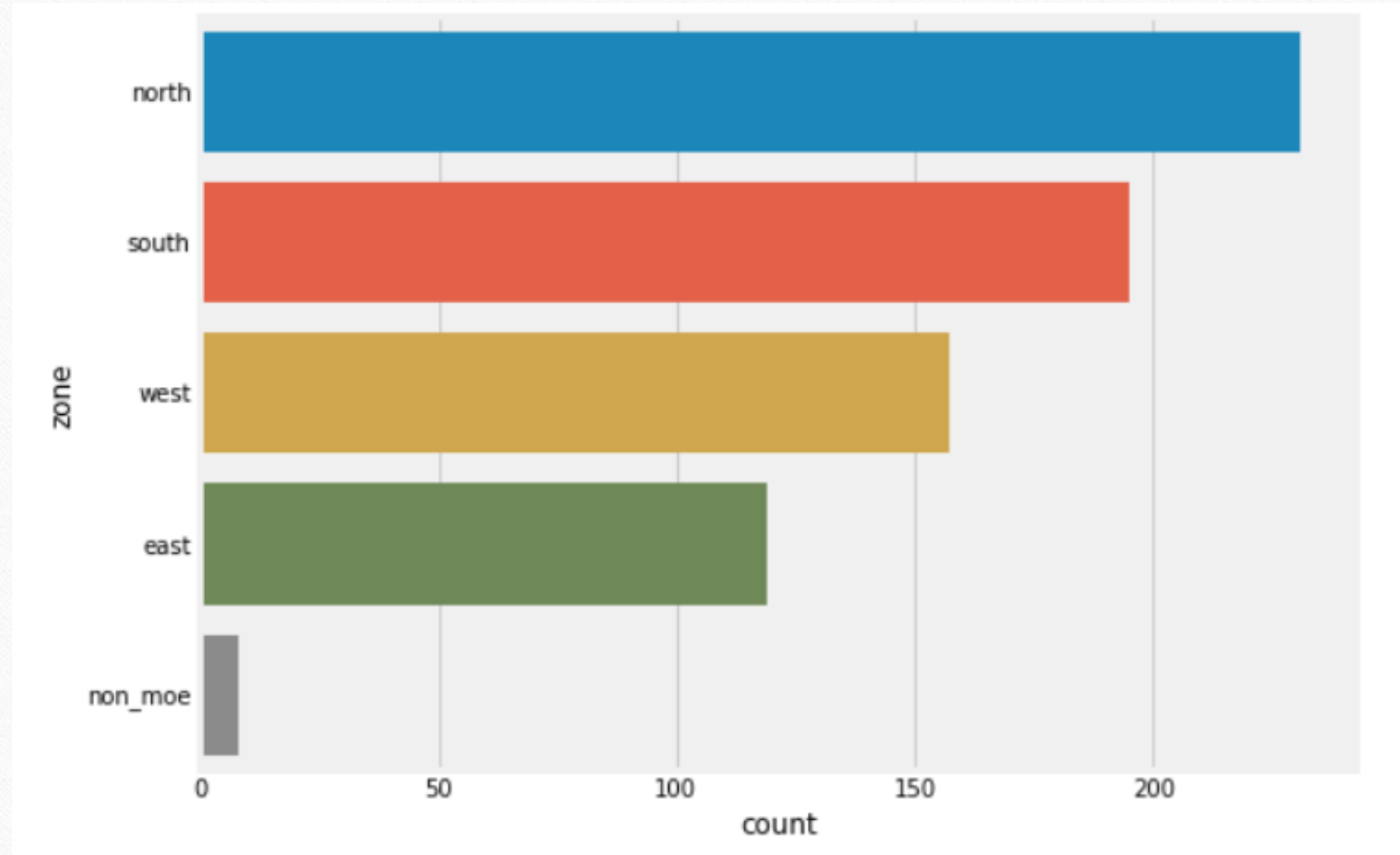
Modelling

Objective 1

Models

Linear Regression

- **Ridge**
- **Lasso**
- **Elastic Net**

Split Into:

- **All Features**
- **Business Units**
- **Consultants**
- **Zones**

	R² Score (Train)	R² Score (Validation)	RMSE (Train)	RMSE (Validation)
Baseline (All Features)	0.489088	0.562626	3388.069290	3357.637086
Lasso (All Features)	0.452751	0.552987	3506.484458	3394.430364
Ridge (All Features)	0.480780	0.563607	3415.504590	3353.866434
Elastic Net (All Features)	0.255492	0.280693	4089.912316	4305.903271
Baseline (Entity)	0.119470	0.051100	4447.863007	4945.580090
Lasso (Entity)	0.112078	0.056427	4466.492811	4931.679422
Ridge (Entity)	0.119282	0.050539	4448.337114	4947.042968
Elastic Net (Entity)	0.090943	0.051438	4519.339756	4944.699843
Baseline (Consultant)	0.132197	0.135346	4415.601806	4720.936564
Lasso (Consultant)	0.132130	0.135687	4415.773527	4720.006020
Ridge (Consultant)	0.132105	0.134826	4415.837144	4722.356910
Elastic Net (Consultant)	0.086963	0.060551	4529.220591	4920.889870
Baseline (Zone)	0.050958	0.002728	4617.660689	5070.068002
Lasso (Zone)	0.050123	0.010510	4619.692256	5050.248789
Ridge (Zone)	0.050905	0.005380	4617.790087	5063.323554
Elastic Net (Zone)	0.037834	0.021018	4649.478797	5023.360621
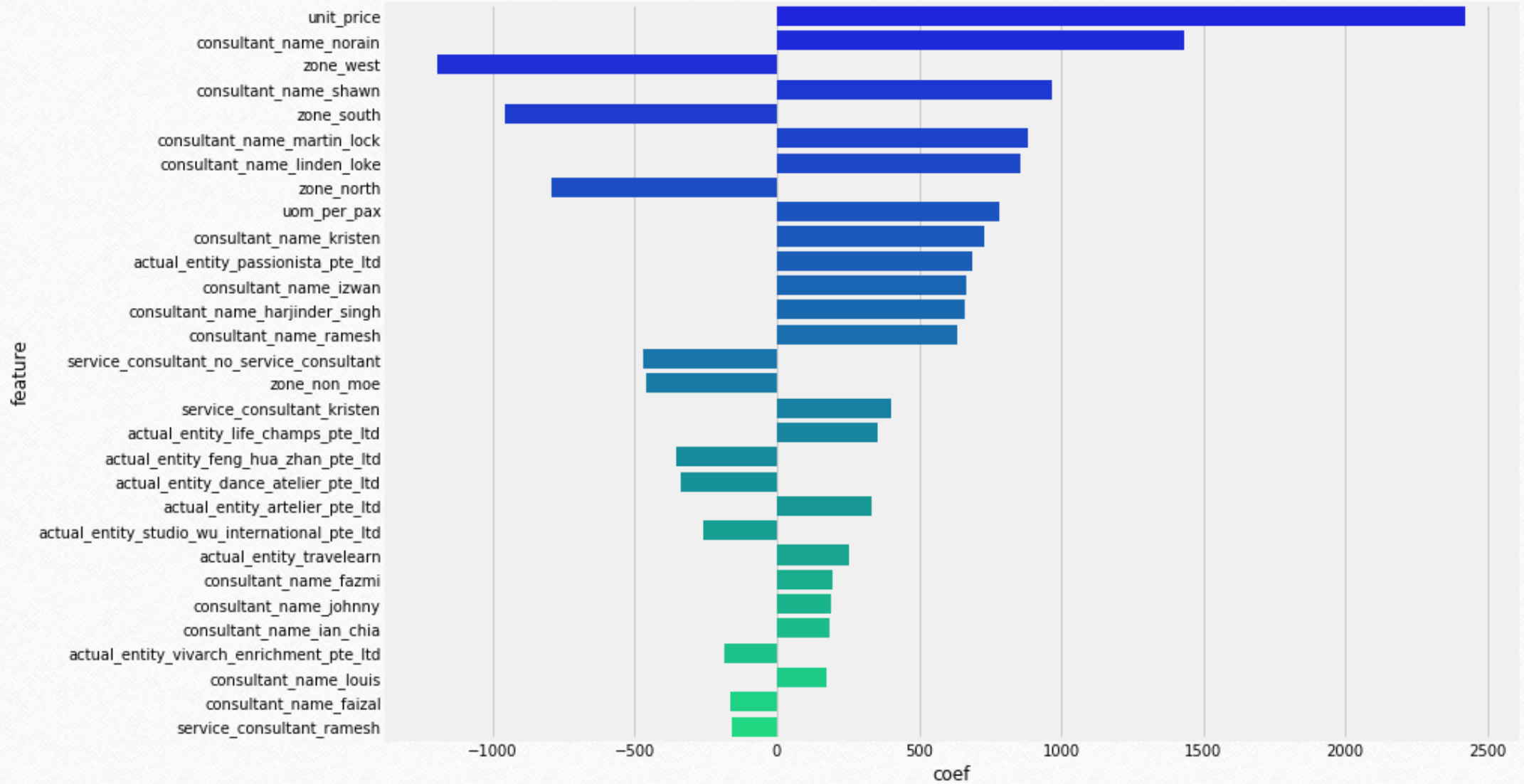
Modelling

Objective 2

	R² Score (Train)	R² Score (Validation)	RMSE (Train)	RMSE (Validation)
Optimised Ridge	0.480780	0.563607	3415.504590	3353.866434
Decision Tree Regressor	0.701613	0.579496	2589.221795	3292.243260
Random Forest Regressor	0.798746	0.666979	2126.431973	2929.834953
Support Vector Regressor	0.411421	0.486549	3636.485453	3637.947090
AdaBoost (Decision Tree Regressors)	0.925296	0.647073	1295.537461	3016.125130
AdaBoost (Random Forest Regressors)	0.920516	0.654208	1336.347576	2985.482482
Gradient Boosting Regressor	0.857589	0.666774	1788.755612	2930.733040
Extreme Gradient Boosting Regressor	0.840017	0.670740	1895.904214	2913.242730

Discussion

Business Implications

Objective 1

- **Per Pax**

- **Lagging Entities**

- **Align Pricing Strategies**
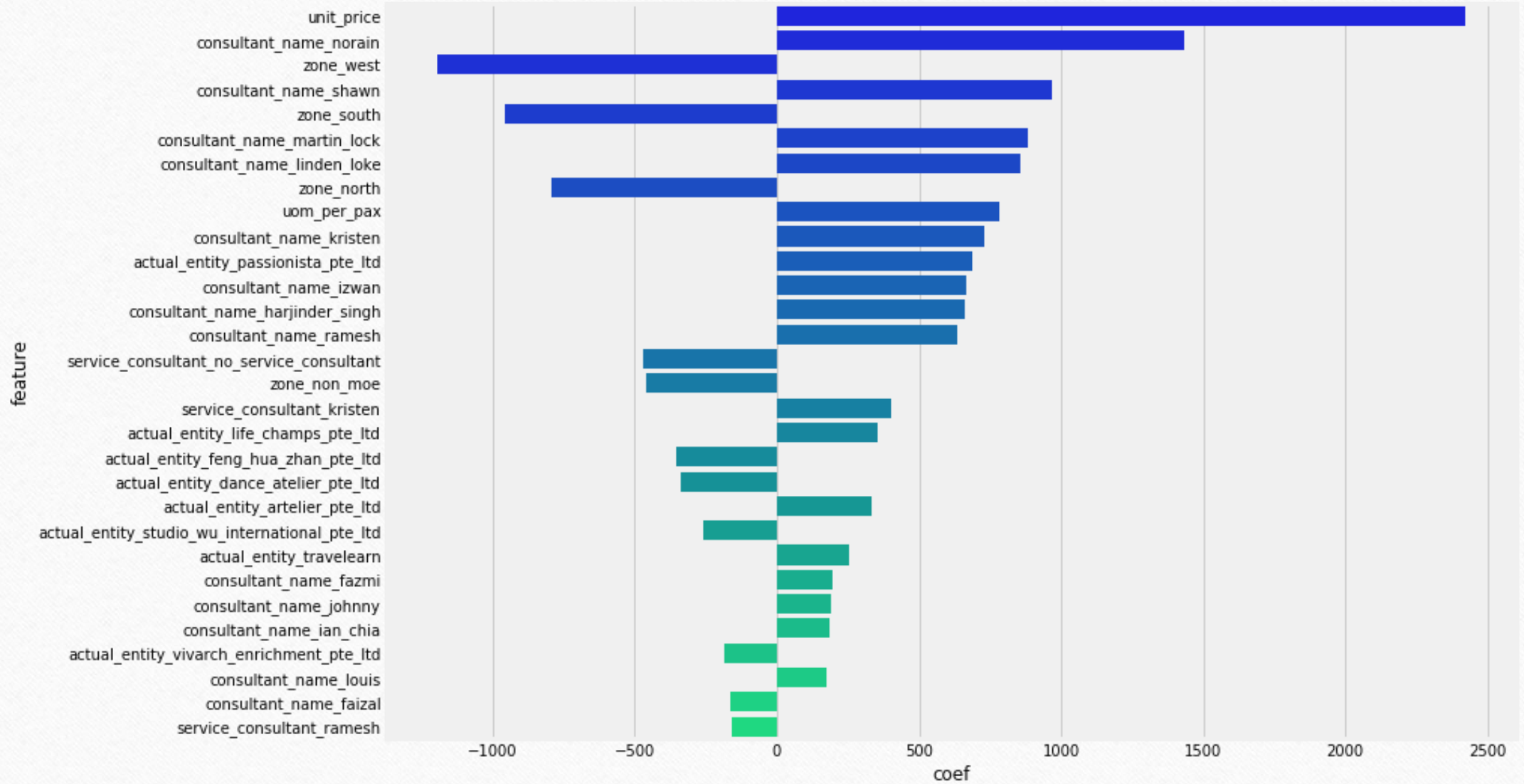
Business Implications

Objective 1

- **Per Pax**
- **Lagging Entities**
- **Align Pricing Strategies**

Objective 2

- **Modest guide for <u>high-value</u> programmes only**

Future Directions

Yup, limitations here too.

Future Directions

- **Small Dataset: Get more data**

- **Limited Tuning: AWS**

- **Differentiating UOM**

Questions