



Location sales performance analysis

Restaurant chain

Built a comprehensive model to analyze demographic variables and site characteristics to identify key sales performance factors to lead the decisions for new location openings and modifications to existing locations

Location sales performance analysis for a restaurant chain

Situation

- Client wanted to understand the key factors that would impact the sales performance of the restaurants in order to scout for new locations and make changes to existing locations.
- Partnered with the client to build a comprehensive model to identify the key performance factors across different locations.

Accordion Value Add

- Analyzed the POS transactional data to identify the key performance metrics; Also analysed 21 demographic variables and 72 restaurant site characteristics.
- Identified the key demographic and site characteristics that would likely impact the restaurant performance using advanced algorithms such as Random Forest and CART (Classification and Regression Tree).
- Summarized the key actions that could be piloted by restaurant segment in order to improve the sales performance.

Impact

- Leveraged insights to guide new location openings and modifications to existing locations (e.g., add to-go parking etc.) for higher sales performance.

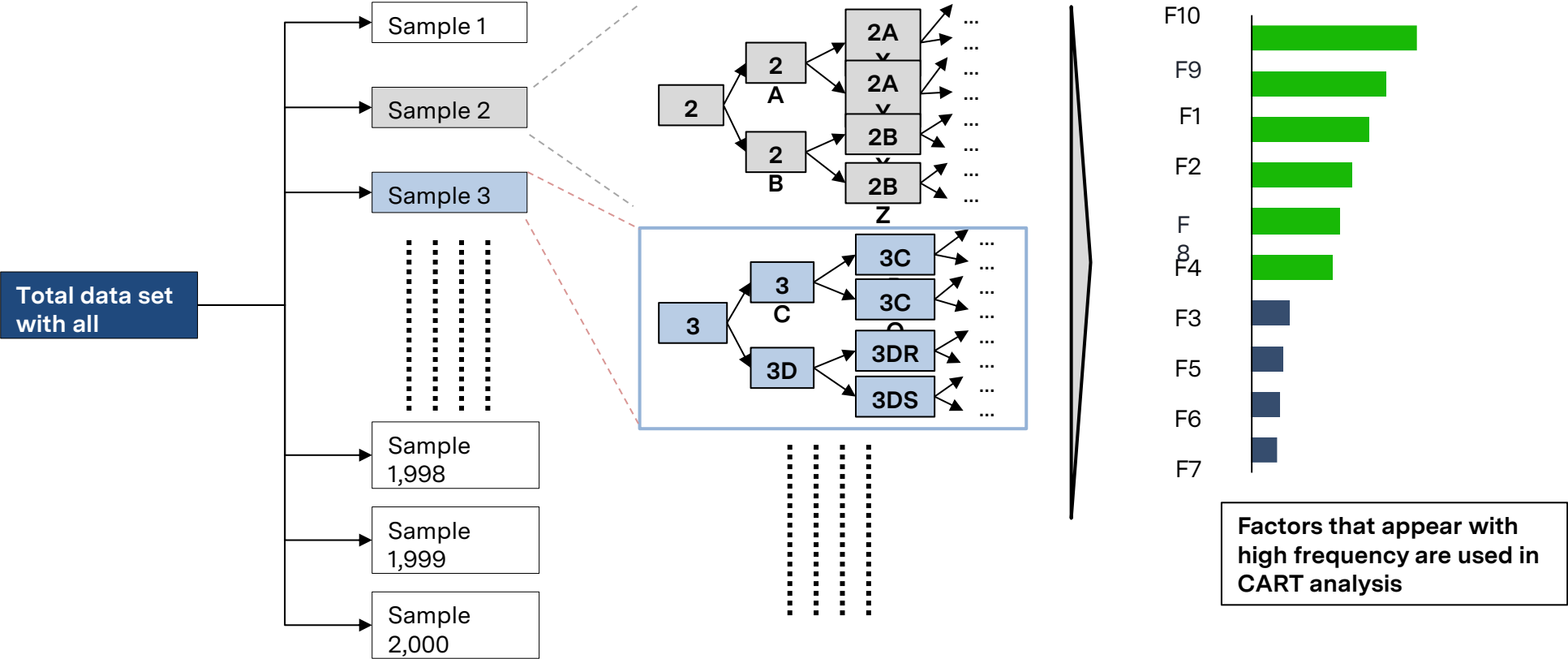
Exhibrandom forest algorithm - Methodology

Objective: To identify the key determining factors using Random forest algorithm (n= 2,000)

Randomly created 2,000 training samples using combinations of data points (restaurants) and variables (demographic and site characteristics)

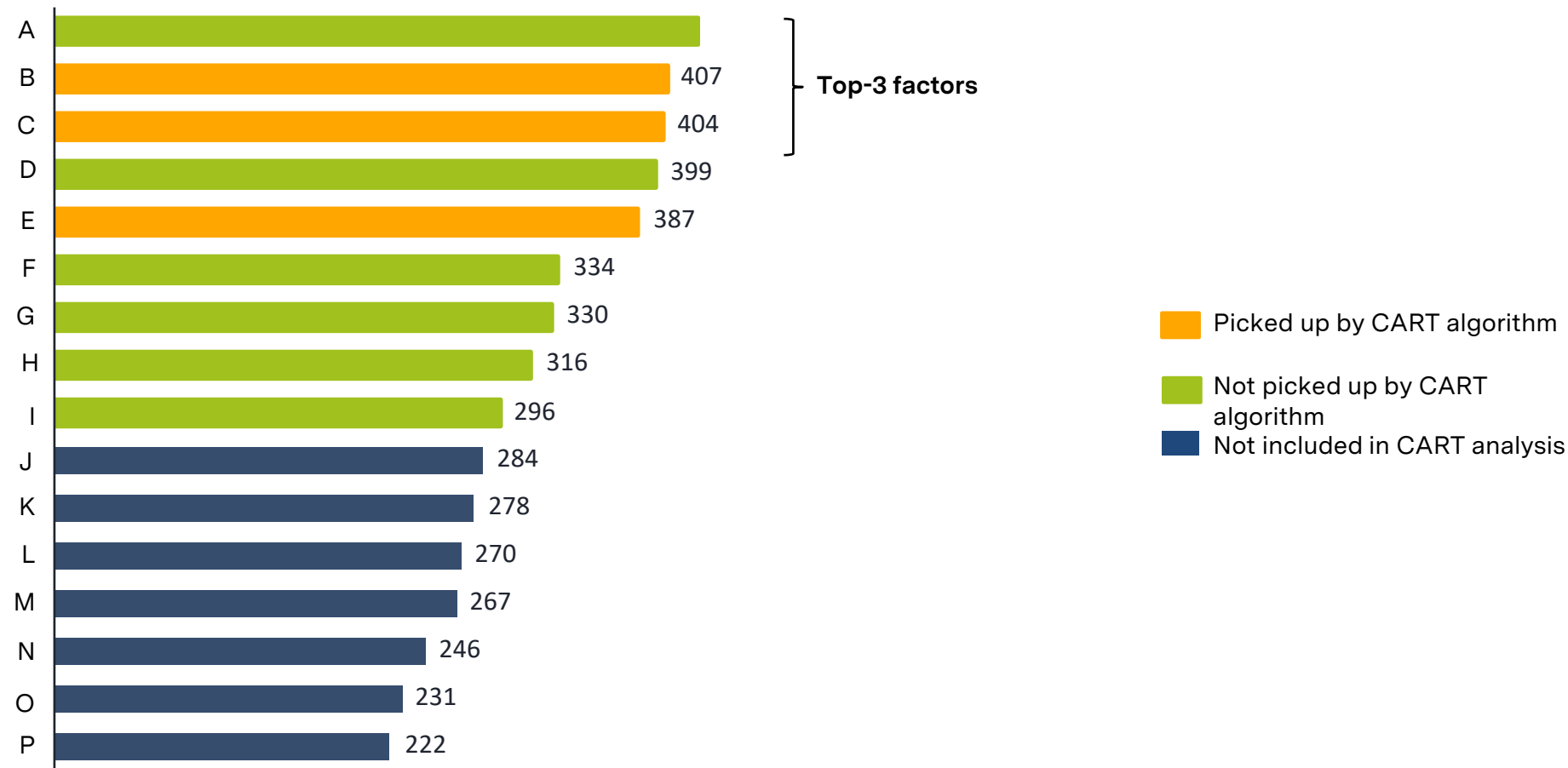
Generated a decision tree for each sample set (Sample 1, Sample 2, Sample 3 etc.) based on variables (A,B,C,P,Q, X, Y etc.) selected in a sample

Plotted the frequency at which each factor (F1, F2 etc.) appears on aggregate in the 2,000 decision trees that RF algorithm generates



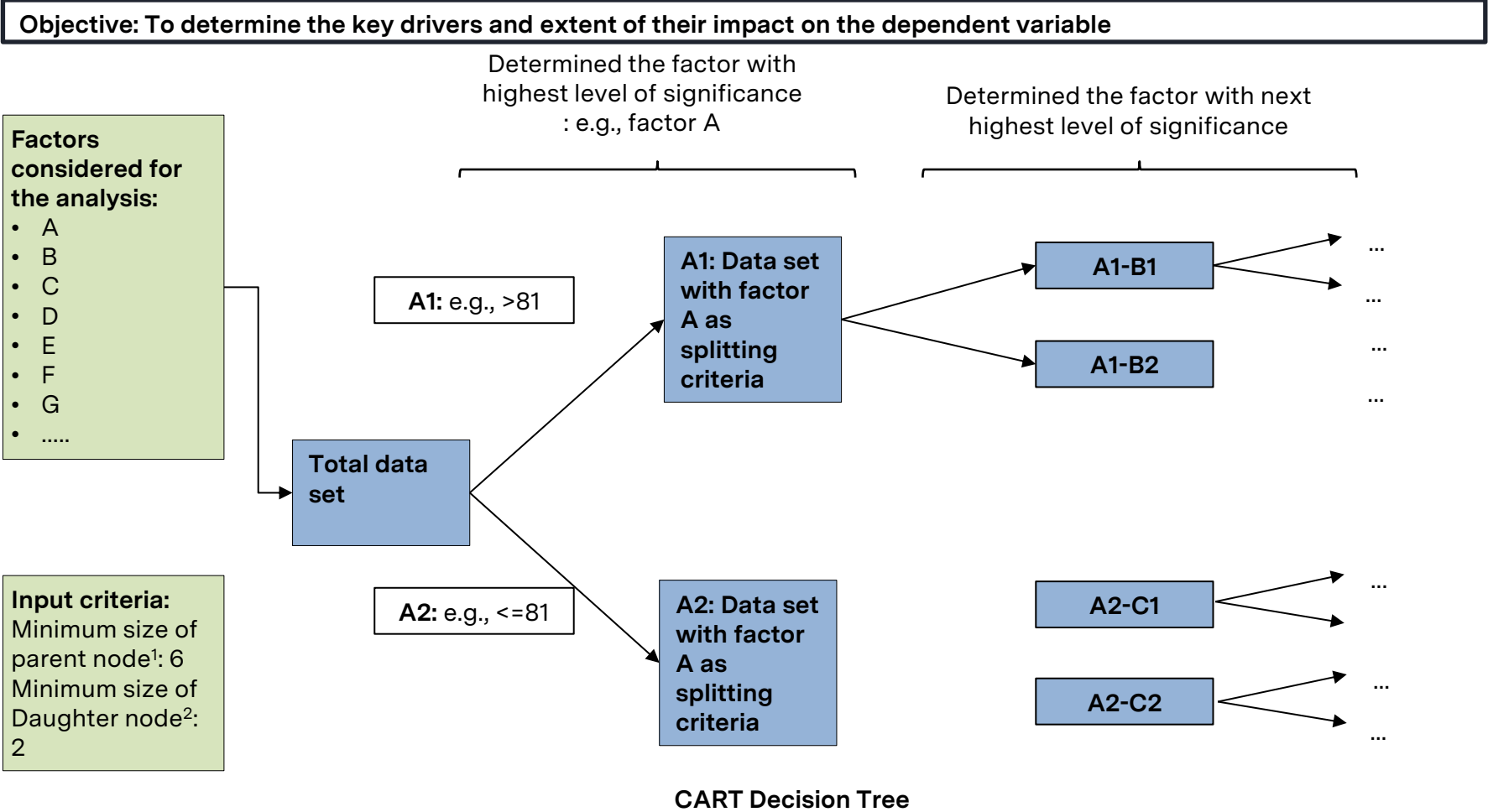
Random forest analysis suggest factors “a”, “b” and “c” impact revenue per 1,000 population

Frequency¹ of each demographic factor (Random forest algorithm), #



1 Frequency at which each demographic variable appeared as a determining factor when random forest algorithm plots 500 different decision trees

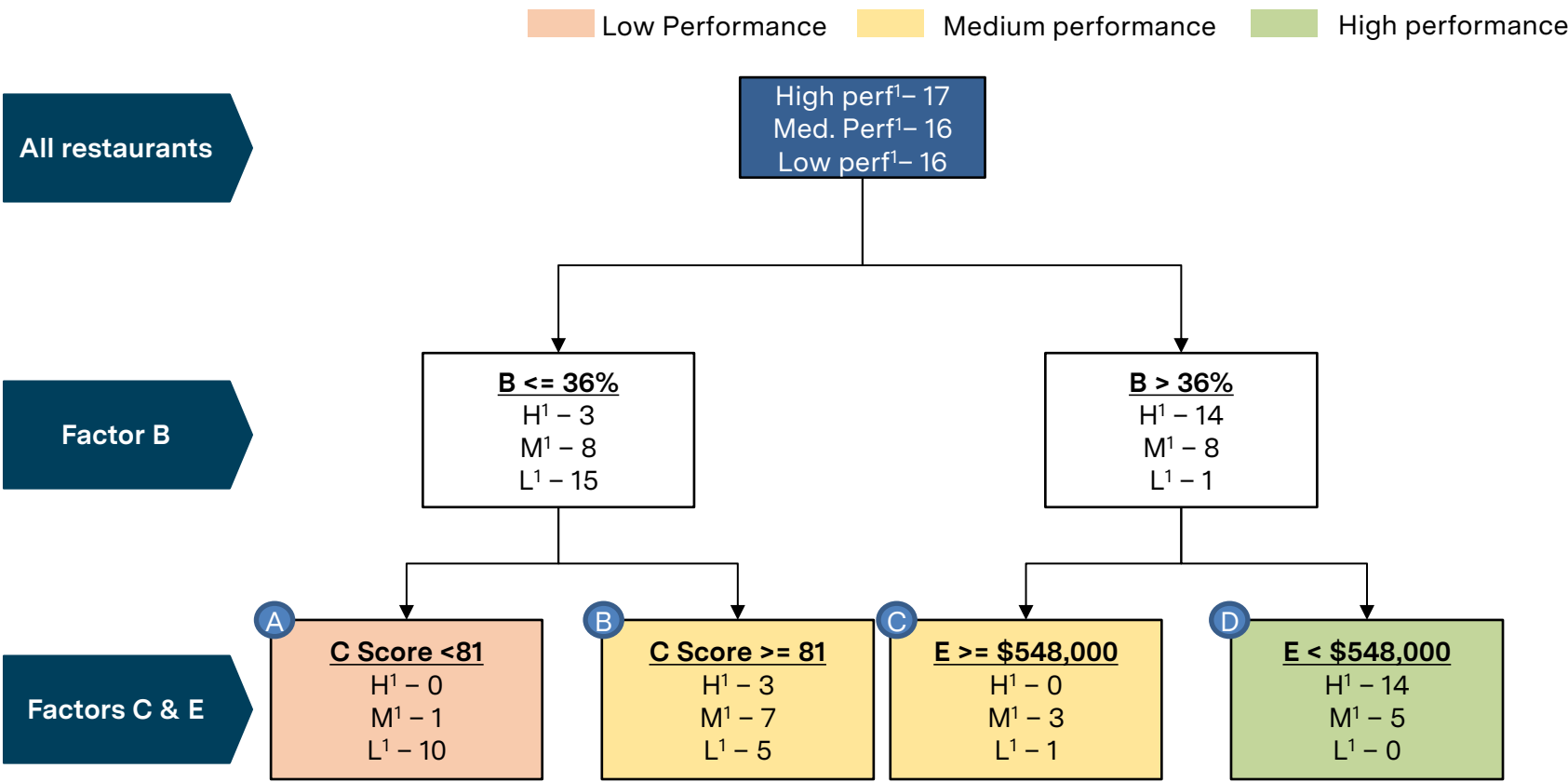
Classification and regression tree (cart) - Methodology



1 Minimum size of data set required for further bifurcation to the next step
2 Minimum size of data set required for a daughter node to be created/ its parent node to be split

Cart analysis indicates that “b”, “c” and “e” impact revenue per 1,000 population

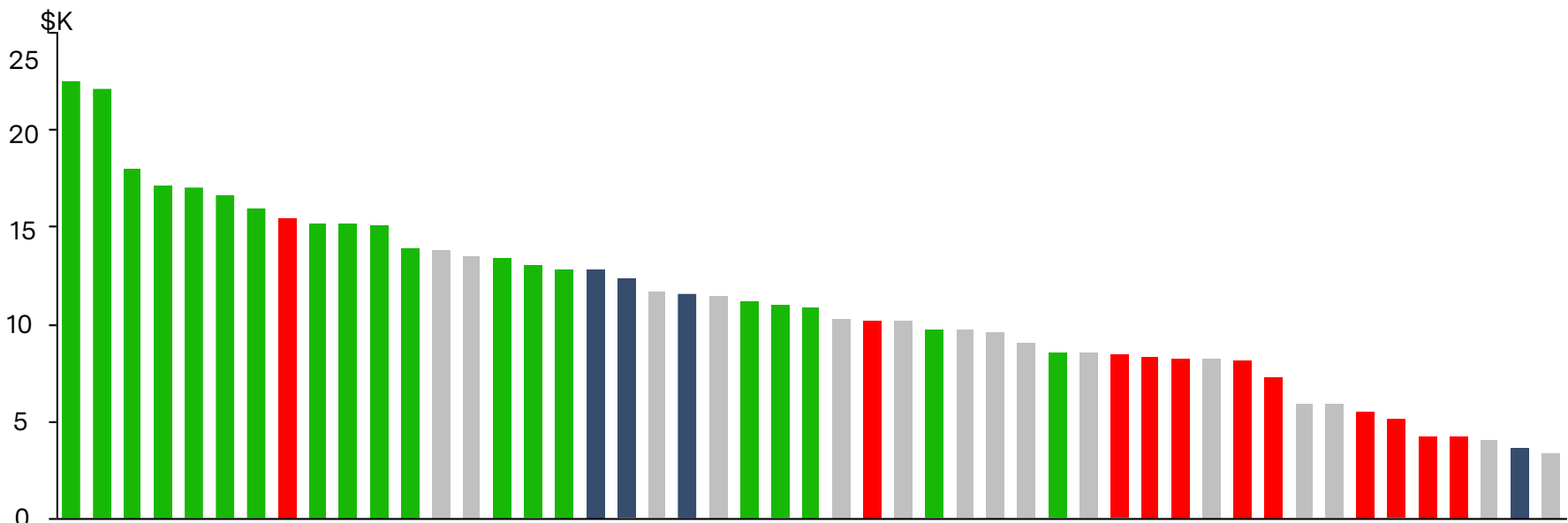
Number of restaurants by performance¹, #



¹ Restaurant are categorized as High, Medium and Low performance based on the percentile (top 33 as H, bottom 33 as L and the remaining as M) of dependent variable - revenue per 1,000 population. The number represents the count of restaurants.

Final restaurants categorization based on performance drivers

‘Restaurant revenue per 1,000 population’ by restaurant and CART terminating nodes, \$K



- Segment A (Low performance): % Bachelors+ education <= 0.36, JI Score < 81
- Segment B (Medium performance): % Bachelors+ education <= 0.36, JI Score >= 81
- Segment C (Medium performance): % Bachelors + education > 0.36, Median Home Value >= \$548,000
- Segment D (High performance): % Bachelors+ education > 0.36, Median home value < \$548,000

Restaurant locations

- **Similar analysis (RF & CART) was done on ‘site characteristics’** and accordingly, restaurants were segmented
- **Key learning was that** restaurants with good **patio visibility from primary road** in areas with **low-to-medium restaurant density** are likely to have higher ‘revenue