Section 3: Week 5: Mining Theories Through Data Problem

Nate Bachmeier

TIM-8130: Data Mining

February 9th, 2020

North Central University

Imagine Boston 2030

As the city of Boston continues to grow, the Mayor has recognized that not all citizens are equally gaining those rewards. The dichotomy of the metropolitan area has led to distinct clusters, disproportionally filled with wealth or crime. Nearly one in five residents live in poverty and does not have a path to improve their situation. Instead, Walsh asks the community to imagine a different Boston, one that within a generation creates access to affordable housing, regulates climate change, and wage inequality (Kovatch et al., 2017). This transformational process not only raise the standard of life for Bostonians, it also encourages investment through tourism and economic expansion. Meeting their ambitious goals requires a strategy to track progress and ensure the progressive agenda is unfolding correctly. The Boston Planning and Development Agency (2017) publishes several Key Performance Indicators (KPI) around each of the major pillars of the plan, in addition to estimating January 2000 to present. This approach produces a highly efficient hierarchical reporting structure that aligns the mission with specific questions and supporting evidence. For instance, the objective *Reduce Housing Cost Burden for Bostonians* qualifies success as low-income residents are not living paycheck to paycheck, and quantifies it through the ratio of wages to housing costs. According to their chart, the ratio has held steady for low-income residents between 45-50% since 2000. When a population does not have disposable income, they are at high risk that minor incidents can become life-impacting. Consider the cascade of events from a 150$ auto-part failure, a broken leg, or a sick child—each scenario forcing the person to take on debt, further compounding their problems, and getting further from the finish line. These human problems need solutions, but what moves the needle the most in the least amount of time? Merely writing a blanket check or applying broad policy can be ineffective without addressing the crux of the issue.

## Improving Housing Affordability

The improvement plan suggests that the price of housing is too high, and this limits access to a broad audience. While their evidence supports this observation, does it tell the entire story? Wikipedia (2020) states that Boston is 89.63 square miles, of which land encompasses 48.42, and this needs to hold approximately 700,000 residents. According to the US Census, the population is consistently increasing at 1.25% annually (ZipAtlas, 2020). With the supply of land being finite and the continuous increase in demand, one needs to reassess the initial statement. That reassessment requires an agreement whether housing prices (a) are too high or (b) are fair. Redfin (2020) specializes in listing housing quotes and supports exporting query results into Comma Separated Value (CSV) files. An analysis of 351 homes in the Boston area shows that the median listing time is five days, and the 95th percentile is sixteen days. If houses are selling within one to two weeks, that suggests that prices are fair and a different knob needs to turn (e.g., wages or commutability). Reviewing descriptive statistics shows the housing quartiles as (600k, 900k, 1.3M, 7.3M) and highlights the diversity price ranges available. These ranges naturally form in alignment with the demographic distribution of the city. For instance, in the financial district (02110), housing is at the 75th percentile with median income at 152,000$. Roxbury (02119), a predominately black neighborhood, has a median income of 27,000 and housing 25th percentile.