

TRANSPORT FINDINGS

An Empirical Taxonomy of Common Curb Zoning Configurations in Seattle

Chase P. Dowling¹  ^a, Thomas Maxner² ^b, Andisheh Ranjbari³ ^c¹ Physical and Computational Sciences Directorate, Pacific Northwest National Laboratory, Seattle, WA, ² Civil and Environmental Engineering, University of Washington, Seattle, WA, ³ Civil and Environmental Engineering, Pennsylvania State University, State College, PA

Keywords: curb zoning, k-modes, unsupervised clustering, data science, transportation

<https://doi.org/10.32866/001c.32446>

Findings

This work applies an unsupervised clustering algorithm to blockface zoning data to identify typical curb configurations in a city. Data is obtained via the city of Seattle's (Washington, USA) open data portal. To compare the distribution of blockfaces of varying length, all blockfaces are normalized where each zone type is presented as a percentage of the total blockface length in an order-preserving format. Common zone sequences are identified via k-modes clustering, where an optimal choice of k is cross-validated, quantifying the number of curb configurations to represent the majority of Seattle's blockfaces. All documented code and data are open source and available at <https://github.com/pnnl/curbclustering>.

1. Questions

In Seattle, WA, there are 2697 unique blockfaces,¹ for which high resolution data is maintained and made publicly available. Each blockface possesses distinct curb zoning configurations, comprised of anywhere between one to a dozen of 60 unique zone types and features, such as paid parking, commercial vehicle loading, passenger pick-up/drop-off, bus lane, bike parking, curb bulbs, crosswalk, driveway, and so on. This zoning data exists in areas where on-street parking fees are actively collected—called paid areas—largely in the core business district of Seattle.

In studies aiming to simulate the impact of different curb zoning configurations on traffic congestion, it is critical to have a tractable number of typical curb configurations from real, existing layouts in order to simulate in a reasonable amount of time. Alternatively, a municipality may wish to *categorize* blockfaces by curb zoning configuration pattern to compare curb performance data (e.g. occupancy, productivity, efficiency, and accessibility). Thus, in both cases we ask, what might a typical cross section of real curb configurations in the core business district of a city like Seattle look like? This question is similar in spirit to that of Raman and Roy (2019), which takes an expert-informed hierarchical approach to similarly derive a taxonomy for urban land use. In the

a chase.dowling@pnnl.gov

b tmaxner@uw.edu

c ranjbari@psu.edu

¹ In this work we adopt the Seattle Department of Transportation's definition of a "blockface" to be the length of city infrastructure abutting a roadway, bounded between two roadway intersections (SDOT 2021). This work is concerned with the length of curb, or raised boundary between the roadway and pedestrian sidewalk, and its zoning along individual blockfaces.

case of curb zoning along individual blockfaces, this might also be achieved through expert selection of typical curb zoning configurations of a blockface. The data-driven approach proposed in this study can be used to supplement such expert selection.

2. Methods

We utilize an unsupervised learning algorithm called k -modes clustering (Huang 1998), which is similar to the better-known k -means method (Hartigan and Wong 1979), but with a dissimilarity measure designed for categorical variables (Cao et al. 2012), originally developed for analyzing sequential categorical data such as gene sequences (Goodall 1966), but also amenable to curb zoning types. For a specified k , the k -modes algorithm finds the top k vectors that minimize a distance to all sample vectors in the training dataset. The resulting top k modes are representative of distinct clusters of sample vectors, with cluster membership determined by the closest mode. The parameter k is chosen through cross-validation by holding out portions of the available training data and finding the smallest k that largely minimizes the within-cluster variation in this hold-out set (also called the “elbow method”). We utilize basic matching dissimilarity, as implemented in (Vos 2015). For two vectors \mathbf{u} and \mathbf{v} of length d , where each element attains categorical values, matching dissimilarity δ is defined as,

$$\delta(\mathbf{u}, \mathbf{v}) = \sum_{i=1}^d \mathbf{1}[u_i \neq v_i], \quad (1)$$

where $\mathbf{1}$ denotes the indicator vector, with value 1 where the bracketed condition is true and 0 otherwise. We’ve chosen this measure of dissimilarity between two sets of categorical variables for a number of reasons: 1) its simplicity, 2) successful use in categorical data clustering (Goodall 1966), and 3) its sensitivity to the ordering of values when vectors u and v are ordered, specific to how we have chosen to represent curb zoning data.

Seattle does not demarcate individual spaces along a blockface, so we suppose a blockface is broken into a vector of incremental lengths, with each unit of length corresponding to a unique zoning type. Because blockfaces are of varying length in Seattle, distributed as illustrated in [Figure 1](#), we normalized these incremental lengths to percentages of total blockface length to make direct comparisons between their constituent curb zones. For example, for a sample vector $\mathbf{x} = \langle x_1, x_2, \dots, x_{100} \rangle$, each element (e.g. x_1) denotes the zoning type for 1% of the total length of a particular blockface, in sequence, with respect to the direction of travel along the roadway immediately adjacent.

Blockface zoning data are collected from Seattle’s open data portal (SDOT 2019), and normalized to equal-length vectors. For each vector, zoning types are represented as a percentage of the blockface’s total length, in an order-preserving (with respect to direction of traffic) format. The 60 unique zoning

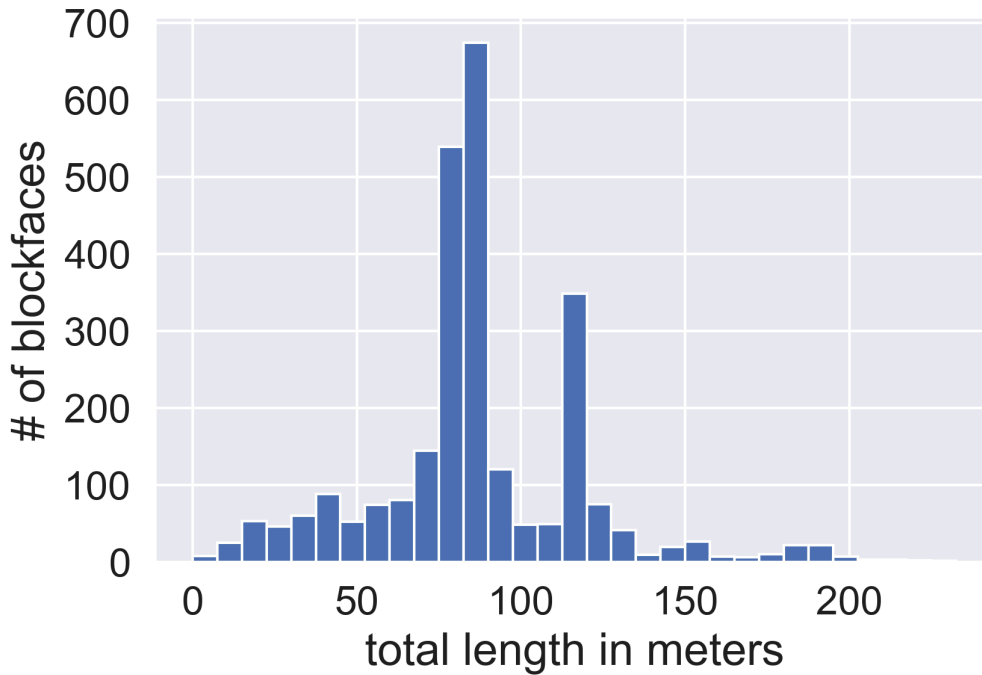


Figure 1. Distribution of blockface lengths in Seattle, WA in increments of 7.26m (25 feet, a common quantization for Seattle). Zoning data, recorded by the city in feet, is accurate to within approximately 3 cm, or 0.1 feet.

types are condensed to a set of 16 zoning labels (e.g. combining bus waiting zones and bus stops to just bus zones, variations on commercial loading, etc.)². We select an appropriate k via cross-validation, minimizing within-cluster variation in a held-out set comprising 10% of the training data. We performed clustering on four sub-samples of the Seattle data, as noted below, but for the sake of brevity, only one of these experiments (#4) is presented in this paper.

1. All paid area curb zonings by blockface, regardless of blockface's total length (2679 blockfaces), both with the full set of all 60 curb zoning labels, and with the condensed set of 16 curb zoning labels.
2. All paid area curb zonings by blockface of similar unnormalized length (i.e. 83.82 - 91.44m (275-300 feet)) (674 blockfaces)
3. Curb zonings near the downtown core, regardless of total blockface length (1695 blockfaces)
4. Curb zonings near the downtown core by blockfaces of similar unnormalized length (i.e. 83.82 - 91.44m (275-300 feet)) (449 blockfaces) **The result of this experiment is presented in this paper**

² All experiments are conducted with the condensed label set unless otherwise specified. The condensed label set—i.e. which zoning types are combined—is fully documented in this work's associated code.

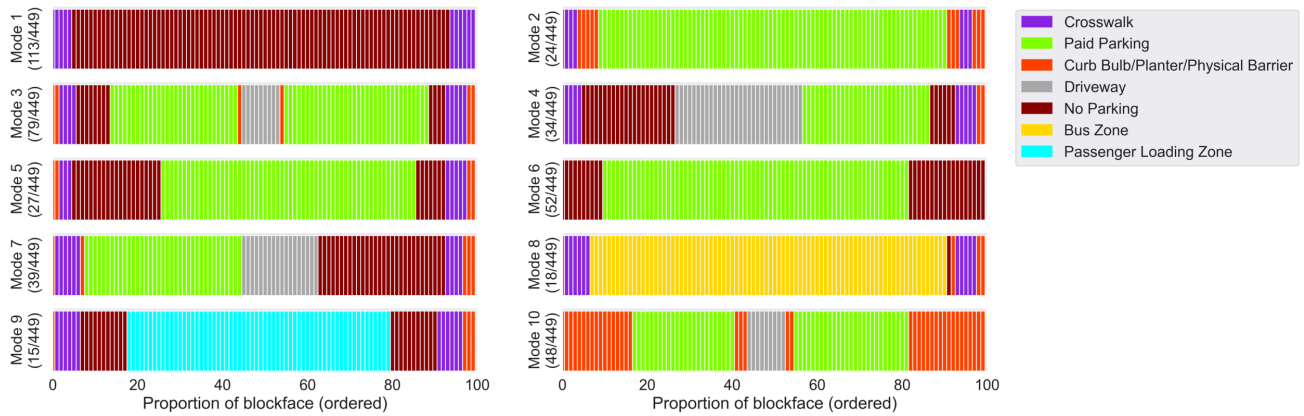


Figure 2. Top 10 blockface (83-91 meters long) configurations in the core downtown of Seattle. Blockfaces per cluster (449 blockfaces total in this experiment’s subsample) is given next to each configuration in parentheses.

Data (and original sources), pre-processing, and documented code to reproduce our results on all four experiments can be found at <https://github.com/pnnl/curbclustering>. In the associated code documenting additional results, we also highlight where dissimilarity measures other than Equation (1) can be substituted and evaluated.

3. Findings

[Figure 2](#) depicts ten (found via cross-validation) typical curb configurations in downtown Seattle for blockfaces between approximately 83 - 91 meters (275-300 feet) long, where each color bar represents 1% of a blockface’s total length. The 449 clustered blockfaces were all non-highway (excluding Interstate 5, Interstate 90, and State Route 99) located in the core business district (within approximately 2 km of 122.3353° W, 47.61484° N, a point roughly in the middle of the core business district illustrated in [Figure 3](#)).

Interestingly, not all clustering experiments yielded all 16 condensed zone types in the top k modes; for example, experiment 3 does not find curb zoning configurations which contain passenger loading zones in the top 12 modes. This result depends on a number of factors: 1. modes identified by k -modes are not unique, and can differ slightly by initialization of the algorithm³ (Cao et al. 2012), 2. choice of dissimilarity measure, and 3. blockfaces represented in the experiment sample. Thus it is critical to consider expert evaluation in tandem with the resulting modes identified by this algorithm.

Curb zoning patterns along blockfaces with the top 10 configurations are illustrated in [Figure 2](#). The results show that the most common configuration in downtown Seattle is Mode 1, representing a no-parking blockface all-along. This is followed by Mode 3, 7, and 8, representing blockfaces with the majority (>50%) of their length being assigned to paid parking, with or without the

³ The random seed of this paper’s experiments is documented and fixed in the accompanying code



Figure 3. (a) The locations of all zoned curb data on blockfaces 83-91 meters (275-300 feet) long in the study area (blue oval), (b) those blockfaces with curb zoning configurations labeled as Mode 3 in Fig. 2, and (c) zooming in, we see these blockfaces are correctly identified as split, according to the curb zoning pattern in Mode 3, by the presence of an alleyway.

presence of a driveway or garage entrance. [Figure 3a](#) shows the locations of all blockfaces 83-91 meters (275-300 feet) long in the study area and [Figure 3b](#) shows the subset of these with curb zoning patterns classified as Mode 2, the second most frequently classified type. As can be seen these correspond with the locations of alleyways splitting non-arterial blockfaces in half, also illustrated in [Figure 3c](#). This indicates that relevant patterns in curb zoning are being detected.

These findings can be used, supplementing expert choices with a data-driven approach, to determine a select number of typical curb zoning configurations in Seattle, for purposes such as developing a microscopic simulation platform to capture a representative breadth of curb activity. The developed clustering methodology can also be applied to any other city for which curb zoning data by length is available. This taxonomic method combined with expert knowledge, can also be used for categorizing and simplifying reporting on curb zoning performance in a city, reducing performance metrics to a limited number of curb zoning patterns per blockface.

Acknowledgements

Pacific Northwest National Laboratory is operated by Battelle Memorial Institute for the U.S. Department of Energy under Contract No. DE-AC05-76RL01830. This work was supported by the U.S. Department of Energy Vehicle Technologies Office.

Submitted: December 09, 2021 AEDT, Accepted: February 03, 2022 AEDT



This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CCBY-SA-4.0). View this license's legal deed at <https://creativecommons.org/licenses/by-sa/4.0> and legal code at <https://creativecommons.org/licenses/by-sa/4.0/legalcode> for more information.

REFERENCES

- Cao, F., J. Liang, D. Li, L. Bai, and C. Dang. 2012. “A Dissimilarity Measure for the K-Modes Clustering Algorithm.” *Knowledge-Based Systems* 26: 120–27.
- Goodall, David W. 1966. “A New Similarity Index Based on Probability.” *Biometrics* 22 (4): 882. <https://doi.org/10.2307/2528080>.
- Hartigan, J.A., and M.A. Wong. 1979. “Algorithm as 136: A k-Means Clustering Algorithm.” *Journal of the Royal Statistical Society*, 100–108.
- Huang, Z. 1998. “Extensions to the K-Means Algorithm for Clustering Large Data Sets with Categorical Values.” *Data Mining and Knowledge Discovery* 2 (3): 283–304.
- Raman, R., and U.K. Roy. 2019. “Taxonomy of Urban Mixed Land Use Planning.” *Land Use Policy* 88: 104102.
- SDOT. 2019. “Paid Area Curbspaces.” <https://data-seattlecitygis.opendata.arcgis.com/datasets/SeattleCityGIS::paid-area-curbspaces/about>.
- . 2021. “Seattle Department of Transportation Blockface Elements Data.” <https://data.seattle.gov/dataset/Blockface/m47t-7pxp/data>.
- Vos, N.J. 2015. *Kmodes Categorical Clustering Library*. <https://github.com/nicodv/kmodes>.