Protecting Montpelier: Strategies for Flood Risk Mitigation

Business Analytics Capstone Project

PROJECT REPORT
FALL 2024

PREPARED BY:
Shraddha Jain, Sana Parab, Sidhant Kumar, Ashikul Kabir, Arjun Malgwa, Ankit Akash

precisely

DREXEL UNIVERSITY
LeBow
College of Business

EXECUTIVE SUMMARY

Montpelier, Vermont faces increasing flood risks, with recent events in 2024 revealing significant vulnerabilities in current risk assessment and insurance practices. These challenges underscore the need for data-driven solutions to address issues like underinsurance, rising claims, and inaccurate risk categorization.

Our Drexel LeBow Capstone Team aimed to address these challenges by providing actionable insights for property insurers. The key objectives were to assess property vulnerability to flood risk, identify underinsured properties, determine opportunities for upselling flood insurance, and develop data-driven recommendations for improved flood risk management.

Using data from 10,311 properties enriched through Precisely's Geoaddressing and Floodrisk APIs, we employed K-means clustering to group properties into five risk categories based on features like proximity to the Winooski River, elevation, and building characteristics. High-risk clusters (Clusters 4 and 5) were identified, providing insurers with a clear understanding of the most vulnerable properties and targeting these for high-margin policies to offset potential losses . Cluster 3, the largest group, represents a key target for customized policy bundles with flood add-ons, enhancing coverage while driving revenue growth for insurers.

We also identified underinsured properties based on insurance-to-value (ITV) ratios, flagging those below 80%, with an average underinsurance amount of $147,689 in Vermont. This presents a significant upselling opportunity for insurers to improve coverage and ensure financial protection for property owners.

Overall, our solutions successfully met the project goals by providing insurers with advanced tools for precise risk assessment, enhanced flood risk categorization, and effective strategies to address underinsurance. These insights equip insurers to make better decisions, improve customer satisfaction, and contribute to a more resilient and financially secure community.

BUSINESS CHALLENGE

Montpelier faces significant challenges due to frequent flooding from the Winooski River, which has led to severe disruptions to properties and infrastructure, particularly in the recent floods of 2023 and 2024 (Montpelier Flood Guide, 2024). The July 2023 flood submerged downtown Montpelier roads, caused extensive transportation disruptions, and led to widespread flooding throughout the Winooski River valley, affecting both residential and commercial areas. Mobile homes, in particular, were severely impacted, with many units condemned due to flood damage. The July 2024 flood further exacerbated these issues, creating significant pressure on insurers to manage claims and mitigate financial impacts.

These floods exposed gaps in traditional risk assessment and claims management, putting pressure on local insurers to manage increasing claims and accurately assess risks. Local insurers often rely on generalized data sources, which fail to capture the localized flood risks that Montpelier faces. This has resulted in inadequate insurance coverage and rising liabilities for insurers. The need for precise data-driven solutions is critical to better understanding property vulnerability, managing risk, and ensuring adequate coverage.

LITERATURE/INDUSTRY REVIEW

According to the Department of Homeland Security, 90% of natural disasters within the United States involve flooding. Floods cause more economic damage, loss of life, and property than any other natural hazard (Department of Homeland Security, 2024). The FEMA National Flood Insurance Program (NFIP) faced similar challenges in assessing flood risk across various geographic regions. The NFIP leveraged geospatial data to classify properties into flood zones and introduced specialized insurance products to manage these risks. Despite improvements, challenges with underinsurance and risk awareness remain (FEMA, 2024).

Zurich Insurance Group implemented geospatial risk assessment methods to enhance flood risk modeling. By integrating elevation data and flood zone classifications, Zurich tailored coverage options and accurately priced policies, resulting in improved customer retention and reduced claims (Zurich Insurance Group, 2024).

DATA REVIEW

Data collection, validation and cleaning

We created a comprehensive dataset for accurate analysis by utilizing data from the address fabric comprising 10,311 properties in Montpelier, Vermont. Additional data was sourced through Precisely’s APIs, focusing on geospatial attributes, property characteristics, and flood risks. The dataset was enriched through multiple API queries, including standardized address formatting using Precisely's Geoaddressing API and adding flood risk attributes with the Floodrisk API. After cleaning and enrichment, the final dataset contained 10,170 unique records, which was further refined to 4,184 unique properties by merging with attributes from Precisely’s propertyAttributesByAddress API.

Key attributes of the dataset included geospatial features such as proximity to the 100-Year Flood Zone, nearest waterbody distance, and elevation above sea level. Property characteristics such as construction type, exterior material, and market value were also incorporated, alongside geographic and flood risk data including FEMA flood zone classifications and proximity to the Winooski River. The data captured in 2024 reflects the most recent flood risks and property assessments.

Data cleaning and preprocessing were crucial for maintaining data quality and consistency. Python was used for data preprocessing, including data cleaning, feature selection, and dataset enrichment. Validation steps ensured no missing values in critical fields, including Unique Property ID, Address Details (City, State, ZIP Code), and Geographic Information (Latitude, Longitude, Location Code). Checks for invalid ZIP codes, incorrect state information, and duplicate rows returned zero issues, ensuring data reliability. Merging was performed to combine multiple datasets into a single enriched dataset, incorporating geospatial, property, and policy-related attributes.

These preprocessing steps ensured that the dataset was clean, consistent, and ready for analysis, ultimately supporting precise risk assessment and providing accurate insights for flood risk management.

DATA ANALYSIS

Solution 1: Clustering Based Risk Assessment

Around **91%** of the properties in our dataset were classified in the FEMA low-risk Zone (Exhibit 1). Considering the fact that 40% of claims actually come from low-to-moderate risk areas, analyzing flood risks based on other factors is crucial, and that's what our clustering model does. To proceed with our analysis, we created a custom variable - ***Distance Factor*** (including 'Year100FloodZoneDistanceFeet', 'DistanceToNearestWaterbodyFeet', 'AddressLocationElevationFeet'). We analyzed risk using three key data segments:

- Distance Factor: proximity to 100-year flood zones, water bodies, and elevation levels.
- Property Characteristics: construction type, exterior walls, etc.
- Geographic and Flood Risk Data: proximity to the Winooski River, flood zones.

To identify the most relevant variables, we did feature engineering using Principal Component Analysis (PCA) to reduce dimensionality, narrowing down to the most impactful features. We also did feature variance analysis to identify features with the highest variance.

We used the K-means clustering algorithm to group properties based on these risk factors. This process segmented the properties into five distinct risk categories, ranging from low to high risk. This grouping allows for targeted recommendations, ensuring resources are allocated where they matter the most. To decide on the optimal number of clusters, we used the Elbow Method. By plotting the within-cluster sum of squares (WCSS) against the number of clusters, we identified a "bend" in the curve at five clusters. (Exhibit 2) This indicates a point of diminishing returns—adding more clusters beyond five wouldn't improve the quality of segmentation meaningfully. We performed correlation analysis (Exhibit 3) to remove any redundant features and selected the following variables for our final model in addition to the Distance Factor: Building Construction Type (Frame), Exterior Walls (Wood Siding) - ***to capture vulnerability to damage***

Flood Zone (AE), Flood Zone (SHX), Total Market Value, Nearest Waterbody (Winooski River) - ***to account for exposure to natural hazards***

Why Five Clusters and Not Six?

Choosing six clusters would result in over-segmentation, creating overly fine groups without meaningful differentiation. The five-cluster model strikes the right balance between capturing variance and maintaining interpretability. (Exhibit 4)

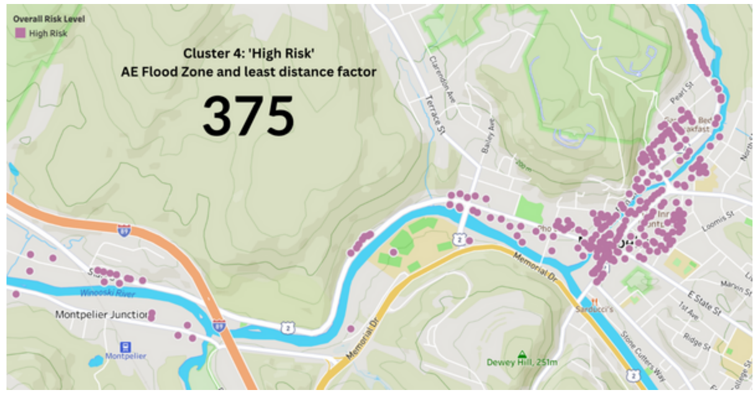
Solution 2: Identifying Underinsured Properties

We identified various properties in our PIF (Policies in Force) dataset which had considerable differences in the Policy value and the property's assessed value. (Exhibit 5)

We created another custom variable: ITV Ratio, which measures the policy value relative to the property's assessed value. (Davies Group, 2024). We used the Policies in Force data for Vermont state and the variables including Policy value, assessed property value, and ITV ratio which we analyzed to identify underinsured properties. Properties with ITV ratios below 80% were flagged as underinsured (Liberty Home Guard, 2024), with an average underinsurance amount of $147,689 in Vermont. We added an automation for automated report generation that fetches underinsured properties along with the underinsured amount. (Exhibit 6). Tableau and MapInfo Pro were used for visual representation of flood risk clusters and underinsured properties.

BUSINESS INSIGHTS

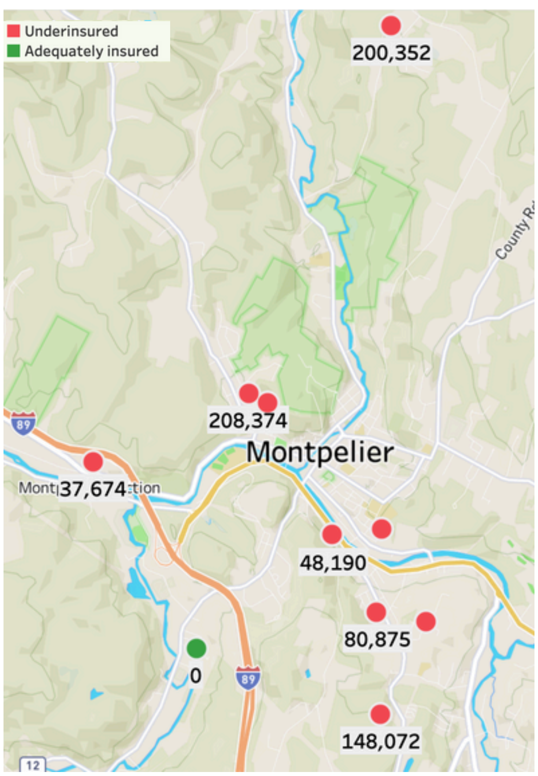
The insights from our clustering model empower insurers to prioritize high-risk zones for tailored policies, explore upselling opportunities in moderate-risk areas, and optimize resources for low-risk properties.



Below are the insights for the clusters for targeted Insurance companies:

- Cluster 4 (High Risk): **375** Properties in AE flood zones with low distance factors, indicating high risk. Focus on high-margin policies and risk-sharing partnerships with reinsurers.
- Cluster 1 (Low Risk): **1175** Properties far from the Winooski River, indicating minimal risk. Opportunity for low-premium loyalty programs and upselling multi-policy packages with flood coverage.
- Cluster 2 (Low-Moderate Risk): **1343** Moderate distance properties with construction attributes that reduce risk. Suitable for loyalty programs and bundling flood coverage with other policies.
- Cluster 3 (Moderate Risk): **1123** Properties near the Winooski River with wooden walls. Largest cluster and key upselling opportunity. Customized policy bundles with flood add-ons are recommended.
- Cluster 5 (Moderate-High Risk): **168** Properties near the Winooski River with significant risk. Recommend risk-sharing partnerships and high-margin policies tailored to cluster-specific risks.

These strategies, combined with data-driven insights, can improve insurers' ability to manage flood risk, enhance customer satisfaction, and ensure financial stability for affected communities.



Underinsurance Identification: The analysis revealed that 9 out of 10 properties in the PIF data for Montpelier were flagged as underinsured, with many exceeding Vermont's average underinsurance value of $147,689. This highlights a significant gap in coverage that increases financial vulnerability for policyholders and unanticipated liabilities for insurers. Addressing these gaps creates an opportunity for revenue growth through targeted upselling strategies. Our automated report also provides a data driven solution to identify these properties and have a focused approach to target these properties to boost coverage and drive revenue.

pbKey	PolicyNo	Policy Value	Property Assessed Value	Underinsured?	Underinsured Amount
P0000N0KCJ3T	479399	89010	137200	1	48190
P0000N0KCH6S	267803	105026	189200	1	84174
P0000N0KCH7Q	267804	105026	313400	1	208374
P0000N0KCMA1	267816	105026	142700	1	37674
P0000N0KCKSK	267815	86025	181400	1	95375
P0000N0KCLOH	498222	95269	40300	0	0
P0000N0KCJQT	267800	139642	367300	1	227658
P0000N0KCLZB	464303	105548	305900	1	200352
P0000N0KCKTJ	443488	113928	262000	1	148072
P0000N0KCJBY	267818	86025	166900	1	80875

Exhibit 6

CONCLUSION AND RECOMMENDATIONS:

Our analysis highlights the transformative role of data-driven approaches in flood risk management. As per Vermont Climate Assessment, Montpelier is projected to have an 80% rise in the likelihood of flooding. (Vermont Climate Action Plan Summary. 2021) This increase is linked to the growing impact of climate change, which is intensifying flood events and challenging insurers to adapt. Our solutions aims to help bridge the gaps in flood risk management, ensure financial stability, and build a more resilient future for Montpelier and beyond.

The recommendations for each cluster based on risk exposure and coverage needs would help insurers as well as policy holders. Strategies such as risk-sharing partnerships and high-margin policies for high-risk properties will improve overall risk management. To sum it up:

Enhanced Risk Assessment: Clustering and ITV analysis refined insurers' ability to allocate resources and price policies accurately.

Improved Policyholder Security: Addressing underinsurance mitigates financial risks for both homeowners and insurers.

Broader Applicability: The methodology serves as a scalable framework, ready for adaptation to other flood-prone regions.

Our further recommendations include:

Tailored Policies: Creating flexible coverage plans with tiered pricing to address diverse risk levels, ensuring both affordability and profitability.

Mitigation Incentives: Encouraging risk-reduction measures with premium discounts and partnerships for flood prevention.

Community Engagement: Educating communities about flood risks and collaborate with governments to bolster resilience efforts.

Appendix

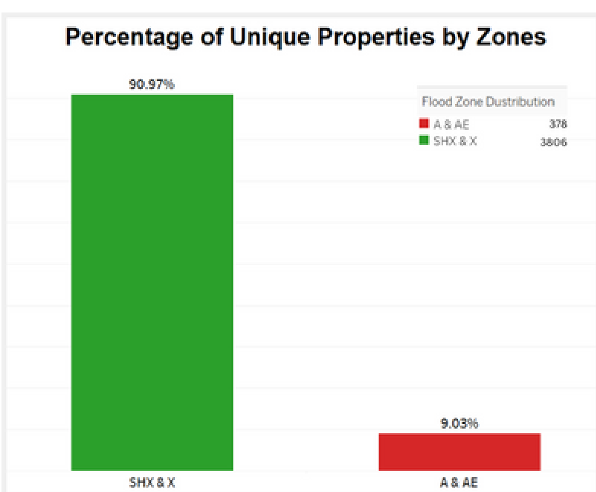


Exhibit 1

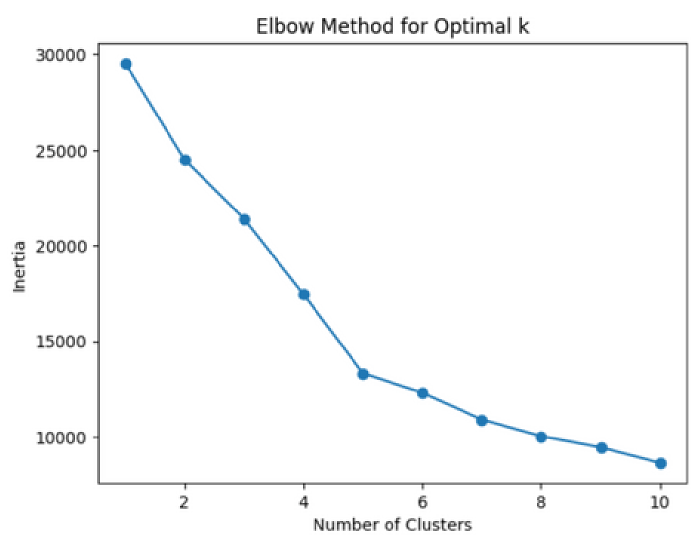


Exhibit 2: Output from the Elbow method

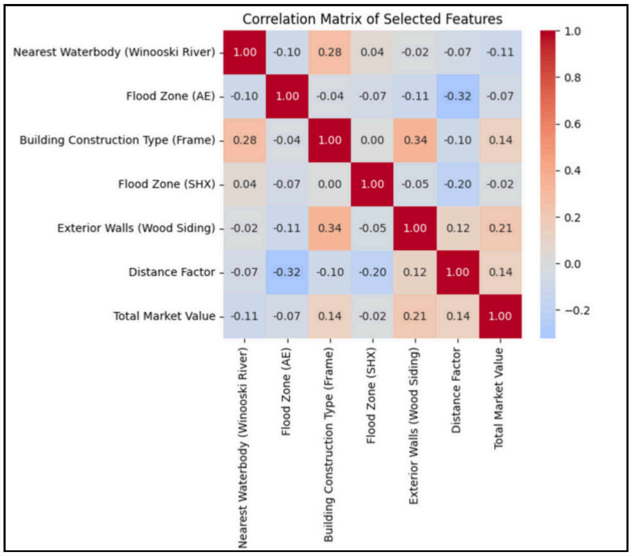


Exhibit 3

Cluster	Proximity to Winooski River	Flood Zone (AE/SHX)	Frame Construction	Exterior Wall: Wood	Distance Factor (ft)	Total Market Value (USD)
Cluster 1	0.51% properties close	AE (0%)	63.7%	99.1%	4115.4	398,256
Cluster 2	16.5% properties close	AE (0%)	24.5%	0.1%	3783.9	323,386
Cluster 3	100% properties close	AE (0%)	97.9%	58.1%	3796.0	332,214
Cluster 4	21.5% properties close	AE (100%)	52.7%	29.6%	871.3	317,906
Cluster 5	44.8% properties close	SHX (100%)	59.8%	35.1%	1016.5	336,554

Exhibit 4

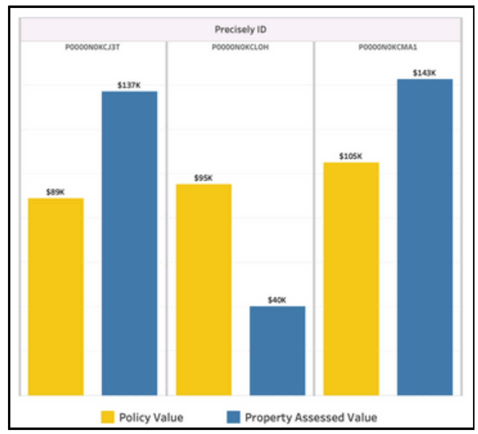


Exhibit 5

References:

Montpelier Flood Guide. (2024). Retrieved from https://www.montpelier-vt.org/610/Flood-Guide

Department of Homeland Security. (2024). Natural Disasters. Retrieved from https://www.dhs.gov/natural-disasters

FEMA National Flood Insurance Program (NFIP). (2024). Flood Risk Management. https://agents.floodsmart.gov/sites/default/files/fema-risk-rating-2.0-fact-sheet-feb-2022.pdf

Zurich Insurance Group. (2024). Flood Risk Assessment Techniques. https://www.zurich.co.uk/media-centre/two-fifths-of-london-firms-at-risk-of-climate-fuelled-flash-floods

Davies Group. (2024). Understanding Insurance-to-Value Challenges. Retrieved from https://davies-group.com/northamerica/knowledge/understanding-insurance-to-value-challenges/

Liberty Home Guard. (2024). Insurance-to-Value. Retrieved from https://www.libertyhomeguard.com/glossary/insurance-to-value/#:~:text=Insurance%20to%20Value%20is%20calculated,of%20100%25%20ensures%20full%20coverage

Vermont Climate Action Plan Summary. (2021). Retrieved from https://climatechange.vermont.gov/sites/climatecouncilsandbox/files/2021-12/VT%20CAP%20Summary_Final_0.pdf