Title: **Clustering Traffic Collision Hotspots in Halifax Regional Municipality**

**Objective:** This project aims to analyze traffic collision data from the Halifax Regional Municipality (HRM) to identify significant patterns and cluster high-risk areas where collisions are more frequent. By utilizing clustering techniques, the goal is to visualize these hotspots and understand the spatial distribution of collisions across the municipality.

**Key Research Questions:**

1. Which geographic areas in HRM experience the highest concentration of traffic collisions?
2. What patterns or similarities exist within the clustered collision areas (e.g., time of day, road conditions, vehicle types)?
3. Can clustering analysis reveal potential correlations between road infrastructure, weather conditions, and the frequency of traffic collisions?

**Proposed Methodology:**

1. **Data Preprocessing:** Clean and prepare the dataset, ensuring all relevant variables (e.g., location, time, weather conditions, type of collision) are available.
2. **Clustering Algorithm Selection:** Apply clustering techniques such as K-Means or DBSCAN to group collision incidents based on location and other relevant factors.
3. **Geospatial Visualization:** Use visual analytics tools like heatmaps or scatter plots to display collision clusters on a map of HRM.
4. **Analysis of Clusters:** Examine the characteristics of each cluster, including collision severity, time of occurrence, and potential external factors (road type, traffic signals).

**Outcome:** The project will generate insights into traffic collision patterns, which could assist city planners and policymakers in implementing targeted safety measures to reduce accidents in high-risk areas.

If you're considering building a web application or using Dash for your visual analytics project around traffic collisions in Halifax, here are some key functionalities you can offer:

**1. Interactive Map Visualization**

* **Clustered Collision Hotspots:** Display an interactive map of HRM with clustering to show traffic collision hotspots. Users can zoom in to see more granular details of each cluster or individual collision points.
* **Filtering by Time/Date:** Allow users to filter the data by specific time ranges (e.g., month, day of the week, time of day) to explore how collision patterns change over time.
* **Weather and Road Condition Overlays:** Enable additional layers on the map that show how factors like weather conditions or road types (e.g., highways vs. city roads) influence collisions.
* **Severity Levels:** Use color coding to indicate the severity of collisions in each cluster, allowing users to quickly identify areas with high-risk collisions.

**2. Collision Data Dashboard**

* **Summary Statistics:** Display summary statistics such as the total number of collisions, most common collision types, and collision severity levels.
* **Trends and Analysis:** Provide line or bar charts to show trends over time, like increases or decreases in collisions across specific periods.
* **Accident Heatmap:** Implement a heatmap view to show the density of accidents across various parts of the city.
* **Geographical and Time-based Analysis:** Enable users to switch between geographical clustering and time-based clustering (e.g., rush hours, weekends) to observe different patterns.

**3. Predictive Insights**

* **Predict Future Collision Hotspots:** Based on historical data, use predictive models to highlight areas that may see a rise in collisions in the future.
* **Weather and Traffic Correlation:** Provide insights on how adverse weather conditions (rain, snow, fog) correlate with the frequency and severity of collisions.

**4. Advanced Filtering Options**

* **Type of Collision Filter:** Users can filter by different types of collisions (e.g., vehicle-pedestrian, vehicle-vehicle, single-vehicle).
* **Vehicle Type:** Offer the ability to filter based on vehicle type involved in collisions (e.g., motorcycles, trucks, bicycles).
* **Time of Day and Week:** Allow for detailed time-based filtering so users can explore patterns based on rush hours, weekends, and holidays.

**5. Detailed Collision Information**

* **Click-to-Expand Collision Details:** When clicking on a specific point in the map or cluster, show detailed information such as the date, time, weather conditions, type of vehicles involved, and severity.
* **Related Incidents:** Show nearby or similar incidents when users select a specific cluster or point on the map.

**6. Recommendations for Action**

* **Suggestions for Traffic Improvements:** Based on identified high-risk areas, the web app could suggest potential safety improvements, such as installing traffic lights, adding speed limits, or improving road lighting.
* **User-driven Reporting:** Users could submit feedback or suggestions on traffic improvements based on the visualized data.

**7. Mobile-Friendly Interface**

* Ensure the web application is responsive, so users can easily access traffic collision data from mobile devices.

**8. Download and Share Data**

* **Export Data:** Allow users to download the visualized data in CSV format for further analysis.
* **Share Insights:** Provide options to share maps or insights through social media or email.

**9. Real-Time Updates**

* **Real-Time Collision Data:** If possible, integrate real-time data to provide up-to-date traffic collision visualizations and alerts.

**Technologies to Use:**

* **Dash by Plotly:** Dash is a great option for creating interactive web applications with minimal code and powerful visualizations. It supports maps, graphs, and real-time updates.
* **Mapbox or Leaflet:** For geospatial visualizations, these libraries integrate well with Dash for creating interactive maps.
* **Backend Integration:** A backend API using Flask or Django can be used to serve data to the Dash app.

By o