

# Unlocking Insights in Calgary Child Care Program Compliance: An Analytical Approach

## Group 39

Sri Harsha Tuttaganti

Venkateshwaran Balu Soundararajan



# Introduction

- ❑ Childcare is essential for early childhood development.
- ❑ Provides a foundation for learning, socialization, and well-being.
- ❑ High-quality childcare significantly impacts a child's growth and future success.



# Objective

- ❖ Utilize the Calgary Childcare Information dataset.
- ❖ Analyze licensed childcare programs.
- ❖ Focus on capacity, compliance status, and inspection history.
- ❖ Ensure high standards of care for children in Calgary resources.



## Target Audience

- Parents: Searching for reliable childcare options.
- Researchers: Analyzing childcare availability and distribution.
- Policymakers: Evaluating childcare needs and planning



# Child Care Information Dataset



[https://data.calgary.ca/Health-and-Safety/Child-Care-Information/qdxh-qngy/about\\_data](https://data.calgary.ca/Health-and-Safety/Child-Care-Information/qdxh-qngy/about_data)

## Overview

- **Dataset Name:** Child Care Information
- **Source:** City of Calgary Open Data Portal
- **Description:**  
This dataset provides information about licensed Child Care facilities in Calgary. It includes details such as facility names, locations, types of care offered, capacity, and contact information. (2014-2023)

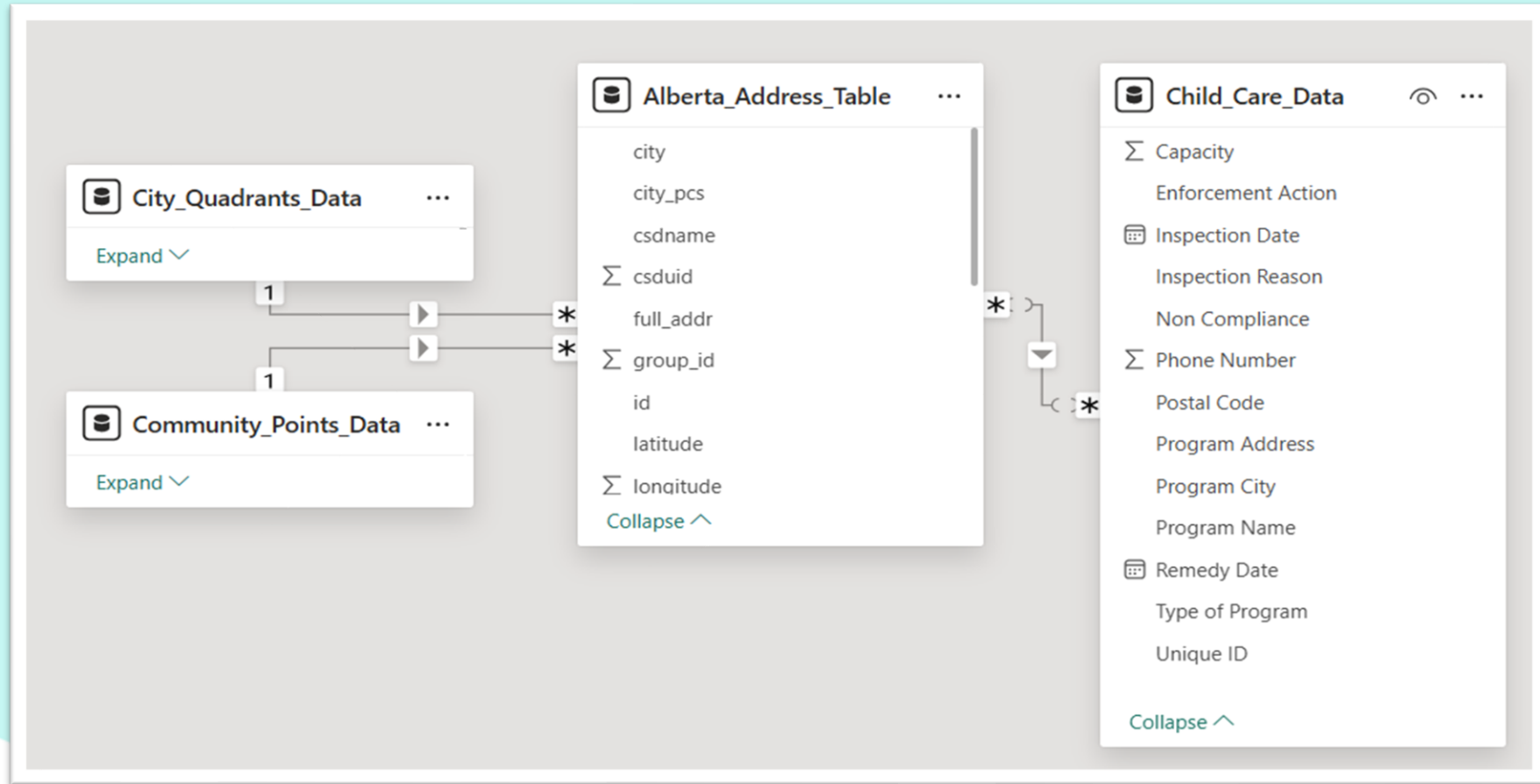


# Data Fields and Its Significance

Field Name	Data Type	Significance
Program Name	Text	Uniquely identifies each program
Type of Program	Text	Categorizes programs for analysis
Program Address	Text	Provides location details
Program City	Text	Identifies the city for location-based analysis
Postal Code	Text	Used for geographic and demographic analysis
Phone Number	Text	Provides a means of contact for inquiries
Capacity	Number	Indicates the program's scale and capacity
Inspection Date	Floating Timestamp	Tracks the timeline of inspections
Inspection Reason	Text	Provides insights into compliance and regulatory focus areas
Non Compliance	Text	Assesses compliance status and identifies areas needing improvement
Enforcement Action	Text	Details corrective actions implemented
Remedy Date	Floating Timestamp	Tracks deadlines for resolving non-compliance issues



# Data Model with Supporting Datasets





# Data Cleaning and Transformation

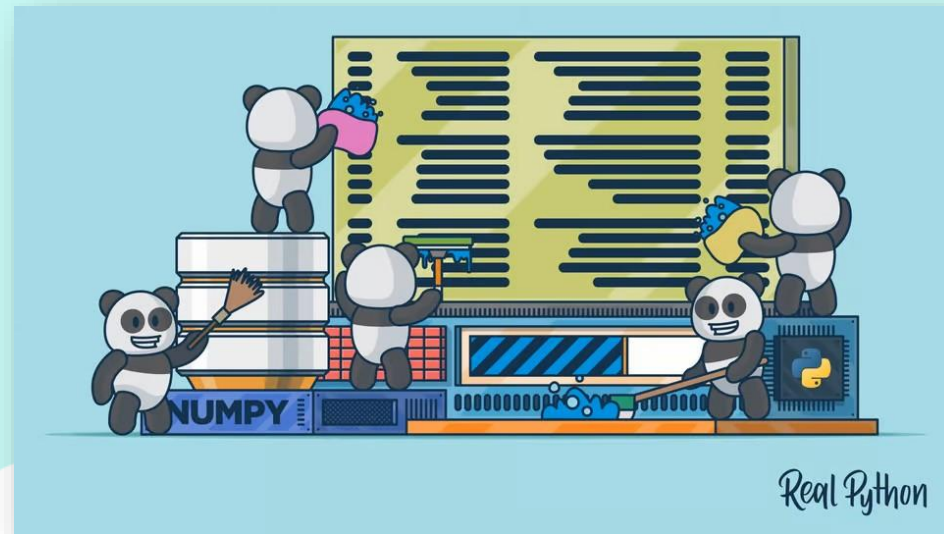
**Step 1 :** Cleaned the data by removing duplicate values.

**Step 2 :** Added a dataset containing latitude and longitude information.

**Step 3 :** Removed duplicates caused by merging location dataset.

**Step 4 :** Addressed normalization issues by cross-checking addresses in both datasets and rectifying discrepancies.

**Step 5 :** Added another dataset to group data into communities and Quadrants.



# Data Preprocessing- Iterative Imputation

**Step 1: Removed rows containing null values**

- Created a cleaned dataset by removing rows with null values.

**Step 2: Split the cleaned dataset**

- Divided the cleaned dataset into test (20%) and train (80%) datasets.

**Step 3: Applied iterative imputation using random forest**

- Used the random forest model for iterative imputation to predict missing values.

**Step 4: Integrated the predicted values**

- Integrated the predicted values back into the original dataset.

**Step 5: Performed imputation on specific columns**

- Applied the imputation process on the "Program Name" and "Type of Program" columns.

**Type of Program column:**

- Accuracy: 0.9126

**Program Name column:**

- Accuracy: 0.9517





# GUIDING QUESTIONS

7

Program Types

1004

Unique Programs

14

Inspection Types

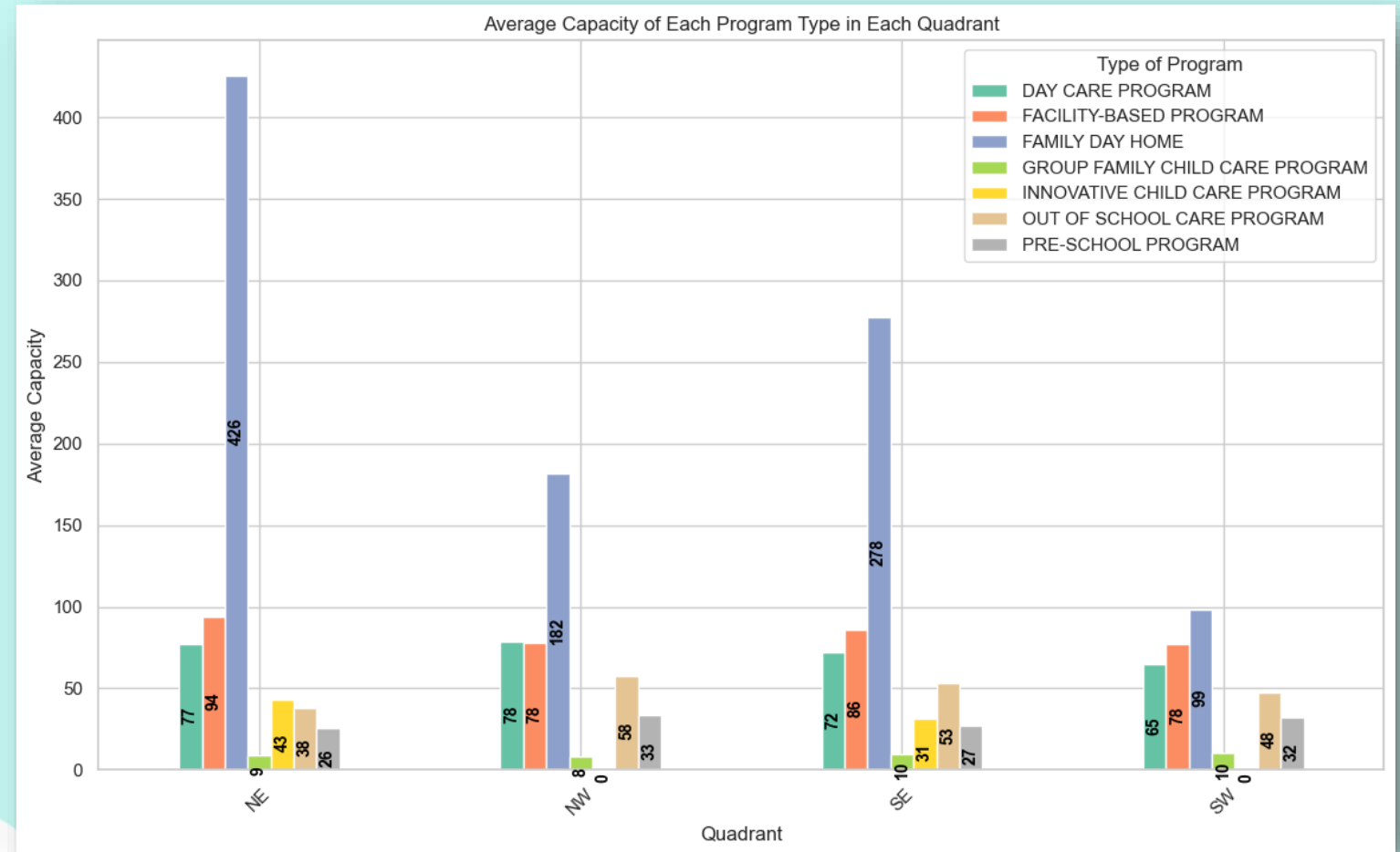


# Program Types and its distribution in Calgary

What is the average capacity of child care programs by its types, how does it vary by its Quadrant

## Insights:

- Highest Average Capacity: Across all quadrants, the FAMILY DAY HOME program consistently has the highest average capacity, indicating its dominance in each region.
- Variation Across Quadrants: The average capacity of program types varies significantly across different quadrants.
- Program Diversity: Each quadrant offers a mix of program types, with notable variation in their capacities.

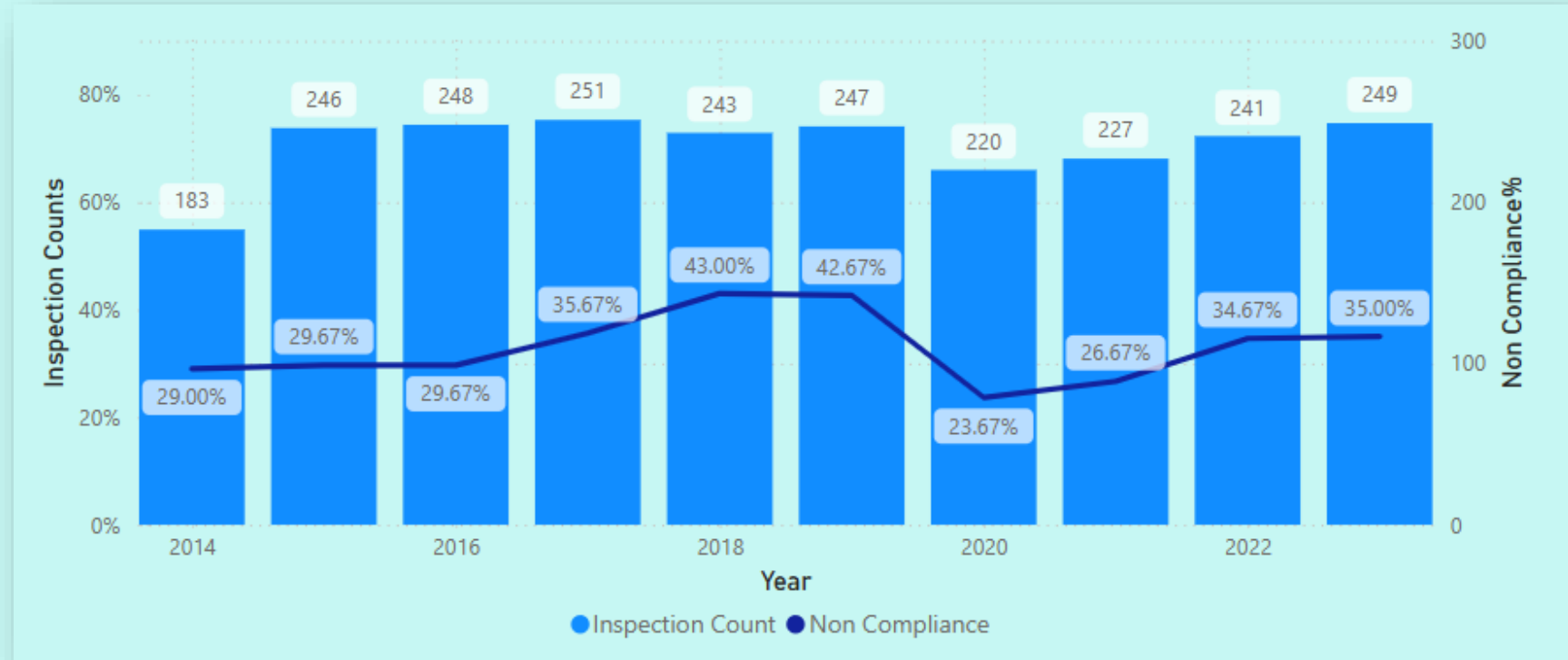


# Year on Year Inspections and its Non- Compliance Trends

How do the average number of inspections and the rate of non-compliance change year-on-year across different childcare programs in Calgary?

## Insights:

- 2016-2019 Increase: There was a noticeable increase in non-compliance percentages from 2016 to 2019.
- 2020 Dip: Both inspection counts and non-compliance percentages dipped significantly in 2020.
- Post-2020 Recovery: There was a recovery in inspection counts post-2020, but non-compliance percentages remain fluctuating.



# Inspections Reasons frequency with Program Type

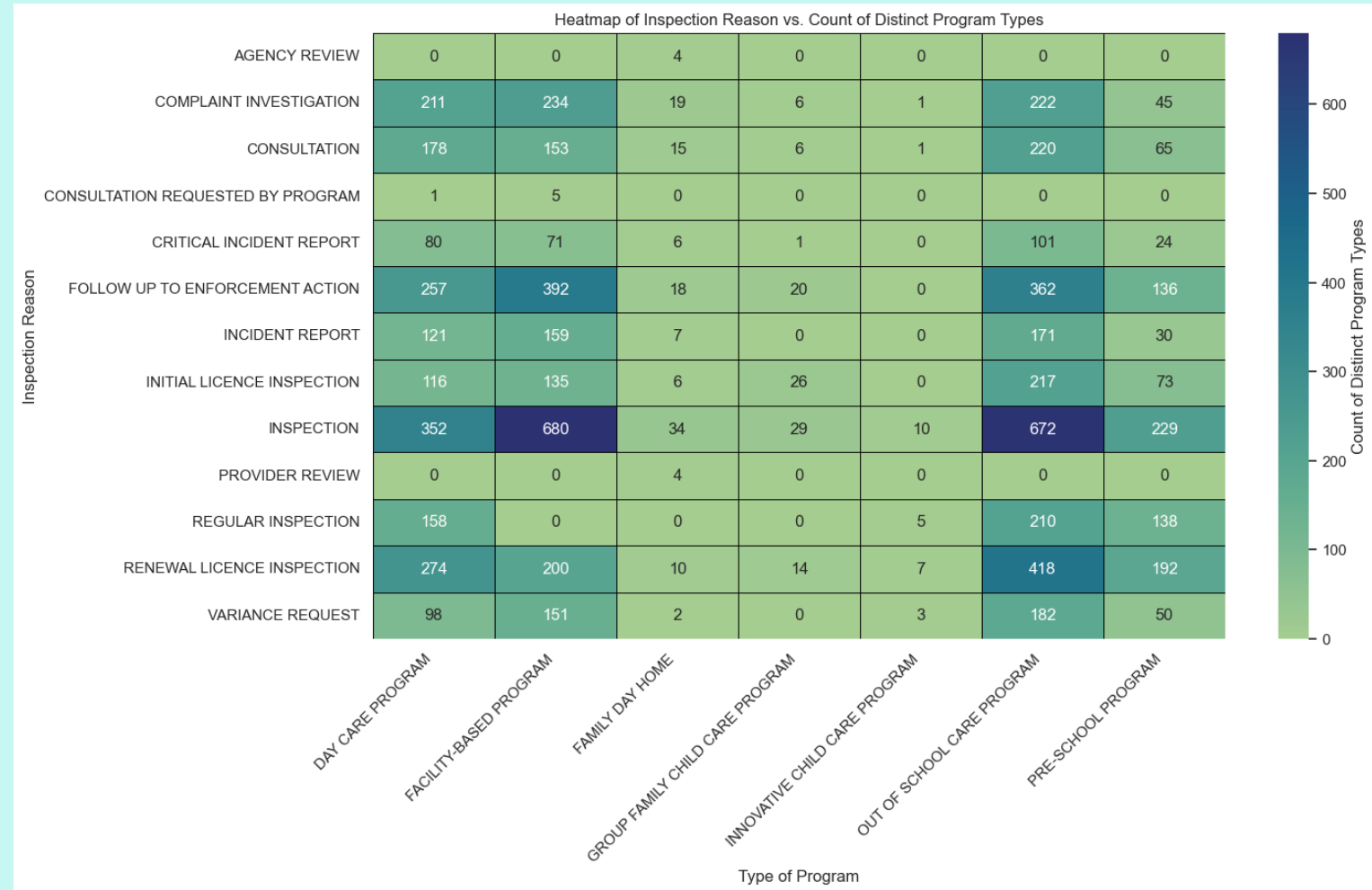
What are the most common reasons for inspections across different types of childcare programs in Calgary?

## Insights:

**High-Frequency Inspections:** **INSPECTION** and **COMPLAINT INVESTIGATION** are common across most program types, indicating regular oversight and response to complaints.

**Program-Specific Trends:** Certain program types like Facility-Based Programs and Family Day Homes are more frequently inspected.

**Focus Areas:** The high counts in **FOLLOW UP TO ENFORCEMENT ACTION** and **RENEWAL LICENCE INSPECTION** suggest these are critical areas needing close monitoring.

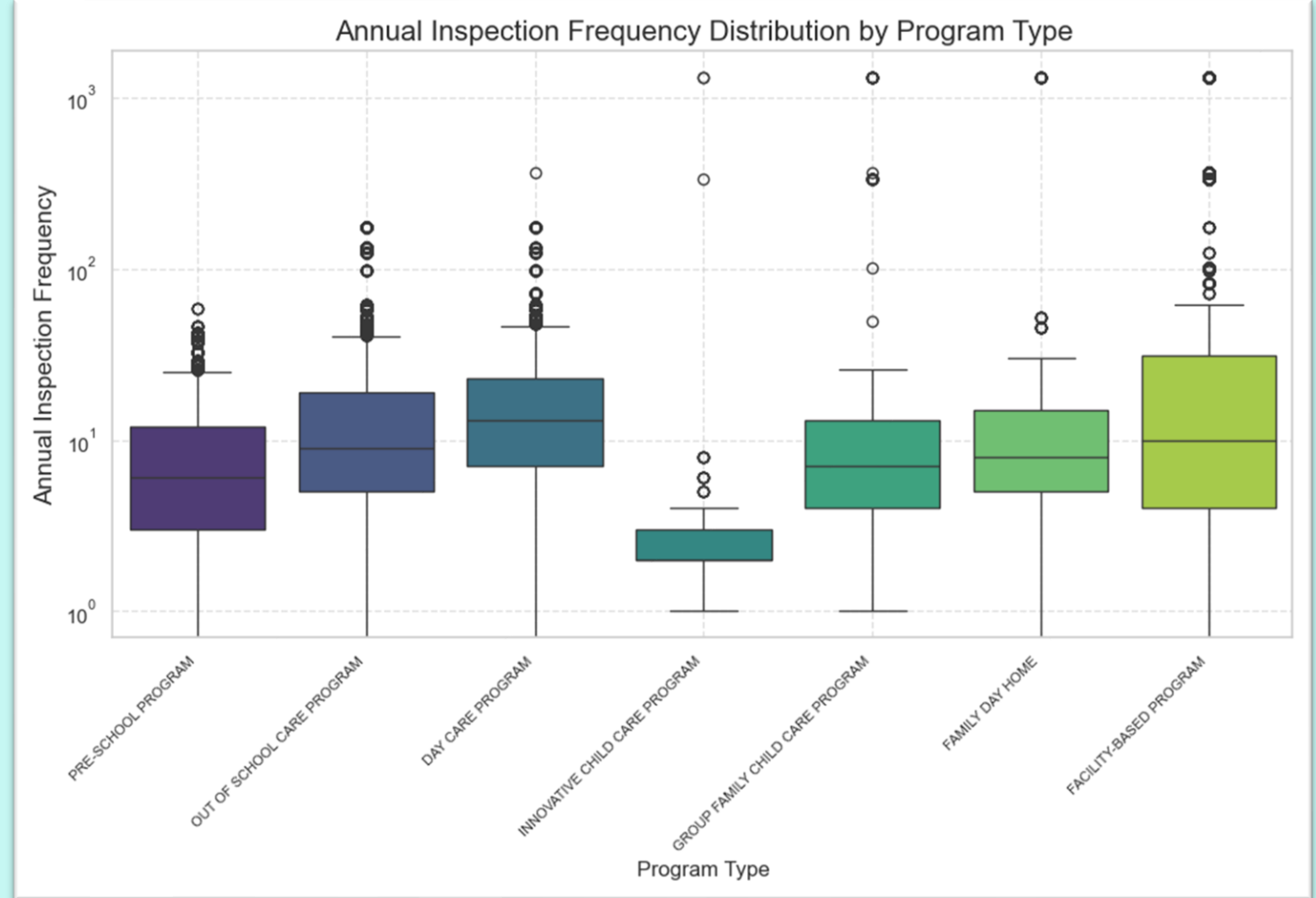


# Inspection Frequency (in Days) V/S Program Type

Is there a correlation between the type of childcare program and the frequency of inspections?

## Insights:

- Frequent Inspections: Innovative Child Care Programs experience more frequent and consistent inspections compared to other program types.
- Less Frequent Inspections: Facility-Based Programs have the least frequent and most variable inspection intervals, indicating less regulatory scrutiny.

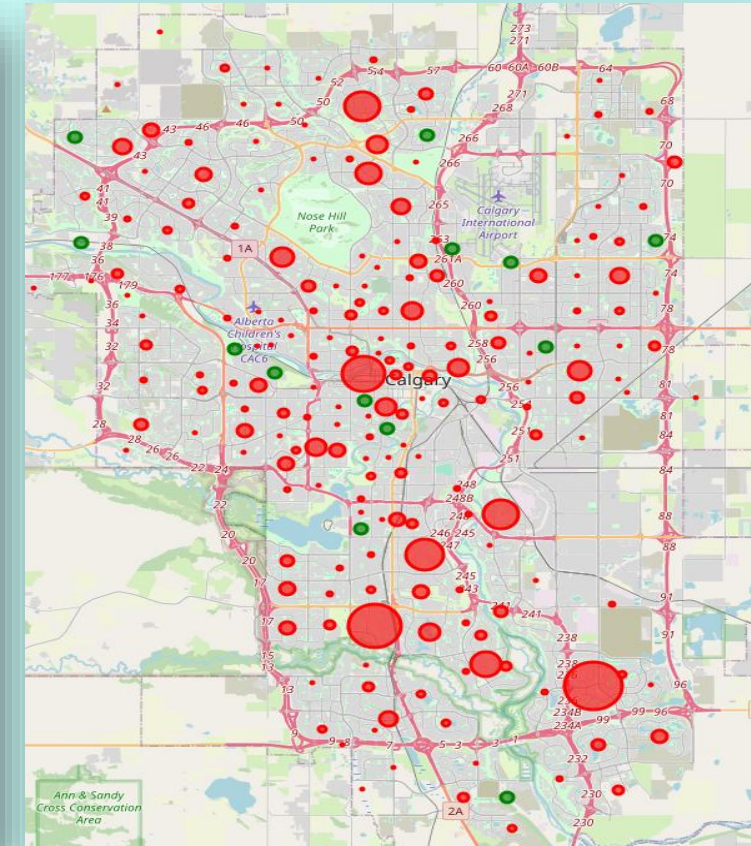
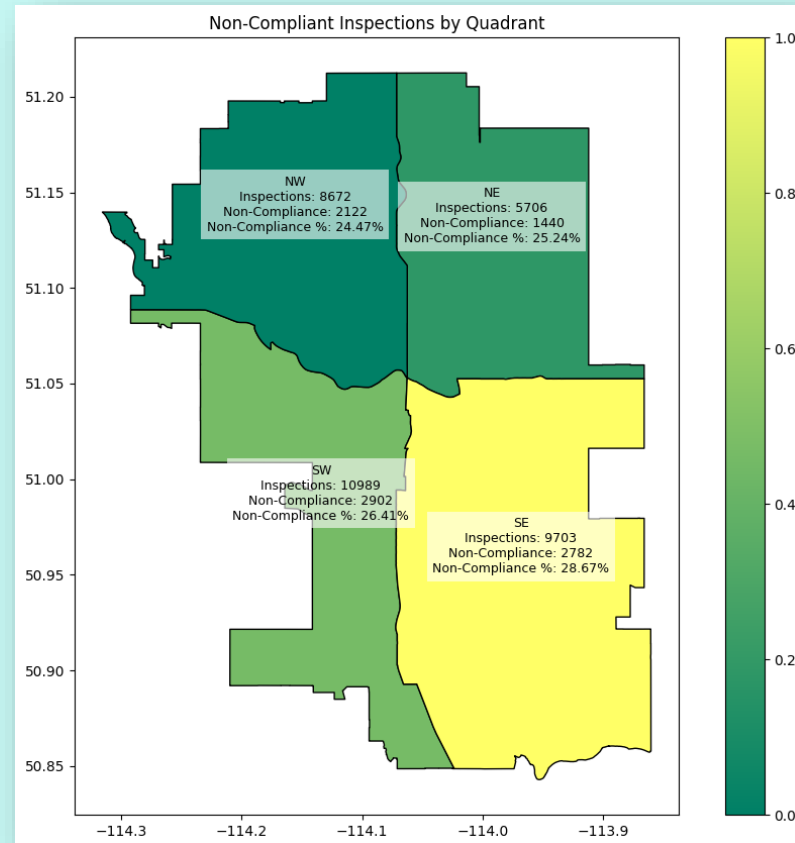


# Understanding the Non-Compliance% spread by Quadrant for better Choices

How is the Non-Compliance Spread across the four Quadrants and Communities of the city ?

## Insights:

- Highest Non-Compliance Rate: The SE (Southeast) quadrant has the highest non-compliance rate at 28.67%, indicating more frequent issues in this area.
- Frequent Inspections: The SW (Southwest) quadrant has the highest number of inspections (10,909), highlighting intensive monitoring.



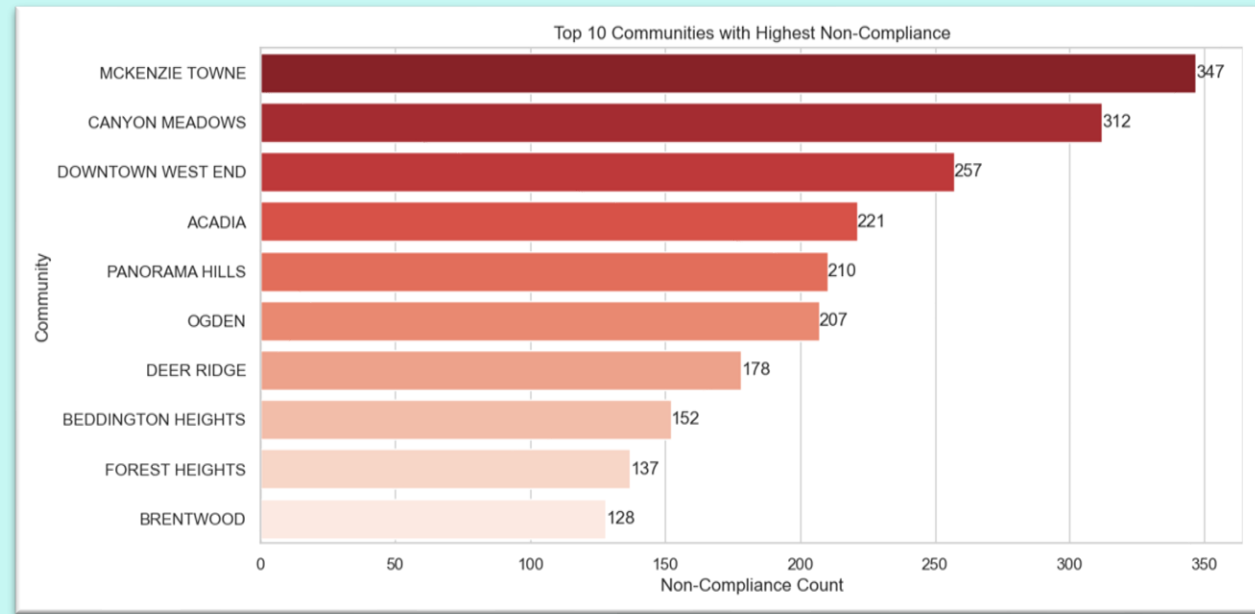
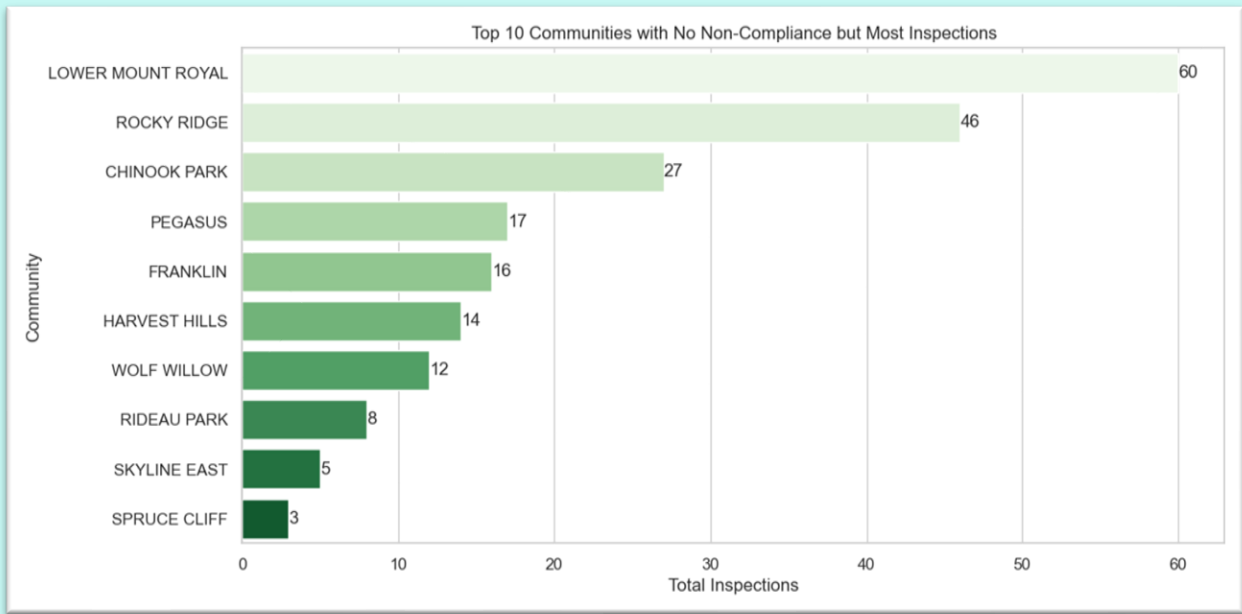
**Legend:**  
● Non-Compliant Community  
● Compliant Community





# Program Inspections and Non-Compliance against Communities

Identify the Top-10 communities with lowest and highest compliance rates

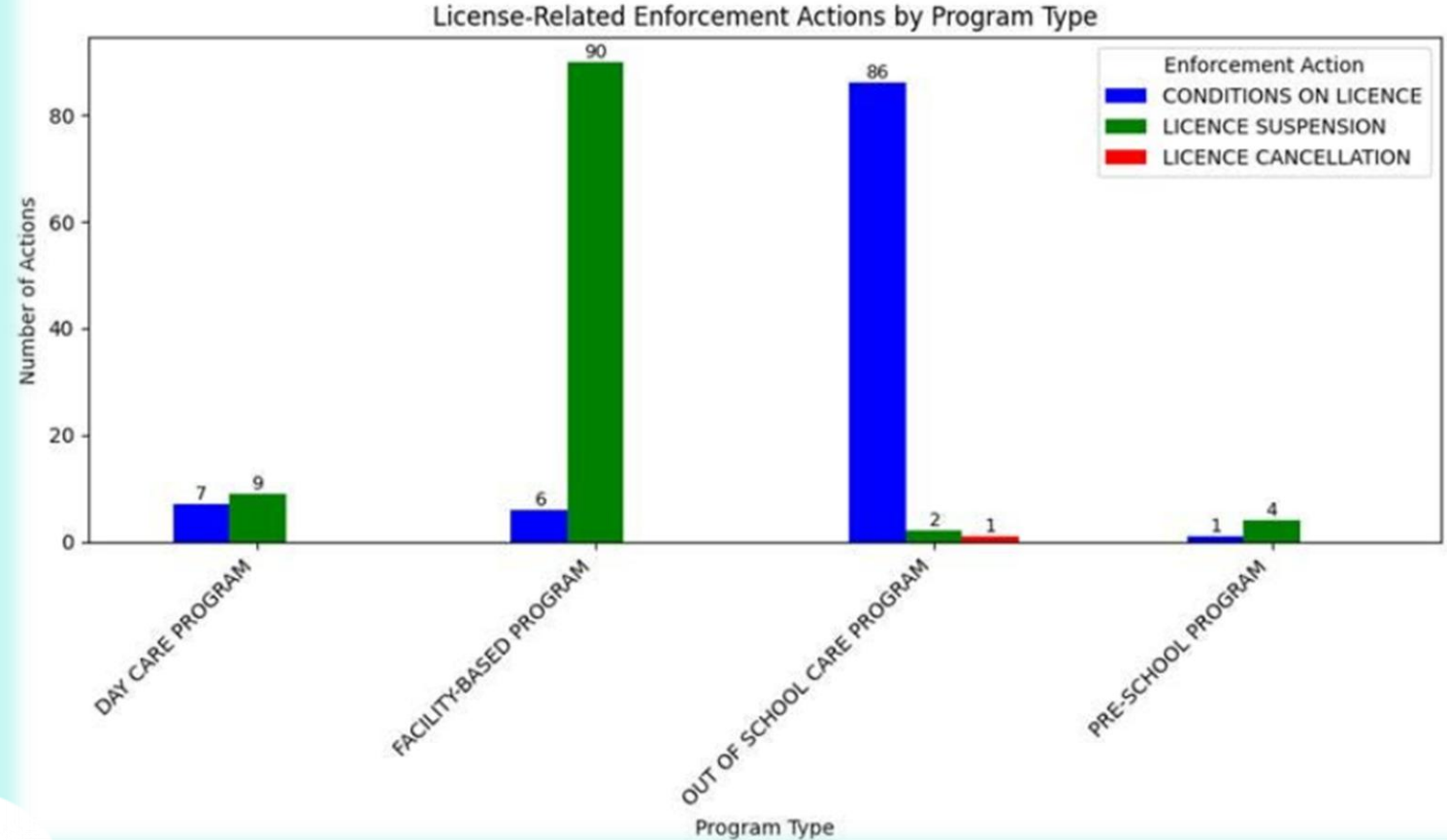


# Analyzing Regulatory Compliance –License Cancellation

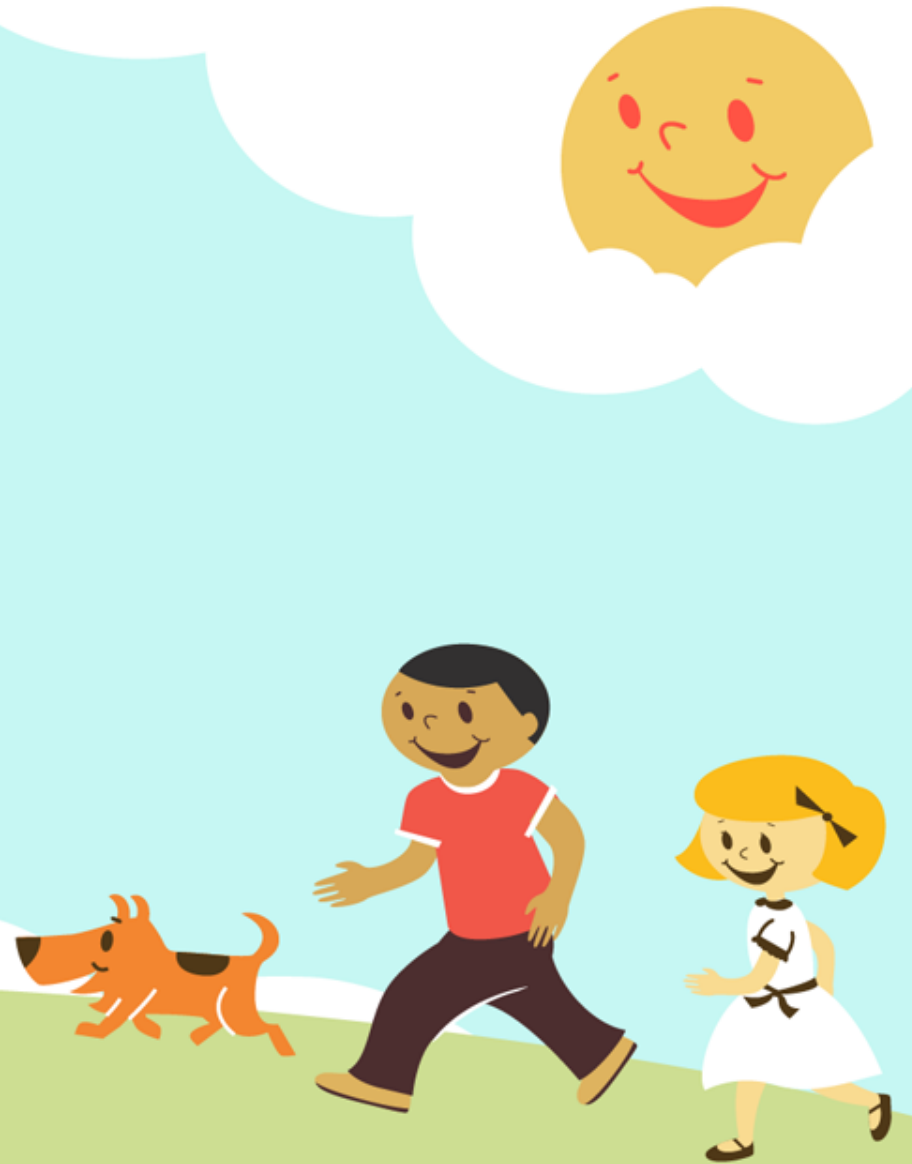
How does the certain program types or locations have more enforcement actions pertaining to licence cancellation and suspension ?

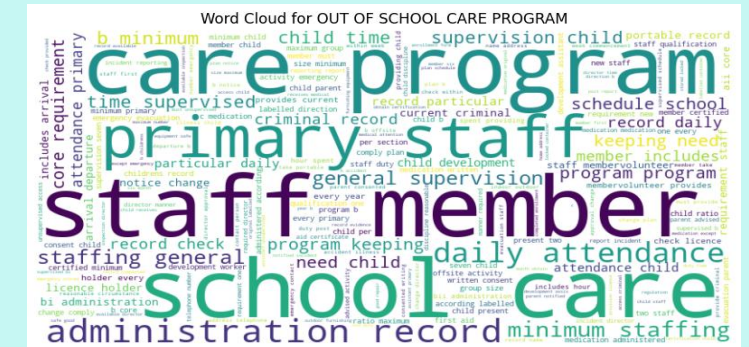
## Insights:

- Facility-Based Programs face significant enforcement actions, primarily suspensions.
- While Pre-School Programs have the fewest actions, indicating relatively better compliance.
- Out of School Care Programs mostly encounter conditions on their licenses.



# MACHINE LEARNING







# Sentimental Analysis on Non - Compliance

**\*\*Sentiment Scores Per Program Type:\*\***

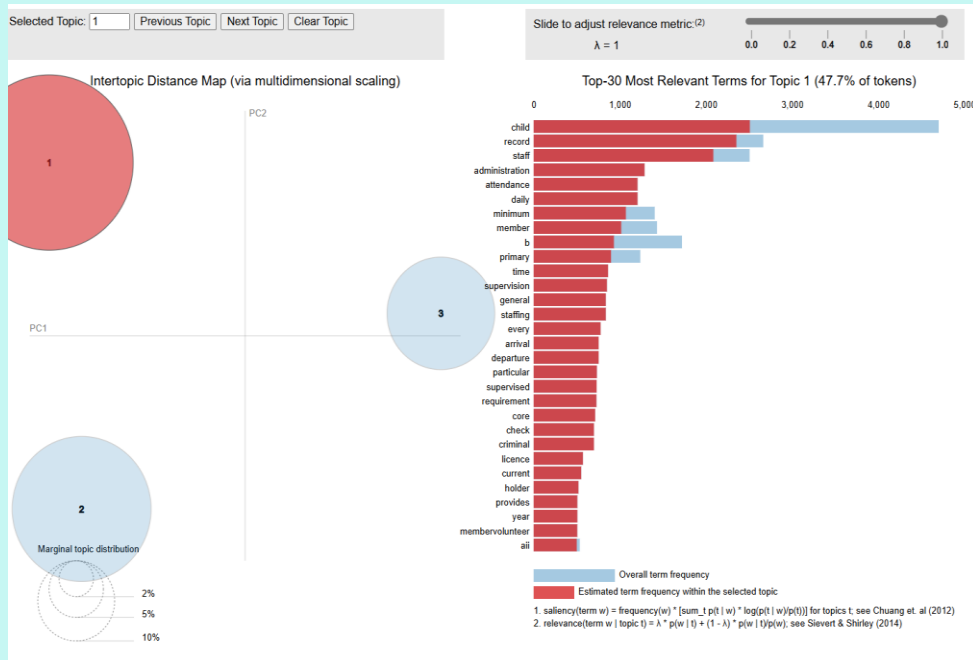
Type of Program	neg	neu	pos	compound	Sentiment Label
DAY CARE PROGRAM	0.032	0.840	0.128	0.113	Positive
FACILITY-BASED PROGRAM	0.040	0.906	0.054	0.001	Positive
FAMILY DAY HOME	0.062	0.887	0.051	-0.017	Negative
GROUP FAMILY CHILD CARE PROGRAM	0.051	0.869	0.079	0.069	Positive
OUT OF SCHOOL CARE PROGRAM	0.035	0.887	0.078	0.049	Positive
PRE-SCHOOL PROGRAM	0.055	0.928	0.017	-0.113	Negative

## Frequent words used in Programs

DAY CARE PROGRAM	: child, program, record, care, day, staff, b, administration, member
FACILITY-BASED PROGRAM	: child, record, b, staff, program, comply, medication, change, notice
FAMILY DAY HOME	: standard, record, child, b, information, home, supervision, childrens, criminal
GROUP FAMILY CHILD CARE PROGRAM	: child, record, care, program, b, group, family, administration, emergency
OUT OF SCHOOL CARE PROGRAM	: child, staff, record, program, care, b, member, minimum, school
PRE-SCHOOL PROGRAM	: record, child, staff, b, criminal, check, program, core, requirement



# LDA Topic Modeling and its insights on Non-Compliance Reasons



## Topic 1: Staffing & Supervision Compliance Issues

### Key Terms:

- child, record, staff, administration, attendance, supervision, requirement, criminal, licence, holder, check

### Non-compliance issues may involve:

- Insufficient supervision of children.
- Failure to maintain proper attendance records.
- Issues with staff qualifications, criminal background checks, and required licensing.

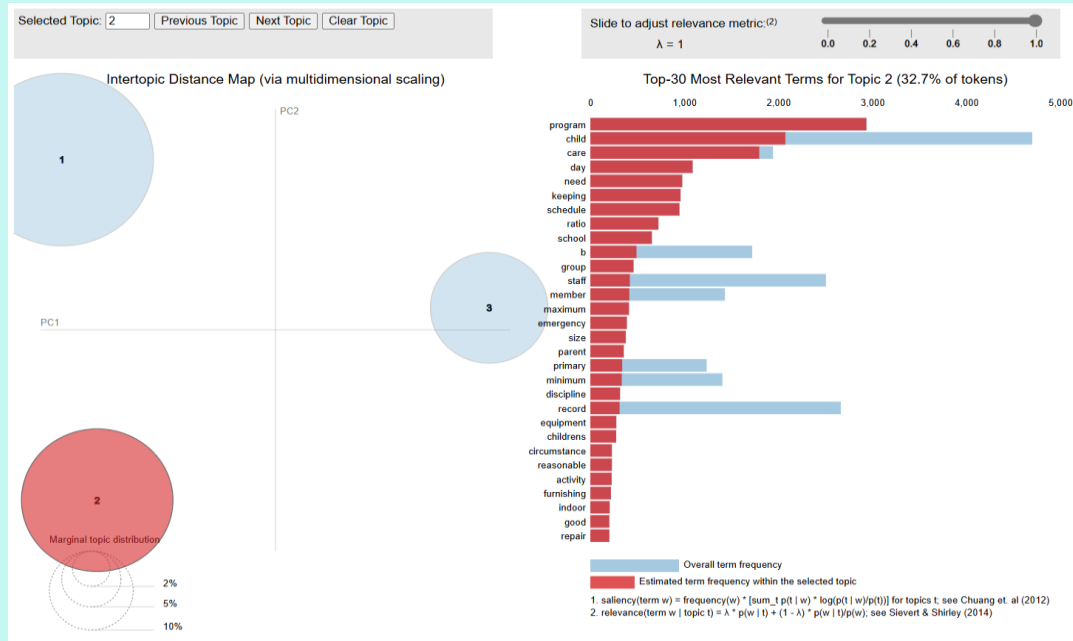
### Suggestion:

- Childcare program should focus on ensuring qualified, well-trained staff for childcare.





# LDA Insights



## Topic 2: Childcare Facility & Safety Regulations

### Key Terms:

- program, child, care, day, keeping, schedule, ratio, school, equipment, children, repair

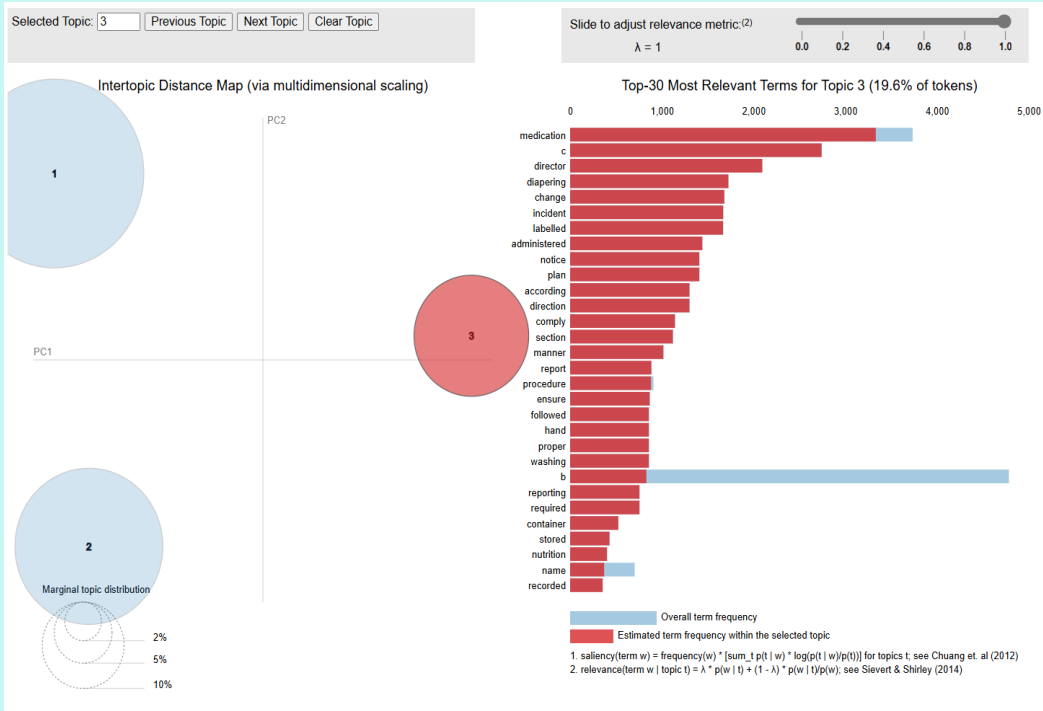
### Non-compliance issues may involve:

- Violation of child-to-staff ratios.
- Poor maintenance of facilities and equipment.
- Failure to follow proper schedules and care programs.

### Suggestion:

- Childcare should prioritize safety, well-maintained environments for child development.





## Topic 3: Medication & Hygiene Compliance

### Key Terms:

- medication, diapering, change, incident, administered, procedure, washing, reporting, nutrition, stored

### Non-compliance issues may involve:

- Improper storage or administration of medication.
- Lack of hygiene standards in diapering and handwashing.
- Failure to report incidents related to medication or health.

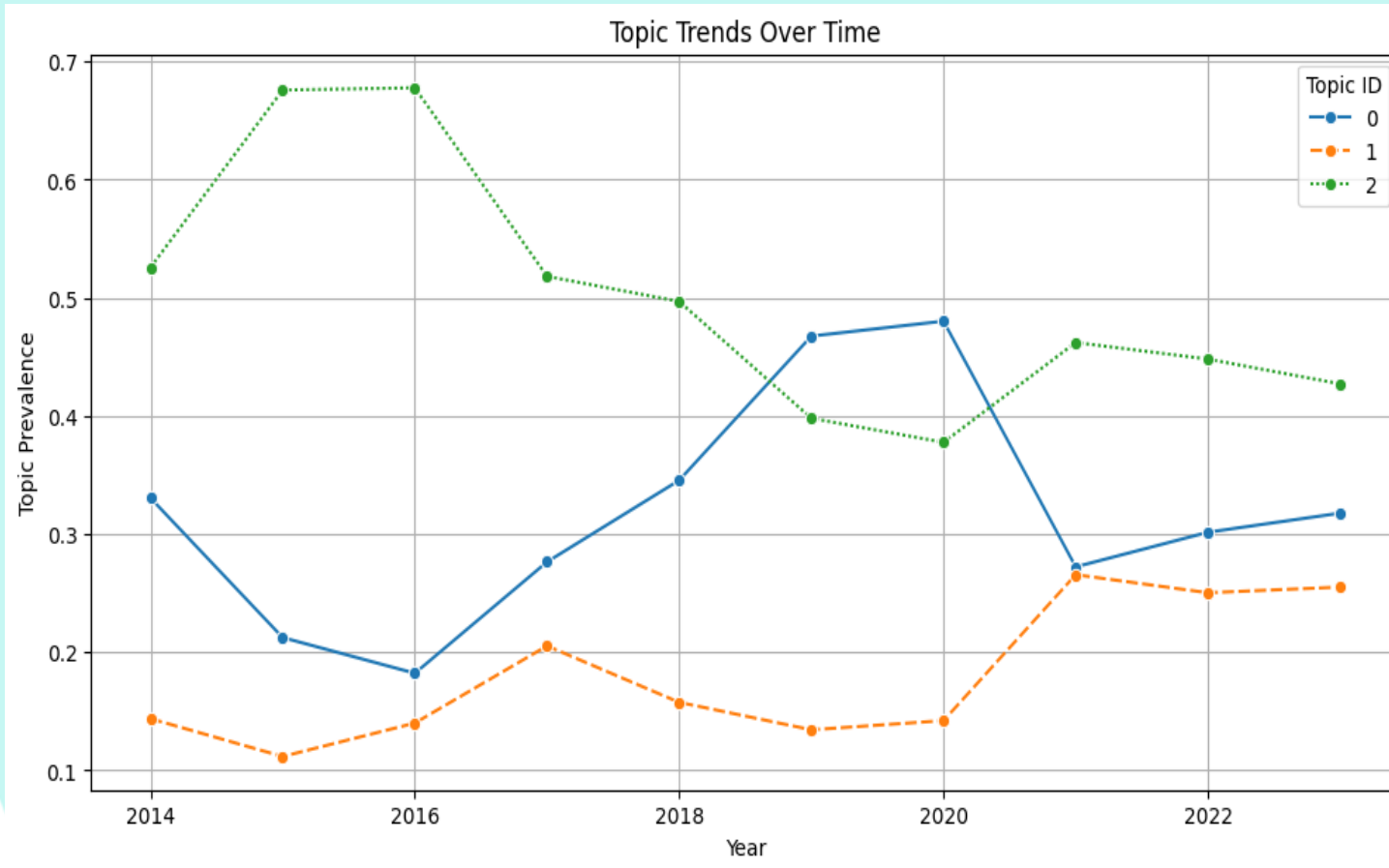
### Suggestion:

- This highlights the importance of strict health and safety protocols in childcare environments.



# LDA Over the Time

Are there trends or patterns in identified topics over time?



Topic 0: program, child, care, day, need

Topic 1: medication, c, director, diapering, change

Topic 2: child, record, staff, administration, attendance



# Predictive Analysis of Non-Compliance:

Can we predict non-compliance likelihood based on program attributes and inspection history, and what features indicate compliance status?

## Data Preparation

**Binary Target Column:** Created a binary target column (Non\_Compliance\_Indicator) to indicate non-compliance.)

### **Handling Missing Values:**

Dropped rows with missing values in selected columns to ensure data quality.

**Feature Selection:** Selected relevant columns for modeling.

## Categorical Encoding

**One-Hot Encoding:** Used one-hot encoding to convert categorical variables into a format suitable for machine learning algorithms.

**Data Splitting:** Split the data into training and testing sets.

## SMOTE

**SMOTE Application:** Applied SMOTE (Synthetic Minority Over-sampling Technique) to oversample the minority class..

## Logistic Regression Model

**Training the Model:** Trained a logistic regression model on the resampled training data.

## Code for Prediction

Code to Encode and Predict



# Model Results and Predictions

## Prediction Scenarios

Class Distribution in Resampled Training Data:

Non\_Compliance\_Indicator

1 24610

0 24610

Name: count, dtype: int64

Confusion Matrix (Resampled Data):

[[3967 2146]

[ 766 1581]]

Classification Report (Resampled Data):

	precision	recall	f1-score	support
0	0.84	0.65	0.73	6113
1	0.42	0.67	0.52	2347
accuracy			0.66	8460
macro avg	0.63	0.66	0.63	8460
weighted avg	0.72	0.66	0.67	8460

### Scenario 1: Group Family Child Program

- Capacity: 10
- Type of Program: Group Family Child Program
- Postal Code: T2M3T4
- Prediction: [0] Likelihood of Non-Compliance: Unlikely

### Scenario 2: Innovative Child Care Program

- Capacity: 6050
- Type of Program: Innovative Child Care Program
- Postal Code: T2M1L9
- Prediction: Predicted likelihood of non-compliance.
- Prediction: [1] Likelihood of Non-Compliance: Likely

```

✓ 0.0s
Prediction for Scenario 1 (Non-Compliance Likelihood): [0]
Prediction for Scenario 2 (Non-Compliance Likelihood): [1]

```



# Conclusion

- ❑ Key Non-Compliance Issues Identified.
- ❑ Non-Compliance Trends by Quadrant.
- ❑ Yearly Trends in Non-Compliance & Inspections.
- ❑ We identified key inspection reasons from LDA model
- ❑ Predictive Modeling on non-compliance likelihood.

# Future Steps

- ❑ **Refine Predictive Models** – Improve non-compliance forecasting accuracy.
- ❑ **Policy Recommendations** – Use trends to guide better regulations.
- ❑ **Community Engagement** – Provide recommendations to Conduct training programs for childcare providers





# References

- [1] Government of Alberta, "Childcare," Alberta.ca, 2025. [Online]. Available: <https://www.alberta.ca/child-care>.
- [2] City of Calgary, "Communities by Sector," Data Calgary, 2025. [Online]. Available: <https://data.calgary.ca/Base-Maps/Communities-by-Sector/e6xg-kaxf>.
- [3] City of Calgary, "City Quadrants," Data Calgary, 2025. [Online]. Available: <https://data.calgary.ca/Base-Maps/City-Quadrants/g8ma-syw>.
- [4] Statistics Canada, "Open Database of Addresses (ODA)," Government of Canada, 2025. [Online]. Available: <https://www.statcan.gc.ca/en/lode/databases/oda>.



# Thank You!

