# Data Science I – Proposal Document

## **Funding Request**

We are seeking funding to conduct an exploratory analysis of the Historical Restaurant Inspection Scores dataset from San Francisco, which contains vital information about restaurant health inspections conducted between **2016 and 2019**. Our objective is to analyze inspection scores, identify trends in restaurant compliance, and assess factors contributing to variations in scores.

This analysis will serve multiple purposes:

1. **Public Health Insight**: By understanding inspection trends and scores, we can provide insights that inform public health officials about compliance levels in the restaurant industry, ultimately contributing to better health outcomes for the community.
2. **Consumer Awareness**: Our findings will be valuable for consumers, enabling them to make informed choices about dining establishments based on their health inspection records.
3. **Policy Recommendations**: The results of this analysis will aid policymakers and restaurant owners in identifying areas for improvement and developing strategies to enhance compliance with health standards.

The dataset contains critical information, including inspection scores, violation types, and restaurant characteristics, which we will leverage to uncover meaningful insights. By utilizing advanced analytical techniques and visualizations, we aim to produce actionable findings that can drive positive change in restaurant health practices and community awareness.

We respectfully request funding to support our research, data analysis tools, and dissemination of our findings to stakeholders and the public.

## Team Members

Sarita Himthani

Zenil Patel

Jhanvi Munim

## **Dataset**

**Source**: <https://data.sfgov.org/Health-and-Social-Services/-Historical-Restaurant-Inspection-Scores-2016-2019/pyih-qa8i/about_data?row_index=0>

**Description**

**Rows**: 54k

**Columns**: business\_id, business\_name, business\_address, business\_city, business\_state, business\_postal\_code, business\_latitude, business\_longitude, business\_location, business\_phone\_number, inspection\_id, inspection\_date, inspection\_score, inspection\_type, violation\_id, violation\_description, risk\_category

**Numeric columns**: business\_latitude, business\_longitude, inspection\_score

**Categorical Columns**: business\_id, business\_name, business\_address, business\_city, business\_state, business\_postal\_code, business\_phone\_number, inspection\_id, inspection\_type, violation\_id, violation\_description, risk\_category

**Date column**: inspection\_date

**Location column**: business\_location

## Research Questions:

1. How do seasonal trends affect restaurant inspection scores?
2. Do inspection scores differ significantly across various neighbourhoods in the city?
3. How do inspection scores vary by restaurant types?
4. Is there a correlation between the number of violations and the inspection score?

## Variables:

**Dependent Variable:**

* **Inspection score:** This is the outcome variable that you should probably examine, as it reflects the quality of the restaurants based on inspection.

**Independent Variables:**

* **Business Location:** The neighbourhood or district where the restaurant is located could influence the inspection scores.
* **Inspection date:** The time of the year or specific date of the inspection might impact scores, possibly due to seasonal variations in food safety practices.
* **Previous inspection scores**: Historical scores from prior inspections may also serve as predictors of current performance.