**DAB501 Group 12 – Baltimore City – Project Proposal**

# Team Introduction

Group No. 12

Section No. 001

**Our Team**

Ikram Patel

Rituraj Borah

Vikrant Siwach

# Background/Motivation

Founded in 1700’s, Baltimore City is one of the major cities in Maryland with a population of 570,000 (estimated in 2021). The city’s council provides its employed citizens with various services. To continue providing these services, the council has been collecting and maintaining a large warehouse of data like name of employees, agency they are hired in, their designations, annual salary and so on.

This dataset is stored and kept available at their [government portal](https://data.baltimorecity.gov/) and can be extracted through ADP system. This data can be used to identify and analyse workforce related issues or to check the proportion of residents working at specific agencies. The data also reveals the employment cycle of city residents.

# Problem Statement

**The City Council and their Financial department** generally require insights on their city’s workforce, salary distribution, employment rate and agencies that are located in/near the city’s premises. Having a visual representation on the employment rate can help them in focusing the areas that need improvement and gaining better employee retention and recruitment.

Having an insight on salary distributions across different agencies and fiscal years can depict a trend in pay-grade distribution among these years. Such insights will be a valuable tool for the city council’s decision makers to better understand the workforce of city and focus on parts that will further establish a strong foundation in terms of employment. This will also help the government in establishing a trust among its residents.

# Project Proposal

An Analysis of this dataset will provide us with insights and trends on Baltimore city’s salary distribution across the different agencies and fiscal years. We will also be comparing hiring duration to the salary distributions to get an insight into variation of pay-grade range among these years. Comparing agencies with gross salaries of citizens will help council to know how salaries are distributed among these agencies.

We will be using R for data manipulation, analysis and visualization to provide our target audience with accurate and efficient insights of their Baltimore city’s workforce.

# Analysis Questions

1. How is the Gross Salary distributed across total Employees?
2. How many numbers of employees were hired during the period 1900 to 2024?
3. What number of employee’s work under each agency?
4. What is the annual salary of employees working in different agencies who were hired between the years 1900 and 2020?
5. What is the Gross Salary of employees between fiscal years 2011 and 2021?
6. What is the Annual Salary of employees hired between the years 1900 and 2024 ?
7. How the employees working under different agencies are segregated across the annual salary range between the financial year 2011 and 2021?
8. How the number of employees vary across the hiring date and Fiscal years.

# Dataset Description

The Dataset is acquired from [Baltimore city’s government portal](https://data.baltimorecity.gov/).

The data includes information on employment of est. 150,000 residents of Baltimore city (Maryland). The data talks about employee annual and gross salary from fiscal year 2011-2021 and the agencies (private and government) they got hired in. It also depicts their date of hiring and the department they are working in.