Project Report

# GitHub URL

(insert URL here)

# Abstract

(Short overview of the entire project and features)

# Introduction

(Explain why you chose this project use case)

# Dataset

(Provide a description of your dataset and source. Also justify why you chose this source)

# Implementation Process

(Describe your entire process in detail)

# Results

(Include the charts and describe them)

# Insights

(Point out at least 5 insights in bullet points)

Machine Learning:

* **Describe what kind of prediction you could perform in the future using machine learning and/or deep learning?**

A prediction that could be made using machine learning or deep learning on the data, or similar data would be as a predictor for whether staff would need pay raises in the coming future to coincide with inflation. If inflation and cost of living data was input via unsupervised learning, it could predict a down turn in the Arlington county economy that could signal to the county officials to put a hold on recruitment for a predicted amount of time (Sotayo, Slide 12, 2023).

Supervised learning would be best suited as it is labelled. As this data looks at demographic data most predictions would involve groups of people, whether that be by gender or race, using the regression method would work best in obtaining predictions (Data Camp, 2023).

* **Would you use classification or regression methods?**

In respect to the dataset I examined during this project. The appropriate statistical model I would apply is that of regression (Sotayo, A, Slide 10, 2023). This could be applied to predict the trajectory of the state employees pay scale. Whether it is likely to increase or decrease. The regression model would suit best as it assigns a continuous variable, which is likely to be the result of a model applied to this dataset (Data Camp, 2023). It would be interesting to also pull data from cost of living and inflation in the Arlington county to see if the pay scale of the employees was to coincide with rising cost of living.

It would also be interesting to run a regression model to see if the pay differences between genders and particular races is likely to increase or decrease in the future. It would be interesting to pull demographic and population data from a census of Arlington county. This could allow for a deeper insight to inequalities that are present for county employee.

# References

(Include any references if required)