

# Kaggle wars 25



Guransh Chugh

# Algorithm used and Accuracy

XGBoost

Target Variable

Data\_Value = Mortality Rate

Ideally Mortality Rate should be low

Linear Regression

```
model = xb.XGBRegressor(n_estimators=100, learning_rate=0.1, random_state=42)
model.fit(X_train, y_train, eval_set=[(X_val, y_val)], early_stopping_rounds=10, verbose=True)
y_pred = model.predict(X_test)

rmse = mean_squared_error(y_test, y_pred, squared=False)
r2 = r2_score(y_test, y_pred)

print(f"Root Mean Squared Error (RMSE): {rmse}")
print(f"R2: {r2}")
```

Root Mean Squared Error (RMSE): 0.01498641514600849  
R<sup>2</sup>: 0.9997763903871116

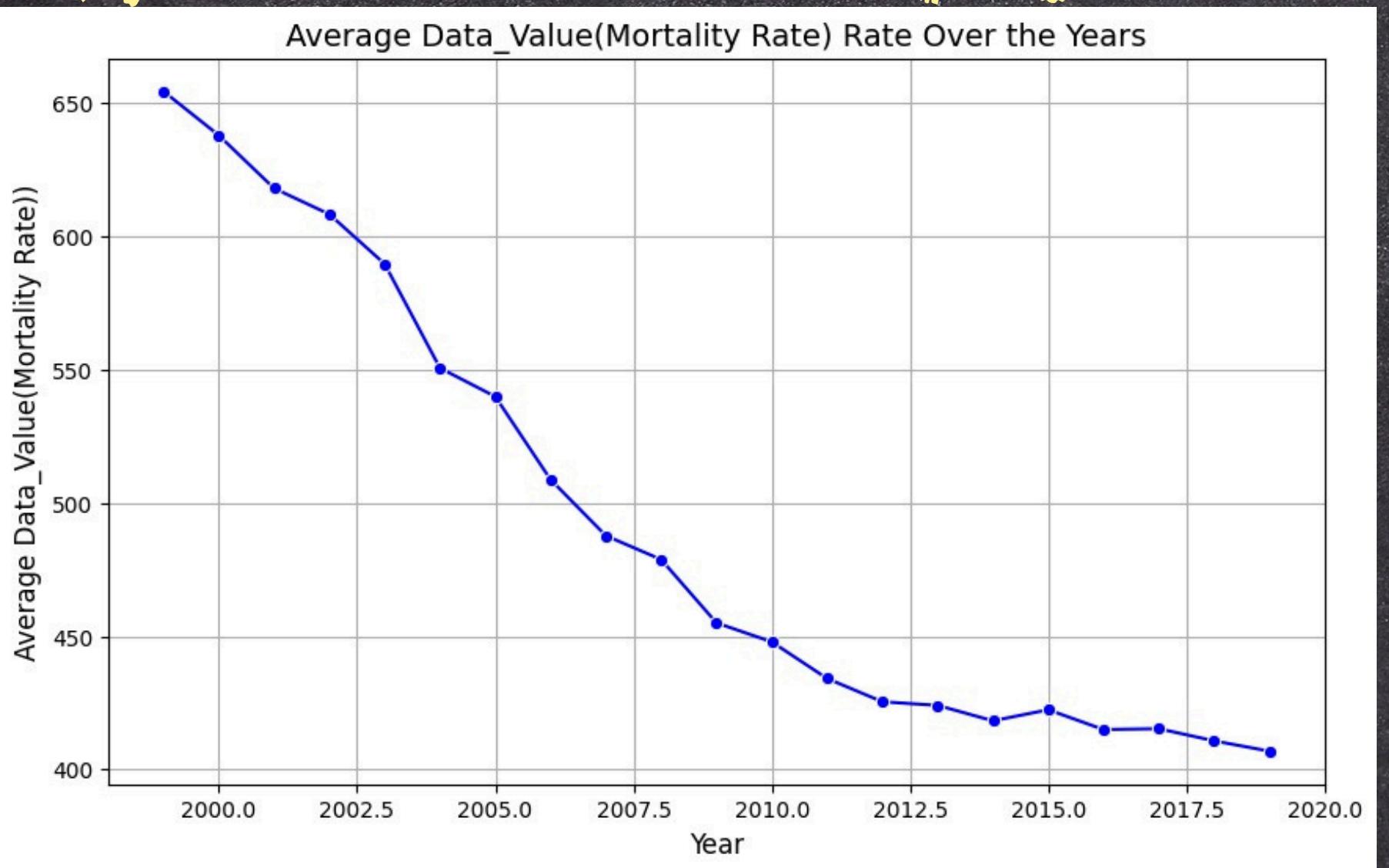
```
lr = LinearRegression()
lr.fit(X_train,y_train)
y_pred = lr.predict(X_test)
rmse = mean_squared_error(y_test, y_pred, squared=False)
r2 = r2_score(y_test, y_pred)

print(f"Root Mean Squared Error (RMSE): {rmse}")
print(f"R2: {r2}")
```

Root Mean Squared Error (RMSE): 0.014832329052385896  
R<sup>2</sup>: 0.9997809649304483

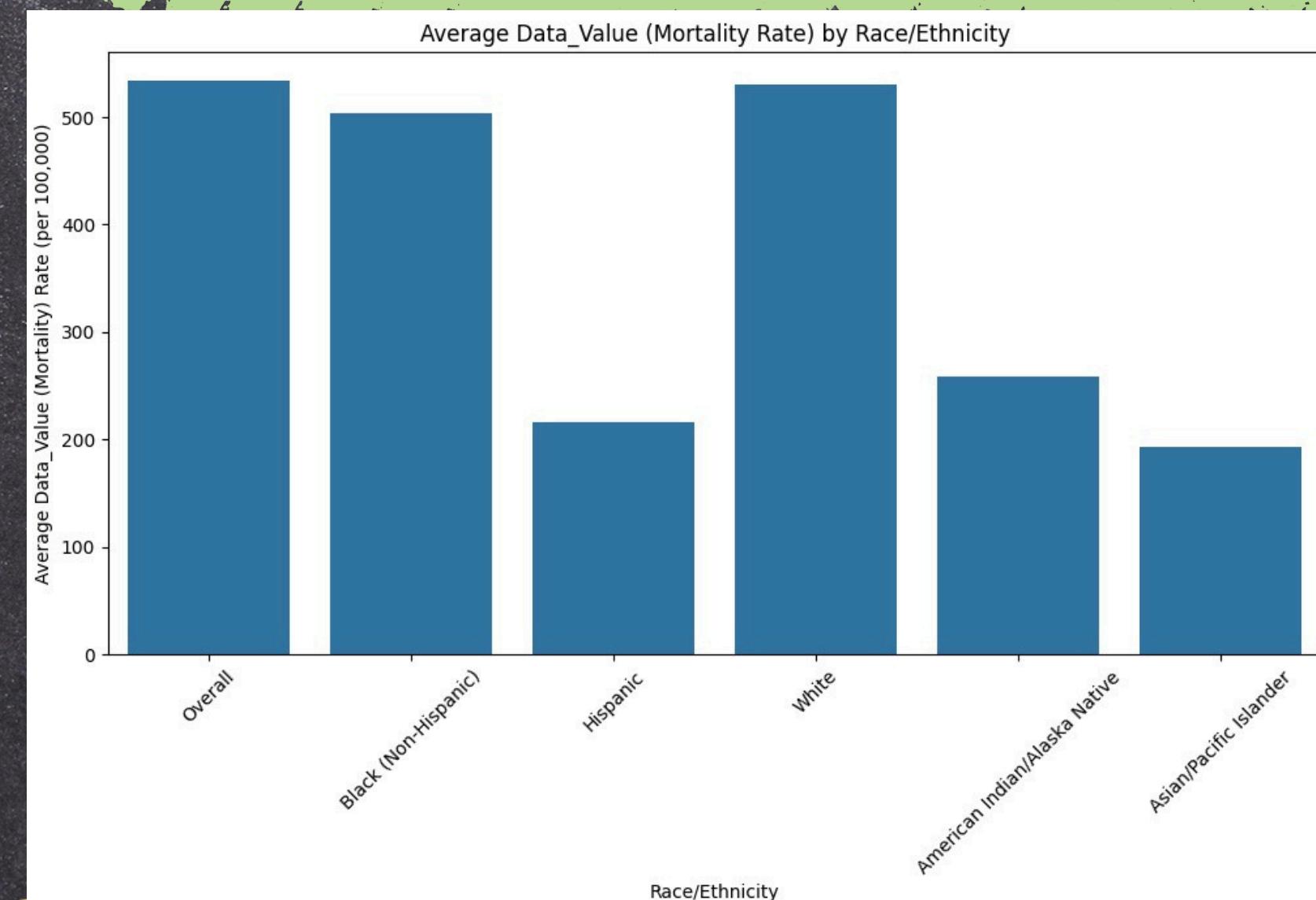
# DATA VISUALIZATION AND INSIGHTS

## Average Data\_Value over years



The drop in average mortality rates over the years shows that public health efforts are working. However, ongoing efforts are needed to keep up and speed up this progress.

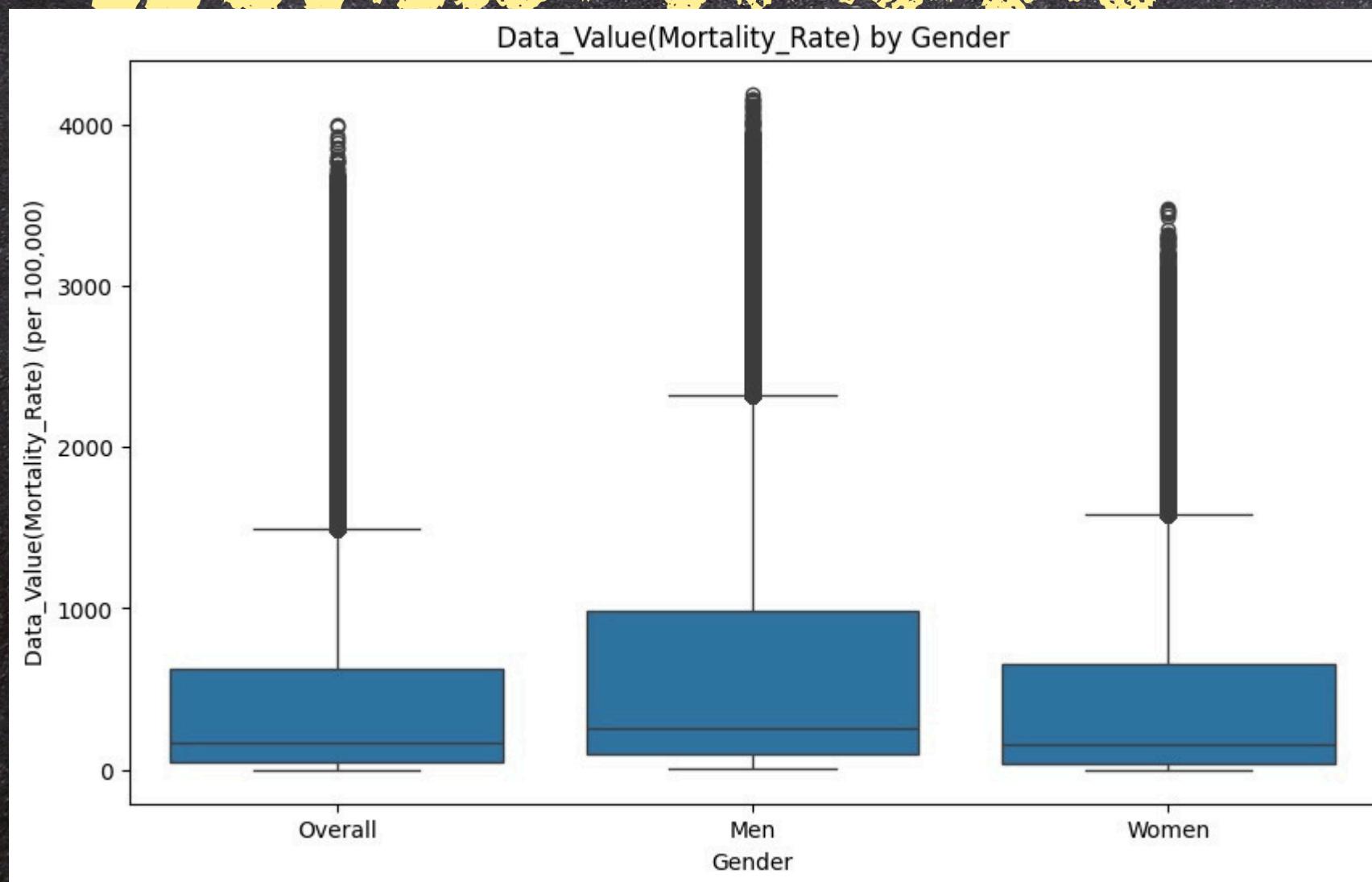
## Average Data\_Value by Ethnicity



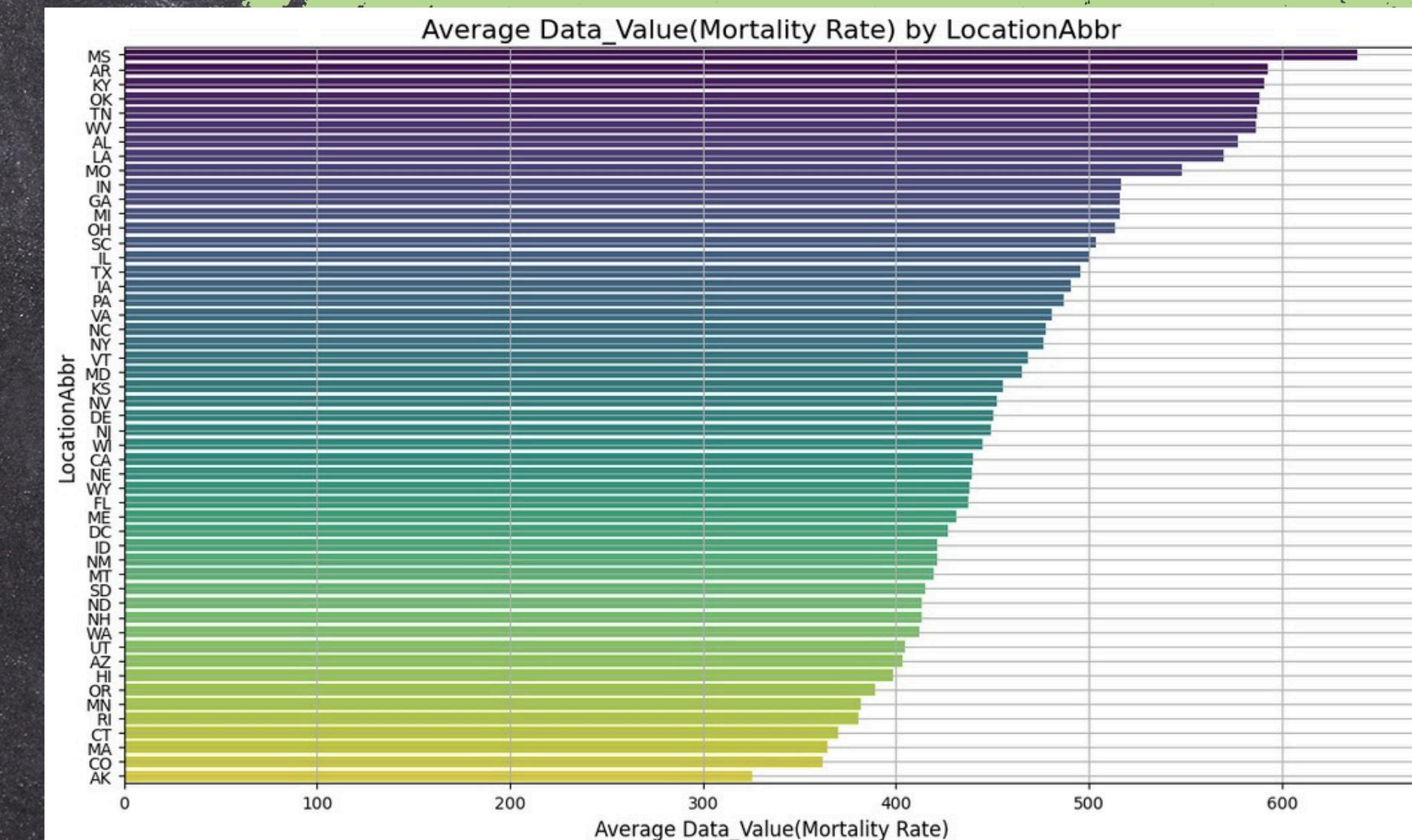
The differences in mortality rates between ethnic groups highlight the need for focused health programs and policies to improve health outcomes, especially for communities with higher rates.

# DATA VISUALIZATION AND INSIGHTS

## Mortality Rate by Gender



## Mortality Rate by LocationAbbr

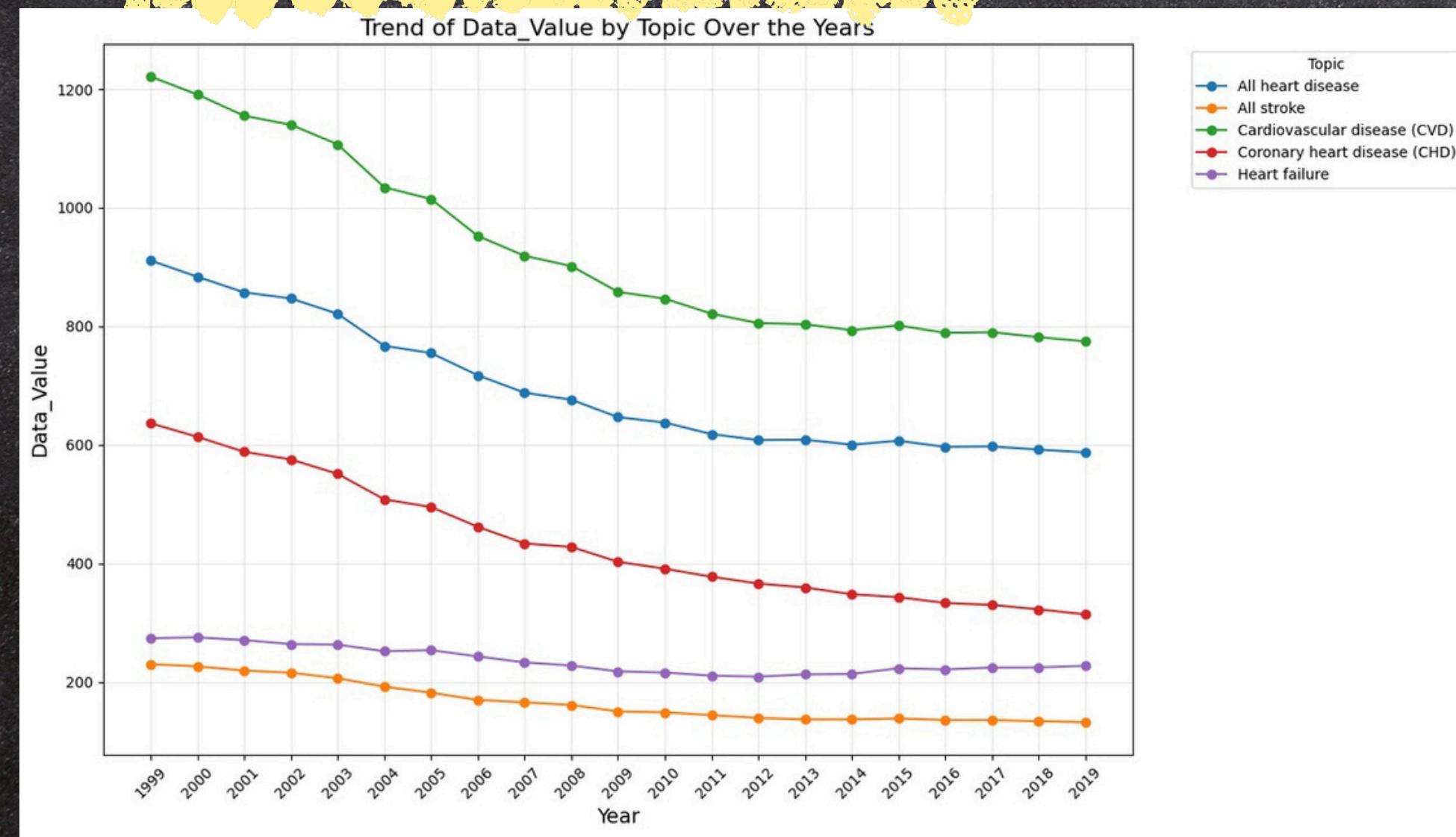


Mortality rates show a higher prevalence in males compared to females. This gender disparity suggests potential biological, lifestyle, or healthcare access factors influencing heart disease outcomes across different genders.

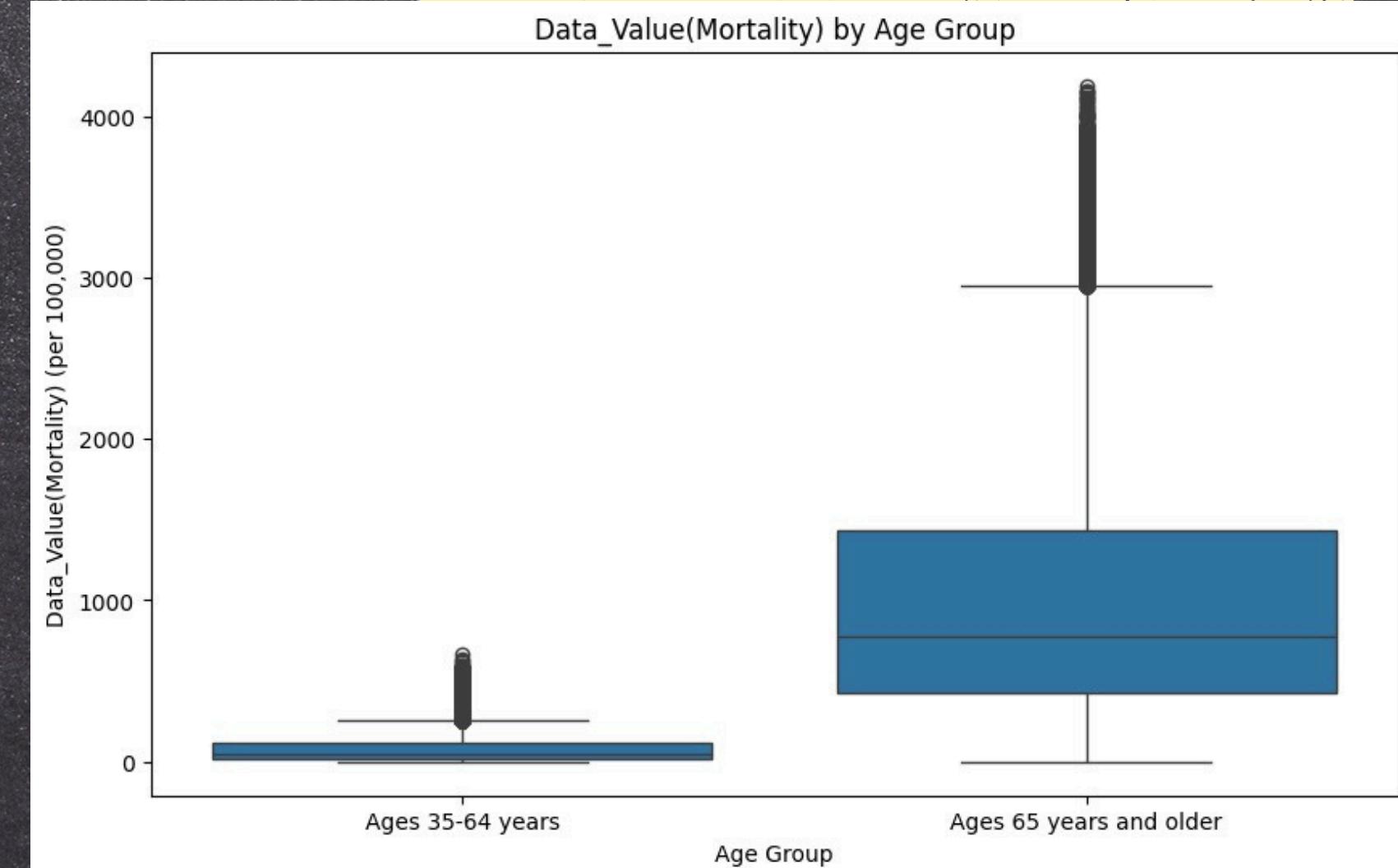
Identify the city with the highest mortality rate to prioritize healthcare resource allocation, targeted intervention programs, health education campaigns, improved medical facilities, and preventive care measures for effective mortality reduction.

# DATA VISUALIZATION AND INSIGHTS

## Data\_Value Rate trend



## Data\_Value vs Age Group



Identifying topics with increasing trends over the years can help guide research focus, ensuring that resources are directed toward emerging public health concerns that require urgent attention and intervention.

Since the data indicates that individuals above 65 years are most affected, targeted health campaigns, preventive care programs, and specific resource allocations could significantly reduce mortality risks in this age group.



**THANK YOU**