Report for Individual Coursework 5DATA004W Data Science Project Lifecycle

Name: Marukh Khan
Student Id: w1971684
Dataset: Hate Crime 2017-2025 (from https://catalog.data.gov/dataset/hate-crimes-2024)

Link to Streamlit app: https://hate-crimes-dashboard.streamlit.app/
Link to GitHub repository: https://github.com/mk123x/Hate-Crimes-Dashboard-DSPL

Aims and Objectives

The dataset I chose is Hate Crime 2017-2025 from data.gov. It contains records of hate crimes in Austin, Texas in 2017-2025, including incident types, locations, and victim categories.

The project aim is to create an interactive dashboard with key insights to help policymakers visualise and explore trends in hate crimes, identify groups affected and understand geographical distribution. By visualising these trends over time, we can better understand shifts in societal attitudes and the geographical areas most affected by hate crimes. This could help law enforcement agencies in targeting resources and also building public awareness.

The objectives of this project:

- Create insightful visualisations
- Summarise key insights such as the most targeted groups and common bias trends
- Allow user data exploration through dashboard interactivity
- Support decision making for policy makers and law enforcement agencies
- Maintain high standards of performance and usability
- Document and test the system effectively to ensure reliability
- Organise and present the project deliverables on GitHub

Development Methodology

For this coursework I will use an iterative development method. This will allow me to break the work down into smaller tasks which is ideal for creating an interactive dashboard, as it requires building and testing components.

I will begin by understanding the dataset and creating basic visualisations, which will help me discover initial insights and decide which features are most important to include.

After this I will use Streamlit to enable user interaction including a sidebar filter and charts which will allow users to explore the data with different dimensions, such as year, bias or victim ethnicity.

As I build each feature, I will test it to ensure usability and if necessary, adjust visualisations or interactivity to improve user experience. After completing the dashboard and confirming its performance I will finalise the GitHub repository and finish the report.

The need for testing and feedback is crucial for the success of my dashboard and this project. The interactive nature of the dashboard means that it is important to test all components regularly to ensure they work as expected. Testing and getting feedback will help identify and resolve any performance issues early.

Requirements

Functional requirements:

- Display key summary statistics
- Users can select a specific year and view hate crime data
- Flexible filtering options
- Bar chart for bias motivation frequency

Non-functional requirements:

- The dashboard should be responsive and user-friendly
- Data should load within 5 seconds
- The design must be clear and consistent
- Ensure the app can work smoothly across various browsers
- The dashboard should maintain performance without lag
- Provide clear error messages for users when something goes wrong (for example there is no data available for the selected filter).

Test cases

Test cases are used to verify a particular feature to make sure my application works correctly. They ensure the app is working correctly and handles any errors.

TC1	Title: Year filter test
Description	To verify that the dashboard correctly shows the data when the 'All' filter is selected in the Select Year filter in the sidebar
Steps and input data	Open the dashboard Set the year filter in the sidebar to 'All'. Read the summary statistics and graphs
Dependencies	1. Streamlit to build the web app 2. Pandas to load, read and manipulate the dataset 3. Matplotlib to plot the line charts and bar charts that visualise trends
Expected result	The dashboard should show data from all years in the visualisations

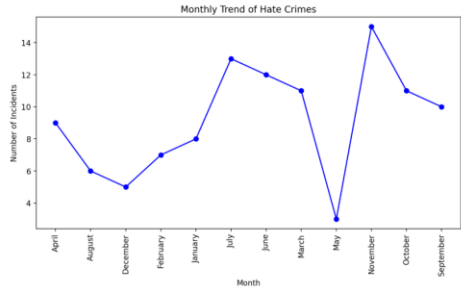| TC2	**Title: Visualising hate crime incidents by zip codes where the location is Park/Playground**
Description	This test case checks whether the dashboard correctly filters and displays hate crime incidents by zip code whether the 'Offence Location' is set to 'Park/Playground'. This makes sure that the location filter is working as expected and ensures the right bar chart has the correct data
Steps and input data	1. Open the Streamlit dashboard 2. Set 'Offence Location' to 'Park/Playground' 3. Set the other filters to 'All' 4. Read the zip code bar chart
Dependencies	Streamlit to build the web app Pandas to load, read and manipulate the dataset Matplotlib to plot the line charts and bar charts that visualise trends Filters in the sidebar
Expected result	The dashboard should update and show the bar chart for zip code where incidents occured at 'Park/Playground' locations There should be no error when selecting this option

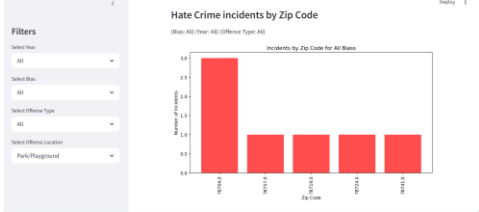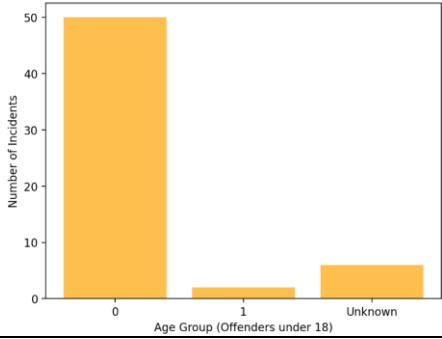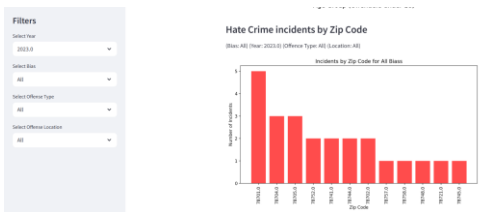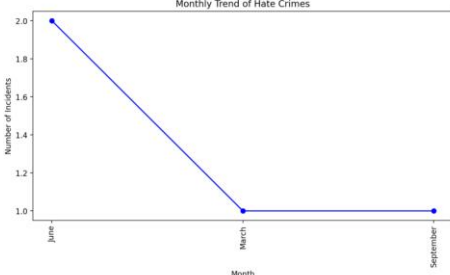| TC3	**Title: Visualising hate crimes where the bias is Anti-Black**
Description	The test case checks that the dashboard correctly filters and displays hate crime incidents where the 'Bias' filter is set to 'Anti-Black'. This will ensure that the bias filter works as expected
Steps and input data	1. Open the Streamlit dashboard 2. Set the 'Bias' to 'Anti-Black' 3. Read visualisations and verify the correct data is displayed
Dependencies	Streamlit to build the web app Pandas to load, read and manipulate the dataset Matplotlib to plot the line charts and bar charts that visualise trends Filters in the sidebar
Expected result	The dashboard should filter and only show visualisations for hate crime incidents with 'Anti-Black' bias

TC4	**Title: Visualising hate crime incidents by year in 2023**
Description	This test case will check whether the dashboard correctly filters and displays hate crime data for 2023. This will ensure the year filter is working properly and that the visualisations update correctly from the select year
Steps and input data	1. Open the Streamlit dashboard 2. Set the 'Year' filter in the sidebar to 2023 3. Set the other filters to 'All' 4. Read visualisations and verify all data shown is from 2023
Dependencies	Streamlit to build the web app Pandas to load, read and manipulate the dataset Matplotlib to plot the line charts and bar charts that visualise trends

	Year filter in the sidebar
Expected result	The expected result is that the dashboard should show hate crime data only for 2023 in all visualisations. No errors should occur when selecting 2023 in the year filter

TC5	**Title: Combining multiple filters**
Description	In this test case we are testing the dashboard to handle and correctly display results when multiple filters are applied at once. This test case is testing incidents from year 2020, bias is Anti-Black, and the offence type is Criminal Mischief, regardless of location
Steps and input data	1. Open the Streamlit dashboard 2. Set the 'Year' filter in the sidebar to '2020' 3. Set the 'Bias' filter in the sidebar to 'Anti-Black' 4. Set the 'Offence Type' filter in the sidebar to 'Criminal Mischief' 5. Set the 'Location' filter to 'All' 6. Read visualisations and verify all data shown is correct to the filters
Dependencies	Streamlit to build the web app with functioning sidebar filters Pandas to load, read and manipulate the dataset Matplotlib to plot the line charts and bar charts that visualise trends
Expected result	The dashboard should display only the hate crime incidents that match the selected filters (2020.0, Anti-Black, Criminal Mischief). No error or irrelevant entries should show.

Test Log

TC	Date	Executed by	Actual result	Pass/ Fail	Notes
TC1	01/05 /25	Marukh Khan	Dashboard displayed data across all years as expected 	Pass	The sidebar filter works without any issue
TC2	01/05 /25	Marukh Khan	The zip code bar chart didn't update to reflect the 'Park/Playground' locations	Fail	The location filter may not be applied correctly, I need to check filter names in the code and verify

					the values match dataset Update: There was a mistake typo in my code and I have fixed and retested and is now working
TC3	01/05/25	Marukh Khan	Visualisations filtered to only Anti-Black bias cases **Offender Age Group Distribution** (Bias: Anti-Black or African American) (Year: All) (Offence Type: All) (Location: All) 	Pass	Data correctly filtered by bias
TC4	01/05/25	Marukh Khan	Dashboard showed data just for 2023 	Pass	The year filter returns correct results
TC5	02/05/25	Marukh Khan	Dashboard displayed data only for Bias: Anti-Black or African American, Year: 2020, Offence Type: Criminal Mischief, and the Location of Offence: All Showing data where Bias: Anti-Black or African American, Year: 2020.0, Offence Type: Criminal Mischief, and the Location of Offence: All **Number of Incidents Over Time** (Bias: Anti-Black or African American) (Year: 2020.0) (Offence Type: Criminal Mischief) (Location: All) 	Pass	Combination of multiple filters work correctly