**Data Science Project Lifecycle – Individual Coursework**

Bora Yavuz – W19489519

Aims

Using the information provided in the Great British Driving Licence dataset, I am aiming to create an interactive dashboard that will assist motor insurance companies, and individual drivers, in understanding which areas in the UK should be considered high risk for insurance companies. For this purpose, the dashboard must be visually appealing, easy to understand, and accessible to anyone regardless of technical ability, as well as being thoroughly tested to ensure that errors do not ruin the user experience in the final version.

Objectives

* Create an interactive dashboard which clearly presents UK Driving Licence data in a way that is easy to analyse
* Ensure that the visualisations are of high quality, include interactive elements, and are easy to use through a well-made test plan
* Provide evidence of the progression of the Streamlit dashboard, documenting all previous versions of the dashboard in a GitHub Repository
* Follow an appropriate development methodology that will guide me to completing and achieving the aim of this project
* Organise and clean the dataset so it can be used effectively to demonstrate trends and patterns with multiple data visualisations

Development Methodology

During this data science project, I will be following an iterative and incremental software development methodology, allowing me to receive feedback on my dashboard early, ensuring that any issues are fixed before moving on to another visualisation. For example, when making a visualisation related to the amount of penalty points drivers have, I will plan the visual, design it with high quality elements, develop it with interactive features, and finally test it to compare its functionality to the defined requirements and aims of the project. If my visualisation differs from the aims of the project, the early feedback will give me the opportunity to immediately change the visual to realign with the objectives, meaning that I will stay on track and won’t have to make major edits after completing the dashboard.

Requirements

|  |  |  |  |
| --- | --- | --- | --- |
| **Functional or Non-Functional** | **Requirement Level** | **Title** | **Requirement** |
| Functional | User | Interactivity | Users shall be able to interact with the dashboard to extract their desired data using filters |
| Non-Functional | System | Performance | The application must apply filters and respond to user requests, updating visualisations within 5 seconds |
| Functional | User | Data Analysis | Users will be able to analyse UK driving licence data based on variables such as age, gender, and location of drivers |
| Non-Functional | System | Security | Application must use unbiased data to ensure risk assessments completed by insurance companies are accurate |
| Non-Functional | System | Usability | Application shall display consistent visual quality regardless of which browser is used to access Streamlit |
| Non-Functional | User | Usability | Users shall be able to use all the interactive features of the application even with minimal computer experience |
| Functional | User | Geographic Visualisation | Users shall view regional distributions of driving licence holders on an interactive UK map |
| Functional | System | Page Layout | Application shall correctly show all the visualisations in an easy-to-read format |
| Non-Functional | System | Accessibility | Application must meet the Web Content Accessibility Guidelines to ensure accessibility for all users |
| Functional | User | Downloading Visuals | Users shall be able to download visualisations in common formats such as PNG for offline analysis or reporting |

Test Cases

|  |  |
| --- | --- |
| TC1 Title: Interactivity | Test Case Title: Test that the user can change age group and other filters, and the dashboard will remain the same once filters are returned to default |
| Description | Users shall be able to interact with the dashboard to extract their desired data using filters |
| Steps and input data | 1. Open dashboard, ensure loads default version 2. Remove data by deselecting data groups with filters 3. Re-apply all filters to default state to ensure it returns to original state without issues |
| Dependencies | * Dashboard must load correctly * Filter options must be visible and available * Data must exist for each option, so each filter displays data |
| Expected Result | Dashboard shows user the data they are searching for, providing them with insights and keeping the dashboard consistent once user returns it to its default version, so they can find other insights |

|  |  |
| --- | --- |
| TC2 Title: Data Analysis | Test Case Title: Test that the user can analyse UK driving licence data by selecting variables such as age, gender, and location |
| Description | Users will be able to analyse UK driving licence data based on variables such as age, gender, and location of drivers |
| Steps and input data | 1. Open the dashboard 2. Ensure users can see Gender data 3. Ensure users can see Age data 4. Ensure users can see Location data 5. Ensure users can filter through specific groups in these variables such as “Male” for Gender, “18-20” for Age, and “London” for Location |
| Dependencies | * Dashboard must load correctly * Filter options must be visible and available * Data must exist for each variable |
| Expected Result | Dashboard shows user the data they are searching for, giving them the opportunity to see all available data on variables from the dataset, such as Age, Gender, and Location |

|  |  |
| --- | --- |
| TC3 Title: Geographic Visualisation | Test Case Title: Test that the user can view Driving Licence data regionally on a clear UK map with county boundaries |
| Description | Users shall view regional distributions of driving licence holders on an interactive UK map |
| Steps and input data | 1. Open dashboard 2. Ensure County boundaries appear correctly 3. Ensure Locations can be highlighted and emphasised 4. Hover over County for more insights on each region |
| Dependencies | * Dashboard must load correctly * UK map must be interactive and accurate once opened * Data must exist for each region |
| Expected Result | Dashboard shows user regional distributions and related data with chosen variables on an interactive map of the UK that clearly displays where each region starts and ends |

|  |  |
| --- | --- |
| TC4 Title: Page Layout | Test Case Title: Test that the dashboard is displayed with visualisations in logical order and any variables are explained for ease of use |
| Description | Application shall correctly show all the visualisations in an easy-to-read format |
| Steps and input data | 1. Open dashboard 2. Ensure initial finding visualisations appear first 3. More in-depth visualisations should appear after with explanations where necessary |
| Dependencies | * Dashboard must load correctly * Headers and captions for visualisations must also load at the same time * Visualisations must be displayed and work even with filters applied |
| Expected Result | Dashboard displays a collection of visualisations, described and ordered to provide users with data and key insights that will be understood no matter the experience level of the user |

|  |  |
| --- | --- |
| TC5 Title: Downloading Visuals | Test Case Title: Test that the dashboard allows visualisations to be downloaded in PNG format, including visualisations that have had data filtered out |
| Description | Users shall be able to download visualisations in common formats such as PNG for offline analysis or reporting |
| Steps and input data | 1. Open dashboard 2. Apply filters to update a visualisation 3. Save image as PNG to computer 4. Open image, and ensure it is saved as PNG 5. Import image into a word document (common for offline reports) |
| Dependencies | * Dashboard must load correctly * Visualisations must be displayed and work even with filters applied * User must have the ability to click on visualisations |
| Expected Result | Opening the saved image shows a PNG image of the chosen updated visual, displaying all features of the visual exactly as it is displayed in the dashboard |

Test Log

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **TC** | **Date** | **Executed by** | **Actual Result** | **Pass/Fail** | **Notes** |
| 1 | 29/04/2025 | Bora Yavuz | Age Groups filtered out and returned to same position in bar charts. Licence Type filter works as intended | Pass | Licence Type bar chart removing one type and adding it back puts it at the end of the three types, however with 3 distinct colours the data can still be compared |
| 2 | 29/04/2025 | Bora Yavuz | Licence Type and Penalty Points by Age Group data displayed clearly with insights visible in the charts | Pass | Gender and Location data not made for the first two visuals, but will appear later in the dashboard |
| 4 | 29/04/2025 | Bora Yavuz | Weighted Total explained and both visualisations split by a header describing the chart below | Pass | I will add a caption below the Licence Type chart, as the Y axis ranges from 0 to 1.75 with 1e6, which means 500,000 for 0.5 |
| 5 | 29/04/2025 | Bora Yavuz | Both visualisations correctly saved as PNG images with Age 40-44 removed and tested import into a Word document | Pass |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |