February - April 2025

Rawad - w19885025

Bhavjot - w1972756

Mourad - w1960946

Kelvin - w18984891

Mohammad - w1996549

Data Science Project Lifecycle - Group Coursework

Enhancement Report

Understanding Communities In Need

**Introduction to the Project** (Majority: *Rawad,* Support: *Bhavjot*)

The global sneaker industry is a major contributor to environmental degradation, Responsible for 1.4% of global greenhouse gas emissions and almost as much as the aviation sector's 2.5%. Every year, 20 billion pairs of shoes are produced, yet 300 million end up in landfills, contributing significantly to waste and pollution. Coupled with this is the social issue of Fuller accessibility, particularly for individuals facing poverty and homelessness.

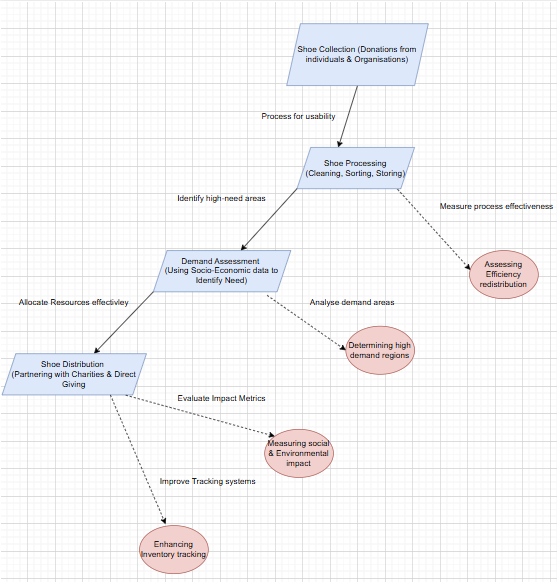
Resole, founded in 2018 by Musa Nsubuga, addresses these dual challenges through its innovative “Reduce, Reuse, Recycle” model. Operating from its Brixton Village Hub, Resole has redistributed over 25,000 pairs of trainers worth £2 million to vulnerable groups, including the homeless, refugees, and low-income families. The organisation partners with charities, shelters, and soup kitchens, bridging environmental sustainability and social welfare.

**Problem Domain**

Resole faces several operational challenges including:

* Assessing the efficiency of its redistribution process.
* Determining which regions have the highest demand for footwear.
* Enhancing inventory tracking and management systems.
* Measuring the social and environmental impact of its initiatives.

These challenges highlight the need for systematic analysis and data-driven solutions to enhance Resole’s operations and maximise its social impact.



The flowchart above outlines the current processes, problems, and main objectives that Resole undertakes. The fundamental aim is to distribute shoes effectively and to those who need them. This process encounters problems mentioned previously. Our job is to gain insights from data to explain and coordinate how to counter these issues, thus making the process easier and more efficient.

**Objectives**

This project seeks to:

* Analyse socio-economic, demographic, and resource accessibility data to identify high-need areas across the UK.
* Model deprivation levels to predict future demand and optimise distribution strategies.
* Develop a dashboard to visualise key findings, enabling data-driven decision-making.

By addressing these objectives, the project aims to provide actionable insights that strengthen Resole’s capacity to support communities, reduce environmental waste, and promote sustainability within the footwear industry.

***Project Planning and Risk Assessment*** (Majority: *Bhavjot*, Support: *Kel, Mourad, Mo*)

***Project Deliverables***

The key deliverables for this project include:

1. **Data Analysis Report**

* A structured report containing insights on shoe demand trends, storage challenges, and priority groups.
* Visualisations such as Graphs/Charts/Tables to illustrate key findings.

1. **Data Cleaning and Processing Framework**

* A documented Approach for handling missing/incomplete data.
* Steps for ensuring data quality before analysis.

1. **Recommendations for Resole**

* Practical suggestions to improve efficiency in shoe distribution and storage.
* Potential collaborations with organisations to enhance their operations.

1. **Final Presentation**

* A summary of findings and recommendations presented to stakeholders/lecturer.

***Team Roles & Responsibilities***

For Team Roles & Responsibilities, our team followed a highly collaborative approach, all members actively contributed across the project’s phases, including **planning, research, data analysis, modeling, visualisation, and report writing**.

* ***Project Coordinator*** - Ensured effective communication, and timeline management to keep the project on track.
* ***Research Contributor*** - Conducted background research, gathered insights, and structured the problem domain for a data-driven approach.
* ***Data Analyst*** - Assisted in data collection, cleaning, and analysis, as well as ensuring the quality and accuracy of the data for further processing.
* ***Modelling & Insights Developer*** - Contributed to modelling techniques and provided insights based on the data findings.
* ***Report & Documentation Contributor*** - Participated in writing and structuring the report, ensuring clarity, and alignment with project objectives.

**Key Responsibilities:**

* Communication & Coordination
* Completing assigned work while supporting others.
* Deadline Management - Meeting project milestones on time
* Problem–Solving - Identifying challenges and refining solutions
* Feedbacks & Reviews

***Work Breakdown Structure (WBS)***

1. Project Planning & Initial Research

* Defining project scope and objectives.
* Identify key stakeholders (Resole, Charities, Homeless peoples, Soup kitchens)
* Challenges faced by Resole

1. Data Collection & Preparation

* Gathering data sets for analysis
* Conduct Interview with Resole’s Representatives.
* Clean and Preprocess collected data, handling missing values.

1. Data analysis & Modelling

* Analyse trends affecting footwear
* Models/Graphs to predict future shoe demand
* Assess the efficiency of Resole’s current redistribution process

1. Dashboard development & Recommendations

* Design and develop a dashboard for visualisations
* Provide solutions for enhancing inventory tracking & management.

1. Final Report & Presentation

* Document containing summary of findings and analysis
* Visualisations to support insights.
* Prepare and submit final report & presentation.

***Risk & Mitigation Plan***

| Risk | Likelihood | Impact | Mitigation Strategy |
| --- | --- | --- | --- |
| Incomplete Data | High | High - Incorrect Insights leading to poor decision making. | Requesting additional datasets from Resole, use external sources, use assumptions if necessary. |
| Time Constraints | High | High - Project delays and even failing the project | Prioritise critical tasks, set milestones |
| Technical Challenges | Medium | High - Delays in project completion | Use simple models and review with lecturers if needed |
| Lack of Group Engagement | Medium | High - Reduced productivity and uneven workload distribution | Schedule regular check ups, set clear individual responsibilities. |
| Conflicts in Decisions | Low | Medium - Disagreements can slow the progress | Establish clear decisions, encourage discussions. |

***Interview Insights: Understanding Resole’s Needs***

To gain a deeper understanding of Resole’s operations and challenges, we conducted an interview with a Resole representative, who provided valuable insights into the difficulties of footwear redistribution, storage limitations, and the need for a more data-driven approach to optimize distribution.

One of the challenges I faced was organising the interview responses in a clear and structured way without losing any important details. Some answers were quite broad, so it took some effort to interpret and condense the key points while keeping everything accurate. I also had to ensure there wasn’t too much repetition and that the information flowed logically.

| **Questions asked** | **Answers** |
| --- | --- |
| What specific key deliverables would be most valuable to Resole's objectives? | **Cleaning**, it needs to be more effective, **financially** and in 5 to 10 years people should donate a lot more shoes. |
| Are there any seasonal or temporal trends in demand for footwear (e.g., higher demand during winter months or back-to-school periods) that we should account for in our analysis? | * In the summertime we get more parents. * People like to save them as memories * We send more shoes to childrens. * Festive seasons, mostly after christmas * From december to february |
| Does Resole collaborate with any other organisations or government bodies that could provide additional data or insights for this analysis? | They are sponsored by Crep Protect so they have free cleaning supplies |
| How does Resole currently identify and prioritise individuals in need of footwear, and are there specific groups (e.g., children, homeless individuals, low-income families) that receive priority in distribution? | * The key service users for Resole are primarily homeless individuals, as they are the most common group in need of shoes. * The second priority is young people, helping keep them away from crime by providing shoes that promote education and personal well-being. * Refugees are the third priority group, as they often lack access to basic necessities. |
| Have you encountered any significant challenges when it comes to predicting shoe demand | Challenges are in storage and manpower (team capacity) and software issues.  Sending email and phone calls, which is tedious, (it's a long process, it is tiring)  The location also must be feasible for them too. |
| How should we handle missing or incomplete data? | By keeping the critical information |
| Are there specific socio-economic indicators we should focus on? | * Homeless * Young people * Refugees |
| What are the most important insights we need to provide in our analysis? | Anything that helps Resole to expand and provide shoes in priority order of:   * Homeless * young people * Refugees |
| On the easy read word document, why is there an asylum seeker category for employment if ‘asylum seekers’ also are included in the unable to work category? | How it was input by the user in the survey, |
| Will we be able to receive data from the locations that act as the distributors such as soup kitchens, charities? | Forms are given to them and they fill it out. Resole may be able to provide this data from their partnerships |

**Additional Insights from Resole Interview:**

***Key Service Users and Prioritization***

* People have jobs but are still unable to buy shoes, which is a common issue addressed by Resole.

***Operational Challenges***

* Storage and Cleaning:
  + About 30% of shoes are cleaned every month, and 70% is redistributed.
  + Cleaning process: Shoes are cleaned using a brush, steam cleaner, and a solution. The process varies depending on the condition of the shoes.
  + Shoes are checked for condition, and if they are not usable, they are recycled through incineration.
  + Resole has two hubs for managing operations.
  + Bigger brands charge premium prices for shoes, while Resole works to take shoes away from landfills.
  + The cleaning process can take 3-4 hours depending on the shoe condition. Resole also uses a solution for cleaning, with a minimum price of £20 for the cleaning service.
* Inventory and Distribution:
  + Shoes are stored manually in a storage space.
  + Resole works with food banks and soup kitchens to help identify and prioritize those in need of shoes. They use forms filled out by these partners to gather information on the types and sizes of shoes needed.
  + The organization also sends out appeals for children’s shoes to increase supply.
  + Shoes are distributed based on size, with a focus on size 7 to 11 for adults.
  + Resole’s ideal outcome is a more financially, environmentally, and time-efficient system that involves working with other communities to make a difference and increase awareness.
  + Resole has a circular system for intake and outtake, but they don’t have enough shoes to meet the demand.

***Cleaning Process and Manual Effort***

* Cleaning and Storage:
  + The cleaning process involves brushes, steam cleaning, and applying a solution based on shoe condition.
  + Shoes that are not in usable condition are recycled by incinerating them.
  + Manual labor is involved in both cleaning and storage processes.
  + It can take 3-4 hours (sometimes 2-3 hours) to clean a batch of shoes.

***Cost and Financial Considerations***

* Resole’s goal is to distribute shoes cost-effectively to those in need.
* Packaging for customers paying for cleaning services is important, as it adds a sense of value to the product.
* The minimum cost for cleaning is set at £20.
* Resole has a limited number of people working on cleaning, with a maximum of 4 people involved in the process.

***Outreach and Partnerships***

* Resole works with a range of community organizations, including food banks and soup kitchens, to identify and prioritize individuals in need of shoes.
* They gather information about shoe sizes through forms filled out by these partners.
* Resole regularly works with partners in Manchester, Birmingham, and other cities, though they only have data for London.
* Resole collaborates with sneaker events and receives shoes when people move to new places, which helps increase the quantity and quality of the shoes available for redistribution.
* Shoe donations also come from people who attend sneaker events.
* Resole considers those on universal credit for shoe donations, ensuring that even individuals who are receiving social benefits are not excluded from support.

***Seasonal Trends and Demand***

* There is also a spike in demand during festive seasons.
* During the summer months, Resole receives more requests, particularly from parents for children’s shoes.

***Data and Partner Collaboration***

* Resole uses data from partners like food banks and soup kitchens to determine the sizes of shoes needed.
* Resole has the potential to receive data from their locations, as the food banks and charities use surveys to track their stock and can share this information.

***Data And Methods*** (Majority: *All*)

This section relates to how data was collected , initially explored and how the quality of the data was evaluated.

It also explains the approaches taken in order to model the data , and goes into detail on the processing techniques used.

**Initial Data Collection** (Majority: *Kelvin*, *Mo*)

Collecting data proved to be more challenging than expected. Many sources provided full reports rather than actual datasets, making it difficult to access raw numbers for analysis. Even when datasets were available, they often lacked key details, covered limited time periods, or were locked behind paywalls. Some sources had clunky APIs, vague instructions, or inconsistent formats, requiring extra effort to clean and standardize the data. Technical issues like slow servers, connectivity problems, and mismatched file formats further complicated the process.

To overcome these challenges, we took a multi-pronged approach. We prioritized official government databases and open-data platforms, ensuring the sources were reliable and regularly updated. When dealing with incomplete or outdated datasets, we supplemented them with additional sources, allowing us to fill gaps and track trends over time. For example, even though some datasets only covered previous years, we retained them for comparative analysis with newer data. Where APIs were difficult to navigate, we researched documentation thoroughly and leveraged community forums for support.

The initial datasets we decided to use are below, while allowing us to add more datasets as required:

| **Data Set Name** | **Data Set Source** |
| --- | --- |
| 01pehdemographydata.xlsx | [People experiencing homelessness, England and Wales: Census 2021 – Demography tables - Office for National Statistics](https://www.ons.gov.uk/peoplepopulationandcommunity/housing/datasets/peopleexperiencinghomelessnessdemographytables) |
| 02peheilrdata.xlsx | [People experiencing homelessness, England and Wales: Census 2021 – Ethnic group, national identity, language and religion tables - Office for National Statistics](https://www.ons.gov.uk/peoplepopulationandcommunity/housing/datasets/peopleexperiencinghomelessnessethnicgroupnationalidentitylanguageandreligiontables) |
| 03pehhealthdata.xlsx | [People experiencing homelessness, England and Wales: Census 2021 – Health and disability tables - Office for National Statistics](https://www.ons.gov.uk/peoplepopulationandcommunity/housing/datasets/peopleexperiencinghomelessnesshealthanddisabilitytables) |
| Publication\_tables.xlsx | [Rough Sleeping Data Framework, September 2024 - GOV.UK](https://www.gov.uk/government/publications/rough-sleeping-data-framework-september-2024) |
| Cambridge\_Homeless\_Point-in-Time\_Count\_data\_\_2012-2024\_20250216 | <https://data.cambridgema.gov/General-Government/Cambridge-Homeless-Point-in-Time-Count-data-2012-2/ify2-i22z/about_data> |
| Resole Data | Provided by Resole (located in Blackboard) |

***Data Cleaning and Preprocessing*** (Majority: *Mourad*, Support: *Rawad*)

***References:*** (Everyone)

* <https://resole.org.uk/>
* <https://wasteawaygroup.com/blog/How-Many-Shoe-Are-In-the-Landfill/>
* <https://www.nature.com/articles/d42473-023-00276-5>
* <https://app.diagrams.net/#G1B22uYfLsBRiwRPylMyMuWfS7yJtCQwpE#%7B%22pageId%22%3A%22G7Msy1v34nbQPyZi0fvw%22%7D>