# 3. Data Collection

# 3.1 Survey Design and Structure

The data for this study were collected via an online survey focused on various aspects of wildland fire management in South Africa. The survey was carefully designed to gather both qualitative and quantitative information, targeting critical areas of wildland fire management practices. Its primary aim was to explore respondents' perceptions, experiences, and knowledge concerning wildland fire incidents, response strategies, recovery efforts, and readiness measures. Additionally, it collected personal and professional demographic data to contextualize responses across different societal and occupational groups.

The survey included a mix of open-ended and multiple-choice questions to facilitate both statistical analysis and in-depth qualitative insights. Open-ended questions allowed for detailed, narrative responses on specific topics, while multiple-choice questions provided structured, comparable data. This mixed-methods approach enabled the survey to capture a comprehensive view of respondents' experiences and perceptions related to wildland fire management.

Below is a detailed breakdown of the key areas covered by the survey:

## Review and Analysis

This section focused on understanding the broader impacts of wildland fires and analyzing specific aspects of fire management strategies in South Africa. Questions in this section aimed to capture respondents' insights on:

* **The negative impacts of wildland fires**: Respondents were asked to identify the three most unwanted effects of wildfires, with responses ranging from loss of life to destruction of biodiversity and economic assets.
* **Positive outcomes**: The survey inquired about the three most desirable effects of prescribed burns or fire management strategies, focusing on ecological benefits such as species diversity or land regeneration.
* **Success factors and areas for improvement**: Respondents provided their views on the most successful aspects of fire management in South Africa and identified areas that need improvement. This included both managerial practices and ecological outcomes.
* **Knowledge sharing**: The survey asked for aspects of fire management in South Africa that could be shared with other African nations, encouraging a comparative analysis of management practices across regions.

## Risk Reduction

The risk reduction category of the survey explored respondents’ perspectives on the causes and prevention of wildfires. This section aimed to identify:

* **Causes of wildfires**: Respondents were asked to specify the top two causes of human-originated wildfires, with options such as arson, agricultural burns, or negligent use of fire.
* **Preventive measures**: Open-ended questions inquired about measures that can mitigate the risk of wildfires. Respondents shared their thoughts on firebreaks, community awareness programs, and policy measures designed to reduce fire risks.
* **Intentional fires**: The survey sought respondents' opinions on the potential motivations behind intentional fires (ignitors), exploring underlying reasons such as land clearing, cultural practices, or malicious intent.

## Readiness

This section of the survey evaluated the level of preparedness within communities and organizations to respond to wildland fires. Respondents provided information on:

* **Training programs**: The survey asked about the existence and efficacy of national training programs or certification processes for fire management personnel. Respondents shared their experiences or knowledge about the availability and quality of such programs.
* **Sufficiency of personnel**: Questions were aimed at assessing whether there were enough trained professionals to handle wildland fires in the respondents' communities. This section provided insight into the operational readiness of local fire management agencies.
* **Professional training**: The respondents also discussed whether wildland fire suppression brigades were adequately trained and equipped to handle fires effectively. This involved gauging the overall level of skill and readiness within different regions.

## Response

The response section of the survey delved into the operational capacity of communities and organizations when dealing with fire incidents. This section captured information on:

* **Incident command systems**: Respondents were asked about the utilization of formal Incident Command Systems (ICS) during fire incidents in their communities. ICS is a standardized, on-scene, all-hazards incident management approach, and its presence was an indicator of organized fire management.
* **Equipment and resources**: The survey sought to understand the types of firefighting equipment and resources available in the community, such as fire trucks, water tenders, helicopters, or personal protective equipment. Respondents indicated whether such resources were government-provided or locally sourced.
* **Operational challenges**: Respondents also identified gaps in equipment, personnel protection, and tactical resources that may hinder the effectiveness of fire response operations in their areas.

## Recovery

This category focused on the strategies and support systems available for community and ecosystem recovery after wildland fire events. Key themes included:

* **Community support programs**: Respondents were asked whether there were formal programs in place to assist communities in recovering from the effects of damaging wildfires. This included financial aid, mental health support, and rebuilding efforts.
* **Ecosystem restoration**: The survey examined the availability of restoration guidelines or initiatives aimed at rehabilitating fire-affected ecosystems. Respondents shared their knowledge about biodiversity recovery efforts and whether they adhered to standardized environmental restoration practices.

## Demographic Information

In addition to the thematic questions about fire management, the survey collected detailed demographic data to enrich the analysis. This information helped profile the respondents and provided context to their responses. The demographic data collected included:

* **Gender and race**: This allowed for an analysis of how experiences and opinions varied across different demographic groups.
* **Geographical location**: Respondents were asked about the province and district in which they primarily resided, providing insight into regional differences in fire management experiences and challenges.
* **Occupation**: Respondents provided details about their professional background, particularly if they worked in fire management, environmental conservation, or related fields.
* **Demography**: The collected demographic data were synthesized into a "Demography" field, which combined race, gender, and location to allow for a quick and comprehensive view of the respondent’s background.

The survey design facilitated comprehensive data collection on various aspects of wildland fire management. By structuring questions around Review and Analysis, Risk Reduction, Readiness, Response, and Recovery, the survey encompassed both operational and ecological views on fire management in South Africa. Including demographic details enabled analysis of how fire management experiences and challenges differed across societal groups and regions.

# 3.2 Approach to data Processing

## Step 1: Data Collection and Initial Cleaning

Data collection was conducted via an online survey with South African respondents, focusing on wildland fire management. The survey gathered demographic data and open-ended responses on fire incidents, preparedness, and impacts. Participants shared their views on fire management strategies and the ecological effects of fires in their communities.

After collection, the data was exported into a structured spreadsheet, containing respondent information, survey answers, and metadata, like response start and end dates. The initial processing step involved cleaning this raw dataset to ensure usability, removing irrelevant columns and handling missing or inconsistent values. Columns such as "IP Address" and optional contact information were excluded to retain focus on core responses. Incomplete entries were either discarded or filled with placeholder values to maintain data integrity. The cleaned dataset, now standardized, included demographic information and survey responses for further analysis.

## Step 2: Categorization of Survey Responses

The second step of the process involved categorizing the responses into predefined categories based on the subject of each question. The survey questions were divided into six primary categories: *Response*, *Readiness*, *Recovery*, *Risk Reduction*, *Review & Analysis*, and *Miscellaneous*. Additionally, there was an *Additional Insight* category for questions that fell outside of the predefined categories.

The categorization process was guided by a keyword-based mapping strategy. Each survey question was analyzed against a predefined set of keywords associated with the different categories. For example, questions involving fire management personnel, equipment, and firefighting tactics were categorized under "Response," while questions related to training programs or preparedness efforts were classified under "Readiness." This mapping required a deep understanding of the survey structure and content.

This step also involved developing an algorithm to detecting key phrases within the questions. The algorithm worked by scanning each question, matching it against a list of predefined keywords, and assigning the appropriate category based on the content of the question. This categorization allowed the responses to be organized by theme, making it easier to analyze the data within each respective field. The output of this step was a newly categorized dataset that included the respondent ID, the category of the question, the question itself, and the corresponding response.

## Step 3: Merging Demographic and Categorized Data

The third and final step involved integrating the demographic information of each respondent with the categorized survey responses. The demographic information included fields such as race, gender, province, and occupation. The goal was to enrich the categorized data by including these demographic details, allowing for a more comprehensive analysis of the responses.

To achieve this, the cleaned dataset (which contained demographic information) was merged with the categorized dataset (which contained categorized survey responses). This merge was based on the respondent ID, ensuring that each response was linked to the corresponding respondent’s demographic data. A key challenge in this step was ensuring that the demographic fields, such as race, gender, and province, were appropriately aligned with the responses. The solution was to harmonize these fields, ensuring consistent formatting and handling any missing data carefully. In cases where respondents did not provide certain demographic details, placeholder values were used to maintain the structure of the data.

To further enhance the analysis, a new field, "Demography," was created by concatenating the race, gender, and province information into a single field. This allowed for a quick view of the respondent's demographic profile alongside their responses. The final dataset consisted of the following columns: *Respondent ID*, *Category*, *Question*, *Response*, *Race*, *Gender*, *Province*, *Occupation*, and *Demography*. This structure enabled detailed analysis across multiple dimensions, such as understanding how responses varied by demographic group or region.

# 3.3 Analysis Strategy

An analysis strategy was designed to extract insights from the Survey by focusing on the relationships between respondents’ demographic information and their responses across different categories of questions. This section details the methodological approach used for the analysis, focusing on the comparative and correlation-based frameworks applied to the dataset.

## 3.3.1 Approach to Comparative Analysis

The comparative analysis aims to understand how different demographic groups responded to the survey's various questions. The focus is on identifying patterns and trends across groups such as gender, race, province, and occupation, and how these groups perceive and interact with the topic of wildland fire management.

The comparative analysis begins by filtering the dataset to isolate responses to a particular question. Once the relevant data is extracted, the responses are further grouped based on a selected demographic, allowing for the comparison of answers across different respondent profiles. This helps to highlight the influence of factors such as location or occupation on attitudes toward wildland fire management.

For example, if analyzing responses to the question, “What are the top two causes of human-originated wildfires?” the data would be grouped by the selected demographic (e.g., province). The responses would then be compared across provinces to reveal whether geographic location has a significant impact on perceptions regarding wildfire causes. The results of these analyses are visually represented using different types of charts, such as bar charts or pie charts, depending on the data structure.

## 3.3.2 Approach to Correlation Analysis

Correlation analysis, in contrast to comparative analysis, aims to uncover potential relationships between responses to different survey questions. This analysis is valuable for identifying underlying patterns between different aspects of wildland fire management, as perceived by the respondents.

The correlation process starts by selecting two questions that are hypothesized to have some form of interdependency. For instance, responses to a question about the sufficiency of firefighting equipment may be correlated with responses regarding the preparedness of fire brigades. The responses to both questions are merged, allowing for the examination of how often certain responses to one question co-occur with specific responses to another.

Once the data is merged, a contingency table (or crosstab) is constructed to cross-classify the responses to the two questions. The strength of the relationship between the responses is visually represented using heatmaps. The correlation analysis provides insights into how different aspects of wildland fire management are interlinked in the perceptions of the respondents. For example, it can reveal whether respondents who believe that equipment is inadequate are also more likely to view community preparedness as insufficient. This type of analysis offers a more nuanced understanding of the challenges and opportunities in fire management.

## 3.3.3 Integration of Demographic Data in Analysis

A critical aspect of the analysis involves integrating respondents’ demographic data to provide additional context to their answers. The demographic data collected, including race, gender, province, and occupation, was systematically integrated into the analysis to allow for richer interpretations of the results.

In the comparative analysis, demographic information acts as the key factor for segmenting the data. This segmentation allows for comparisons between different groups and helps identify how personal and professional backgrounds influence perspectives on wildland fire management. For example, examining responses from individuals in different occupations, such as firefighters versus foresters, reveals how roles within the fire management ecosystem shape perceptions.

In the correlation analysis, demographic data does not directly affect the calculation of correlations but helps in understanding the broader context of the relationships between different responses.

## 3.3.4 Response Distribution and Top-N Analysis

There were more than 7000 records of data. In order to manage large datasets and focus on the most significant findings, the analysis includes a Top-N approach. This involves selecting the top N most frequent responses for analysis, which ensures that the most impactful or commonly held views are prioritized in the analysis. The function selects the top responses based on total mentions across all groups (e.g., Gender) and then filters for visualization accordingly. So Instead of selecting the top N rows, the function calculates the total number of mentions for each response across all groups and selects the top N responses based on the overall count. This ensures that the top N responses are truly the most frequently mentioned across all categories. The Top-N responses provide a clearer picture of general sentiment without being overwhelmed by outliers or less common opinions.

For example, in a question where respondents are asked about the most frequent causes of wildfires, only the top 20 responses might be visualized, allowing for more focused insights on the primary drivers of wildfires as perceived by the majority of respondents.

# 3.4 Using Streamlit for Exploration and Visualization

The survey findings were analyzed through a web-based application using Streamlit to enable interactive exploration and visualization. Streamlit's flexible and user-friendly platform allows users to dynamically filter and analyze data by selecting categories, questions, and demographic attributes via an intuitive sidebar. This real-time engagement facilitates quick insights based on specific interests or concerns for stakeholders, researchers, and decision-makers.

The app supports comparative or correlation analyses across various survey questions, offering tailored visualizations such as bar charts, pie charts, heatmaps, and line charts. Users can choose the type of analysis and chart to best represent the relationships and trends they wish to explore. This interactivity is crucial for presenting complex datasets like the South Africa Wildland Fire Survey in an accessible manner, enabling users to gain actionable insights without needing advanced technical expertise.

The analytical approach is customizable, accommodating various combinations of questions, demographics, and analysis types. Users can focus on specific survey questions or examine relationships between multiple questions. The option to switch between comparative and correlation analysis ensures a wide range of insights can be uncovered, aligning with research goals.

This flexibility supports the broader aim of understanding respondent perceptions and deriving actionable insights to inform policy and operational strategies in wildland fire management. The data-driven approach helps stakeholders identify key areas for improvement, assess existing programs' efficacy, and prioritize future research and interventions.

A detailed step-by-step guide on how to use the Streamlit app for exploring and visualizing the data can be accessed here <https://github.com/desmond-lartey/Wildfire-Assessment/blob/Fires/README.md>. This guide will assist users in fully leveraging the interactive features of the app to perform their own customized analyses.

# 4. Findings and Analysis

Survey responses revealed significant insights into perceptions and experiences of wildland fire management across South Africa, categorized into five thematic areas: Review and Analysis, Risk Reduction, Readiness, Response, and Recovery. Examination of these areas showed notable regional, demographic, and gender-based differences in experiences and perceptions related to wildland fires.

## 1. Review and Analysis

Our findings show that 58% of respondents either "somewhat agree" or "strongly agree" that wildland fires significantly harm their communities, indicating widespread acknowledgment of these detrimental effects. Notably, heightened concern was observed in the Western Cape province, where 11 out of 19 participants "strongly agree" about the severe negative impacts of wildland fires. This suggests that certain areas in South Africa are especially vulnerable to wildland fires, likely due to local environmental conditions or higher fire frequency.

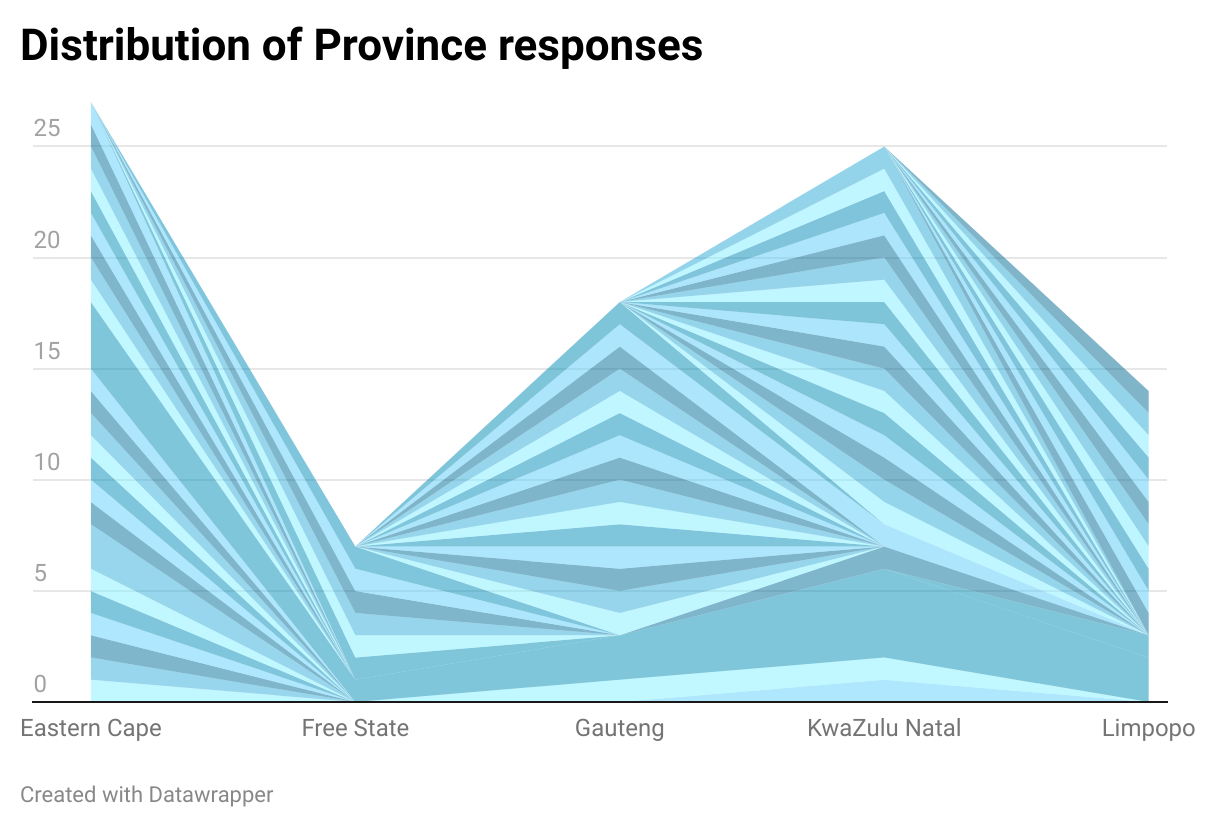
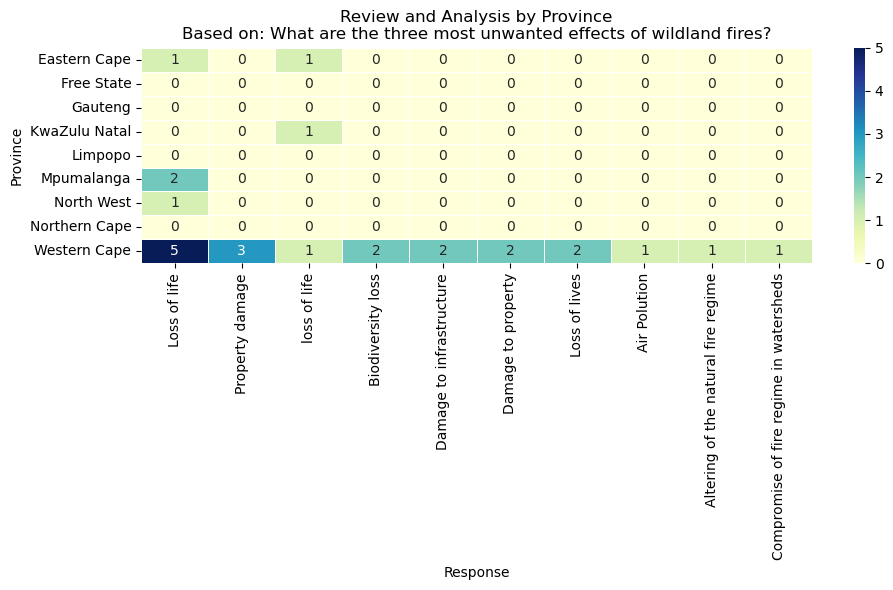


Figure xx: This heatmap shows responses to the survey question about unwanted effects of wildland fires across different provinces.

Demographically, 46% of Black respondents expressed concern about the adverse effects of wildland fires, possibly due to greater exposure or heightened awareness. Additionally, male respondents were more likely than females to "strongly agree" about the negative impacts of wildfires, indicating that men may have had more direct or negative experiences with such events.

## 2. Risk Reduction

Our analysis identified arson as the leading cause of human-originated wildfires, cited by 17 respondents, followed by negligence, mentioned by 7 respondents. When queried about intentional fires, 41 respondents pointed to arson as the main motive, while grazing and fuel management were also significant, with 16 and 11 mentions, respectively. These results indicate that although arson is a major issue, fire management practices related to grazing and fuel management also play a role.

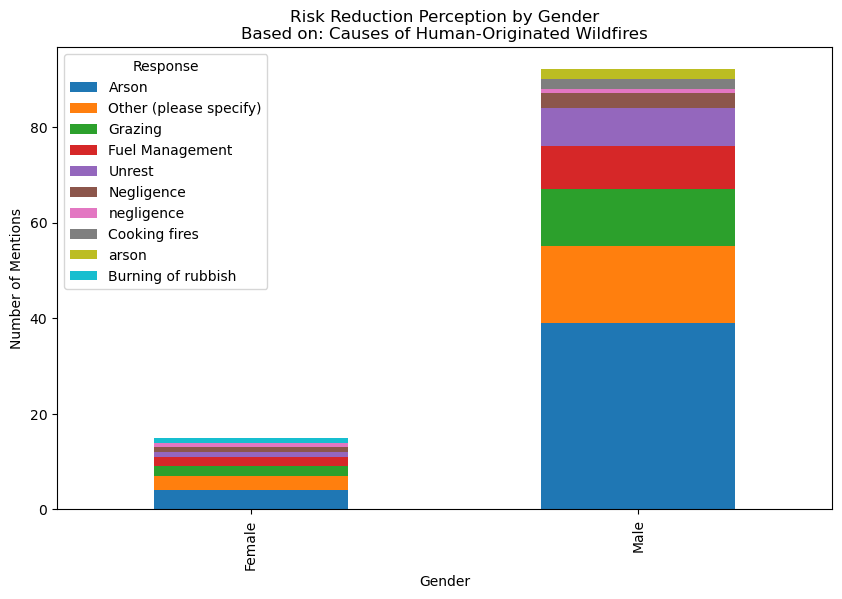


Figure xx: This chart shows the distribution of fire causes across gender

Female respondents identified arson as a primary concern for intentionally set fires, reflecting gender-specific safety concerns or experiences. This suggests differing perceptions of fire causes between genders.

## 3. Readiness

When evaluating the readiness of personnel to manage wildland fires, our results indicate mixed perceptions. While 29% of respondents "somewhat agree" that there are sufficient personnel to deal with wildland fires, 23% "strongly disagree," highlighting a significant divide in the perceived preparedness. Regional analysis shows that the Western Cape once again stands out, with 7 out of 19 respondents "strongly disagreeing" about personnel readiness, pointing to localized concerns in fire management capacity.



Figure xx: This radar chart shows perceptions of readiness across different racial groups based on survey responses to fire management preparedness

Regarding brigade training, we found that 42 respondents "somewhat agree" and 33 "strongly agree" that wildland fire suppression brigades are professionally trained and equipped. This positive outlook on brigade readiness was more prevalent among Black and Coloured respondents, who were more likely to express confidence in the brigade’s preparedness compared to White respondents. This demographic difference suggests varying levels of engagement or exposure to fire response efforts across communities.

## 4. Response

Our findings on the response capabilities to wildfires show that a substantial number of respondents, 36 "strongly agree" and 36 "somewhat agree," believe that a Formal Incident Command System is utilized in their communities. This indicates a generally positive perception of the structured mechanisms in place for managing fire incidents. However, we also noted regional disparities, with respondents in the Western Cape showing stronger agreement (8 out of 19 "strongly agree") compared to mixed perceptions in Eastern Cape and Gauteng.

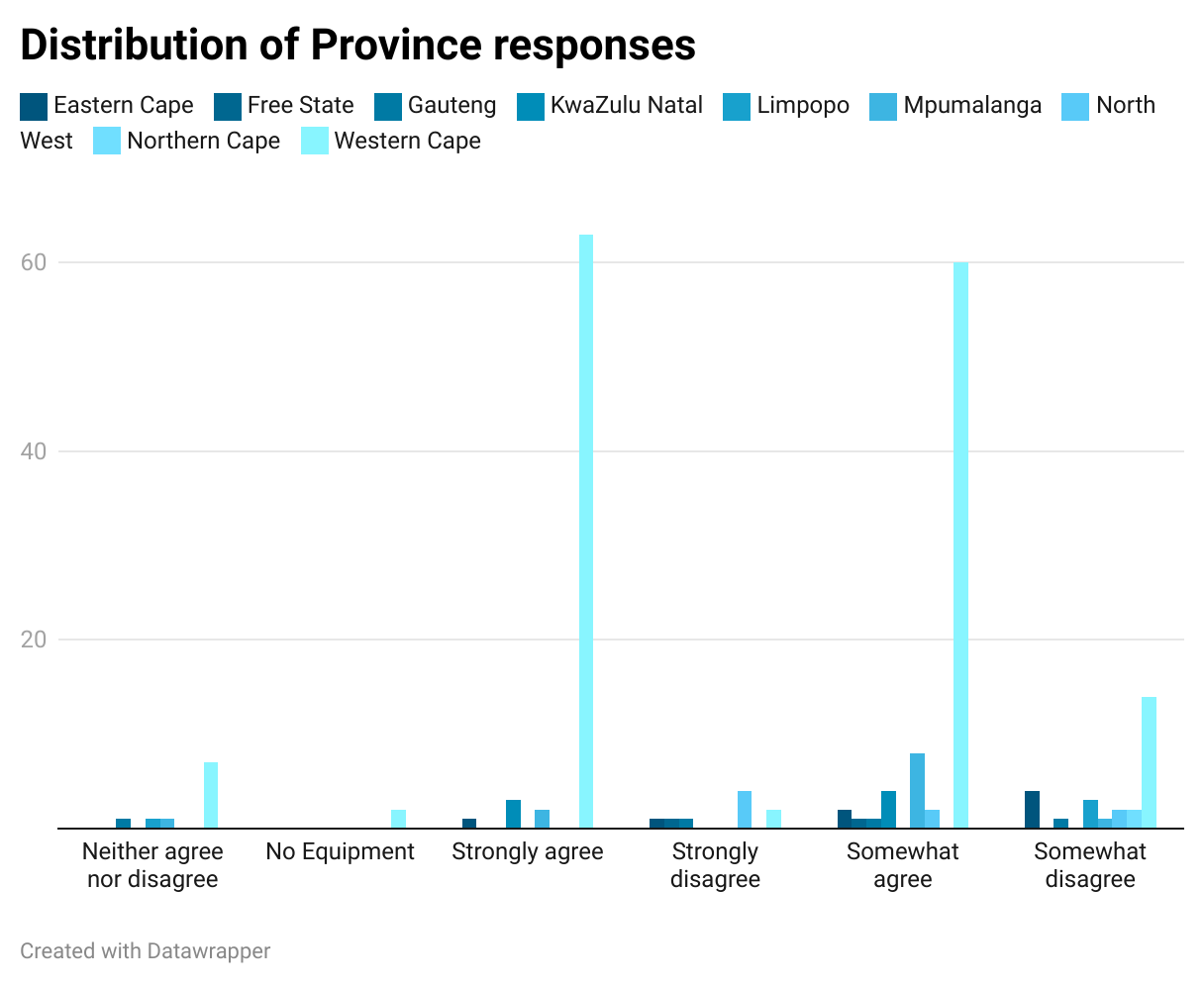
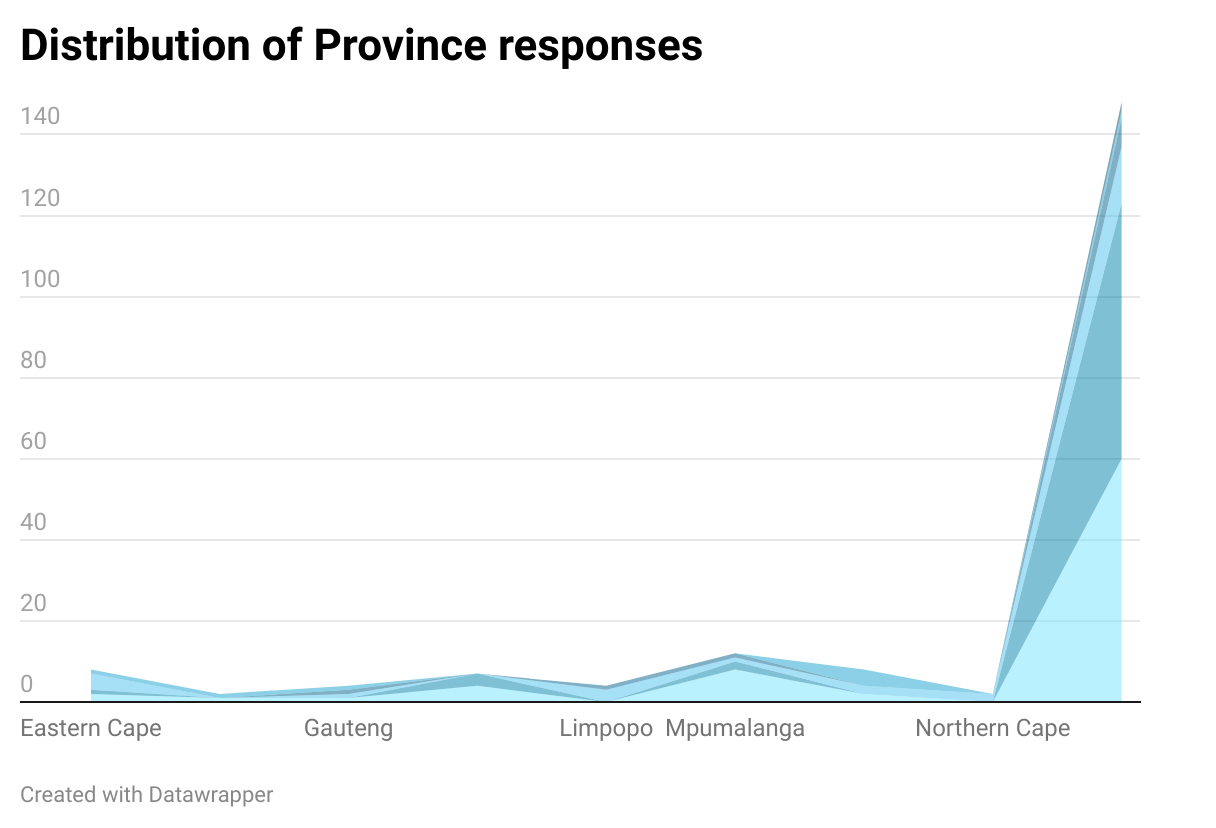
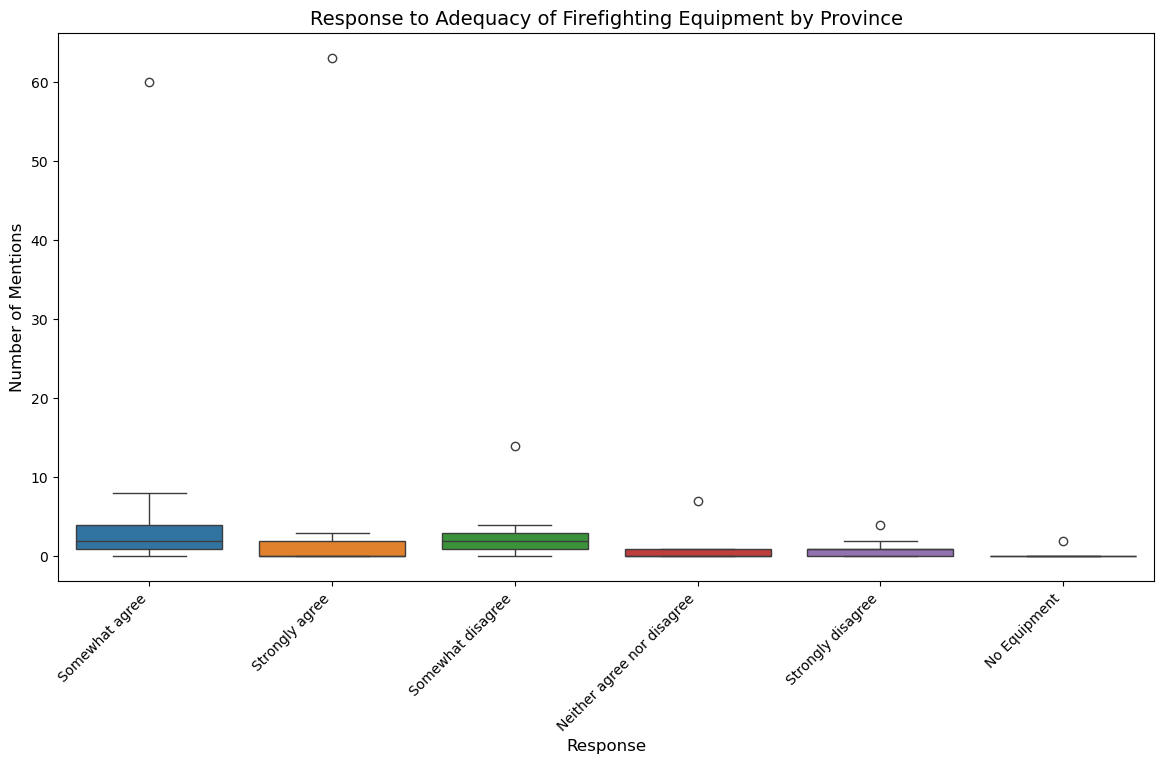


Figure xx: The box plot shows the spread of responses regarding equipment sufficiency across different provinces

In terms of equipment provision, 75 respondents either "strongly agree" or "somewhat agree" that firefighters are adequately equipped. However, 17 respondents expressed concerns about the quality or availability of equipment, suggesting that while most respondents view the equipment as sufficient, there are areas where improvements may be necessary. Gender-based analysis revealed that female respondents were slightly more optimistic about the provision of equipment compared to males, possibly indicating different levels of exposure to equipment inadequacies or a more positive perception of firefighting efforts among women.

## 5. Recovery

When it comes to recovery efforts following wildland fires, we found that 42 respondents either "somewhat agree" or "strongly agree" that there are restoration guidelines in place for ecosystems affected by wildland fires. However, a substantial proportion of respondents (32) expressed concerns, either "somewhat disagreeing" or "strongly disagreeing," suggesting that recovery programs may not be uniformly effective or adequately communicated.

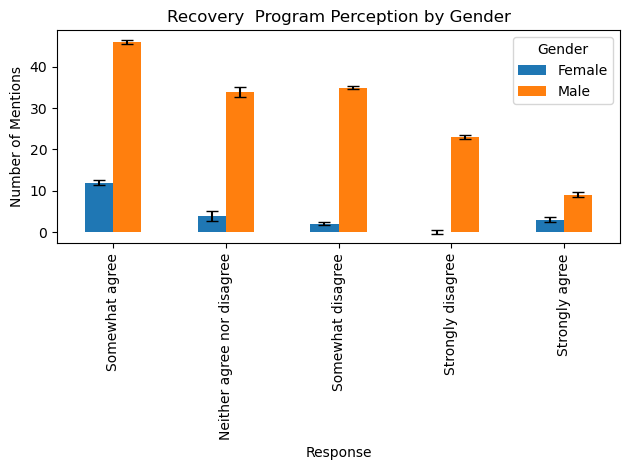


Figure: This chart shows how the effectiveness of recovery programs are perceived across gender

Regional data suggests disparities in recovery perceptions, particularly in the Western Cape, where responses were mixed regarding the presence of support programs and restoration guidelines after damaging wildfires. In terms of gender, male respondents were more likely to "strongly disagree" about the effectiveness of recovery programs compared to females. This may indicate that men in the survey have had more direct experiences with post-fire recovery efforts and found them lacking.

## 6. Summary of Key Insights

Our findings offer a detailed understanding of perceptions related to wildland fire management in South Africa. The Western Cape is notably concerned about personnel readiness, response efficiency, and recovery efforts. Demographic analysis reveals differences between Black and White respondents, with significant gender-based differences regarding equipment provision and arson concerns. Despite a general acknowledgment of wildland fire challenges, regional and demographic disparities necessitate targeted fire management strategies that address specific vulnerabilities and perceptions across South Africa.

# 5. Statistical Inferences

The statistical analysis of data across various demographic groups (gender and province) reveals important insights into perceptions of wildfire management in South Africa. The aim is to understand the perception of the 5r’s under the different groups.

The ANOVA test for general response counts by gender yielded an **F-statistic of 3.882** and a **P-value of 0.077**, suggesting variability between genders but not statistical significance at the 5% level. The high P-value implies that observed differences may be due to random chance rather than genuine gender-based response patterns. The Chi-Square test for gender in the general category produced a **Chi2 value of 8.669** and a **P-value of 0.123**, indicating no significant association between gender and general response counts. With a degree of freedom of 5, this suggests gender does not significantly affect the distribution of responses in this category.

Table xxx: Statistical inferences of the Wildland fire survey data

|  |  |  |  |
| --- | --- | --- | --- |
| Statistical Test | Category | Demographic | Score |
| ANOVA Test | General | Gender | F-Statistic: 3.882, P-Value: 0.077 |
| Chi-Square Test | General | Gender | Chi2: 8.669, P-Value: 0.123 |
| Descriptive Statistics | General | - | Count: 1347, Mean: 3.097, Std: 15.85, Min: 1, Max: 399 |
| Chi-Square Test | Review and Analysis | Gender | Chi2: 449.509, P-Value: 0.524 |
| Chi-Square Test | Review and Analysis | Province | Chi2: 4049.697, P-Value: 0.258 |
| Chi-Square Test | Risk Reduction | Gender | Chi2: 66.651, P-Value: 0.656 |
| Chi-Square Test | Risk Reduction | Province | Chi2: 559.136, P-Value: 0.983 |
| Chi-Square Test | Readiness | Gender | Chi2: 6.122, P-Value: 0.190 |
| Chi-Square Test | Readiness | Province | Chi2: 42.344, P-Value: 0.104 |
| Chi-Square Test | Response | Gender | Chi2: 8.669, P-Value: 0.123 |
| Chi-Square Test | Response | Province | Chi2: 114.016, P-Value: 0.00000000490 |
| Chi-Square Test | Recovery | Gender | Chi2: 10.395, P-Value: 0.034 |
| Chi-Square Test | Recovery | Province | Chi2: 62.006, P-Value: 0.001 |
| Chi-Square Test | Miscellaneous | Gender | Chi2: 135.484, P-Value: 0.638 |
| Chi-Square Test | Additional Insights | Gender | Chi2: 624.447, P-Value: 0.006 |
| Chi-Square Test | Miscellaneous | Province | Chi2: 1352.106, P-Value: 0.058 |
| Chi-Square Test | Additional Insights | Province | Chi2: 5917.040, P-Value: 5.05e-35 |

In the Review and Analysis category, the Chi-Square test showed no significant association between gender **(Chi2 = 449.509, P = 0.524)** or province **(Chi2 = 4049.697, P = 0.258)** and the number of responses. Similarly, for Risk Reduction, there was no significant link between gender **(Chi2 = 66.651, P = 0.656)** or province **(Chi2 = 559.136, P = 0.983)** and response counts. In the Readiness category, results were not significant for gender **(Chi2 = 6.122, P = 0.190)** or province **(Chi2 = 42.344, P = 0.104)**.

In the Response category, Chi-Square tests showed significant differences in response counts across provinces **(Chi2 = 114.016, P = 0.00000000490)**, but not by gender **(Chi2 = 8.669, P = 0.123),** indicating provincial variation but no gender difference. The Recovery category demonstrated significant differences for both gender **(Chi2 = 10.395, P = 0.034)** and province **(Chi2 = 62.006, P = 0.001)**, indicating that both factors influenced perceptions or engagement with recovery efforts.

The Miscellaneous category showed no significant differences based on gender (Chi2 = 135.484, P = 0.638) or province (Chi2 = 1352.106, P = 0.058). Lastly, the Additional Insights category revealed statistically significant differences in response counts by both gender **(Chi2 = 624.447, P = 0.006)** and province **(Chi2 = 5917.040, P = 5.05e-35),** indicating that both demographic factors significantly influenced the distribution of responses in this category.

The results indicate that in categories like Review and Analysis, Risk Reduction, and Readiness, response counts did not show statistically significant differences based on gender or province. This implies that respondents from diverse backgrounds generally share similar views on these aspects of wildfire management. Consequently, policies in these areas could be designed more broadly to address general perceptions of wildfire impacts and management efforts.

Significant differences were identified in the Recovery, Response and Additional Insights categories. Statistical analysis in the Recovery category revealed notable variations in perceptions based on gender and province, indicating regional disparities in recovery support or differences in how men and women are impacted by or involved in recovery processes. This insight is vital for policymakers, suggesting that recovery strategies should be tailored to address specific concerns of different groups, ensuring equitable distribution of resources considering regional vulnerabilities and gendered experiences. Similarly, the Additional Insights category showed significant differences across gender and province, reflecting diverse regional challenges or varying levels of engagement with wildfire management. This underscores the need for localized policies that address the unique concerns and priorities of different communities.