XN Project: Final Report

Executive Summary:

Kids At Risk Action (KARA), through "AMERICA'S CHILDREN IN 100 CHARTS," is dedicated to raising public awareness of the issues impoverished children confront. This research showcases the use of an extensive dataset on American cases of missing children, spanning the years 1930–2017. The dataset offers data regarding missing, mistreated, and neglected children in an easy-to-understand visual format and evaluates the efficacy of Child Protective Services (CPS). An analysis of the dataset yields significant insights into the circumstances and demographics of missing children. It comprises information on names, dates of birth, sexes, races, physical attributes, and case information. KARA uses images to make the information more impactful and enlightening in order to influence public opinion and legislative action.

Business Problem:

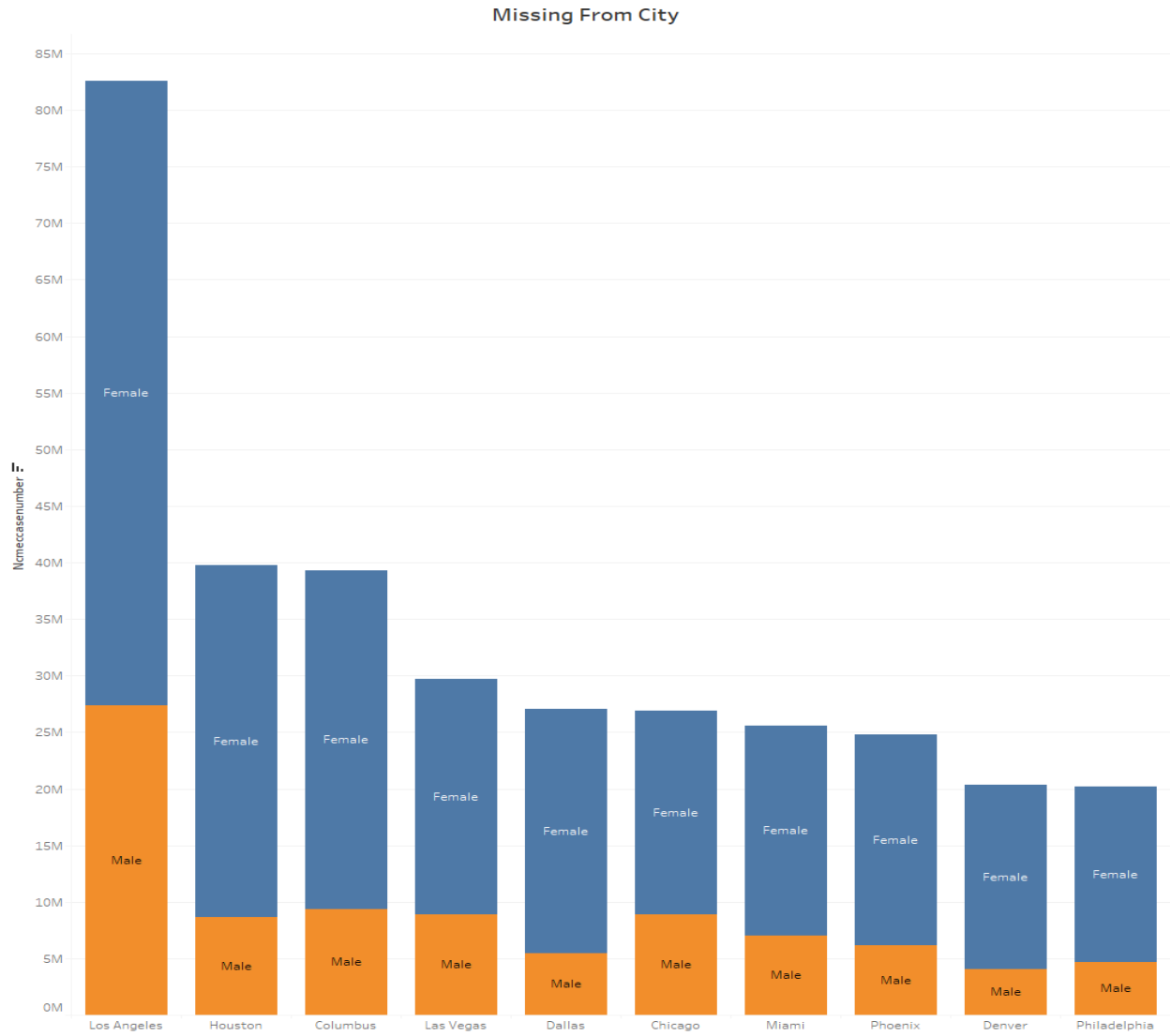
Using data on missing children and "AMERICA'S CHILDREN IN 100 CHARTS," what is the best way for Kids At Risk Action (KARA) to inform the public and policymakers about the challenges faced by kids who are deemed to be at risk? How may KARA make use of this information to enhance the effectiveness of Child Protective Services (CPS)? Moreover, how might KARA assist these marginalized groups in managing and reducing maladaptive behaviors such as aggression, crime, subpar academic achievement, and psychological issues?

Clear Concise Flow:

We started the data preparation and collection part of our project, making sure the data came from reputable sources. After that, a thorough data purification procedure that involved eliminating rows with null values was carried out, which reduced the size of the dataset from 2834 to 2503 entries. After that, in order to find important patterns and trends that would direct our further investigations, we conducted exploratory data analysis, or EDA. With the use of visualization technologies like Tableau/Power BI/Python, we were able to produce shareable and comprehensible dashboards that helped people grasp the data more deeply. Our main objective was to pinpoint the essential measures that are necessary to comprehend the overall performance and direct subsequent actions. Ultimately, we combined our research results into succinct and understandable visuals to offer Kids At Risk Action (KARA) useful information.

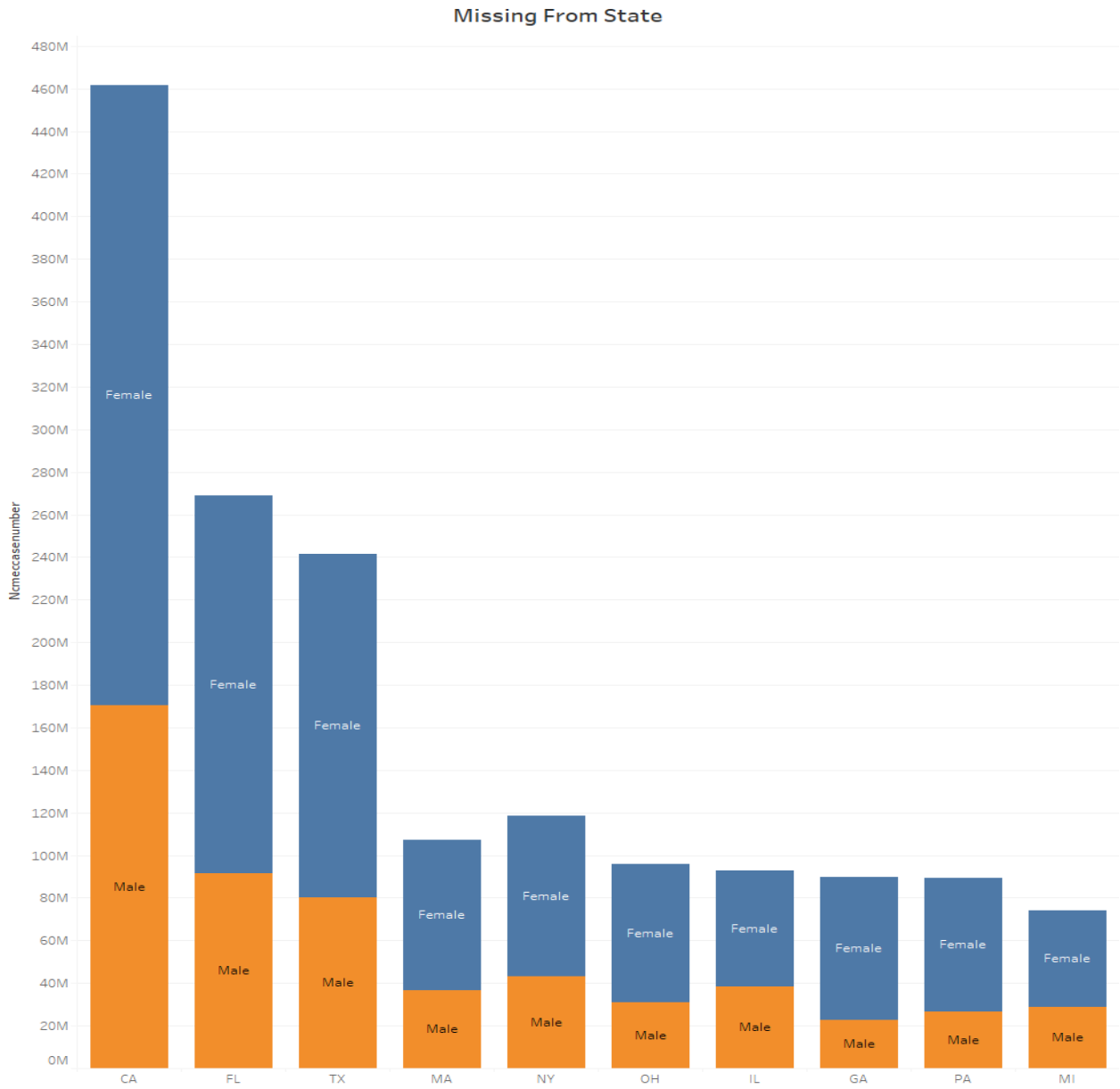
The main findings of the group analysis:

In the analysis of missing persons by city, Los Angeles holds the highest number of cases, with females significantly outnumbering males. Other cities like Houston, Columbus, and Las Vegas also show more missing females than males. Cities such as Dallas, Chicago, Miami, Phoenix, Denver, and Philadelphia have similar levels of missing persons, with females consistently being the majority.



1. Missing From State

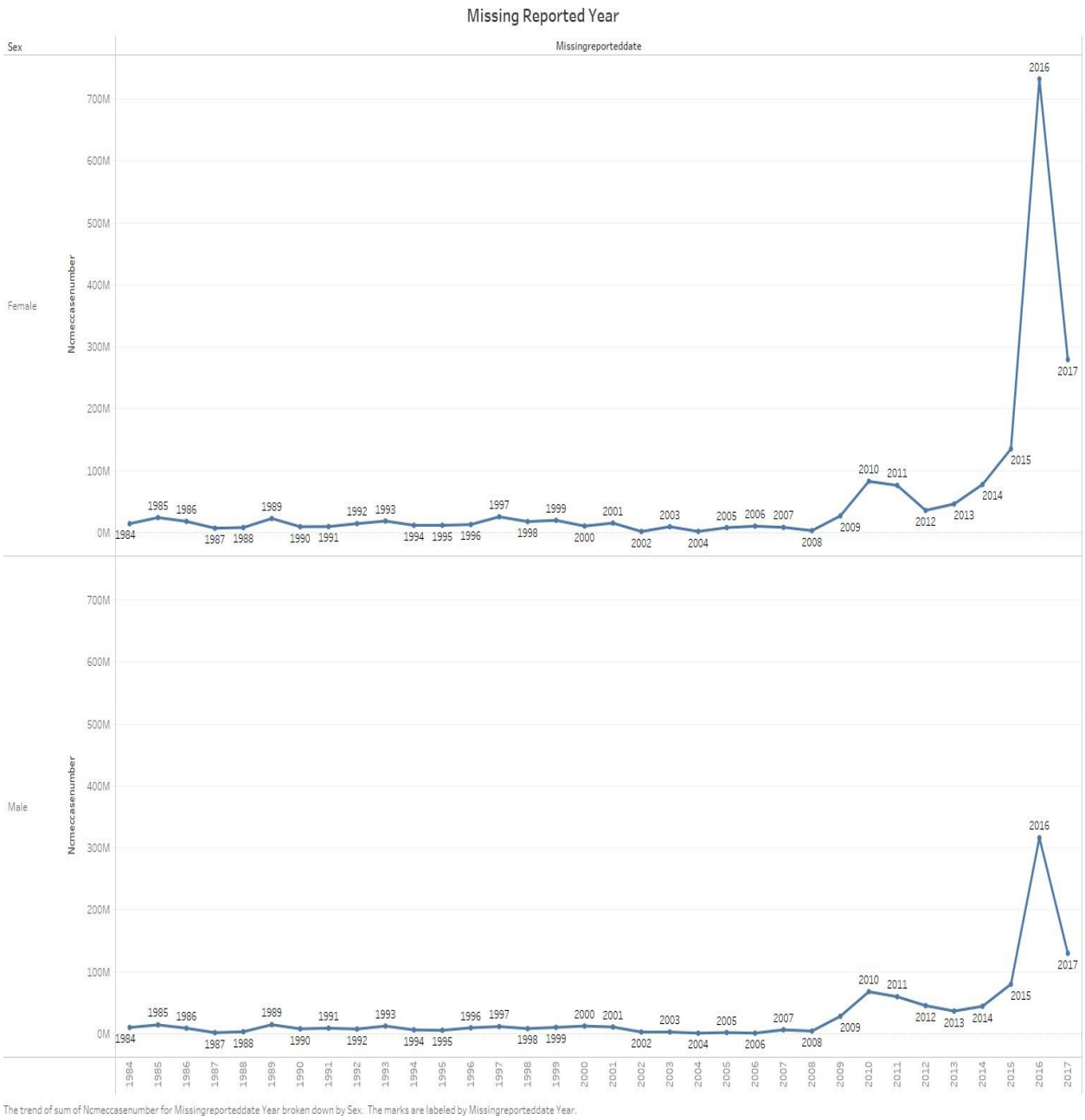
At the state level, California has the highest number of missing persons cases, with females significantly outnumbering males. Florida and Texas follow, also showing a majority of missing females. Other states like Massachusetts, New York, Ohio, Illinois, Georgia, Pennsylvania, and Michigan have similar levels of missing persons, with females consistently being the majority.



Sum of Ncmeccasenumber for each Missingfromstate. Color shows details about Sex. The marks are labeled by Sex. The view is filtered on Missingfromstate, which keeps 10 of 53 members.

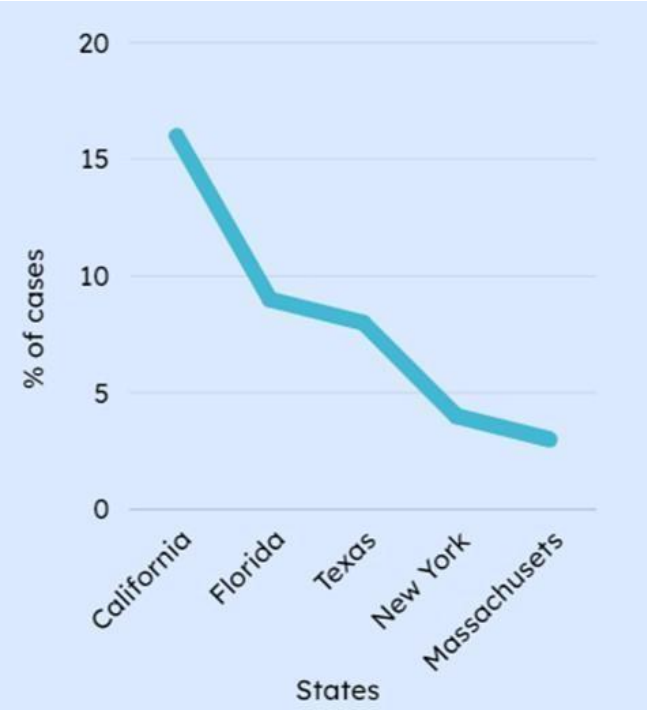
2. Missing Reported Year

The trend of missing persons remained relatively stable and low from 1984 to 2009. However, there was a noticeable increase starting in 2010 and 2011, followed by a dramatic spike from 2014, peaking in 2016. The numbers dropped again in 2017 but remained higher than the pre-2010 levels. Throughout these years, females consistently had more missing persons than males.



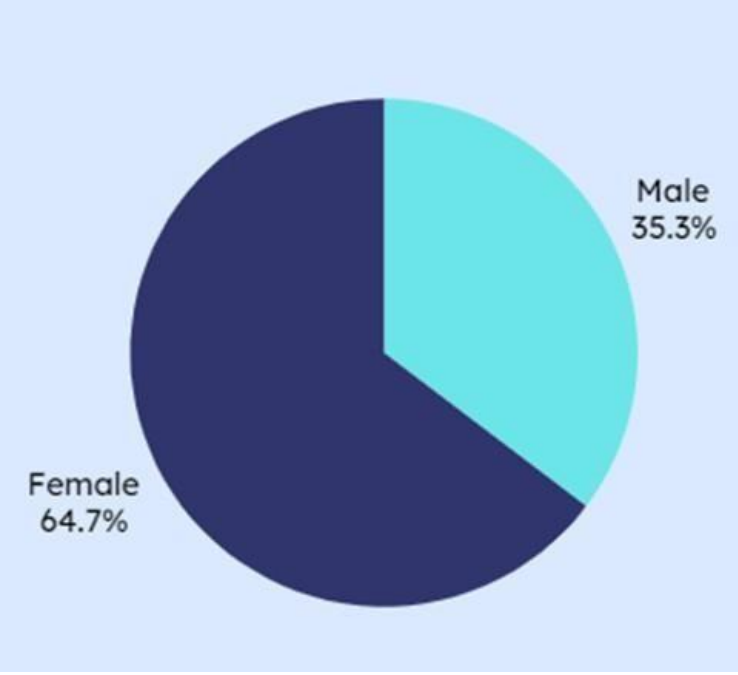
3. Top 5 cities where kidnapping is prevalent

The research findings reveal substantial variations in the prevalence of these issues across different states. The image demonstrates that states like California have a considerably higher percentage of cases, approximately 16%, compared to states like Massachusetts, which has a significantly lower percentage of around 3%. This discrepancy highlights the necessity for a customized, data-driven approach that takes into account the regional disparities in risk factors and the specific service requirements of each area.



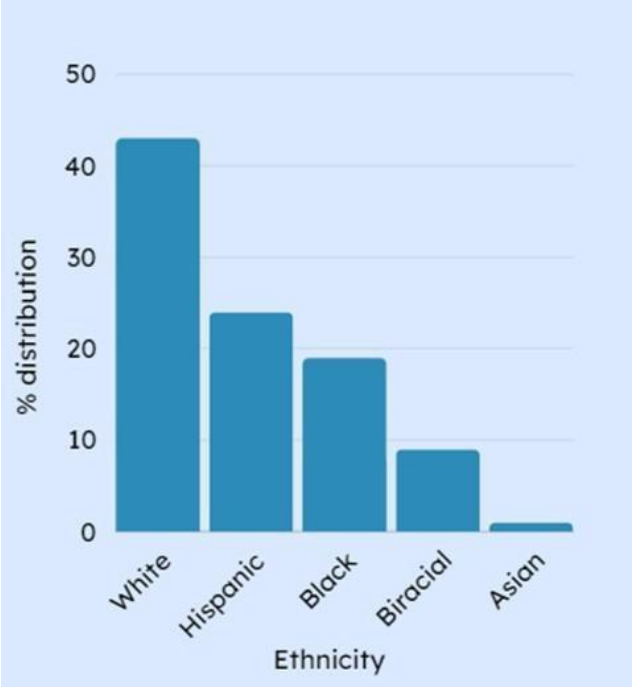
4. Ratio of Sex of children kidnapped

The data presented in this image indicates that a significant majority, approximately 64.7%, of the children affected by these issues are female, while 35.3% are male. This gender disparity suggests that policies and strategies need to take into account gender-specific risk factors and that interventions may need to be tailored differently to address the unique needs of boys and girls effectively.



5. Distribution of ethnicity in cases

This Image sheds light on the ethnic distribution of the affected children, revealing that White children constitute the largest group at approximately 45%, followed by Hispanic children at around 25%, Black children at around 18%, and smaller percentages for other ethnic groups. This data underscores the significance of taking into account cultural and socioeconomic factors that may vary across different ethnic communities when developing the predictive model and designing interventions. These considerations are crucial to ensure that the proposed solutions are tailored and effective for diverse populations.

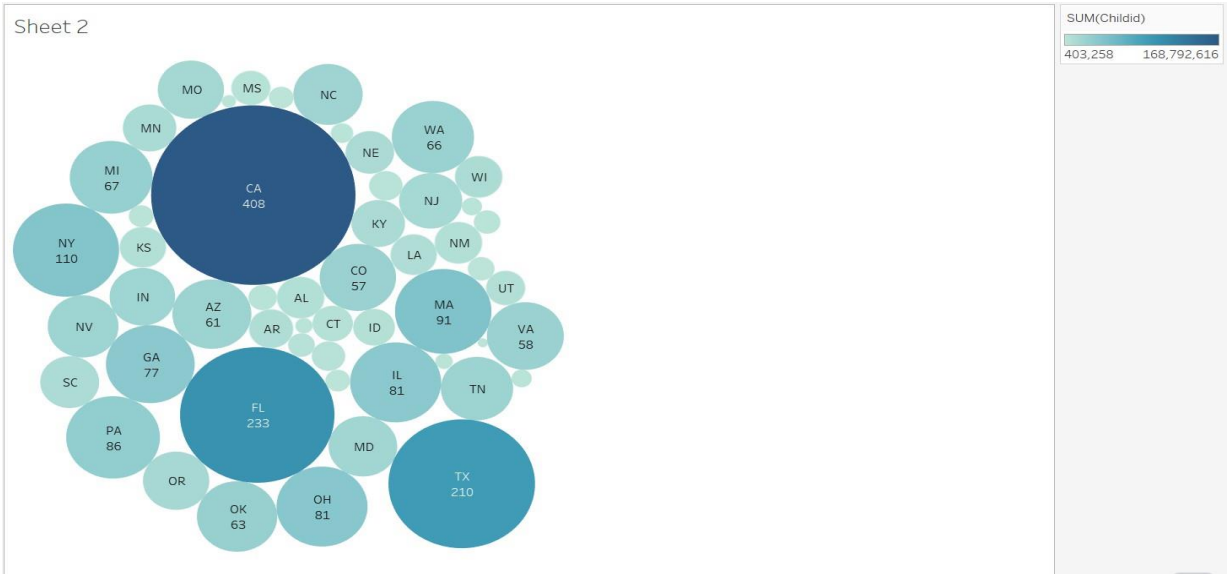


6. Case incident types

The bar chart illustrates the distribution of child cases based on different incident types. It reveals that the most prevalent issue is "Endangered Runaway," with 1,354 cases, far surpassing other categories. "Lost Injured Missing" follows with 541 cases, highlighting another significant concern. Family-related incidents include 277 "Family Abduction" and 222 cases "Non-Family Abduction." Lastly, "Section 5779" accounts for 109 cases. The data underscores that runaway incidents are the most frequent, indicating a critical area for intervention and support.

7. Number of cases

The bubble chart highlights the top three states with the highest cases. California leads with 408 cases, followed by Florida with 233 cases, and Texas with 210 cases. These three states have significantly higher case counts compared to others, indicating a regional concentration of incidents. This suggests that California, Florida, and Texas are critical areas for addressing the issues being measured.



A how-to kind of guide that explains how to use the dashboard and or model the group created and is delivering:

Our dashboard, created using Tableau, visualizes missing person data. Here's how to navigate and use it effectively:

1. Navigate:

- Locate the specific chart or visualization you are interested in (e.g., bar charts for "Missing From City" and "Missing From State," or line charts for "Missing Reported Year").

2. Interaction:

- Hover Over Elements: Hover over bars or lines to see the detailed information. This will display the combined case numbers from the dataset.

- Tooltips: Utilize tooltips that appear when hovering to get quick insights into the data points.

3. Filtering and Sorting:

- Use any available filters or sorting options to customize the view. You can focus on specific cities, states, or years depending on your interest.

4. Zoom and Pan:

- If the dashboard includes interactive features for zooming and panning, use these to explore different data sections more closely.

5. Legends and Labels:

- Refer to the legends and labels on the charts to understand the color coding and categories used in the visualizations.

6. Export and Share:

- Use Tableau's built-in options to export visualizations or share the dashboard with others for collaborative analysis.

Tool Used

We created these visualizations using Tableau. Tableau's interactive features allow us to explore the data efficiently and gain insights into the patterns and trends in cases of missing persons.

Analysis and Synthesis of The Data:

In order to identify important trends and changes over time that could guide KARA's operational plans, our investigation looked at the overall data directions. We divided the data on the missing children by gender, ethnicity, and region and performed a demographic analysis. This gave us insight into the groups that were most impacted and served as a foundation for focused treatments. The analysis brought to light important findings that contributed to a better comprehension of the data and its consequences for child welfare. To ensure compelling and clear communication, we used Tableau/Power BI/Python to visually exhibit these insights. Our all-inclusive methodology offered a full summary of the data, bolstering KARA's strategic planning and facilitating more efficient resource and intervention prioritization.

Recommendations and Findings:

Findings:

1. Geographic Discrepancies in Missing Children Cases:
    - Prevalence variance: Across states, there is a notable variance in the prevalence of incidents involving missing children, according to the KARA dataset. For example, Massachusetts has a lesser percentage (3%), whereas California has a higher percentage (16%) of cases.
    - Implication: This emphasizes the necessity for region-specific techniques to deal with the problem, acknowledging that many nations would need specialized methods depending on their particular difficulties and resources.
2. Gender Disparity in Kidnapping Cases:
    - Gender Ratio: According to the data, there are 35.3% male and 64.7% female missing children.
    - Implication: Gender-sensitive policies and interventions that target certain risk factors and needs that are particular to boys and girls are required. This could entail developing preventative strategies and specialized support networks for each gender.
3. Ethnic Distribution of Missing Children:
    - Ethnic Distribution: White people account for 45% of missing children, followed by Hispanic people (25%), Black people (18%), and lower proportions of other ethnic groups.
    - Implication: To meet the varied requirements of various ethnic communities, interventions must be socioeconomically and culturally appropriate. Comprehending the distinct obstacles encountered by every faction can facilitate the development of more efficient and comprehensive resolutions.

Recommendations:

1. Enhanced Public Awareness and Education:
    - Use Visual Data: To inform the public and policymakers on the seriousness and complexities of the problems faced by at-risk children, make use of the extensive dataset and visualizations from "AMERICA'S CHILDREN IN 100 CHARTS".
    - Conduct Campaigns: Create public awareness campaigns that draw attention to the unique difficulties faced by high-prevalence states and disproportionately impacted populations, such as women and particular ethnic groups.
2. Improving CPS Effectiveness:
    - Data-Driven Strategies: By using the information to spot patterns and trends in child welfare cases, CPS will be able to better allocate resources and create interventions that are specifically targeted.
    - Training and Development: Make an investment in providing CPS employees with gender-sensitive and culturally competent training so they can better meet the specific needs of various populations.
3. Policy Advocacy and Legislative Action:
    - Informing Policymakers: Advocate for policies that alleviate regional inequities and give high-risk communities sufficient support by presenting the findings to legislators.
    - Gender-Specific Policies: Promote the creation and execution of gender-specific laws that address the underlying causes of these differences and acknowledge the increased risk that female youngsters face.
4. Community and Stakeholder Engagement:
    - Collaborative study Initiatives: To improve the general knowledge and strategy for working with children who are at risk, collaborate with advocacy organizations and academic institutions to carry out additional study and exchange insights.
    - Stakeholder Feedback: In order to improve tactics and guarantee that interventions are workable and efficient, it is important to consistently solicit feedback from stakeholders, such as communities and organizations that assist children who are at risk.
5. Long-Term Impact Studies:
    - Intervention Effectiveness: To assess the long-term success of child welfare policies and programs, carry out longitudinal studies. This will enable modifications and enhancements based on factual data.
    - Advanced Analytics: Use predictive modeling and machine learning to spot new trends and possible intervention points. This will enable you take proactive measures to handle problems before they get out of hand.
6. Policy Impact Analysis:
    - Assessing Advocacy Outcomes: To better help children who are at-risk, assess the practical effects of policy changes that are the result of advocacy activities on a regular basis. This involves identifying areas that require improvement and successful tactics.
7. Enhanced Data Visualization:
    - Captivating Presentations: Make use of cutting-edge data visualization tools to convey findings in a way that is easy to grasp, compelling, and applicable to the public and policymakers.

By putting these suggestions into practice, KARA will be able to make better use of the data on missing children in order to safeguard and assist vulnerable groups by increasing awareness, influencing legislation, and enhancing the effectiveness of Child Protective Services.

Future Research:

Expanding upon our preliminary discoveries, subsequent investigations will explore crucial domains to augment our comprehension and efficacy in tackling concerns associated with children who are at-risk. In order to evaluate the long-term impacts of particular initiatives, we will perform longitudinal impact studies to monitor the effectiveness of present child welfare laws and practices across time. To more accurately detect trends and possible intervention areas, advanced data analysis techniques will be applied, such as machine learning for pattern recognition and predictive modeling. To enable more thorough and in-depth studies, collaborative research efforts with academic institutions and advocacy groups will be developed to pool resources and expertise. Furthermore, policy impact analysis will assess how advocacy-driven policy changes actually affect the wellbeing of children who are at-risk and pinpoint effective approaches.

Relevant Information:

From the beginning, our project has been a team effort, with each member contributing to different parts of the research and analysis. This has allowed us to handle difficult challenges inside the child welfare system using a multifaceted approach. We have followed stringent ethical norms throughout the project, especially when managing sensitive data pertaining to vulnerable populations and upholding the confidentiality and integrity of our research approach. Our research emphasis and techniques have been refined as a result of ongoing analysis and stakeholder feedback, which has improved our ability to solve the complex problems inside the child welfare system. The project's course has been greatly shaped by stakeholder involvement, which has guaranteed that our efforts closely correspond with the requirements of children who are at risk and the organizations who assist them. Last but not least, we've added new research and papers to our list of resources, making sure our conclusions are supported by the most up-to-date and thorough data accessible.

References

- Gray, J. (2017).Missing Children in the US. data.world. https://data.world/jamesgray/missing-children-in-the-us
- NCMEC Data. (n.d.). Missing Children Analysis. NATIONAL CENTER FOR MISSING & EXPLOITED CHILDREN. https://www.missingkids.org/ourwork/ncmecdata