

Abstract geometric lines in the top-left corner of the slide, consisting of several thin, light brown lines that intersect to form various polygons and shapes.

UNDERSTANDING DELINQUENCY: ANALYZING CATEGORIES, AGE, GENDER, ENVIRONMENT, AND GEOGRAPHICAL DISTRIBUTION

Team 7



MOTIVATION

This project examines delinquent behaviors and their causes through analysis of factors such as age, gender, environment, and location to identify patterns and trends, with the goal of identifying potential areas for intervention and prevention.

Understanding delinquency is essential for creating effective strategies to prevent and address delinquent behavior. Analyzing various factors such as categories, age, gender, environment, and geographical distribution can provide a comprehensive understanding of the problem and help in developing targeted interventions.

By understanding different categories of delinquent behavior, such as drug abuse, property crimes, or violent offenses, we can identify the underlying causes and risk factors specific to each category. Similarly, analyzing age and gender differences can help in designing prevention programs that are age and gender-appropriate.

Environment and geographical distribution also play a crucial role in delinquency. For instance, delinquent behavior is often linked to the neighborhood's socioeconomic status, family environment, and access to community resources. Analyzing these factors can help in identifying the root causes of delinquent behavior and developing targeted interventions.

Overall, understanding delinquency through analyzing different categories, age, gender, environment, and geographical distribution can help in creating evidence-based policies and programs that effectively prevent and address delinquent behavior.

Two thin orange lines intersect on the left side of the slide. One line is horizontal, and the other is diagonal, crossing it.

LITERATURE REVIEW

Delinquent behaviours can be influenced by various factors, including age, gender, socioeconomic status, and environmental conditions

Studies conducted by Farrington (1998), Moffitt (1993), and Sampson & Laub (1993) have demonstrated the impact of these factors on the manifestation and prevalence of delinquency

The geographical distribution of delinquency has been linked to cultural, economic, and social differences among countries, as shown by Piquero (2008)

This project aims to further explore the relationships between these factors and delinquency categories

STORY

The five graphs provided present different aspects of delinquent behavior and its distribution across various demographic and geographic categories.

The first graph, a bar graph, shows the distribution of delinquent categories and their frequencies. This graph provides an overview of the most common types of delinquent behavior and the frequency with which they occur.

The second graph, a packed bubble plot, shows the distribution of delinquent categories by age group. This graph allows us to see which age groups are most likely to engage in specific types of delinquent behavior.

The third graph, a heat map, shows the distribution of delinquents by age and surroundings. This graph helps us understand the relationship between delinquent behavior and the environment in which it occurs.

The fourth graph, a text table, shows the distribution of delinquents by age, gender, and surroundings. This table provides a more detailed breakdown of delinquent behavior across different demographic categories.

The fifth graph, a choropleth map, shows the count of delinquent categories across different countries. This graph allows us to see how delinquent behavior is distributed across different parts of the world and identify areas where it may be more prevalent.

Overall, these graphs provide a comprehensive picture of delinquent behavior and its distribution across different demographic and geographic categories, which can help policymakers and researchers better understand and address the issue of delinquency.



VISUALIZATIONS

VISUALIZATION-1

Bar graph- The distribution of delinquent categories and their frequencies.

VISUALIZATION-2

Packed bubble plot- The distribution of delinquent categories by age group.

VISUALIZATION-3

Heat map- The distribution of delinquents by age and surroundings.

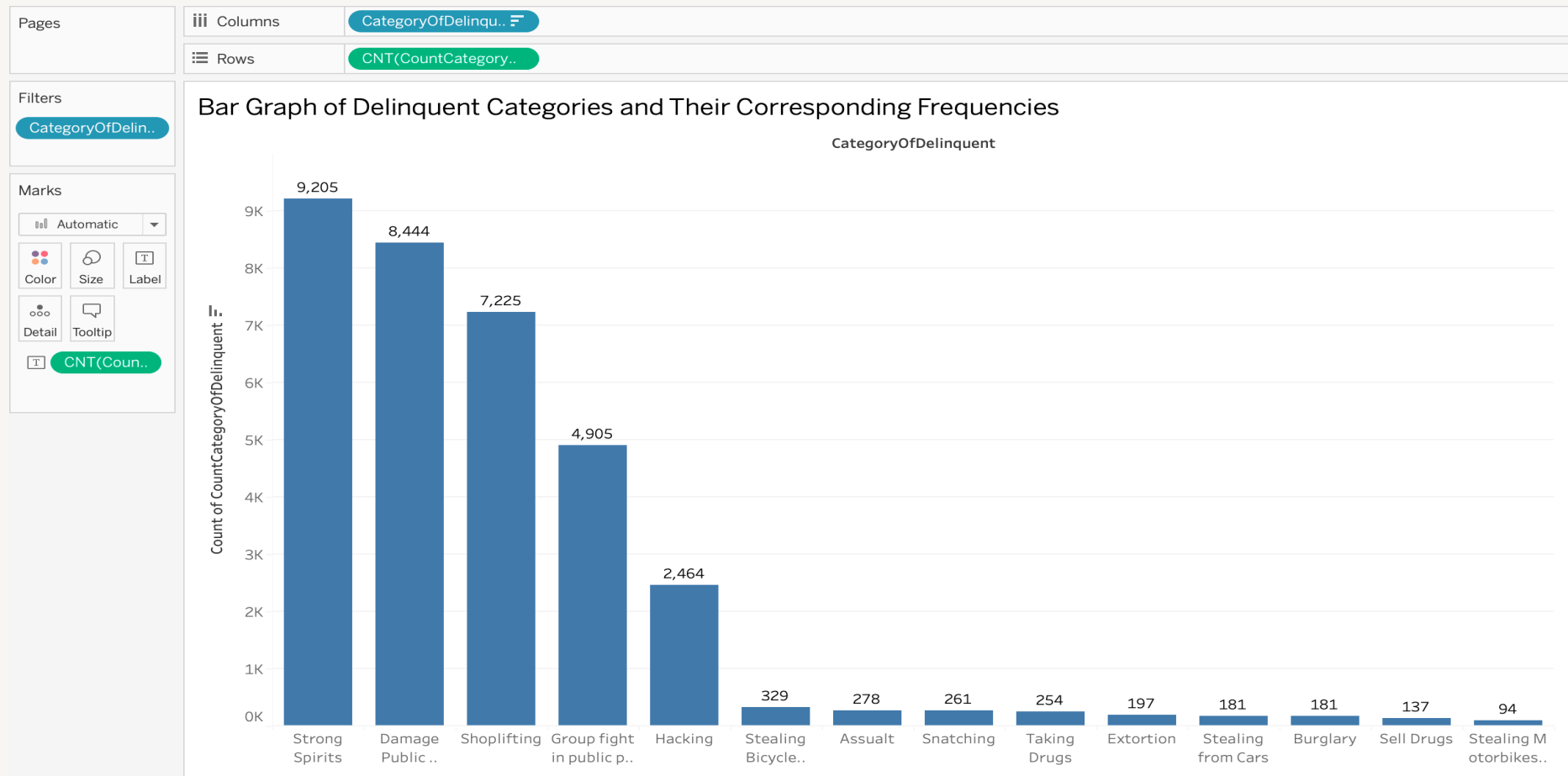
VISUALIZATION-4

Text table- The distribution of delinquents by age, gender, and surroundings.

VISUALIZATION-5

Choropleth- The count of delinquent categories across different countries.

VISUALIZATION 1



A series of thin, light brown lines crisscrossing the left side of the slide, creating an abstract geometric pattern.

VISUALIZATION 1 DETAILS

MARK:

Bars

CHANNEL:

Position(Vertical), length

DERIVED COLUMNS:

CategoryOfDelinquent:

used these columns to derive this column by segregating each type of act into a delinquency category

Vandltp, Shopltp, Burgltp, Bictltp, Cartltp, Hackltp, Carbltp, Snatltp, Extoltp, Gfigltp, Asltltp, Drudltp, Spirltp, Hashltp, Xtcltp, Lhcltp

CountCategoryOfDelinquent:

used these columns to derive this column and to get counts of each delinquent category

Vandltp, Shopltp, Burgltp, Bictltp, Cartltp, Hackltp, Carbltp, Snatltp, Extoltp, Gfigltp, Asltltp, Drudltp, Spirltp, Hashltp, Xtcltp, Lhcltp



WHY MARK AND CHANNEL

MARK:

The mark used in the bar chart is bars. Bars are used to represent discrete or categorical data where the height of the bar represents the value of the category. In this case, the categories are the different delinquent categories, and the height of the bar represents the frequency of delinquency in each category. The bars provide a clear and concise representation of the frequency distribution of the delinquent categories.

CHANNEL:

The channel used in the bar chart is position (vertical) and length. Position is used to represent the value of the data point along the y-axis, and length is used to represent the magnitude of the value. In this case, the length of the bar represents the count of delinquents in each category. The vertical position of the bars also makes it easy to compare the frequencies of the different delinquent categories.

HOW WERE THE DERIVED COLUMNS DERIVED?

CategoryOfDelinquent DERIVED COLUMN

```
ELSEIF [Shopltlp] = 1  
THEN "Shoplifting"  
  
ELSEIF [Burgltlp] = 1  
THEN "Burglary"  
  
ELSEIF [Bictltlp] = 1  
THEN "Stealing Bicycle, Moped or Scooter"  
  
ELSEIF [Cartltlp] = 1  
THEN "Stealing Motorbikes or Cars"  
  
ELSEIF [Hackltlp] = 1  
THEN "Hacking"  
  
ELSEIF [Carbtltp] = 1  
THEN "Stealing from Cars"  
  
ELSEIF [Snattltlp] = 1  
THEN "Snatching"  
  
ELSEIF [Extoltlp] = 1  
THEN "Extortion"  
  
ELSEIF [Gfigltlp] = 1  
THEN "Group fight in public place"  
  
ELSEIF [Asltltlp] = 1  
THEN "Assault"
```

CountCategoryOfDelinquent DERIVED COLUMN

CountCategoryOfDelinquent

```
IF ([Asltltlp] = 1 OR [Burgltlp] = 1 OR [Vandltlp] = 1 OR [Shopltlp] = 1  
OR [Burgltlp] = 1 OR [Bictltlp] = 1 OR [Cartltlp] = 1 OR [Hackltlp] = 1 OR  
[Carbtltp] = 1 OR [Snattltlp] = 1 OR [Extoltlp] = 1 OR [Gfigltlp] = 1 OR  
[Drudltlp] = 1 OR [Spirltp]=1 OR [Hashltlp]=1 OR [Xtcltp]=1 OR [Lhclt  
THEN "YES"  
ELSE "NO"  
END
```

Effectiveness - The visualization showed which delinquent behavior was committed the most.

Expressiveness - Bar charts show positive and negative values, but only positive values are used for the delinquent count in this case.

In this visualization, we have taken the overall view of all the different types of delinquencies and their frequencies.

As we can see the highest is strong spirits at a count of 9205 and the least is Stealing motorbikes with a count of 94.

We have chosen a bar graph as it makes it easier to find which delinquency category has the highest count.

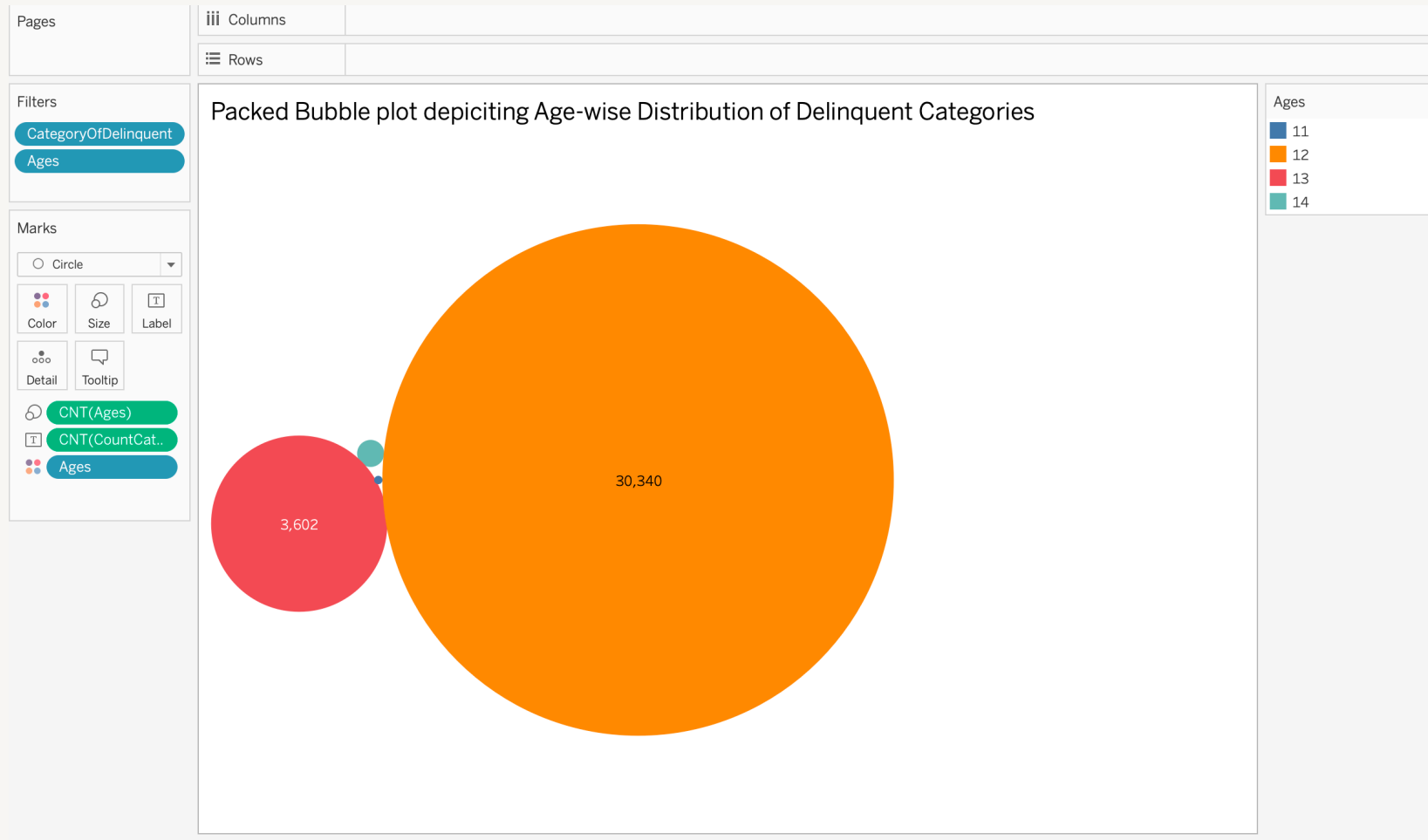
LITERATURE REVIEW

The literature review supporting the use of Delinquent Categories and Their Corresponding Frequencies can be found in various criminology studies. For example, a study by Kempf-Leonard (1995) showed that juvenile delinquency can be categorized into few categories. This study provides a framework for understanding delinquent behavior and can help researchers and practitioners in the field of criminology to better understand the nature and prevalence of delinquency.

One study that supports the use of delinquent categories is "Understanding and Treating Adolescent Substance Abuse: Why the Multiple Problem Youth Require Multiple Services" by Henggeler et al. (1997). Another study by Elliott et al. (1985) categorized juvenile delinquents into different groups. The authors found that these categories were useful in predicting future delinquent behavior and could help identify at-risk youth who may benefit from early intervention.

The use of a bar chart to visualize Delinquent Categories and Their Corresponding Frequencies is suitable because it clearly displays the frequency of delinquency in each category and allows for easy comparison between the different categories. The bars provide a visual representation of the distribution of delinquent behavior, making it easy to identify which categories have a higher or lower frequency of delinquency. This can be useful for researchers and practitioners in the field of criminology who are interested in understanding the prevalence and nature of delinquency in different categories. Overall, the use of a bar chart to visualize Delinquent Categories and Their Corresponding Frequencies is a simple and effective way to communicate complex information in a clear and concise manner.

VISUALIZATION 2





VISUALIZATION 2 DETAILS

MARK:

Circles

CHANNEL:

Area, color

DERIVED COLOUMNS:

Ages: we have used Agegroup column to get the data for each age group



WHY MARK AND CHANNEL

MARK:

The mark used in the Packed Bubble chart is circles. Circles are effective in displaying data that has a hierarchy, such as in this case where we have different age groups and different categories of delinquency within each age group. The size of each circle represents the frequency of delinquency within an age group for a particular category, making it easy to compare the frequency of delinquency across different categories within each age group.

CHANNEL:

The channels used in this visualization are area and color. Area is used to represent the frequency of delinquency within each age group for a particular category, while color is used to distinguish between different categories of delinquency.

HOW WERE THE DERIVED COLUMNS DERIVED?

Ages

Ages

```
IF [Agegroup]=0 then "11"  
ELSEIF [Agegroup]=1 then "12"  
ELSEIF [Agegroup]=2 then "13"  
ELSEIF [Agegroup]=3 then "14"  
ELSEIF [Agegroup]=4 then "15"  
ELSEIF [Agegroup]=5 then "16"  
ELSEIF [Agegroup]=6 then "17"  
ELSEIF [Agegroup]=7 then "18+"  
ELSE "None"  
END
```


Effectiveness - The visualization showed the relationship between age group and delinquency

Expressiveness - Allows easy comparison of multiple dimensions in a single chart.

In this visualization, we have taken the overall view of all the different types of delinquencies and their frequencies as per age.

As we can see the highest is age 12 with a count of 30340 and the least is age 11 with a count of 8.

We have a Bubble plot as it helps in the visualization of which age group is the highest.

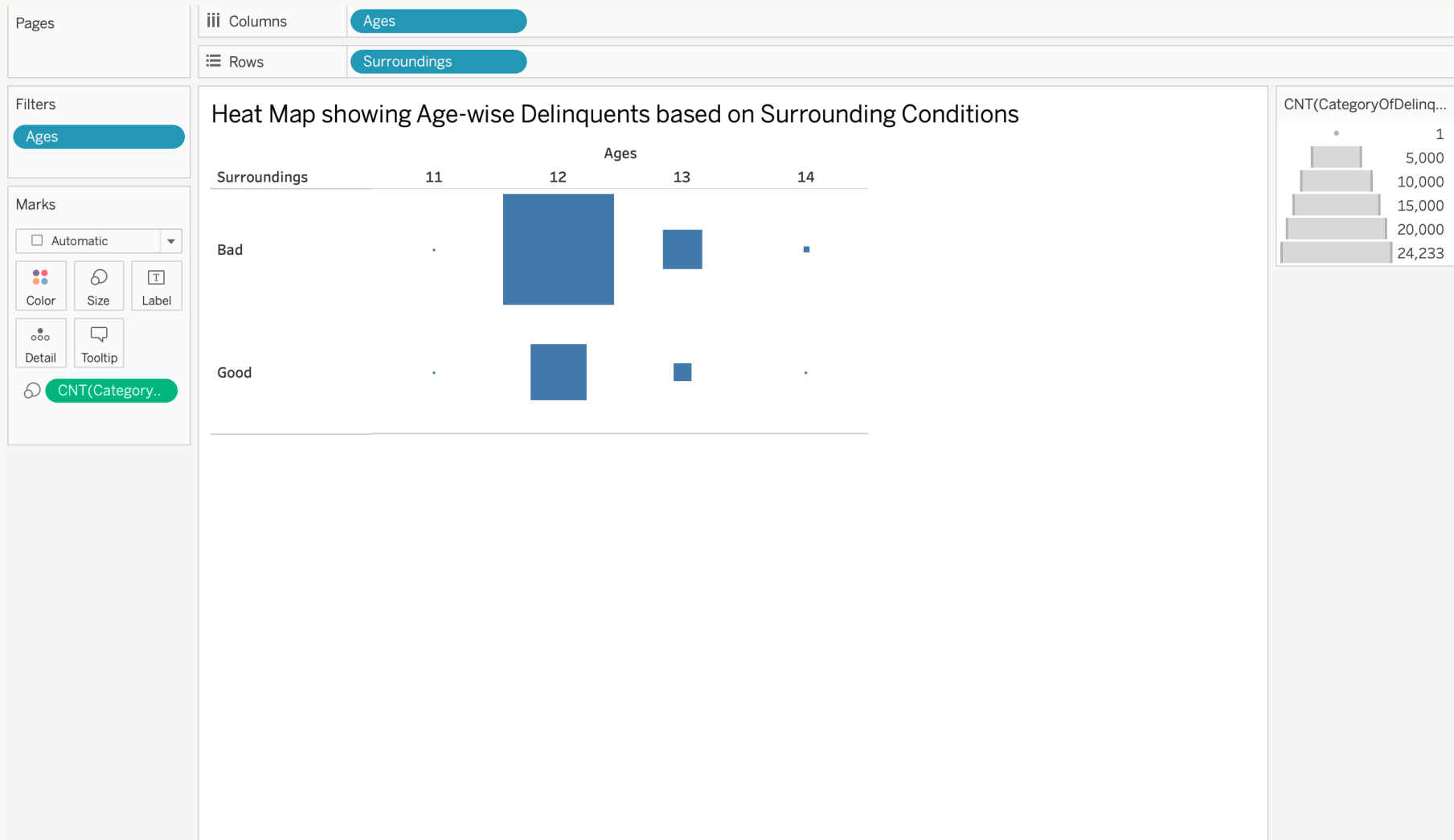
LITERATURE REVIEW

The use of packed bubble plots to represent hierarchical data has been well established in literature. One study by Hierarchical Edge Bundles by Danny Holten (2006) showed that packed bubble plots can be used to represent hierarchical data with high accuracy and precision, while also being aesthetically pleasing and easy to interpret.

In the context of delinquent categories and their corresponding ages, a packed bubble plot can be a useful tool to visualize the distribution of different delinquent categories within each age group. The size of the bubble can represent the relative frequency of the delinquent category, while the color can represent the type of delinquent category. This visualization can help identify trends and patterns in the distribution of delinquent categories across different age groups.

Therefore, using a packed bubble plot to represent the age-wise distribution of delinquent categories is a suitable visualization technique for this type of data, and can provide valuable insights into the distribution of different types of delinquent behavior across different age groups.

VISUALIZATION 3





VISUALIZATION 3 DETAILS

MARK:

Rectangles

CHANNEL:

Size, Position (grid)

DERIVED COLUMNS:

Surroundings: we have used Nhood05, Nhood06, Nhood07, Lifeev06, Lifeev07, Lifeev08, Lifeev02, Lifeev01, Lifeev03 to get good and bad environment.



WHY MARK AND CHANNEL

MARK:

The mark used in this heatmap is rectangles. Rectangles are used because they provide a clear and consistent shape for each cell in the grid, making it easier for the viewer to compare values across the different age and surroundings.

CHANNEL:

Size: By varying the size of the rectangles, we can encode a quantitative variable (e.g., the frequency of delinquents) in a visual way. Larger rectangles represent a higher frequency, while smaller rectangles represent a lower frequency.

Position (grid): By arranging the rectangles in a grid pattern, we can encode two categorical variables (e.g., age group and surrounding conditions) in a clear and organized way. This allows us to easily compare the frequency of delinquents across different age groups and surrounding conditions.

Effectiveness - The visualization showed if the surroundings were good or not

Expressiveness - Can show large complex datasets in an easy-to-understand visual format.

In this visualization, we have considered good and bad surroundings to find out what kind of environment was affecting that kind of behavior.

As we can see the highest in Bad environment is Age 12 with a Count of 24, 333 and the least is Age 11 with a Count of 7.

As we can see the highest in Good environment is Age 12 with a Count of 6 and the least is Age 11 with a Count of 1.

We have used heatmap as it helps in visualization of the highest in each type of environment and becomes easier to compare.

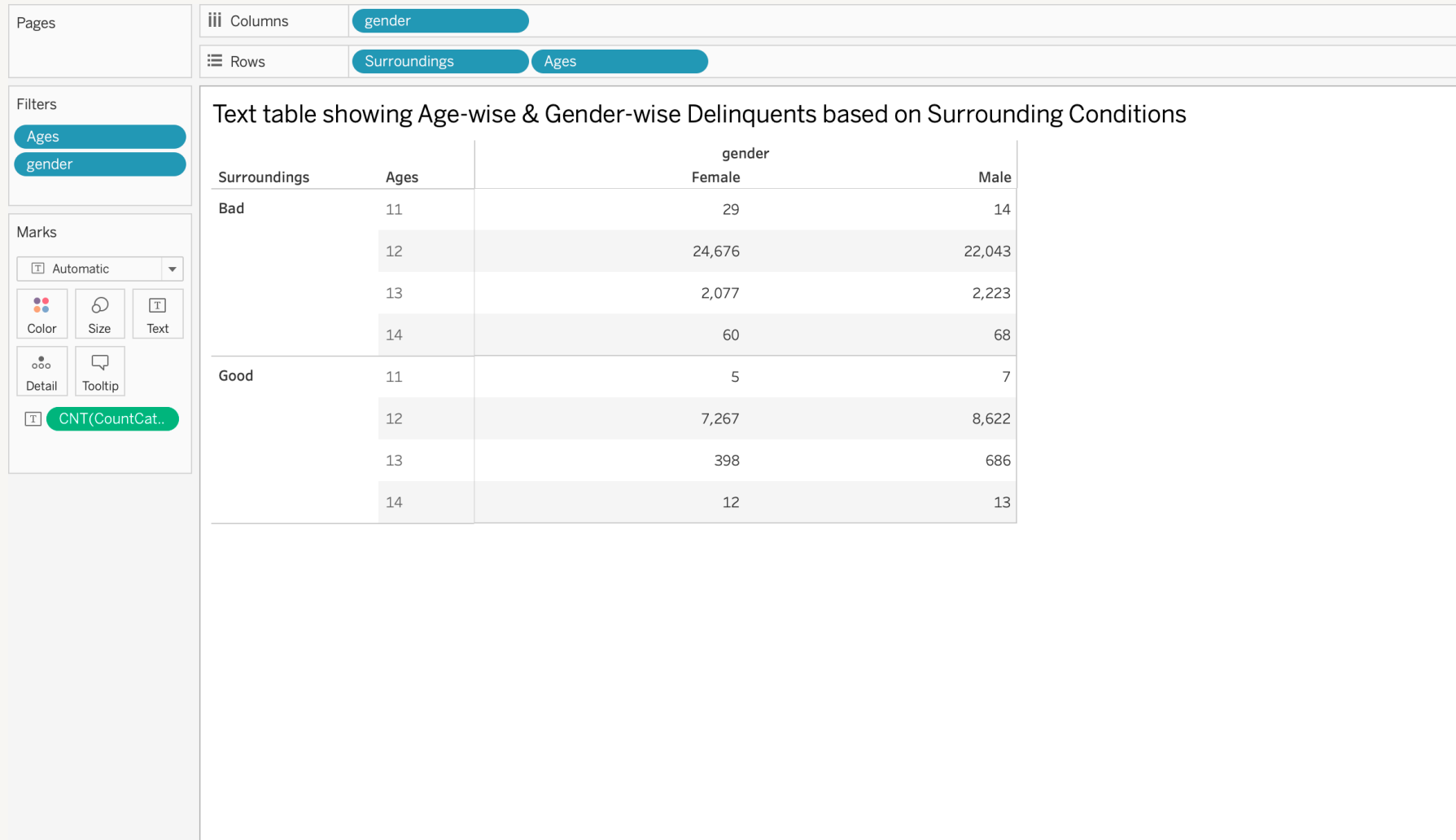
LITERATURE REVIEW

There is literature supporting the use of heatmaps to visualize patterns in multidimensional data. Heatmaps are particularly useful when the data has many rows and columns, making it difficult to identify patterns using tables or simple bar charts. They allow viewers to quickly identify patterns and relationships between variables.

For example, a study by Porst et al. (2020) used heatmaps to visualize patterns in environmental quality data across different neighborhoods in Germany. Another study by Chen et al. (2019) used heatmaps to identify patterns in crime data across different neighborhoods in a city. In both cases, the use of heatmaps allowed the researchers to identify patterns and relationships that would have been difficult to identify using other types of charts or tables.

In this case, the heatmap is suitable for visualizing patterns in delinquency rates across different age groups and environmental conditions. It allows viewers to quickly identify which age groups are more likely to engage in delinquent behavior under different environmental conditions, providing insights that may be difficult to obtain using other types of charts

VISUALIZATION 4





VISUALIZATION 4 DETAILS

MARK:

Text

CHANNEL:

Position (grid)



WHY MARK AND CHANNEL

MARK:

For the text table visualization, we used the mark as text to represent the numerical data in a tabular format. Text tables are a common way to display data in a structured format that is easy to read and interpret. Using a text mark allows us to display the data in a compact and organized manner, which is particularly useful when we need to show large amounts of data in a limited space.

CHANNEL:

The channel used for this visualization includes position and color. We use the position channel to organize the data into rows and columns, making it easy for the user to compare and analyze the data.

Effectiveness - Delinquency based on surroundings with the criteria of Age and Gender

Expressiveness - Simply shows data.

In this visualization we have considered good and bad surroundings to find out what kind of environment was affecting that kind of behavior along with that we have considered gender as a factor.

As we can see the highest in Bad environment is Age 12 with a Count of 24, 676 for females and males, the Count is 22043.

As we can see the least Bad environment is Age 11 with a Count of 29 for females and males, and the Count is 14 for Age 14.

As we can see the least in a Good environment is Age 12 with a Count of 7267 for females and males the Count is 8622.

As we can see the least Good environment is Age 11 with a Count of 5 for females and males the Count is 11 for Age 7.

We have used a Text table as it makes it easier to compare as we are considering three factors. Other visualizations will make it difficult to understand.

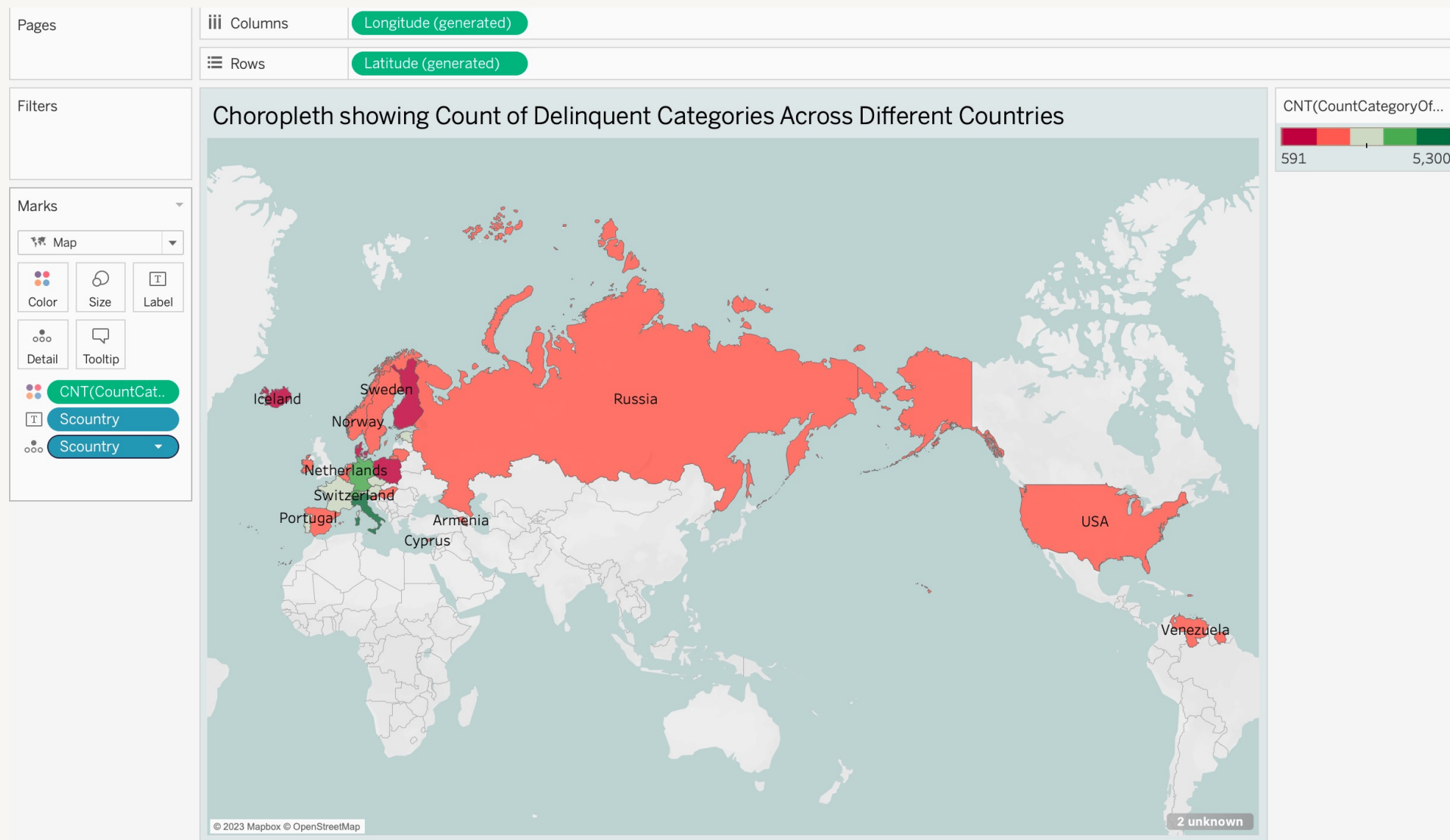
LITERATURE REVIEW

The use of a text table is a common way to represent data in a tabular format. It is suitable for presenting large amounts of information in a compact and easy-to-read format. In the case of this visualization, it allows for a quick comparison of the number of delinquents across age groups and genders based on their surrounding conditions.

Research studies have shown that the environment a person grows up in can have a significant impact on their behavior and development. For example, a study conducted by the National Bureau of Economic Research found that growing up in a high-poverty neighborhood is associated with an increased likelihood of criminal behavior (Sharkey, 2018). Another study published in the Journal of Research in Crime and Delinquency found that exposure to violence and crime in the neighborhood can lead to an increase in delinquent behavior among adolescents (Hill et al., 2001).

Therefore, by using a text table to show the number of delinquents across age groups and genders based on their surrounding conditions, this visualization provides a clear and concise representation of the impact of environmental factors on delinquent behavior. It allows for easy comparison and analysis of the data, which can help in identifying patterns and trends.

VISUALIZATION 5





VISUALIZATION 5 DETAILS

MARK:

Regions

CHANNEL:

Color, Position (geographical)



WHY MARK AND CHANNEL

MARK:

Regions or Geographical areas are commonly used marks for Choropleth maps as they help to convey the data associated with specific geographic regions, such as countries, states, or counties.

CHANNEL:

Color is used to represent the magnitude or density of the data associated with a specific geographic region, while Position on the map allows viewers to easily identify and compare the data between different regions.

Effectiveness - The visualization shows the count of a count of delinquency in countries

Expressiveness – Choropleth maps can be visually impactful, using color to emphasize key points and draw attention to presented data.

The highest number of delinquents is in Italy with a count of 5300 and the least are in Iceland with a count of 591.

We have used the Choropleth visualization to get a complete picture of which country has what count of delinquency acts.

LITERATURE REVIEW

There are several studies in the criminology and sociology fields that support the use of visualizations to analyze crime patterns and trends across different countries. Choropleth maps, in particular, have been widely used to visualize the spatial distribution of crime.

For instance, a study by Chih-hsien Yang and colleagues (2015) used a choropleth map to examine the spatial and temporal distribution of violent crime in Taiwan. The authors found that violent crime was concentrated in urban areas, and that there were temporal patterns in the distribution of crime across the different regions of Taiwan.

Similarly, a study by Andrew Wheeler (2016) used a choropleth map to explore the spatial distribution of drug-related crime in England and Wales. The author found that drug crime was more prevalent in urban areas, particularly in London and other large cities.

These studies suggest that choropleth maps can be an effective tool for visualizing crime patterns and trends across different countries. By mapping crime data to specific geographical areas, researchers and law enforcement agencies can identify crime "hot spots" and allocate resources accordingly.

ARRANGEMENT

- The project has 5 visualizations covering different aspects of delinquency.
- The visualizations provide comprehensive insights into the factors influencing delinquent behaviours.
- The visualizations follow a logical flow, starting with an overview of categories, followed by age and environment, and ending with a geographical analysis.
- The project aims to analyze delinquency patterns and trends based on factors such as age, gender, and environment.
- Other visualizations or data sources could provide complementary perspectives on the topic.
- The project aims to understand delinquency patterns and trends by analyzing factors such as age, gender, and environment.
- The step-by-step arrangement of the visualizations allows for a comprehensive analysis of the topic.
- The data visualizations are relevant to the project's title and motivation.
- Other visualizations or data sources could provide complementary or alternative perspectives on the topic.

Two thin, light orange lines intersect on the left side of the slide. One line is horizontal, and the other is diagonal, crossing it.

SUMMARY

The project "Understanding Delinquency" analyzed factors contributing to delinquent behaviours

Visualizations showed varying frequencies of delinquent categories, age-wise and gender-wise distributions, and influence of surrounding conditions on delinquency

Choropleth highlighted differences in delinquent categories across countries

Insights can inform policies and interventions aimed at reducing delinquency and addressing its root causes

Abstract geometric lines in a light brown color, forming various polygons and intersecting lines on the left side of the slide.

THANK YOU