Analyzing Young Adults' Choice of Behaviour Based on Their Environment

Lakshmi Satvika Nekkanti, Mythri Muchamari,Nilisha Makam Prashantha, Sangamithra Murugesan, Sally Wei

Data 230

Seungjoon Lee

December 9, 2022

Table of Contents

Literature Survey

The literature survey performed by Marripedia, titled “Effects of Family Structure on Crime” formed the inspiration for this project [1]. This article presents empirical evidence suggesting that many young men and women from broken families tend to have a much weaker sense of connection with their neighborhood and are more inclined toward committing a crime. They also claimed that strong parental bonds would significantly decrease the chance of the child committing an act of violence. Then, they attempted to draw parallels between the rise in violent crime and families abandoned by fathers.

The way they conveyed the article's narrative interested us since the data they used was very similar to what we will be working on. As a result, we decided to incorporate a similar story into our project. This study, however, has a few shortcomings.

One of the shortcomings was that they hadn't analyzed the children who lived in other kinds of family structures such as foster care, children with both parents, orphanages, etc. Additionally, they also hadn't tried to study the impact of neighborhoods on young adults and their activities.

Thus, taking these shortcomings as inspiration, the motivation and story for the project were crafted.

Motivation

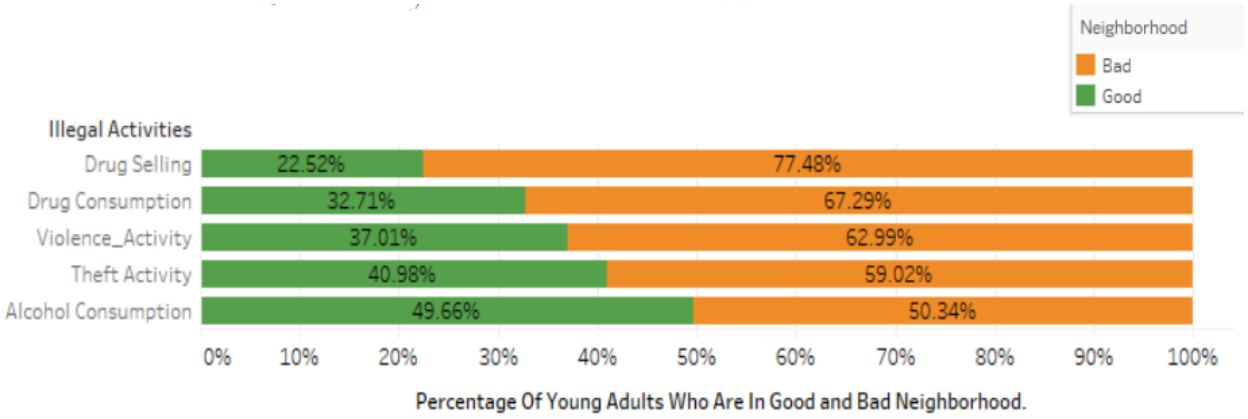
The world shapes us around us. The most crucial time in a person's life is when they are adolescent-aged young adults. It shapes a person's abilities and character to determine their path for the rest of their lives. Based on the literature review we completed and the dataset we were given, Our goal is to visually examine how a young adult's decision to engage in delinquent

behavior is influenced by external factors, including family, neighborhood ambiance, friends, and peers. Additionally, we can examine how these characteristics influence their decision-making in the future.

Visualizations

Visualization 1

Figure 1: How Neighborhood Impacts A Person's Propensity Towards Illegal Activities



Mark: Horizontal stack of lines.

Channel: Size (length) for values, vertical position for the category of illegal activity, **color/hue** for the neighborhood category (good/bad).

Task: Part to whole relationship for each category

Reason for choosing this marks and channels:

2 categorical features and 1 derived quantitative feature.

Stacked bar Charts are best used when showing comparisons between categories. Typically, the bars are proportional to the values they represent and can be plotted either horizontally or vertically.

Instead of just comparing the precise proportion of which answer was more famous for each group, we wanted to highlight the overall distribution of answers. Thus we opted for a normalized stacked bar graph which is the best. The length mark is a horizontal stack of lines, which represents the part of the whole comparison. We have used channels such as color/hue to distinguish between bad and good neighborhoods and length to define the percentage of quantity features.

Motivation:

It is said that, **"we are what we are surrounded by".** Some neighborhoods also provide opportunities for learning and engaging in violence. If a person is surrounded with a bad neighborhood or surrounded by a bad environment , then that person's propensity towards illegal activities like drugs, alcohol, drug selling, violence activities and theft activities will be more. In this visualization, we want to analyze the same.

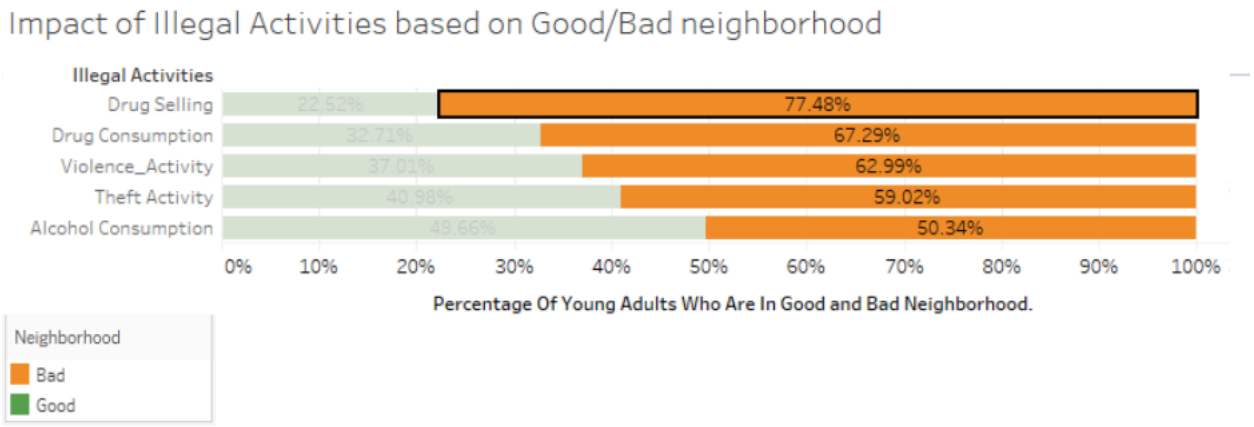
Derived attributes:

Derived attributes involved in this visualization are:

- **'Drug selling' :** Question number 67(drudltp) is about any attempt to sell drugs in the past. Question number 51,52, and 53 (hashltp, xtcltp, and lhcltp), which includes LSD, coke, XTC, speed, hash, and marijuana, have been used. Young adults who answered yes for any types are considered as doing that particular illegal activity.
- **'Violence activity' :** Question number 54 (vandltp), which includes damaging something on purpose. Young adults who answered yes for any type are considered as doing that particular illegal activity.

- ‘**Theft Activity’:** Questions 55, 56, 57, and 58(shopltp, shopltp, bictltp, cartltp) include stealing something from a shop, breaking into buildings on purpose, and stealing scooter, bicycle, stealing motorbike, and car.
- **‘Alcohol consumption’:** Questions 49,50 (beerltp, spirltp) are considered, which includes drinking beer or wine or any of strong spirits. Young adults who answered yes for any type are considered as doing that particular illegal activity.
- **‘Bad neighborhood’:** For bad neighborhoods, question number 47.5(nhood5),47.6(nhood6) 47.7(nhood7), 47.8(nhood8) were considered. Young adults who answer "not true" or "not at all true" for good neighborhoods or "true" or" very true" for **"bad neighborhoods'** are considered as living around **bad neighborhoods.**
- **Good Neighborhood**: Question number 47.2(nhood2), 47.4(nhood4), 47.10(nhood10), 47.12(nhood 12) were considered. So, young adults who answer “true” ,“very true” for a good neighborhood, or “not at all true” or“not true” for a bad neighborhood are considered **“young adults staying in a good neighborhood”.**

Figure 2: Impact of Illegal Activities Based on Good/Bad Neighborhood

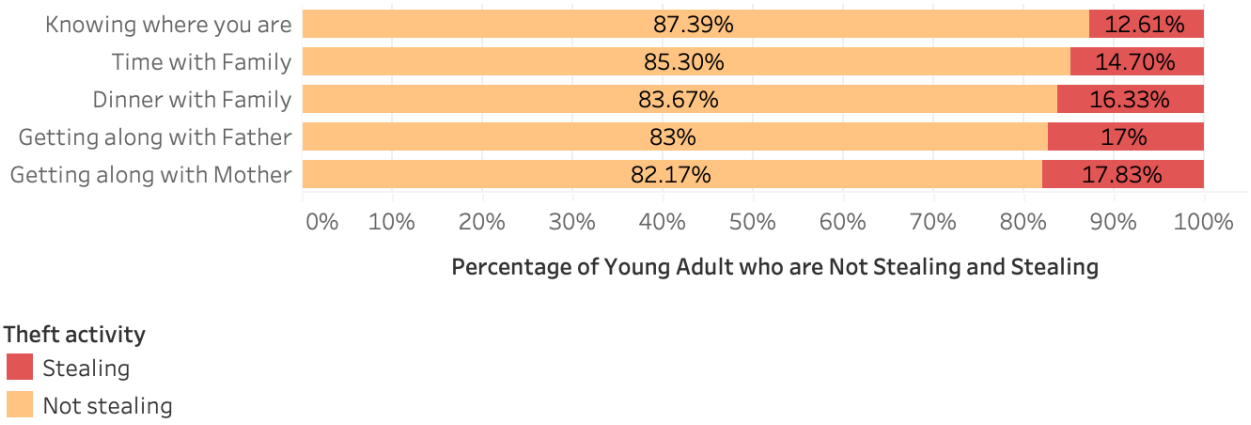


Observations from this visualization

The majority of people who are doing illegal activities are from the "bad neighborhood". So, there must be a significant influence of bad neighborhoods on young adults that drives them to indulge in illegal activities.

Visualization 2

Figure 3: How Family Bonding Impacts A Person's Propensity To Steal



Mark: Horizontal stack of Line.

Channel: Size (length) for values, Y Position (Vertical) for Family bonding category, color (hue) for category of Theft activity (stealing/not stealing), and Label to represent percentage.

Reason for choosing this marks and channels:

The idea that we wanted to visualize has 2 categorical features and 1 derived quantitate feature. Stacked bars with color channel will visually help to see the difference between the categories. Ideally, a normalized stacked bar would be efficient to normalize the data for showing

the difference in percentage, which is what we aim to visualize. It helps in understanding the overall comparison.

Motivation:

The motive of this visualization is to identify how a family bonding will impact a young adult's action toward theft activity. Here we consider the young adults' who have said yes to having a good family bonding. Out of this, we visually observe the percentage difference between the people who have committed theft activity and those who have not committed it. This is performed by implementing the "Total stealing" derived attribute with the color channel.

Derived attributes:

The derived data for this visualizations are as follows

- **'Knowing where you are':** This feature is derived from question 20 of the questioner. The people who have answered "always" are considered as their parents know where they are.
- **'Time with Family':** This feature is derived from question 18 of the questioner. The people who have answered "once/week" or "more than once/week" are considered to spend more time with their family.
- **'Dinner with Family':** This feature is derived from question 19 of the questioner. A young adult who eats dinner more than 5 times per week is considered as having dinner with family.
- **'Getting along with Father':** This feature is derived from question 16 of the questioner. The young adults who say they get along "very well" or "rather well" are considered to get along well with their father.
- **'Getting along with mother':** This feature is derived from question 17 of the questioner.

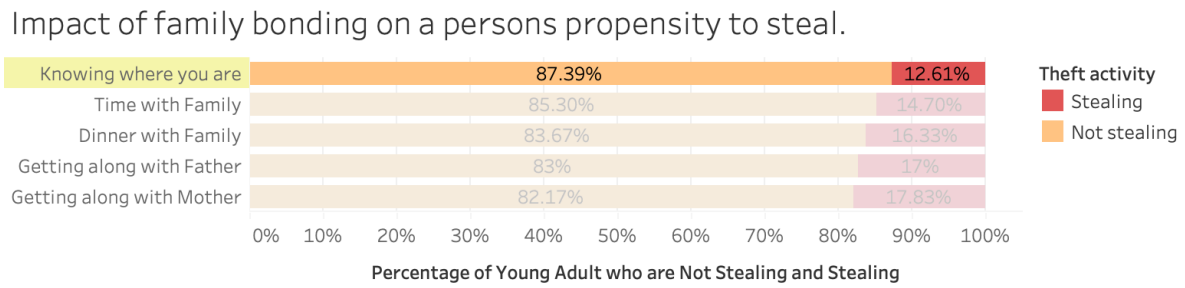
The young adults who say they get along ”very well” or ”rather well” are considered to get along well with their mother.

- **‘Total stealing’:** This feature is derived from a series of theft-related questions like question number 55, 56, 57, and 58 of the questioner. If a young adult has said yes to any one of the mentioned crimes, then he/she is considered to have performed a theft activity.

Observations from the visualization:

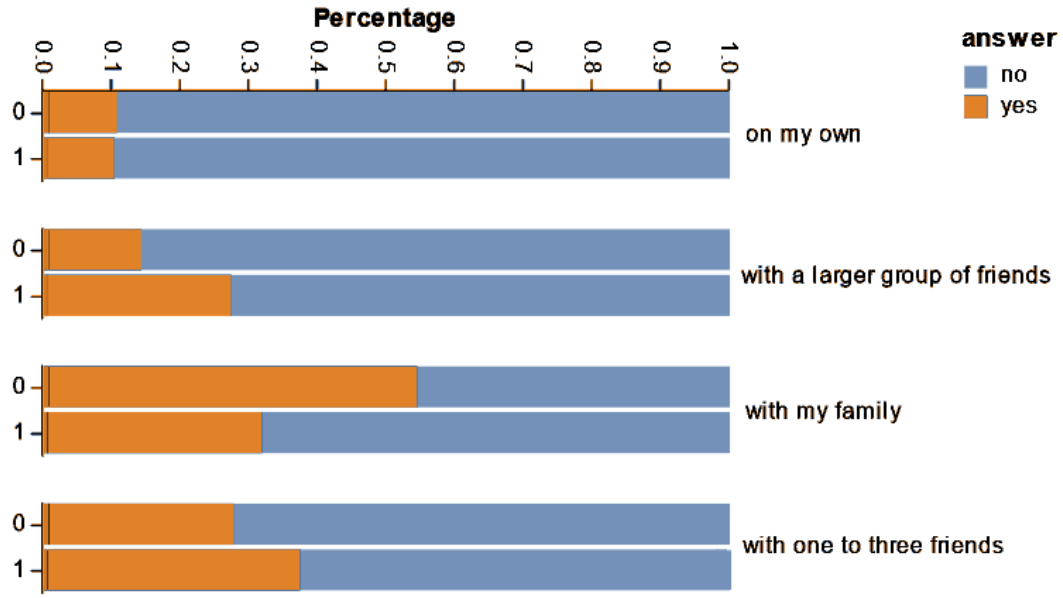
From the visualization, we can conclude that young adults who have good family bonding tend to commit less theft activity. From Figure 4, we can clearly see that 87.39% of young adults who have said that their family is aware of where they are, perform very less theft activity. Hence, positive family bonding reduces delinquency among young adults.

Figure 4: **Impact of Family Bonding on a Person’s Propensity to Steal**



Visualization 3

Figure 5: How Time Spent With Family Versus With Friends Affects Delinquent Behavior



Mark: Line

Channel: Size (length) for proportions; vertical position for categories; color (hue) for categories

Reason for choosing this marks and channels:

We wanted to show the overall spread of answers rather than just compare the exact percentage of which answer was more prevalent for each group, so we chose a normalized stacked bar graph over a regular bar graph. Using a bar graph instead of a series of donut charts or pie charts makes them easier to compare to each other, as well.

Since there are two categories to split the data by–the question and history of criminal activity but we wanted to prioritize ease of comparison across groups, we chose to use vertical position and the gestalt principles of perception to visually group the categories together.

Motivation:

We wanted to look at how the people surrounding an adolescent might affect their propensity for crime, and chose to look at whether they spent most of their time with family or with friends. More time spent with friends and less time spent with family could relate to the previous visualization, where we found that family bonding could relate to a reduced propensity for crime.

Derived attributes:

For this visualization, we needed to show the distributions of yes/no answers to each question (on my own/with a larger group of friends/with my family/with one to three friends), divided into groups based on the participant's history of criminal activity (0 for none, 1 for any).

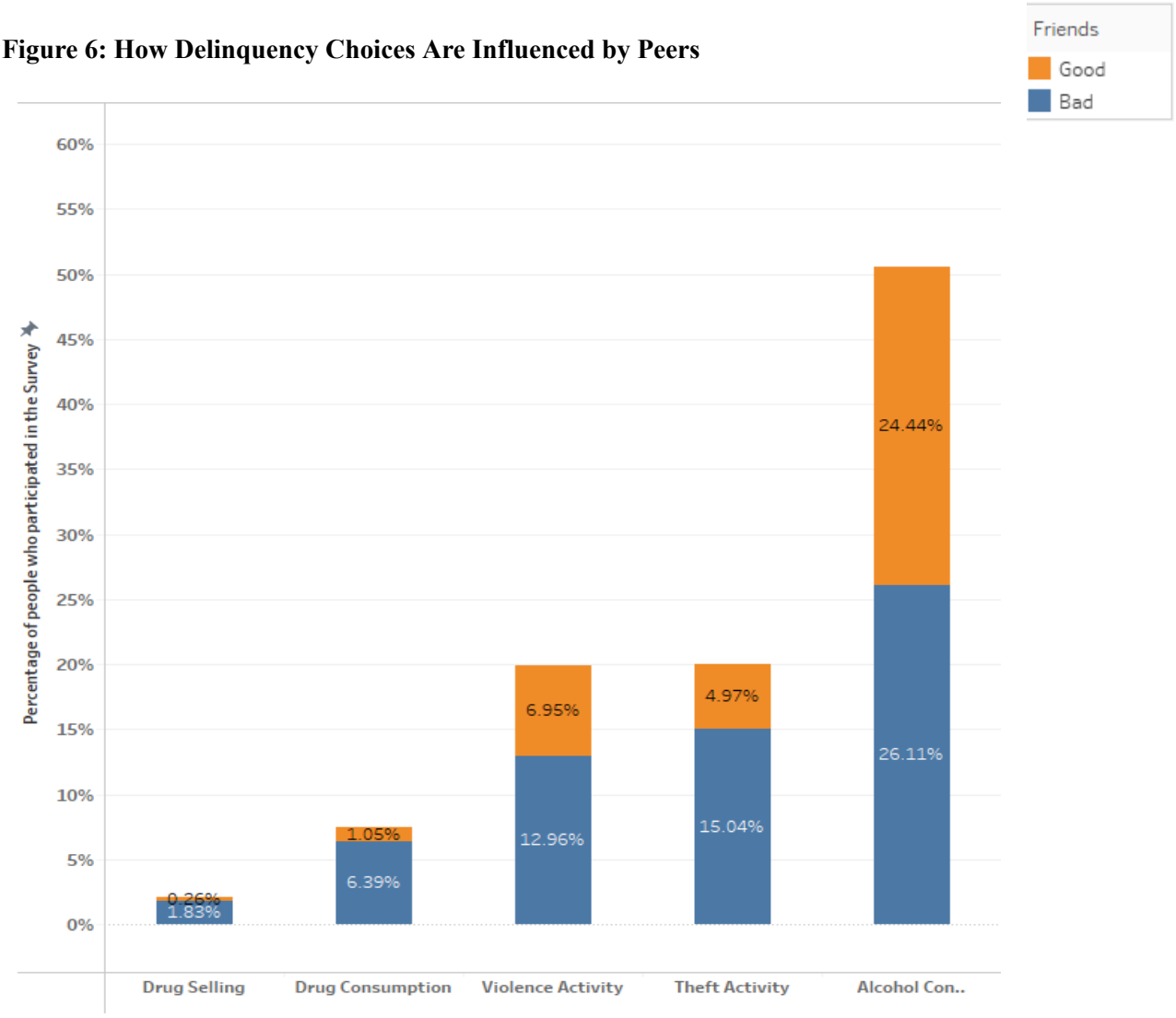
Criminal history (0 - no criminal history, 1 - has criminal history) where having a criminal history means having a positive response to committing one or any of the following in the past: alcohol consumption (Q49, 50), theft (Q55-58), violent activity (54), drug use (Q51-53), or drug selling (Q67).

Observations from this visualization:

We observed that a greater percentage of those with criminal history reported spending most of their time with friends, either in a small group (one to three friends) or in a larger group. On the other hand, a greater percentage of those who did not have any criminal history reported spending most of their time with family. An equal percentage in both groups (with and without criminal history) reported spending their time alone. We can conclude that time spent with peers could possibly be related to encouraging criminal activity in adolescents while spending more time with family might act as a deterrent.

Visualization 4

Figure 6: How Delinquency Choices Are Influenced by Peers



Mark: Vertical stack of lines.

Channel: In the above figure, we have two categorical and one quantitative features

Size(length) for quantitative feature (Percentage of young adults committing illegal activities).

Position (horizontal)for 1 category which is illegal activities (aligned vertically).

Position (Vertical) for 1 category of good/bad influence of friends (unaligned vertically).

Color(hue) to differentiate between good influence of friends and bad influence of friends.

Task: Part to whole relationship for each category.

Reason for choosing marks and channels:

In the above figure there are two categorical and one quantitative features.

Bar charts are usually known to be good at visualizing the size difference between two values.For the above problem a stacked bar chart was used to differentiate between bad influence of friends and good influence of friends for every single category. To use a normal bar graph would have made it look too complicated and many in number so a stacked bar chart was chosen to show the quantities of five different illegal activities. Orange and blue colors were used (Color/hue) where blue represents the Bad influence of friends and orange color indicates the good influence of friends .They were stacked one on top of other to easily understand the number of good and bad influences.

Motivation:

The main message is what kind of circle a person indulging himself in illegal activities such as drugs, alcohol, and stealing prefers to hang out with. Here, five illegal activities were considered on the x-axis. On the y-axis, the percentage of people who participated in the survey was considered. A person who is involved in drug selling tends to have about 1.83% bad influence and 0.26% of good influence around them. It is clear that someone involved in drug selling has the least good influence around them. While talking about someone who is addicted to alcohol is surrounded by 26.11 % of bad friends and 24.44% of good friends. It is clear that out of all the illegal activities, someone who indulges themselves in alcohol has the highest percentage of bad people around them. When we compare the rest of the illegal activities like Drug consumption, Violence, and Theft activities, they seem to hang out with more than fifty percent of a bad influence.

Derived attributes:

Derived attributes involved in this visualization are

All the illegal activities in this visualization i.e.,

- **'Drug selling' :** Question number 67(drudltp) is taken and has records of any past acts of drug selling.
- **'Drug consumption':** Question number 51,52 and 53 (hashltp, xtcltp and lhcltp) were considered as they involve consumption of LSD,coke, XTC, speed, hash ,marijuana and so on. Young adults, who answered yes for any of the types, are considered as doing that particular illegal activity were considered to be 1 and the ones who did not commit to it were considered as 0(zero)
- **'Violence activity':** Question number 54 (vandltp) involves vandalism. Damaging someone else's property on purpose, and so on. People who did the above-mentioned activities were confirmed to be guilty and put in the category "1". The rest of them are considered to be zero.
- **'Theft Activity':** Question number 55, 56, 57, and 58(shopltp, shopltp, bictltp, cartltp) include stealing, shoplifting from a store, breaking into buildings on purpose, stealing motor vehicles such as cars, scooters, and so on. If the young adults committed at least one of these activities, then they were considered to be 1, and if they were involved in neither of those, they are considered to be zero.
- **'Alcohol consumption':**Questions 49,50 (beerltp, spirltp) are considered, which includes drinking beer or wine or any of strong spirits. Young adults who commit any of the above activities are considered involved in an illegal activity.
- **For categorizing good influence of friends and bad influence of friends** Question 48 (delpdr.delpsl,delpbhu,delpex,delpas) were considered, which include friends who

consumed weed, has and so on., friends who were a part of shoplifting, friends who purposefully break into other houses, friends who threaten others with weapons, knives, Friends who hurt others with objects like sticks and knives.

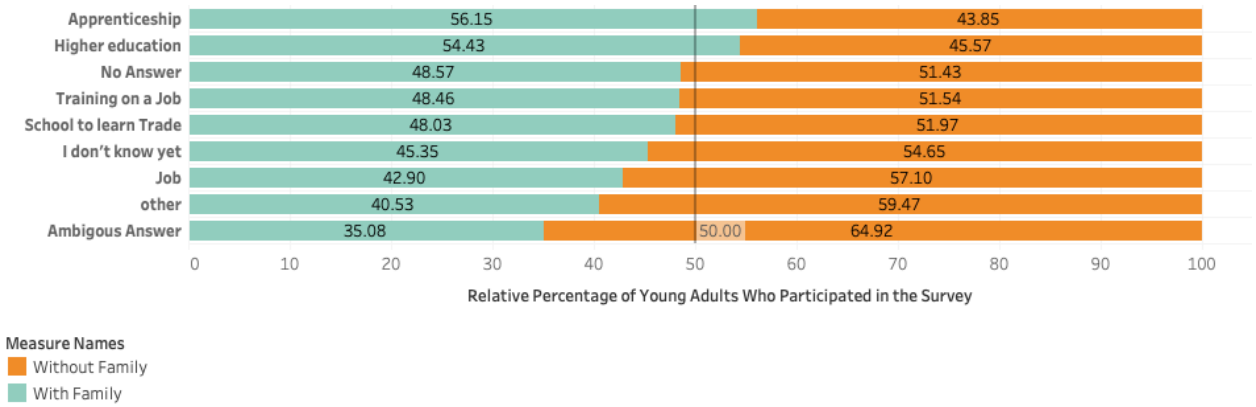
- **Bad Influence of Friends:** If a friend has previously committed at least one of the above-mentioned crimes, they are considered as a bad influence of a friend. i.e '1'(Yes)
- **Good Influence of Friends** :A friend who has not taken part in any of the activities, he/she is considered a good influence. If all of them are '0'(No), then they are good friends.

Observation from this visualization:

To conclude, it is very clear after seeing young adults who are involved in these illegal activities are more likely to hang out with people similar to them, i.e., with a bad influence. So it is very important to choose a good friend circle as their influence on your life is very impactful.

Visualization 5

Figure 7: How Family Situation Affects the Future Choices After School



Mark: Horizontal stack of Line.

Channel: Categorical type of data in X position, Percentage (numerical) of people who participated in the survey in the Y position. Size, and color (hue) were some redundant channels used to represent categorical variables and finally the labels were another redundant encoding used to represent the percentage of each category.

Reason for choosing this marks and channels:

To understand and evaluate three or more categorical variables, the best method to visualize them was a bar chart. Bar charts help understand the parts of the whole relationship in an efficient manner. Stacked bars, on the one hand, provide good visualizations, but on the other, the disadvantage is that if all the stacks don't start from a common point, it is very difficult to compare between categories. Hence, a normalized stacked bar is the most efficient method to represent this problem.

Color (hue) and label are used as an additional channel that enhances a user's sense of understanding.

Motivation:

The previous visualizations described how a young adult's delinquent behavior might have been impacted by their families, neighborhood, friends, or leisure time. Now, using this visualization, we want to find answers for how a young adult's future choice of decision after school is impacted by the family structure they come from.

Derived attributes :

To achieve the motivation for the graph, we first categorized the different families into 'with family' and 'without family' derived fields. These fields are formed by grouping all the responses from question 46. The count of responses with option 1 is said to be living with family, and the count of responses with options > =7 and < = 9 are taken as living with other families.

Now, both these columns will be derived into ratios which can be achieved by the lambda function (where each row value is multiplied by 100 and divided by the total of that column). In this case, using the below code, derived columns were achieved.

```
df_joined['WITH_FAMILY_RATIO'] = df_joined.apply(lambda row:
row.WITH_FAMILY_PCT*100/(row.WITH_FAMILY_PCT+row.WITHOUT_FAMILY_PCT),
axis=1)
df_joined['WITHOUT_FAMILY_RATIO'] = df_joined.apply(lambda row:
row.WITHOUT_FAMILY_PCT*100/(row.WITH_FAMILY_PCT+row.WITHOUT_FAMILY_P
CT), axis=1)
```

Observations from this visualization:

- We can see about 54.43% of adults who choose higher education come from their own families and are comparatively higher than other family adults who choose the same. This can be because the adults might have strong support financially or in any way from their parents.
- We can also observe that adults from the Foster family choose choices like "school to learn trade", "job" is comparatively 1 - 2% more than adults from their own family. This might be because being financially independent would be the first priority any person coming from this background would consider. Since people with their own parents might not have that as a compulsion which is why they think about upskilling themselves.
- Another interesting insight we can see is that 16.01% of adults who have chosen the “don’t know” option come from a foster family; the numbers say that they have a higher percentage. This might be because they might not have got proper guidance from elders to understand which option might be better for them.

Thus, the conclusion from this visualization is that young adults who are coming from foster families should be guided better; there might be many options, like career counseling in schools which can be taken advantage of. If the students are well deserved, they should be given the opportunities with a scholarship to pursue their goals.

Arrangement

We arranged our visualizations from more general topics to more specific topics, starting with the general impact of the neighborhood on a person's propensity for crime in Figure 1. The next visualization continues with looking at the neighborhood environment and breaks down the illegal activities in good and bad neighborhoods. Figures 3 and 4 explore how the family situation affects crime, with Figure 3 looking at family structure and Figure 4 looking at the level of family bonds. Figure 5 and Figure 6 are about how friends and peers affect the propensity for crime. While Figure 6 focuses on how peer influence affects the types of criminal activity undertaken. Finally, Figure 7 looks into long-term decision-making through the lens of family circumstances.

Summary

In conclusion, there is a definite connection between an individual's environment and those engaged in illicit actions. Young adults who are exposed to harmful influences, live in undesirable neighborhoods, and don't spend enough time with their families are more likely to engage in illegal activity. In addition, having a dysfunctional family situation may influence their decision-making in the future. It also affects a young adult's decision to make future choices. So,it is very important to choose your surroundings wisely.

References

[1] “Effects of Family Structure on Crime” Marripedia. Retrieved Dec 2022. URL:

https://marripedia.org/effects_of_family_structure_on_crime