Social media marketing of alcohol in the night time economy (NTE) and its relationship with assault related injuries

Aayush Negi

4th November 2022

School of Mathematics,

Cardiff University

A dissertation submitted in partial fulfilment of the

requirements for **Data Science and Analytics**

by taught programme, supervised by **James White**.

CONTENTS

List of Tables:

List of Figures:

EXECUTIVE SUMMARY

Social media sites are gaining traction in the nightclub industry as a marketing tool, and any alcohol-related content posted on these sites has the potential to reach a large proportion of adolescents and young adults. Many studies have linked social media alcohol advertising to alcohol consumption. However, it is unclear whether the presence of social media alcohol marketing influences assault-related injuries. This study examines the role of social media marketing of alcohol in the night time economy (NTE) and its relationship with assault related injuries in Cardiff during 2019 and to compare Poisson, Negative Binomial, and Zero Inflated Poisson models on this data, to determine the preferred count regression model.

Twitter was used to gather counts of all tweets and counts of tweets promoting alcohol from five nightclubs in Cardiff. Assault related injuries were the number of patients with assault related injuries coming into the University Hospital Wales, Cardiff. Poisson, Negative Binomial, and Zero Inflated Poisson were used to analyse all tweets posted by nightclubs, alcohol promotional posts and weekends on the prevalence of assault related injuries. The performance of the regression models was evaluated by using log-likelihood comparison.

The study showed that tweets posted by nightclubs and weekends have a positive association with assault related injuries, while alcohol promotions were shown to be negatively associated with assault related injuries. Weekend and tweets by nightclubs are statistically significant as $p<0.05$ and alcohol promotion is statistically not significant $p>0.05$ in affecting the number of assault injury attendances. A nightclub posts 5 tweets per day on average. A total of 1,872 tweets and 686 assault-related injuries from the 2019/2020 school year were examined. 43% of all tweets promote alcohol. The negative binomial model provided the best fit to the data having the highest Log-Likelihood value (-643). Males account for 77% of the 686 assault-related injuries, while females account for 22% hence the dependent variable was split into male assault data and female assault data, and NB2 was used to analyse the association of male assault data and female assault data with all tweets posted by nightclubs, alcohol promotional posts and weekends. The analysis of results shows that the variables, tweets by nightclubs and weekend have a positive effect on both male and female assault related injuries in Cardiff, whereas alcohol promotions have a negative effect on both. Weekend and tweets by nightclubs are statistically significant as $p<0.05$ in affecting the number of male assault injury attendances.

Nightclubs use a variety of social networking sites and techniques to advertise alcohol and boost customer traffic. These could undercut efforts to alter drinking-related social norms, particularly the normalisation of daily consumption. Nightclubs use images, stories and personalised ads to market alcohol and increase customer traffic however little research has been done into these methods for marketing. Future research should consider all of these methods as well as the various platforms used by young people, such as Facebook, Instagram, Snapchat, and Tiktok.

1 INTRODUCTION

Alcohol is frequently in the spotlight due to social concerns, particularly issues like binge drinking and underage drinking (Advertising Standards Authority | Committee of Advertising Practice, 2019). Due to the political significance of these concerns, elements that may influence how much alcohol we consume have been thoroughly investigated, and as a result, the UK has some of the strongest alcohol advertising regulations in the world (Advertising Standards Authority | Committee of Advertising Practice, 2019). According to UK regulations governing the marketing of alcoholic beverages, promotional marketing materials cannot indicate, support, or encourage excessive alcohol consumption. Alcohol should not be associated with actions or settings where drinking is dangerous or inappropriate. Alcohol cannot be portrayed as having the power to alter one's mood, physical health, or behaviour. Any character who is shown drinking or playing an important role must be at least 25 years old (Advertising Standards Authority | Committee of Advertising Practice, 2019a). However, according to a study of university students in Canada (Paradis et al., 2020), the majority of well-known bars post alcohol-related content on Facebook and Instagram in violation of the rules governing the promotion of alcoholic beverages, and there is a strong correlation between students' preferences for drinking establishments and these establishments' tendency to publish pictures on social media that are against the advertising laws for alcoholic beverages (Paradis et al., 2020).

There is growing evidence that alcohol marketing makes it more likely that young people start drinking and raise how much they consume overall and on any given occasion (Stautz et al., 2016). Adolescents are more likely to start drinking alcohol due to alcohol advertising and promotion, and if they already do, it raises the possibility that they will drink more (Anderson et al., 2009). The alcohol business has modified its advertising strategies in recent years and now places a significant amount of its money and resources on digital and internet media (Federal Trade Commission. Self-regulation in the Alcohol Industry, 2014). The primary goal of social media marketing for alcohol promotional activities is to promote alcoholic products among social network users and their networks. Any user-generated content, such as a tweet, event comment, status update, wall post, group message, or video, that refers to or discusses an alcoholic beverage amplifies messages about the beverage with other users in the user's social network. Talking about the alcohol product and expressing support for the brand fosters a friendly environment among users (Mart, 2011). Although promotion is the primary social media activity of nightclubs, in certain cases non-promotional activities such as uploading photos, handling booking inquiries, and conducting unofficial rival and customer research are also part of nightclubs' social media communications (Nevin et al., 2012). Higher rates of binge drinking are related to the availability of large quantities of alcohol, cheap sale prices, and frequent promotions and advertisements (Wickham, 2012). Exposure to alcohol-related advertising is regularly linked to the onset of adolescent drinking and the consequences of alcohol (Atkinson et al., 2020). Social media frequently features alcohol-related information that can both reflect and affect offline drinking habits. Posting alcohol-related information online has been connected to greater rates of alcohol intake, cravings, alcohol-related issues, and clinical alcohol use disorders (Westgate and Holliday, 2016).

Violence in the nightlife is frequently linked to alcohol (van Amsterdam et al., 2019, Pedersen, Copes and Sandberg, 2016). Alcohol use by perpetrators was found to be present in 47% of all violent occurrences and 58% of stranger incidents, according to the 2001/02 British Crime Survey (Allen et al., 2003) and people who frequent bars and pubs are far more likely than non-users to become victims of such crimes (Povey and Allen, 2003). Nearly half (44%) of all assaults are alcohol-related (equivalent to almost one million annually) and one in five takes place near a bar, pub or other licensed drinking venue (Chaplin, Flatley and Smith, 2011). In late-night entertainment districts, incidents frequently occur in small, well-defined groups and are occasionally connected to specific locations, as the venues close violence peaks (Finney, 2004).

Interpersonal violence poses a serious hazard to health on a global scale (Krug et al., 2002), it is the leading cause of death among teenagers and young adults (World Health Organization, 2021), and exposure to interpersonal violence makes people more susceptible throughout their lives to emotional, behavioural, and physical health issues (Mercy et al.,2017). Health services are burdened with high expenditures associated with treating the effects of violence on physical and mental health (Butchart, 2008). According to Crime Survey for England and Wales (CSEW) which measures the general trend in violent crime, there were 1.5 million violent offences in the year ending March 2022 (Office for National Statistics, 2022) according to an English survey 67% of assaults occur at night (Chaplin, Flatley and Smith, 2011). Alcohol has been associated with violent or aggressive behaviour (National Institute on Alcohol Abuse and Alcoholism, 2015, Caces, Stinson & Harford 1991).

Many studies have been done to study the effect of social media on alcohol consumption (Brunborg, Skogen and Burdzovic Andreas, 2022, Ng Fat, Cable and Kelly, 2021); however, little research has been done between the effects of promotional activity of night time economy on social media platforms and its relation to the assaults. This project aims to study the relationship of promotional activity of night time economy and its relationship to assault related injury in Cardiff.

2 LITERATURE REVIEW

2.1 Factors affecting assault related violence

This section includes relevant material on topics which affect assault related violence including the night time economy (NTE), alcohol consumption, violence and the impact of social media on alcohol use.

2.1.1 The Night Time Economy

The NTE has no established definition (Wickham, 2012). For this study, it may be characterised as commercial activity that takes place between 6 p.m. and 6 a.m. involving the purchase of alcoholic beverages for consumption in pubs, clubs, bars, and restaurants (Wickham, 2012). Cardiff, the capital of Wales, has seen its population increase exponentially during the past ten years (Hardwick, 2018). Foot traffic in the City Centre has surged by 61% since 2009, and it currently attracts over 300 events annually, making it a well-known worldwide destination (Hardwick, 2018). Cardiff City Centre draws up to 40,000 people on weekends and up to 120,000 on days when there are significant events. The expansion of the late-night industry in Cardiff City Centre places a significant strain on police resources (Hardwick, 2018).

Alcohol use/consumption can be advantageous in the NTE in several ways. For instance, it creates economic activity and employment, it may facilitate social interaction, and it is a rewarding hobby that many people cherish. However, there may be drawbacks involved (Wickham, 2012). Public disturbance, vandalism, drunk driving, major attacks (including killings), and sexual assaults are some of the issues brought on by the night-time economy. These issues are at their worst on weekends (Tomsen, 2005).

In 53% of violent occurrences recorded by the 2013/14 Crime Survey of England and Wales, victims believed the offender(s) to be under the influence of alcohol. This equates to over 704,000 violent "alcohol-related" events (Office for National Statistics, 2015b). Crime related to alcohol may have a significant negative impact on the community's overall sense of safety, even those who have never personally experienced the negative effects of alcohol or engaged in antisocial behaviour are concerned and see it the same way (Nicholas 2006).

2.1.2 Alcohol consumption and violence

Alcohol affects sensory perception and motor control, which produces behavioural effects including clumsiness and sluggish reaction times. It interferes with cognitive and emotional functions, leading to issues including imperfect perceptual information encoding, poor problem-solving, and decreased fear of antisocial behaviour. These modifications raise the probability of aggressiveness in response to unpleasant stimuli (Giancola et al, 2003).

Even though alcohol-related violence has always been a problem in urban societies, there has been a significant increase in the number, capacity, and popularity of entertainment district bars and nightclubs in the UK over time as a result of the significant investment made by the alcohol and entertainment industries (Warburton and Shepherd, 2006).

The National Violence Surveillance Network (NVSN), which is now in its 21st year, consists of 158 emergency departments (EDs), minor injury units (MIUs), and walk-in clinics across England and Wales. In 2020, there were predicted to be 119,111 victims of violence (81,453 men and 37,658 women), down from 175,764 in 2019 (122,134 men and 53,630 women) (Sivarajasingam et al., 2022). Regular pub and bar patrons are far more likely to be victims of violent crime than infrequent users (Povey and Allen, 2003). This substantial decline in violence was attributed to COVID-19-related countrywide lockdowns and social gathering prohibitions inside and outside of residences, bars, nightclubs, and restaurants in 2020 (Sivarajasingam et al., 2022).

While frequent drinking has also been linked to an increased risk of violence in other studies on teenage populations, the clearest and most well-established relationship between alcohol use and violent behaviour is the use of significant amounts of alcohol on a single occasion (Caces, Stinson & Harford 1991). The correlation between excessive drinking and conflicts after drinking is significantly complicated by drinking frequency and amount. It was shown that drinking in public places other than your house was strongly connected with a higher risk of arguments among women the next day. The relationship between drinking frequency and alcohol-related aggressiveness in men was altered by the usual drinking location, with frequent drinkers in public places away from home having the highest risk of hostility (Wells et al., 2005).
According to a study done by Shepherd and Brickley, when people consume more than 8 or 10 units of alcohol in one sitting, both the danger of being injured in an attack and the chance of participating in one rise dramatically (Shepherd and Brickley, 1996). Public violence and disorder levels are correlated with the concentration of bars and clubs in a region, with more establishments being linked to higher levels of public disorder and violence (Campbell et al., 2009). According to a study in UK, some locations are riskier than others for experiencing violence. For instance, just a few licenced establishments were linked to more than half of the violent occurrences that occurred in or around pubs and nightclubs in Southampton city centre in 1980 (Ramsay, 1982).
Because a small number of outlets frequently account for most injuries sustained in licenced premises (Homel and Clark, 2004), researching popular venues in an area can provide insight into factors influencing violence in that area. The level of violence varies by drinking establishment (both inside and outside). This difference has been linked to management style. When excessive and/or rapid drinking is encouraged, such as during "happy hours," there is an increased risk of intoxication (Homel and Clark, 2004).

2.1.3 Social Media and Alcohol Consumption:

Global social media users have increased by more than 10% in the last year, with 424 million new users joining the platform in 2021 (Kemp, 2022) creating a new environment where young people might be exposed to information relating to alcohol. Social media sites are diverse, but many of their features are similar. Users typically create an account, connect to a network of other individuals or groups, and use the site to share ideas, photographs, videos, news stories, and other content (Kietzmann et al., 2011). Businesses use social media to promote their products and services. Most websites include mechanisms for expressing approval or disapproval of content; as a result, users can form their own opinion of a post or video while also seeing how many others have expressed approval. The multidirectional and user-generated content communication of social media distinguishes it from traditional mass media (Kaplan and Haenlein, 2010). Young adults use social networking sites like Facebook extensively for everyday friendship and socialising. Businesses and brands have used Facebook strategically to embed their alcohol marketing into the social networking friendship activities of young adults, blurring the lines between user and alcohol brand-generated content (Niland et al., 2016). Youngsters are exposed to words and images that promote alcohol consumption through online representations of drinking on personal accounts as well as unrestricted alcohol promotion on social media platforms. Such online demonstrations of alcohol use have been linked to dangerous drinking and offline alcohol use (Moreno and Whitehill, 2014).

The purpose of social media marketing for bars and clubs is to build relationships with patrons on these social networks to draw attention to their goods and services. Although occasionally users may proactively seek a nightclub's page on social media, most of the time users will click on links in comments made by nightclubs that are shown in a stream on their homepage (Nevin et al., 2012). To draw in more consumers, nightclubs use price-based promotions that provide inexpensive and discounted beverages (Nevin et al., 2012). Advertising is the primary social media activity of nightclubs (Nevin et al., 2012). The multi-platform marketing methods used by the alcohol business now heavily they rely on social media platforms. This might lead to the development of toxic online environments where people are encouraged to consume alcohol to excess learn about alcohol (Atkinson et al., 2020). Using social media as a marketing tool is gaining traction in the nightclub industry (Nevin et al., 2012). The amount of nightclub advertising activity on social media platforms, and Facebook in particular, seems to be quite high (Miletsky, 2010, Nevin et al., 2012). All the nightclubs studied by Nevin et al., (2012) had active Facebook accounts and considered Facebook to be the most popular and effective social media marketing website. Participants in the study were also active Facebook users, with the majority describing Facebook as the hub of their social media activity.

According to one study, alcohol brands and clubs promote drinking on Facebook by using contextual associations, offering good deals, and fostering positive attitudes toward alcohol promotions. Alcohol brands and clubs also encourage young consumers to participate in Facebook promotions and to co-create social media conversations (Moraes, Michaelidou and Meneses, 2014).

A nightclub's main social media activity is advertising and promoting events (Nevin et al., 2012). A combination of overt and covert sales tactics are used by clubs, such as alcohol advertising, special deals, speed drinking gadgets, bartenders' strategic intimacy, flirtation, and encouragement to buy more, as well as architectural elements that discourage moderate drinking while hastening the purchase and consumption of alcohol (Tutenges and Bøhling, 2019). Many drinking establishments, especially those that cater to youthful clientele, actively participate in encouraging and guiding behaviour so that they may be seen as staged environments for consumption where people are enticed to excessively consume alcohol (Tutenges and Bøhling, 2019).

2.2 Data type and Models used for Data Analysis

This section explains the characteristics of data and the statistical models used to evaluate the relationship between assault related injuries and twitter data.

2.2.1 Count Variable

A count variable is a parameter with discrete values (0, 1, 2, ...) that represents the number of times an event has occurred during a predetermined amount of time. An event can only occur once, hence a count variable can only take on positive integer values or zero (Coxe, West and Aiken, 2009).

A data collection of counts contains the qualities listed below (Date,2021):

- *Whole number data:* The data points are non-negative integers.
- *Skewed Distribution:* The frequency distribution may be highly skewed since the data may have many data points for a small number of values.
- *Sparsity:* The data may be scant because it represents a reflection of a rare event.
- *Rate of occurrence:* To build a model, it may be assumed that events occur at a given rate, which is what causes this data to be generated. Over time, the incident rate might fluctuate.

2.2.2 Poisson regression

Regular, recurrent health-related events like the number of hospital stays, pregnancies, or doctor visits are of particular interest to health researchers. These discrete events are reported using non-negative integer or count data. For the study of this data, techniques like Poisson regression and negative binomial regression would be more effective due to the infrequent number of events leading to a non-normally distributed and positively skewed dataset (Hutchinson and Holtman, 2005).The mean for a Poisson distribution is denoted by λ. The value of λ is affected by a vector of explanatory factors, sometimes referred to as predictors, regression variables, or regressors. The Poisson Regression model's task is to fit the observed counts Y to the regression matrix X using a link-function that represents the rate vector as a function of the regression coefficients β and X (the regression matrix), respectively. (Date, 2021).

Poisson regression's assumptions are as follows*:*

- Counts are Y-values.
- Counts must be positive integers, or whole numbers, with a value of 0 or above (0, 1, 2, 3, etc.). Because the Poisson distribution is discrete, the method will not work with fractions or negative values.
- There must be a Poisson distribution for counts. As a result, the mean and variance ought should be equal.
- There must be continuous, dichotomous, or ordinal explanatory variables. Independent observations are required. (Stephanie, 2016).

Some limitations of Poisson Model are:

- Overdispersion is a phenomenon that occurs when the conditional variance is bigger than the conditional mean, and it can cause Poisson regression to underperform in certain circumstances (Hilbe, 2011).
- When the number of zeros in a model is significantly large, the data does not fit a Poisson distribution well (Weaver et al., 2015).

Problems with excess zeros and overdispersion can be addressed by two typical extensions of Poisson regression.

1. Negative Binomial Model
2. Zero-Inflated Model

2.2.3 Negative Binomial regression

The parameter of a Poisson distribution is λ, which is also the mean and variance of the distribution. A count distribution will often have a variance that is greater than its mean. When this occurs with data that we presume (or hope) is Poisson distributed, depending on whether the variance is lower or bigger than the mean, we say we have under- or overdispersion. Negative binomial regression is one strategy for dealing with this problem. The Poisson distribution and the negative binomial distribution both represent the likelihoods of occurring for whole numbers larger than or equal to 0. The variance is not equal to the mean, unlike the Poisson distribution. This shows that it could work well as an approximation for modelling counts with variability that differs from the mean. If our count is, for example, random variable Y from a distribution with a negative binomial, then Y's variance is var(Y)=μ+μ^2/k (Ford, 2016).

The Generalized Linear Model class provided by the Python **statsmodels**[1] module is used to implement the NB2 model. We will use the NB2 model instead of NB1 as the NB2 model is slightly more computationally efficient and has a higher log likelihood association than NB1 (Cameron and Trivedi, 1986).

For NB2 model: variance = mean * α mean^2.

According to Cameron and Trivedi (2013), utilising a method they call auxiliary OLS regression without a constant, they came up with a creative way to compute α as shown in Fig 2 (Date, 2021). For negative binomial model X and Y remain the same however λ = the event rate vector. One of the main characteristics of count-based data sets is the vector λ. The n rates [λ_0, λ_1, λ _2, ..., λ _n] it contains values correspond to the n observed counts in the counts vector y. It is presumable that the rate λ _i for observation 'i' drives the actual observed count y_i in the counts vector y as shown in Fig:1. There is no column named "λ" in the provided data. Instead, during the training phase of the regression model, the inferred variable λ vector is computed (Date, 2021).

We adjusted the Poisson regression model to our data set to determine λ_i. This provides us with the entire rate vector = λ = [λ_1, λ_2, λ_3, …, λ_n] corresponding to all n observations in the data set (Date, 2021a).

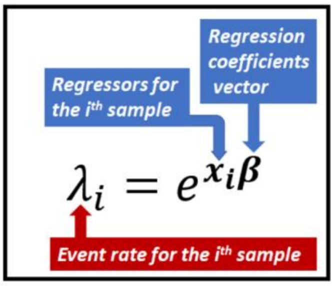
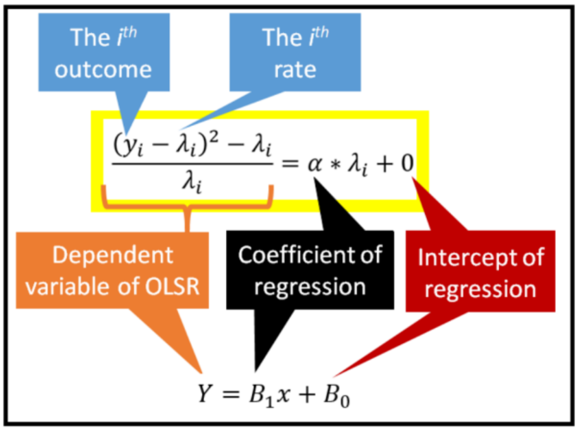


Figure 1:Exponential link function of the Poisson regression model (Date, 2021)

[1] https://www.statsmodels.org/stable/index.html



[2]

Figure 2:Auxiliary OLS regression to find α for the NB2 model (Date, 2021a)

2.2.4 Zero-Inflated Model

Most uses of count data in health economics and other social sciences suffer from the alleged "*extra zeros problem*" which means that the proportion of observations in the sample with zero counts is significantly higher when compared to common count models (Staub and Winkelmann, 2012). According to (Lambert, 1992) "*Zero-inflated Poisson (ZIP) regression is a model used for count data with excess zeros. It assumes that with probability p he only possible observation is 0, and with probability 1 - p, a Poisson(λ) random variable is observed*''. For instance, defects could be almost hard to produce when production equipment is properly aligned. The Poisson(λ) distribution, however, may apply when it is misaligned and defects appear. The mean number of defects λ in the imperfect state and the likelihood of the ideal, zero-defect condition may both be affected by covariates (Lambert, 1992). The Zero Inflated Poisson model is based on the idea that a secondary process determines whether a count is zero or non-zero. When a count is found to be non-zero, the standard Poisson process takes over and calculates the actual non-zero value of the count using the PMF of the Poisson process (Date, 2021). According to (Date, 2021) ZIP regression model consists of three parts as shown in Fig 3.

'' *ϕ_i is calculated by estimating it as a function of regression variables X. This is usually done by transforming the y variable to a binary 0/1 random variable y' (y_prime) which takes the value 0 if the underlying y is 0, and 1 in all other cases. Then we fit a Logistic regression model on the transformed y'. We then train the Logistic regression model on the data set [X, y'] and it yields a vector of fitted probabilities μ_fitted= [μ_1, μ_2, μ_3,…,μ_n], (because that's what a Logistic regression model does).*'' (Date, 2021c) (Fig 4)

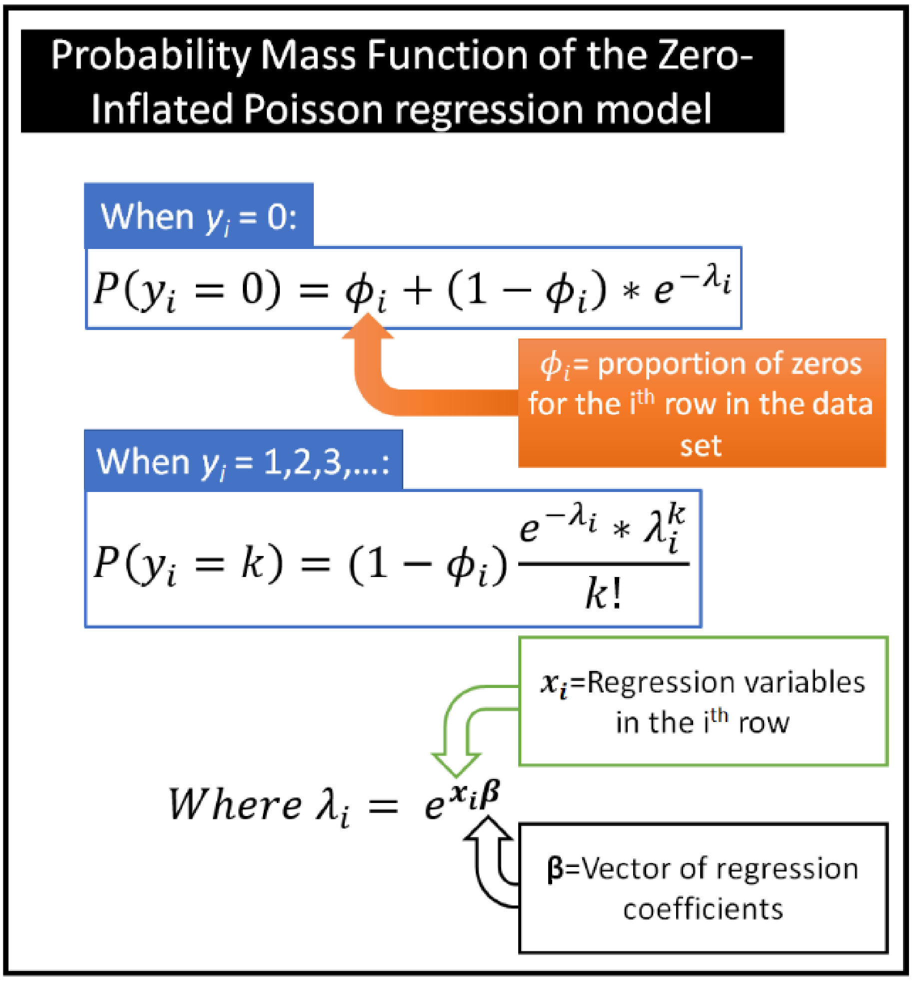


Figure 3:Probability Mass Function of the ZIP model (Date,2021c)

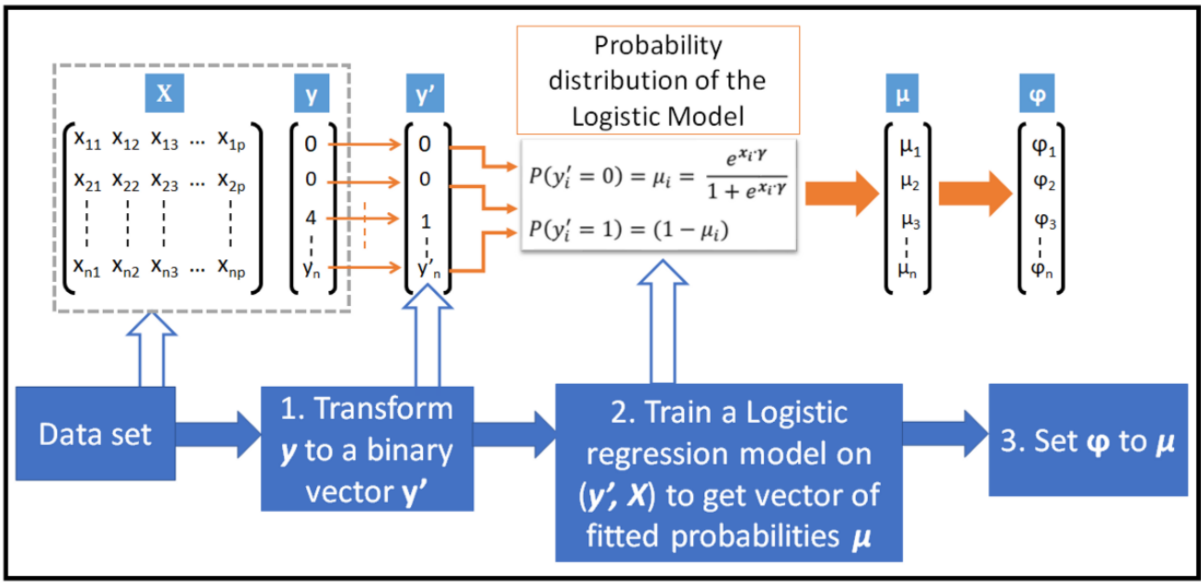


Figure 4: The training sequence for estimating excess zeros parameter ϕ in a ZIP model (Date, 2021c)

3 METHODOLOGY

3.1 Overview

The purpose of this study is to investigate the impact of alcohol advertising on social media platforms by NTE on the Accidents and Emergency (A&E) data for assault-related injuries in Cardiff. For this project, we decided to select the establishments based on their popularity and internet presence. For this experiment, we chose nightclubs as a representative of NTE because over a quarter (16,529) of the 58,885 offences that occurred at or near nightclubs in the year leading up to July 2019 were classified as violent or sexual offences by police in England, Wales, and Northern Ireland (Cachia, 2019).

The study was conducted in two phases.

Phase 1: The goal of this phase was to identify the most suitable social media platform to study. This involved collecting and analysing different types of promotions run by the Night-Time Economy on social media accounts including Twitter, Facebook, and Instagram. Twitter was selected after a thorough analysis of all platforms since its data was readily available. I examined 3336 tweets from Cardiff nightclubs that were sent between the years of 2019 and 2020 on twitter. The data was then divided into 3 categories, using an algorithm implemented in python, based on regular expressions.

1. Data containing alcohol promotion.
2. Data containing retweets or posts meant for specific individuals.
3. Data containing all other posts made by nightclub the majority of which were advertisements for the location and events.

After removing 1464 rows containing retweets, the data was then compared to a manually labelled dataset containing 15 rows. As the manually labelled dataset was unbalanced, we used precision, recall and F1 score for measuring performance instead of accuracy.

Phase 2: After completion of Phase 1, the methods of extraction and categorization of twitter data were applied to the remaining 4 clubs. After getting all the data from clubs and obtaining secondary A&E data for assault-related injury the dataset was analysed. Following data analysis, the final models also included variables weekend, a binary variable (True or False), promo_sum, which represents the total number of alcohol-related promotions, and all_sum, which represents all other tweets excluding retweets. With these variables we used Poisson, Negative Binomial NB2 and Zero Inflated Poisson models to fit the observed counts Y which represents the number of assaults to the matrix of regression values X.

3.2 Tools Setup and Data Collection:

Twitter data:

A developer account was set up in twitter in order to use twitter API. The Twitter API (Tweepy) makes it possible for us to access Twitter in novel and sophisticated ways, use it to study Tweets, Direct Messages, users, and other important Twitter resources. The dependencies and related packages were also installed.

The combination of users, API keys, access tokens, authorisation key etc. is used to access the online services. To extract, modify, and further analyse the outputs from the services' APIs, Python is used.

Data collection from twitter was approved by the Ethics committee. Data collected by twitter does not have any identifiers except from the nightclubs handles which were anonymised. Twitter data was stored in excel and csv files.

Assault Data:

The A&E (Attendances and Emergency) data of University Hospital Wales is secondary data provided by Professor Simon C Moore – School of Dentistry and the Emergency Department. A&E data is anonymised and aggregated data that contains no patient identifiable information. Data is generated as part of patients' usual care, anonymised and made available to support public health surveillance. Neither the generation of this data nor use in research require patient consent or ethical approval (https://www.hra.nhs.uk/covid-19-research/guidance-using-patient-data/). The number of assault-related admissions and the number of male assault-related injury cases, with respect to date are both included in the A&E data.

3.3 PHASE 1: Data Selection

Cardiff nightclubs were chosen using data from Google. A study was conducted to identify the nightclubs in Cardiff with the most active social media accounts and study their presence on social media.

Data was collected by manually analysing popular clubs in Cardiff and going through their Twitter, Instagram and Facebook accounts. The objective of the study was to find the most active platform used to promote alcoholic beverages to the audience.

After deciding on the platform to use for data collection and selecting 5 nightclubs for the study, we obtained data from the Twitter API for the nightclubs, including tweets, likes, time, and retweet data for 2019. The data was then saved to an excel file and manually labelled to differentiate between posts containing drinks promotions and other promotional posts. After manually labelling the data and loading into the data frame, I wrote an algorithm to find keywords such as “drink”, “drinks”, and “£” using Regex to see how well it identifies promotional posts from all the tweets (Fig 5). Tweets were classified into 3 categories: Yes, No and None. “Yes”, had drinks promotions, “No” were all tweets made by nightclubs excluding retweets and “None” were retweets. While initially running the data, I found that there were some outliers because retweets were also included in the dataset, so retweets were identified using “RT @” and sentences beginning with “@” and then they were removed. Fig 6 shows the dataset after removing retweets and combining both manually labelled tweets and tweets labelled by algorithm.

```python
for tweet in new_Df['tweets']:

    if 'RT @' in tweet or re.findall('\A@', tweet):
        Promo.append('None')

    elif 'RT @' not in tweet:
            #
            if  re.findall('drink', tweet, flags=re.IGNORECASE) or re.findall('£', tweet):
                Promo.append('Yes')
            else:
                Promo.append('No')

# adding labelled data to data frame
new_Df['Promo']= Promo

#printing new data frame:
new_Df
```

Figure 5: Code for classifying tweets

tweets	likes	time	Promotior	Promo
ðŸ¥° Love is in the air this weekend at l	0	2020-02-1	Yes	Yes
ðŸSouthern Comfort | February VIP o	2	2020-02-1	Yes	Yes
Upgrade your night to VIP this evening	0	2020-02-1	Yes	Yes
Single & ready to mingle? Check c	0	2020-02-1	Yes	Yes
Special offer for FRIDAY's Valentines p	0	2020-02-1	Yes	Yes
Who's joining us tomorrow for our Val	0	2020-02-1	Yes	Yes
Celebrating Valentines day with your g	1	2020-02-1	Yes	Yes
Grab your Â£5 tickets for NYE now! ðŸ	0	2019-12-2	No	Yes
Merry Christmas to all our staff, prom	0	2019-12-2	No	No
VIP Booths available for NYE from just	0	2019-12-2	Yes	Yes
Joel Corry, FooR & Freddie Mercu	0	2019-12-2	No	Yes
Catch Joel Corry in our Main Room on	1	2019-12-2	No	Yes
Â£5 tickets available for NYE! Entry by	0	2019-12-2	No	Yes

Figure 6:Twitter dataset

After labelling the data set, the algorithm was checked for performance. Accuracy is not a good measure of efficacy if the dataset is not balanced so as we have unbalanced data in our testing dataset, to measure the efficiency of the algorithm we use Precision, Recall and F1 score. **Precision** is the proportion of accurately anticipated positive observations to all positively expected observations. **Recall** (Sensitivity) - The ratio of accurately anticipated positive observations to all observations is known as recall. **F1 score**- The weighted average of Precision and Recall. Therefore, both false positives and false negatives are included while calculating F1 score (Exsilio Solutions, 2016). After calculating the performance measures, the number of promotional and non-promotional tweets from 2019–20 were computed to study the effect of tweets from that time frame on the assaults. Due to covid being most active during 2020 and 2021 the data for the year 2019 was chosen for the study.

3.4 PHASE 2 : Data Analysis

After the data was cleaned, transformed and collated into a single data frame, we had to determine which regression model fits the data best and whether there is indeed a relationship between promotional data and assault related injuries. Correlation and linear regression are the two most often used methods for examining the connection between two quantitative variables. Regression expresses the relationship as an equation, whereas correlation assesses the strength of the linear link between two variables (Bewick, Cheek and Ball, 2003). The number of assault-related admissions in the A&E department is a count data point, hence we will use Poisson regression. In our data set, the regression variable X matrix consists of the variables weekend, promo_sum, and all_promo. The number of assaults is represented by the observed counts vector y in Fig 7, Python **statsmodels** package was used to to configure and fit all 3 models.

date	promo_sum	weekend	all_sum	assaults
01/01/2019	< 5	FALSE	< 5	< 5
02/01/2019	< 5	FALSE	< 5	< 5
03/01/2019	< 5	FALSE	< 5	< 5
04/01/2019	< 5	FALSE	< 5	< 5
05/01/2019	< 5	TRUE	< 5	< 5
06/01/2019	< 5	TRUE	< 5	< 5
07/01/2019	< 5	FALSE	< 5	< 5
08/01/2019	< 5	FALSE	< 5	< 5
09/01/2019	< 5	FALSE	< 5	< 5
10/01/2019	< 5	FALSE	< 5	< 5
11/01/2019	< 5	FALSE	< 5	6
12/01/2019	5	TRUE	7	< 5
13/01/2019	0	TRUE	< 5	< 5
14/01/2019	0	FALSE	< 5	< 5
15/01/2019	0	FALSE	< 5	< 5

Regression Variable matrix X

Observed Counts vector y

Figure 7: Few values of the dataset

Poisson Model:
The steps to implement a Poisson Model are:

1. Determine the dependent variable y, which in our model represents injuries caused by assaults.
2. Identify the regression factors that will affect the observed numbers.
3. Use Python **statsmodels package GLM class** configure and fit the Poisson Regression model on the data.
4. Use a goodness-of-fit measure to determine model fit.

Negative Binomial Model (NB2):
The steps to implement NB2 model are:

1. Determine the dependent variable y, which in our model represents injuries caused by assaults.
2. Identify the regression factors that will affect the observed numbers
3. Analyze the data and fit the Poisson regression model. This will provide us with the fitted rates vector λ save it as a new column (Fig 9).
4. Add a column to store the dependent variable of the OLS regression and use it to configure and fit the OLSR model (Fig 7).
5. Obtain the value of alpha derived from the OLSR model using **aux_olsr_results.params** function and use it in the model.
6. Use **NegativeBinomial** class in the **statsmodels package** we train and fit the NB2 model.
7. Use a goodness-of-fit measure to determine model fit.

```
#Add the λ vector as a new column called 'lam' to the Data Frame of the training data set
df_train['lam'] = poisson_training_results.mu

#add a derived column called 'OLS_DEP' to the pandas Data Frame. This new column will store the values of the dependent variable
df_train['OLS_DEP'] = df_train.apply(lambda x: ((x['assaults'] -x['lam'])**2 -x['lam']) / x['lam'], axis=1)

#use patsy to form the model specification for the OLSR
olsr_expr = """OLS_DEP ~ lam -1"""

#Configure and fit the OLSR model
olsr_results = smf.ols(olsr_expr, df_train).fit()
```

Figure 8: OLSR model

```
[1.45053803 1.45053803 1.36068062 1.48827956 2.10062085 2.35738971
 1.45053803 1.56673415 1.48827956 1.52700308 1.46007698 2.61247997
 2.29760828 1.38696328 1.45053803 1.45053803 1.38696328 1.51702689
 2.10062085 2.29760828 1.58656342 1.74675658 1.63854913 1.67019904
 1.58656342 2.46544627 2.29760828 1.52700308 1.8743554  1.71365592
 2.31585095 1.99813515 2.57845585 2.1969078  1.48827956 1.62784419
 1.67019904 1.70246029 1.74675658 3.06620254 2.2540691  1.59699689
 1.59699689 1.52700308 1.68118251 1.88668143 2.80331902 2.15527687
 1.62784419 1.56673415 1.62784419 1.83883677 1.85092923 2.96892214
 2.40292461 1.3261749  1.56673415 1.56673415 1.38696328 1.48827956
```

Figure 9: First few values of fitted λ vector

Zero Inflated Poisson Model (ZIP):
Steps to implement Zero Inflated Poisson model:

1. Determine the dependent variable y, which in our model represents injuries caused by assaults.
2. Identify the regression factors that will affect the observed numbers
3. Use **statsmodels's ZeroInflatedPoisson** class, we built and trained a ZIP regression model on the data set.
4. Use a goodness-of-fit measure to determine model fit.

Goodness-of-fit: A goodness-of-fit test, in general, refers to evaluating how closely the observed data match the anticipated model after fitting. This idea will be used to assess the model's fitness. The **summary()** method on the **statsmodels GLMResults** class has some goodness-of-fit statistics to evaluate whether the regression model was able to successfully fit the data or not.

Log-Likelihood*:* A technique to gauge a regression model's quality of fit is by looking at its log-likelihood value. The better a model matches a dataset, the greater the log-likelihood value. For a particular model, the log-likelihood value can be between -infinity and infinity. Though mainly useless, the actual log-likelihood number for a specific model might be helpful when comparing two or more models (Zach, 2021).

Deviance and Pearson Chi square: Used to compare to the values in the Chi square table to determine if the model correctly describes the data or not.

Next Steps:
As males account for 77% of the 686 assault-related injuries recorded in Cardiff during 2019. We split the data on assaults depending on gender and ran the Negative Binomial NB2 model on male and female assaults separately after determining that the model had the best match for our data. The results of our experiments are presented in the next chapter.

4 RESULTS

4.1 PHASE 1:

It was observed that Facebook has the greatest number of alcoholic drinks promotional posts in all the social media platforms. 70 percent of the clubs in the research have alcohol promotional posts in Facebook. It has also been observed during the research that majority of the promotions are done in the stories of Facebook and Instagram as shown in Table 1. Most nightclubs have a Twitter presence, though it is small in comparison to their Facebook activity that is suggested in the literature (Nevin et al., 2012).

Club	Twitter	Facebook	Instagram
X	5	0	1
Y	2	3	1
Z	0	1	2
A	1	3	3
B	0	0	1
C	0	3	3
D	1	1	0
E	0	3	NA
F	0	0	0
G	3	2	2
Total:	12	16	13

Table 1: Number of alcoholic promotions by Nightclubs on different platforms

Facebook is the best source to collect data as observed during this analysis. However, there are some challenges which arise in collecting Facebook data. Facebook has no set API to retrieve public account posts from the website unlike twitter. Due to restrictions on collecting data from Facebook the next best data source twitter was finalised.

Results of the performance of labelling algorithm:

Scores:

Precision	0.667
Recall	1
F-Measure	0.800

Table 2: Performance matrix of the labelling algorithm

Category	No of Tweets
Alcohol Promotional	810
Other tweets	1062
All tweets	1872

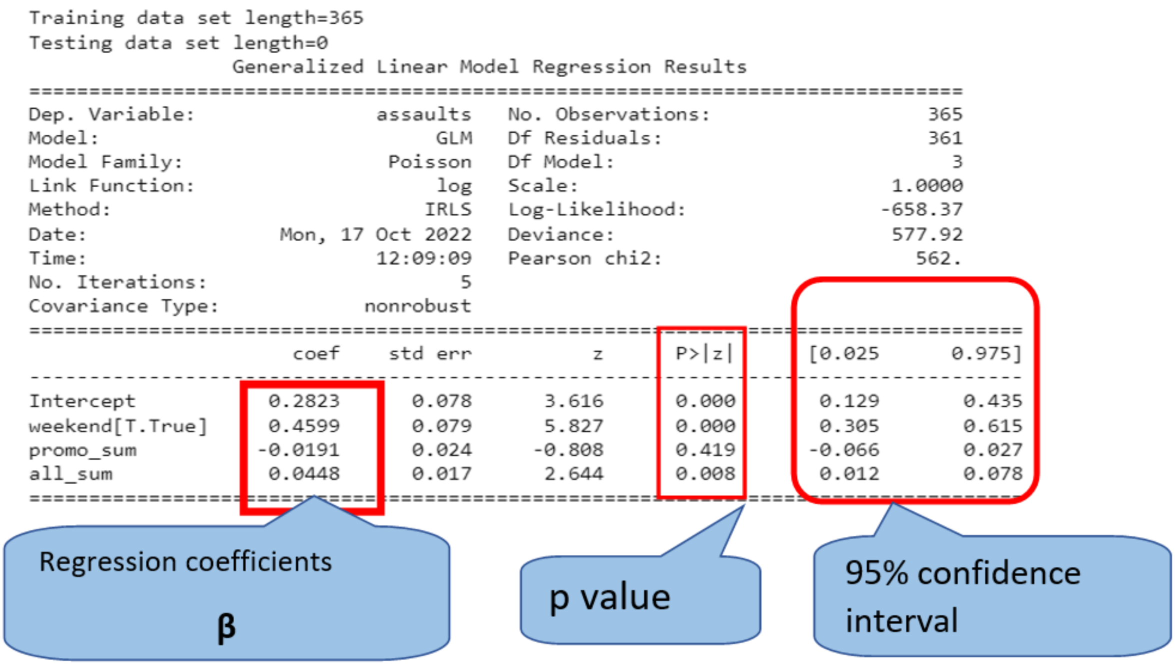
Table 3: Twitter data statistics

Category	Total
Assault related injuries	686
Male	530
Female	156

Table 4: Assault data statistics

4.2 PHASE 2

4.2.1 Poisson model statistics:



```
Training data set length=365
Testing data set length=0
                 Generalized Linear Model Regression Results
==============================================================================
Dep. Variable:                assaults   No. Observations:                  365
Model:                             GLM   Df Residuals:                      361
Model Family:                  Poisson   Df Model:                            3
Link Function:                     log   Scale:                          1.0000
Method:                           IRLS   Log-Likelihood:                -658.37
Date:                 Mon, 17 Oct 2022   Deviance:                       577.92
Time:                         12:09:09   Pearson chi2:                     562.
No. Iterations:                      5
Covariance Type:             nonrobust
===================================================================================
                      coef    std err          z      P>|z|      [0.025      0.975]
-----------------------------------------------------------------------------------
Intercept           0.2823      0.078      3.616      0.000       0.129       0.435
weekend[T.True]     0.4599      0.079      5.827      0.000       0.305       0.615
promo_sum          -0.0191      0.024     -0.808      0.419      -0.066       0.027
all_sum             0.0448      0.017      2.644      0.008       0.012       0.078
===================================================================================
```


Figure 10: Training summary for the Poisson Model

Variance	3.095
Mean	1.879

Table 5: Variance vs Mean of Assault related injuries

Variable	Coefficient	P value	CI (95%)	Log Likelihood
Weekend	0.4599	< 0.001	0.305, 0.615	-658
Promo_sum	-0.191	0.419	-0.066, 0.027	
All_sum	0.0448	0.008	0.012, 0.078	

Table 6: Poisson Model Results

As we can see from Table 6:

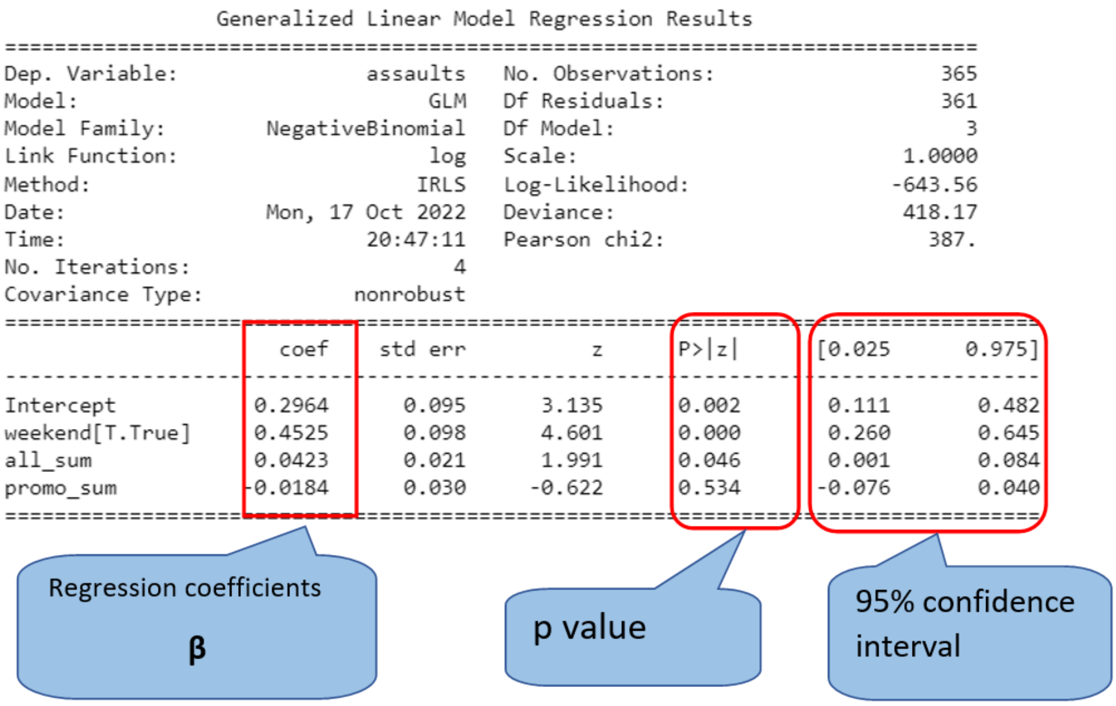
- Weekend and all tweets by nightclubs are statistically significant as p<0.05 and alcohol promotion is statistically not significant p>0.05 in affecting the number of assault injury attendances.
- The weekend is positively associated with assault data. (0. 459, 95% confidence interval [CI], 0.305, 0.615).
- Number of alcohol promotions is negatively associated with the number of assault injury attendances (-0.0184, 95% confidence interval [CI], -0.076, 0.04).
- Number of tweets made by clubs is positively associated with the number of assault injury attendances (0.04, 95% confidence interval [CI], 0.01, 0.08).
- As shown in Table 5 variance > mean hence data is over dispersed.

We can see from Fig 10 that the values of Deviance and Pearson chi2 are 577.9 and 562, respectively, we may use any of them to compare to the values in the Chi square table to determine if the model correctly describes the data or not. We check up the value in the Chi square table for p=0.05 and Degrees of freedom of residuals = 361 at 95% confidence level (p=0.05). (DF Residuals = [DF model - No. of Observations]).

The value is 406.304

Pearson chi2 is greater than the Chi square table value hence the model fails goodness of fit test.

4.2.2 Negative Binomial model NB2 statistics:



```
                 Generalized Linear Model Regression Results
==============================================================================
Dep. Variable:                assaults   No. Observations:                  365
Model:                             GLM   Df Residuals:                      361
Model Family:         NegativeBinomial   Df Model:                            3
Link Function:                     log   Scale:                          1.0000
Method:                           IRLS   Log-Likelihood:                -643.56
Date:                 Mon, 17 Oct 2022   Deviance:                       418.17
Time:                         20:47:11   Pearson chi2:                     387.
No. Iterations:                      4
Covariance Type:             nonrobust
===================================================================================
                     coef    std err          z      P>|z|      [0.025      0.975]
-----------------------------------------------------------------------------------
Intercept          0.2964      0.095      3.135      0.002       0.111       0.482
weekend[T.True]    0.4525      0.098      4.601      0.000       0.260       0.645
all_sum            0.0423      0.021      1.991      0.046       0.001       0.084
promo_sum         -0.0184      0.030     -0.622      0.534      -0.076       0.040
===================================================================================
```


Figure 11: Training summary for NB2 model

Variable	Coefficient	P value	CI (95%)	Log Likelihood
Weekend	0.4525	< 0.001	0.260, 0.645	-643
All_sum	0.0423	0.046	0.001, 0.084	
Promo_sum	-0.0184	0.534	-0.076, 0.04	

Table 7: NB2 model statistics

As we can see from above Table 7:

- Weekend and all tweets by nightclubs are statistically significant as p<0.05 and alcohol promotion is statistically not significant p>0.05 in affecting the number of assault injury attendances.
- The weekend is positively associated with assault data. (0.452, 95% confidence interval [CI], 0.260, 0.645)
- Number of tweets made by nightclubs is positively associated with the number of assault injury attendances (0.04, 95% confidence interval [CI], 0.01, 0.08)
- Number of alcohol promotions is negatively associated with the number of assault injury attendances (-0.0184, 95% confidence interval [CI], -0.076, 0.04)

As shown in Table 7 the Log Likelihood value of NB2 model is -643 which is better than the value for Poisson model which is -658. This means the NB2 model fits the data better than the Poisson model.

Pearson chi2 value is 387 (Fig 11) which is less than the Chi square table value hence the model passes goodness of fit test and is a good fit for the data.

4.2.3 Zero Inflated Poisson Model statistics:

```
                     ZeroInflatedPoisson Regression Results
==============================================================================
Dep. Variable:                 assaults   No. Observations:                  365
Model:              ZeroInflatedPoisson   Df Residuals:                      361
Method:                             MLE   Df Model:                            3
Date:                  Wed, 19 Oct 2022   Pseudo R-squ.:                 0.02643
Time:                          12:14:06   Log-Likelihood:                -649.01
converged:                        False   LL-Null:                       -666.63
Covariance Type:              nonrobust   LLR p-value:                 1.083e-07
==========================================================================================
                             coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------------------
inflate_Intercept         -1.2978      0.443     -2.928      0.003      -2.167      -0.429
inflate_weekend[T.True]   -1.3174      0.791     -1.665      0.096      -2.868       0.233
inflate_promo_sum         -0.0186      0.162     -0.115      0.908      -0.335       0.298
inflate_all_sum           -0.0714      0.117     -0.608      0.543      -0.301       0.159
Intercept                  0.4948      0.084      5.911      0.000       0.331       0.659
weekend[T.True]            0.3382      0.083      4.052      0.000       0.175       0.502
promo_sum                 -0.0208      0.025     -0.842      0.400      -0.069       0.028
all_sum                    0.0373      0.018      2.100      0.036       0.002       0.072
==========================================================================================
```

Figure 12:Training Summary for the Zero Inflated Model

Variable	Coefficient	P value	CI (95%)	Log Likelihood
Weekend	0.3382	0.000	0.175, 0.502	-649
Promo_sum	-0.0208	0.400	-0.069, 0.028	
All_sum	0.0373	0.036	0.002, 0.072	

Table 8: ZIP statistics

As we can see from above Table 8:

- Weekend and all tweets by nightclubs are statistically significant as $p<0.05$ and alcohol promotion is statistically not significant $p>0.05$ in affecting the number of assault injury attendances.
- The weekend is positively associated with assault data. (0.338, 95% confidence interval [CI], 0.175, 0.502)
- Number of alcohol promotions is negatively associated with the number of assault injury attendances (-0.02, 95% confidence interval [CI], -0.069, 0.028)
- Number of tweets made by clubs is positively associated with the number of assault injury attendances (0.037, 95% confidence interval [CI], 0.002, 0.072).
- As shown by Fig 12, the ZIP model's training procedure was unable to converge on the data set which means the model is not a good fit for the data.
- The Log Likelihood value of ZIP model as shown in Table 8 is -649.

Model name	Log-Likelihood
Poisson	-658
Negative Binomial NB2	-643
Zero Inflation Poisson ZIP	-649

Table 9: All model Goodness of fit data

As shown in Table 9, the NB2 model has a higher Log -Likelihood than the other two models, indicating that it fits the data better.

4.2.4 NB2 Model for Male assault data:

```
                 Generalized Linear Model Regression Results
==============================================================================
Dep. Variable:                   male   No. Observations:                  365
Model:                            GLM   Df Residuals:                      361
Model Family:        NegativeBinomial   Df Model:                            3
Link Function:                    log   Scale:                          1.0000
Method:                          IRLS   Log-Likelihood:                -575.02
Date:                Thu, 20 Oct 2022   Deviance:                       404.20
Time:                        14:34:43   Pearson chi2:                     381.
No. Iterations:                     5
Covariance Type:            nonrobust
===================================================================================
                      coef    std err          z      P>|z|      [0.025      0.975]
-----------------------------------------------------------------------------------
Intercept          -0.0392      0.105     -0.372      0.710      -0.246       0.167
weekend[T.True]     0.4882      0.108      4.512      0.000       0.276       0.700
promo_sum          -0.0249      0.032     -0.769      0.442      -0.088       0.039
all_sum             0.0564      0.023      2.427      0.015       0.011       0.102
===================================================================================
```

Figure 13: Training Summary for the males data NB2 Model

Variable	Coefficient	P value	CI (95%)	Log Likelihood
Weekend	0.4882	0.000	0.276, 0.700	-575
Promo_sum	-0.0249	0.442	-0.088, 0.039	
All_sum	0.0564	0.015	0.011, 0.102	

Table 10: Male NB2 statistics

As shown in Table 10:

- Weekend and all tweets by nightclubs are statistically significant as p<0.05 and alcohol promotion is statistically not significant p>0.05 in affecting the number of assault injury attendances.
- The weekend is positively associated with assault data. (0.4882, 95% confidence interval [CI], 0.276, 0.700)
- Number of alcohol promotions is negatively associated with the number of assault injury attendances (-0.024, 95% confidence interval [CI], -0.088, 0.039).
- Number of tweets made by clubs is positively associated with the number of assault injury attendances (0.056, 95% confidence interval [CI], 0.011, 0.102).
- Deviance and Pearson chi2 are 404 and 381 (Fig 13) which are lower than the Chi square table value which is 406 indicating that model is a good fit.

4.2.5 NB2 Model for Female assault data:

```
                 Generalized Linear Model Regression Results
==============================================================================
Dep. Variable:                 female   No. Observations:                  365
Model:                            GLM   Df Residuals:                      361
Model Family:        NegativeBinomial   Df Model:                            3
Link Function:                    log   Scale:                          1.0000
Method:                          IRLS   Log-Likelihood:                -313.63
Date:                Thu, 20 Oct 2022   Deviance:                       301.00
Time:                        14:32:28   Pearson chi2:                     380.
No. Iterations:                     5
Covariance Type:            nonrobust
===================================================================================
                      coef    std err          z      P>|z|      [0.025      0.975]
-----------------------------------------------------------------------------------
Intercept          -0.9139      0.176     -5.183      0.000      -1.260      -0.568
weekend[T.True]     0.3131      0.188      1.669      0.095      -0.055       0.681
all_sum            -0.0106      0.042     -0.255      0.798      -0.092       0.071
promo_sum           0.0084      0.058      0.145      0.885      -0.105       0.122
===================================================================================
```

Figure 14: Training Summary for the Females data NB2 Model

Results:

Variable	Coefficient	P value	CI (95%)	Log Likelihood
Weekend	0.3131	0.095	-0.055, 0.681	-313
All_sum	-0.0106	0.798	-0.092, 0.071	
Promo_sum	0.0084	0.885	-0.105, 0.122	

Table 11: Female NB2 statistics

As shown in Table 11:

- Weekend, all tweets by nightclubs and alcohol promotion are statistically not significant p>0.05 in affecting the number of assault injury attendances.
- The weekend is positively associated with female assault data. (0.313, 95% confidence interval [CI], -0.055, 0.681).
- Number of tweets made by clubs is positively associated with the number of female assault injury attendances (0.04, 95% confidence interval [CI], 0.01, 0.08).
- Number of alcohol promotions is negatively associated with the number of assault injury attendances (-0.0184, 95% confidence interval [CI], -0.076, 0.04).
- For the females assault data Deviance and Pearson chi2 are 301 and 380 (Fig 13) which are lower than the Chi square table value which is 406.

5 DISCUSSION

This experiment aimed to study the relationship of promotional activity of night time economy and its relationship to assault related injury in Cardiff.

It was found that promotions on twitter and weekends have a significant positive effect on assault related injuries in Cardiff. Our results show that posts by nightclubs on twitter have a positive correlation with assault related injuries in males. A further finding was that males were shown to have a higher likelihood of suffering an assault-related injury on the weekends than they did during the work week. These results build on existing evidence of assault-related injuries increased on weekends (Khurana, Prakash and Loder, 2022). Alcohol promotions were discovered to impact negatively on the assault related injuries which means increase in alcohol promotion leads to decrease in assault related injuries.

Twitter data analysed from 5 prominent night clubs in Cardiff shows that out of 1,872 posts made in the year 2019, 43% of the tweets were Alcoholic promotions and 56% Non-alcoholic promotions (Table 2). Consumers do not like nightclubs that solely focus their social media efforts on promotions (Nevin et al., 2012). Nightclubs' attention to promotional activities may indicate that they are underperforming in terms of meeting the needs and expectations of their customers on social media. It may also result in price wars between competitors (Nevin et al., 2012). The literature highlights the associated co-creation and interactive nature of social media marketing (Moraes, Michaelidou and Meneses, 2014), implying that there is significant potential for nightclubs to use social media channels for non-promotional purposes such as consumer research. According to Nevin et al. (2012), some nightclubs admitted to using social media channels for competitor analysis, as one nightclub put it, 'keeping an eye on what they're doing down the road.' This intriguing discovery adds to the literature on social media marketing by identifying a new, non-promotional business activity for which social media can be used (Nevin et al., 2012).

70 percent of the clubs had Facebook postings promoting alcohol which was highest out of all three platform which supports the assertion that Facebook is widely used as a marketing tool (Miletsky, 2010, Nevin et al., 2012). Males are more likely than females to suffer assault injuries (Kwan et al., 2019), males account for 77% of the 686 assault-related injuries recorded in 2019 according to Cardiff A&E data, while females account for 22%.

Because social media has such a broad reach into the lives of young people, it has the potential to strongly influence their decisions (Moreno and Whitehill, 2014). According to the growing body of literature on social media and alcohol, researchers can look at the role of social media in alcohol consumption in two ways (Moreno and Whitehill, 2014). First, as demonstrated by studies that link online content to offline behaviour or demonstrate links between online and offline alcohol consumption patterns, social media can serve as a source of information about the individual user's behaviour. Second, according to such behavioural models and new theoretical frameworks, social media can be a source of influence on behaviour (Moreno and Whitehill, 2014).

The research's biggest limitation is the sample size, which is too small. Five out of the ten nightclubs that took part in the research have a presence on Twitter. A similar finding was drawn from Nevin et al. (2012) that twitter may not yet be a viable social media channel for nightclubs in the midlands and west of Ireland as usage among consumers appears to be relatively low. Another limitation is that the study only focuses on the text data in the form of tweets, however, it has been found by various studies (Paradis et al., (2020), Nevin et al. (2012)) that apart from text promotions, social networking sites also use pictures for alcohol promotions and to boost customer traffic. People are constantly exposed to, and frequently influenced by, images of their peers having fun while drinking, with little representation of negative outcomes (Jones et al., 2017). Limitations of the current approach also include the possibility that social media may be used in response to the volume of customers entering a store or business during a specific period of time rather than to drive customers to the premises. There are no falsifiability checks that serve as a negative control in cases where an association is not expected; for example, an association is not expected between tweets by restaurants and assault related injuries.

Future research:

Although the study establishes a relationship between weekend and twitter promotions with assault related injuries in male data from Cardiff in the year 2019 it would be interesting to see how these relationships behave over a longer period and across different platforms such as Facebook, Instagram, Snapchat and TikTok. Nightclubs use images, stories and personalised ads to market alcohol and increase customer traffic however little research has been done into these methods for marketing. Future research should consider all these methods. Restaurants act as negative controls where we do not expect to see an association. If there were more time available first step would be to include falsifiability checks. Another aspect to investigate in the future could be to study the effect of the potential confounding effect of the number of visits and length of stay at the retail venue on the assault related injuries.

6 CONCLUSION

Three regression count data models were used in this study to determine the effects of alcohol promotions, tweets by nightclubs, and weekends on assault-related injuries in Cardiff. Secondary data obtained from Twitter and University Hospital Wales for the 2019/2020 academic year was used to study the relationship. The dependent variable is the number of visits to the Assault and Emergency Department for assault-related injuries, while the independent variables are the effects of alcohol promotions, tweets by nightclubs, and the weekend. The analysis results show that the variables, tweets by nightclubs and weekend have a positive effect on assault related injuries in Cardiff, whereas alcohol promotions have a negative effect on assault related injuries in Cardiff. The study's findings also show that Poisson regression was unsuitable for modelling the data due to overdispersion, so other models such as Negative Binomial NB2 and Zero Inflated Poisson ZIP were applied to the count data. The result shows that the Negative Binomial NB2 model is the best fit for the assault-related injury data. According to the goodness of fit of model results, the NB2 model was the best fit, with the highest log-likelihood value of -643. Nightclubs use a variety of social networking sites and techniques to advertise alcohol and boost customer traffic. These could undercut efforts to alter drinking-related social norms, particularly the normalisation of daily consumption. More research is needed in the future to broaden our understanding of alcohol content across social media platforms.

7 REFERENCES:

1. Advertising Standards Authority | Committee of Advertising Practice (2019). Alcohol. [online] Asa.org.uk. Available at: https://www.asa.org.uk/topic/alcohol.html [Accessed 2022].
2. Advertising Standards Authority | Committee of Advertising Practice (2019a). 18 Alcohol. [online] Asa.org.uk. Available at: https://www.asa.org.uk/type/non_broadcast/code_section/18.html.
3. Allen, J., Nicholas, S., Salisbury, H. and Wood, M.(2003). Nature of burglary, vehicle and violent crimein C. Flood-Page and J. Taylor (eds). Crime in Englandand Wales 2001/2002: Supplementary Volume. HomeOffice Statistical Bulletin 01/03. London: Home Office
4. Anderson, P., de Bruijn, A., Angus, K., Gordon, R. and Hastings, G. (2009). Impact of Alcohol Advertising and Media Exposure on Adolescent Alcohol Use: A Systematic Review of Longitudinal Studies. *Alcohol and Alcoholism*, [online] 44(3), pp.229–243. doi:10.1093/alcalc/agn115.
5. Atkinson, A. M., Ross-Houle, K. M., Begley, E. & Sumnall, H. (2017). An exploration of alcohol advertising on social networking sites: an analysis of content, interactions and young people's perspectives. Addiction Research & Theory, 25(2), 91-102. DOI: 10.1080/16066359.2016.1202241
6. Barry, A.E., Valdez, D., Padon, A.A. and Russell, A.M. (2018). Alcohol Advertising on Twitter—A Topic Model. *American Journal of Health Education*, 49(4), pp.256–263. doi:10.1080/19325037.2018.1473180.
7. Bewick, V., Cheek, L. and Ball, J. (2003). Statistics review 7: Correlation and regression. *Critical Care*, [online] 7(6), pp.451–459. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC374386
8. Brunborg, G.S., Skogen, J.C. and Burdzovic Andreas, J. (2022). Time spent on social media and alcohol use among adolescents: A longitudinal study. Addictive Behaviors, 130, p.107294. doi:10.1016/j.addbeh.2022.107294.
9. Butchart, A. (2008). New manual for estimating the economic costs of injuries due to interpersonal and self-directed violence. *Injury Prevention*, 14(2), pp.143–143. doi:10.1136/ip.2008.018580.
10. Caces, F. M., Stinson, F. S. & Harford, T. C. (1991) Alcohol use and physically risky behavior among adolescents. Alcohol Health and Research World, 15, 228–233
11. Cachia, A. (2019). The nightclub crime hotspots where thousands of crimes take place. [online] WalesOnline. Available at: https://www.walesonline.co.uk/news/wales-news/nightclub-crime-hotspots-thousands-crimes-16938777 [Accessed 12 Oct. 2022].
12. Cameron A. C. and Trivedi P. K. (2013). Regression Analysis of Count Data, Second Edition, Econometric Society Monograph No. 53, Cambridge University Press, Cambridge.
13. Cameron, A.C. and Trivedi, P.K. (1986). Econometric models based on count data. Comparisons and applications of some estimators and tests. Journal of Applied Econometrics, 1(1), pp.29–53. doi:10.1002/jae.3950010104.
14. Campbell, C.A., Hahn, R.A., Elder, R., Brewer, R., Chattopadhyay, S., Fielding, J., Naimi, T.S., Toomey, T., Lawrence, B. and Middleton, J.C. (2009). The Effectiveness of Limiting Alcohol Outlet Density As a Means of Reducing Excessive Alcohol Consumption and Alcohol-Related Harms. American Journal of Preventive Medicine, 37(6), pp.556–569. doi:10.1016/j.amepre.2009.09.028.
15. Chaplin, R., Flatley, J. and Smith, K., 2011. Crime in England and Wales 2010/11. Home Office Statistical Bulletin, 10(11).

16. Coxe, S., West, S.G. and Aiken, L.S. (2009). The Analysis of Count Data: A Gentle Introduction to Poisson Regression and Its Alternatives. Journal of Personality Assessment, 91(2), pp.121–136. doi:10.1080/00223890802634175.
17. Date, S. (2021a). The Negative Binomial Regression Model. [online] Time Series Analysis, Regression and Forecasting. Available at: https://timeseriesreasoning.com/contents/negative-binomial-regression-model/ [Accessed 20 Oct. 2022].
18. Date, S. (2021b). The Poisson Regression Model. [online] Time Series Analysis, Regression and Forecasting. Available at: https://timeseriesreasoning.com/contents/poisson-regression-model/.
19. Date, S. (2021c). The Zero Inflated Poisson Regression Model. [online] Time Series Analysis, Regression and Forecasting. Available at: https://timeseriesreasoning.com/contents/zero-inflated-poisson-regression-model/.
20. Edwards, A. (2010). *Evaluation of the Cardiff Night-Time Economy Co-ordinator (NTEC) post*. [online] orca.cardiff.ac.uk. Available at: https://orca.cardiff.ac.uk/id/eprint/78193/ [Accessed 10 Oct. 2022].
21. Evans, G. (2012). Hold back the night:Nuit Blancheand all-night events in capital cities. *Current Issues in Tourism*, 15(1-2), pp.35–49. doi:10.1080/13683500.2011.634893.
22. Exsilio Solutions (2016). Accuracy, Precision, Recall & F1 Score: Interpretation of Performance Measures - Exsilio Blog. [online] Exsilio Blog. Available at: https://blog.exsilio.com/all/accuracy-precision-recall-f1-score-interpretation-of-performance-measures/.
23. Federal Trade Commission. Self-regulation in the Alcohol Industry. (2014). [online] Available at: https://www.ftc.gov/system/files/documents/reports/self-regulation-alcohol-industry-report-federal-trade-commission/140320alcoholreport.pdf.
24. Finney, A. (2004). *Violence in the night-time economy: Key findings from the research*
25. Ford, C. (2016). *Getting started with Negative Binomial Regression Modeling | University of Virginia Library Research Data Services + Sciences*. [online] data.library.virginia.edu. Available at: https://data.library.virginia.edu/getting-started-with-negative-binomial-regression-modeling/.
26. Giancola PR, White HR, Berman ME et al (2003) Diverse research on alcohol and aggression in humans. Alcohol Clin Exp Res27(2): 198–208
27. Griffiths, R., & Casswell, S. (2010). Intoxigenic Digital Spaces? Youth, Social Networking Sites and Alcohol Marketing. Drug & Alcohol Review 29, 525–530.
28. Hardwick, J. (2018). Cardiff City Centre Night-time Economy Policing overview. [online] South Wales Police. Available at: https://cardiff.moderngov.co.uk [Accessed 2022].
29. Hilbe, J.M. (2011). Negative Binomial Regression. [online] Google Books. Cambridge University Press. Available at: https://books.google.co.uk/books?hl=en&lr=&id=0Q_ijxOEBjMC&oi=fnd&pg=PR11&ots=IK61UvxXL6&sig=0iPhqVQQwnyOAXJt80ys7hGSqJ8&redir_esc=y#v=onepage&q&f=false [Accessed 17 Oct. 2022].
30. Homel R, Clark J (2004). The prediction and prevention of violence in pubs and clubs. Crime Prevention Studies. Criminal Justice Press, New York. 3: 1–46
31. Hutchinson, M.K. and Holtman, M.C. (2005). Analysis of count data using poisson regression. Research in Nursing & Health, 28(5), pp.408–418. doi:10.1002/nur.20093.
32. Jones, P., Hillier, D. and Turner, D. 1999. Towards the '24 hour city. Town & Country Planning, 68: 164–165.
33. Kaplan, A.M. and Haenlein, M. (2010). Users of the world, unite! the Challenges and Opportunities of Social Media. Business Horizons, [online] 53(1), pp.59–68. Available at: https://doi.org/10.1016/j.bushor.2009.09.003.

34. Kemp, S. (2022). Digital 2022: Global overview report. [online] DataReportal. Available at: https://datareportal.com/reports/digital-2022-global-overview-report.
35. Khurana, B., Prakash, J. and Loder, R.T. (2022). Assault related injury visits in US emergency departments: An analysis by weekday, month and weekday-by-month. Chronobiology International, pp.1–10. doi:10.1080/07420528.2022.2065285.
36. Kietzmann, J.H., Hermkens, K., McCarthy, I.P. and Silvestre, B.S. (2011). Social media? Get serious! Understanding the Functional Building Blocks of Social Media. Business Horizons, 54(3), pp.241–251. doi:10.1016/j.bushor.2011.01.005.
37. Krug, E.G., Mercy, J.A., Dahlberg, L.L. and Zwi, A.B. (2002). The world report on violence and health. *The Lancet*, 360(9339), pp.1083–1088. doi:10.1016/s0140-6736(02)11133-0.
38. Kwan, K., Wiebe, D., Cerda, M. and Goldman-Mellor, S. (2019). Repeat assault injury among adolescents utilizing emergency care: A statewide longitudinal study. The Journal of emergency medicine, [online] 57(2), pp.254–262. doi: 10.1016/j.jemermed.2019.05.030.
39. Lambert, D. (1992). Zero-Inflated Poisson Regression, with an Application to Defects in Manufacturing. Technometrics, 34(1), p.1. doi:10.2307/1269547.
40. Lewis, S. (2012). *Social Media's Effects on Nightlife*. [online] Bar & Restaurant. Available at: https://www.barandrestaurant.com/operations/social-medias-effects-nightlife.
41. Lovatt, A. and O'Connor, J. (1995). Cities and the Night-time Economy. *Planning Practice & Research*, 10(2), pp.127–134. doi:10.1080/02697459550036676.
42. Lyons, A.C., MCcreanor, T., Hutton, F., Goodwin, I., Moewaka Barnes, H., Griffin, C., Kerryellen, V., O'Carroll, A.D. Niland, P., & Samu, L. (2014). Flaunting it on Facebook: Young adults, drinking cultures and the cult of celebrity. Wellington, New Zealand: Massey University School of Psychiatry.
43. M.A. and Whitehill, J.M. (2014). Influence of Social Media on Alcohol Use in Adolescents and Young Adults. Alcohol research : current reviews, [online] 36(1), pp.91–100. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4432862/.
44. Mart, S.M. (2011). Alcohol Marketing in the 21st Century: New Methods, Old Problems. *Substance Use & Misuse*, 46(7), pp.889–892. doi:10.3109/10826084.2011.570622.
45. Mercy, J.A., Hillis, S.D., Butchart, A., Bellis, M.A., Ward, C.L., Fang, X. and Rosenberg, M.L. (2017). Interpersonal Violence: Global Impact and Paths to Prevention. 3rd ed. [online] PubMed. Available at: https://www.ncbi.nlm.nih.gov/books/NBK525208/#:~:text=Go%20to%3A- [Accessed 30 Sep. 2022].
46. Miletsky, J. (2010) Principles of Internet Marketing: New Tools and Methods for Web Developers, Boston, MA: Course Technology.
47. Moewaka Barnes, H., McCreanor, T., Goodwin, I., Lyons, A., Griffin, C. & Hutton, F. (2016). Alcohol and social media: drinking and drunkenness while online. Critical Public Health, 26:1, 62-76, DOI: 10.1080/09581596.2015.1058921.
48. Moraes, C., Michaelidou, N. and Meneses, R.W. (2014). The use of Facebook to promote drinking among young consumers. Journal of Marketing Management, [online] 30(13-14), pp.1377–1401. doi:10.1080/0267257x.2014.909512.
49. National Institute on Alcohol Abuse and Alcoholism (2015). Alcohol, Violence, and Aggression - Alcohol Alert No. 38-1997. [online] Nih.gov. Available at: https://pubs.niaaa.nih.gov/publications/aa38.htm.
50. Nevin, F. and Torres, A.M., 2012. Club 2.0: Investigating the Social Media Marketing Strategies of Nightclubs in the Midlands and West of Ireland. *Irish Journal of Management, Volume 32, Number 1, ISSN 1649-248X*

51. Ng Fat, L., Cable, N. and Kelly, Y. (2021). Associations between social media usage and alcohol use among youths and young adults: findings from Understanding Society. Addiction, 116(11). doi:10.1111/add.15482.
52. Nicholas, R., 2006. ittee Identifying and responding to problematic licensed premises-A guide for police.
53. Niland, P., McCreanor, T., Lyons, A.C. and Griffin, C. (2016). Alcohol marketing on social media: young adults engage with alcohol marketing on facebook. Addiction Research & Theory, 25(4), pp.273–284. doi:10.1080/16066359.2016.1245293.
54. Office for National Statistics (2015). Violent crime and sexual offences - alcohol-related violence - office for national statistics. [online] Ons.gov.uk. Available at: https://www.ons.gov.uk/peoplepopulationandcommunity/crimeandjustice/compendium/focusonviolentcrimeandsexualoffences/2015-02-12/chapter5violentcrimeandsexualoffencesalcoholrelatedviolence.
55. Office for National Statistics (2015b). *Violent crime and sexual offences - alcohol-related violence - office for national statistics*. [online] Ons.gov.uk. Available at: https://www.ons.gov.uk/peoplepopulationandcommunity/crimeandjustice/compendium/focusonviolentcrimeandsexualoffences/2015-02-12/chapter5violentcrimeandsexualoffencesalcoholrelatedviolence.
56. Office for National Statistics (2022). *Crime in England and Wales - Office for National Statistics*. [online] www.ons.gov.uk. Available at: https://www.ons.gov.uk/peoplepopulationandcommunity/crimeandjustice/bulletins/crimeinenglandandwales/yearendingmarch2022.
57. Office for National Statistics (2022). *Crime in England and Wales - Office for National Statistics*. [online] www.ons.gov.uk. Available at: https://www.ons.gov.uk/peoplepopulationandcommunity/crimeandjustice/bulletins/crimeinenglandandwales/yearendingmarch2022.
58. Povey, D. and Allen, J. (2003). Violent crime in England and Wales. Crime in England and Wales 2002/2003. Home Office Statistical Bulletin 07/03. London: Home Office.
59. Povey, D. and Allen, J. (2003). Violent crime in England and Wales. Crime in England and Wales 2002/2003. Home Office Statistical Bulletin 07/03. London: Home Office
60. Ramsay, M. (1982). City-Centre Crime - The Scope for Situational Prevention | Office of Justice Programs. [online] www.ojp.gov. Available at: https://www.ojp.gov/ncjrs/virtual-library/abstracts/city-centre-crime-scope-situational-prevention [Accessed 1 Oct. 2022].
61. Shepherd, J. and Brickley, M. (1996). The relationship between alcohol intoxication, stressors and injury in urban violence. British Journal of Criminology 36 (4): 546-566.
62. Sivarajasingam, V., Guan, B., Page, N., Moore, S. and Shepherd, J. (2022). *Violence in England and Wales in 2021: an accident and emergency perspective*. [online] orca.cardiff.ac.uk. Available at: https://orca.cardiff.ac.uk/id/eprint/149614/ [Accessed 6 Oct. 2022].
63. Staub, K.E. and Winkelmann, R. (2012). CONSISTENT ESTIMATION OF ZERO-INFLATED COUNT MODELS. Health Economics, 22(6), pp.673–686. doi:10.1002/hec.2844.
64. Stautz, K., Brown, K.G., King, S.E., Shemilt, I. and Marteau, T.M. (2016). Immediate effects of alcohol marketing communications and media portrayals on consumption and cognition: a systematic review and meta-analysis of experimental studies. *BMC Public Health*, [online] 16(1), p.465. doi:10.1186/s12889-016-3116-8.
65. Stephanie (2016). Poisson Regression / Regression of Counts: Definition. [online] Statistics How To. Available at: https://www.statisticshowto.com/poisson-regression/.
66. Tierney, J. and Hobbs, D. (2003). Alcohol-related Crime and Disorder Data: Guidance for Local Partnerships. Home Office Online Research Report. London: Home Office.

67. Tomsen, P.S. (2005). *Safety and Pleasure in the Night-Time Economy*. [online] papers.ssrn.com. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3073517 [Accessed 2 Oct. 2022].
68. Tutenges, S. and Bøhling, F. (2019). Designing drunkenness: How pubs, bars and nightclubs increase alcohol sales. *International Journal of Drug Policy*, 70, pp.15–21. doi:10.1016/j.drugpo.2019.04.009.
69. van Amsterdam, J.G.C., Ramaekers, J.G., Verkes, R.-J., Kuypers, K.P.C., Goudriaan, A.E. and van den Brink, W. (2019). Alcohol- and drug-related public violence in Europe. European Journal of Criminology, 17(6), pp.806–825. doi:10.1177/1477370819828324.
70. Warburton, A.L. and Shepherd, J.P. (2006). Tackling alcohol related violence in city centres: effect of emergency medicine and police intervention. *Emergency Medicine Journal*, [online] 23(1), pp.12–17. doi:10.1136/emj.2004.023028.
71. Weaver, C.G., Ravani, P., Oliver, M.J., Austin, P.C. and Quinn, R.R. (2015). Analyzing hospitalization data: potential limitations of Poisson regression. Nephrology Dialysis Transplantation, [online] 30(8), pp.1244–1249. doi:10.1093/ndt/gfv071.
72. Wells, S., Graham, K., Speechley, M. and J. Koval, J. (2005). Drinking patterns, drinking contexts and alcohol-related aggression among late adolescent and young adult drinkers. *Addiction*, [online] 100(7), pp.933–944. doi:10.1111/j.1360-0443.2005.001121.x.
73. Westgate, E.C. and Holliday, J. (2016). Identity, influence, and intervention: The roles of social media in alcohol use. Current Opinion in Psychology, 9, pp.27–32. doi:10.1016/j.copsyc.2015.10.014.
74. Wickham, M. (2012). *Alcohol Consumption in the Night-Time Economy*. ISBN 978-1-84781-521-7.
75. Wickham, M. (2012). GLA Economics - Alcohol consumption in the night time economy. London: Greater London Authority.
76. World Health Organization (2021). Adolescents: health risks and solutions. [online] Who.int. Available at: https://www.who.int/news-room/fact-sheets/detail/adolescents-health-risks-and-solutions.
77. Zach (2021). How to Interpret Log-Likelihood Values (With Examples). [online] Statology. Available at: https://www.statology.org/interpret-log-likelihood/

78. Paradis, C., Zhao, J., Joy-Goatley, S. and Stockwell, T. (2020). Original quantitative research - What popular bars post on social media platforms: a case for improved alcohol advertising regulation. Health Promotion and Chronic Disease Prevention in Canada : Research, Policy and Practice, [online] 40(5-6). Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7367423/ [Accessed 30 Oct. 2022].

79. Jones, S.C., Pettigrew, S., Biagioni, N., Daube, M., Chikritzhs, T., Stafford, J. and Tran, J. (2017), "Young adults, alcohol and Facebook: a synergistic relationship", Journal of Social Marketing, Vol. 7 No. 2, pp. 172-187.