

# EXPLORATORY DATA ANALYSIS AND DATA VISUALISATION

SUSHMA HALLIKERE HUNDI SHAMBHU

## INTRODUCTION

Exploring the Colchester policing dataset in 2023 requires analyzing crime instances and temperature variations to better understand their relationship and obtain useful insights for the police. Crime is a common problem in metropolitan areas, and understanding crime patterns is critical for law enforcement organizations to distribute resources effectively to prevent them. In 2023, Colchester, like many other cities, had various kinds of street-level crime occurrences. Analyzing this data could provide useful insights into crime trends, allowing law enforcement to make improved decisions. The datasets 'crime23.csv' and 'temp2023.csv' give street-level crime occurrences and daily climatic data from a weather station in Colchester, respectively. This study aims to conduct an in-depth data analysis of Colchester's policing dataset in 2023. This analysis, employing data visualization techniques, will provide a more thorough understanding of the association between crime events and temperature fluctuations, enabling law enforcement to make better decisions.

The datasets provided comprise important information that must be carefully organized and preprocessed before analysis. The 'context' column has been dropped and the 'date' column was converted into Date format as part of the initial steps of data preprocessing. The analysis will look at the spatial and temporal trends of crime instances, correlate crime occurrences with various temperature conditions, and identify the most common crime categories. By visualizing and analyzing the data, I intend to provide an accurate representation of street-level crime events in Colchester in 2023.

Through the analysis of temperature data, I can potentially identify relationships between variations in temperature and crime rates. Data visualization will serve a significant part in identifying insights while presenting them in an understandable form. The provided datasets provide an ample amount of information that can be used to predict crime occurrences under various temperature circumstances. Through this analysis, I will delve into the data to uncover trends, patterns, and associations, providing useful information for the authorities in Colchester.

## Importing the Libraries

# ANALYSIS OF CRIME DATA

## Load the Dataset

```
## 'data.frame': 6878 obs. of 12 variables:
## $ category      : chr "anti-social-behaviour" "anti-social-behaviour" "anti-social-behaviour" "anti-social-behaviour" ...
## $ persistent_id : chr "" "" "" ...
## $ date          : chr "2023-01" "2023-01" "2023-01" "2023-01" ...
## $ lat           : num 51.9 51.9 51.9 51.9 51.9 ...
## $ long          : num 0.909 0.902 0.898 0.902 0.895 ...
## $ street_id     : int 2153366 2153173 2153077 2153186 2153012 2153379 2153105 2153541 2152937 2153107 ...
## $ street_name    : chr "On or near Military Road" "On or near " "On or near Culver Street West" "On or near Ryegate Road" ...
## $ context        : logi NA NA NA NA NA NA ...
## $ id             : int 107596596 107596646 107595950 107595953 107595979 107595985 1075966
03 107596291 107596305 107596453 ...
## $ location_type  : chr "Force" "Force" "Force" "Force" ...
## $ location_subtype: chr "" "" "" ...
## $ outcome_status : chr NA NA NA NA ...
```

```
##           category persistent_id      date      lat      long street_id
## 1 anti-social-behaviour 2023-01 51.88306 0.909136 2153366
## 2 anti-social-behaviour 2023-01 51.90124 0.901681 2153173
## 3 anti-social-behaviour 2023-01 51.88907 0.897722 2153077
## 4 anti-social-behaviour 2023-01 51.89122 0.901988 2153186
## 5 anti-social-behaviour 2023-01 51.89416 0.895433 2153012
## 6 anti-social-behaviour 2023-01 51.88050 0.909014 2153379
##           street_name context      id location_type
## 1 On or near Military Road NA 107596596 Force
## 2 On or near                 NA 107596646 Force
## 3 On or near Culver Street West NA 107595950 Force
## 4 On or near Ryegate Road   NA 107595953 Force
## 5 On or near Market Close   NA 107595979 Force
## 6 On or near Lisle Road    NA 107595985 Force
##   location_subtype outcome_status
## 1 <NA>
## 2 <NA>
## 3 <NA>
## 4 <NA>
## 5 <NA>
## 6 <NA>
```

Removing the 'context' column from the crime dataset, which is null, results in a more organized dataset and makes subsequent analysis more efficient.

## Descriptive Statistics

```
##   category      persistent_id       date        lat
## Length:6878      Length:6878      Length:6878    Min.  :51.88
## Class :character Class :character Class :character 1st Qu.:51.89
## Mode  :character Mode  :character Mode  :character Median :51.89
##                                         Mean   :51.89
##                                         3rd Qu.:51.89
##                                         Max.   :51.90
##   long       street_id     street_name      id
## Min.  :0.8793  Min.  :2152702  Length:6878    Min.  :107582824
## 1st Qu.:0.8964 1st Qu.:2153025  Class :character 1st Qu.:109309182
## Median :0.9014 Median :2153158  Mode  :character Median :111497486
## Mean   :0.9030 Mean   :2153877                    Mean   :111301793
## 3rd Qu.:0.9088 3rd Qu.:2153365                    3rd Qu.:113746477
## Max.   :0.9246 Max.   :2343256                    Max.   :115699577
##   location_type  location_subtype outcome_status
## Length:6878      Length:6878      Length:6878
## Class :character Class :character Class :character
## Mode  :character Mode  :character Mode  :character
## 
## 
```

The crime dataset's descriptive statistics outline the frequency and distribution of crime incidents by category, as well as information on the period illustrated geographical coordinates, and the street names where the crimes occurred, allowing for a more complete understanding of the dataset's structure and the types of crimes recorded.

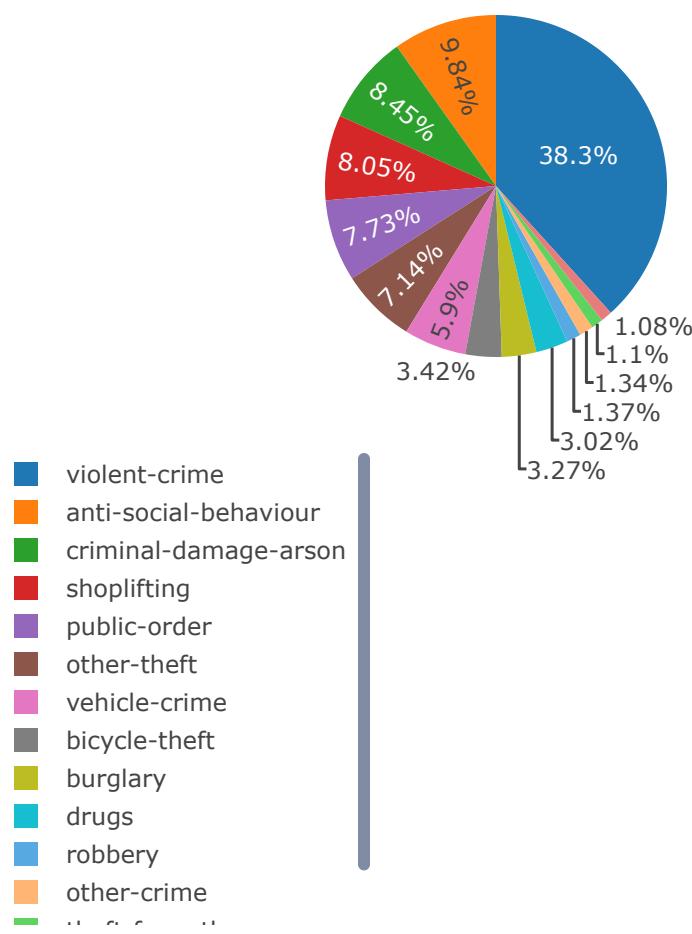
## Two-way table for Category Frequency

##	2023-01-01	2023-02-01	2023-03-01	2023-04-01	2023-05-01
## anti-social-behaviour	46	49	21	53	67
## bicycle-theft	20	14	19	16	16
## burglary	17	22	14	22	15
## criminal-damage-arson	59	37	52	63	64
## drugs	14	17	21	21	22
## other-crime	7	5	6	15	3
## other-theft	48	37	35	38	42
## possession-of-weapons	3	3	11	5	7
## public-order	45	42	58	51	37
## robbery	8	7	8	7	7
## shoplifting	76	31	51	40	51
## theft-from-the-person	6	7	12	7	5
## vehicle-crime	65	15	21	29	24
## violent-crime	237	181	226	207	226
##	2023-06-01	2023-07-01	2023-08-01	2023-09-01	2023-10-01
## anti-social-behaviour	52	76	71	90	68
## bicycle-theft	14	15	21	37	26
## burglary	26	14	20	18	31
## criminal-damage-arson	42	42	33	47	45
## drugs	15	17	7	25	19
## other-crime	11	12	9	7	6
## other-theft	41	51	41	34	49
## possession-of-weapons	3	8	5	8	6
## public-order	36	40	41	45	52
## robbery	17	6	5	8	9
## shoplifting	59	33	57	33	43
## theft-from-the-person	6	9	5	7	3
## vehicle-crime	45	25	16	20	26
## violent-crime	196	236	219	263	209
##	2023-11-01	2023-12-01			
## anti-social-behaviour	39	45			
## bicycle-theft	27	10			
## burglary	11	15			
## criminal-damage-arson	53	44			
## drugs	13	17			
## other-crime	5	6			
## other-theft	37	38			
## possession-of-weapons	8	7			
## public-order	45	40			
## robbery	5	7			
## shoplifting	39	41			
## theft-from-the-person	4	5			
## vehicle-crime	56	64			
## violent-crime	221	212			

The two-way table for category frequency over 12 months produces a table depicting the distribution of various crime categories over 12 months. The data illustrates how crime instances vary over time, with each row indicating a specific crime category and each column denoting a month.

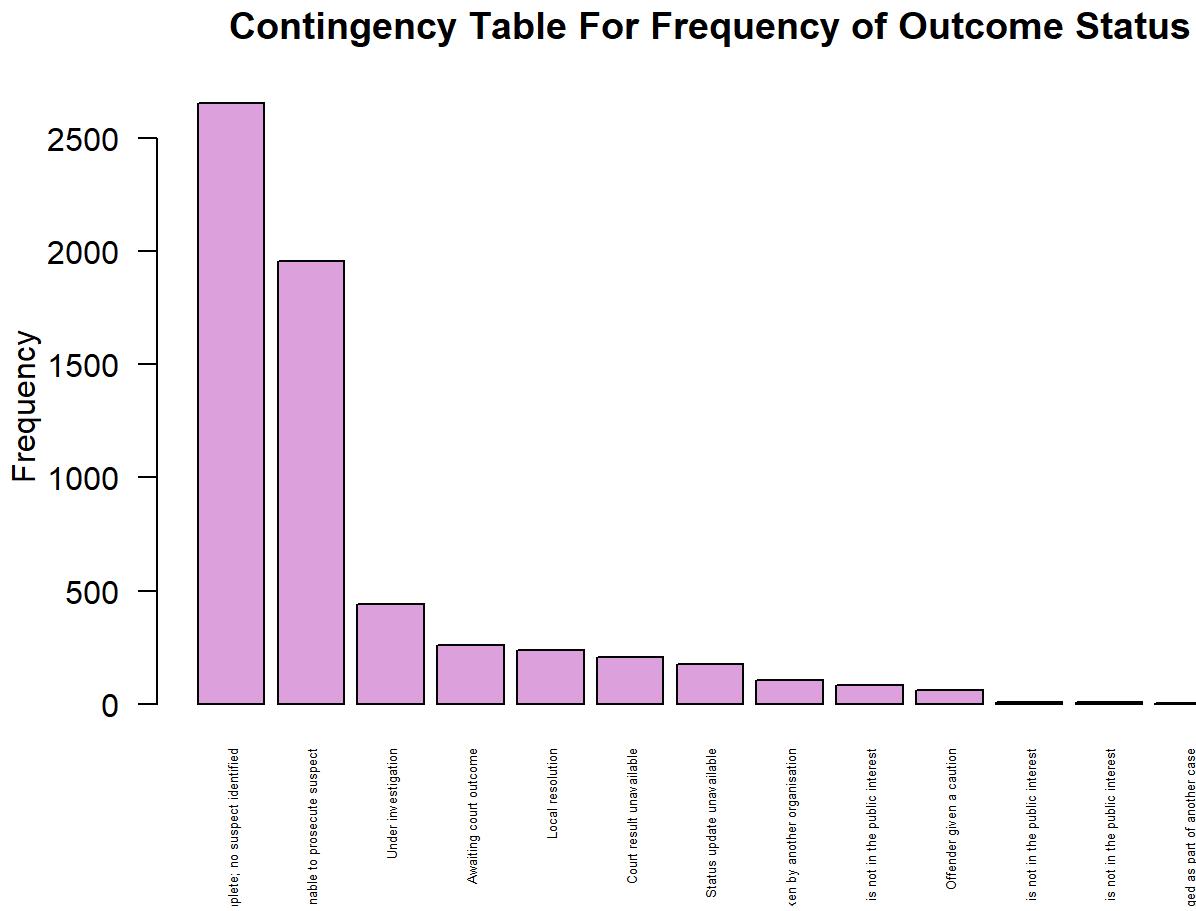
# Pie Chart of Frequency of Category For 2023

## Cross-Tabulation For Frequency of Category



The pie chart demonstrating crime categories in Colchester shows a wide range of reported incidents. Criminal damage/arsenal emerges as the most common offence, accounting for roughly 18% of all reported offences, highlighting worries about property vandalism and destruction. Anti-social behaviour follows closely behind, accounting for around 21% of instances, emphasising issues with community cohesiveness and public nuisance offences. Other theft offences account for around 15% of reported crimes, highlighting continued concerns about property theft and loss prevention. Public order offences and violent crimes, while less common, remain important, accounting for around 16% and 10% of reported cases, respectively. Together, these data provide a comprehensive picture of crime trends in Colchester, allowing stakeholders to develop targeted initiatives to improve public safety and address community concerns.

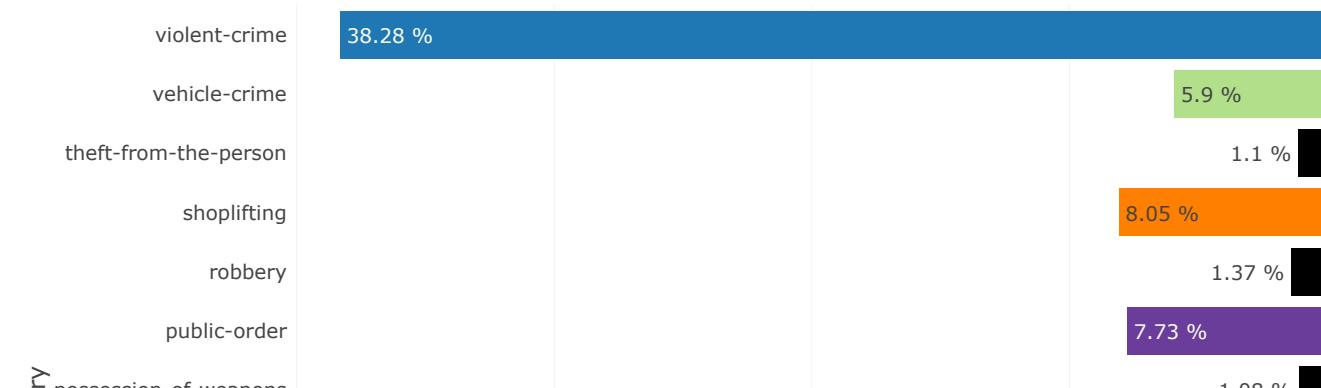
# Bar Graph of Frequency of Outcome Status For 2023

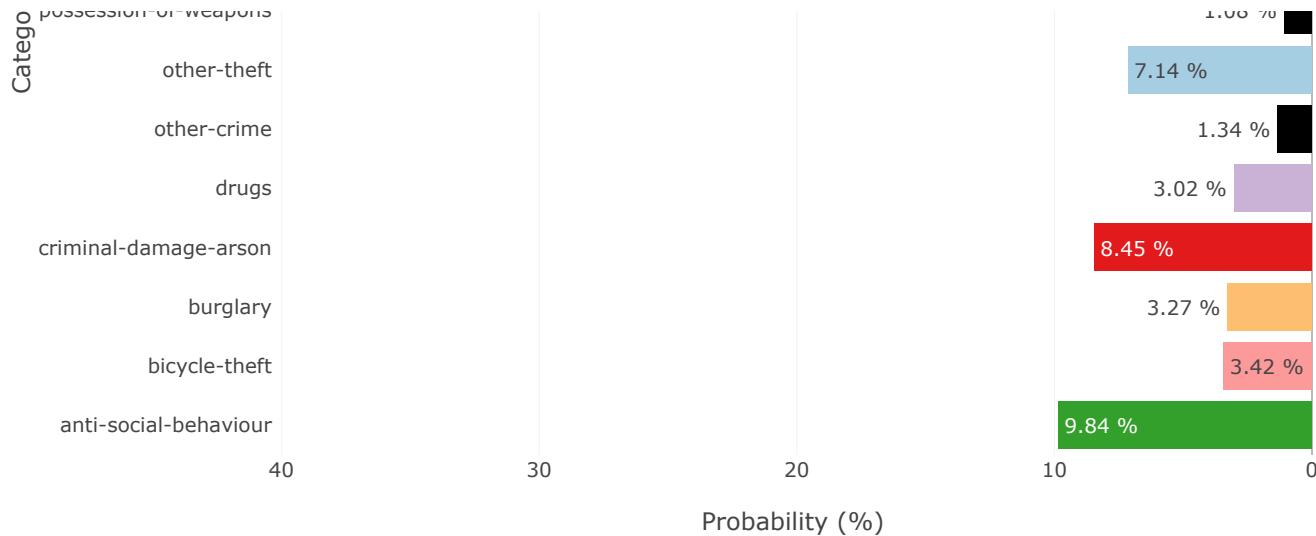


The bar graph representing crime outcomes in Colchester, England, during 2023 gives information about the status of recorded events. The most common outcome, “Investigation complete; no suspect identified,” which includes 2,656 cases, highlights the difficulties law enforcement faced in identifying criminals and completing investigations. Following closely behind is the conclusion “Unable to prosecute suspect,” which has 1,959 cases and suggests that prosecution is hindered by insufficient evidence or a lack of public interest. The graph also indicates other important outcomes, such as cases under investigation (4,39), those awaiting court outcomes (2,60), and cases concluded locally (2,39). Overall, these findings provide insight into the complexity involved in dealing with crime and implementing justice, giving significant insights for law enforcement agencies and legislators in improving investigative processes and strengthening the community.

## Probability of Each Category Frequency

Probability of Crime Category Frequency

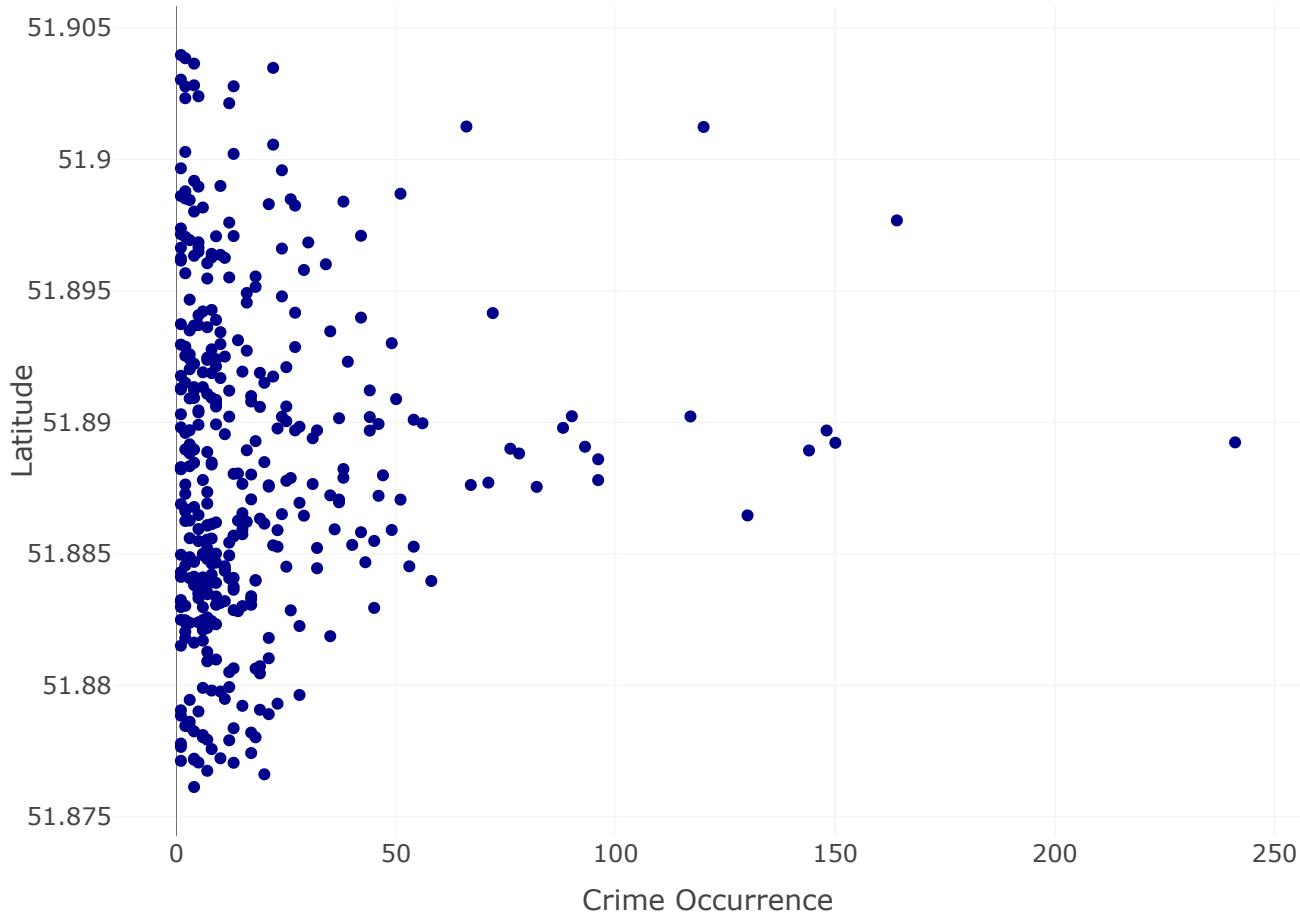




Horizontal bar charts are useful for comparing probabilities across multiple categories. The graph shows the likelihood of each crime category occurring. "Violent Crime" has the highest probability of 38.28%, followed by "Anti-social Behaviour" (9.84%) and "Criminal Damage/Arson" (8.45%). "Other crime" and "possession of weapons" had the lowest probabilities, at 1.08%. This figure indicates that the most likely crime categories are violent crime, anti-social behavior, and criminal damage/arson, which could be attributable to the impact of these sorts of crimes in the area or the reporting rates for these crimes.

## Grouping of Latitude and Longitude

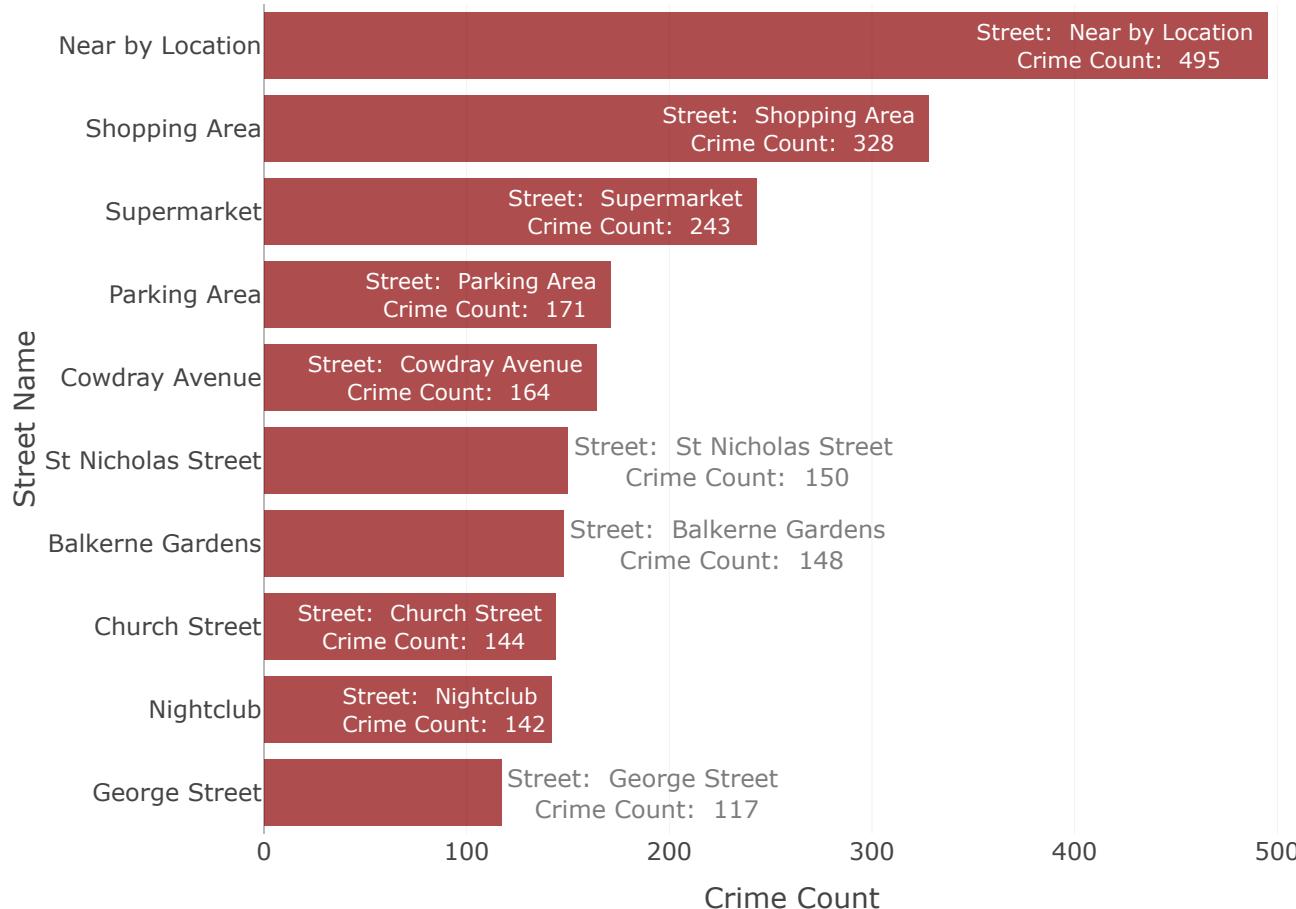
Latitude and Longitude of Crime Frequency



The graph shows the crime frequency at various latitudes and longitudes. The x-axis indicates the crime count, and the y-axis shows the latitude. Each point on the scatter plot represents a specific location, and its color corresponds to the number of crimes committed there. The particular inference from this scatter plot is that the darker blue spots represent places with a high crime rate over the land areas.

## Analysis of the frequency of Crime Category over Different Streets

Top 10 Streets with Most Crime Frequency



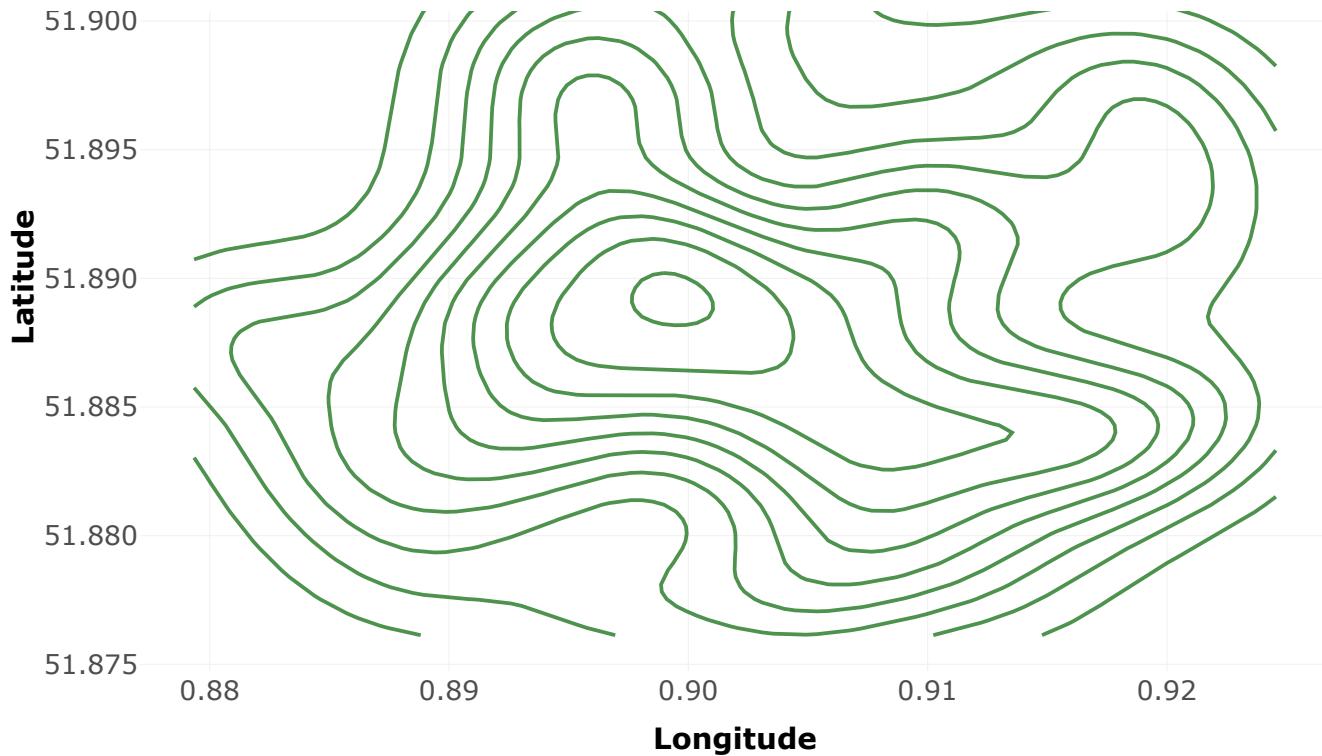
The graph illustrates the top ten streets with the highest crime rates. The y-axis depicts the street name, while the x-axis reflects the crime count. "Nearby Location" which is closer to the "Force" or "BTP" has the highest crime frequency, with 495 offenses. The "Supermarket" (243 crimes), the "Shopping Area" (328 crimes), and the "Parking Area" (171 crimes) are followed, and "George Street" is the street on the list with the fewest crimes (117 crimes). This graph is useful for determining which streets in a city have the most crime. This information can help authorities focus their patrols and residents make wise decisions regarding where they want to reside and stroll. It is vital to remember that this only includes the top ten streets with the highest crime rates. There could be more streets in the city with a high crime rate. In addition, the frequency of crime on a certain street may change over time.

## Visualization of Areas of More Concentrated Crimes

Crime Frequency over Latitude and Longitude

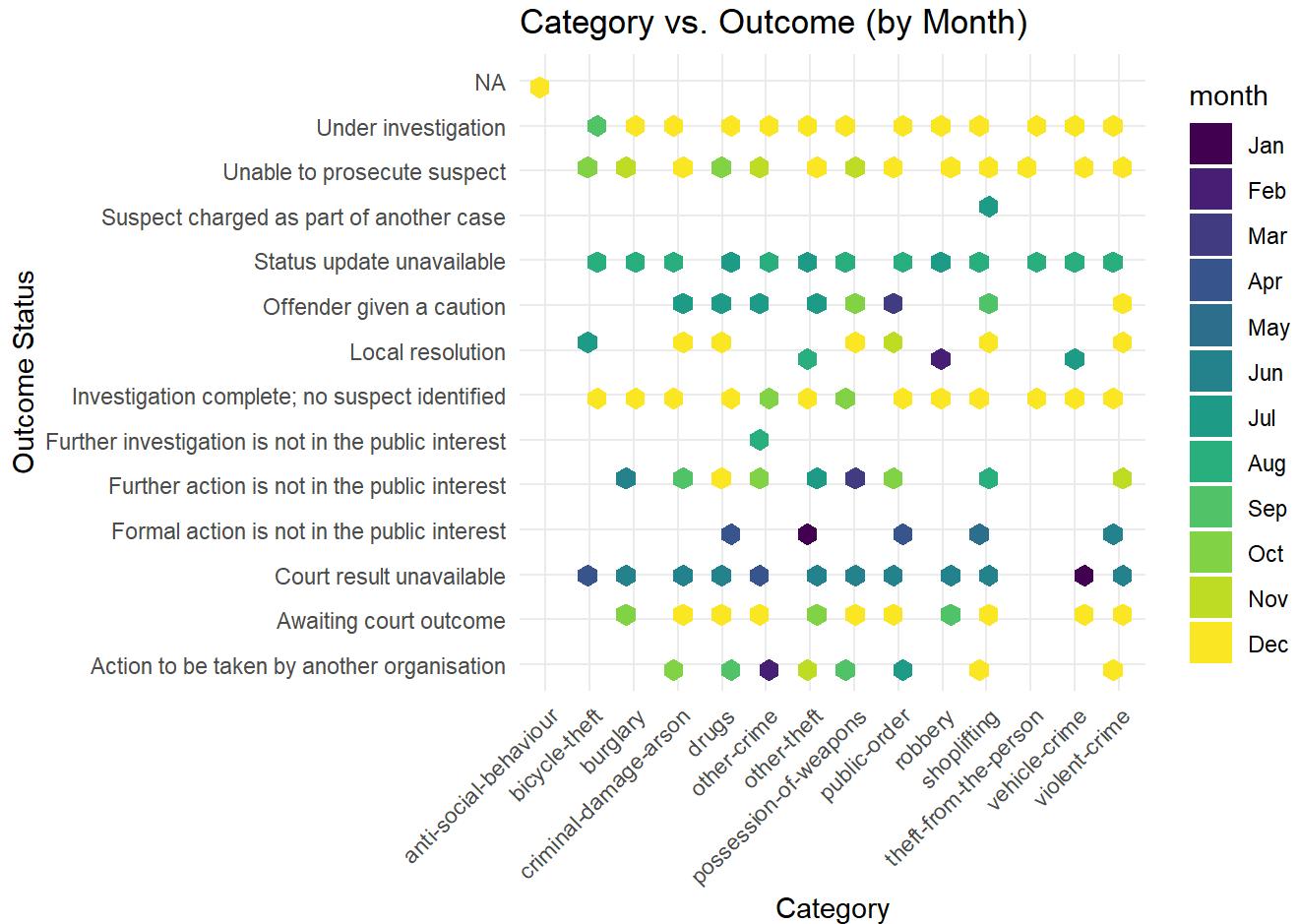
51.905





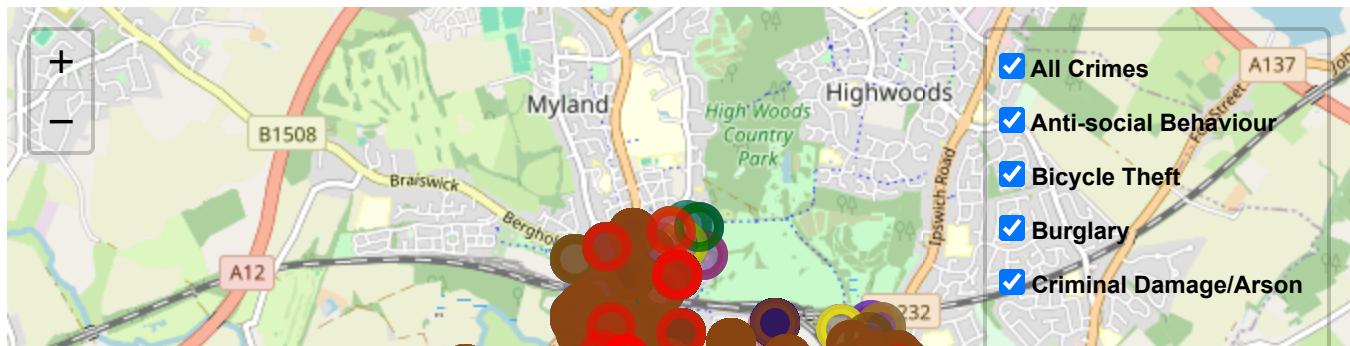
The density map depicts the distribution of crime frequency across latitude and longitude. Darker green regions imply higher crime concentrations. Clusters of criminal activity are seen throughout the area, indicating locations with greater crime rates. Darker green zones may suggest high-density areas such as city centers or areas with vibrant nightlife, but specific sites such as bus or train stations may also contribute to concentrated criminal activity. Density plots, such as this one, are useful for visualizing the distribution of data over a continuous range and identifying locations that require specific crime prevention strategies.

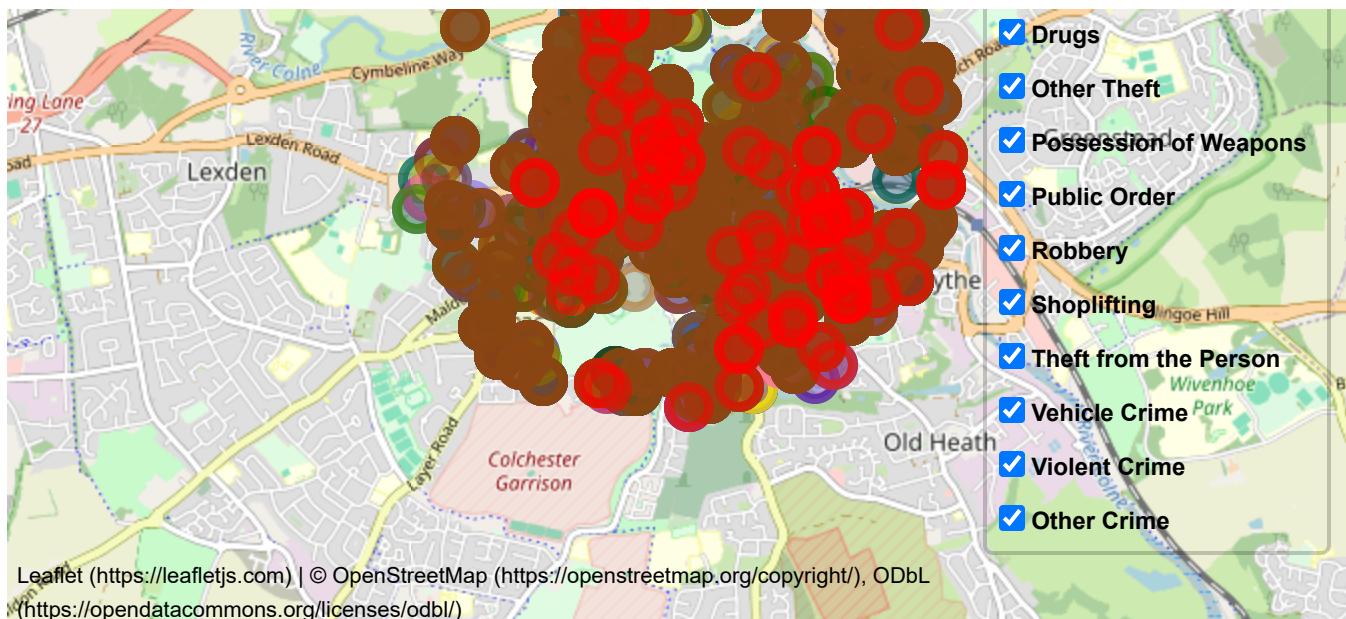
# Visualization of Relationship between Category vs. Outcome over the Month



Hex bin plots are excellent for visualizing correlations between categorical variables and detecting patterns in data that other forms of plots may not reveal. The hexagon's colour corresponds to the month, with lighter colours representing earlier months of the year (January) and darker colours signifying later months (December). The graph highlights stealing and burglary as the most common crime categories across all months, as evidenced by the bigger clusters of hexagons on the left side of the plot. Some crime categories also appear to follow a seasonal trend, such as an increase in crime during the summer months (July-August), which could be linked to increased outdoor activity or alcohol use. In addition, certain outcome statuses are more prevalent in specific crime categories. For example, "Under investigation" is a common outcome state for theft and shoplifting, whereas "Awaiting court outcome" is a common outcome status for drugs and other crimes. Overall, the hex bin plot is an effective tool for visualizing the association between crime category, outcome status, and month in 2023.

## Geospatial Analysis



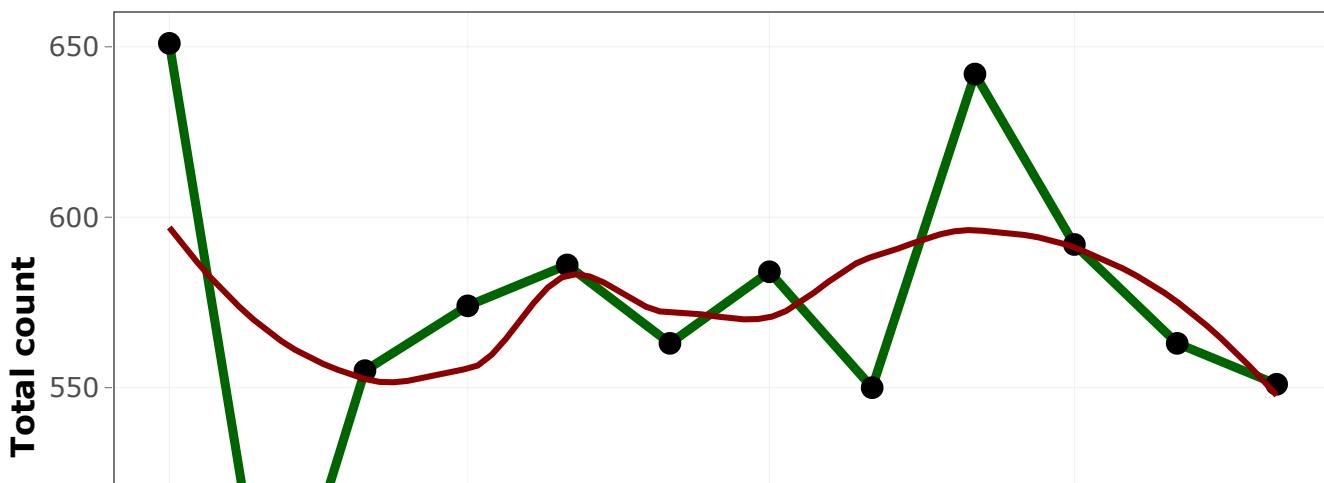


The map illustrates the distribution of reported crimes in Colchester, England in 2023. Each circle on the map symbolizes a distinct form of crime, with the color signifying the crime and the size representing the number of reported incidences. The visualization shows that train stations, supermarkets, deserted areas, Abbey Field, Berechurch Road, and the surroundings of Colchester Cemetery have the greatest crime rates. The color coding of the circles shows that the most common types of crime in Colchester are violent crime, anti-social behavior, criminal-arson-damage, shoplifting, and theft. This visualization provides essential insights for law enforcement organizations to optimize patrol approaches, as well as for citizens to make informed decisions about their own safety and living arrangements.

Several different solutions can be considered when resolving the crime issues identified by the visualization. These include enhancing police presence in high-risk locations to prevent criminal activity and improve the likelihood of capturing criminals. Improving illumination in public places can also help reduce crime by making it more difficult for thieves to go undetected. Furthermore, promoting community policing programs and conducting social initiatives targeted at addressing root causes such as poverty and a lack of education can help with long-term crime prevention efforts. The visualization is an effective tool for understanding and tackling crime in Colchester, paving the way for a better and more secure city.

## Time Series Analysis for Crime Frequency

### Overall Crime Frequency Over 2023





The time series plot illustrates the variability in overall crime frequency throughout 2023 in Colchester, England, demonstrating an evident seasonal pattern with high in the summer and declines in the winter. The crime rate was highest in July, and lowest in January. The smoothed trend line on the map helps to discover underlying trends that are disguised by month-to-month variability, demonstrating a year-over-year increase in crime rates. This information can help law enforcement agencies allocate resources and develop patrol techniques, as well as help residents make informed choices about safety measures. Time series charts are useful tools for understanding crime dynamics and anticipating future trends in metropolitan environments.

The crime dataset study provides useful insights into crime patterns and dynamics in Colchester, England, in 2023. Several significant conclusions arise from visualizations such as maps, time series plots, and scatter plots. There is evidence of seasonal variation, with greater crime rates seen in the summer and lower rates in the winter. There are some neighborhoods with greater crime rates than others, like Abbey Field and Berechurch Road. The most common offenses include violent crime, anti-social behavior, criminal-arson-damage, and shoplifting. Law enforcement can use these findings to improve patrol strategy and resource allocation, potentially lowering crime rates. Furthermore, community initiatives and social programs focused on tackling underlying causes of crime, such as poverty and a lack of education, could supplement law enforcement operations. Overall, by harnessing data-driven insights and executing focused interventions, Colchester can help to create a safer and more secure environment for its citizens.

# ANALYSIS OF TEMPERATURE DATA

## Load the Dataset

```
## 'data.frame': 365 obs. of 18 variables:
## $ station_ID : int 3590 3590 3590 3590 3590 3590 3590 3590 3590 3590 ...
## $ Date       : chr "31-12-2023" "30-12-2023" "29-12-2023" "28-12-2023" ...
## $ TemperatureCAvg: num 8.7 6.6 9.9 9.9 5.8 9.8 12.5 10 9.6 10 ...
## $ TemperatureCMax: num 10.6 9.7 11.4 11.5 10.6 12.7 14.3 12 10.8 12.6 ...
## $ TemperatureCMin: num 4.4 4.4 6.9 4 3.9 6.3 9.5 8.4 8.1 8.1 ...
## $ TdAvgC      : num 7.2 4.2 6 7.5 3.7 7.6 10.1 7 6.5 6.2 ...
## $ HrAvg        : num 89.6 85.5 77.2 84.6 86.4 86.9 85.3 81.5 81.2 78.2 ...
## $ WindkmhDir   : chr "S" "WSW" "SW" "SSW" ...
## $ WindkmhInt   : num 25 22.7 32.8 32.2 13.2 23.5 34.1 32.7 34.1 37.5 ...
## $ WindkmhGust  : num 63 50 61.2 70.4 37.1 46.3 72.3 61.2 68.6 77.8 ...
## $ PresslevHp   : num 999 1007 1004 1003 1016 ...
## $ Precmm       : num 6.2 0.4 0.8 2.8 2 4.4 0.8 0.8 0 2 ...
## $ TotClOct    : num 8 4.6 6.5 6.8 4 6.5 7.8 5 8 7.5 ...
## $ lowClOct    : num 8 6.5 6.7 7.1 6.9 7.4 7.8 6.7 8 7.5 ...
## $ SunD1h       : num 0 1.1 0.1 0 3.2 0 0 2.9 0 1.4 ...
## $ VisKm        : num 26.3 48.3 26.7 25.1 30.1 45.8 61.8 72.9 69.4 34.3 ...
## $ PreselevHp  : logi NA NA NA NA NA NA ...
## $ SnowDepcm   : int NA ...
```

```
##   station_ID     Date TemperatureCAvg TemperatureCMax TemperatureCMin TdAvgC
## 1      3590 31-12-2023          8.7          10.6          4.4    7.2
## 2      3590 30-12-2023          6.6          9.7          4.4    4.2
## 3      3590 29-12-2023          9.9         11.4          6.9    6.0
## 4      3590 28-12-2023          9.9         11.5          4.0    7.5
## 5      3590 27-12-2023          5.8         10.6          3.9    3.7
## 6      3590 26-12-2023          9.8         12.7          6.3    7.6
##   HrAvg WindkmhDir WindkmhInt WindkmhGust PresslevHp Precmm TotClOct lowClOct
## 1  89.6         S      25.0      63.0     999.0    6.2     8.0    8.0
## 2  85.5       WSW      22.7      50.0    1006.9    0.4     4.6    6.5
## 3  77.2        SW      32.8      61.2    1003.6    0.8     6.5    6.7
## 4  84.6       SSW      32.2      70.4    1003.2    2.8     6.8    7.1
## 5  86.4        SW      13.2      37.1    1016.4    2.0     4.0    6.9
## 6  86.9       WSW      23.5      46.3    1006.2    4.4     6.5    7.4
##   SunD1h VisKm PreselevHp SnowDepcm
## 1     0.0  26.3        NA       NA
## 2     1.1  48.3        NA       NA
## 3     0.1  26.7        NA       NA
## 4     0.0  25.1        NA       NA
## 5     3.2  30.1        NA       NA
## 6     0.0  45.8        NA       NA
```

Removing the 'PreselevHp' column from the temperature dataset, which is null, results in a more organized dataset and makes subsequent analysis more efficient.

## Descriptive Statistics

```

##      station_ID      Date      TemperatureCAvg TemperatureCMax
##  Min.   :3590  Length:365      Min.   :-2.60  Min.   : 1.70
##  1st Qu.:3590  Class  :character  1st Qu.: 7.20  1st Qu.:10.60
##  Median :3590  Mode   :character  Median :10.40  Median :14.20
##  Mean   :3590                  Mean   :10.92  Mean   :15.13
##  3rd Qu.:3590                  3rd Qu.:15.80 3rd Qu.:20.00
##  Max.   :3590                  Max.   :23.10  Max.   :30.40
##
##      TemperatureCMin      TdAvgC      HrAvg      WindkmhDir
##  Min.   :-6.200  Min.   :-4.400  Min.   :43.10  Length:365
##  1st Qu.: 3.200  1st Qu.: 4.400  1st Qu.:75.60  Class  :character
##  Median : 6.300  Median : 7.600  Median :81.70  Mode   :character
##  Mean   : 6.365  Mean   : 7.578  Mean   :81.25
##  3rd Qu.:10.600 3rd Qu.:11.200 3rd Qu.:87.90
##  Max.   :16.300  Max.   :17.500  Max.   :97.90
##
##      WindkmhInt      WindkmhGust      PresslevHp      Precmm
##  Min.   : 6.20  Min.   :13.00  Min.   : 967.4  Min.   : 0.000
##  1st Qu.:12.00  1st Qu.:31.50  1st Qu.:1006.3  1st Qu.: 0.000
##  Median :16.10  Median :38.90  Median :1014.3  Median : 0.000
##  Mean   :16.81  Mean   :40.87  Mean   :1013.6  Mean   : 1.866
##  3rd Qu.:20.20  3rd Qu.:48.20  3rd Qu.:1021.7  3rd Qu.: 1.150
##  Max.   :37.50  Max.   :98.20  Max.   :1045.1  Max.   :33.600
##                               NA's   :27
##      TotClOct      lowClOct      SunD1h      VisKm      SnowDepcm
##  Min.   :0.000  Min.   :1.800  Min.   : 0.000  Min.   : 3.60  Min.   :1
##  1st Qu.:3.600  1st Qu.:5.800  1st Qu.: 1.150  1st Qu.:22.70  1st Qu.:1
##  Median :5.100  Median :6.700  Median : 4.700  Median :31.50  Median :1
##  Mean   :4.988  Mean   :6.443  Mean   : 5.127  Mean   :32.11  Mean   :1
##  3rd Qu.:7.000  3rd Qu.:7.400  3rd Qu.: 8.050  3rd Qu.:41.50  3rd Qu.:1
##  Max.   :8.000  Max.   :8.000  Max.   :15.400  Max.   :72.90  Max.   :1
##                               NA's   :13    NA's   :82    NA's   :364

```

The summary statistics of temperature data for 2023 in Colchester give some significant insights into the weather conditions throughout the year. The average daily temperature was around 10.92°C, with fluctuations ranging from -2.60°C to 23.10°C. Humidity levels were moderate, with an average dew point temperature of 7.578°C and a relative humidity of 81.25%. Wind speeds averaged 16.81 km/h, with rare gusts up to 98.20 km/h. Atmospheric pressure averaged 1013.6 hPa, with precipitation averaging 1.866 mm. Cloud cover and visibility were variable, with an average sunshine duration per hour of 5.127 hours and an average visibility of 32.11 km. Interestingly, snow depth was continuously reported as 1 cm throughout the year, indicating scant snow accumulation. These statistics provide vital insights into Colchester's weather patterns and conditions in 2023, establishing the framework for future research and visualization of seasonal trends.

## Two - way table of Weather Conditions over Month

```

##      Month Weather Condition No of days the weather is cold
## 1     April    Cold           29
## 2     August   Cold            0
## 3    December  Cold           27
## 4    February Cold           26
## 5    January  Cold           29
## 6     July    Cold            0
## 7     June    Cold            0
## 8    March    Cold           23
## 9     May    Cold             5
## 10   November Cold           24
## 11   October  Cold            9
## 12   September Cold            0
## 13   April   Moderate        1
## 14   August  Moderate       31
## 15   December Moderate        4
## 16   February Moderate        2
## 17   January  Moderate        2
## 18     July  Moderate       31
## 19     June  Moderate       30
## 20   March   Moderate        8
## 21     May  Moderate       26
## 22   November Moderate        6
## 23   October  Moderate       22
## 24   September Moderate       30
## 25   April   Hot              0
## 26   August  Hot              0
## 27   December Hot              0
## 28   February Hot              0
## 29   January  Hot              0
## 30     July  Hot              0
## 31     June  Hot              0
## 32   March   Hot              0
## 33     May  Hot              0
## 34   November Hot              0
## 35   October  Hot              0
## 36   September Hot              0

##      No of days the weather is not cold
## 1                  1
## 2                  2
## 3                  3
## 4                  4
## 5                  5
## 6                  6
## 7                  7
## 8                  8
## 9                  9
## 10                 10
## 11                 11
## 12                 12

```

```

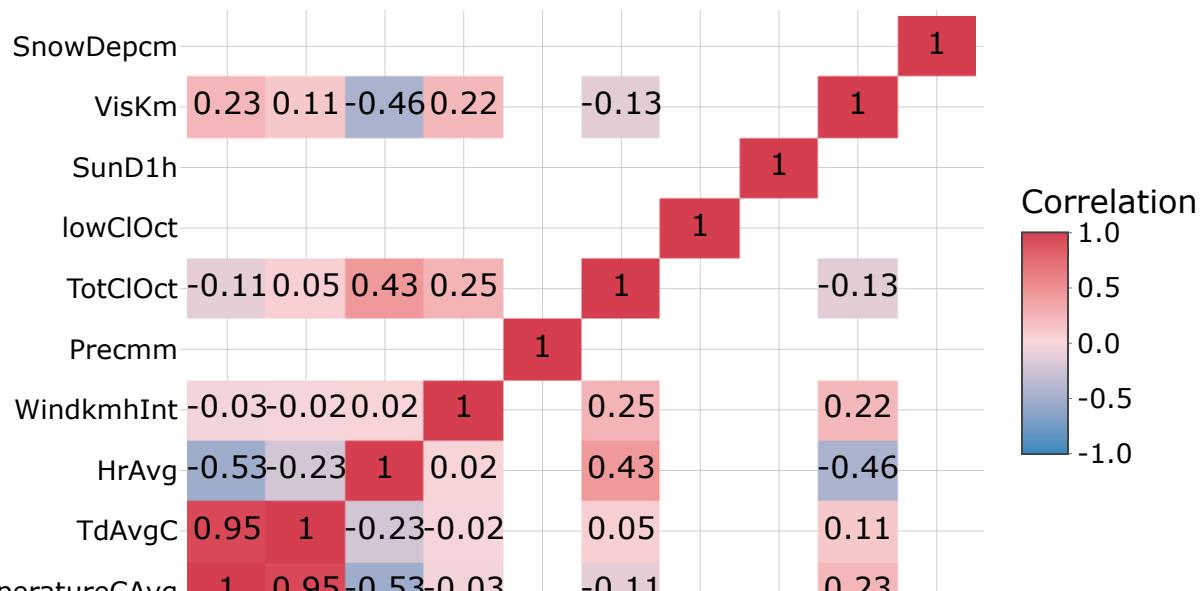
## 13          13
## 14          14
## 15          15
## 16          16
## 17          17
## 18          18
## 19          19
## 20          20
## 21          21
## 22          22
## 23          23
## 24          24
## 25          25
## 26          26
## 27          27
## 28          28
## 29          29
## 30          30
## 31          31
## 32          32
## 33          33
## 34          34
## 35          35
## 36          36

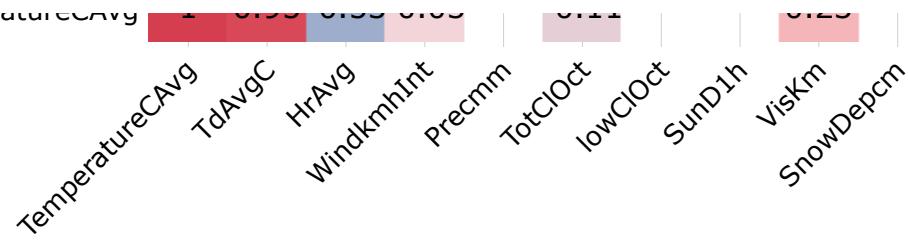
```

A two-way table analysing weather conditions in Colchester throughout various months in 2023 reveals interesting trends in temperature distribution throughout the year. January, February, and December have the highest proportion of cold days, showing that lower temperatures are more common during the winter. As the temperature shifts between seasons, however, transitional months such as March, April, and November exhibit a combination of days that are chilly and days that are not. This analysis provides insight into seasonal fluctuations and trends in temperature distribution, helping to improve our understanding of Colchester's climatic dynamics.

## Correlation Analysis for Temperature Variations

### Correlation Matrix of Average Temperature Change

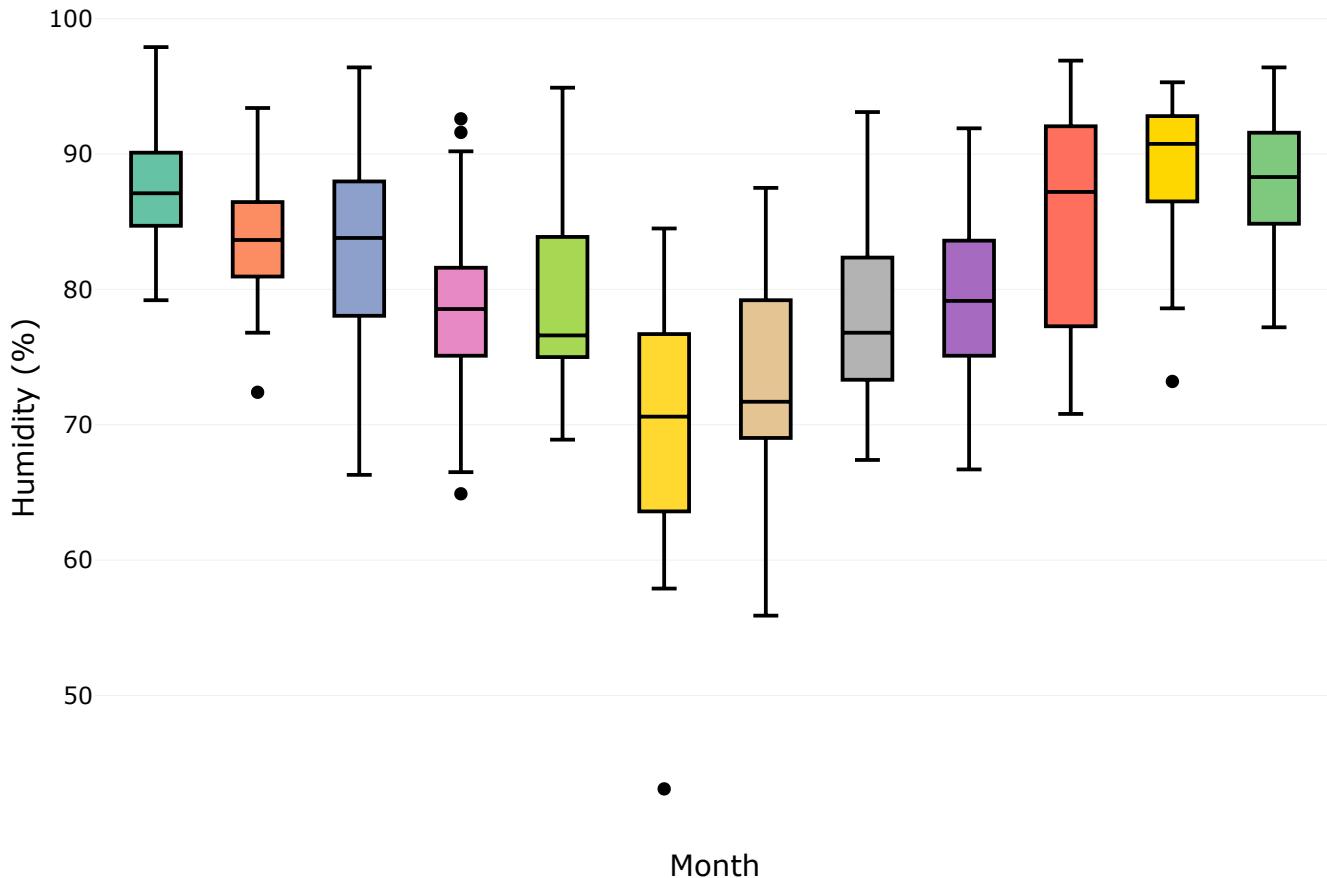




The correlation matrix shows the similarities between several weather indicators gathered in Colchester, England in 2023. Key observations show substantial connections between temperature and dew point temperature, demonstrating that they are interdependent due to moisture content in the air. Furthermore, temperature has a negative link with humidity, and wind speed has a negative correlation with sunshine length, indicating weather dynamics related to gloomy conditions. Precipitation correlates positively with cloud cover, indicating that cloud formation and moisture saturation contribute to precipitation occurrences. In contrast, snow depth has a negative association with temperature, showing that snow melts as temperatures rise. Overall, the correlation matrix helps in comprehending the complex relationships between different weather characteristics, which assists climate researchers and weather forecasters in anticipating local climate patterns.

## Humidity Distribution over 12 months

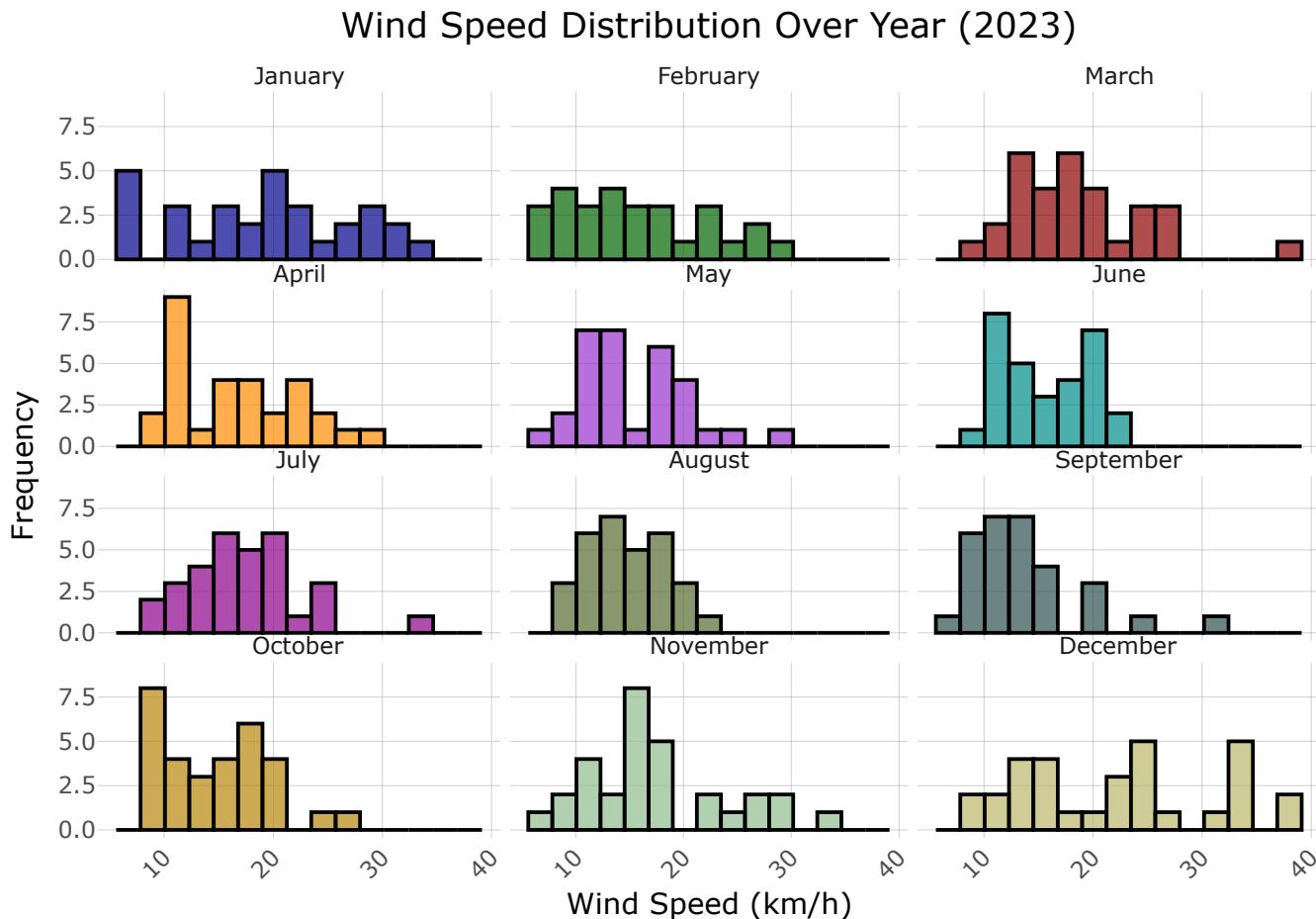
Average Humidity Distribution Over 12 Months



The box plot analysis demonstrates significant seasonal patterns. Humidity levels rise significantly during the summer months, particularly June through August, correlating with higher temperatures and increasing evaporation rates. Humidity levels fall during the winter months, particularly from December to February, corresponding to cooler temperatures and lower moisture content in the air. The average humidity is highest in

July, and lowest in December. The box plot also shows a broader range of humidity values throughout the summer months, indicating greater unpredictability than the more steady levels seen in the winter. Overall, the box plot provides useful information about the seasonal dynamics of humidity in Colchester, which aids in understanding weather patterns and their effects throughout the year.

## Wind Speed Distribution



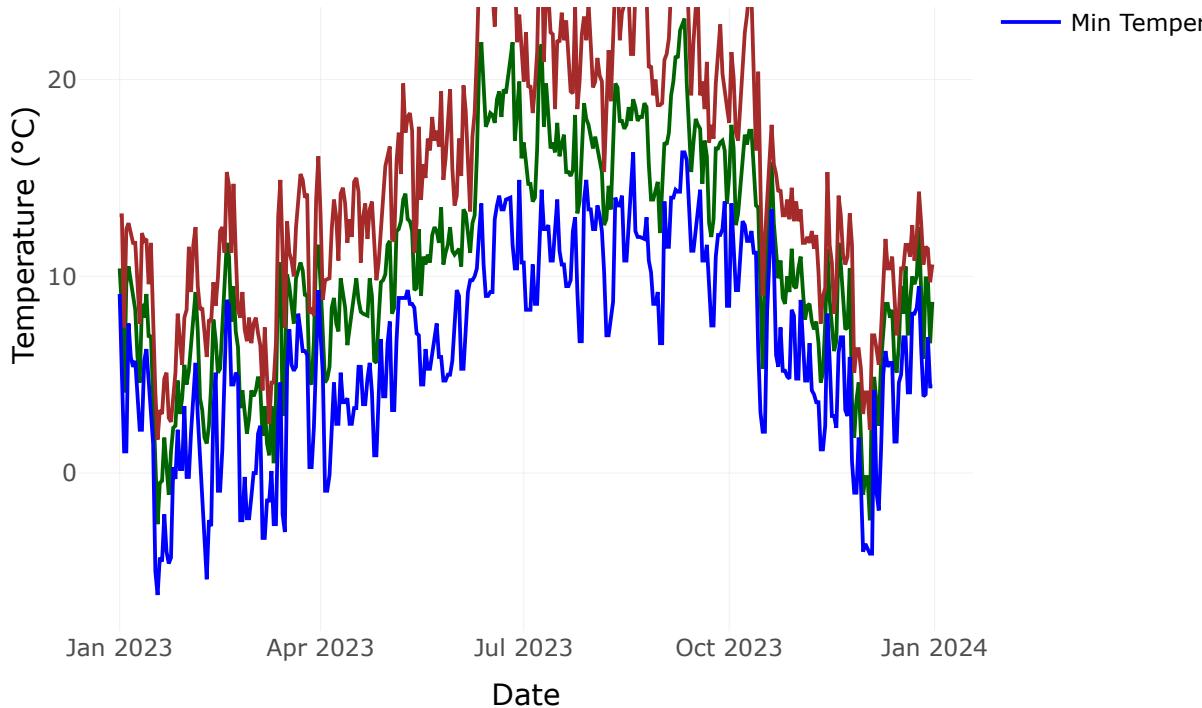
The faceted histograms provide informative observations on seasonal and monthly fluctuations. Notably, wind speed distributions change between months, with some months presenting symmetrical distributions and others exhibiting skewness towards higher speeds. For example, the histogram for July shows a right-skewed distribution, indicating an increase in the occurrence of greater wind speeds that month. Furthermore, a continuous tendency appears, with stronger wind speeds during the winter months (December–February) than the summer months (June–August), which corresponds to usual weather patterns in temperate locations. Furthermore, the histograms highlight the diversity and intensity of wind conditions throughout the colder months by showing a greater spread of bars, which indicates a broader range of wind speeds in winter.

## Temperature Trend for the Year 2023

### Temperature Trend in 2023



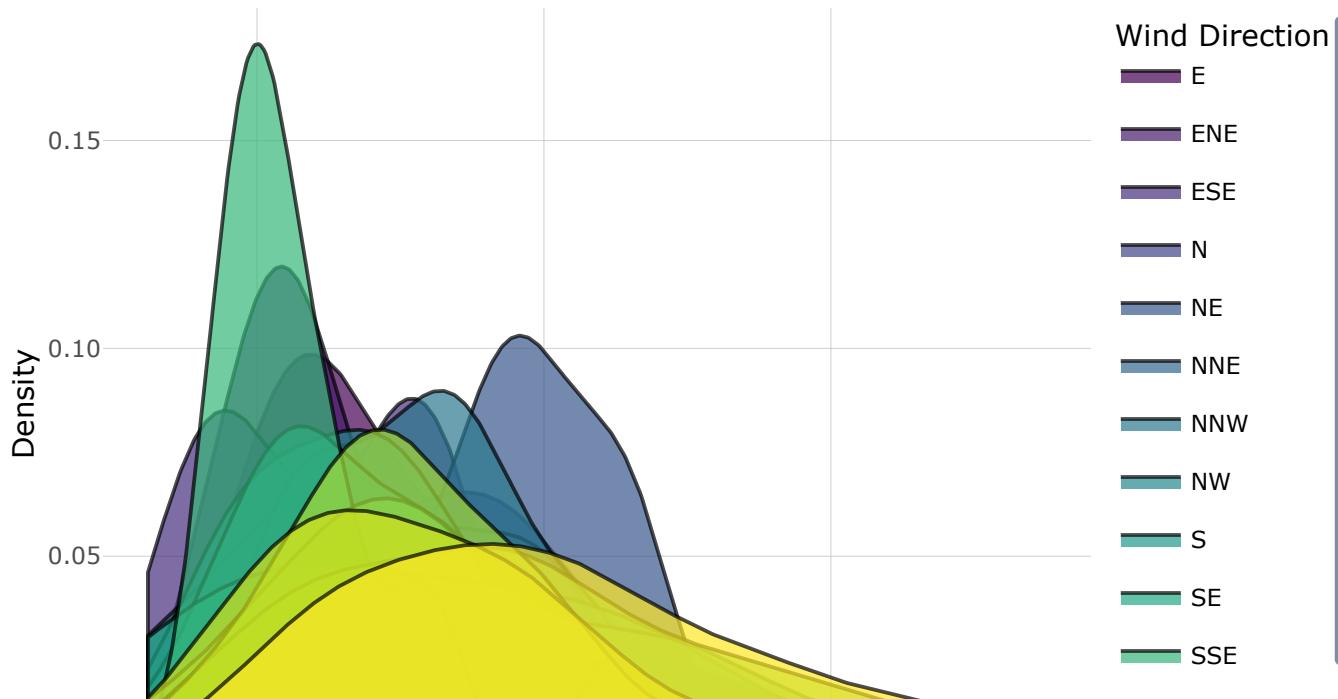
Min Temperature

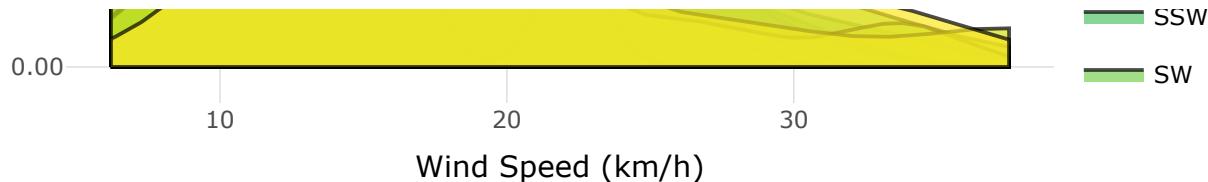


Line graphs successfully capture temporal trends, offering useful information about temperature dynamics throughout the year. The line graph shows the average, maximum, and minimum daily temperatures in Colchester, England, during the year 2023. In particular, a strong seasonal pattern appears, with temperatures rising in the summer (June-August) and decreasing in the winter (December-February). The graph shows that the highest temperature consistently exceeds the average temperature, and the minimum temperature consistently goes below the average, as expected given the computation method. Furthermore, the daily temperature range varies, with larger ranges reported in the summer months compared to narrower ranges in the winter. This pattern demonstrates how sunshine length and seasonal variations affect temperature fluctuations.

## Analysis of Distribution of Wind Direction

### Analysis of Distribution of Wind Direction

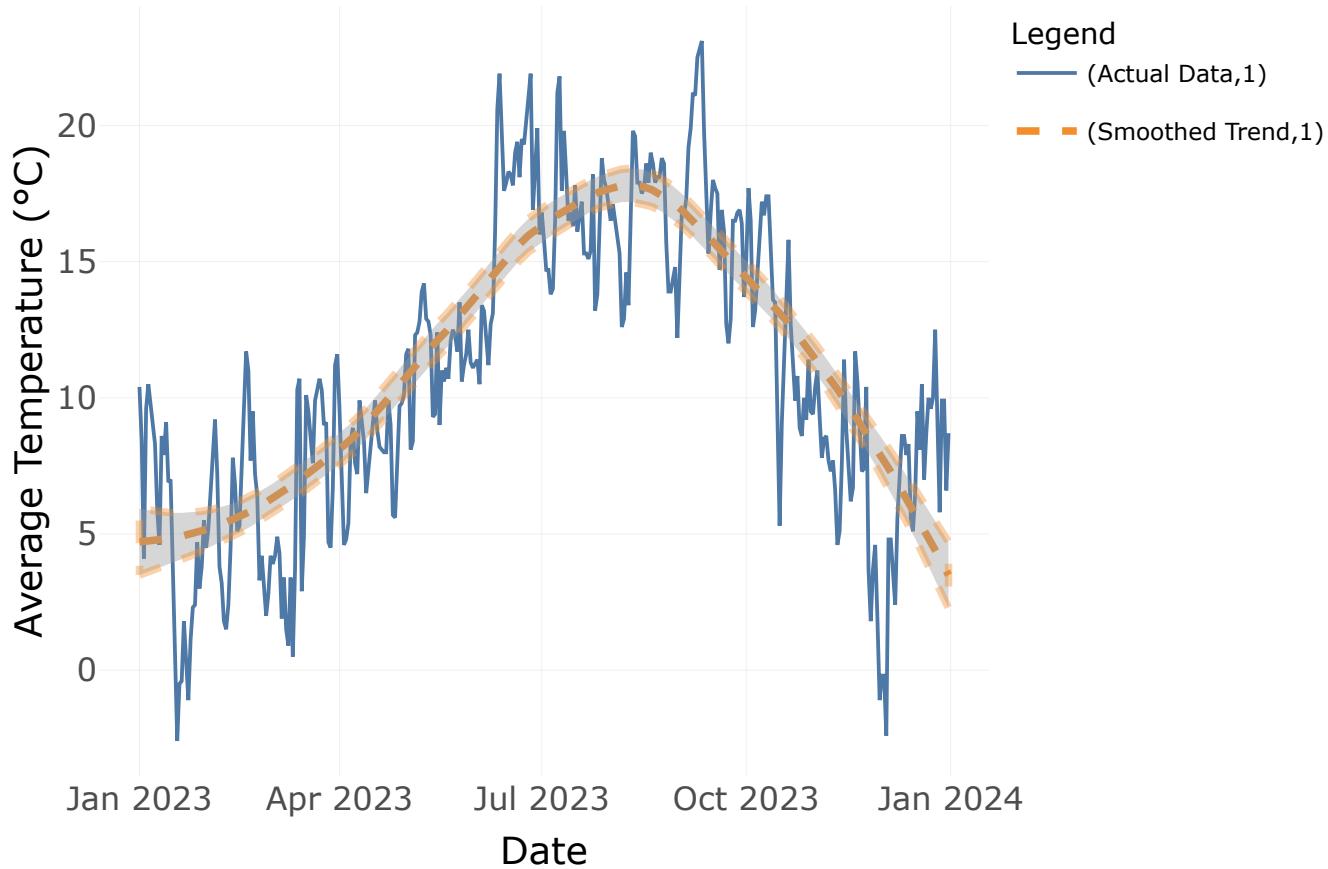




The density graph displays the distribution of wind speed and direction in Colchester for the year 2023. Wind speeds typically fall below 20 km/h, with the maximum density concentrated at the lower end of the speed spectrum. As wind speed increases, density eventually decreases. Furthermore, the figure shows a lack of a prevailing wind direction, with colors reflecting multiple wind directions uniformly dispersed across the plot. This implies that wind in Colchester comes from a variety of directions throughout the year. Density plots are useful for visualizing continuous data distributions, with darker hues representing larger concentrations and lighter shades suggesting lower concentrations.

## Time Series Analysis for Average Temperature

### Average Daily Temperature Over 12 Months



The time series plot shows the average daily temperature in Colchester, throughout the year, using both real data and a smoothed trend line. Temperatures tend to rise in the summer (June-August) and fall in the winter (December-February), according to observations. The smoothed trend line aids in recognizing overall trends, indicating that the highest temperatures occurred in July or August and the lowest temperatures in December or January. Time series plots are useful for visualizing trends across time, and using a smoothed trend line helps in identifying underlying patterns within data variations.

## CONCLUSION

In conclusion, the analysis of crime data for 2023 showed several noteworthy findings. Higher crime rate locations were found, including Abbey Field, Berechurch Road, and the area surrounding Colchester Cemetery. Robbery, violent crime, shoplifting, and anti-social behavior were the most frequently reported offenses. Furthermore, a distinct seasonal pattern appeared, with crime rates increasing in the summer and decreasing in the winter. These findings imply that environmental factors, such as temperature changes, may influence crime occurrence.

Similarly, an examination of temperature data for 2023 in Colchester revealed unique seasonal trends. The temperature was highest in the summer and lowest in the winter, which corresponded with recorded changes in crime rate occurrence. In addition, humidity levels followed a regular pattern, with higher humidity in the summer and lower humidity in the winter. Wind speeds were also found to be lower in the summer and greater in the winter, with no clear dominating wind direction throughout the year.

The association between temperature and crime rates in Colchester with a deeper examination reveals a link between greater crime rates resulting in colder temperatures, other factors that may play a role are also considered. Summertime factors such as greater outdoor activity and extended daylight hours may also have an impact on crime rates, independent of temperature. Furthermore, the analysis's limitations, such as a lack of demographic and socioeconomic data, underscore the need for more extensive research to better understand the intricate interplay between environmental factors and crime occurrence. Despite these limitations, the findings of this analysis can help to inform targeted actions and crime prevention measures in Colchester, resulting in a safer community.

## REFERENCES

- [1] Essex University Moodle 2024, Course Page, viewed 23 April 2024, <https://moodle.essex.ac.uk/course/view.php?id=12150> (<https://moodle.essex.ac.uk/course/view.php?id=12150>).
- [2] R Core Team 2021, R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, <https://www.R-project.org/> (<https://www.R-project.org/>).
- [3] Wickham, H. 2016, ggviz: Interactive Grammar of Graphics, R package version 0.4.7, <https://CRAN.R-project.org/package=ggviz> (<https://CRAN.R-project.org/package=ggviz>).
- [4] Kassambara, A. 2017, ggpubr: 'ggplot2' Based Publication Ready Plots, R package version 0.4.0, <https://CRAN.R-project.org/package=ggpubr> (<https://CRAN.R-project.org/package=ggpubr>).
- [5] Sievert, C. 2020, Interactive web-based data visualization with R, plotly, and shiny, Chapman and Hall/CRC, ISBN 978-0367563837.
- [6] Cheng, J., Karambelkar, B., & Xie, Y. 2021, leaflet: Create Interactive Web Maps with the JavaScript 'Leaflet' Library, R package version 2.0.3, <https://CRAN.R-project.org/package=leaflet> (<https://CRAN.R-project.org/package=leaflet>).