**Spatial and Temporal Distribution of Missing Incidents (in Cheshire)**

Contents:

[Intro 2](#_Toc77684960)

[Lit Review 2](#_Toc77684961)

[Research Questions 2](#_Toc77684962)

[Methods 2](#_Toc77684963)

[Datasets and Data Carpentry 2](#_Toc77684964)

[Models/Analysis 2](#_Toc77684965)

[Limitations 2](#_Toc77684966)

[Ethics 2](#_Toc77684967)

[Analysis 2](#_Toc77684968)

[Temporal Distribution (RQ 1) 2](#_Toc77684969)

[General Trend of Weekly Calls 2](#_Toc77684970)

[Call Origin 2](#_Toc77684971)

[Grade 2](#_Toc77684972)

[Response Time 2](#_Toc77684973)

[Calls Attended 2](#_Toc77684974)

[Final Classification 2](#_Toc77684975)

[Sensitivity Analysis (RQ 2) 2](#_Toc77684976)

[Spatial and Temporal Distribution (RQ 3) 2](#_Toc77684977)

[General trend of MP across LSOAs 2](#_Toc77684978)

[Quantile Classification 2](#_Toc77684979)

[LISA 2](#_Toc77684980)

[IMD (RQ 4) 2](#_Toc77684981)

[Total count from LSOAs in each IMD decile 2015:2020 2](#_Toc77684982)

[Mapping IMD 2](#_Toc77684983)

[Non-Spatial Regression: Missing Rate and Deprivation Decile 2](#_Toc77684984)

[Spatial Regression: Missing Rate and Deprivation 2](#_Toc77684985)

[The distribution across grade, origin and response time? (4a) 2](#_Toc77684986)

[Mental Health (RQ 5) 2](#_Toc77684987)

[Total count from LSOAs in each Mental Health decile 2](#_Toc77684988)

[Mapping Mental Health 2](#_Toc77684989)

[Non-Spatial Regression; Missing Rate and Mental Health 2](#_Toc77684990)

[Spatial Regression: Missing Rate and Mental Health 2](#_Toc77684991)

[The distribution across grade, origin and response time? (5b) 2](#_Toc77684992)

[Covid-19 Distribution 2](#_Toc77684993)

[General trend of MP 2](#_Toc77684994)

[Association with IMD (this is a maybe if the years match) 2](#_Toc77684995)

[Association with Mental Health (this is a maybe if the years match) 2](#_Toc77684996)

[Discussion 2](#_Toc77684997)

[Conclusion 2](#_Toc77684998)

# Intro

Aims and objectives:

This paper uses Calls for Service provided by Cheshire Police to understand the spatial and temporal patterns of missing incidents.

* The premise for this paper is two-fold;
  + Literature on missing persons has primarily focused on the qualitative aspects that hone in on the narrative of missing people in order to shape police practise (Fyfe et al., 2015; Parr and Fyfe, 2013). Yet, there is an absence of quantitative analysis which can prove essential in the understanding of missing incident trends across time and space
  + Secondly, geospatial analysis of calls for service is paramount in the safeguarding of vulnerable people by providing a decision-making tool for law enforcement, while also promoting the use of crime research

This paper focuses on how specific risks faced by missing persons can be predicted from a combination of individual characteristics and spatial-temporal analysis. The definition of ‘vulnerability’ in this paper refers to both deprivation and mental health (explained in lit review)

……..

# Lit Review

In the UK alone, missing persons are reported every two minutes, yet there is a large amount of ambiguity that surrounds why people go missing, how they are reported, how they can be found, who the most common groups to go missing are and how police deal with the initial response (Taylor et al., 2019). Some of this confusion stems from a lack of an implicit definition of ‘missing persons’ among police and missing persons agencies.

……..

# Research Questions

1. How has the handling of missing person calls changed from 2015-2020 over, grade, origin, response time and classification?
2. What effect do changes in origin, grade, time and classification have on the rate of missing person calls?
3. Spatially what areas are associated with missing person calls? (rural/urban dived)
4. What is the spatial association between going missing and levels of deprivation?
   1. And how do these vary grade, response time and call origin?
5. What is the spatial association between going missing and mental health?
   1. And how do these vary across grade, response time and call origin?

# Methods

## Datasets and Data Carpentry

* Calls for Service
* IMD
* Mental health
* Census (population statistics)
* LSOA/LA lookup table
* Variables Used/data manipulation
* Study area

## Models/Analysis

* Time series models
* Sensitivity analysis
* Spatial Autocorrelation Maps
* Spatial Regression
* (Justify the use of each one, referring to RQs)

## Limitations

* Problems with spatial data
* Problems with MH dataset
* Problems with IMD/census
* missing person vs missing incident
* repeated calls

## Ethics

* All cleared by UoM/Cheshire police (cite reference number)
* All data anonymised

PRop

# Analysis

### **I. Spatial and temporal patterns in missing incidents**



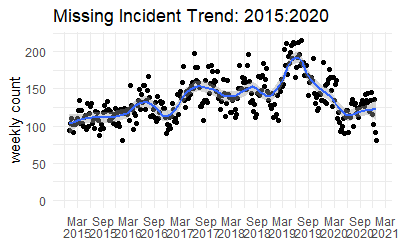
5

. The clusters appears to me in the most populated areas, which would be expected as an increase in population leads to an increase in missing person incidents. We can also run a LISA map on the residential population by dividing the number of missing dividing by the residential population and multiplying by 1000 – does the lisa map still hold?

### Temporal Distribution (RQ 1)

This section will look specifically at the changing nature of missing person calls from 2015:2020 and how these trends vary between grade, origin, response time and final classification. For each relationship, anova/t-tests will be run to test the strength of each relationship

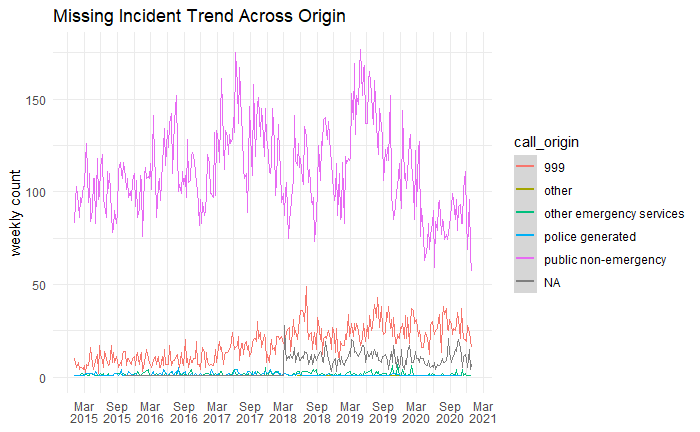
##### General Trend of Weekly Calls



**II. Police Response to Missing Incidents**

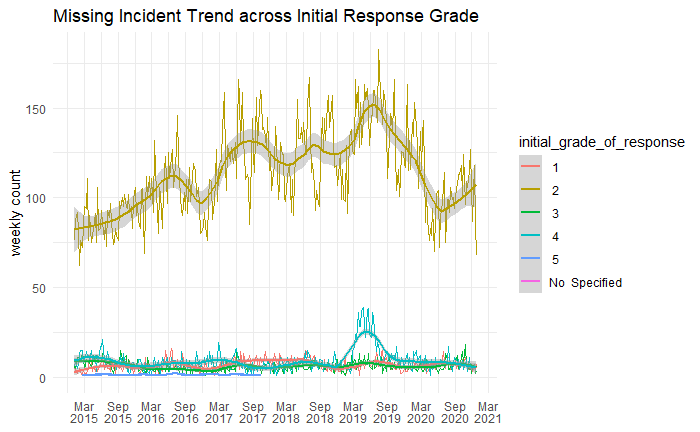
##### Call Origin

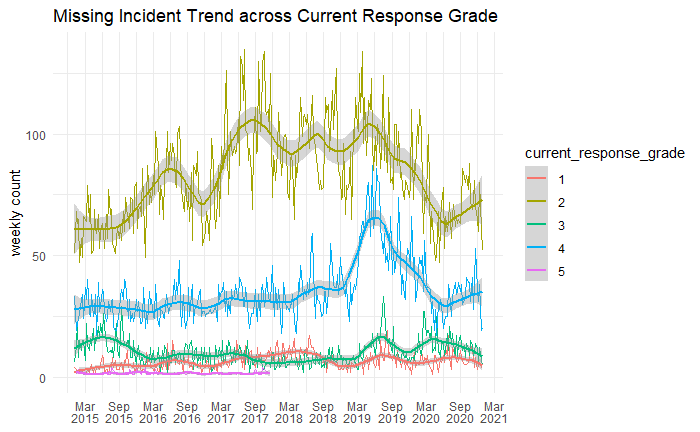
* Welch Anova Test

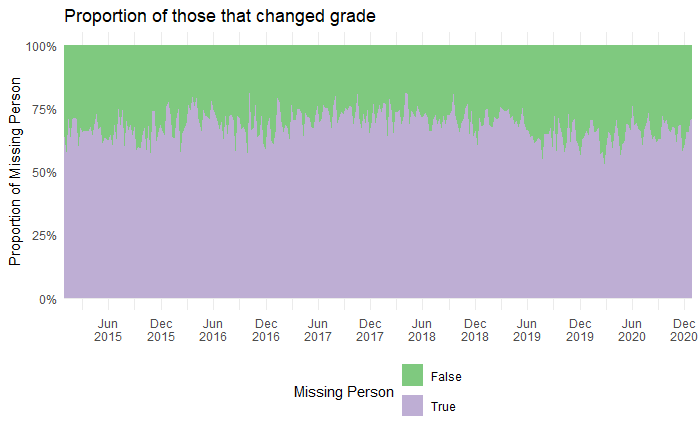


##### Grade

* Welch Anova Test

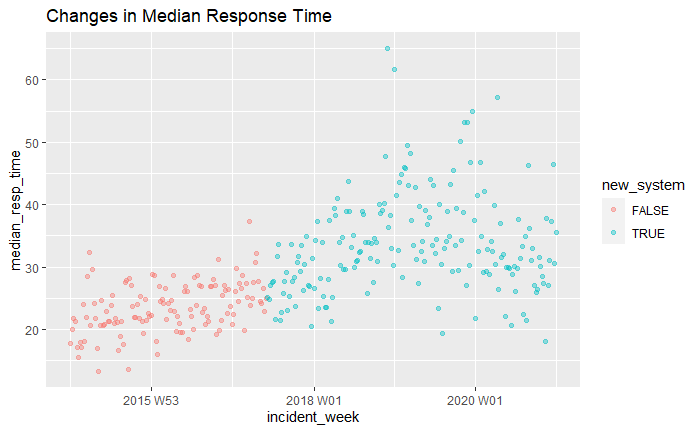




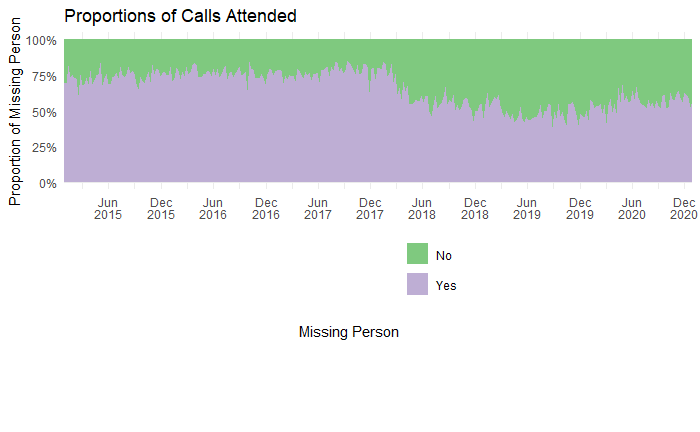


##### Response Time

* Linear regression

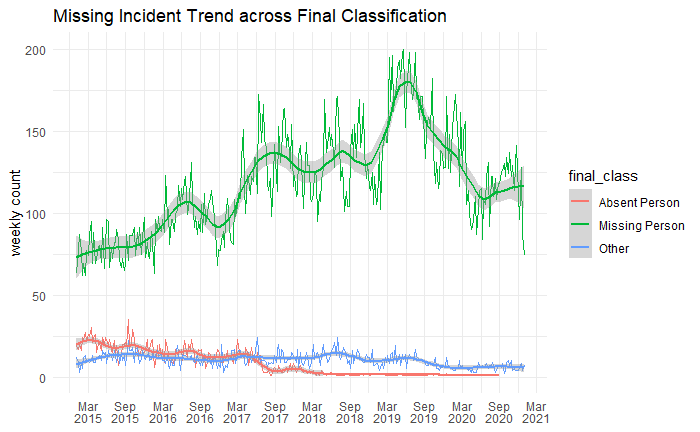


##### Calls Attended

* T-test

##### Final Classification

* Anova



With each graph summarise and why this has led you on to study the spatial distribution

### Sensitivity Analysis (RQ 2)

In order to answer which of these predictor variables are most important or the most influential (i.e., the effect that changes in origin, grade, time, classification have on the missper rate).

A sensitively analysis will also give insight into the robustness of the population size estimates against unobserved heterogeneity (used to account for the overdispersion in Poisson Regression)

* Start with a poison regression, including interactions for unobserved confounders
* Then run the sensitivity analysis

### 

##### 

##### 

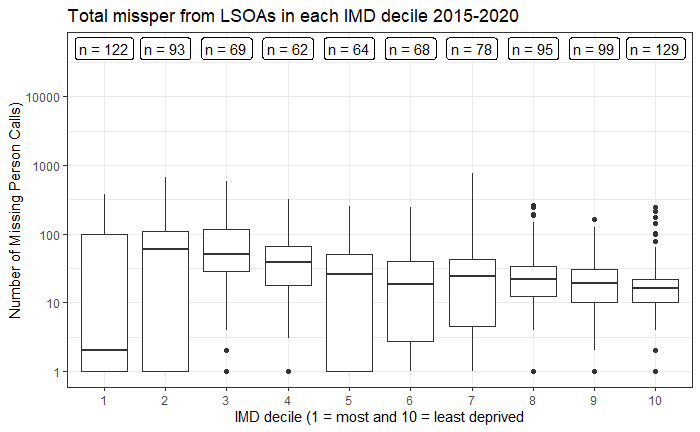
##### 

**III. Environmental/ neighbourhood level correlates of missing incidents**

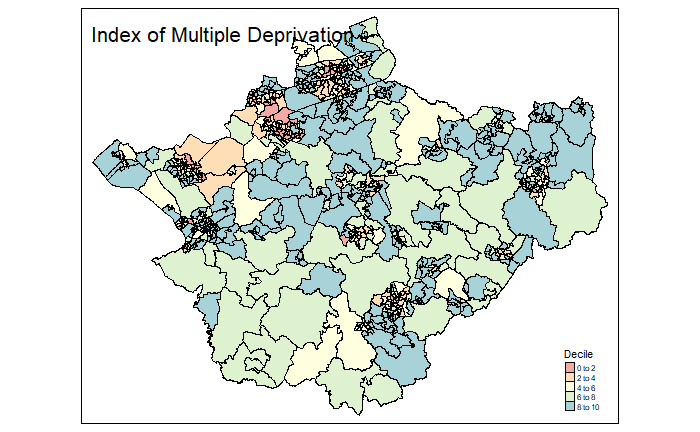
### IMD (RQ 4)

* Following the literature, vulnerability in missing incidents can be classed within levels of poverty and economic instability. Therefore, the IMD was used to explore this association through conducting a spatial regression
* Additionally, this section examines the distribution of IMD rates and the proportion of those LSOAs classed as high risk, the response time and the call origin

##### Total count from LSOAs in each IMD decile 2015:2020



##### Mapping IMD



##### Non-Spatial Regression: Missing Rate and Deprivation Decile



Using the morans test for regression residuals (i - 63.36, p = 2.2e-16), we obtain a statistically significant value for Moran's I so we need run a spatial regression model

##### Spatial Regression: Missing Rate and Deprivation

* RLM lag produces higher results so I will use that.
* results: the spatial lag parameter rho is significant (<2.22e-16). The AIC reduces to 85111 from 85987 (lm) therefore a better fit

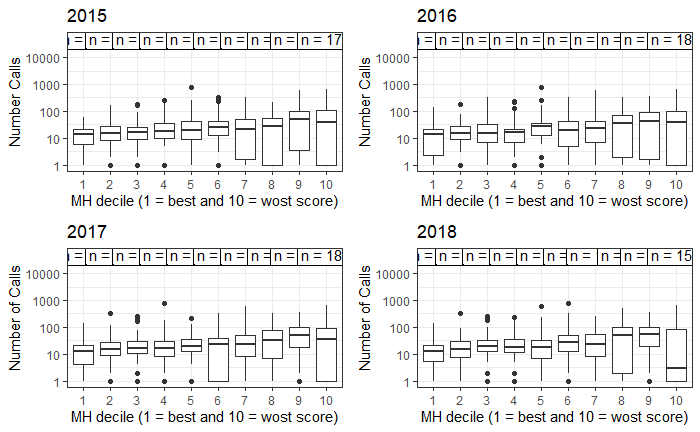
##### The distribution across grade, origin and response time? (4a)

* This section will aim to show the association between differences in IMD rank and the differences in my predictor variables (origin, response time, classification etc) across LSOAS

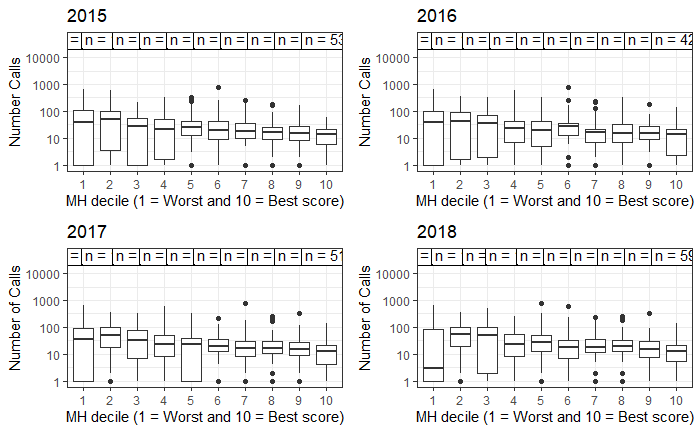
### Mental Health (RQ 5)

* Following the literature, vulnerability in missing incidents can also be linked to mental health. Therefore, the Mental Health dataset was used to explore this association through a spatial regression
* Additionally, this section examines the distribution of Mental Health rates and the proportion of those LSOAs classed as high risk, the response time and the call origin

##### Total count from LSOAs in each Mental Health decile



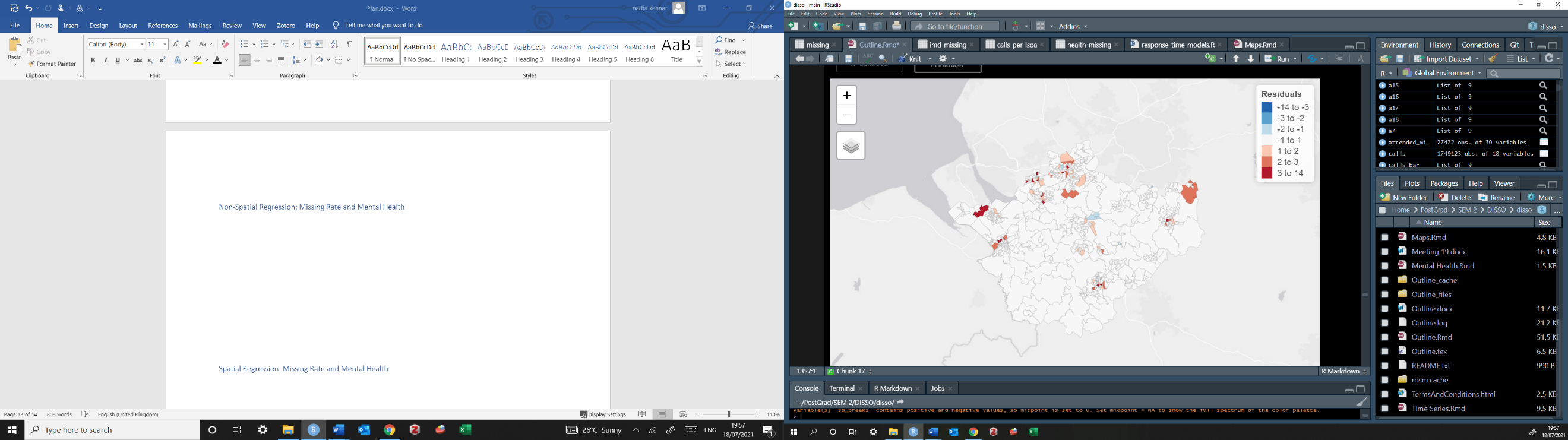
##reverse ones



* There are more missing persons from those LSOAs with increased mental health problems (including NHS-Mental health-related hospital attendances, Prescribing data – Antidepressants, QOF - depression, and DWP - Incapacity benefit and Employment support allowance for mental illness)
* Study the means between these variances are also significant, all tests are significant

##### Mapping Mental Health

##### Non-Spatial Regression; Missing Rate and Mental Health



##### Spatial Regression: Missing Rate and Mental Health

* Again, the non-spatial regression highlights some area of over prediction and under prediction. The Morans I statistic deviate (4.2135 and p = 2.515e-05). As test is significant, we can run a spatial regression
* in order to decide to run a error or lagged model, we run a lagrange multiplier test. in this case lag is has the higher robust test so we choose this
* Again, results highlight a significant rho test, and the AIC has reduced from the linear model (12696 - 12728)

##### The distribution across grade, origin and response time? (5b)

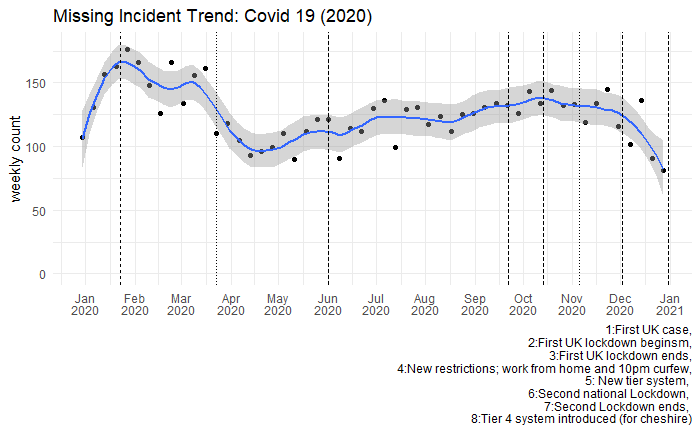
* This section will aim to show the association between differences in Mental Health rank and the differences in my predictor variables (origin, response time, classification etc) across LSOAS

### **IV. COVID and Missing Incidents**

### Covid-19 Distribution

It is also important to note how these results changed during covid-19 when there was a distribution to both policing and everyday routines

##### General trend of MP



…..

# Discussion

* Detail major findings and the importance
* Draw on some limitations of the findings
  + Generalisability etc
* Possibly suggestions for future research

Theme 1

Theme 1 exposed missing incidents spatial patterns, using Lisa maps and Moran I statistics we were able to witness a positive spatial autocorrelation where high levels of missing incidents are associated with other high areas of missing incidents. Concluding that missing incidents do cluster, as would the majority of any other crime or criminal activity – these results were consisted across the rates of missing incidents and not just the counts. Additionally, theme 1 also exposed that missing incident trends are seasonal in nature. The use of seasonal variation and decomposition methods allowed to expose the trend, seasonality and noise components of the data. Typically, we have an increasing trend of missing incidents from 2015, following a decline in 2020 as a result of the pandemic changing the daily routines of both society and police response. The seasonally adjusted modifies the effect of seasonal influences which provide for more meaningful comparisons; we can conclude that missing incident reports happen most frequently over the summer months. This is interesting as it compares to what literature says. From this, a SARIMA model was conducted which compared the actual counts of missing incidents compared to the predicted rates following the first UK Covid-19 case concluding that the calls were received were far less then expected

Theme 2 explored the police response to missing persons providing both descriptive and temporal trends. Despite literature stating most calls are handled by 999, the results from this paper show that PNE handle the majority of calls, additionally most of the calls are graded at level 2. The results of the temporal trends highlighted a consistency in the allocation of police response with the main call origin and grade not sifting too much. However, median response time highlight an increasing trend which may be a result of new policy introduced in 2018. Additionally, the classification of ‘absent person’ reduced from 2018. Its also interesting to note how the incident policy response from gmp (2019) were consistent with the average response time from each grade and happened much quicker then possible expected.

Theme 3 explores the neighbourhood covariates where missing incidents are more prone in higher deprived LSOAs and higher mental health LSOAs. Additionally, the spatial regressions highlighted that these are clustered among the Urban LSOAS, whereas rural provide insignificant Moran statistics. Missing incidents make up 2.4% of calls from the whole calls for service dataset, yet results highlight that these are spatially clustered in 18.8% of our study space (the 103 urban LSOAs) of which are also high in deprivation statistics and mental health concerns. There is a clear relationship between the environmental correlates and the distribution of missing incidents

# Conclusion

* Summarise the whole paper while restating RQs
* What contributions this paper has made
* State future directions for policy and research

Questions

1. For the covid-19 trends, would it be better to just include these in the temporal models (like we did with N8), or include these as its own separate models as seen here?
   1. You can include those trend models – think about what question this section is answering? What would you like to model? I think there’s already a lot there so you probably want to keep your COVID section concise. But the forecast models are good because they help you talk about what was different during covid/lockdown compared to what we would expect if there was no COVID based on the forecast.
2. For the sensitivity analysis, is it correct in assuming you model these using a rate rather than the count data (i.e. missing person call rate)?
   1. This is relevant for all the analysis – think about when it’s appropriate to use rate and when it’s appropriate to use count. I raised this in some of my comments, but overall think about when you use count, can you be certain you’re not just measuring ‘more people’ in general?
3. For the IMD, I’ve ran the results on each domain of deprivation as well (not included here?) – I think it might be more useful than just providing a holistic image – opinions?
   1. This is up to you, you can say that you think certain domains of deprivation are more relevant than others – you will have to justify why you think this – and then run those. You could include each domain separately in one model like MP ~ domain 1 + domain 2 + domain 3 etc, but think about justifying this as well!
4. I’m unsure if research question 4a and 5b truly add depth to this paper, partly because the results may not seem generalisable due to being centralised to Cheshire – I could however adda rural/urban element which may explain the changes in origin, response time and grade over the most/least deprived areas – opinions?
   1. I think there is definitely some reflection to do on the research questions, I made a few comments for you throughout so that should hopefully answer this question. The urban/rural distinction could be interesting! But this could be something to include in the environmental correlates section. For example, you might want to create an urban/rural variable for your LSOAs to then include in your model with the IMD and MH options, and then you can see if urban/rural location has an importance, alongside deprivation or mental health.
   2. Generally I think that seeing that missing incidents cluster is important finding. You can tie this into your literature review around crime concentration/ environmental crim – where you contribute that you find that Missing Incidents concentrate. To make it more broadly relevant, make sure you interpret the clusters. You might say oh these are the major cities in the area, etc – but gain think about whether it might be worth mapping RATE here, so you’re not just mapping where people are, and interpreting that as where people go missing…!
5. Is it possible to run a small multiple across lsoas and month, I’ve had a look and it sems there are over 600 LSOAS with more than one month of no calls reported? (And over 56 LSOAS with more than one year of no calls)
   * Following the labs from crime mapping it seems inefficient has this involves manually coding for all these LSOAS – are you aware of any other method for this?
   * Not necessary but would show a clearer trend across the months/years
   * For this one, think about what this would show. I think instead of going into spatio temporal variation, which like you say I think we have a little too much noise, it might be worth to think about other temporal variations, maybe around seasonality, or other units of fluctuation. I think it’s better to focus on what you already have and flesh it out more, engage more with the questions you have, and fill in the gaps.

*For the rural area: THIS IS WHEN USING NUMERICC DECILE NOT FACTOR DUH!!!*

*The Moran I for residuals on just the rural areas are (Moran I = 3.38 p = 0.0007), the Moran I for residuals in the urban areas are (Moran 1 = 2.79, P = 0.005192)*

*The Lagrange test was used to establish the hypothesis about parameters in a likelihood framework, under the rural model the standard error and lag models are both significant as below 0.05 (error p = 0.000096, lag p = 0.0011). The robust error model (p =0.7) was not significant so the robust lag (p = 0.002) proves the correct method (Luc Anselin, 2008).*

*The spatial autoregressive parameter rho is significant (Rho: 0.36468, P: 0.0018). the spatial lag parameter, Rho, is also positive and statistically significant. When missing incident rate in surrounding rural areas increase, so does the missing incident rare of each LSOA, even when adjusting for deprivation statistics and median age. The lag model has an AIC of 889.81 whereas the linear model with no lag has an AIC of 897.6 indicating a model with lag is a better fit*

*In a spatial lag model, the coefficients focus on the short-run impact of xi on yi, rather than the remaining effect. This results in two situations; the first being the direct impact of an observation’s predictor on its own outcome and secondly the indirect impact of an observation of the neighbour’s predictor on its outcome. We can obtain simulated distributions of the various impacts using Monte Carlo (Lesage and Pace, 2009)*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | *Direct* | *P* | *Indirect* | *P* | *Total* | *P* |
| *IMD Decile* | *-2.14* | ***0.03*** | *-1.09* | *0.12* | *-3.23* | *0.04* |
| *Median Age* | *-0.59* | *0.19* | *-0.3* | *0.27* | *-0.89* | *0.2* |

*A negative indirect effect is considered a spatial benefit, these indicate variables that lead to a reduction in missing incident in neighbouring rural LSOAs with decreased IMD decile score and decreased age*

*A positive indirect effect represents negative externality, where neighbour rural LSOA result in an increase*

*….*