Time Series Analysis of United Kingdom Traffic Accidents

Brian Cervantes Alvarez

March 14, 2024

Introduction

This study focuses on identifying patterns in the frequency and severity of traffic accidents during winter months and holiday seasons in the United Kingdom. Specifically, we investigate variations in fatal, serious, and minor accidents to understand how seasonal factors influence traffic safety. Figure 1 plotted daily accidents by severity level, which revealed an unexpected dip in accident numbers during winter, prompting further investigation into underlying patterns and potential confounding variables.

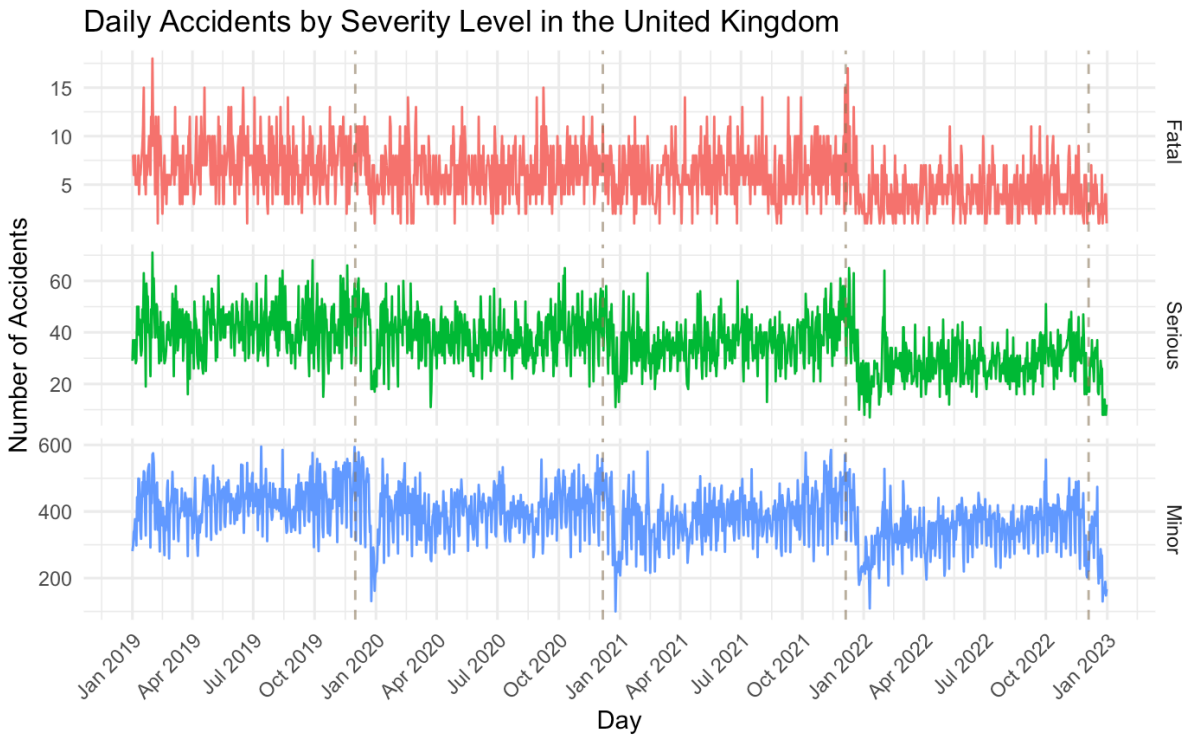

Figure 1: Time Plot

Methodology

Our methodology began with data wrangling and utilized the `dplyr` and `lubridate` packages to clean the data. We filtered out irrelevant rows and ensuring proper data structure formatting of columns. Following data wrangling, we employed various visualizations using base `R` to explore temporal trends, seasonality, and patterns in accident severity. Specifically, we conducted an in-depth examination of daily trends in accident severity, plotting the number of accidents by severity level against time. This enabled us to discern fluctuations over time and potential seasonal patterns, offering valuable insights into the inherent dynamics of the data.

Time Series Analysis

We performed detrending on the time series of fatal accident severity to assess stationarity and temporal dependencies, aiming to remove systematic trends or patterns from the data to make it amenable to formal statistical modeling. Following detrending, we conducted autocorrelation and partial autocorrelation analyses to evaluate temporal dependencies in the detrended series, helping to identify significant correlations at various lag intervals, which informed our choice of modeling approach. Finally, we employed formal statistical tests, such as the Augmented Dickey-Fuller (ADF) test, to assess the stationarity of the detrended series, providing critical insights into the suitability of modeling approaches and the need for further data transformations. Reference Figure 2 for more information

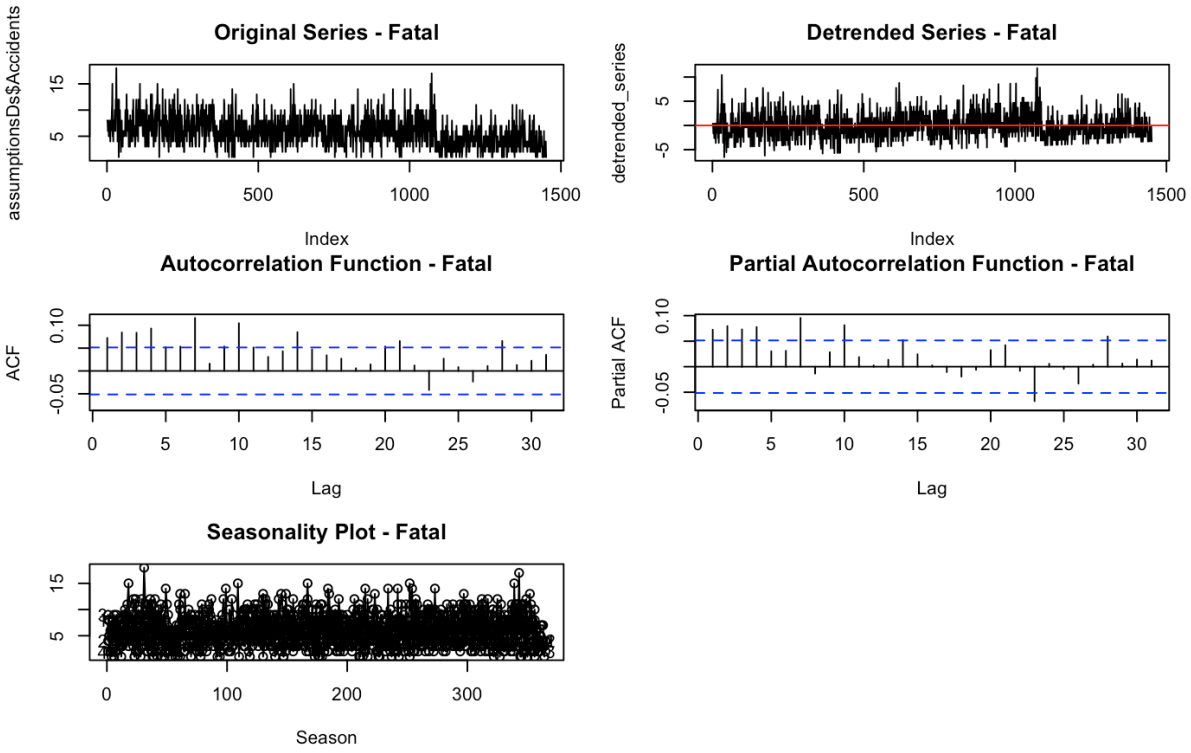

Figure 2

Results

1. **Original and Detrended Series**: The visual inspection indicated no strong trend in the original series for fatal accidents, confirmed by a relatively stable mean in the detrended series.
2. **Stationarity Test**: The Augmented Dickey-Fuller test on the detrended series yielded a test statistic of -8.2276 and a p-value of 0.01, strongly suggesting the series is stationary. This indicates that the series does not have a unit root and supports the suitability of modeling without the need for further differencing.
3. **Seasonality Plot**: Did not reveal clear patterns of seasonality, indicating the complexity of identifying seasonal effects on fatal accidents.
4. **ACF and PACF Analysis**: Both autocorrelation and partial autocorrelation functions showed no significant correlations for lags up to 30, suggesting the absence of strong temporal dependencies after detrending.

Discussion

The initial exploratory analysis suggested a counterintuitive decrease in accidents during winter, leading to a hypothesis of potential confounding variables such as weather conditions, public awareness campaigns, and holiday effects. The lack of significant autocorrelation in the ACF and PACF plots, along with the ADF test's confirmation of stationarity, indicates that simple temporal models may not capture the complexity of the underlying processes affecting fatal accidents.

This suggests that while seasonality and trend components are not dominant in the fatal accidents series, other external factors or non-linear patterns might play a significant role. The absence of clear autocorrelation and seasonality could imply that factors beyond the scope of this time series analysis impact accident rates, such as changes in road safety laws, traffic volume, or driver behavior.

Conclusions

The analysis reveals that fatal traffic accidents in the United Kingdom do not exhibit strong seasonal or trend patterns but are influenced by more complex factors. The stationarity of the series post-detrending supports the use of ARIMA models for forecasting, although the lack of significant autocorrelation points to the potential need for incorporating external variables or exploring non-linear modeling approaches to better understand the dynamics at play.

Future research should consider integrating additional data sources, such as weather conditions, traffic flow information, and details on road safety initiatives, to develop a more comprehensive model of traffic accident occurrences. Additionally, exploring machine learning models that can capture complex interactions between various predictors might offer deeper insights into the factors driving variations in accident severity over time.

Reference

Gibin, W. O., & Sheen. (2023). ***Road Accident Casualties***. Kaggle. https://doi.org/10.34740/KAGGLE/DSV/7292741