



When Do We Need Emergency Services?

Analysing Vic Roads crash data

August 2023

Andre Medina

Why?

How many emergency services do we need?

Should we reallocate the current distribution?

Where should we send Ambulances & Police in the moment?



How?

The crash data!

- Find key metrics
- Group the crashes
- Number of times 000 arrived
- Create a statical model

Major steps



Investigating
data structure



General data
wrangling



Viewing initial
trends



Further data
cleaning



Final statistical
model



Investigating Structure & Wrangling Data

Initial Data Structure

Pre-Crash data

- ▶ Time of accident
- ▶ Atmospheric condition
- ▶ Road condition
- ▶ Local Government Area (LGA)
- ▶ Light Condition

Post-Crash data

- ▶ Severity
- ▶ Number of people involved
- ▶ Accident Type
- ▶ Police Needed
- ▶ Ambulance Needed

Simplifying Data

Pre-Crash data

- ▶ Day of the week
- ▶ Part of the day
- ▶ Atmospheric condition
- ▶ Road condition
- ▶ LGA / Region

Post-Crash data

- ▶ Police Needed
- ▶ Ambulance Needed

Ambulance Needed

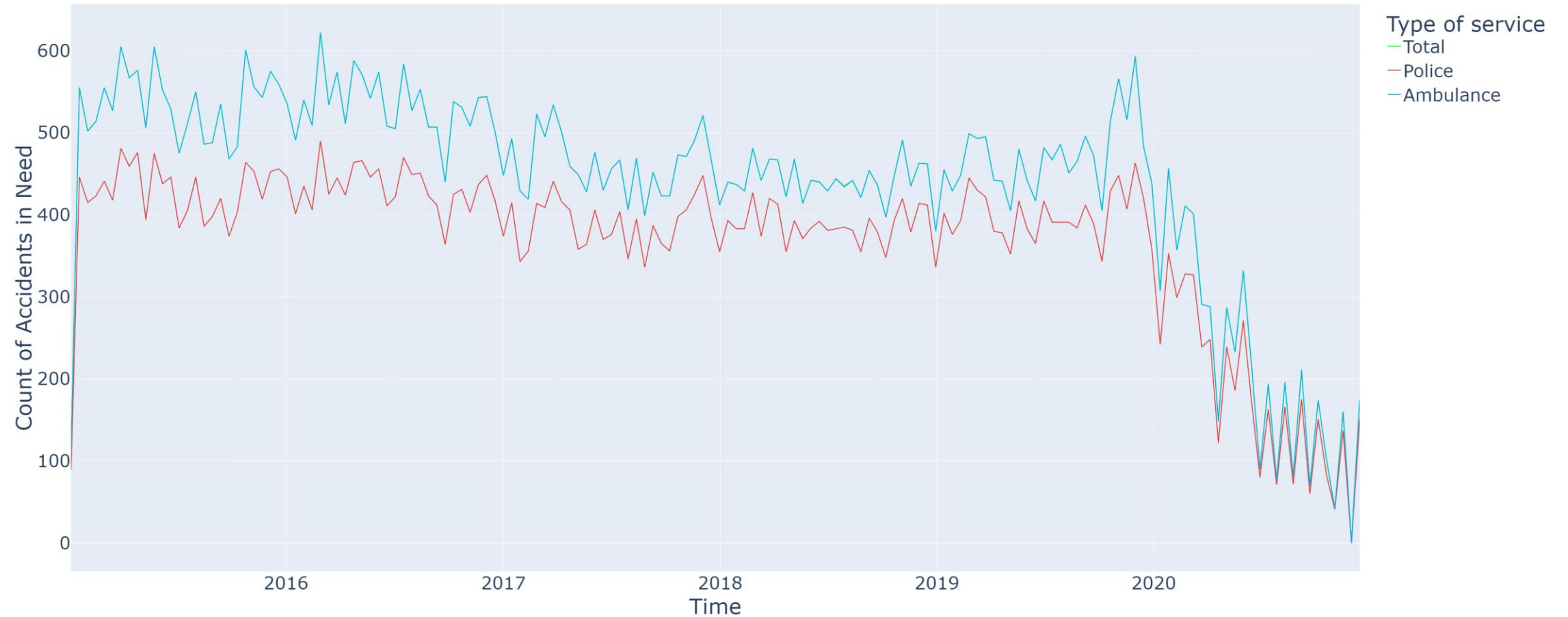


- ▶ Only one ambulance per crash
- ▶ Each person of a crash had their injury severity recorded
- ▶ If someone was taken to the hospital, it was recorded

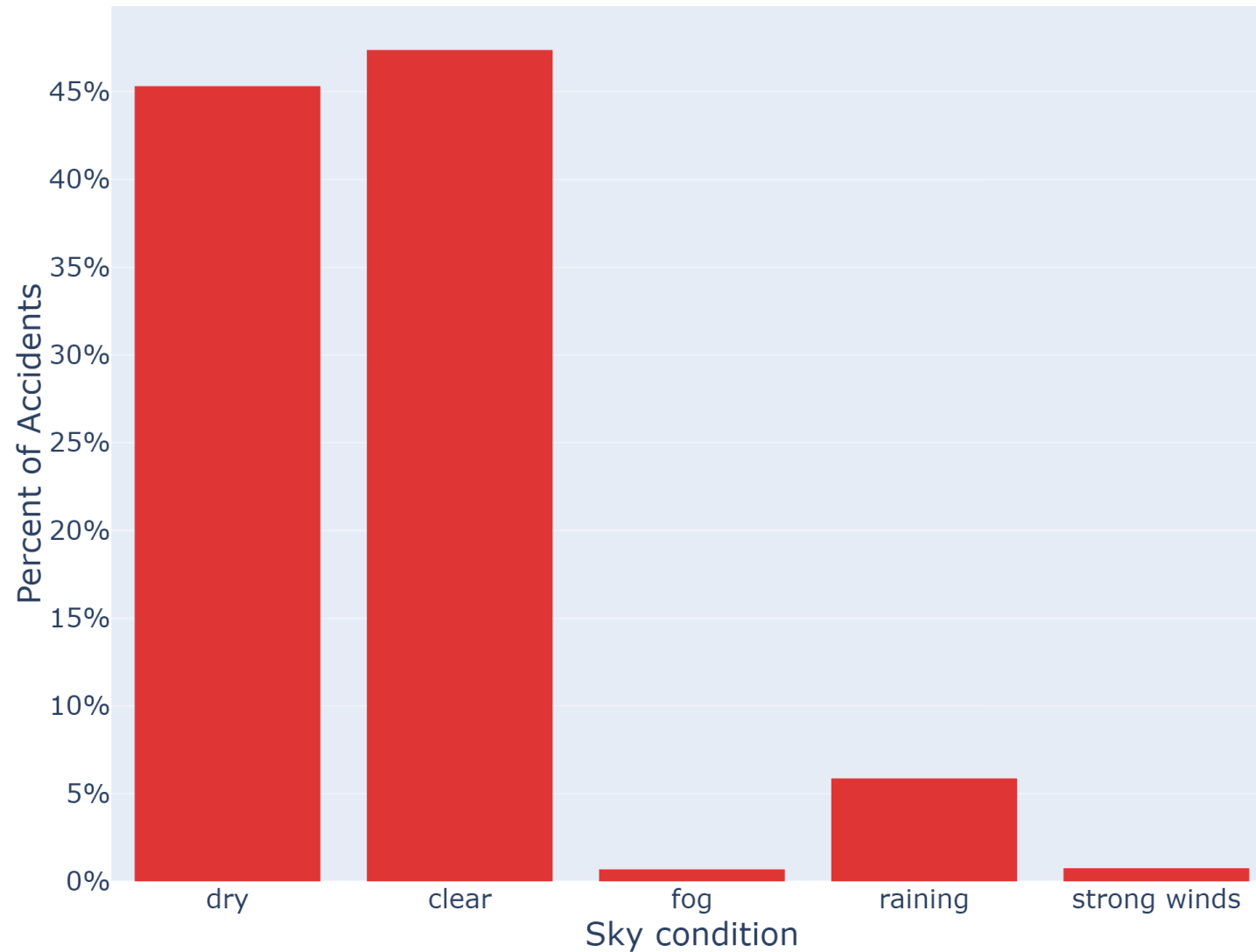
The background of the slide features a blurred image of financial market data. It includes several candlestick charts with white and black bars, overlaid with various colored trend lines (red, blue, and green). A silver pen is positioned diagonally across the upper right portion of the image. The overall color palette is dominated by reds and blues, with a semi-transparent white area in the center where the text is located.

Viewing Initial Trends

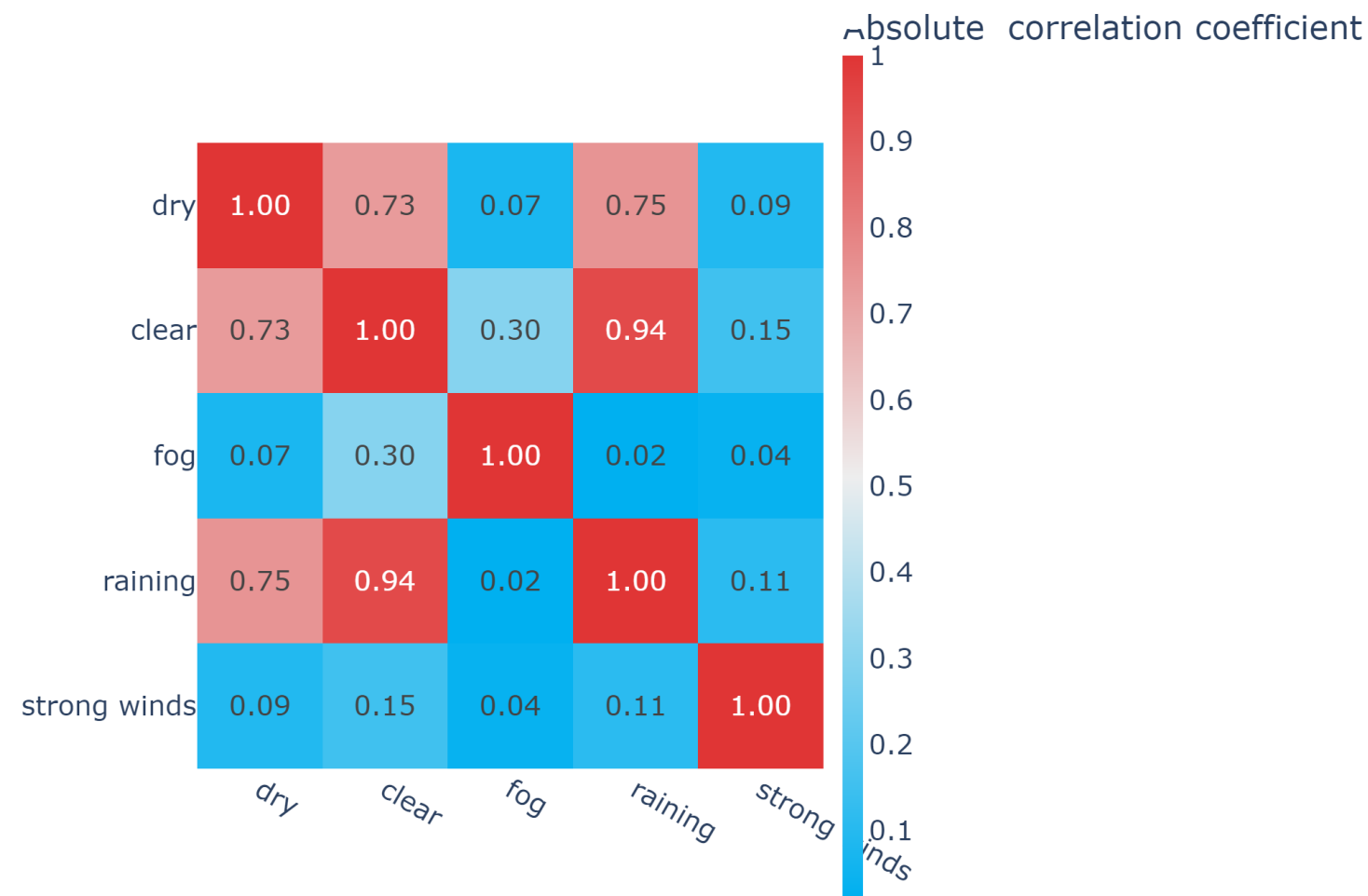
Emergency Callouts Over Time (Fortnightly)



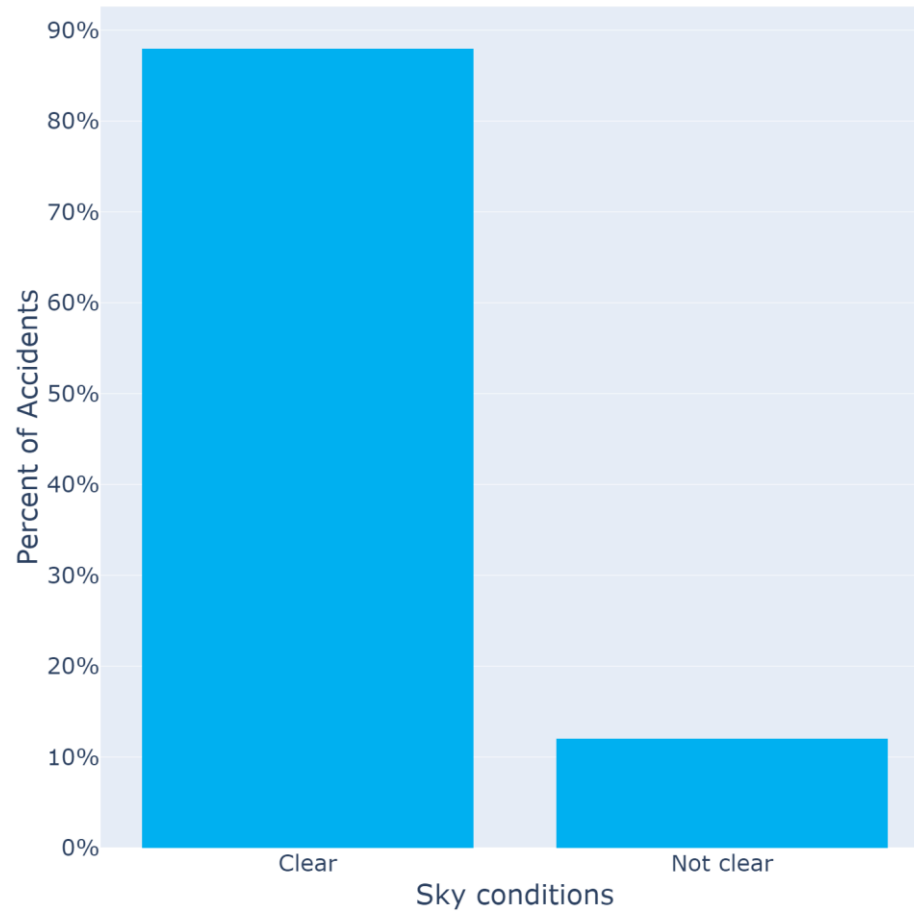
Crashes by Road and Sky Conditions



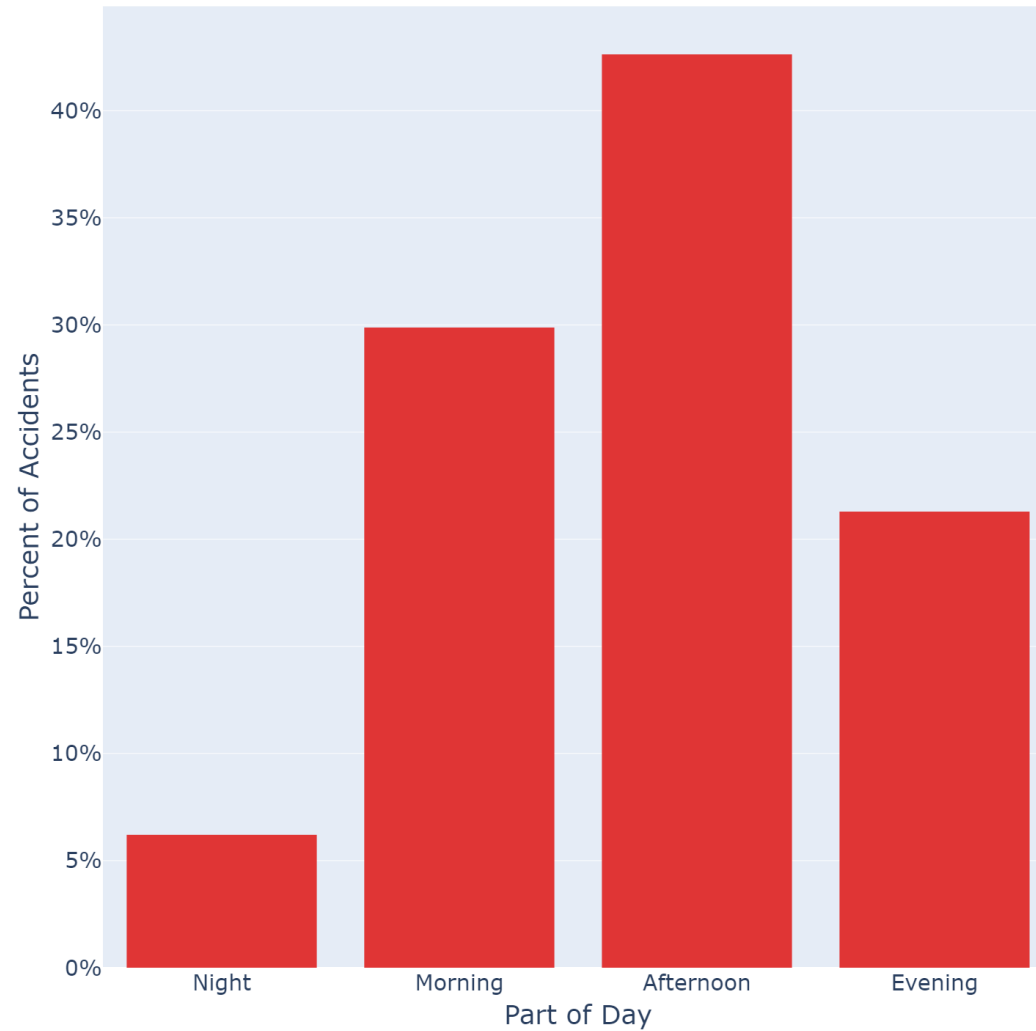
Correlation of Conditions



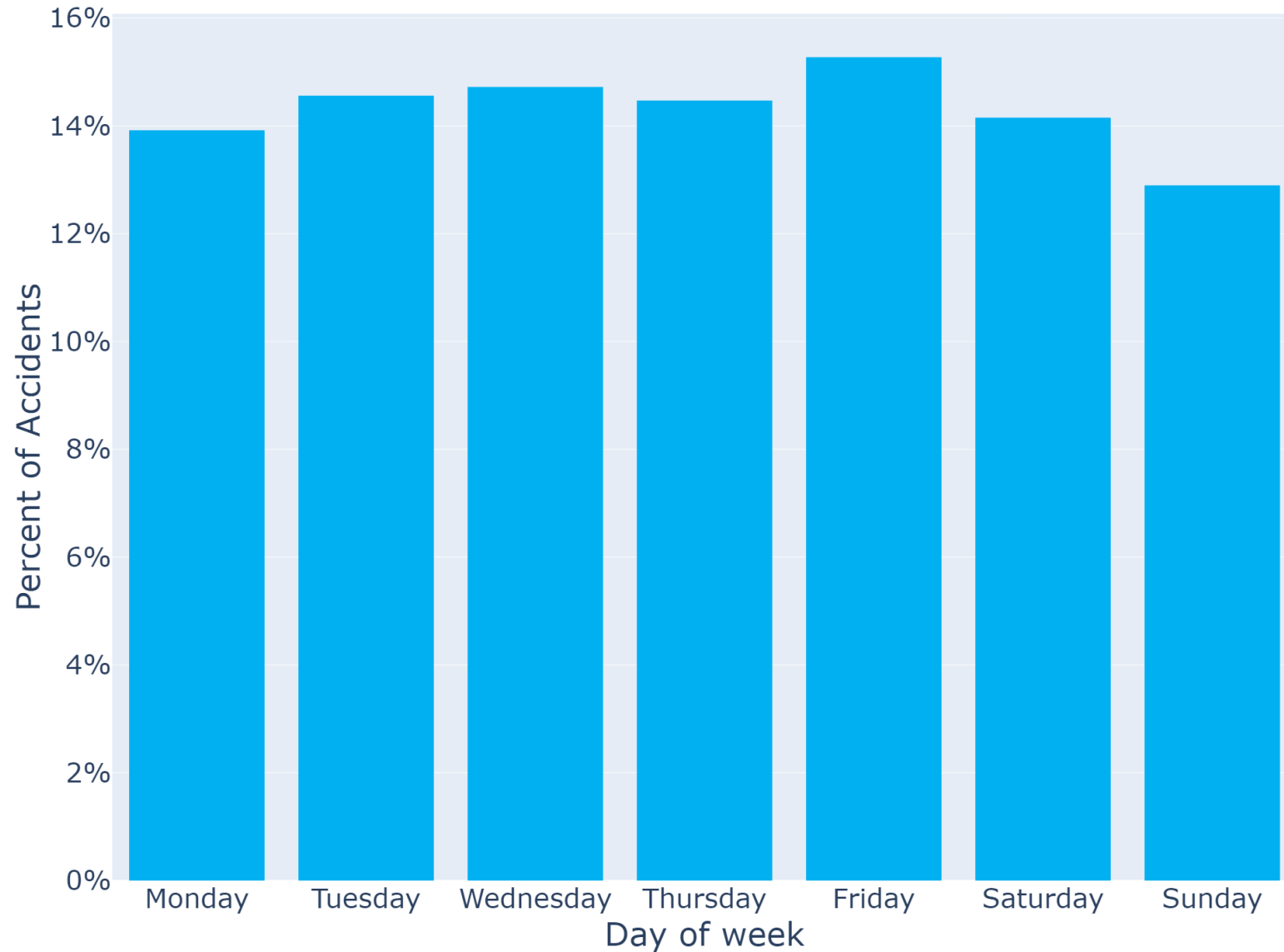
Emergency Callouts by Sky Conditions



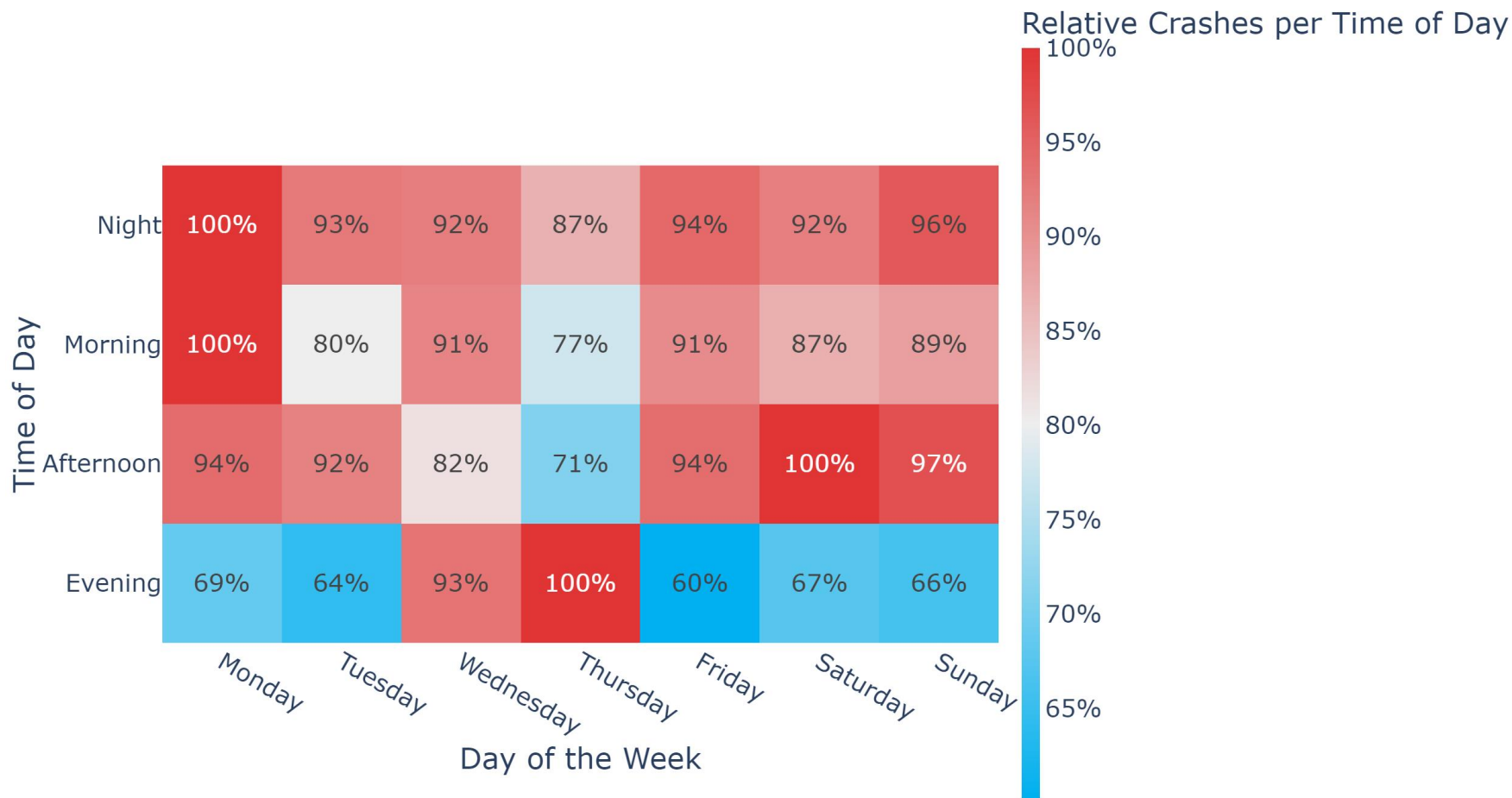
Emergency Callouts by Part of Day



Emergency Callouts Over Time



Crashes by Day





Further Cleaning

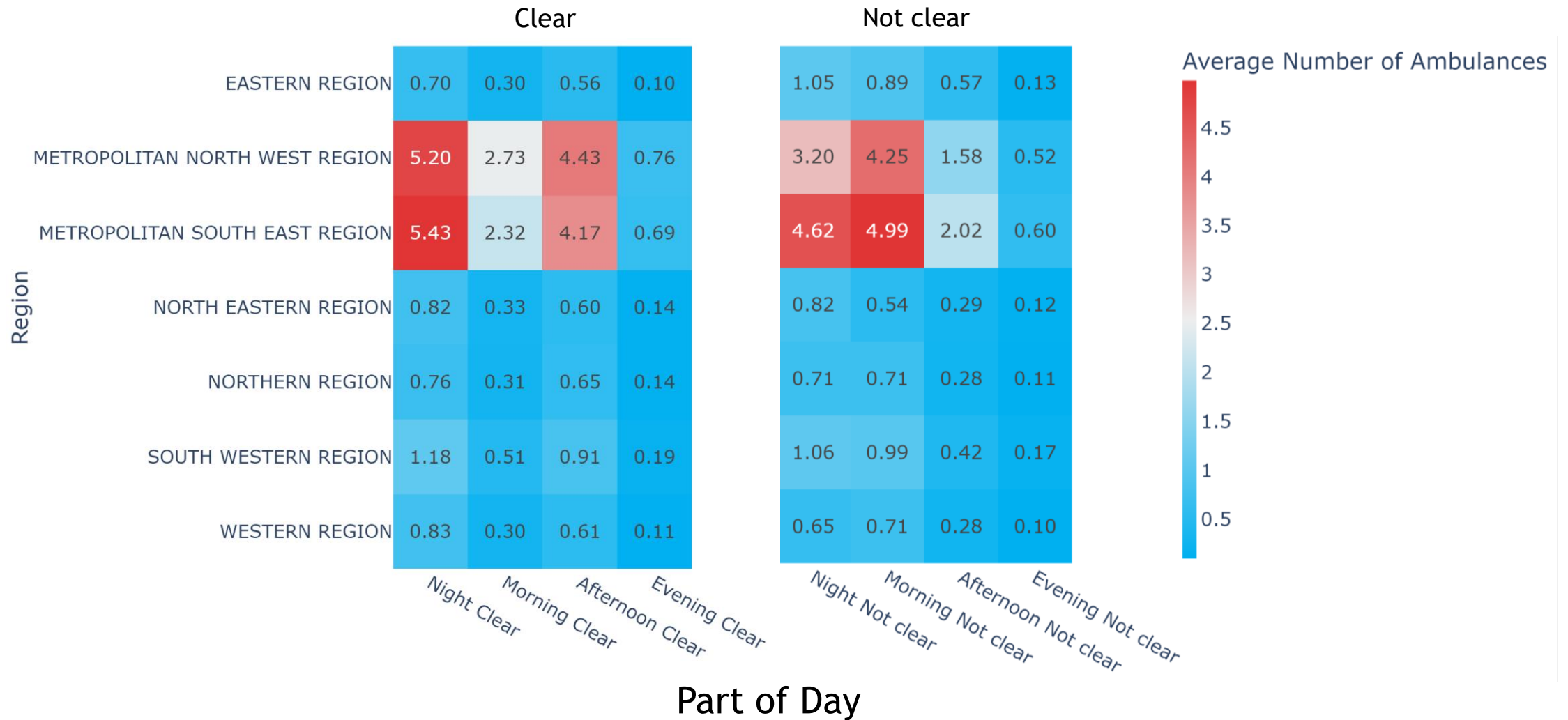
Target Format

- ▶ For each combination of
 - ▶ Sky condition
 - ▶ Region
 - ▶ Day of the week
 - ▶ Part of the day
- ▶ Calculate average number of emergency services

Potential Issues

- ▶ Regional areas have smaller number of samples
- ▶ Data is already split up into rainy days and clear days
 - ▶ Infrequent rainy days will skew data
 - ▶ Use probability of a rainy day to compensate
- ▶ Data is both summated and averaged over time

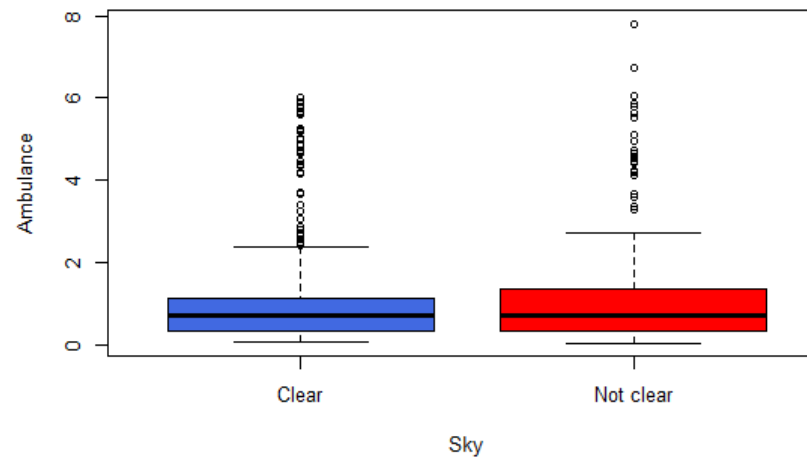
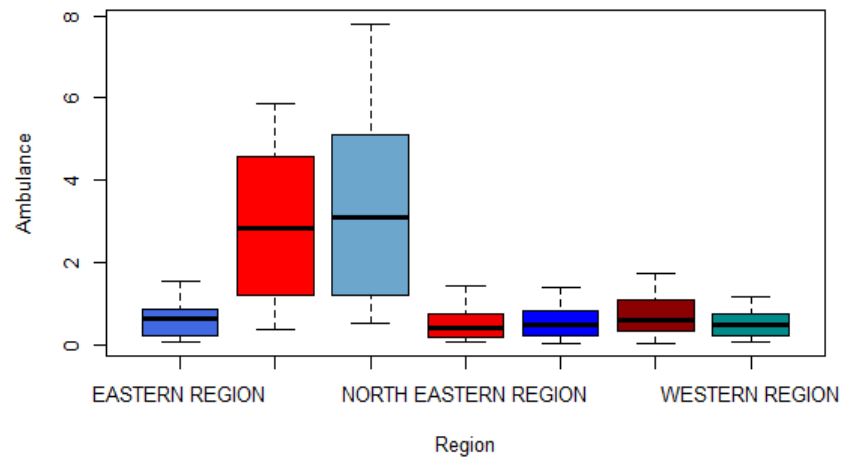
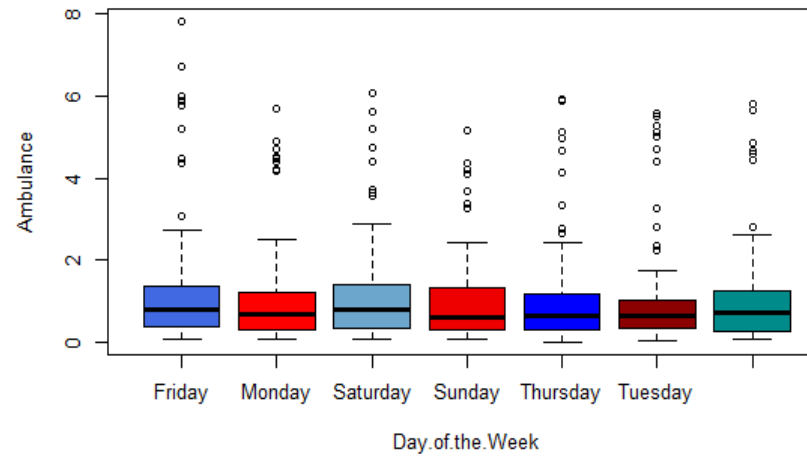
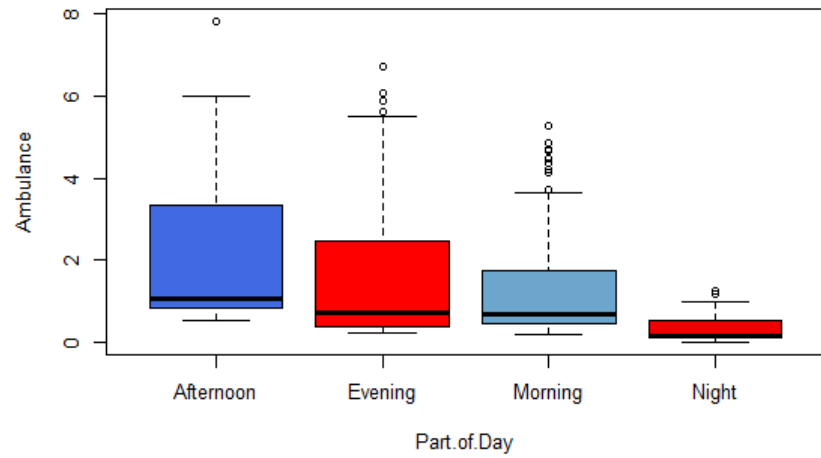
Cleaned Data Averaged Over Days



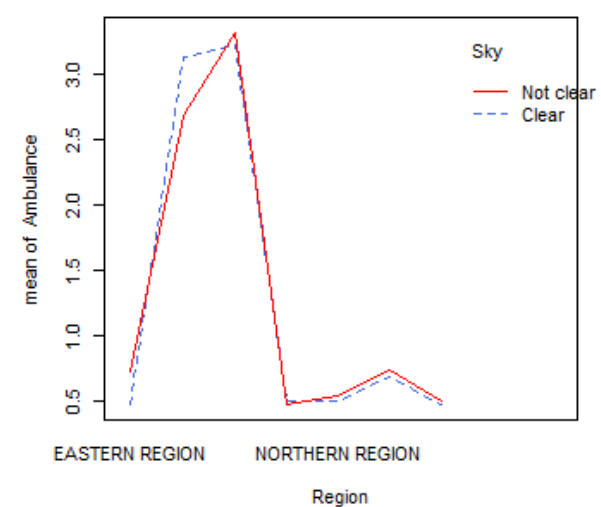
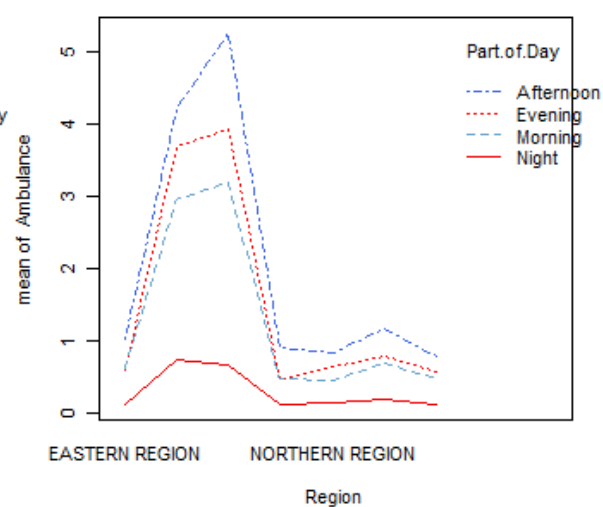
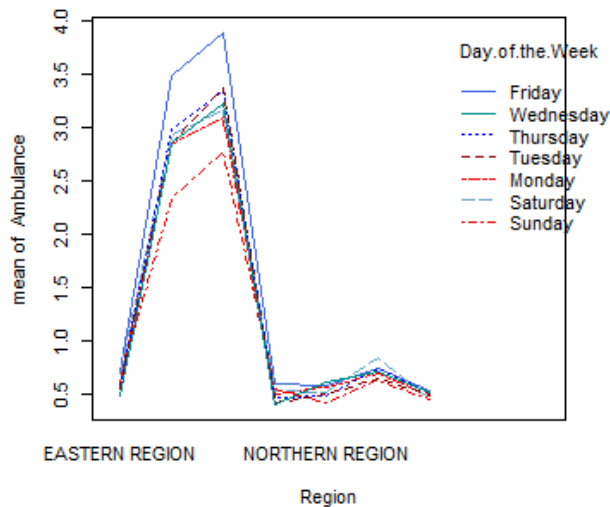
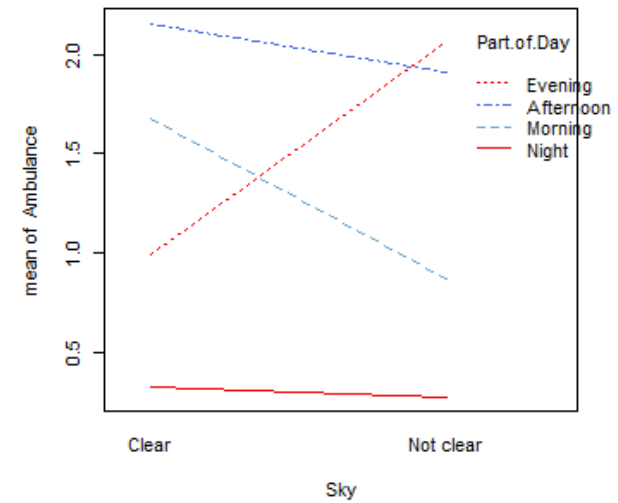
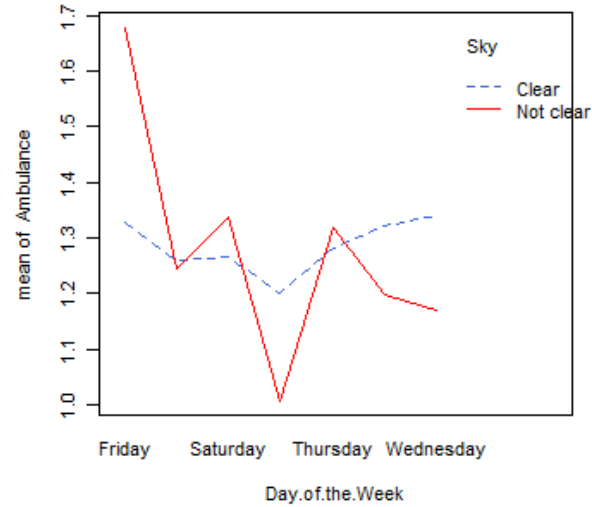
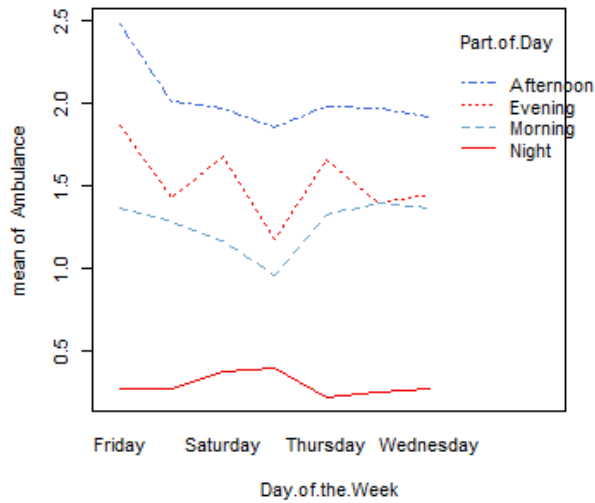
The background features a collage of semi-transparent bar charts and geometric shapes. On the left, a bar chart shows data for quarters Q2 and Q3. In the center, another bar chart shows data for Q3 and Q4. The right side is decorated with overlapping triangles in shades of red and blue. The text 'Final Statistical Model' is centered in a bold, dark red font.

Final Statistical Model

Distribution of Features



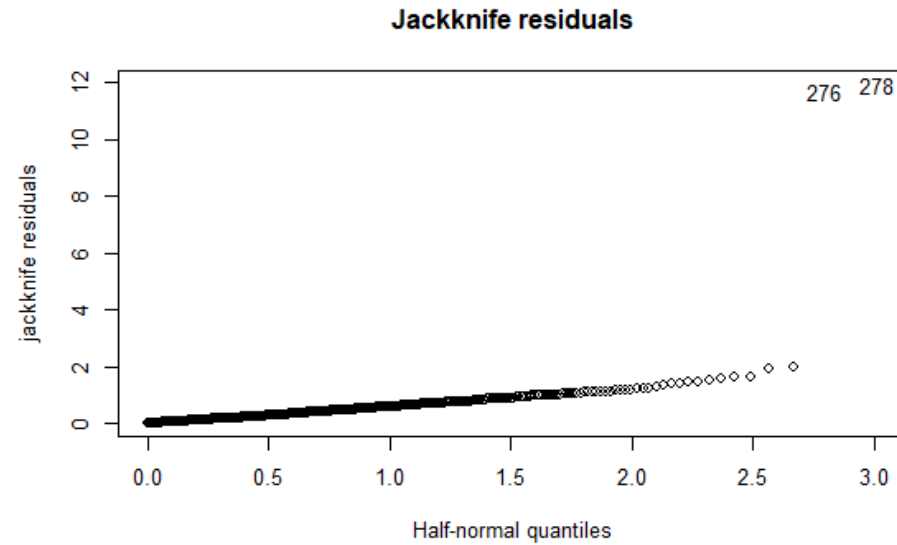
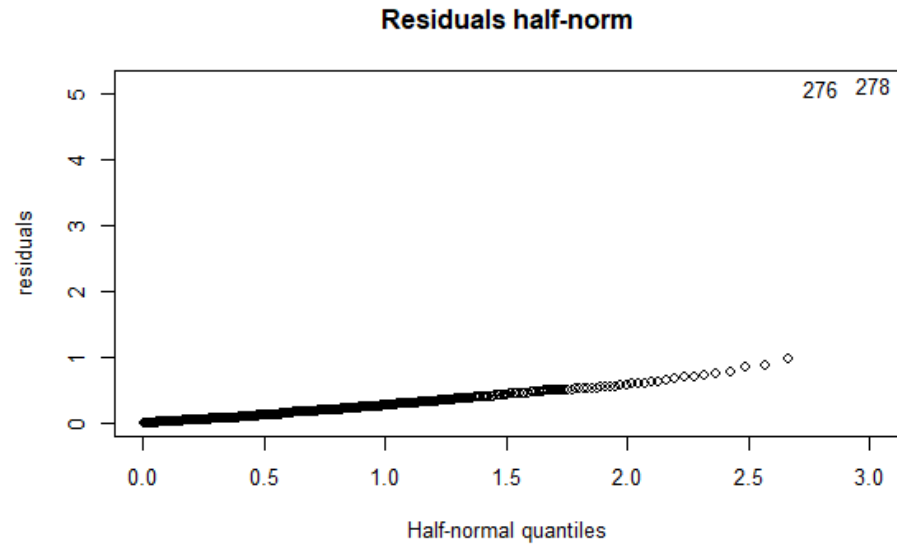
Interaction of Features



Model of choice: Gamma General Linear Model (GLM)

- ▶ Statistical model allows for prediction and understanding inner workings
- ▶ Good for continuous data
- ▶ Good for data between 0 and infinity

Oh oh! outliers!

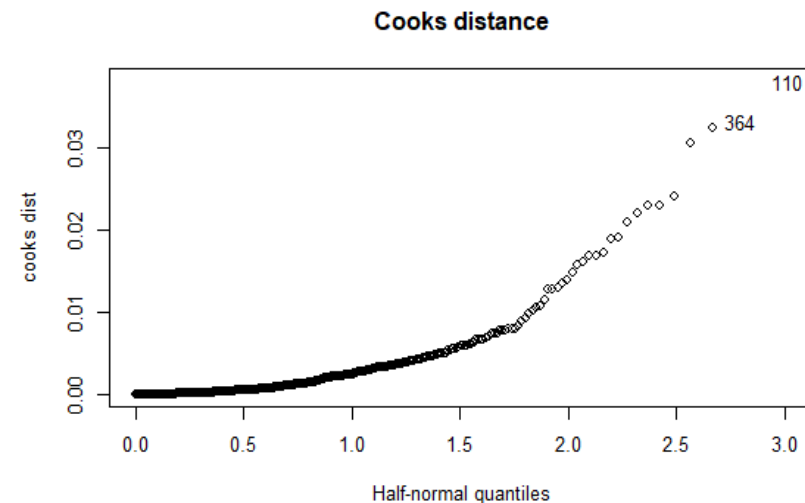
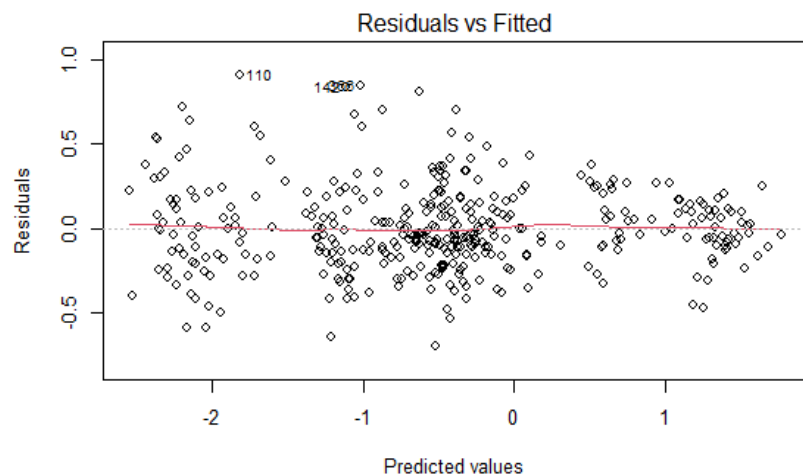
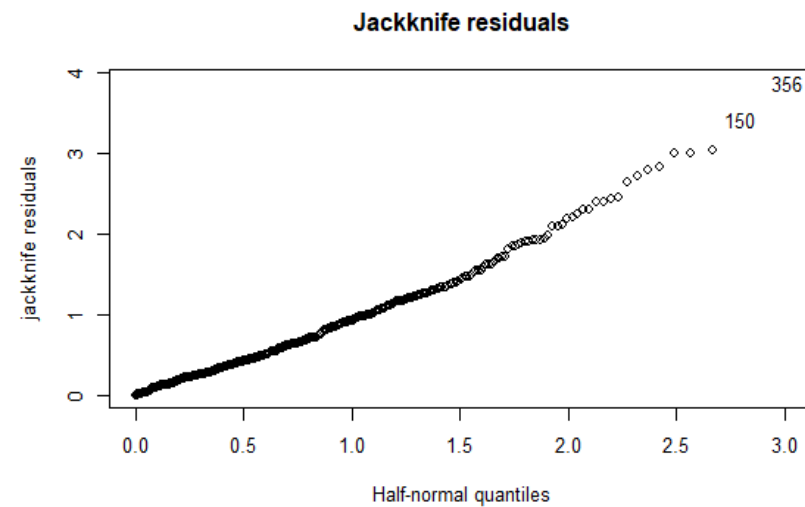
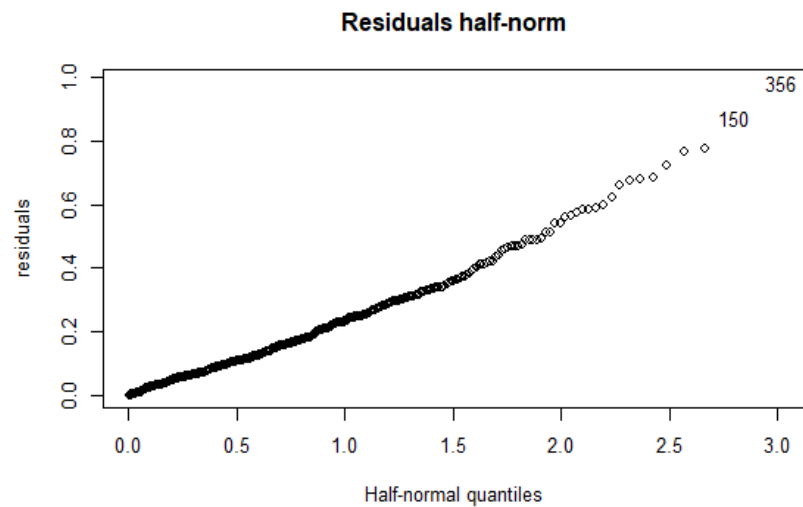


Model Formula

Gamma GLM

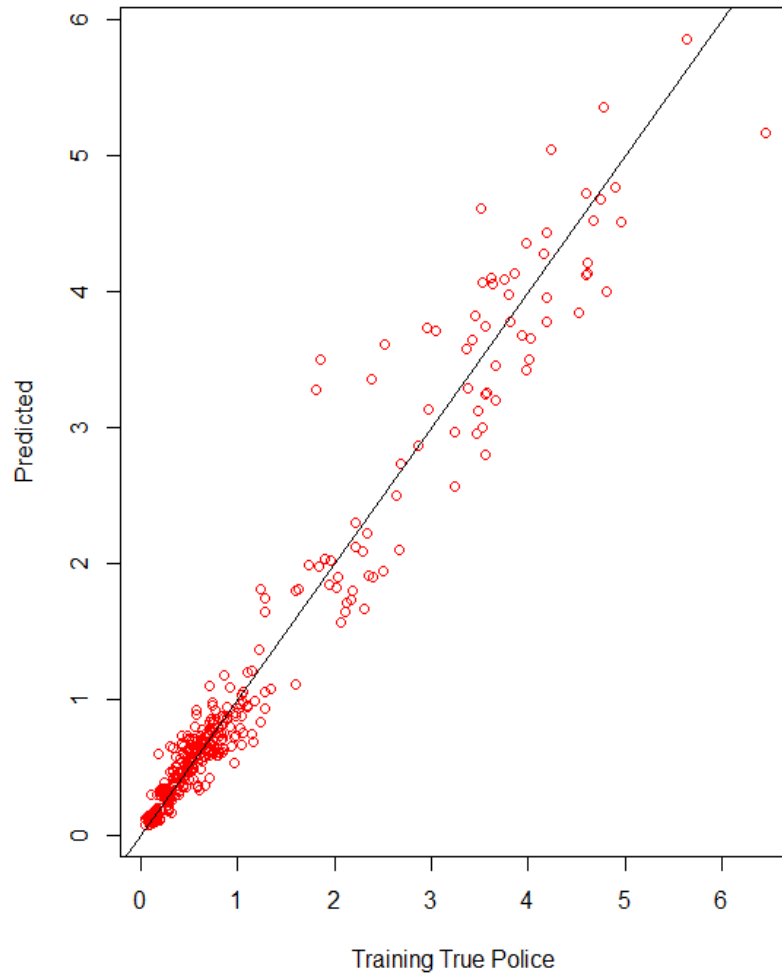
- ▶ Region
 - ▶ Day of the Week
 - ▶ Part of Day
 - ▶ Sky condition
-
- ▶ 2-way interaction between
 - ▶ Day of the Week
 - ▶ Part of Day
 - ▶ Sky condition

Model Diagnostic Plots

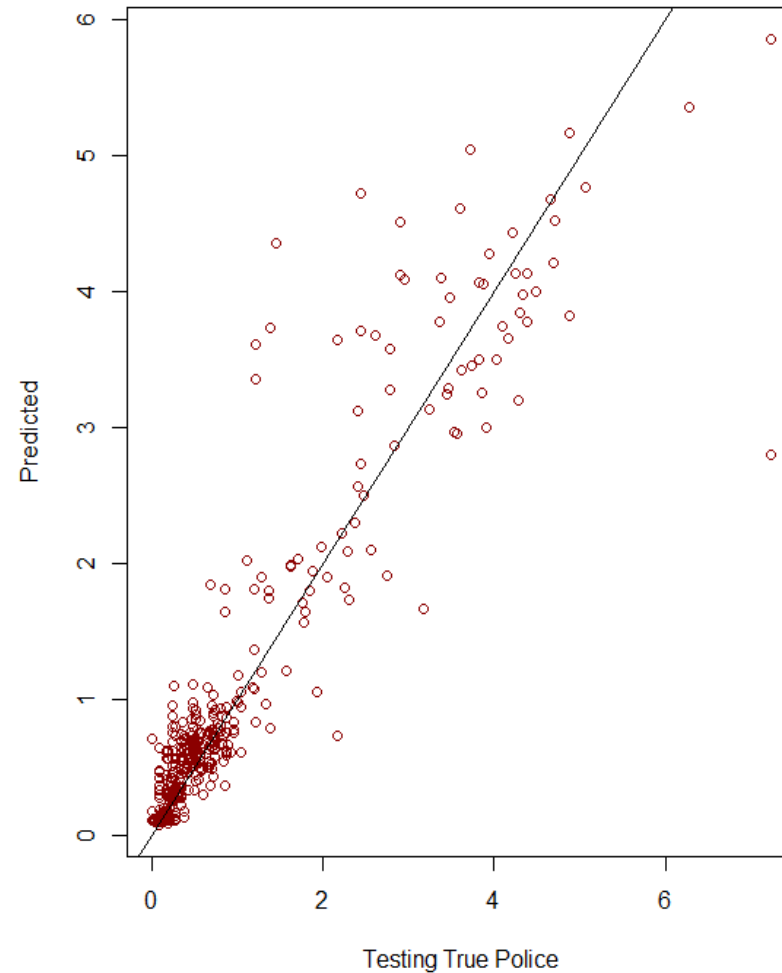


Predicting Police Demand

Training Data

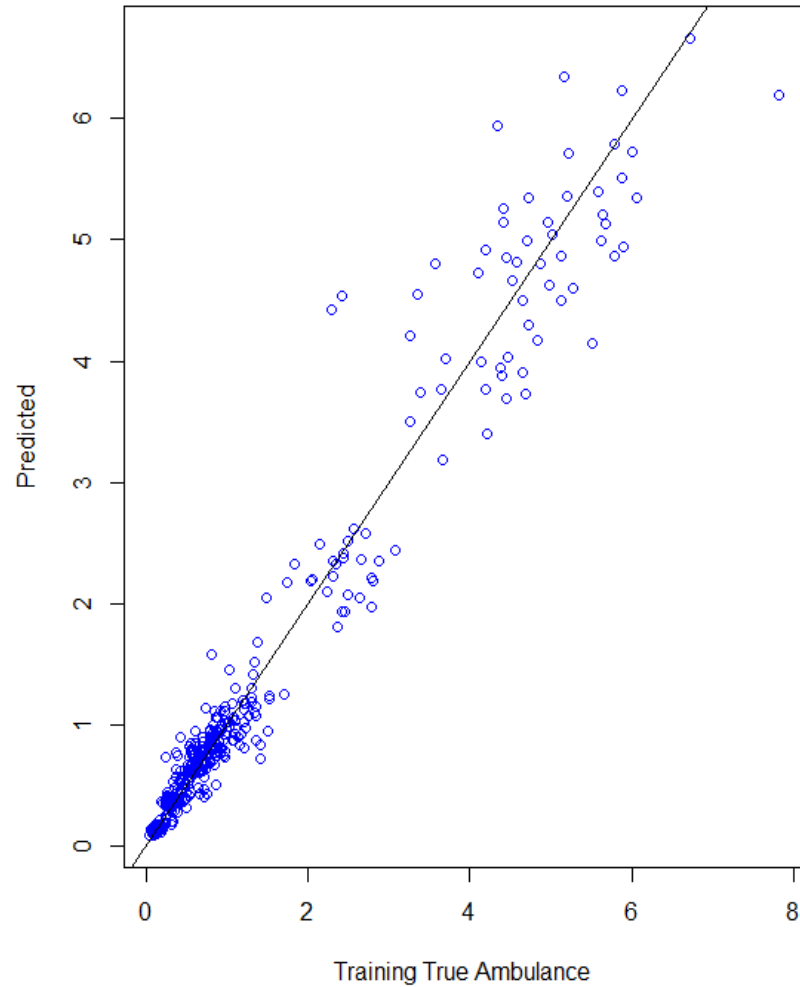


Testing Data

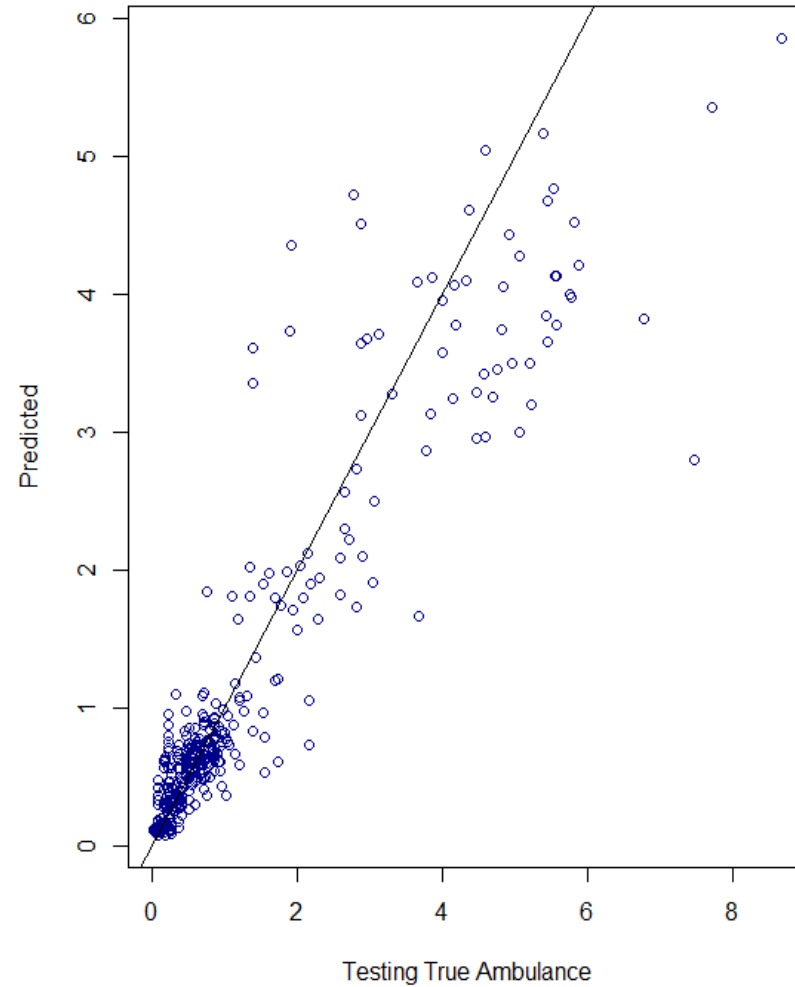


Predicting Ambulance Demand

Training Data



Testing Data



Statistical results

- ▶ Metropolitan areas greatly increase the chance of need an ambulance
- ▶ Evenings and nights decreases risk
 - ▶ Except Saturday and Sunday night
- ▶ Mornings increased risk
- ▶ Weather being not clear in the evening increased risk quite a bit
 - ▶ Whereas raining in the morning decreased risk
- ▶ Raining itself wasn't significant

The background is a complex geometric composition. On the left, a dark red area contains a pattern of overlapping white question marks. A large, white, 3D-style question mark is centered in the middle of the slide. To the right, there are several overlapping triangular and polygonal shapes in shades of red and blue. The word "Questions!" is written in a red, sans-serif font, centered over the large white question mark.

Questions!