# Walmart Analysis

USING R STUDIO

# Brief

- A quick project to use R Studio to generate data visualisations within R Studio.

- The .csv data file was opened in Microsoft Excel to check for null values, to assign the correct format to the columns and to check for duplicates.

- R Studio was then used to write the code that would allow for the production of four data visuals:
  - Histogram
  - Boxplot
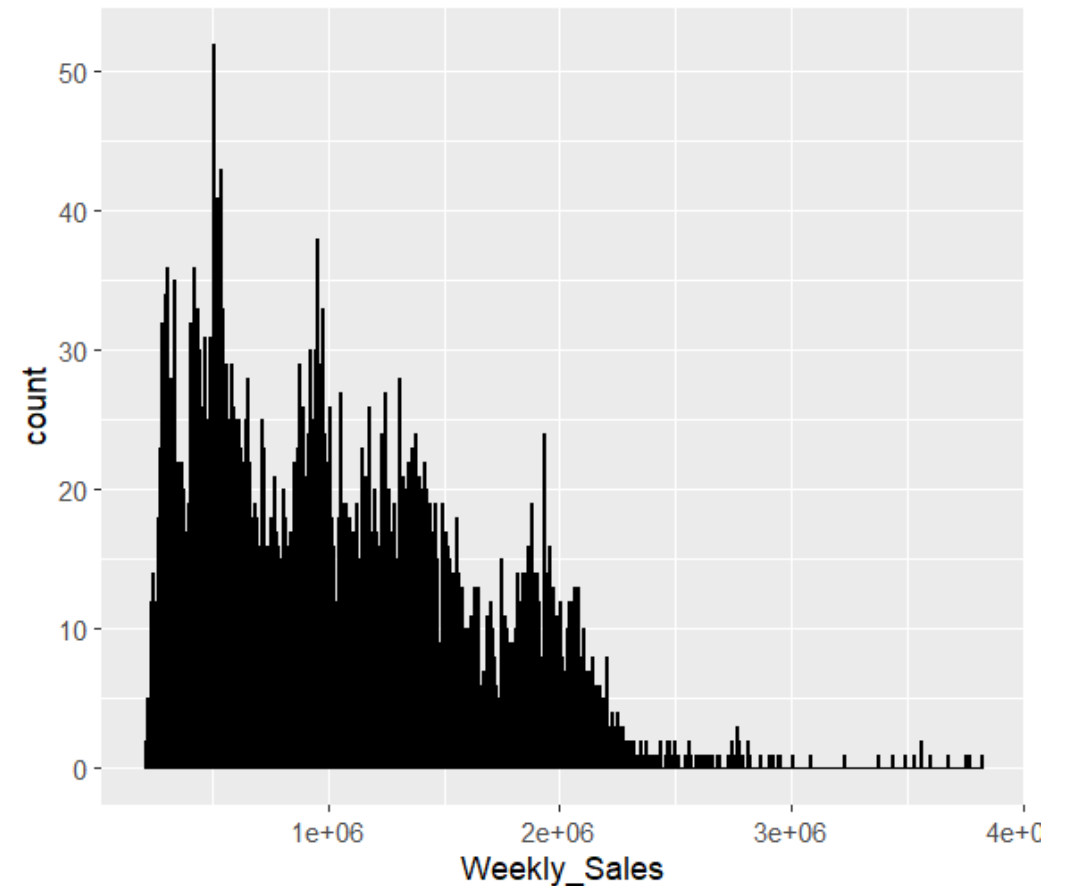  - Scatter Plot
  - Time Series Plot

# R Studio
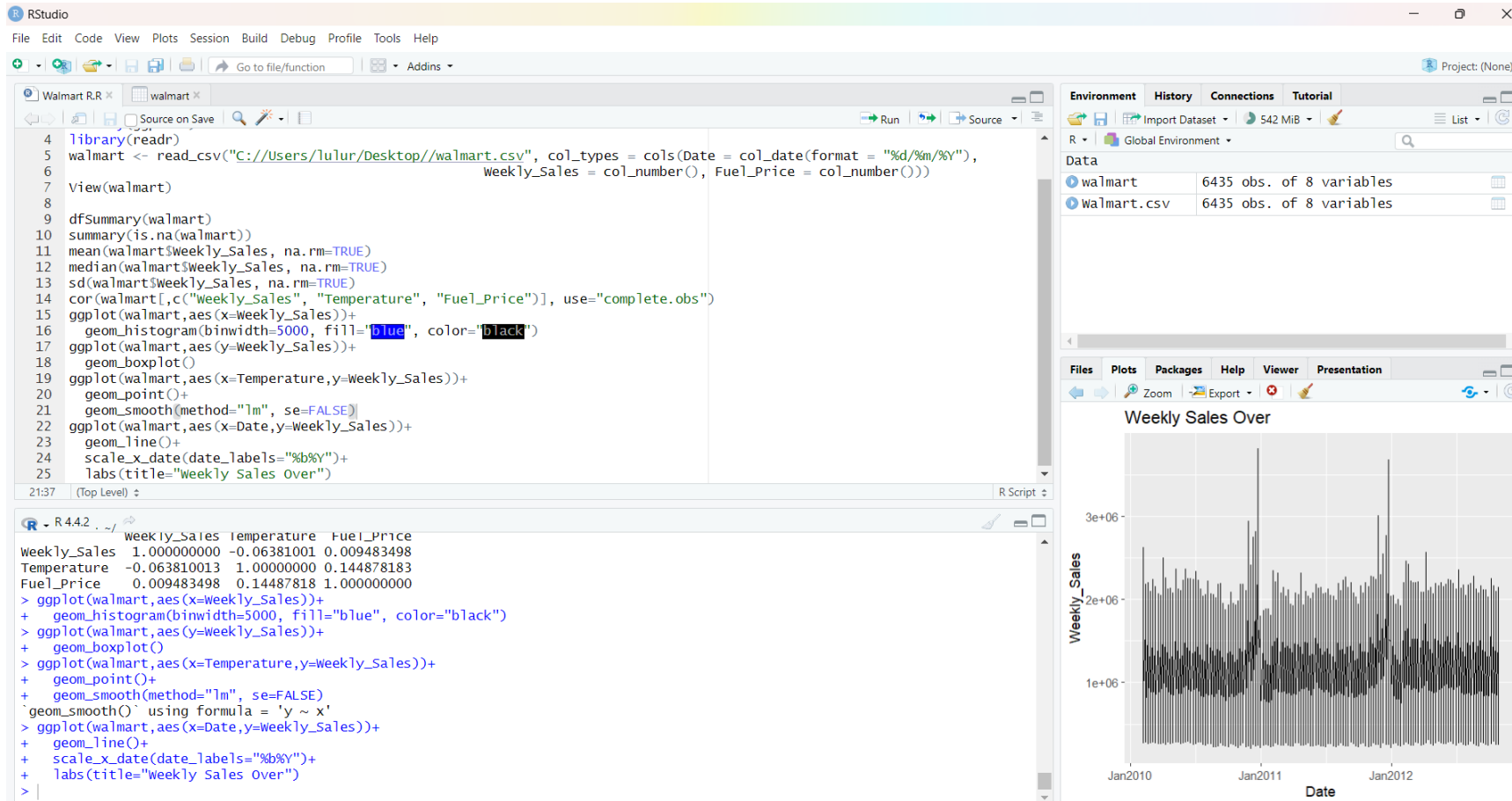
# Data in R Studio
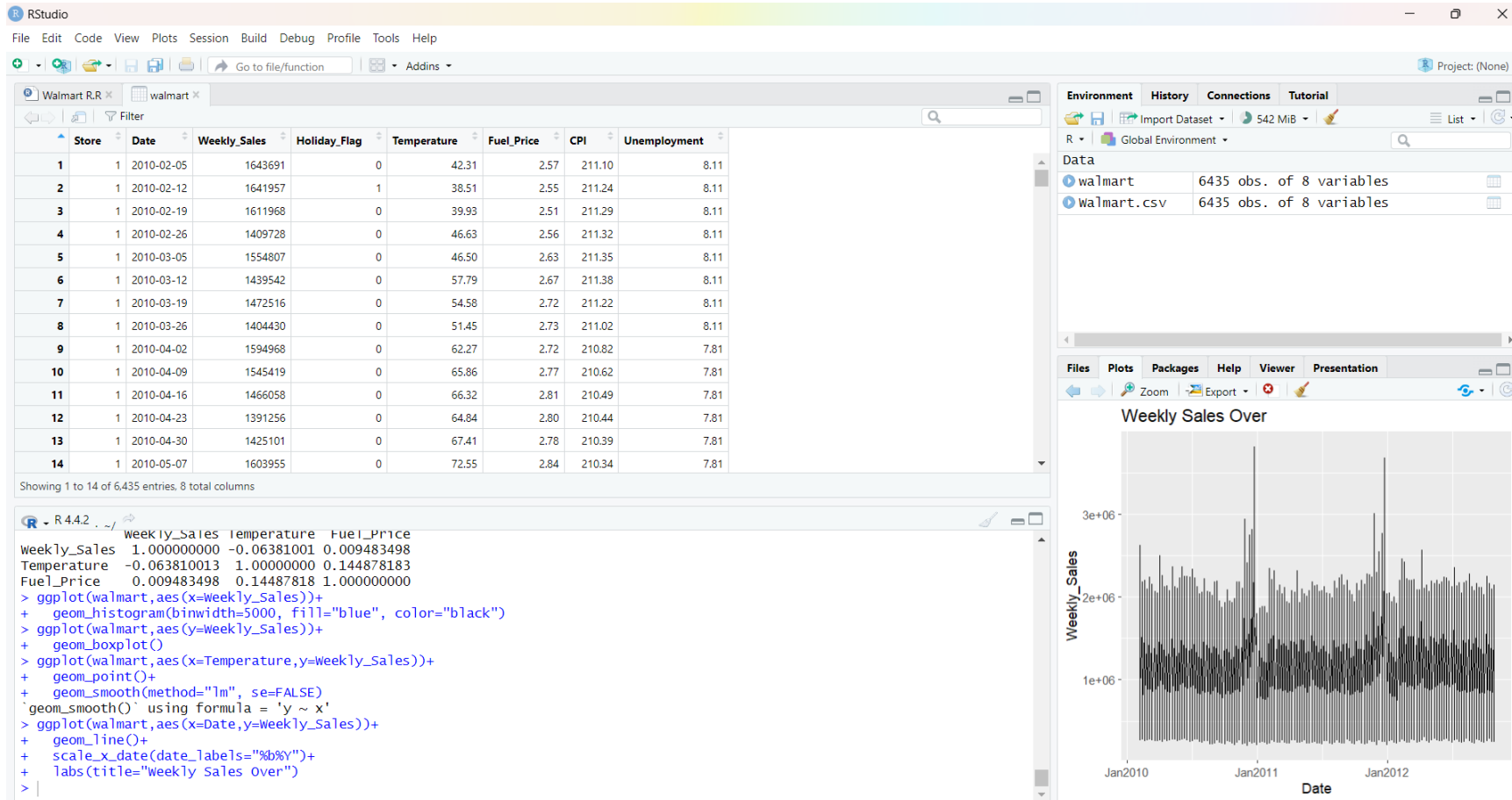
# Histogram

- The histogram is skewed to the right which indicates that a number of outliers, visualised on the right hand side of the graph as a 'tail' are pulling the mean away from the medium.

- When using R Studio to calculate the mean the value was £1,046,965 compared to the lower median value of £960,746.
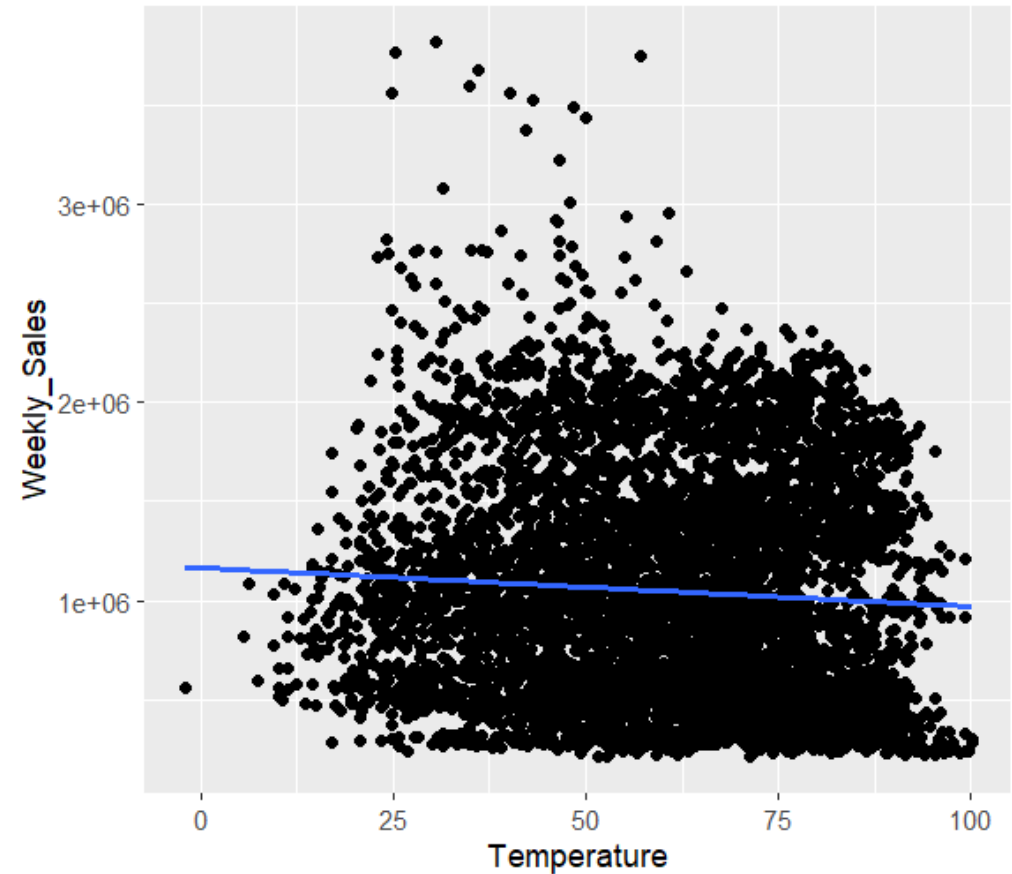
# Box Plot

■The outliers in the tail section of the Histogram can be more clearly seen in the Box Plot Diagram.
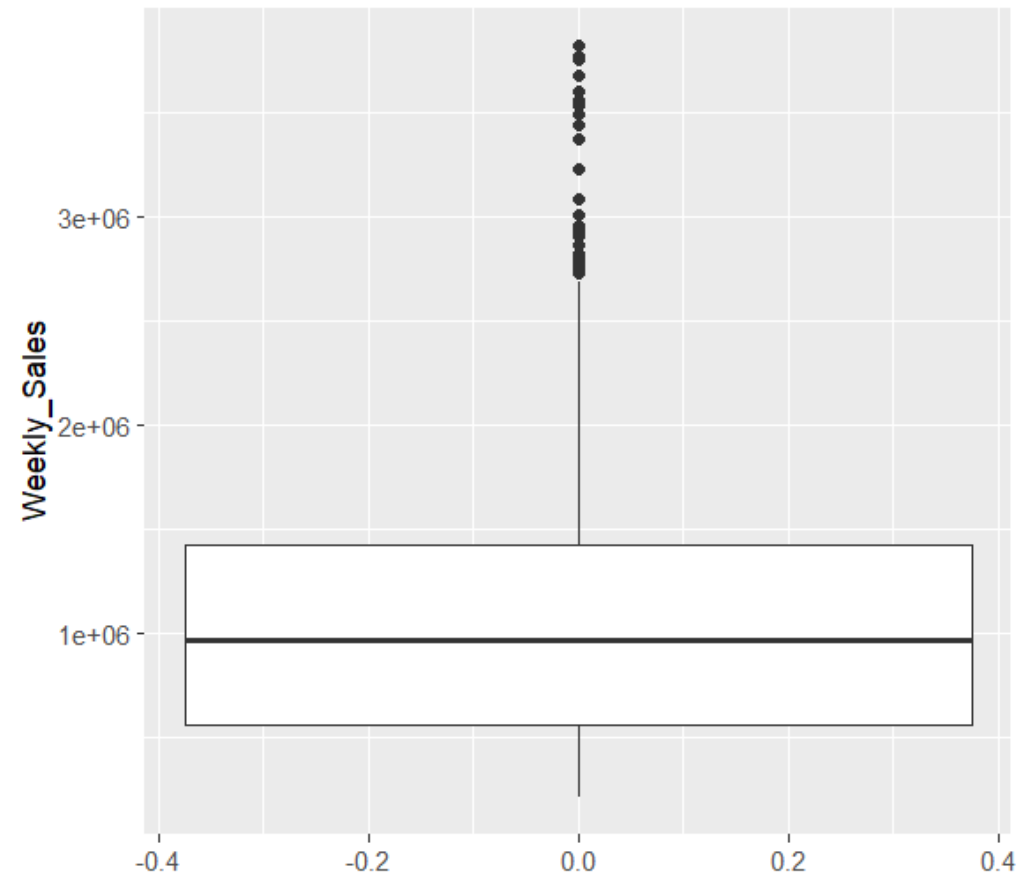
# Scatter Plot

- The Scatter Plot shows there is a very weak negative correlation between weekly sales and temperature, showing that there is virtually no relationship between the impact of temperature on sales.

# Time Series

- The Time Series indicates that in the lead up to the American holiday season (November-December) show quite strong spikes in sales.

- This indicates that the weeks leading up to the holiday period show that Walmart customers increase their spending at Walmart.



Weekly Sales Over
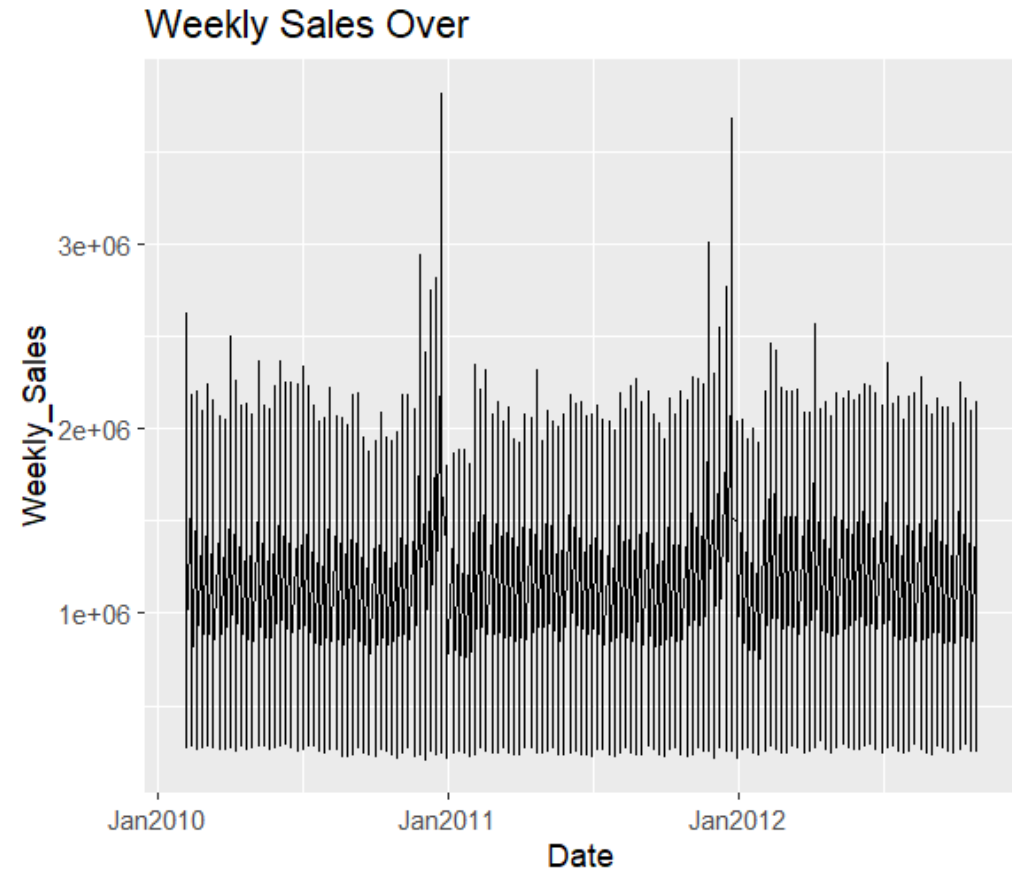
# Recommendations

- It was be to treat the holiday period as an outlier for the data and analyse this information separate to draw conclusions on what offers, if any the company would like to run.

- It is recommended that the period outside of holiday period be treated as separate and to run offers and recommendations based on that information.