Descriptive Statistics:

Table 1 presents an overview of weekly shopping habits, covering spending, queuing time, family size, store type, shopping method, and household income.
Explanation:
1. Spend (Spending): The average spending is approximately £103.59, with a range of £46.79 to £159.71. The typical spending is £104.12.
2. Qtime (Time Spent Queuing): On average, queuing time is around 6.90 units, with a typical queuing time of 7 units. The standard deviation is approximately 2.71.
3. Famsize (Family Size): The average family size is around 3.95, with a standard deviation of 1.97.
4. Income: The average household income is approximately £37,745.97, with a range of £24,739.49 to £53,183.45. The typical income is £37,529.36, with a standard deviation of £4442.34.

These statistics provide a quick overview, summarizing the average (Mean), middle value (Median), variability (Standard Deviation - SD), minimum (Min), maximum (Max) values, and the count of observations for each variable.

```
Descriptive Statistics:
> print(descriptive_table)
  Variable       Mean     Median          SD         Min         Max Count Mode
1    Spend   103.5933   104.1174   16.547940    46.78897   159.7126   500   NA
2    Qtime     6.8960     7.0000    2.706998     0.00000    16.0000   500   NA
3  Famsize     3.9540     4.0000    1.972731     0.00000    16.0000   500   NA
4    Store         NA         NA          NA          NA         NA   500    1
5 Shoptype         NA         NA          NA          NA         NA   500    0
6   Income 37745.9747 37529.3561 4442.342859 24739.49204 53183.4513   500   NA
```

Table 1: Descriptive Statistics for Spend, Qtime, Famsize, and Income and mode for shoptype and store

The analysis (Table 2) delves into the distribution of shopping types (Shoptype) within the dataset, shedding light on the frequency and proportion of each category. In-store shopping (type 0), is the most prevalent, occurring 370 times with a proportion of 74%. Online collection (type 1) follows with a frequency of 123, making up to 24% of the dataset. And online home delivery (type 2), occurs seven times, constituting 0.14% of our observations.

```
Shoptype Table:
> print(shoptype_table)
  Shoptype Frequency Proportion.Var1 Proportion.Freq
1        0       370              0           0.740
2        1       123              1           0.246
3        2         7              2           0.014
>
```

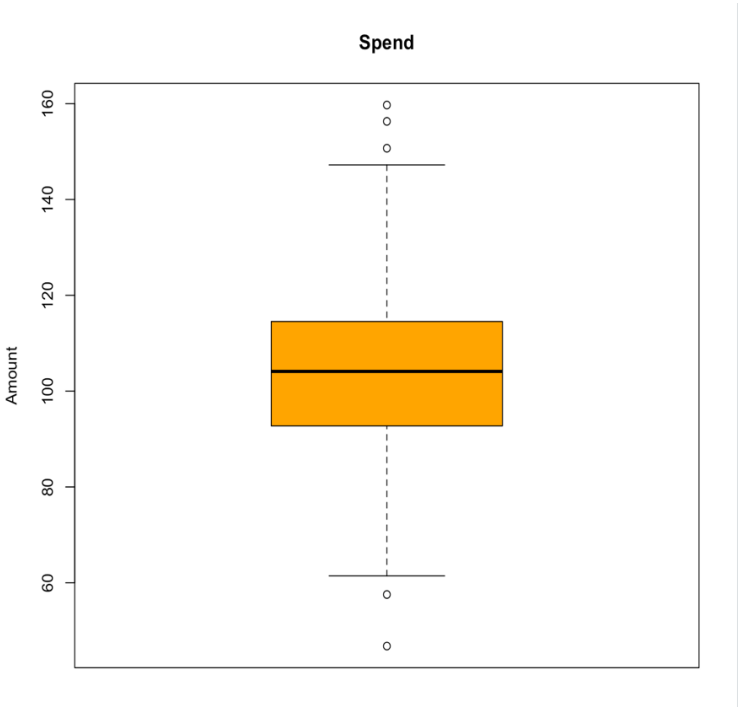
Table 2: Descriptive statistics for shoptype

Similarly for store types (store) (Table 3) urban stores (type 1) makes up 34% of the data distribution followed by rural stores (type 2) at 33% and city centre stores(type 3). The frequency distribution is almost the same here for each category.

```
Store Table:
> print(store_table)
  Store Frequency Proportion.Var1 Proportion.Freq
1     1       170              1            0.34
2     2       165              2            0.33
3     3       165              3            0.33
```

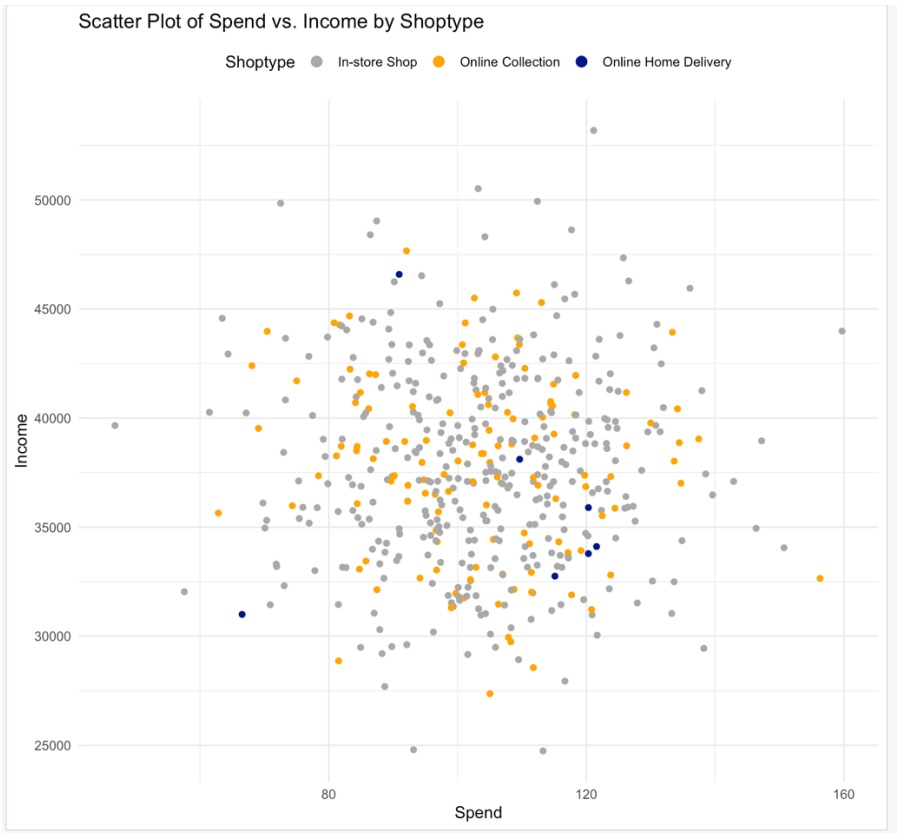
Table 3: Descriptive statistics for store

Visualising Spend: Boxplot



The box plot makes it clear that a majority (50%) of the spending data falls within the range of 80-120 pounds, and the typical spending amount, or median, is around 100 pounds. The highest spending observed was more than 140 pounds, while the lowest recorded spending was 60 pounds. It's important to note that there are a few instances, represented as outlier dots on the graph, where spending deviates from this general pattern.

Relationship between spend and income:



The scatter plot highlights a clear relationship between income and spending for all three shopping types. Despite online home deliveries being less frequent compared to in-store and online collection, all the three categories consistently align with the income range of 30,000 to 45,000. Moreover, spending for all shopping categories tends to concentrate between 80 to 120 pounds.

ANOVA: Testing association between shoptype and spend

Hypothesis:
Null Hypothesis (H0): There is no significant difference in spend across the levels of shoptype
Alternative Hypothesis (H1): There is a significant difference in spend across the levels of shoptype

```
             Df Sum Sq Mean Sq F value Pr(>F)
shoptype      2     94   47.25   0.172  0.842
Residuals   497 136549  274.75
```

The statistical test looked at spending differences among various shopping types. The outcome, with a p-value of 0.842, suggests there isn't enough evidence to claim a significant spending disparity between the different shopping methods. In conclusion, the results don't show a clear trend indicating that one type of shopping significantly affects the amount people spend.

Corelation: Testing association between qtime and spend

```
> # Choose pair of variables
> correlation_variables <- c("qtime", "spend")
>
> # Calculate correlation
> correlation_result <- cor(weeklyshop[, correlation_variables])
>
> # Extract the correlation coefficient
> correlation_coefficient <- correlation_result[1, 2]
>
> # Print correlation result
> cat("Correlation between", correlation_variables[1], "and", correlation_variables[2], ":", correlation_coefficient, "\n")
Correlation between qtime and spend : 0.0743494
> install.packages("ggplot2")
```

The correlation coefficient between the time spent queuing at checkout/ collection point (qtime) and the amount spent during the weekly supermarket shopping is 0.0743494. This positive correlation suggests a minor tendency for individuals who spend more time queuing to also spend more. However, the correlation is not very strong, indicating that the relationship is not highly impactful. In conclusion there is a connection between queuing time and spending, but it's not significant enough to reliably predict one variable based on the other for any individual.