Spring 2021 **Firstname M. Lastname**

University of Central Florida
College of Business

QMB 6911
Capstone Project in Business Analytics

Solutions: Problem Set #5

0.1 Introduction

This note summarizes the findings in the script `Tractor_Data_Vis.R`, which analyzes the prices of fly reels, the dependent variable in the `TRACTOR7.csv` dataset. The output includes several plots of the dependent variable against the explanatory variables.

The primary goal is to determine the relative value of John Deere tractors compared to others. A secondary consideration is to determine the time of year in which to sell a tractor.

0.2 Histogram and Density of Log. Fly Reel Prices

0.2.1 All Tractors Together

Start with the log of sale prices because that seemed more promising, given that prices were highly skewed.

```
hist(tractor_sales[, 'log_saleprice'],
    main = 'Histogram and Density of Log. Tractor Prices',
    xlab = 'Price',
    col = 'red',
    probability = TRUE)
rug(tractor_sales[, 'log_saleprice'])
lines(density(tractor_sales[, 'log_saleprice']),
     col = 'blue',
     lwd = 3)
```

Figure 1 is a histogram of the logarithm of tractor prices, along with a rug plot and a kernel density estimate, generated by the code block above. After taking logs, we can see that the distribution is approximately sym-

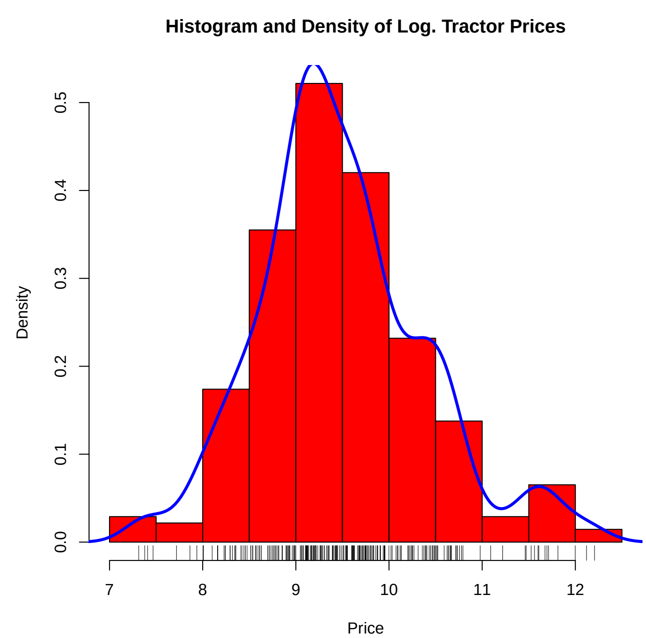

Figure 1: Relative Histogram of Tractor Prices

metric, with some bunching in the upper tail.

0.2.2 Comparison By Make

Now we investigate the value of John Deere tractors compared to other brands. Figure 2 shows the kernel density estimate of the sale prices of John Deere tractors in green and that of the other brands in red. The distribution of sale prices of John Deere tractors has several modes and is skewed to the right, with the highest mode lower than that for other brands.

```
library(sm)
sm.density.compare(tractor_sales[, 'log_saleprice'],
                   tractor_sales[, 'johndeere'],
                   xlab = "Log. of Sale Price",
                   lwd = 3,
                   col = c('red','green'))
title(main = "Log. of Sale Price by Brand")
legend('topright', c('Other', 'John Deere'),
       fill = c('red','green'),
       cex = 0.75)
```

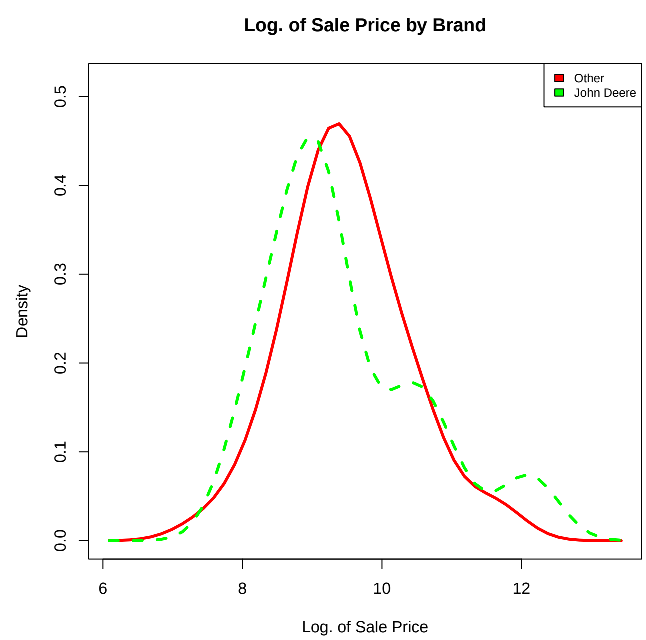

Figure 2: Densities of Log. Tractor Prices by Brand

0.2.3 Comparison By Season of Sale

Figure 3 shows the densities of the logarithm of sales price, separated by the season of the year in which they were sold. The figure was generated by the following code.

```
plot(density(tractor_sales[tractor_sales[, 'fall'] == 1,
   'log_saleprice']),
    col = 'orange',
    lwd = 3,
    xlim = c(min(tractor_sales[, 'log_saleprice']),
           max(tractor_sales[, 'log_saleprice'])),
    main = 'Log. of Sale Price by Season of Sale')
# Plot the rest.
lines(density(tractor_sales[tractor_sales[, 'spring'] == 1,
   'log_saleprice']),
    col = 'green',
    lwd = 3)
lines(density(tractor_sales[tractor_sales[, 'summer'] == 1,
   'log_saleprice']),
     col = 'yellow',
     lwd = 3)
lines(density(tractor_sales[tractor_sales[, 'winter'] == 1,
   'log_saleprice']),
     col = 'blue',
     lwd = 3)
legend('topright',
      c('Spring', 'Summer', 'Fall', 'Winter'),
      fill = c('green', 'yellow', 'orange', 'blue'),
      cex = 0.65)
```

We see that the distribution of sales is similar during summer and fall (yellow and orange, respectively) with some bunching in the upper tail. We also see more variance in the sale price in winter, shown in blue. In the spring, shown in green, the mode of the distribution is higher.

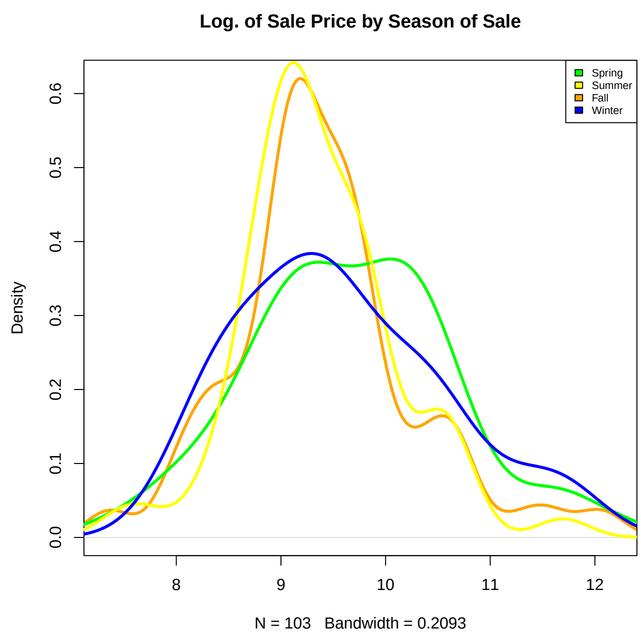

Figure 3: Relative Histogram of Tractor Prices

0.3 Sales Volume By Brand and Season of Sale

Figure 4 shows a spinogram of the number of sales by brand and the time of year the tractor is sold. It is generated by the following code.

```
# Create a table and plot it in a spinogram.
counts <- table(tractor_sales[, 'season'],
            tractor_sales[, 'JD'])

# Plot the spinogram.
spine(counts,
     main = "Spinogram of Sales by Brand and Season")
```

It appears that sales of John Deere tractors are fairly evenly distributed throughout the year, with more sales of John Deere tractors in the winter months and a fewer sales in the summer.

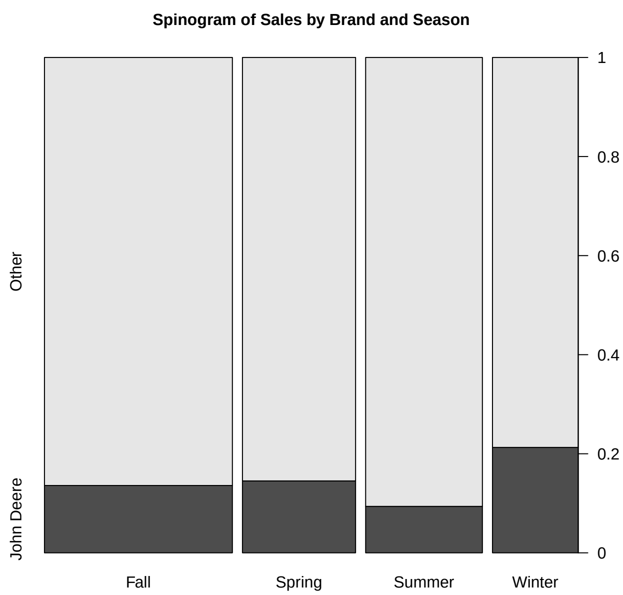

Figure 4: Sales Volume of Tractor Prices by Brand and Season of Sale