Expedia Case Analysis on Customer Sensitivity to Hotel Price Changes

Prarabdh Bhatia

Introduction

The analysis involved examining how changes in hotel prices affect booking likelihood and the number of nights booked. Several models were used to quantify this relationship, and coefficients were interpreted to express the percentage change in booking likelihood or nights booked for a $100 increase in price. The key results and insights are summarized below.

Price Sensitivity Across Models

1. Diagram: Model 1A (Booking Likelihood):

   - The coefficient for `PricePerNight` is -0.0007498. This implies that for every $1 increase in price, the booking likelihood decreases by 0.07498%.

     $-0.0007498 \times 100 = -0.07498$ or $-7.498\% -0.0007498 \times 100 = -0.07498$ or -7.498%
   - Therefore, for a $100 increase in price, the likelihood of a booking decreases by 7.498%.

2. Diagram: Model 1B (Booking Likelihood):

- The coefficient is -0.000760, meaning that a $1 increase in price reduces booking likelihood by 0.076%.

  `-0.000760×100=-0.0760 or -7.60%-0.000760×100=-0.0760 or -7.60%`

- For a $100 price increase, the reduction in booking likelihood is 7.60%.

3. Diagram: Model 1C (Booking Likelihood):

- With a coefficient of -0.0008138, the booking likelihood decreases by 8.138% for a $100 increase in price.

  `-0.0008138×100=-0.08138 or -8.138%-0.0008138×100=-0.08138 or -8.138%`

4. Diagram: Model 2A (Nights Booked):

- The coefficient is -0.002891, indicating that a $1 increase in price reduces the number of nights booked by 0.2891%.

  $-0.002891 \times 100 = -0.2891$ or $-28.91\% -0.002891 \times 100 = -0.2891$ or -28.91%

- For a $100 price increase, the reduction in nights booked is 28.91%.

5. Diagram: Model 2B (Nights Booked):

- With a coefficient of -0.0029206, the number of nights booked decreases by 29.21% for a \$100 increase in price.

  $-0.0029206\times100=-0.2921$ or $-29.21\%-0.0029206\times100=-0.2921$ or -29.21%

6. Diagram: Model 2C (Nights Booked):

- The coefficient is -0.003078, meaning that a \$100 price increase reduces the number of nights booked by 30.78%.

  $-0.003078\times100=-0.3078$ or $-30.78\%-0.003078\times100=-0.3078$ or -30.78%

Regional and Income-Based Price Sensitivity

The analysis explored how price sensitivity varies by customer income level and travel destination:

By Region

1. Hawaii:
   - The price sensitivity coefficient is -0.000518, meaning that customers in Hawaii are less responsive to price increases compared to other regions. A \$100 price increase reduces booking likelihood by 5.18%.
2. Las Vegas and Miami:
   - These regions are highly sensitive to price changes, with coefficients of -0.00109 and -0.00125, respectively. For a \$100 price increase, booking likelihood drops by 10.9% in Las Vegas and 12.5% in Miami.
3. Washington DC:
   - This region shows the least sensitivity to price changes, with a coefficient of -0.000185, corresponding to a 1.85% reduction in booking likelihood for a \$100 increase in price.

By Customer Income
1. Low-Income Customers:
- The coefficient is -0.00338, indicating that these customers are highly sensitive to price changes. A \$100 increase leads to a 33.8% reduction in the number of nights booked.

2. Medium-Income Customers:
   - With a coefficient of -0.00303, these customers are somewhat less sensitive, with a 30.3% decrease in nights booked for a \$100 price increase.
3. High-Income Customers:
   - High-income customers are the least sensitive to price changes, with a coefficient of -0.00257, resulting in a 25.7% reduction in nights booked for a \$100 price increase.

Comparison with Observational Estimates

When comparing the results of the experimental models with observational estimates, the experimental data generally indicates higher price sensitivity. This suggests that customers might be more responsive to price changes in controlled environments, or that previous models may have underestimated the effect of price increases. Such findings highlight the importance of using experimental data to capture a more accurate understanding of customer behavior.

Implications for Pricing Strategy

The results suggest that Expedia can leverage this information to develop region-specific and income-targeted pricing strategies:

Dynamic Pricing: The data indicates that in regions like Las Vegas and Miami, where price sensitivity is higher, Expedia can adopt a more cautious pricing strategy, particularly during peak travel seasons. Targeted Discounts for Low-Income Customers: Offering discounts or promotions to low-income customers could help retain this price-sensitive segment, particularly in regions like Miami and Las Vegas. Opportunities in Low-Sensitivity Markets: In regions such as Washington DC, and for high-income customers, there is more flexibility to increase prices without drastically affecting bookings.

Limitations and Future Directions

While the models provide valuable insights, there are limitations: Seasonality: The analysis does not account for seasonal fluctuations, which could impact price sensitivity. Future studies could include a time-based variable to better capture these effects. Customer Behavior Segmentation: Additional segmentation by customer behavior (e.g., loyalty program membership, frequent travelers) could enhance the understanding of price sensitivity. External Factors: Events, such as holidays or special events in certain regions, could affect bookings but were not included in the models. Incorporating such factors would provide a more comprehensive analysis.

Conclusion

Overall, the findings highlight that price sensitivity varies significantly by both destination and customer income.

Regions like Las Vegas and Miami, and lower-income customers, show higher sensitivity to price changes, suggesting that pricing strategies should be region- and demographic-specific.

This information can help Expedia optimize pricing to cater to different customer segments, balancing between maximized bookings and revenue.

```r
#install.packages("dplyr")
#install.packages("ggplot2")
#install.packages("psych")
#install.packages("moments")
#install.packages("PASWR2")
#install.packages("pwr")
#install.packages("EnvStats")
#install.packages("OneTwoSamples")
#install.packages("lsr")
#install.packages("tidyr")
#install.packages("reshape2")
#install.packages("DescTools")
#install.packages("MASS")

library(dplyr)
```

```
##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##     filter, lag

## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```r
library(ggplot2)
library(DescTools)
library(MASS)
```

```
##
## Attaching package: 'MASS'

## The following object is masked from 'package:dplyr':
##
##     select
```

```r
load("/Users/prarabdhbhatia/Desktop/Projects/Price_sensitivity/expedia_dataset.Rdata")
head(Expedia)
```

```
##   PricePerNight        Region UserIncome Booked? Nights
## 1           253     Las Vegas      32000       0      0
## 2           241        Hawaii      49000       0      0
## 3           272         Miami      28000       0      0
## 4           241 Washinton DC     107000       1      2
## 5           264     Las Vegas      54000       1      3
## 6           259        Hawaii      28000       0      0
```

Summary of Expedia Data Overall

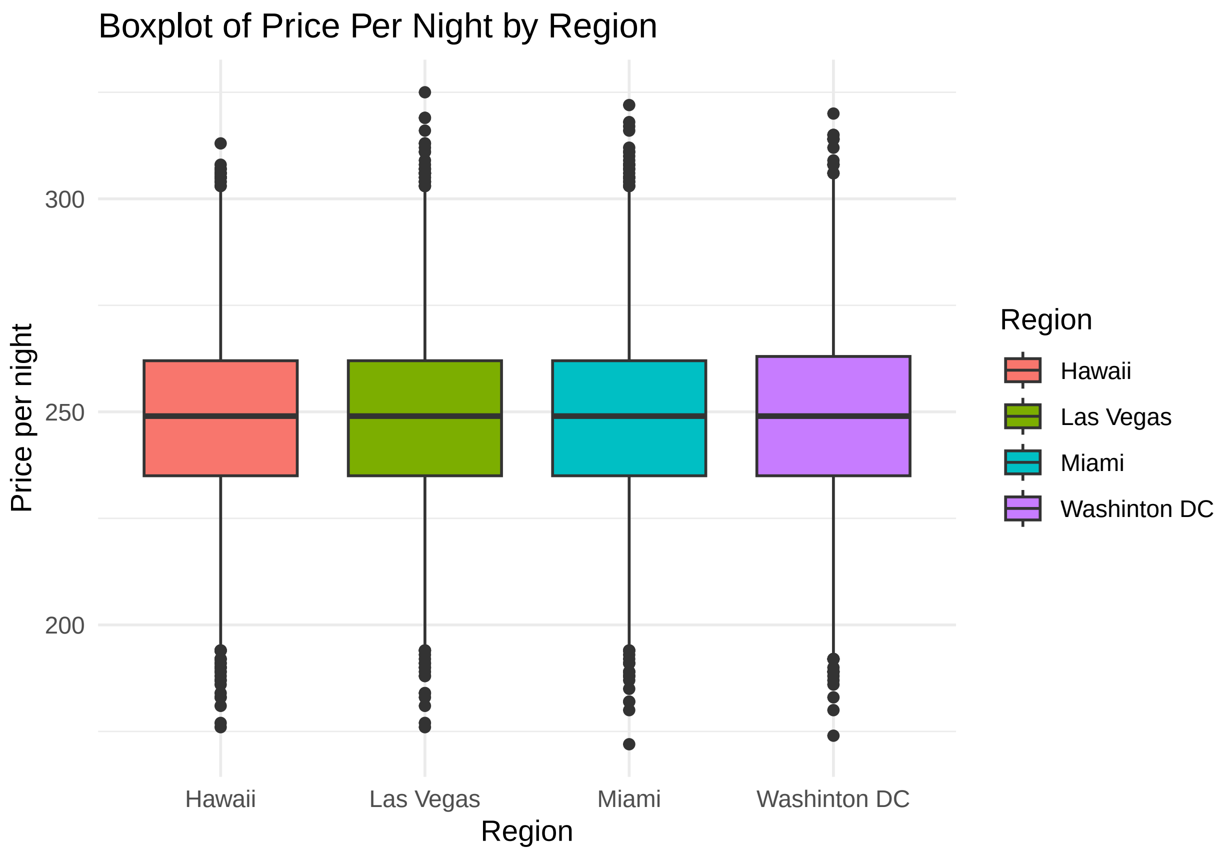
```r
summary(Expedia)
```

```
##  PricePerNight     Region            UserIncome        Booked?
##  Min.   :172.0   Length:25000       Min.   :  4000   Min.   :0.0000
##  1st Qu.:235.0   Class :character   1st Qu.: 31000   1st Qu.:0.0000
##  Median :249.0   Mode  :character   Median : 45000   Median :0.0000
##  Mean   :248.9                      Mean   : 52040   Mean   :0.2598
##  3rd Qu.:263.0                      3rd Qu.: 65000   3rd Qu.:1.0000
##  Max.   :325.0                      Max.   :363000   Max.   :1.0000
##      Nights
##  Min.   :0.0000
##  1st Qu.:0.0000
##  Median :0.0000
##  Mean   :0.7672
##  3rd Qu.:2.0000
##  Max.   :6.0000
```

Boxplot (A)

```r
ggplot(Expedia, aes(x = Region, y = PricePerNight, fill = Region)) +
  geom_boxplot() +
  labs(title = "Boxplot of Price Per Night by Region",
```

```r
       x = "Region",
       y = "Price per night") +
  theme_minimal()
```

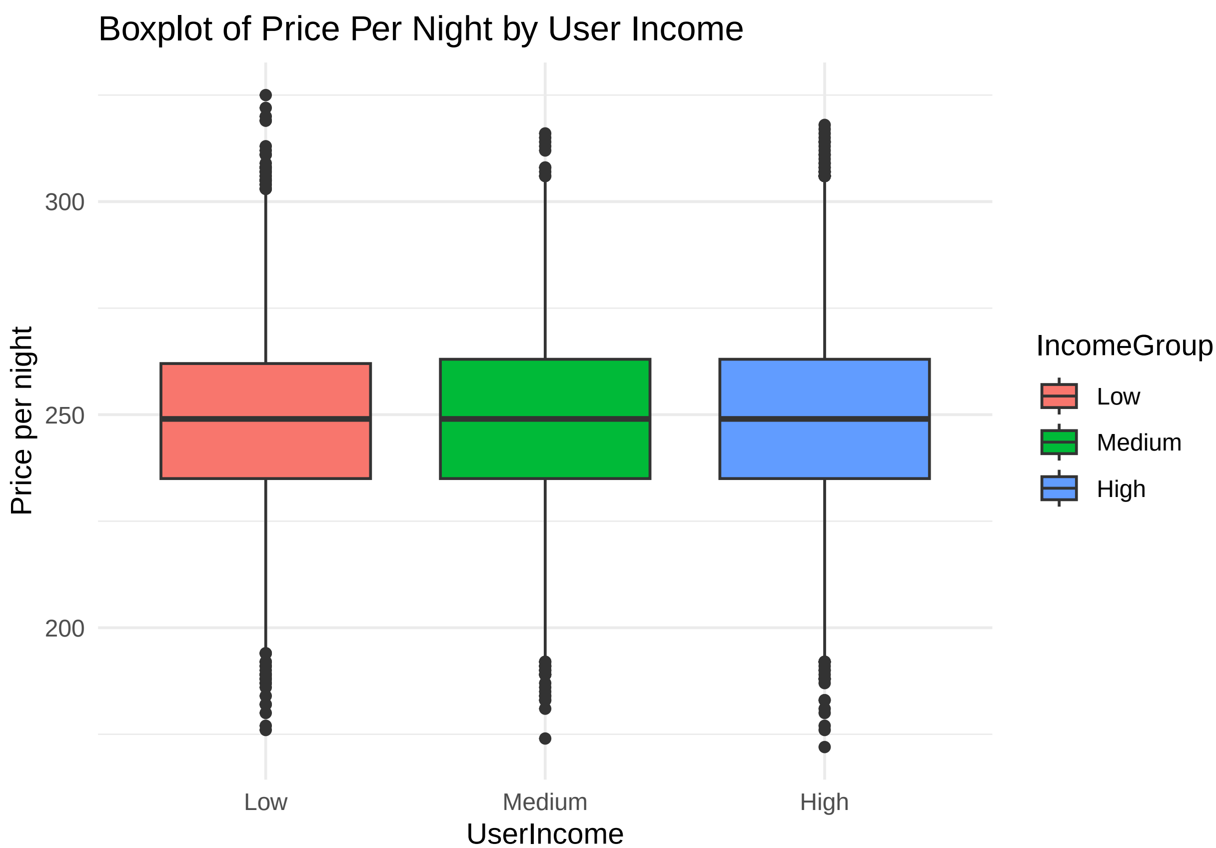



```r
# Calculate means for each region
las_vegas_mean <- mean(Expedia$PricePerNight[Expedia$Region == "Las Vegas"], na.rm = TRUE)
hawaii_mean <- mean(Expedia$PricePerNight[Expedia$Region == "Hawaii"], na.rm = TRUE)
miami_mean <- mean(Expedia$PricePerNight[Expedia$Region == "Miami"], na.rm = TRUE)
washington_mean <- mean(Expedia$PricePerNight[Expedia$Region == "Washinton DC"], na.rm = TRUE)

# Create a data frame with the results
mean_table <- data.frame(
  Region = c("Las Vegas", "Hawaii", "Miami", "Washington DC"),
  MeanPricePerNight = c(las_vegas_mean, hawaii_mean, miami_mean, washington_mean)
)

# Display the table
print(mean_table)
```

```
##          Region MeanPricePerNight
## 1     Las Vegas          248.9000
## 2        Hawaii          248.5662
## 3         Miami          248.9413
## 4 Washington DC          249.1262
```

Boxplot (B)

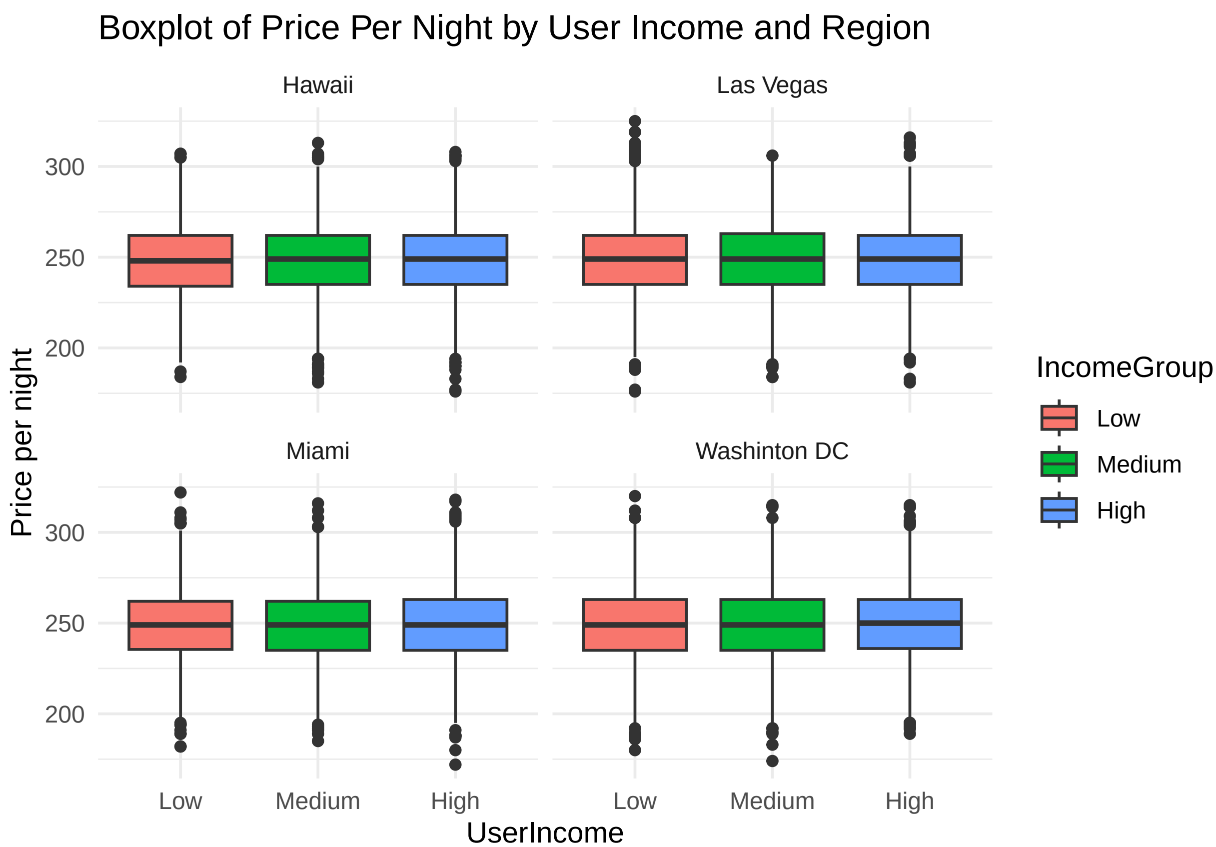
```
Expedia <- Expedia %>%
  mutate(IncomeGroup = cut(UserIncome,
                           breaks = quantile(UserIncome, probs = c(0, 0.33, 0.66, 1), na.rm = TRUE),
                           labels = c("Low", "Medium", "High"),
                           include.lowest = TRUE))

ggplot(Expedia, aes(x = IncomeGroup, y = PricePerNight, fill = IncomeGroup)) +
  geom_boxplot() +
  labs(title = "Boxplot of Price Per Night by User Income",
       x = "UserIncome",
       y = "Price per night") +
  theme_minimal()
```



Boxplot (C)

```
Expedia <- Expedia %>%
  mutate(IncomeGroup = cut(UserIncome,
                           breaks = quantile(UserIncome, probs = c(0, 0.33, 0.66, 1), na.rm = TRUE),
                           labels = c("Low", "Medium", "High"),
```

```r
                           include.lowest = TRUE))

ggplot(Expedia, aes(x = IncomeGroup, y = PricePerNight, fill = IncomeGroup)) +
  geom_boxplot() +
  labs(title = "Boxplot of Price Per Night by User Income and Region",
       x = "UserIncome",
       y = "Price per night") +
  theme_minimal()+
  facet_wrap(~Region)
```

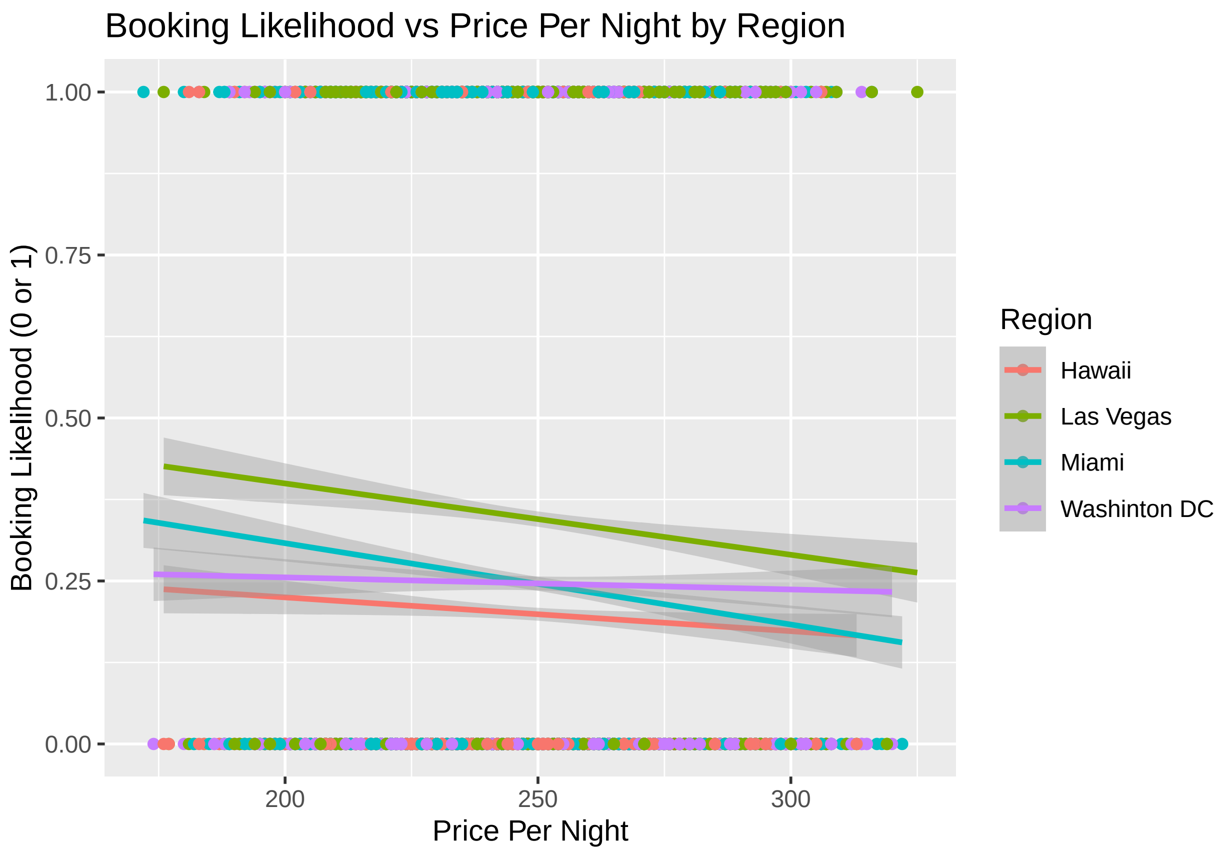


Booked? as Dependent Variable

Graph (1a)

```r
ggplot(Expedia, aes(x = PricePerNight, y = `Booked?`, color = Region)) +
  geom_point() +
  geom_smooth(method = "lm") +
  labs(
    title = "Booking Likelihood vs Price Per Night by Region",
    x = "Price Per Night",
    y = "Booking Likelihood (0 or 1)"
  )
```

```
## `geom_smooth()` using formula = 'y ~ x'
```



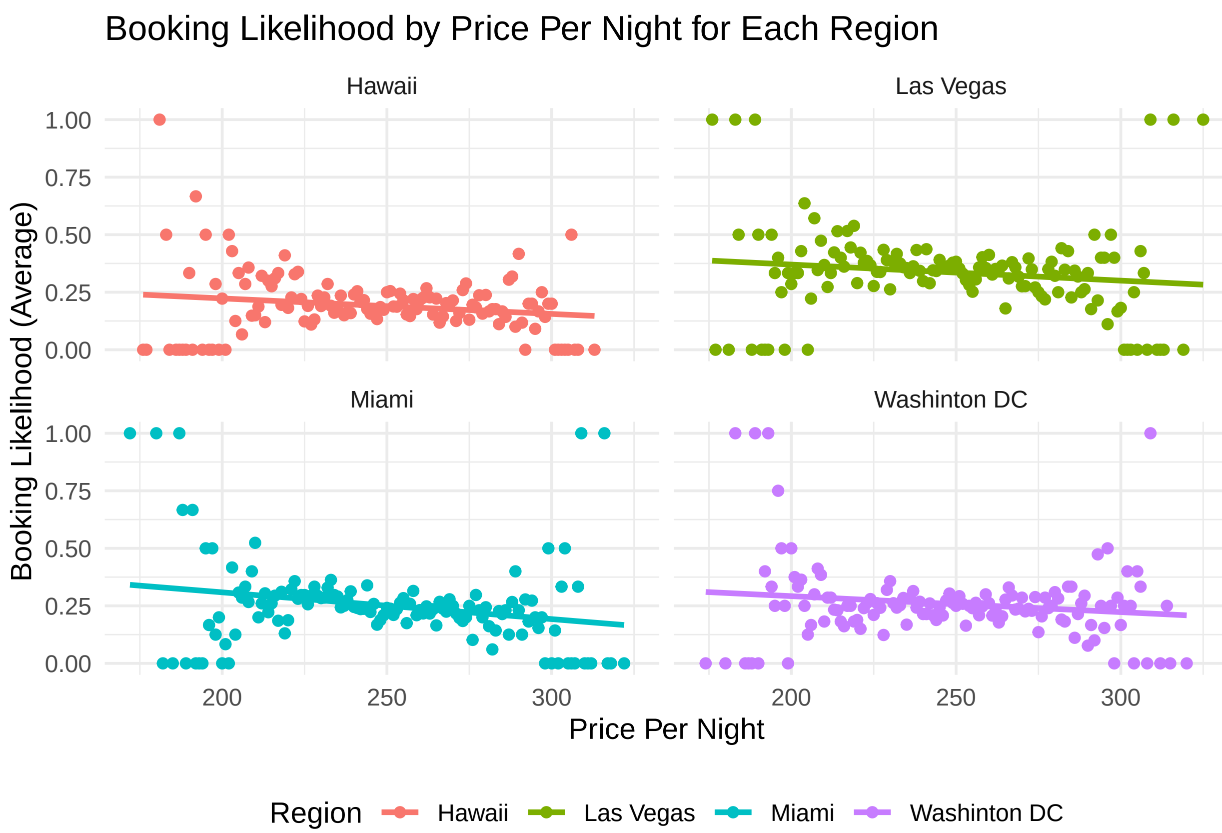
Graph (1b)

```r
data_grouped <- Expedia %>%
  group_by(PricePerNight, Region) %>%
  summarise(AverageBooked = mean(`Booked?`, na.rm = TRUE))
```

```
## `summarise()` has grouped output by 'PricePerNight'. You can override using the
## `.groups` argument.
```

```r
ggplot(data_grouped, aes(x = PricePerNight, y = AverageBooked, color = Region)) +
  geom_point() +
  geom_smooth(method = "lm", se = FALSE) +  # Add a linear regression line without the confidence inter
  labs(title = "Booking Likelihood by Price Per Night for Each Region",
       x = "Price Per Night",
       y = "Booking Likelihood (Average)") +
  theme_minimal() +
  theme(legend.position = "bottom")+
  facet_wrap(~Region)  # Create separate panels for each Region
```

```
## `geom_smooth()` using formula = 'y ~ x'
```




```
  slopes <- Expedia %>%
  group_by(Region) %>%
  do(model = lm(`Booked?` ~ PricePerNight, data = .)) %>%
  summarise(Region = Region,
            Slope = coef(model)[2])  # Extract the slope (coefficient for PricePerNight)

# Print the slopes
print(slopes)
```

```
## # A tibble: 4 x 2
##   Region           Slope
##   <chr>            <dbl>
## 1 Hawaii       -0.000518
## 2 Las Vegas    -0.00109
## 3 Miami        -0.00125
## 4 Washinton DC -0.000185
```

Graph (1c)

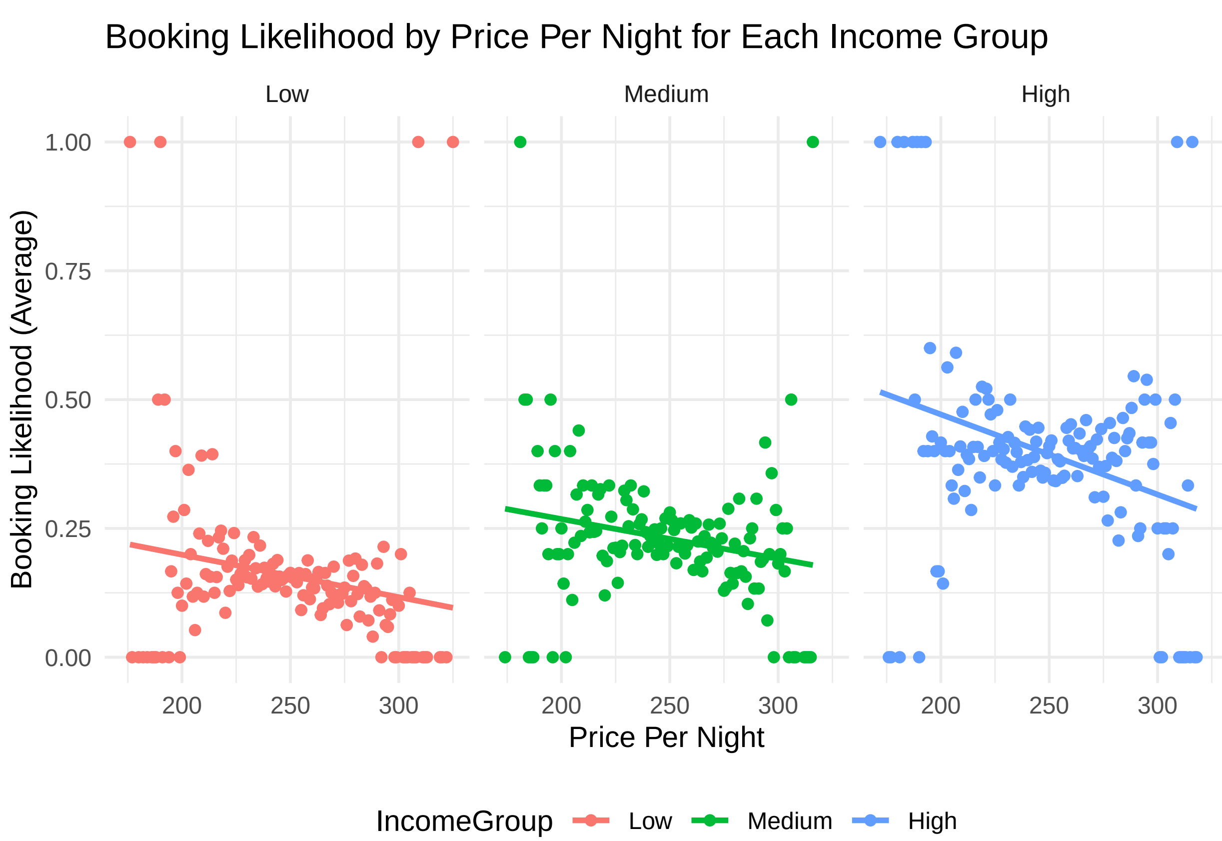
```
# Step 1: Categorize UserIncome into low, medium, and high based on quantiles
Expedia <- Expedia %>%
  mutate(IncomeGroup = cut(UserIncome,
                           breaks = quantile(UserIncome, probs = c(0, 0.33, 0.66, 1), na.rm = TRUE),
```

```r
                              labels = c("Low", "Medium", "High"),
                              include.lowest = TRUE))

# Step 2: Group data by PricePerNight and IncomeGroup, and calculate average booking
data_grouped <- Expedia %>%
  group_by(PricePerNight, IncomeGroup) %>%
  summarise(AverageBooked = mean(`Booked?`, na.rm = TRUE))
```

```
## `summarise()` has grouped output by 'PricePerNight'. You can override using the
## `.groups` argument.
```

```r
# Step 3: Plot the graph between PricePerNight and AverageBooked, colored by IncomeGroup
ggplot(data_grouped, aes(x = PricePerNight, y = AverageBooked, color = IncomeGroup)) +
  geom_point() +
  geom_smooth(method = "lm", se = FALSE) +  # Add linear regression line without confidence interval
  labs(title = "Booking Likelihood by Price Per Night for Each Income Group",
       x = "Price Per Night",
       y = "Booking Likelihood (Average)") +
  theme_minimal() +
  theme(legend.position = "bottom")+
  facet_wrap(~IncomeGroup)
```

```
## `geom_smooth()` using formula = 'y ~ x'
```



```r
# Step 4: Calculate slopes for each IncomeGroup
slopes <- Expedia %>%
  group_by(IncomeGroup) %>%
  do(model = lm(`Booked?` ~ PricePerNight, data = .)) %>%
  summarise(IncomeGroup = IncomeGroup,
            Slope = coef(model)[2])  # Extract the slope (coefficient for PricePerNight)
# Print the slopes
print(slopes)
```

```
## # A tibble: 3 x 2
##   IncomeGroup     Slope
##   <fct>           <dbl>
## 1 Low         -0.00108
## 2 Medium      -0.000798
## 3 High        -0.000475
```

lm Model 1A

```r
summary(lm(`Booked?` ~ PricePerNight, data = Expedia))
```

```
##
## Call:
## lm(formula = `Booked?` ~ PricePerNight, data = Expedia)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.3159 -0.2657 -0.2522  0.7126  0.7973
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.4463765  0.0342230  13.043  < 2e-16 ***
## PricePerNight -0.0007498  0.0001371  -5.471 4.52e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4383 on 24998 degrees of freedom
## Multiple R-squared:  0.001196,  Adjusted R-squared:  0.001156
## F-statistic: 29.93 on 1 and 24998 DF,  p-value: 4.52e-08
```

lm Model 1B

```r
summary(lm(`Booked?` ~ PricePerNight + factor(Region), data = Expedia))
```

```
##
## Call:
## lm(formula = `Booked?` ~ PricePerNight + factor(Region), data = Expedia)
##
```

```
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.4007 -0.2579 -0.2286  0.6342  0.8440
##
## Coefficients:
##                             Estimate Std. Error t value Pr(>|t|)
## (Intercept)                 0.388584   0.034261  11.342  < 2e-16 ***
## PricePerNight              -0.000760   0.000136  -5.586 2.35e-08 ***
## factor(Region)Las Vegas     0.146654   0.007782  18.846  < 2e-16 ***
## factor(Region)Miami         0.047485   0.007782   6.102 1.06e-09 ***
## factor(Region)Washinton DC  0.047146   0.007782   6.058 1.40e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.435 on 24995 degrees of freedom
## Multiple R-squared:  0.01606,    Adjusted R-squared:  0.0159
## F-statistic:   102 on 4 and 24995 DF,  p-value: < 2.2e-16
```

lm Model 1C

```r
summary(lm(`Booked?` ~ PricePerNight + factor(Region) + UserIncome, data = Expedia))
```

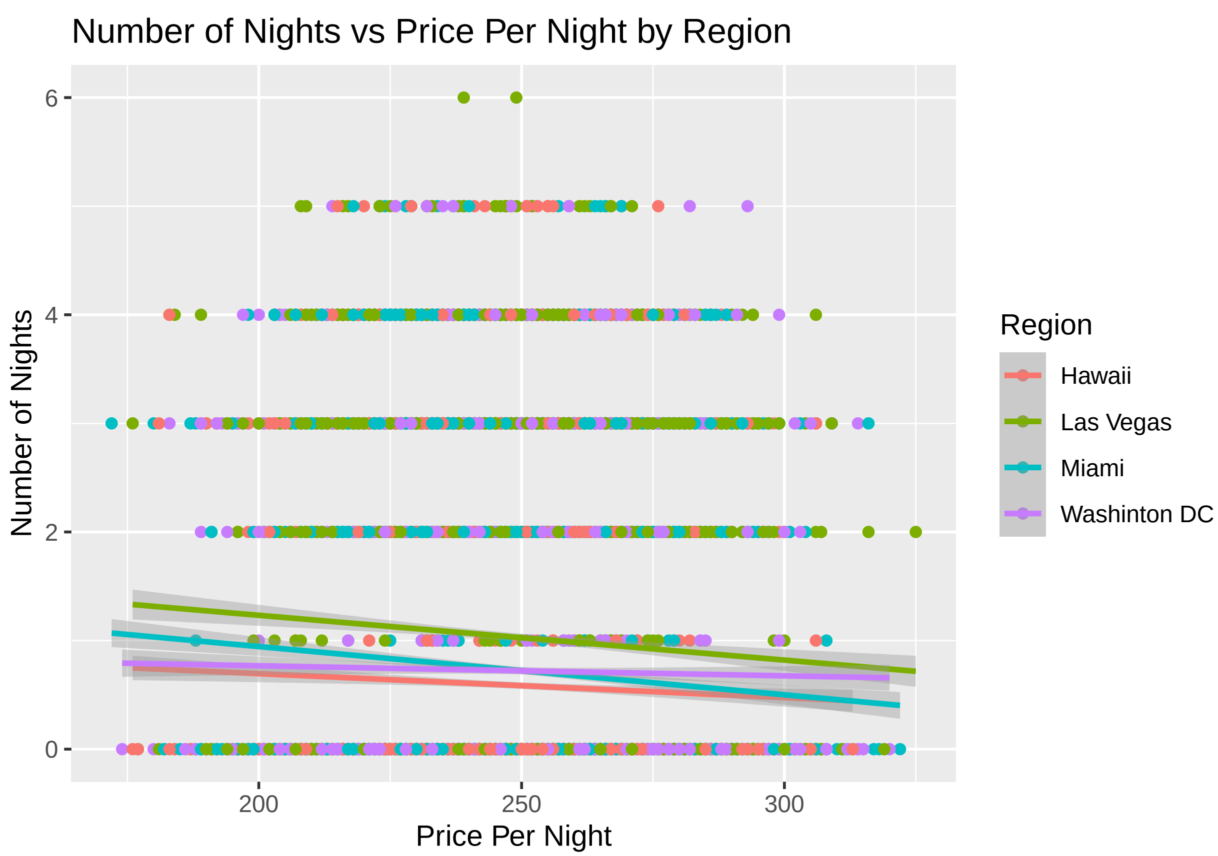
```
##
## Call:
## lm(formula = `Booked?` ~ PricePerNight + factor(Region) + UserIncome,
##     data = Expedia)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.3738 -0.2647 -0.1709  0.3166  0.9561
##
## Coefficients:
##                              Estimate Std. Error t value Pr(>|t|)
## (Intercept)                 1.915e-01  3.307e-02   5.791 7.10e-09 ***
## PricePerNight              -8.138e-04  1.303e-04  -6.246 4.28e-10 ***
## factor(Region)Las Vegas     1.488e-01  7.453e-03  19.965  < 2e-16 ***
## factor(Region)Miami         4.635e-02  7.453e-03   6.219 5.10e-10 ***
## factor(Region)Washinton DC  4.613e-02  7.453e-03   6.189 6.15e-10 ***
## UserIncome                  4.045e-06  8.512e-08  47.517  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4166 on 24994 degrees of freedom
## Multiple R-squared:  0.09758,    Adjusted R-squared:  0.0974
## F-statistic: 540.5 on 5 and 24994 DF,  p-value: < 2.2e-16
```

Nights as Dependent Variable

Graph (2a)

```
ggplot(Expedia, aes(x = PricePerNight, y = Nights, color = Region)) +
  geom_point() +
  geom_smooth(method = "lm") +
  labs(
    title = "Number of Nights vs Price Per Night by Region",
    x = "Price Per Night",
    y = "Number of Nights"
  )
```

```
## `geom_smooth()` using formula = 'y ~ x'
```



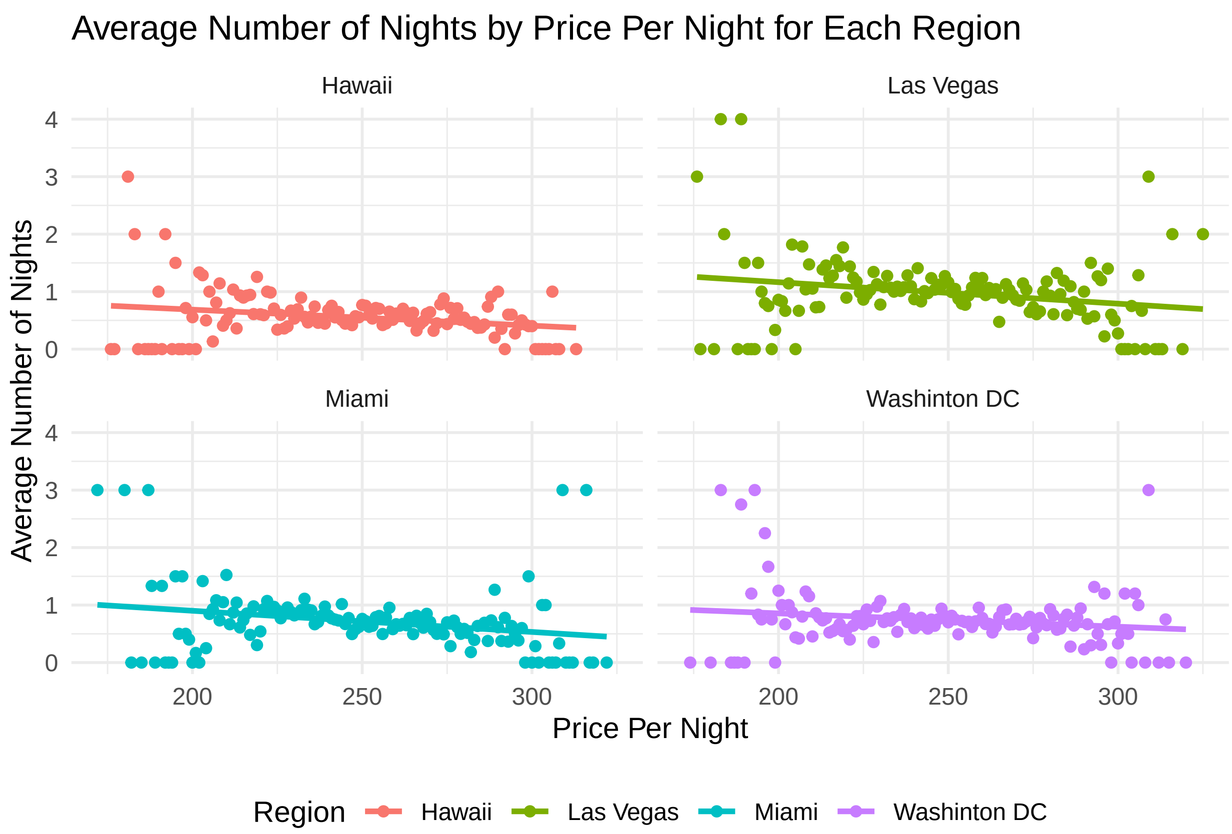
Graph (2b)

```
data_grouped <- Expedia %>%
  group_by(PricePerNight, Region) %>%
  summarise(AverageNights = mean(Nights, na.rm = TRUE))
```

```
## `summarise()` has grouped output by 'PricePerNight'. You can override using the
## `.groups` argument.
```

```
ggplot(data_grouped, aes(x = PricePerNight, y = AverageNights, color = Region)) +
  geom_point() +
  geom_smooth(method = "lm", se = FALSE) +  # Add a linear regression line without the confidence inter
  labs(title = "Average Number of Nights by Price Per Night for Each Region",
       x = "Price Per Night",
       y = "Average Number of Nights") +
  theme_minimal() +
  theme(legend.position = "bottom") +
  facet_wrap(~Region)  # Create separate panels for each Region
```

```
## `geom_smooth()` using formula = 'y ~ x'
```




```
# Calculate slopes for each Region
slopes <- Expedia %>%
  group_by(Region) %>%
  do(model = lm(Nights ~ PricePerNight, data = .)) %>%
  summarise(Region = Region,
            Slope = coef(model)[2])  # Extract the slope (coefficient for PricePerNight)

# Print the slopes
print(slopes)
```

```
## # A tibble: 4 x 2
##   Region           Slope
##   <chr>            <dbl>
## 1 Hawaii       -0.00221
## 2 Las Vegas    -0.00413
## 3 Miami        -0.00444
## 4 Washinton DC -0.000926
```

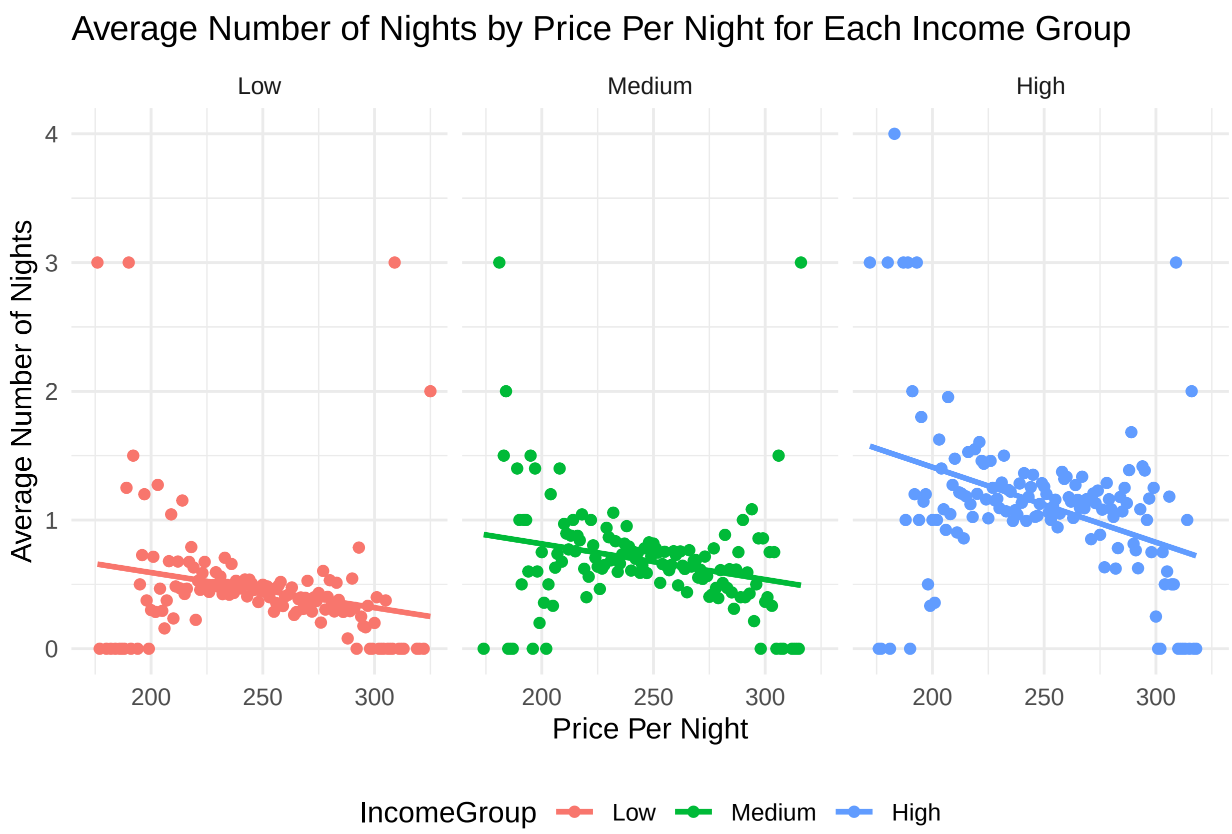
Graph (2c)

```r
# Step 1: Categorize UserIncome into low, medium, and high based on quantiles
Expedia <- Expedia %>%
  mutate(IncomeGroup = cut(UserIncome,
                           breaks = quantile(UserIncome, probs = c(0, 0.33, 0.66, 1), na.rm = TRUE),
                           labels = c("Low", "Medium", "High"),
                           include.lowest = TRUE))

# Step 2: Group data by PricePerNight and IncomeGroup, and calculate average nights
data_grouped <- Expedia %>%
  group_by(PricePerNight, IncomeGroup) %>%
  summarise(AverageNights = mean(Nights, na.rm = TRUE))
```

```
## `summarise()` has grouped output by 'PricePerNight'. You can override using the
## `.groups` argument.
```

```r
# Step 3: Plot the graph between PricePerNight and AverageNights, colored by IncomeGroup
ggplot(data_grouped, aes(x = PricePerNight, y = AverageNights, color = IncomeGroup)) +
  geom_point() +
  geom_smooth(method = "lm", se = FALSE) +  # Add linear regression line without confidence interval
  labs(title = "Average Number of Nights by Price Per Night for Each Income Group",
       x = "Price Per Night",
       y = "Average Number of Nights") +
  theme_minimal() +
  theme(legend.position = "bottom") +
  facet_wrap(~IncomeGroup)
```

```
## `geom_smooth()` using formula = 'y ~ x'
```




```
# Step 4: Calculate slopes for each IncomeGroup
slopes <- Expedia %>%
  group_by(IncomeGroup) %>%
  do(model = lm(Nights ~ PricePerNight, data = .)) %>%
  summarise(IncomeGroup = IncomeGroup,
            Slope = coef(model)[2])  # Extract the slope (coefficient for PricePerNight)

# Print the slopes
print(slopes)
```

```
## # A tibble: 3 x 2
##   IncomeGroup    Slope
##   <fct>          <dbl>
## 1 Low         -0.00338
## 2 Medium      -0.00303
## 3 High        -0.00257
```

lm Model 2A

```
summary(lm(Nights ~ PricePerNight, data = Expedia))
```

```
##
```

```
## Call:
## lm(formula = Nights ~ PricePerNight, data = Expedia)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.9837 -0.7900 -0.7380  1.1522  5.2331
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.4867377  0.1054292  14.102  < 2e-16 ***
## PricePerNight -0.0028909  0.0004222  -6.847 7.72e-12 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.35 on 24998 degrees of freedom
## Multiple R-squared:  0.001872,	Adjusted R-squared:  0.001832
## F-statistic: 46.88 on 1 and 24998 DF,  p-value: 7.718e-12
```

lm Model 2B

```r
summary(lm(Nights ~ PricePerNight + factor(Region), data = Expedia))
```

```
##
## Call:
## lm(formula = Nights ~ PricePerNight + factor(Region), data = Expedia)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.2410 -0.7682 -0.6644  0.9313  4.9693
##
## Coefficients:
##                              Estimate Std. Error t value Pr(>|t|)
## (Intercept)                 1.3139517  0.1055680  12.446  < 2e-16 ***
## PricePerNight              -0.0029206  0.0004192  -6.967 3.32e-12 ***
## factor(Region)Las Vegas     0.4440148  0.0239780  18.518  < 2e-16 ***
## factor(Region)Miami         0.1406153  0.0239782   5.864 4.57e-09 ***
## factor(Region)Washinton DC  0.1360355  0.0239788   5.673 1.42e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.34 on 24995 degrees of freedom
## Multiple R-squared:  0.01633,	Adjusted R-squared:  0.01617
## F-statistic: 103.7 on 4 and 24995 DF,  p-value: < 2.2e-16
```

lm Model 2C

```r
summary(lm(Nights ~ PricePerNight + factor(Region) + UserIncome, data = Expedia))
```

```
##
## Call:
## lm(formula = Nights ~ PricePerNight + factor(Region) + UserIncome,
##     data = Expedia)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -4.0301 -0.7841 -0.5044  0.3240  4.9445
##
## Coefficients:
##                              Estimate Std. Error t value Pr(>|t|)
## (Intercept)                 7.376e-01  1.024e-01   7.206 5.93e-13 ***
## PricePerNight              -3.078e-03  4.033e-04  -7.633 2.38e-14 ***
## factor(Region)Las Vegas     4.503e-01  2.307e-02  19.520  < 2e-16 ***
## factor(Region)Miami         1.373e-01  2.307e-02   5.952 2.69e-09 ***
## factor(Region)Washinton DC  1.331e-01  2.307e-02   5.768 8.11e-09 ***
## UserIncome                  1.183e-05  2.634e-07  44.900  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.289 on 24994 degrees of freedom
## Multiple R-squared:  0.08975,    Adjusted R-squared:  0.08957
## F-statistic: 492.9 on 5 and 24994 DF,  p-value: < 2.2e-16
```