Movie Theater Insights Data Analysis

2024-12-13

Project Resources

```
# Packages used

library(dplyr)
```

```
## Warning: package 'dplyr' was built under R version 4.4.2
```

```
##
## 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
```

```
library(ggplot2)
```

```
## Warning: package 'ggplot2' was built under R version 4.4.2
```

```
library(patchwork)
```

```
## Warning: package 'patchwork' was built under R version 4.4.2
```

```
library(hms)
```

```
## Warning: package 'hms' was built under R version 4.4.2
```

```
library(effsize)
```

```
## Warning: package 'effsize' was built under R version 4.4.2
```

```
library(car)
```

```
## Warning: package 'car' was built under R version 4.4.2
```

```
## Loading required package: carData
```

```
## Warning: package 'carData' was built under R version 4.4.2
```

```
##
## Attaching package: 'car'
```

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

```
# Data Repository

library(readxl)
```

```
## Warning: package 'readxl' was built under R version 4.4.2
```

```
Galaxy_Theatres_Movie_Ticket_Characteristics_For_Data_Analysis_ <- read_excel(
  "Galaxy Theatres Movie Ticket Characteristics (For Data Analysis).xlsx")



# The title for the dataset was changed for simplicity.
# R has trouble reading 'Screening Showtime' as a numeric value, so let's fix that.

MovieTheaterDataSet <- Galaxy_Theatres_Movie_Ticket_Characteristics_For_Data_Analysis_

MovieTheaterDataSet$`Screening_Showtime`<- as_hms(MovieTheaterDataSet$`Screening_Showtime`)
View(MovieTheaterDataSet)
```

Brief Statistics

```
# Brief Summary of DataSet

Base_TP_Stats <- MovieTheaterDataSet %>%
  summarize(
    Min = min(Base_Ticket_Price, na.rm = TRUE),
    Q1 = quantile(Base_Ticket_Price, 0.25, na.rm = TRUE),
    Median = median(Base_Ticket_Price, na.rm = TRUE),
    Mean = mean(Base_Ticket_Price, na.rm = TRUE),
    Q3 = quantile(Base_Ticket_Price, 0.75, na.rm = TRUE),
    Max = max(Base_Ticket_Price, na.rm = TRUE)
  )


Final_TP_Stats <- MovieTheaterDataSet %>%
  summarize(
    Min = min(Final_Ticket_Price, na.rm = TRUE),
    Q1 = quantile(Final_Ticket_Price, 0.25, na.rm = TRUE),
    Median = median(Final_Ticket_Price, na.rm = TRUE),
    Mean = mean(Final_Ticket_Price, na.rm = TRUE),
    Q3 = quantile(Final_Ticket_Price, 0.75, na.rm = TRUE),
    Max = max(Final_Ticket_Price, na.rm = TRUE)
  )


# Counts

DiscountedCount <- MovieTheaterDataSet %>%
  group_by(Discounted_Ticket) %>%
  summarise(Count = n())

DiscountedCount
```

```
## # A tibble: 2 × 2
##   Discounted_Ticket Count
##   <chr>             <int>
## 1 No                 3085
## 2 Yes                  18
```

```
SpecialEventCount <- MovieTheaterDataSet %>%
  group_by(Special_Event_Pricing) %>%
  summarise(Count = n())

SpecialEventCount
```

```
## # A tibble: 2 × 2
##   Special_Event_Pricing Count
##   <chr>                 <int>
## 1 No                     3055
## 2 Yes                      48
```

```
SpecialProgramCount <- MovieTheaterDataSet %>%
  group_by(Special_Program) %>%
  summarise(Count = n())

SpecialProgramCount
```

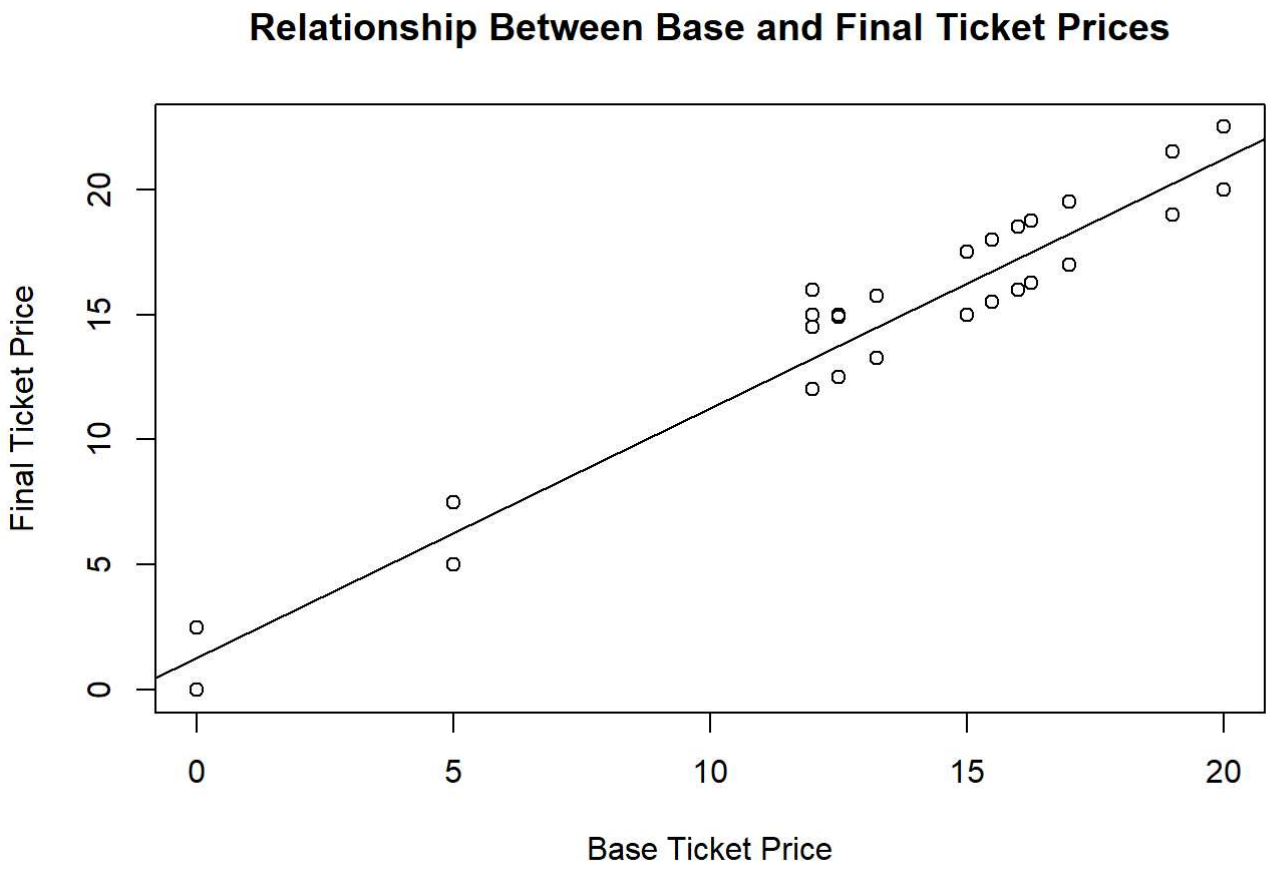
```
## # A tibble: 6 × 2
##   Special_Program  Count
##   <chr>            <int>
## 1 AXCN                12
## 2 Fathom Events        4
## 3 Flashback Cinema    28
## 4 None              3043
## 5 Studio Ghibli        6
## 6 Unique              10
```

Data Visualizations

```
# Single Linear Regression Model

lm_model <- lm(Final_Ticket_Price ~ Base_Ticket_Price, data = MovieTheaterDataSet)
summary(lm_model)
```

```
##
## Call:
## lm(formula = Final_Ticket_Price ~ Base_Ticket_Price, data = MovieTheaterDataSet)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -1.290 -1.255  1.145  1.235  2.744
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        1.28960    0.14522    8.88   <2e-16 ***
## Base_Ticket_Price  0.99722    0.01064   93.76   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.251 on 3101 degrees of freedom
## Multiple R-squared:  0.7392, Adjusted R-squared:  0.7391
## F-statistic:  8790 on 1 and 3101 DF,  p-value: < 2.2e-16
```

```
# Scatterplot of Relationship between Base and Final Ticket Prices

plot(MovieTheaterDataSet$Base_Ticket_Price, MovieTheaterDataSet$Final_Ticket_Price,
xlab = "Base Ticket Price", ylab = "Final Ticket Price", main =
"Relationship Between Base and Final Ticket Prices")
abline(lm_model)
```



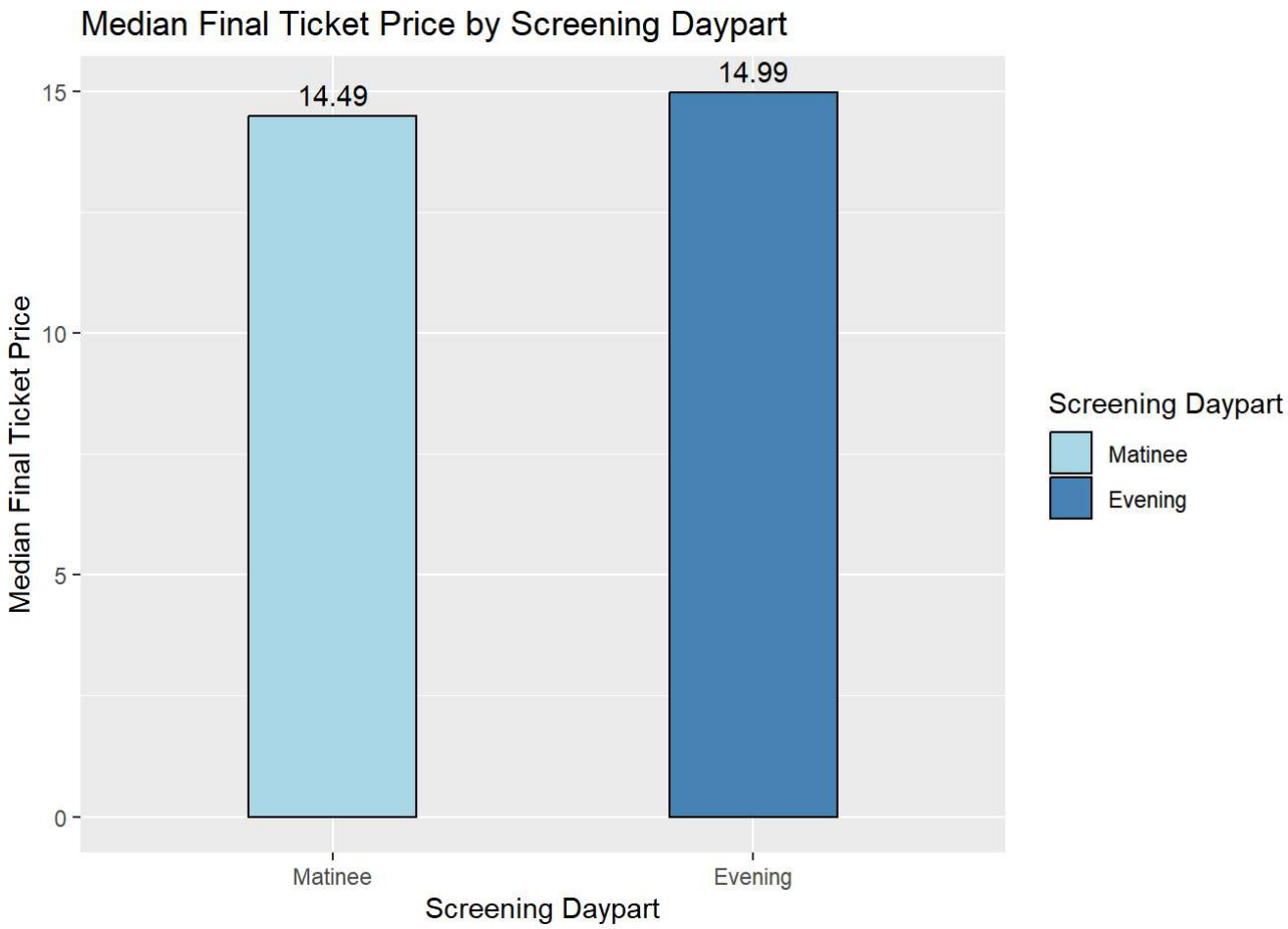

```
# Correlation Coefficient (p)

p <- cor(MovieTheaterDataSet$Base_Ticket_Price, MovieTheaterDataSet$Final_Ticket_Price)
head(p)
```

```
## [1] 0.8597777
```

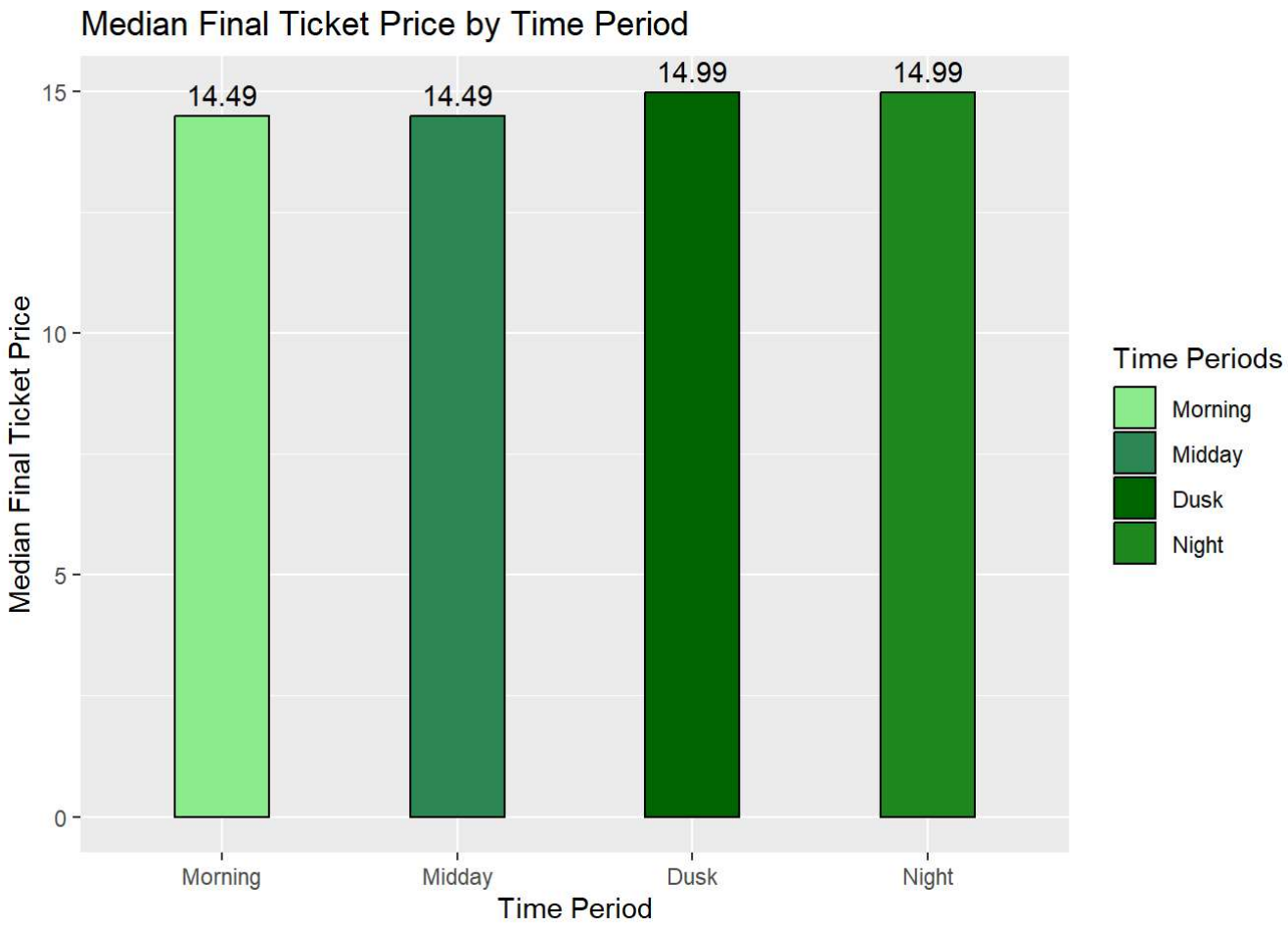
```
# There is a strong positive correlation between Base and Final Ticket Prices.
```

```
#### Bar Charts -- One Categorical Variable Against One Continuous Variable


### Group 1: Time-Related Variables


## Relationship Between Screening Dayparts and the Median Final Ticket Price

# Before visualizing data based on Screening Dayparts,
# we need to manually alter the variable.
MovieTheaterDataSet$Screening_Daypart <- factor(
  MovieTheaterDataSet$Screening_Daypart,
  levels = c("Matinee", "Evening"))


Plot1 <- ggplot(MovieTheaterDataSet,
                aes(x = Screening_Daypart,
                    y = Final_Ticket_Price,
                    fill = Screening_Daypart)) +
  stat_summary(fun = "median", geom = "bar",
               color = "black", width = 0.4,
               position = position_dodge(width = 0.5)) +
  stat_summary(fun = "median", geom = "text",
               aes(label = round(..y.., 2)), vjust = -0.5) +
  scale_fill_manual(values = c("lightblue", "steelblue",
                               "dodgerblue", "darkblue")) +
  labs(title = "Median Final Ticket Price by Screening Daypart",
       x = "Screening Daypart", y = "Median Final Ticket Price",
       fill = "Screening Daypart")

Plot1
```

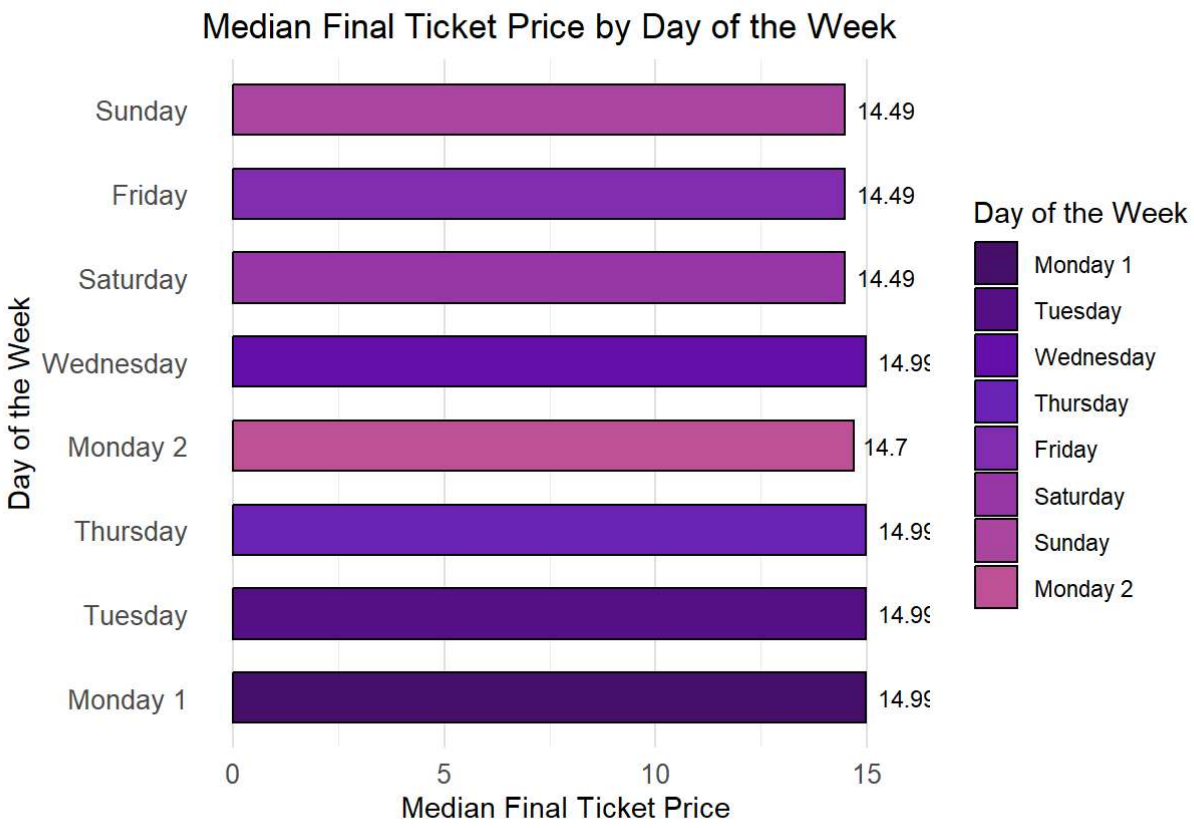
```
## Warning: The dot-dot notation (`..y..`) was deprecated in ggplot2 3.4.0.
## ℹ Please use `after_stat(y)` instead.
## This warning is displayed once every 8 hours.
## Call `lifecycle::last_lifecycle_warnings()` to see where this warning was
## generated.
```



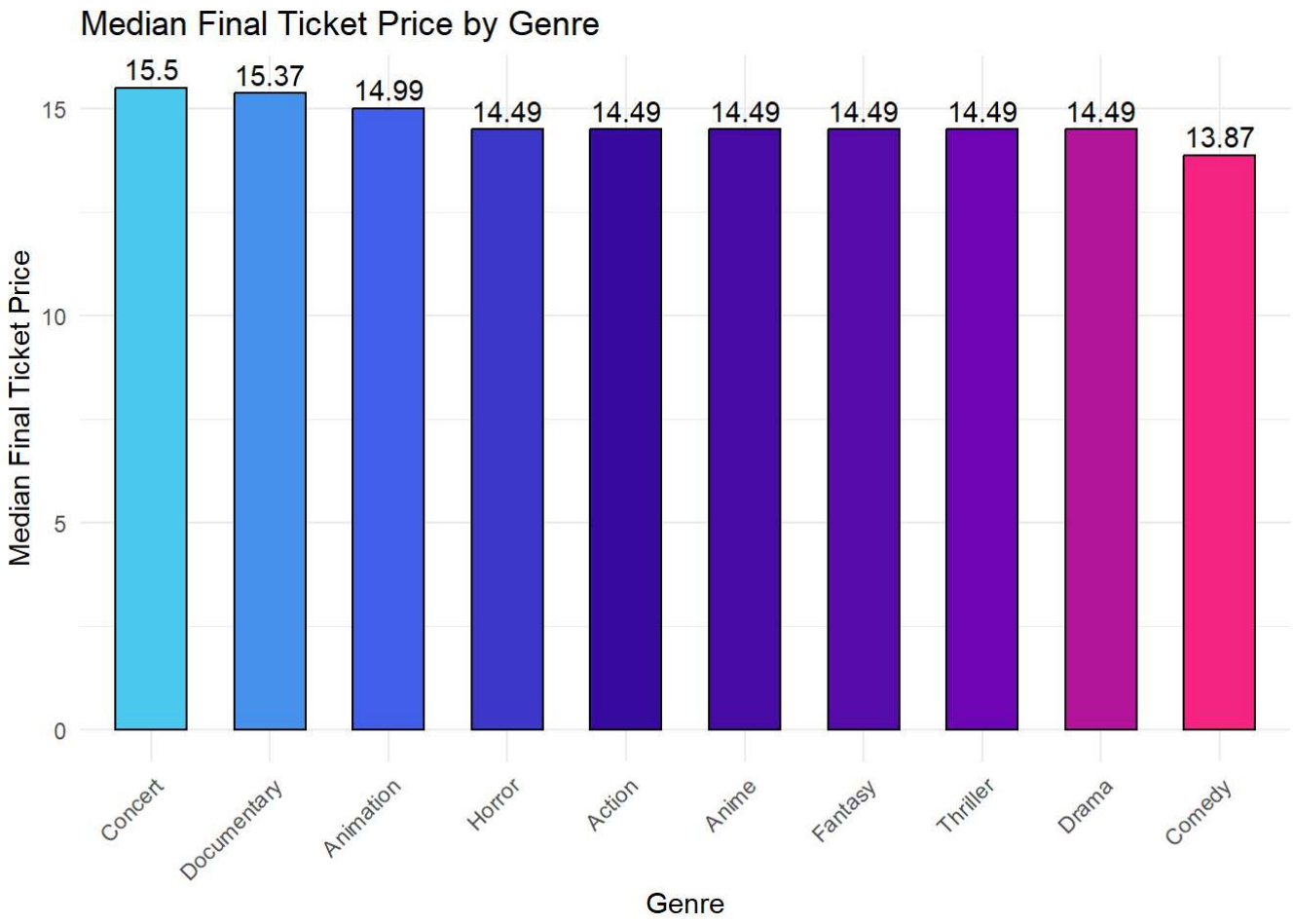

```
## Relationship Between Time Periods and the Median Final Ticket Price

# Before visualizing data based on Time Periods,
# we need to manually alter the variable.
MovieTheaterDataSet$Time_Periods <- factor(
  MovieTheaterDataSet$Time_Periods,
  levels = c("Morning", "Midday", "Dusk", "Night"))


Plot2 <- ggplot(MovieTheaterDataSet,
                aes(x = Time_Periods,
                    y = Final_Ticket_Price,
                    fill = Time_Periods)) +
  stat_summary(fun = "median", geom = "bar",
               color = "black", width = 0.4,
               position = position_dodge(width = 0.5)) +
  stat_summary(fun = "median", geom = "text",
               aes(label = round(..y.., 2)), vjust = -0.5) +
  scale_fill_manual(values = c("lightgreen", "seagreen",
                               "darkgreen", "forestgreen")) +
  labs(title = "Median Final Ticket Price by Time Period",
       x = "Time Period", y = "Median Final Ticket Price",
       fill = "Time Periods")

Plot2
```


Median Final Ticket Price by Time Period
14.49
14.49
14.99
14.99
15
10
5
0
Median Final Ticket Price
Morning
Midday
Dusk
Night
Time Period
Time Periods
Morning
Midday
Dusk
Night

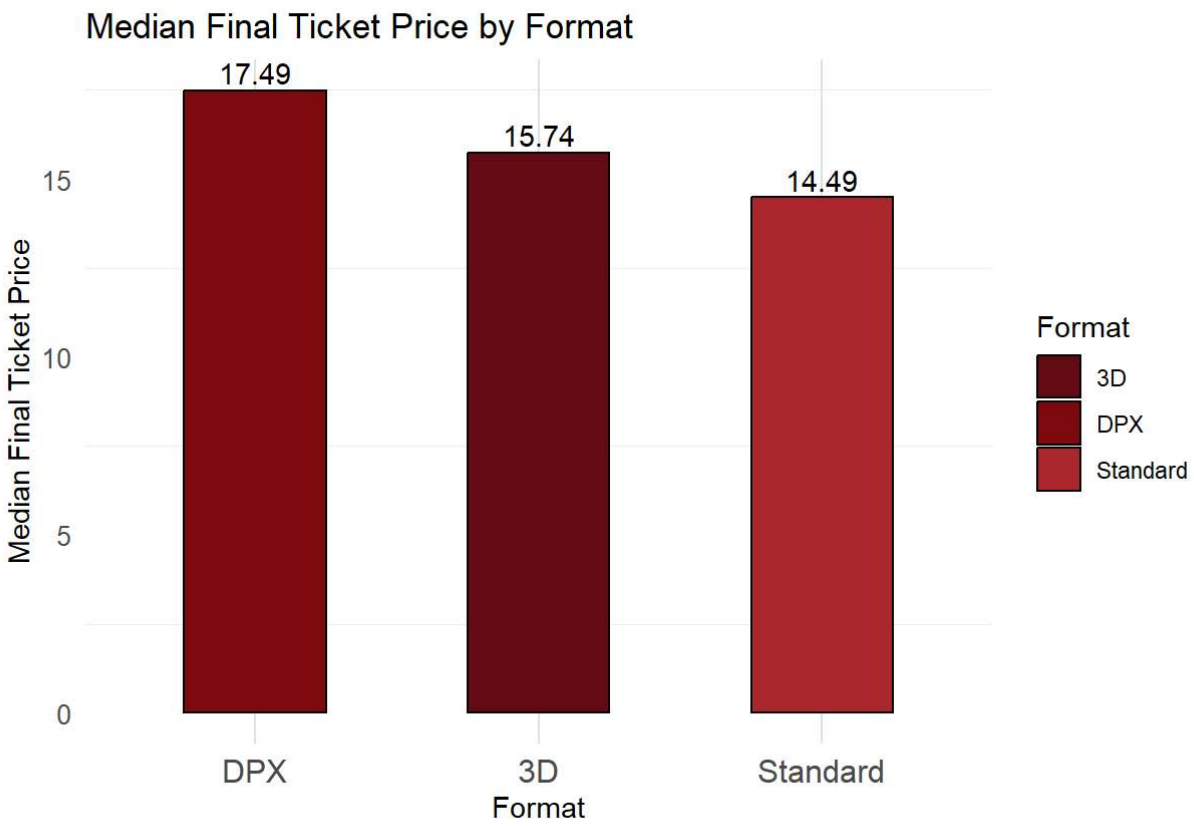
```
## Relationship Between Days of the Week and the Median Final Ticket Price

# Before visualizing data based on Days of the Week,
# we need to manually alter the variable.
MovieTheaterDataSet$Day_of_the_Week <- factor(
  MovieTheaterDataSet$Day_of_the_Week,
  levels = c("Monday 1", "Tuesday", "Wednesday",
             "Thursday", "Friday", "Saturday",
             "Sunday", "Monday 2"))


Plot3 <- ggplot(MovieTheaterDataSet,
                aes(x = reorder(Day_of_the_Week, -Final_Ticket_Price),
                    y = Final_Ticket_Price, fill = Day_of_the_Week)) +
  stat_summary(fun = "median", geom = "bar",
               color = "black", width = 0.6) +
  stat_summary(fun = "median", geom = "text",
               aes(label = round(..y.., 2)), hjust = -0.2,size = 3) +
  scale_fill_manual(values = c("Monday 1" = "#47126b", "Tuesday" = "#571089",
    "Wednesday" = "#6411AD", "Thursday" = "#6D23B6", "Friday" = "#822FAF",
    "Saturday" = "#973AA8", "Sunday" = "#AC46A1", "Monday 2" = "#C05299")) +
  coord_flip() +
  scale_y_continuous(expand = expansion(mult = c(0.05, 0.1))) +
  labs(title = "Median Final Ticket Price by Day of the Week",
       x = "Day of the Week", y = "Median Final Ticket Price",
       fill = "Day of the Week") +
  theme_minimal() +
  theme(axis.text.y = element_text(size = 10),
        axis.text.x = element_text(size = 10),
        panel.grid.major.y = element_blank(),
        panel.grid.major.x = element_line(color = "gray90"),
        plot.margin = unit(c(1, 1, 1, 1), "cm"))


Plot3
```


Median Final Ticket Price by Day of the Week
Sunday
Friday
Saturday
Wednesday
Monday 2
Thursday
Tuesday
Monday 1
14.49
14.49
14.49
14.99
14.7
14.99
14.99
14.99
Day of the Week
0
5
10
15
Median Final Ticket Price
Day of the Week
Monday 1
Tuesday
Wednesday
Thursday
Friday
Saturday
Sunday
Monday 2

```
### Group 2: Movie Characteristics


## Relationship Between Genre and the Median Final Ticket Price

# Before visualizing data based on Genre,
# we must manually alter the variable.
MovieTheaterDataSet$Genre <- factor(
  MovieTheaterDataSet$Genre,
  levels = c("Concert", "Documentary", "Animation", "Horror", "Action",
             "Anime", "Fantasy","Thriller", "Drama", "Comedy"))


Plot4 <- ggplot(MovieTheaterDataSet,
                aes(x = reorder(Genre, -Final_Ticket_Price),
                    y = Final_Ticket_Price, fill = Genre)) +
stat_summary(fun = "median", geom = "bar",
             color = "black", width = 0.6) +
stat_summary(fun = "median", geom = "text",
             aes(label = round(..y.., 2)), vjust = -0.4) +
scale_fill_manual(values = c("#4CC9F0", "#4895EF", "#4361EE","#3F37C9","#3A0CA3",
                             "#480CA8", "#560BAD", "#7209B7", "#B5179E", "#F72585")) +
labs(title = "Median Final Ticket Price by Genre",
x = "Genre",
y = "Median Final Ticket Price",
fill = "Genre") +
theme_minimal() +
theme(axis.text.x = element_text(angle = 45, hjust = 1),
legend.position = "none")


Plot4
```



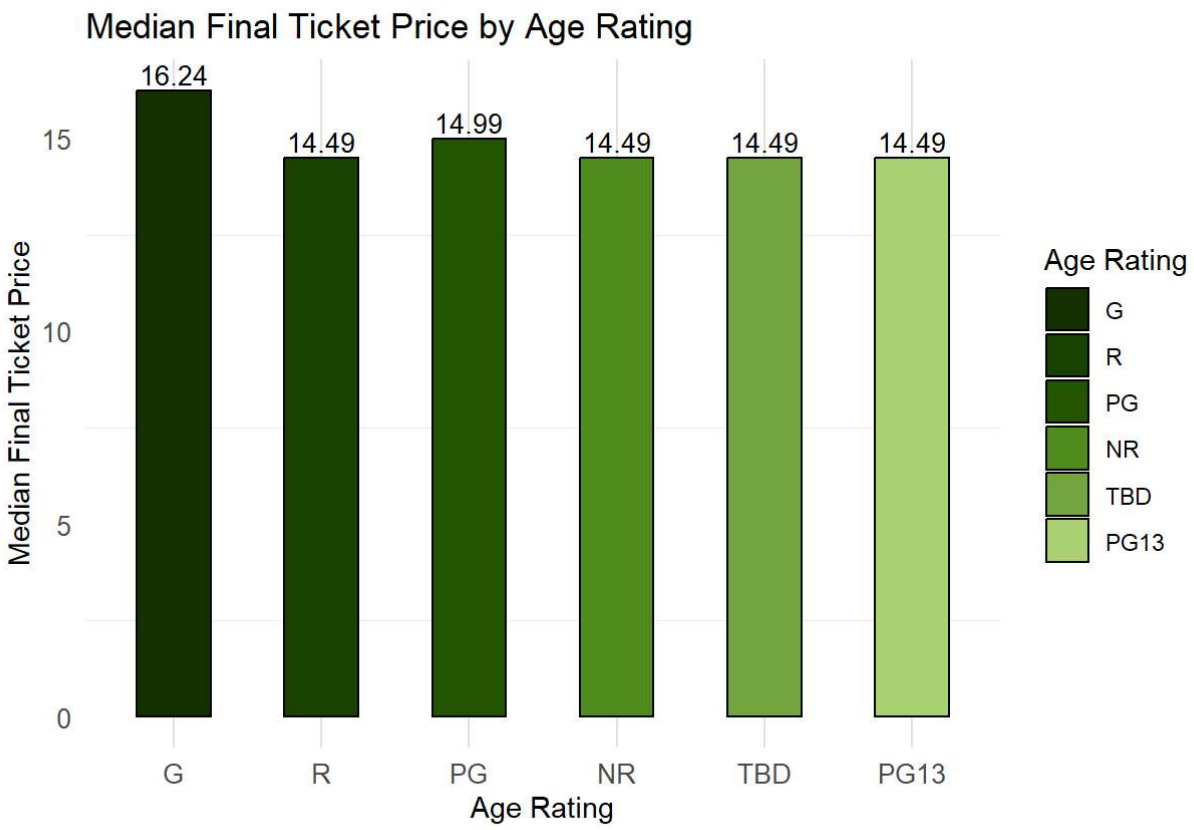

```
## Relationship Between Format and the Median Final Ticket Price


Plot5 <- ggplot(MovieTheaterDataSet,
                aes(x = reorder(Format, -Final_Ticket_Price),
                    y = Final_Ticket_Price, fill = Format)) +
  stat_summary(fun = "median", geom = "bar",
               color = "black", width = 0.5) +
  stat_summary(fun = "median", geom = "text",
               aes(label = round(..y.., 2)), vjust = -0.3) +
  scale_fill_manual(values = c("#640D14", "#800E13", "#AD2831")) +
  labs(title = "Median Final Ticket Price by Format",
       x = "Format",
       y = "Median Final Ticket Price",
       fill = "Format") +
  theme_minimal() +
  theme(legend.position = "right",
    axis.text.x = element_text(size = 12),
    axis.text.y = element_text(size = 10),
    panel.grid.major.y = element_blank(),
    panel.grid.major.x = element_line(color = "gray90"),
    plot.margin = unit(c(1, 1, 1, 1), "cm"))


Plot5
```


Median Final Ticket Price by Format
17.49
15.74
14.49
15
10
5
0
Median Final Ticket Price
DPX
3D
Standard
Format
Format
3D
DPX
Standard

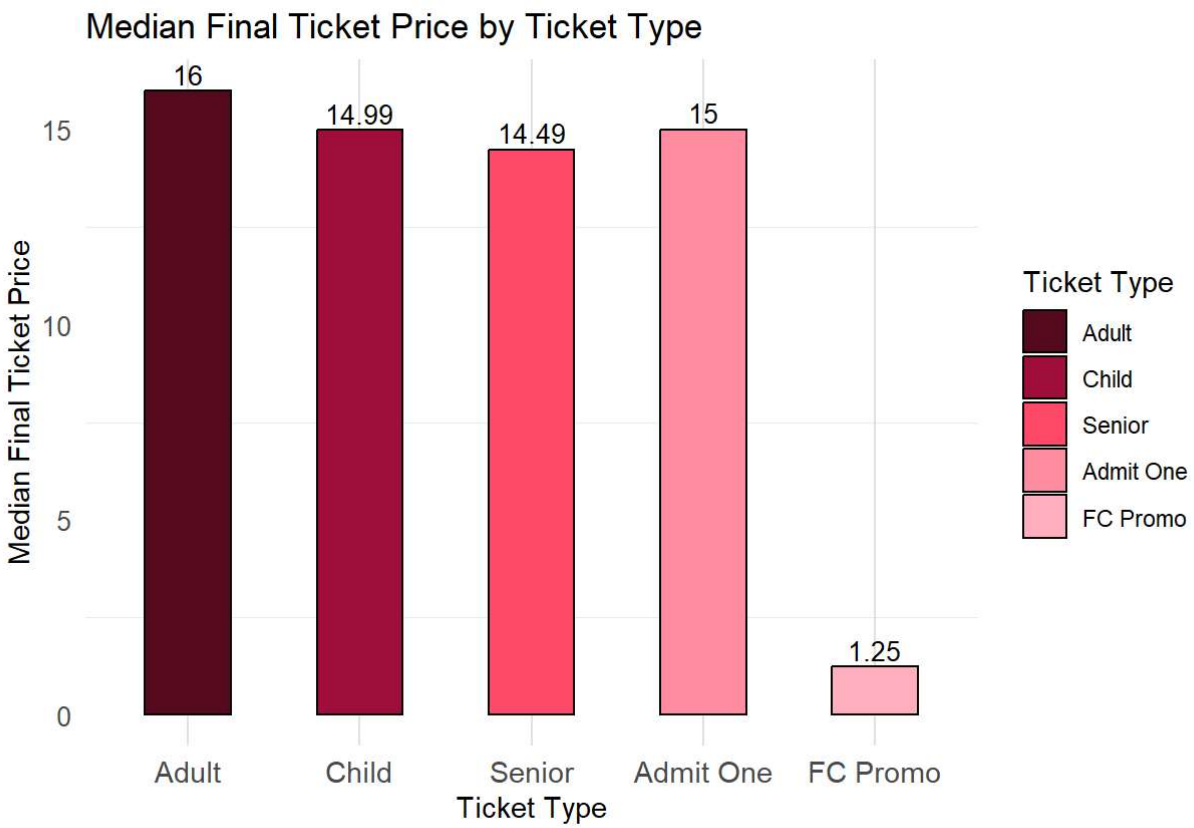
```
## Relationship Between Special Program and the Median Final Ticket Price

# Before visualizing data based on Special Program,
# we must manually alter the variable.

MovieTheaterDataSet$Special_Program <- factor(
  MovieTheaterDataSet$Special_Program,
  levels = c("Flashback Cinema", "None", "AXCN",
             "Studio Ghibli", "Unique", "Fathom Events"))


Plot6 <- ggplot(MovieTheaterDataSet,
                aes(x = reorder(Special_Program, -Final_Ticket_Price),
                    y = Final_Ticket_Price, fill = Special_Program)) +
  stat_summary(fun = "median", geom = "bar",
               color = "black", width = 0.7) +
  stat_summary(fun = "median", geom = "text",
               aes(label = round(..y.., 2)), vjust = -0.2, size = 3) +
  scale_fill_manual(values = c("#FF4800", "#FF6000", "#FF7900",
                               "#FF8500", "#FF9E00","#FFB600")) +
  labs(title = "Median Final Ticket Price by Special Program",
       x = "Median Final Ticket Price",
       y = "Special Program",
       fill = "Special Program") +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1),
        legend.position = "right",
        panel.grid.major.y = element_blank(),
        panel.grid.major.x = element_line(color = "gray90"),
        plot.margin = unit(c(1, 1, 1, 1), "cm"))


Plot6
```


Median Final Ticket Price by Special Program
18.24
17.49
16.24
14.49
14.74
5
15
10
5
0
Special Program
Fathom Events
Unique
Studio Ghibli
None
AXCN
Flashback Cinema
Median Final Ticket Price
Special Program
Flashback Cinema
None
AXCN
Studio Ghibli
Unique
Fathom Events

```
## Relationship Between Age Rating and the Median Final Ticket Price

# Before visualizing data based on Age Rating,
# we must manually alter the variable.

MovieTheaterDataSet$Age_Rating <- factor(
  MovieTheaterDataSet$Age_Rating,
  levels = c("G", "R", "PG", "NR", "TBD", "PG13"))


Plot7 <- ggplot(MovieTheaterDataSet,
                aes(x = reorder(Age_Rating, -Final_Ticket_Price),
                    y = Final_Ticket_Price, fill = Age_Rating)) +
  stat_summary(fun = "median", geom = "bar",
               color = "black", width = 0.5) +
  stat_summary(fun = "median", geom = "text",
               aes(label = round(..y.., 2)),
               vjust = -0.3, size = 3.5) +
  scale_fill_manual(values = c("#143601", "#1A4301", "#245501",
                               "#538D22", "#73A942", "#AAD576")) +
  labs(title = "Median Final Ticket Price by Age Rating",
       x = "Age Rating",
       y = "Median Final Ticket Price",
       fill = "Age Rating") +
  theme_minimal() +
  theme(legend.position = "right",
    axis.text.x = element_text(size = 10),
    axis.text.y = element_text(size = 10),
    panel.grid.major.y = element_blank(),
    panel.grid.major.x = element_line(color = "gray90"),
    plot.margin = unit(c(1, 1, 1, 1), "cm"))


Plot7
```


Median Final Ticket Price by Age Rating
16.24
14.49
14.99
14.49
14.49
14.49
15
10
5
0
Median Final Ticket Price
G
R
PG
NR
TBD
PG13
Age Rating
Age Rating
G
R
PG
NR
TBD
PG13

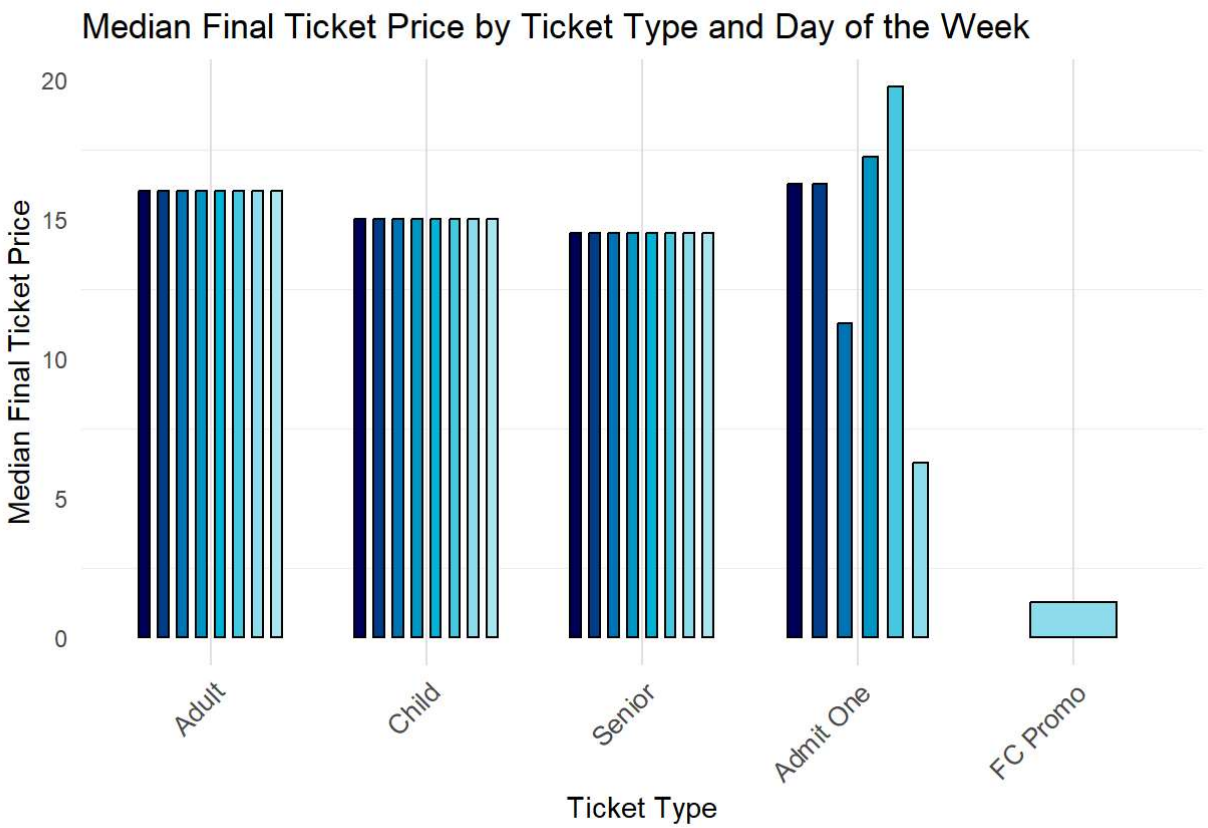
```
### Group 3: Customer Characteristics


## Relationship Between Ticket Type and the Median Final Ticket Price

# Before visualizing data based on Ticket Type,
# we must manually alter the variable.
MovieTheaterDataSet$Ticket_Type <- factor(
  MovieTheaterDataSet$Ticket_Type,
  levels = c("Adult", "Child", "Senior", "Admit One", "FC Promo"))

Plot8 <- ggplot(MovieTheaterDataSet,
                aes(x = reorder(Ticket_Type, -Final_Ticket_Price),
                    y = Final_Ticket_Price, fill = Ticket_Type)) +
  stat_summary(fun = "median", geom = "bar",
               color = "black", width = 0.5) +
  stat_summary(fun = "median", geom = "text",
               aes(label = round(..y.., 2)),
               vjust = -0.3, size = 3.5) +
  scale_fill_manual(values = c("#590D22", "#A4133C", "#FF4D6D",
                               "#FF8FA3", "#FFB3C1")) +
  labs(title = "Median Final Ticket Price by Ticket Type",
       x = "Ticket Type",
       y = "Median Final Ticket Price",
       fill = "Ticket Type") +
 theme_minimal() +
  theme(legend.position = "right",
    axis.text.x = element_text(size = 11),
    axis.text.y = element_text(size = 10),
    panel.grid.major.y = element_blank(),
    panel.grid.major.x = element_line(color = "gray90"),
    plot.margin = unit(c(1, 1, 1, 1), "cm"))

Plot8
```


Median Final Ticket Price by Ticket Type
16
14.99
14.49
15
1.25
15
10
5
0
Median Final Ticket Price
Adult
Child
Senior
Admit One
FC Promo
Ticket Type
Ticket Type
Adult
Child
Senior
Admit One
FC Promo

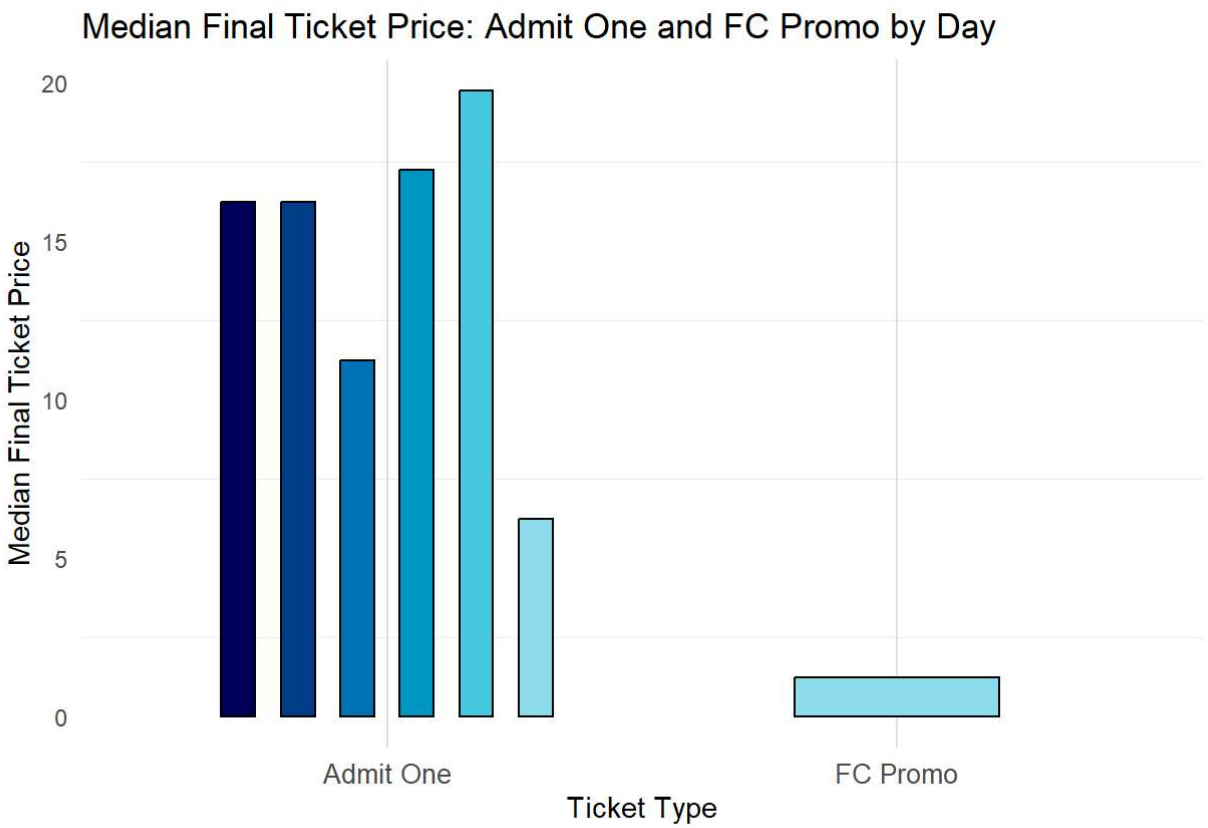
```
## Relationship Between Purchase Method and the Median Final Ticket Price

Plot9 <- ggplot(MovieTheaterDataSet,
                aes(x = reorder(Purchase_Method, -Final_Ticket_Price),
                    y = Final_Ticket_Price, fill = Purchase_Method)) +
  stat_summary(fun = "median", geom = "bar",
               color = "black", width = 0.6) +
  stat_summary(fun = "median", geom = "text",
               aes(label = round(..y.., 2)), vjust = -0.5, size = 4) +
  scale_fill_manual(values = c("Online" = "#6A0DAD",
                               "In-Person" = "#DDA0DD" )) +
  scale_y_continuous(expand = expansion(mult = c(0.05, 0.1))) +
  labs(title = "Median Final Ticket Price by Purchase Method",
       x = "Purchase Method",
       y = "Median Final Ticket Price",
       fill = "Purchase Method") +
  theme_minimal() +
  theme(legend.position = "right",
    axis.text.x = element_text(size = 10),
    axis.text.y = element_text(size = 10),
    panel.grid.major.y = element_blank(),
    panel.grid.major.x = element_line(color = "gray90"),
    plot.margin = unit(c(1, 1, 1, 1), "cm"))


Plot9
```


Median Final Ticket Price by Purchase Method
14.99
12.5
15
10
5
0
Median Final Ticket Price
Online
In-Person
Purchase Method
Purchase Method
In-Person
Online

```
#### Grouped Bar Charts -- Two Categorical Variables Against One Continuous Variable


### Analyzing Price Variations for Days of the Week


# When Days of the Week becomes a constant fill variable,
# both 'Screening_Daypart' and 'Purchase_Method' produce zero insights
# since there is little to no price variation.

# When analyzing 'Ticket_Type' or 'Format',
# new price variation insights are highlighted.


## 'Ticket_Type'

GroupedPlot1 <- ggplot(MovieTheaterDataSet,
       aes(x = Ticket_Type,
           y = Final_Ticket_Price,
           fill = Day_of_the_Week)) +
  stat_summary(fun = "median", geom = "bar",
               position = position_dodge(width = 0.7),
               color = "black",
               width = 0.4) +
  scale_fill_manual(values = c("Monday 1" = "#03045E", "Tuesday" = "#023E8A",
    "Wednesday" = "#0077B6", "Thursday" = "#0096C7", "Friday" = "#00B4D8",
    "Saturday" = "#48CAE4", "Sunday" = "#90E0EF", "Monday 2" = "#ADE8F4")) +
  labs(title = "Median Final Ticket Price by Ticket Type and Day of the Week",
       x = "Ticket Type",
       y = "Median Final Ticket Price") +
  theme_minimal() +
  theme(axis.text.x = element_text(size = 10, angle = 45, hjust = 1),
        axis.text.y = element_text(size = 9),
        panel.grid.major.y = element_blank(),
        panel.grid.major.x = element_line(color = "gray90"),
        legend.position = "none",
        plot.margin = unit(c(1, 1, 1, 1), "cm"))


GroupedPlot1
```


Median Final Ticket Price by Ticket Type and Day of the Week
Median Final Ticket Price
20
15
10
5
0
Adult
Child
Senior
Admit One
FC Promo
Ticket Type

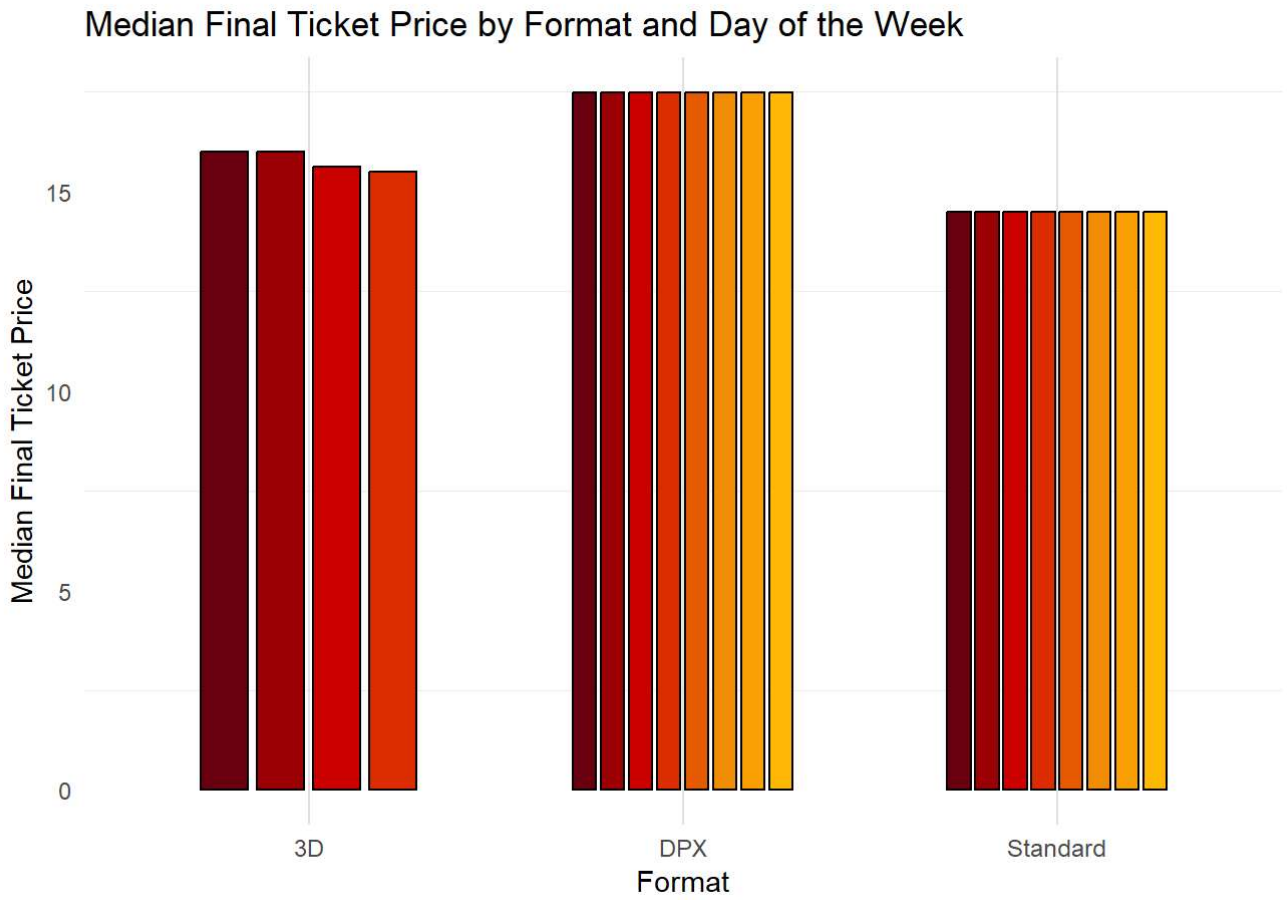
```
# As you can see, outputs 'Admit One' and 'FC Promo'
# exhibit the most price variation.

TicketTypeData <- MovieTheaterDataSet %>%
  filter(Ticket_Type %in% c("Admit One", "FC Promo"))


PVTicketType <- ggplot(TicketTypeData,
                       aes(x = Ticket_Type,
                           y = Final_Ticket_Price,
                           fill = Day_of_the_Week)) +
  stat_summary(fun = "median", geom = "bar",
               position = position_dodge(width = 0.7),
               color = "black", width = 0.4) +
  scale_fill_manual(values = c("Monday 1" = "#03045E", "Tuesday" = "#023E8A",
                               "Wednesday" = "#0077B6", "Thursday" = "#0096C7",
                               "Friday" = "#00B4D8", "Saturday" = "#48CAE4",
                               "Sunday" = "#90E0EF", "Monday 2" = "#ADE8F4")) +
  labs(title = "Median Final Ticket Price: Admit One and FC Promo by Day",
       x = "Ticket Type",
       y = "Median Final Ticket Price",
       fill = "Day of the Week") +
  theme_minimal() +
  theme(axis.text.x = element_text(
    size = 10, angle = 0, hjust = 0.5),
    axis.text.y = element_text(size = 9),
    panel.grid.major.y = element_blank(),
    panel.grid.major.x = element_line(color = "gray90"),
    legend.position = "none",
    plot.margin = unit(c(1, 1, 1, 1), "cm"))

 PVTicketType
```




```
 # Conducting a T-test can allow us to investigate whether the
 # difference between 'Admit One' and 'FC Promo' is statistically
 # significant and not due to random chance.

AdmitOneData <- subset(MovieTheaterDataSet, Ticket_Type == "Admit One")
FCPromoData <- subset(MovieTheaterDataSet, Ticket_Type == "FC Promo")
t.test(AdmitOneData$Final_Ticket_Price,
       FCPromoData$Final_Ticket_Price,
       var.equal = FALSE)
```

```
##
##  Welch Two Sample t-test
##
## data:  AdmitOneData$Final_Ticket_Price and FCPromoData$Final_Ticket_Price
## t = 10.677, df = 43.116, p-value = 1.102e-13
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##   9.057552 13.275781
## sample estimates:
## mean of x mean of y
##  12.41167   1.24500
```

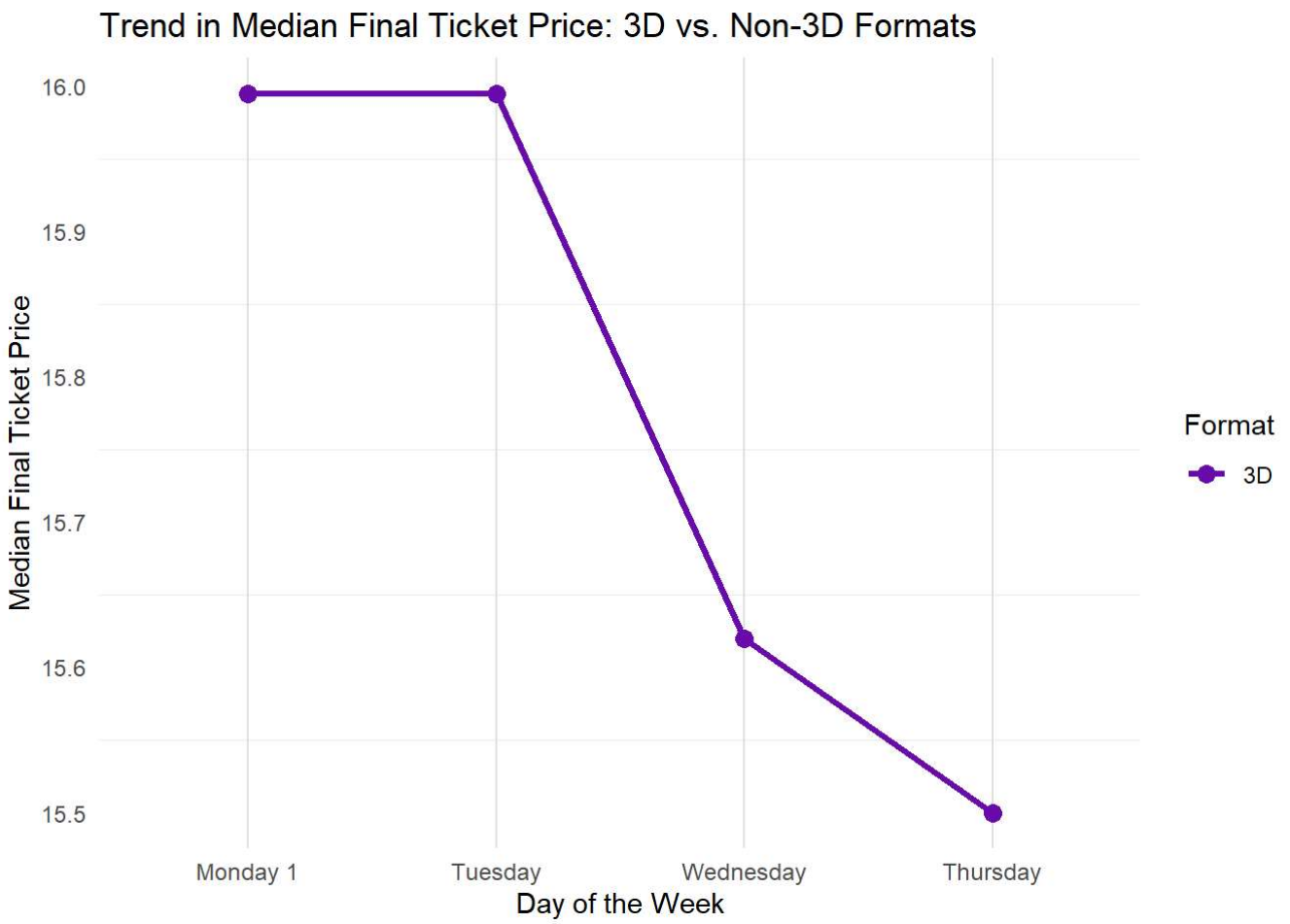
```
## Format

GroupedPlot2 <- ggplot(MovieTheaterDataSet,
                       aes(x = Format,
                           y = Final_Ticket_Price,
                           fill = Day_of_the_Week)) +
  stat_summary(fun = "median", geom = "bar",
               position = position_dodge(width = 0.6),
               color = "black", width = 0.5) +
  scale_fill_manual(values = c("Monday 1" = "#6A040F", "Tuesday" = "#9D0208",
                               "Wednesday" = "#D00000", "Thursday" = "#DC2F02",
                               "Friday" = "#E85D04", "Saturday" = "#F48C06",
                               "Sunday" = "#FAA307", "Monday 2" = "#FFBA08")) +
  labs(title = "Median Final Ticket Price by Format and Day of the Week",
       x = "Format",
       y = "Median Final Ticket Price",
       fill = "Day of the Week") +
  theme_minimal() +
  theme(axis.text.x = element_text(size = 9),
        axis.text.y = element_text(size = 9),
        panel.grid.major.y = element_blank(),
        panel.grid.major.x = element_line(color = "gray90"),
        legend.position = "none",
        plot.margin = unit(c(0.5, 0.5, 0.5, 0.5), "cm"))


 GroupedPlot2
```


Median Final Ticket Price by Format and Day of the Week
Median Final Ticket Price
15
10
5
0
3D
DPX
Standard
Format

```
 # Only output '3D' exhibits meaningful price variation.
 # You will also notice this price variation is
 # day-specific and appears as if the Ticket Price
 # is decreasing over the course of time.

 # A line chart can help visualize the passage of time.

MovieTheaterDataSet$Is3D <-
  ifelse(MovieTheaterDataSet$Format == "3D", "3D", "Non-3D")

FormatData <- MovieTheaterDataSet %>%
  filter(Format %in% c("3D", "Non-3D"))

LC3DPlot <- ggplot(FormatData,
                   aes(x = Day_of_the_Week,
                       y = Final_Ticket_Price,
                       group = Is3D,
                       color = Is3D)) +
  stat_summary(fun = "median", geom = "line", size = 1.2) +
  stat_summary(fun = "median", geom = "point", size = 3) +
  scale_color_manual(values = c("3D" = "#6A0DAD", "Non-3D" = "#DDA0DD")) +
  labs(title = "Trend in Median Final Ticket Price: 3D vs. Non-3D Formats",
       x = "Day of the Week",
       y = "Median Final Ticket Price",
       color = "Format") +
  theme_minimal() +
  theme(axis.text.x = element_text(size = 9),
        axis.text.y = element_text(size = 9),
        panel.grid.major.y = element_blank(),
        panel.grid.major.x = element_line(color = "gray90"),
        legend.position = "right")
```

```
## Warning: Using `size` aesthetic for lines was deprecated in ggplot2 3.4.0.
## ℹ Please use `linewidth` instead.
## This warning is displayed once every 8 hours.
## Call `lifecycle::last_lifecycle_warnings()` to see where this warning was
## generated.
```

```
LC3DPlot
```




```
 # Conducting a T-test can allow us to investigate whether the
 # difference between '3D' and 'Non-3D' is statistically
 # significant and not due to random chance.

FormatDataFiltered <- MovieTheaterDataSet %>%
  filter(Is3D %in% c("3D", "Non-3D"))
t.test(Final_Ticket_Price ~ Is3D,
       data = FormatDataFiltered,
       var.equal = FALSE)
```

```
##
##  Welch Two Sample t-test
##
## data:  Final_Ticket_Price by Is3D
## t = 6.2036, df = 101.37, p-value = 1.212e-08
## alternative hypothesis: true difference in means between group 3D and group Non-3D is not equ
al to 0
## 95 percent confidence interval:
##  1.048530 2.034278
## sample estimates:
##     mean in group 3D mean in group Non-3D
##             16.23458             14.69318
```

Econometrics

```
#### Multi-Linear Regression

### Variables of interest include 'Screening_Daypart', 'Time_Periods', 'Genre', 'Day_of_the_Wee
k', 'Format', 'Ticket_Type', 'Base_Ticket_Price', 'Purchase_Method', 'Discounted_Ticket', 'Speci
al_Event_Pricing', 'Special_Program', and 'Age_Rating'.

# These variables may or may not be included in the Final Regression Model.

# We must make the determination as to whether certain variables of interest might be eliminated
from the model.


### Variables of Concern include: 'Time_Periods', Special_Program'.

# Both 'Time_Periods' and 'Special_Program' are variables that are highly related to other varia
bles: 'Screening_Daypart' and 'Special_Event_Pricing', respectively.

# We can test for multicollinearity within both variables.

# We calculated the correlation between 'Time_Periods' and 'Screening_Daypart':
MovieTheaterDataSet$Time_Periods_Num <- as.numeric(as.factor(MovieTheaterDataSet$Time_Periods))
MovieTheaterDataSet$Screening_Daypart_Num <- as.numeric(as.factor(MovieTheaterDataSet$Screening_
Daypart))
Correlation_tp_sd <- cor(MovieTheaterDataSet$Time_Periods_Num,
                         MovieTheaterDataSet$Screening_Daypart_Num,
                         use = "complete.obs" )
Correlation_tp_sd
```

```
## [1] 0.8898868
```

```
# A correlation of 0.89 indicates there is a strong positive linear relationship between both va
riables. 0.89 > 0.7 threshold for multicollinearity.
# In this instance we would ELIMINATE 'Time_Periods' in favor of 'Screening_Daypart' as the latt
er has better predictive power.


# We must do the same for 'Special_Program' and 'Special_Event_Pricing':
MovieTheaterDataSet$Special_Program_Num <- as.numeric(as.factor(MovieTheaterDataSet$Special_Prog
ram))
MovieTheaterDataSet$Special_Event_Pricing_Num <- as.numeric(as.factor(MovieTheaterDataSet$Specia
l_Event_Pricing))
Correlation_sp_sep <- cor(MovieTheaterDataSet$Special_Program_Num,
                          MovieTheaterDataSet$Special_Event_Pricing_Num,
                          use = "complete.obs")
Correlation_sp_sep
```

```
## [1] 0.289545
```

```
# A correlation of 0.29 indicates a weak positive linear relationship between both variables. 0.29 < 0.7 threshold for multicollinearity.
# While there is No Multicollinearity, the sample sizes within 'Special_Program' are imbalanced and inadequate for the Regression Model.
SpecialProgramCount
```

```
## # A tibble: 6 × 2
##   Special_Program  Count
##   <chr>            <int>
## 1 AXCN                12
## 2 Fathom Events        4
## 3 Flashback Cinema    28
## 4 None              3043
## 5 Studio Ghibli        6
## 6 Unique              10
```

```
# In this instance we would ELIMINATE 'Special_Program' as the variable has no practical signifi
cance.



### Single-Factor ANOVA Tests

VariableOfInterest <- c("Screening_Daypart", "Genre",
                        "Day_of_the_Week", "Format",
                        "Ticket_Type", "Base_Ticket_Price",
                        "Purchase_Method", "Discounted_Ticket",
                        "Special_Event_Pricing", "Age_Rating")

run_anova <- function(variable) {
  formula <- as.formula(paste("Final_Ticket_Price ~", variable))
  anova_result <- aov(formula, data = MovieTheaterDataSet)
  cat("\nANOVA for", variable, ":\n")
  print(summary(anova_result))
}

SingleFactorAnovaOutput <- capture.output(lapply(VariableOfInterest, run_anova))
writeLines(SingleFactorAnovaOutput, "anova_results.txt")




### Multi-Factor ANOVA Tests


# Multi-Factor ANOVA tests can be used to examine
# combined effects variables have on the Final_Ticket_Price.
anova_formula <- Final_Ticket_Price ~ Screening_Daypart + Genre + Day_of_the_Week +
  Format + Ticket_Type + Base_Ticket_Price + Purchase_Method +
  Discounted_Ticket + Special_Event_Pricing + Age_Rating

multi_factor_anova <- aov(anova_formula, data = MovieTheaterDataSet)

summary(multi_factor_anova)
```

```
##                      Df Sum Sq Mean Sq   F value  Pr(>F)
## Screening_Daypart     1    901     901 3.006e+04 < 2e-16 ***
## Genre                 9   1105     123 4.093e+03 < 2e-16 ***
## Day_of_the_Week       7    142      20 6.770e+02 < 2e-16 ***
## Format                2   2707    1354 4.514e+04 < 2e-16 ***
## Ticket_Type           4   6553    1638 5.464e+04 < 2e-16 ***
## Base_Ticket_Price     1   2351    2351 7.841e+04 < 2e-16 ***
## Purchase_Method       1   4754    4754 1.585e+05 < 2e-16 ***
## Discounted_Ticket     1      0       0 6.600e-02 0.79697
## Age_Rating            4      0       0 3.810e+00 0.00429 **
## Residuals          3072     92       0
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
### The Big Picture

## Strongest Predictors (From High to Low F-Values)
# ("Purchase_Method","Base_Ticket_Price",
# "Ticket_Type", "Format", "Screening_Daypart")

## Moderate Predictors
# ("Genre","Day_of_the_Week")

## Marginal Predictors
# ("Age_Rating")

## Insignificant Predictors
# ("Discounted_Ticket",
#  "Special_Event_Pricing(???)")



### Interaction Effects


## Exploring Format:Day_of_the_Week

anova_interaction1 <- aov(Final_Ticket_Price ~ Format * Day_of_the_Week,
                          data = MovieTheaterDataSet)
summary(anova_interaction1)
```

```
##                          Df Sum Sq Mean Sq F value   Pr(>F)
## Format                    2   3240  1620.0 329.239  < 2e-16 ***
## Day_of_the_Week           7    172    24.5   4.981 1.27e-05 ***
## Format:Day_of_the_Week   10     25     2.5   0.506    0.887
## Residuals              3083  15169     4.9
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# The interaction between both variables is not statistically significant.
# Do NOT include in the final regression model.



## Exploring Screening_Daypart:Format

anova_interaction2 <- aov(Final_Ticket_Price ~ Screening_Daypart * Format,
                          data = MovieTheaterDataSet)
summary(anova_interaction2)
```

```
##                             Df Sum Sq Mean Sq F value Pr(>F)
## Screening_Daypart            1    901   901.2  193.18 <2e-16 ***
## Format                       2   3215  1607.3  344.53 <2e-16 ***
## Screening_Daypart:Format     2     42    20.8    4.46 0.0116 *
## Residuals                 3097  14448     4.7
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# The interaction between both variables is statistically significant.
# Include in the final regression model.



## Exploring Ticket_Type:Purchase_Method

anova_interaction3 <- aov(Final_Ticket_Price ~ Ticket_Type * Purchase_Method,
                          data = MovieTheaterDataSet)
summary(anova_interaction3)
```

```
##                                Df Sum Sq Mean Sq  F value Pr(>F)
## Ticket_Type                     4   7082    1771  810.794 <2e-16 ***
## Purchase_Method                 1   4769    4769 2183.595 <2e-16 ***
## Ticket_Type:Purchase_Method     4      0       0    0.034  0.998
## Residuals                    3093   6754       2
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# The interaction between both variables is not statistically significant.
# DO NOT include in the final regression model.



## Exploring Genre:Screening_Daypart

anova_interaction4 <- aov(Final_Ticket_Price ~ Genre * Screening_Daypart,
                          data = MovieTheaterDataSet)
summary(anova_interaction4)
```

```
##                            Df Sum Sq Mean Sq F value Pr(>F)
## Genre                       9   1151   127.9  23.905 <2e-16 ***
## Screening_Daypart           1    854   854.4 159.650 <2e-16 ***
## Genre:Screening_Daypart     8     96    12.0   2.235 0.0224 *
## Residuals                3084  16504     5.4
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# The interaction between both variables is statistically significant.
# Include in the final regression model.



## Exploring Special_Event_Pricing:Day_of_the_Week

anova_interaction5 <- aov(Final_Ticket_Price ~ Special_Event_Pricing * Day_of_the_Week,
                          data = MovieTheaterDataSet)
summary(anova_interaction5)
```

```
##                                      Df Sum Sq Mean Sq F value  Pr(>F)
## Special_Event_Pricing                 1   1278  1278.5 255.718 < 2e-16 ***
## Day_of_the_Week                       7    117    16.7   3.333 0.00153 **
## Special_Event_Pricing:Day_of_the_Week 5   1767   353.3  70.675 < 2e-16 ***
## Residuals                          3089  15444     5.0
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# The interaction between both variables is statistically significant.
# Include in the final regression model.


## Final Interaction Terms
# Statistically significant: (Screening_Daypart:Format,
#                             Genre:Screening_Daypart,
#                             Special_Event_Pricing:Day_of_the_Week)
# NOT statistically significant (Exclude): (Format:Day_of_the_Week,
#                                           Ticket_Type:Purchase_Method)



### Final List of Variables


## Continuous Variables
# Base_Ticket_Price


## Categorical Variables
# Screening_Daypart
# Genre
# Day_of_the_Week
# Format
# Ticket_Type
# Purchase_Method
# Special_Event_Pricing
# Discounted_Ticket
# Age_Rating


## Interaction Terms
# Screening_Daypart:Format
# Genre:Screening_Daypart
# Special_Event_Pricing:Day_of_the_Week



### Final Regression Model


FinalRegressionModel <- lm(Final_Ticket_Price ~
                             Screening_Daypart + Genre + Day_of_the_Week +
                             Format + Ticket_Type + Base_Ticket_Price +
                             Purchase_Method + Special_Event_Pricing +
                             Discounted_Ticket + Age_Rating +
                             Screening_Daypart:Format +
                             Genre:Screening_Daypart +
                             Special_Event_Pricing:Day_of_the_Week,
                           data = MovieTheaterDataSet)
```

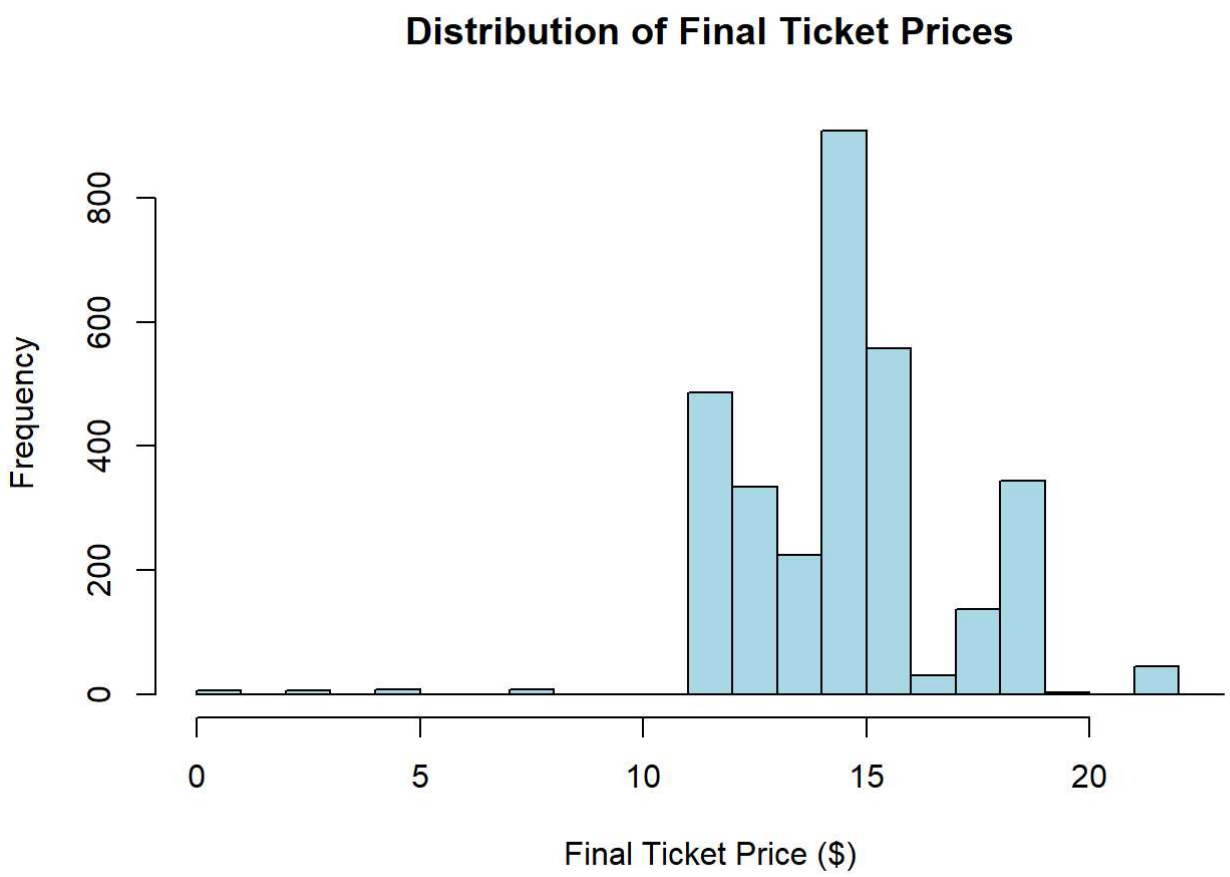
```
summary(FinalRegressionModel)
```

```
##
## Call:
## lm(formula = Final_Ticket_Price ~ Screening_Daypart + Genre +
##     Day_of_the_Week + Format + Ticket_Type + Base_Ticket_Price +
##     Purchase_Method + Special_Event_Pricing + Discounted_Ticket +
##     Age_Rating + Screening_Daypart:Format + Genre:Screening_Daypart +
##     Special_Event_Pricing:Day_of_the_Week, data = MovieTheaterDataSet)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.3971 -0.0156 -0.0015  0.0128  3.5980
##
## Coefficients: (5 not defined because of singularities)
##                                                Estimate Std. Error t value
## (Intercept)                                   3.849e-01  9.865e-02   3.902
## Screening_DaypartEvening                      3.385e-02  5.104e-02   0.663
## GenreDocumentary                             -2.205e-01  9.485e-02  -2.324
## GenreAnimation                               -2.251e-01  6.661e-02  -3.379
## GenreHorror                                  -2.353e-01  6.988e-02  -3.367
## GenreAction                                  -2.303e-01  6.516e-02  -3.534
## GenreAnime                                   -1.735e-01  1.236e-01  -1.404
## GenreFantasy                                 -2.205e-01  6.588e-02  -3.346
## GenreThriller                                -1.429e-01  6.373e-02  -2.243
## GenreDrama                                   -2.242e-01  6.876e-02  -3.261
## GenreComedy                                  -2.373e-01  6.010e-02  -3.949
## Day_of_the_WeekTuesday                        3.786e-05  1.279e-02   0.003
## Day_of_the_WeekWednesday                      3.577e-03  1.293e-02   0.277
## Day_of_the_WeekThursday                       2.030e-02  1.286e-02   1.578
## Day_of_the_WeekFriday                         1.401e-02  1.264e-02   1.108
## Day_of_the_WeekSaturday                       1.413e-02  1.274e-02   1.109
## Day_of_the_WeekSunday                         1.411e-02  1.288e-02   1.095
## Day_of_the_WeekMonday 2                       2.347e-02  1.317e-02   1.782
## FormatDPX                                     7.091e-03  2.900e-02   0.245
## FormatStandard                               -2.228e-02  2.659e-02  -0.838
## Ticket_TypeChild                             -1.859e-02  1.251e-02  -1.487
## Ticket_TypeSenior                            -6.241e-03  1.347e-02  -0.463
## Ticket_TypeAdmit One                         -1.012e-02  1.085e-01  -0.093
## Ticket_TypeFC Promo                          -3.469e-02  1.501e-01  -0.231
## Base_Ticket_Price                             9.888e-01  4.134e-03 239.168
## Purchase_MethodOnline                         2.476e+00  6.107e-03 405.386
## Special_Event_PricingYes                             NA         NA      NA
## Discounted_TicketYes                         -3.138e-02  7.386e-02  -0.425
## Age_RatingR                                   1.608e-02  1.230e-02   1.308
## Age_RatingPG                                  7.524e-03  1.327e-02   0.567
## Age_RatingNR                                 -2.074e-01  3.070e-02  -6.756
## Age_RatingTBD                                -3.885e-02  1.146e-01  -0.339
## Age_RatingPG13                                       NA         NA      NA
## Screening_DaypartEvening:FormatDPX           -1.495e-02  4.209e-02  -0.355
## Screening_DaypartEvening:FormatStandard      -1.814e-02  3.946e-02  -0.460
## Screening_DaypartEvening:GenreDocumentary     1.343e-01  1.443e-01   0.930
## Screening_DaypartEvening:GenreAnimation      -5.088e-03  3.374e-02  -0.151
## Screening_DaypartEvening:GenreHorror          3.694e-01  4.538e-02   8.139
```

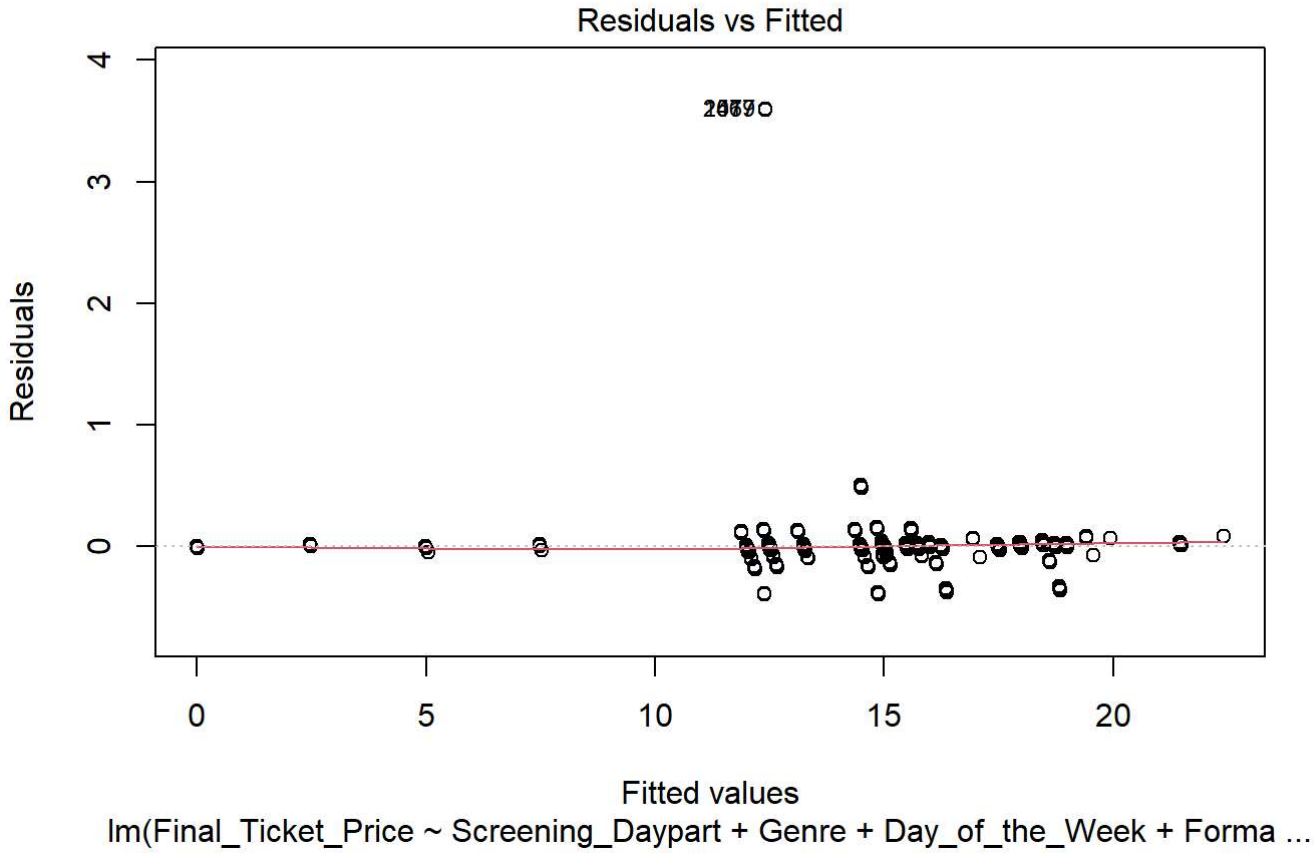
```
## Screening_DaypartEvening:GenreAction              -4.432e-03  3.337e-02  -0.133
## Screening_DaypartEvening:GenreAnime               -5.448e-03  5.063e-02  -0.108
## Screening_DaypartEvening:GenreFantasy             -8.878e-03  3.314e-02  -0.268
## Screening_DaypartEvening:GenreThriller            -8.561e-02  4.346e-02  -1.970
## Screening_DaypartEvening:GenreDrama               -4.686e-03  5.435e-02  -0.086
## Screening_DaypartEvening:GenreComedy                      NA         NA      NA
## Day_of_the_WeekTuesday:Special_Event_PricingYes   -1.428e-02  1.578e-01  -0.091
## Day_of_the_WeekWednesday:Special_Event_PricingYes -3.206e-02  1.109e-01  -0.289
## Day_of_the_WeekThursday:Special_Event_PricingYes  -1.823e-01  1.322e-01  -1.379
## Day_of_the_WeekFriday:Special_Event_PricingYes            NA         NA      NA
## Day_of_the_WeekSaturday:Special_Event_PricingYes  -2.132e-02  1.456e-01  -0.146
## Day_of_the_WeekSunday:Special_Event_PricingYes    -8.419e-02  1.247e-01  -0.675
## Day_of_the_WeekMonday 2:Special_Event_PricingYes          NA         NA      NA
##                                                  Pr(>|t|)
## (Intercept)                                      9.75e-05 ***
## Screening_DaypartEvening                         0.507272
## GenreDocumentary                                 0.020175 *
## GenreAnimation                                   0.000737 ***
## GenreHorror                                      0.000768 ***
## GenreAction                                      0.000415 ***
## GenreAnime                                       0.160535
## GenreFantasy                                     0.000828 ***
## GenreThriller                                    0.024993 *
## GenreDrama                                       0.001124 **
## GenreComedy                                      8.04e-05 ***
## Day_of_the_WeekTuesday                           0.997637
## Day_of_the_WeekWednesday                         0.782095
## Day_of_the_WeekThursday                          0.114655
## Day_of_the_WeekFriday                            0.267807
## Day_of_the_WeekSaturday                          0.267532
## Day_of_the_WeekSunday                            0.273564
## Day_of_the_WeekMonday 2                          0.074861 .
## FormatDPX                                        0.806836
## FormatStandard                                   0.402133
## Ticket_TypeChild                                 0.137216
## Ticket_TypeSenior                                0.643250
## Ticket_TypeAdmit One                             0.925739
## Ticket_TypeFC Promo                              0.817201
## Base_Ticket_Price                                 < 2e-16 ***
## Purchase_MethodOnline                             < 2e-16 ***
## Special_Event_PricingYes                               NA
## Discounted_TicketYes                             0.670908
## Age_RatingR                                      0.191128
## Age_RatingPG                                     0.570814
## Age_RatingNR                                     1.70e-11 ***
## Age_RatingTBD                                    0.734637
## Age_RatingPG13                                         NA
## Screening_DaypartEvening:FormatDPX               0.722428
## Screening_DaypartEvening:FormatStandard          0.645757
## Screening_DaypartEvening:GenreDocumentary        0.352257
## Screening_DaypartEvening:GenreAnimation          0.880123
## Screening_DaypartEvening:GenreHorror             5.72e-16 ***
```

```
## Screening_DaypartEvening:GenreAction              0.894328
## Screening_DaypartEvening:GenreAnime               0.914306
## Screening_DaypartEvening:GenreFantasy             0.788819
## Screening_DaypartEvening:GenreThriller            0.048971 *
## Screening_DaypartEvening:GenreDrama               0.931295
## Screening_DaypartEvening:GenreComedy                    NA
## Day_of_the_WeekTuesday:Special_Event_PricingYes   0.927873
## Day_of_the_WeekWednesday:Special_Event_PricingYes 0.772468
## Day_of_the_WeekThursday:Special_Event_PricingYes  0.168116
## Day_of_the_WeekFriday:Special_Event_PricingYes          NA
## Day_of_the_WeekSaturday:Special_Event_PricingYes  0.883615
## Day_of_the_WeekSunday:Special_Event_PricingYes    0.499632
## Day_of_the_WeekMonday 2:Special_Event_PricingYes        NA
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1701 on 3057 degrees of freedom
## Multiple R-squared:  0.9952, Adjusted R-squared:  0.9952
## F-statistic: 1.422e+04 on 45 and 3057 DF,  p-value: < 2.2e-16
```

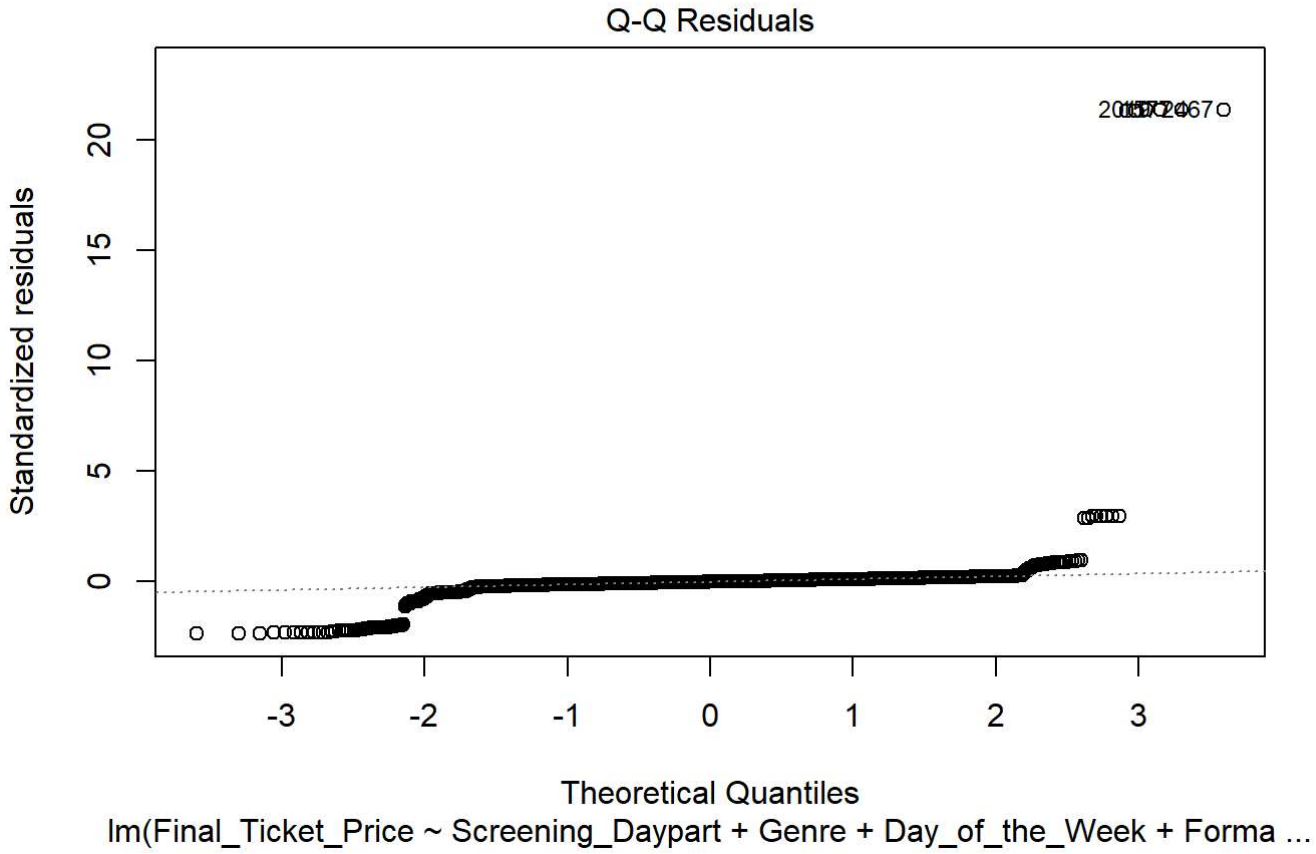
```
## Model Accuracy

# Basic Distribution of Final_Ticket_Price
hist(MovieTheaterDataSet$Final_Ticket_Price,
     breaks = 20,
     col = "lightblue",
     border = "black",
     main = "Distribution of Final Ticket Prices",
     xlab = "Final Ticket Price ($)",
     ylab = "Frequency")
```




```
# Residual vs. Fitted Plot
plot(FinalRegressionModel, which = 1)
```

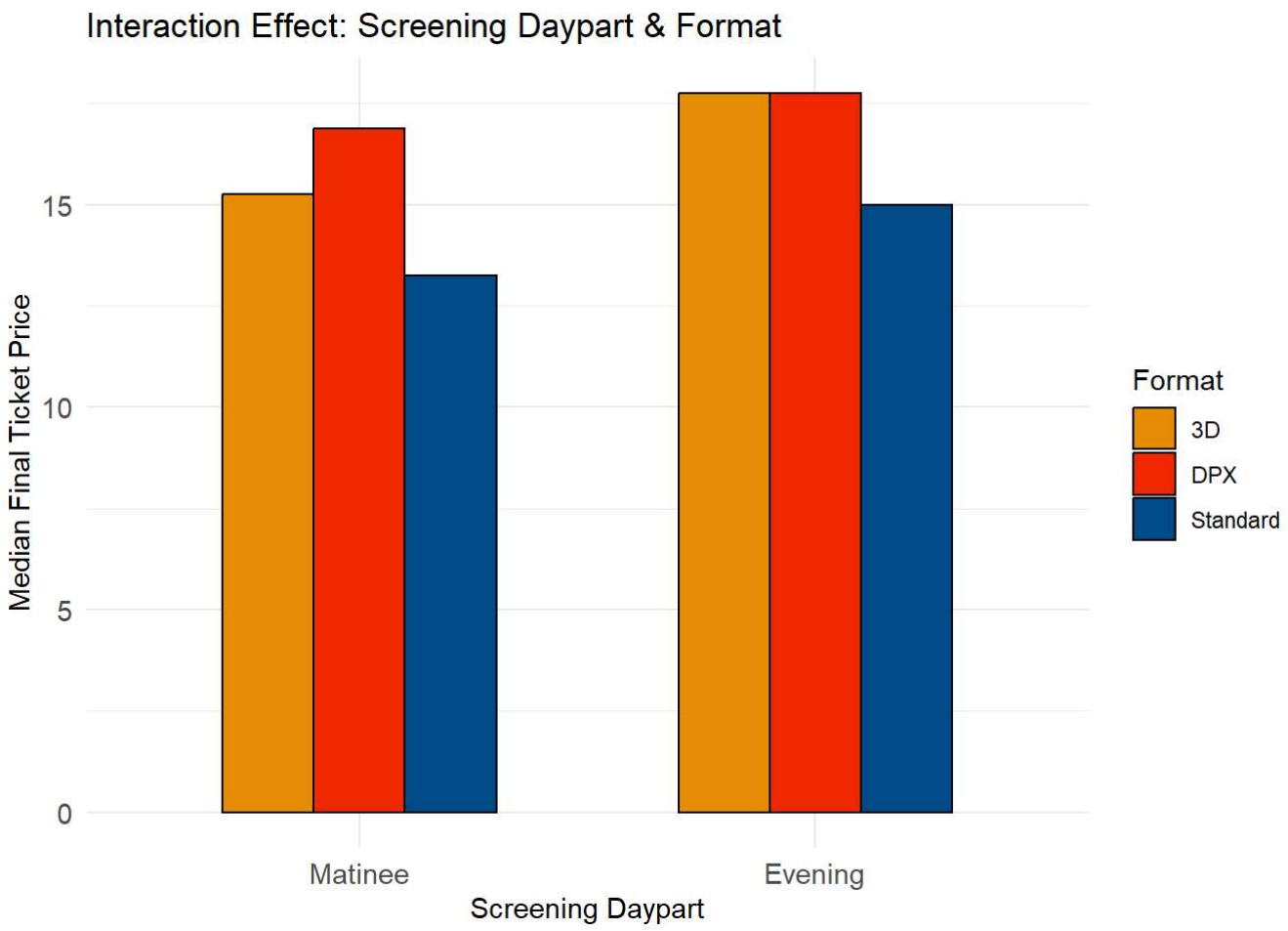



```
# QQ Plot
plot(FinalRegressionModel, which = 2)
```

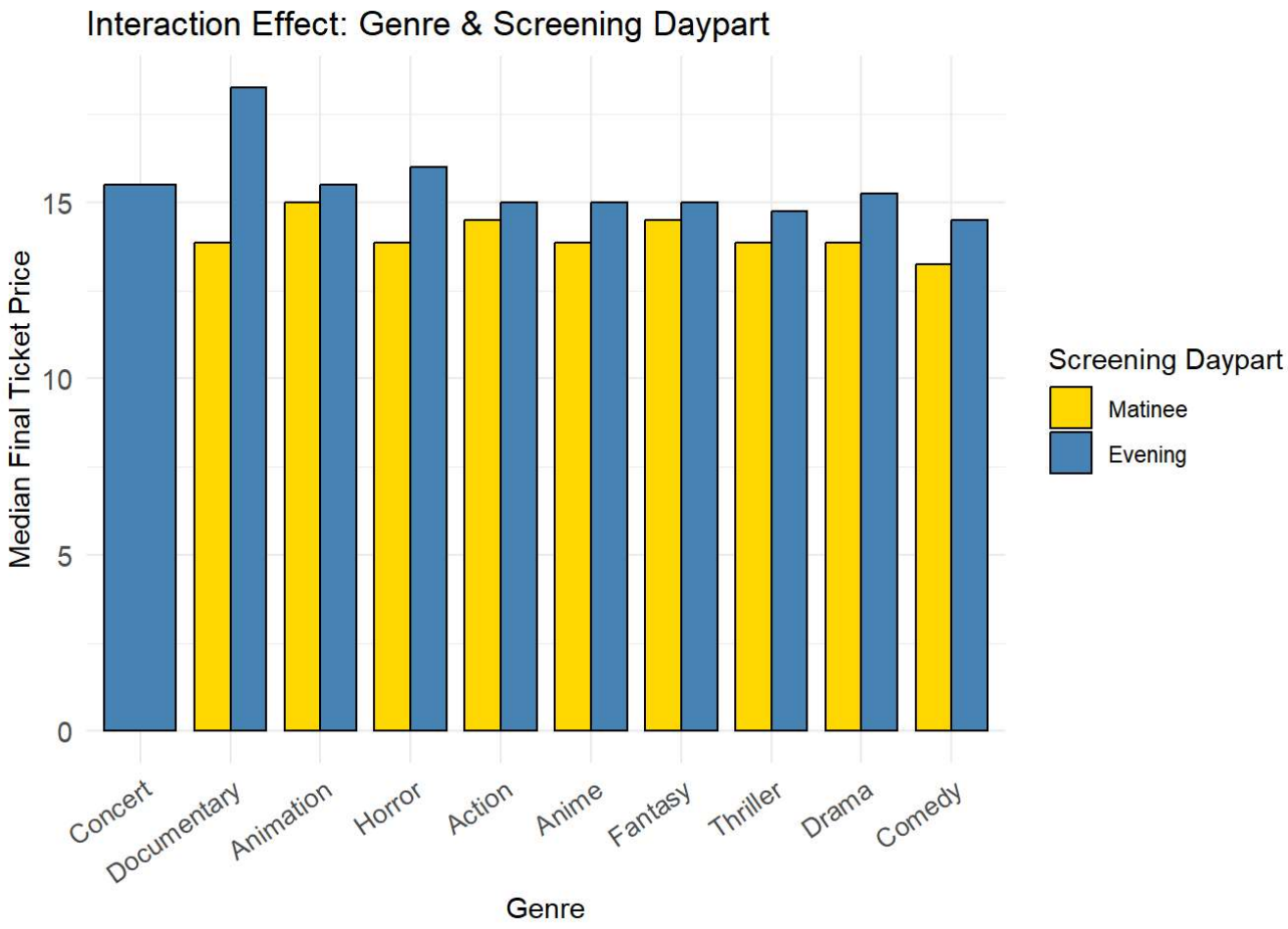



```
# Predicted vs Actual Plot
MovieTheaterDataSet$Predicted_Price <- predict(FinalRegressionModel, newdata = MovieTheaterDataS
et)

ggplot(MovieTheaterDataSet, aes(x = Final_Ticket_Price, y = Predicted_Price)) +
  geom_point(alpha = 0.5, color = "blue") +
  geom_abline(slope = 1, intercept = 0, linetype = "dashed", color = "red") +
  labs(title = "Predicted vs. Actual Ticket Prices",
       x = "Actual Final Ticket Price",
       y = "Predicted Final Ticket Price") +
  theme_minimal()
```

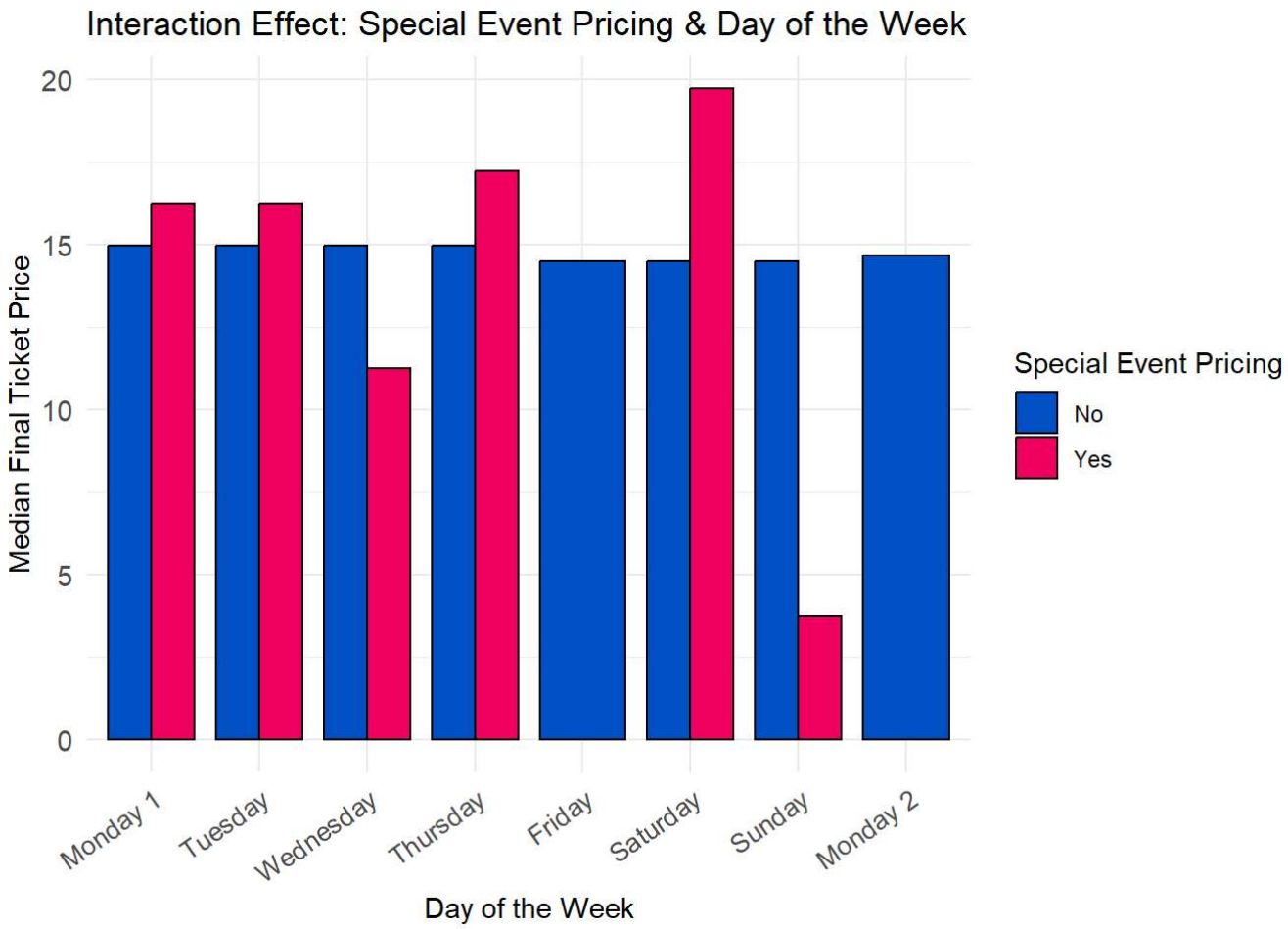

```
## Other Visualizations

# Interaction Effect Between Screening Daypart & Format
InteractionPlot1 <- ggplot(MovieTheaterDataSet,
                           aes(x = Screening_Daypart,
                               y = Final_Ticket_Price,
                               fill = Format)) +
  stat_summary(fun = "median", geom = "bar",
               position = "dodge", color = "black", width = 0.6) +
  scale_fill_manual(values = c("Standard" = "#004E89",
                               "3D" = "#E89005",
                               "DPX" = "#F42B03")) +
  labs(title = "Interaction Effect: Screening Daypart & Format",
       x = "Screening Daypart",
       y = "Median Final Ticket Price",
       fill = "Format") +
  theme_minimal() +
  theme(axis.text.x = element_text(size = 11),
        axis.text.y = element_text(size = 11),
        legend.position = "right")

InteractionPlot1
```



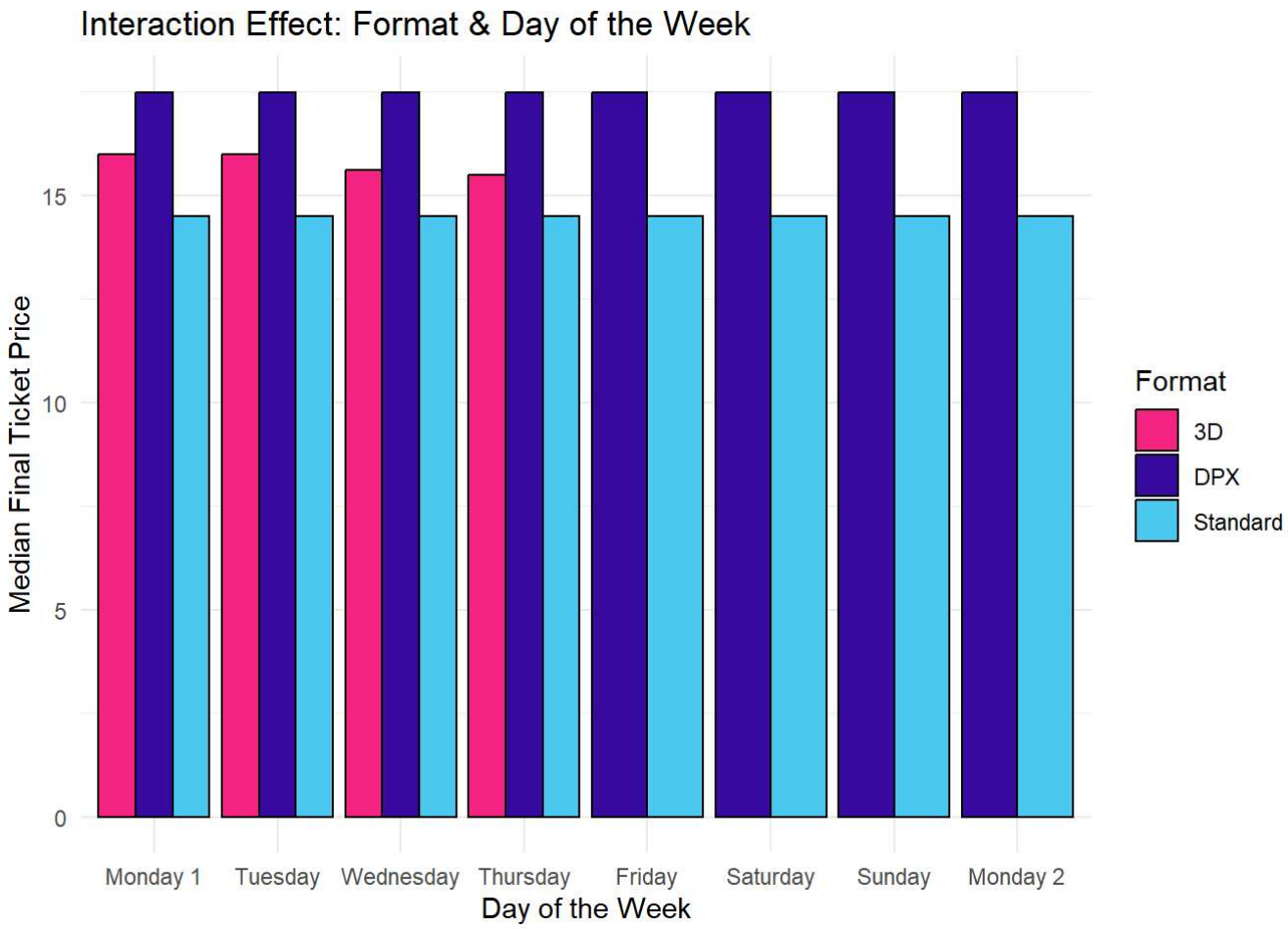

```
# Interaction Effect Between Genre & Screening Daypart
InteractionPlot2 <- ggplot(MovieTheaterDataSet,
                           aes(x = Genre,
                               y = Final_Ticket_Price,
                               fill = Screening_Daypart)) +
  stat_summary(fun = "median", geom = "bar",
               position = "dodge", color = "black", width = 0.8) +
  scale_fill_manual(values = c("Matinee" = "#FFD700",
                               "Evening" = "#4682B4")) +
  labs(title = "Interaction Effect: Genre & Screening Daypart",
       x = "Genre",
       y = "Median Final Ticket Price",
       fill = "Screening Daypart") +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 35, hjust = 1, size = 10),
        axis.text.y = element_text(size = 11),
        legend.position = "right")

InteractionPlot2
```




```
# Interaction Effect Between Special Event Pricing & Day of the Week
InteractionPlot3 <- ggplot(MovieTheaterDataSet,
                           aes(x = Day_of_the_Week,
                               y = Final_Ticket_Price,
                               fill = Special_Event_Pricing)) +
  stat_summary(fun = "median", geom = "bar",
               position = "dodge", color = "black", width = 0.8) +
  scale_fill_manual(values = c("Yes" = "#F20262",
                               "No" = "#0052C4")) +
  labs(title = "Interaction Effect: Special Event Pricing & Day of the Week",
       x = "Day of the Week",
       y = "Median Final Ticket Price",
       fill = "Special Event Pricing") +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 35, hjust = 1, size = 10),
        axis.text.y = element_text(size = 11),
        legend.position = "right")

InteractionPlot3
```




```
# Interaction Effect Between Format & Day of the Week
InteractionPlot4 <- ggplot(MovieTheaterDataSet,
       aes(x = Day_of_the_Week,
           y = Final_Ticket_Price,
           fill = Format)) +
  stat_summary(fun = "median", geom = "bar",
               position = "dodge", color = "black") +
  scale_fill_manual(values = c("Standard" = "#4CC9F0",
                               "3D" = "#F72585",
                               "DPX" = "#3A0CA3")) +
  labs(title = "Interaction Effect: Format & Day of the Week",
       x = "Day of the Week",
       y = "Median Final Ticket Price",
       fill = "Format") +
  theme_minimal()

InteractionPlot4
```


Interaction Effect: Format & Day of the Week
Median Final Ticket Price
0
5
10
15
Monday 1
Tuesday
Wednesday
Thursday
Friday
Saturday
Sunday
Monday 2
Day of the Week
Format
3D
DPX
Standard