

ETE3-2.R

Pranab Rai-2447137

2025-01-02

```
# --- ANOVA Analysis ---
print("ANOVA Analysis:")

## [1] "ANOVA Analysis:"

# Load your data
df <-
read.csv("C:\\Users\\prana\\OneDrive\\Desktop\\2trimester\\R\\ETE3\\test2.csv")

View(df)

# 1. One-Way ANOVA: Total Amount by Pickup Hour
print(" 1. One-Way ANOVA: Does the average total amount vary significantly
across different hours of the day?")

## [1] " 1. One-Way ANOVA: Does the average total amount vary significantly
across different hours of the day?"

print("Hypotheses:")

## [1] "Hypotheses:"

print("H0: The average total amount is the same for all hours of the day.")

## [1] "H0: The average total amount is the same for all hours of the day."

print("Ha: At least one hour of the day has a different average total
amount.")
)

## [1] "Ha: At least one hour of the day has a different average total
amount."

anova_one_way_result <- aov(total_amount ~ factor(pickup_hour), data = df)
summary_table <- summary(anova_one_way_result)
print(summary_table)

##               Df Sum Sq Mean Sq F value Pr(>F)
## factor(pickup_hour) 23  19136    832.0    1.302  0.157
## Residuals        696 444859    639.2

print("Result: The average total amount does not vary significantly across
hours of the day.")
```

```
## [1] "Result: The average total amount does not vary significantly across
hours of the day."

#-----

# 2. One-Way ANOVA: Total Amount by Day of Week

print("2. One-Way ANOVA: Does the average total amount vary significantly
across different days of the week?")

## [1] "2. One-Way ANOVA: Does the average total amount vary significantly
across different days of the week?"

print("Hypotheses:")

## [1] "Hypotheses:"

print("H0: The average total amount is the same for all days of the week.")

## [1] "H0: The average total amount is the same for all days of the week."

print("Ha: At least one day of the week has a different average total
amount.")

## [1] "Ha: At least one day of the week has a different average total
amount."

anova_one_way_day <- aov(total_amount ~ day_of_week, data = df)
summary_table_day <- summary(anova_one_way_day)
print(summary_table_day)

##              Df Sum Sq Mean Sq F value Pr(>F)
## day_of_week    6   5123    853.8    1.327  0.243
## Residuals   713 458872    643.6

print("Result: The average total amount does *not* vary significantly across
days of the week.")

## [1] "Result: The average total amount does *not* vary significantly across
days of the week."

#-----

# 3. One-Way ANOVA: Trip Distance by Day of Week

print("3. One-Way ANOVA: Does the average trip distance vary significantly
across different days of the week?")

## [1] "3. One-Way ANOVA: Does the average trip distance vary significantly
across different days of the week?"

print("Hypotheses:")
```

```
## [1] "Hypotheses:"

print("H0: The average trip distance is the same for all days of the week.")
## [1] "H0: The average trip distance is the same for all days of the week."

print("Ha: At least one day of the week has a different average trip
distance.")

## [1] "Ha: At least one day of the week has a different average trip
distance."

anova_one_way_distance <- aov(trip_distance ~ day_of_week, data = df)
summary_table_distance <- summary(anova_one_way_distance)
print(summary_table_distance)

##           Df Sum Sq Mean Sq F value Pr(>F)
## day_of_week   6      51   8.502    0.408  0.874
## Residuals  713  14856  20.836

print("Result: The average trip distance does *not* vary significantly across
days of the week.")

## [1] "Result: The average trip distance does *not* vary significantly
across days of the week."

#-----

# 4. Two-Way ANOVA: Total Amount by Hour of Day and Day of Week
print("4. Two-Way ANOVA: Does the average total amount vary significantly by
both hour of day and day of week?")

## [1] "4. Two-Way ANOVA: Does the average total amount vary significantly by
both hour of day and day of week?"

print("Hypotheses:")

## [1] "Hypotheses:"

print("H0a (Main effect of Hour): The average total amount is the same for
all hours of the day.")

## [1] "H0a (Main effect of Hour): The average total amount is the same for
all hours of the day."

print("H1a: At least one hour of the day has a different average total
amount.")

## [1] "H1a: At least one hour of the day has a different average total
amount."

print(" ")

## [1] " "
```

```

print("H0b (Main effect of Day of Week): The average total amount is the same
for all days of the week.")

## [1] "H0b (Main effect of Day of Week): The average total amount is the
same for all days of the week."

print("H1b: At least one day of the week has a different average total
amount.")

## [1] "H1b: At least one day of the week has a different average total
amount."

print(" ")

## [1] " "

print("H0c (Interaction effect): There is no interaction between hour of day
and day of week on the average total amount.")

## [1] "H0c (Interaction effect): There is no interaction between hour of day
and day of week on the average total amount."

print("H1c: There is an interaction between hour of day and day of week on
the average total amount.")

## [1] "H1c: There is an interaction between hour of day and day of week on
the average total amount."

anova_two_way_result <- aov(total_amount ~ factor(pickup_hour) * day_of_week,
data = df)
print(summary(anova_two_way_result))

##
## factor(pickup_hour)          Df Sum Sq Mean Sq F value Pr(>F)
## day_of_week                 6   5276   879.3    1.387  0.218
## factor(pickup_hour):day_of_week 138 89596   649.2    1.024  0.419
## Residuals                   552 349988   634.0

print("Result: The average total amount does not vary significantly across
hours of the day.")

## [1] "Result: The average total amount does not vary significantly across
hours of the day."

print("Result: The average total amount does not vary significantly across
days of the week.")

## [1] "Result: The average total amount does not vary significantly across
days of the week."

#-----
# 5. Two-Way ANOVA: Trip Distance by Passenger Count and Day of Week

```

```

print("5. Two-Way ANOVA: Does the average trip distance vary significantly by
both passenger count and day of week?")

## [1] "5. Two-Way ANOVA: Does the average trip distance vary significantly
by both passenger count and day of week?"

print("Hypotheses:")

## [1] "Hypotheses:"

print("H0a (Main effect of Passenger Count): The average trip distance is the
same for all passenger counts.")

## [1] "H0a (Main effect of Passenger Count): The average trip distance is
the same for all passenger counts."

print("H1a: At least one passenger count has a different average trip
distance.")

## [1] "H1a: At least one passenger count has a different average trip
distance."

print("H0b (Main effect of Day of Week): The average trip distance is the
same for all days of the week.")

## [1] "H0b (Main effect of Day of Week): The average trip distance is the
same for all days of the week."

print("H1b: At least one day of the week has a different average trip
distance.")

## [1] "H1b: At least one day of the week has a different average trip
distance."

print("H0c (Interaction effect): There is no interaction between passenger
count and day of week on the average trip distance.")

## [1] "H0c (Interaction effect): There is no interaction between passenger
count and day of week on the average trip distance."

print("H1c: There is an interaction between passenger count and day of week
on the average trip distance.")

## [1] "H1c: There is an interaction between passenger count and day of week
on the average trip distance."

anova_two_way_distance <- aov(trip_distance ~ factor(passenger_count) *
day_of_week, data = df)
summary_table_two_way_distance <- summary(anova_two_way_distance)
print(summary_table_two_way_distance)

##               Df Sum Sq Mean Sq F value Pr(>F)
## factor(passenger_count)      6    224   37.36   1.781  0.100
## day_of_week                  6     51    8.43   0.402  0.878

```

```
## factor(passenger_count):day_of_week  22    261   11.88   0.566   0.946
## Residuals                        685  14371   20.98

print("Result: The average trip distance does not vary significantly across
passenger counts.")

## [1] "Result: The average trip distance does not vary significantly across
passenger counts."

print("Result: The average trip distance does not vary significantly across
days of the week.")

## [1] "Result: The average trip distance does not vary significantly across
days of the week."

#-----
```