

Analytics for Big Data

DIVVY DATA ANALYSIS

PREPARED BY –

AJAY SRIVATS
SHARADIND PEDDIRAJU
UTKARSH SINGH



TODAY'S TOPICS

OUR DISCUSSION FLOW

Problem Setting

Data Description

Data Analysis and Visualization

Key Takeaways

Business Solutions



PROBLEM SETTING



In this Project, We've conducted a series of exploratory data analysis and Modeling on Chicago Divvy bicycle sharing data. The goal of this project is to :

- **Visualize the bicycle sharing data**
- **Try to find some interesting patterns and relations behind the data**
- **Business Optimization Suggestions**



DATA DESCRIPTION

.....

ATTRIBUTES-

TRIP_ID
STARTTIME
STOPTIME
BIKEID
TRIPDURATION
FROM_STATION_NAME
TO_STATION_NAME
FROM_STATION_ID
TO_STATION_ID
GENDER
BIRTHYEAR
USERTYPE
CUSTOMER
SUBSCRIBER

DATA ANALYSIS AND VIZUALIZATION

DATA PREPROCESSING

Combine bicycle,
station, and weather
data



DATA CLEANING

Exclude trip duration
that is less than 2
minutes and more than
1 hour

Extracting Useful
information from
variables such as Date,
Time, etc..

Removing Variables
with lot of Missing
Values



VIZUALIZATION AND ANALYSIS

Visualized all potentially
correlated variables

Plotted Trip Distribution
using various time
variables like date, year,
month, timestamps.

Plotted Trip Distribution
against Temperature
and Weather data to
gain further interesting
Insights.

KEY TAKEAWAYS

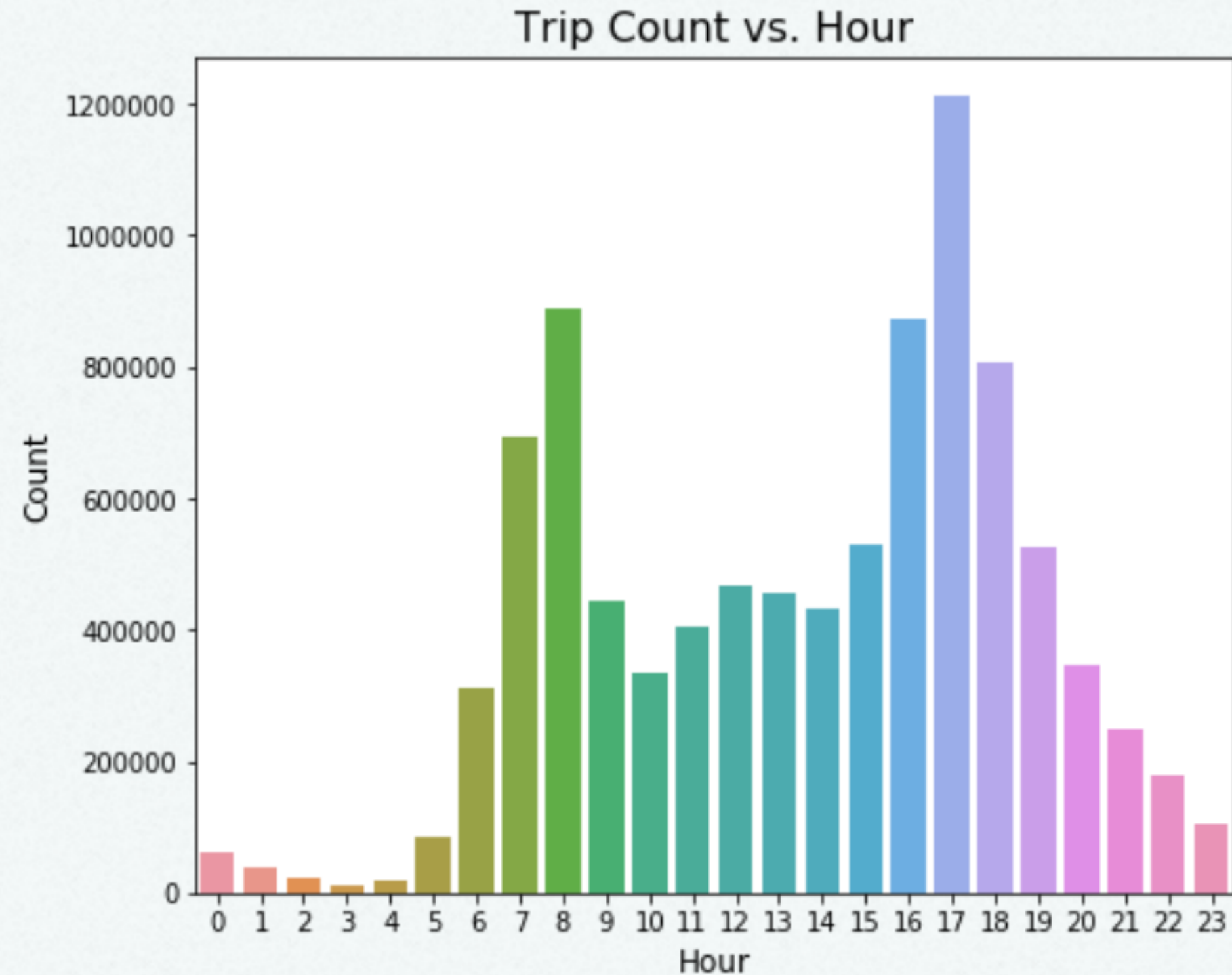


The chart we see here is a bimodal – skyscraper,

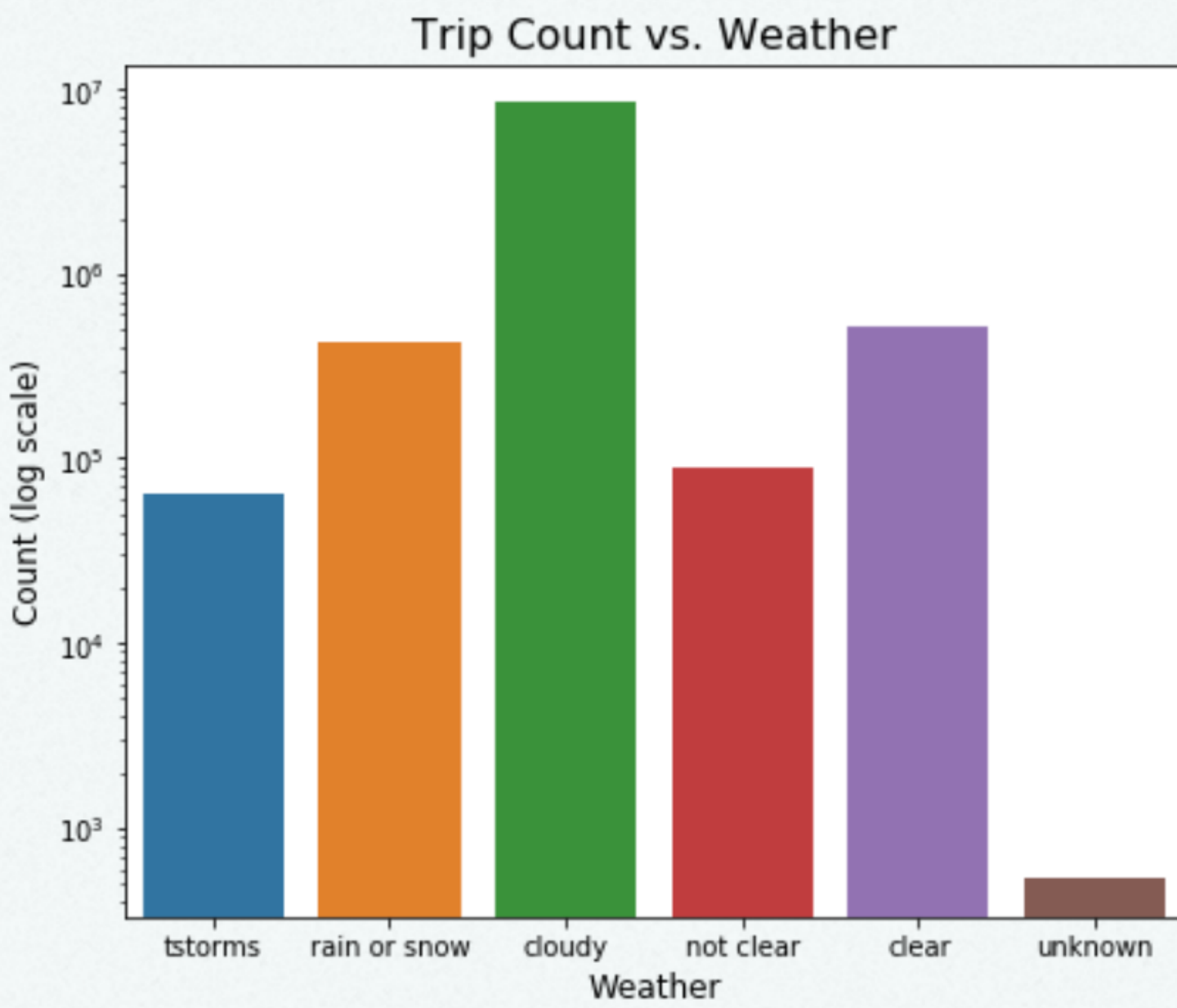
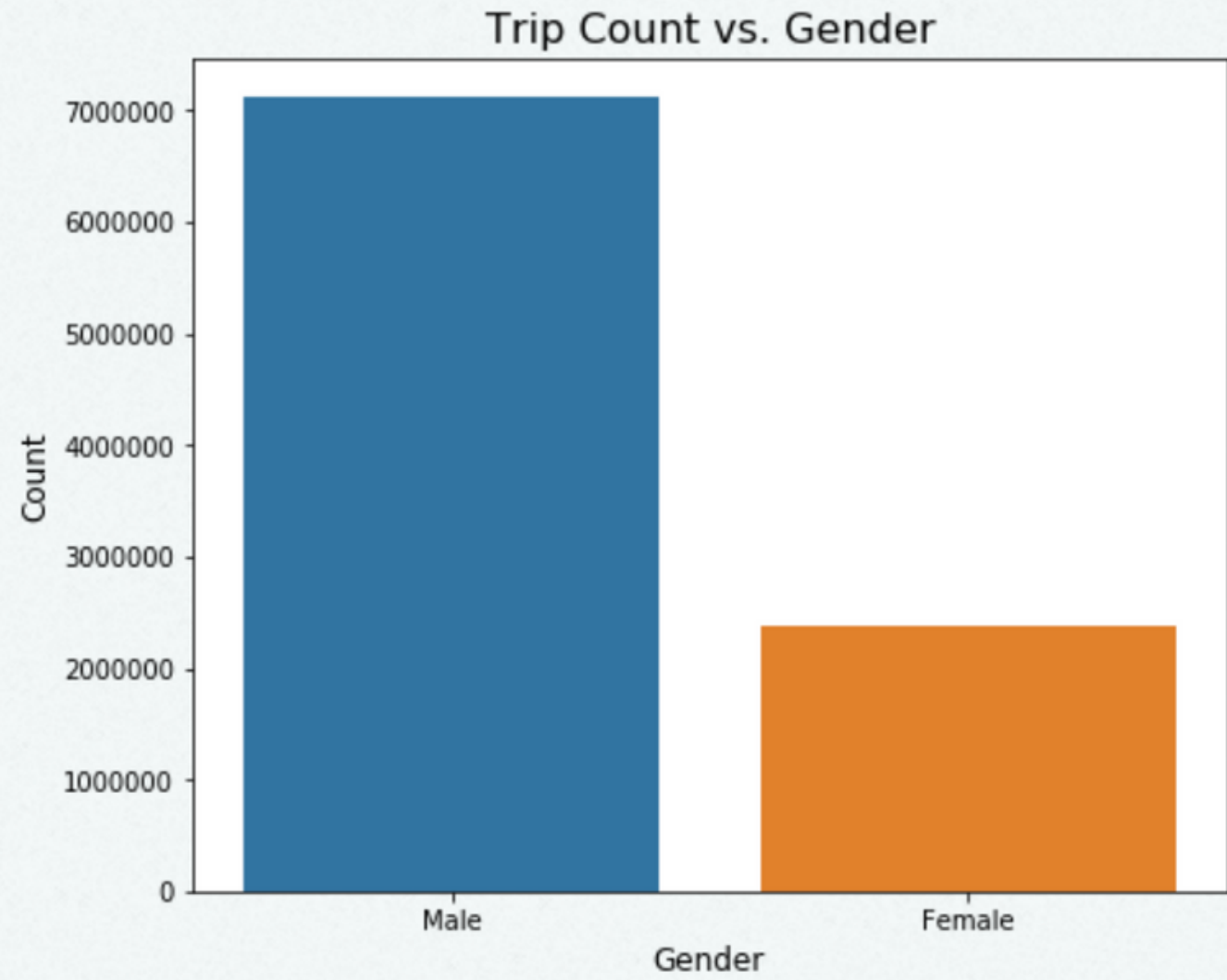
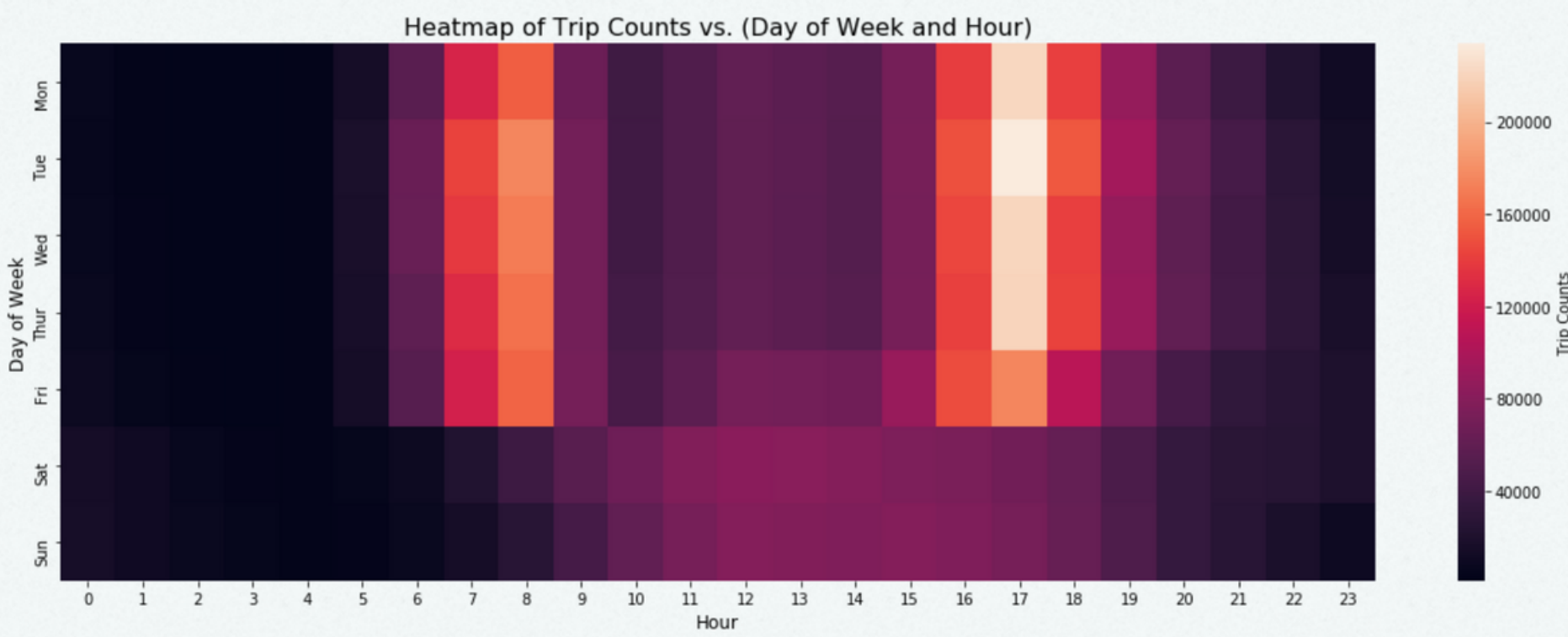
This is an indication that the distribution seen is a mixture of two normal distributions.

Hence we split the time durations under consideration into two charts that split at mid day.

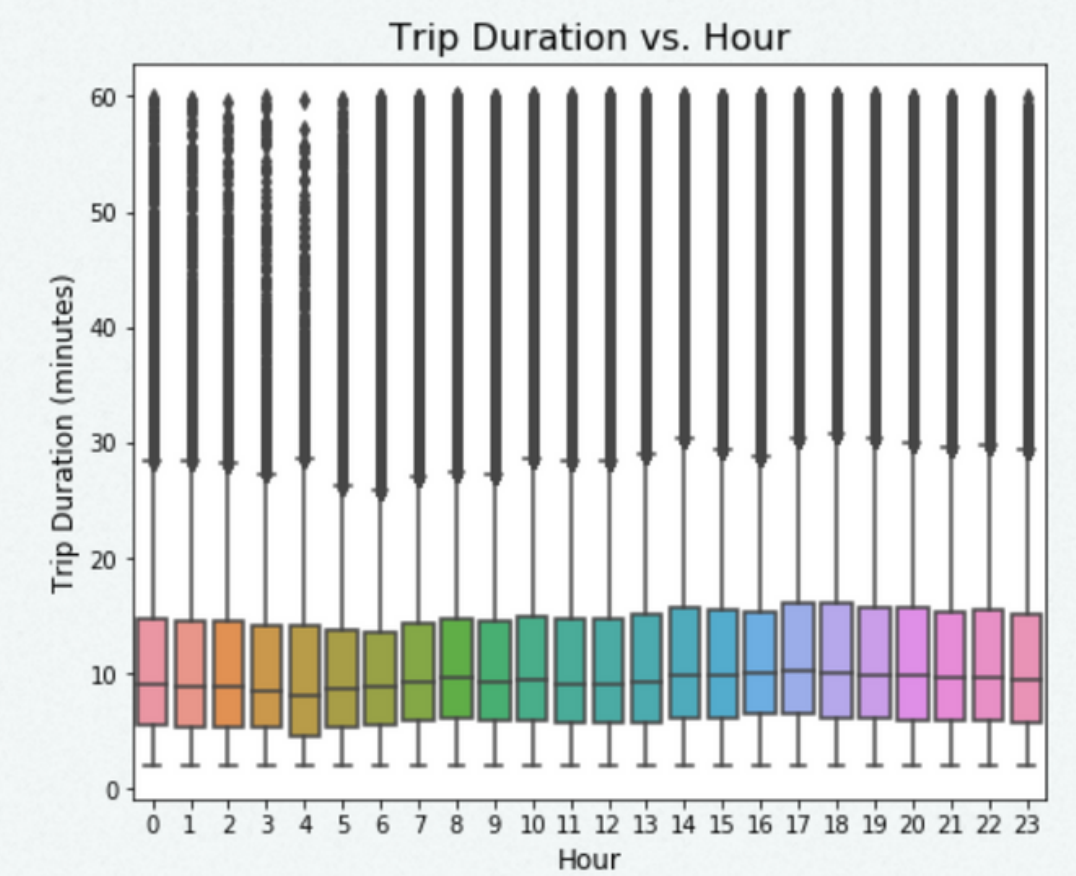
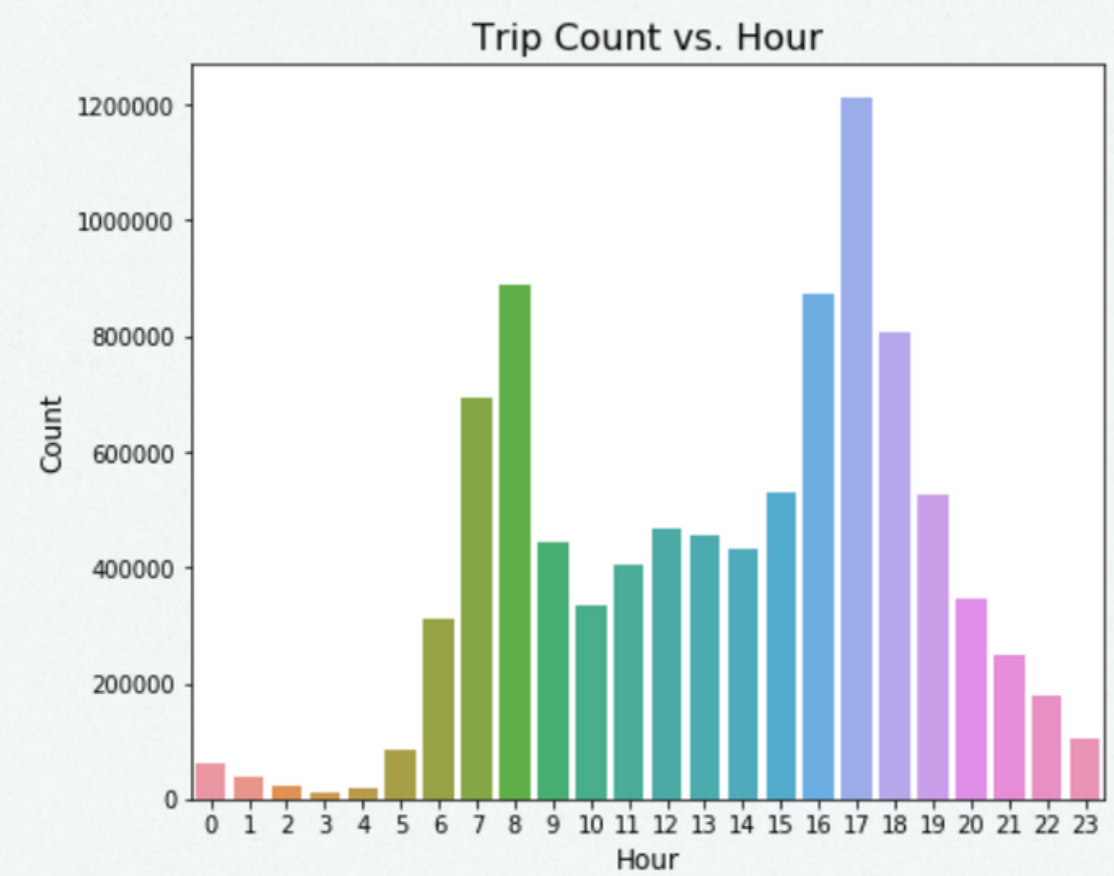
Trip Distribution vs. Hour



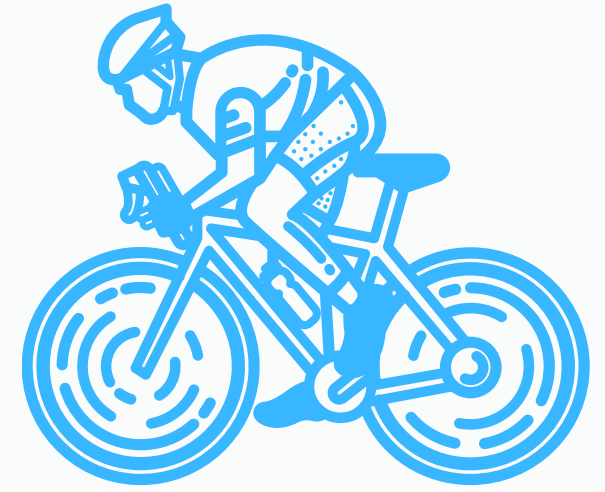
Heat map of Trip Counts vs. Day of Week and Hour



Trip Distribution vs. Hour



BUSINESS SOLUTIONS



- **IN 2018, THERE IS INCREASING DEMAND FOR SHARING BICYCLES. SINCE THE POPULATION OF CHICAGO IS RELATIVE CONSTANT, WE CAN ASSUME THAT PEOPLE TEND TO LIVE HEALTHIER AS TIME GOES ON.**
- **DURING RUSH HOURS (~8:00 AM AND ~5:00 PM), THERE ARE INCREASING DEMAND FOR SHARING BICYCLES DURING WEEKDAYS, BUT ON WEEKENDS, THERE IS NO SUCH CLEAR TREND.**
- **MOST USERS PURCHASED THE ANNUAL MEMBERSHIP. BUT THEIR TRIP DURATIONS ARE RELATIVELY SHORTER COMPARED WITH ORDINARY CUSTOMERS AND DEPENDENT. THOSE WHO PURCHASED 24-HOUR PASS TEND TO HAVE LONGEST TRIP.**
- **PEOPLE USE SHARING BICYCLE MORE FREQUENTLY IN WEEKDAYS THAN WEEKENDS, BUT THE AVERAGE TRIP DURATION IS LONGER IN WEEKENDS THAN WEEKDAYS.**

THANK YOU