

The background of the slide is a light gray gradient, decorated with numerous realistic water droplets of various sizes. Some droplets are large and prominent, while others are small and subtle, scattered across the top and bottom edges of the slide.

# **TITLE:**

# **EXPLORATORY DATA ANALYSIS**

## **– TITANIC DATASET**

NAME: RITHYA REDDY

DATA ANALYST INTERNSHIP – ELEVATE LABS

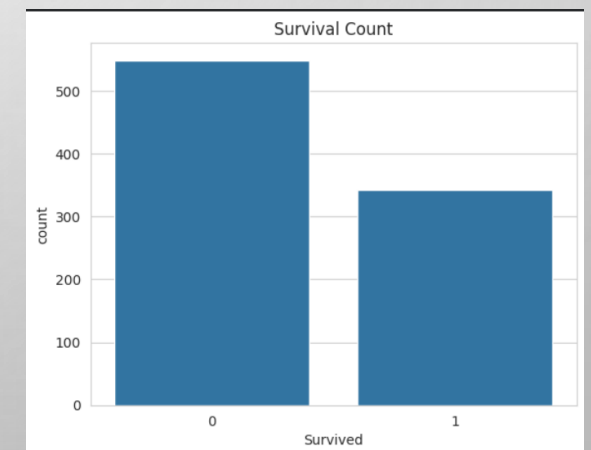
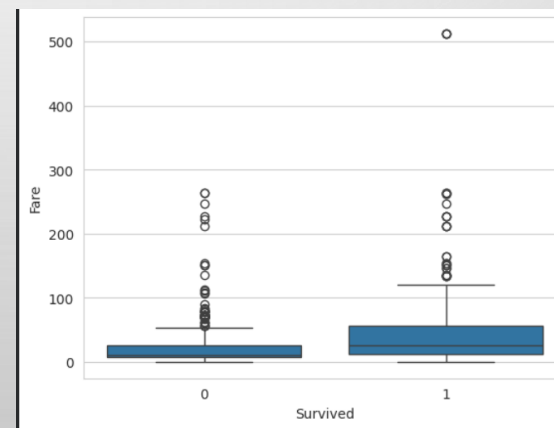
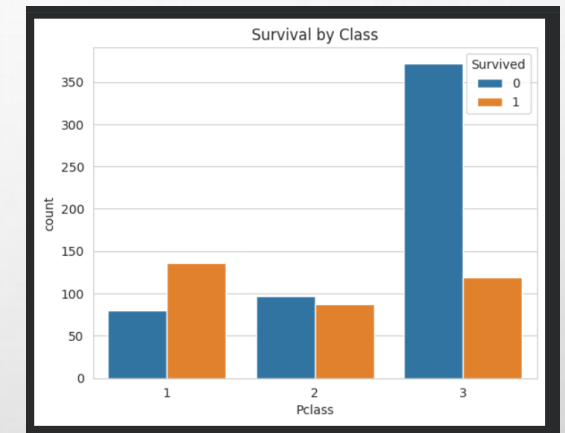
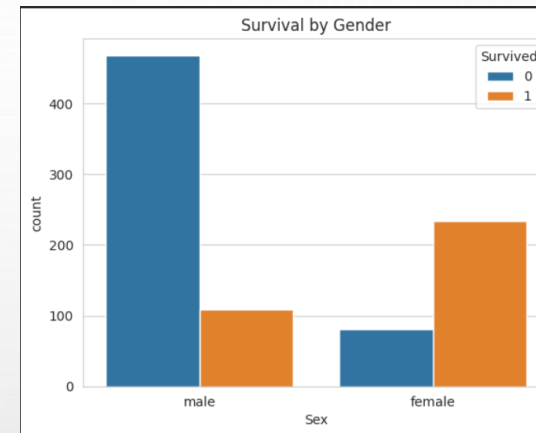
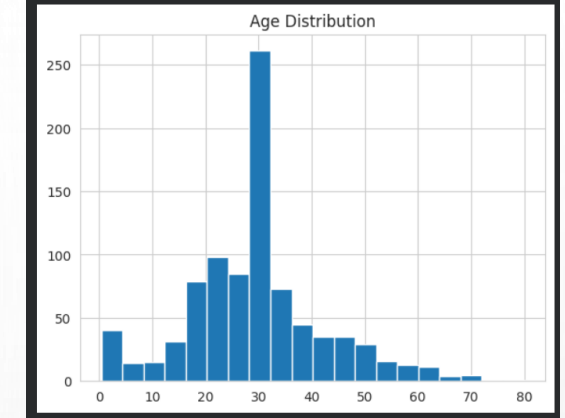
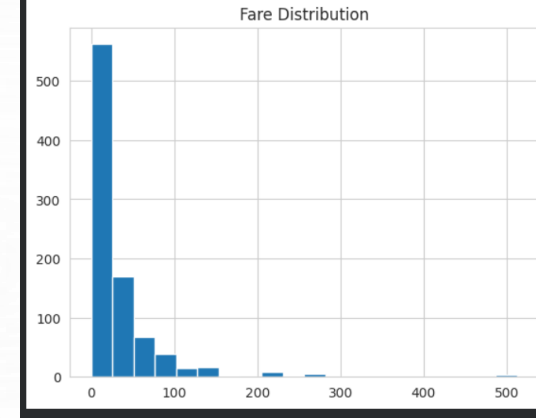
TASK 5

# DATASET OVERVIEW

- CONTENT:
  - TOTAL OBSERVATIONS: 891 PASSENGERS
  - TOTAL FEATURES: 12 (BEFORE CLEANING)
  - TARGET VARIABLE: SURVIVED (0 = NOT SURVIVED, 1 = SURVIVED)
  - DATASET TYPE: BINARY CLASSIFICATION
- **STATISTICAL SUMMARY:**
  - AVERAGE AGE  $\approx$  29.7 YEARS
  - AVERAGE FARE  $\approx$  32.20
  - SURVIVAL RATE  $\approx$  38%

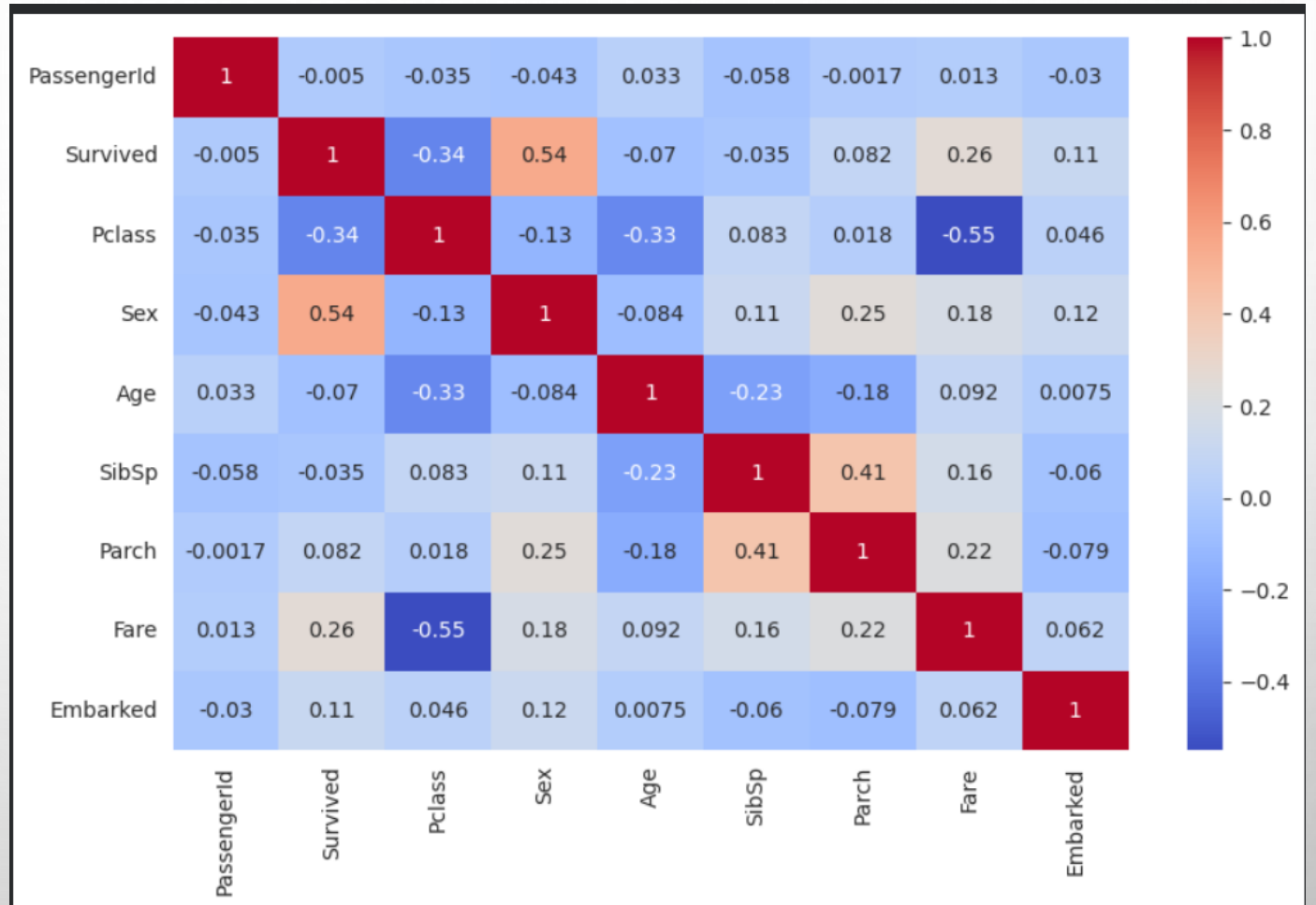
# HISTOGRAMS (UNIVARIATE ANALYSIS)

- SURVIVAL RATE IS AROUND 38%.
- MAJORITY PASSENGERS WERE FROM 3RD CLASS.
- MOST PASSENGERS WERE AGED 20–40 YEARS.
- FARE IS POSITIVELY SKEWED WITH OUTLIERS.
- MOST PASSENGERS TRAVELED ALONE.
- MOST EMBARKED FROM SOUTHAMPTON (S).



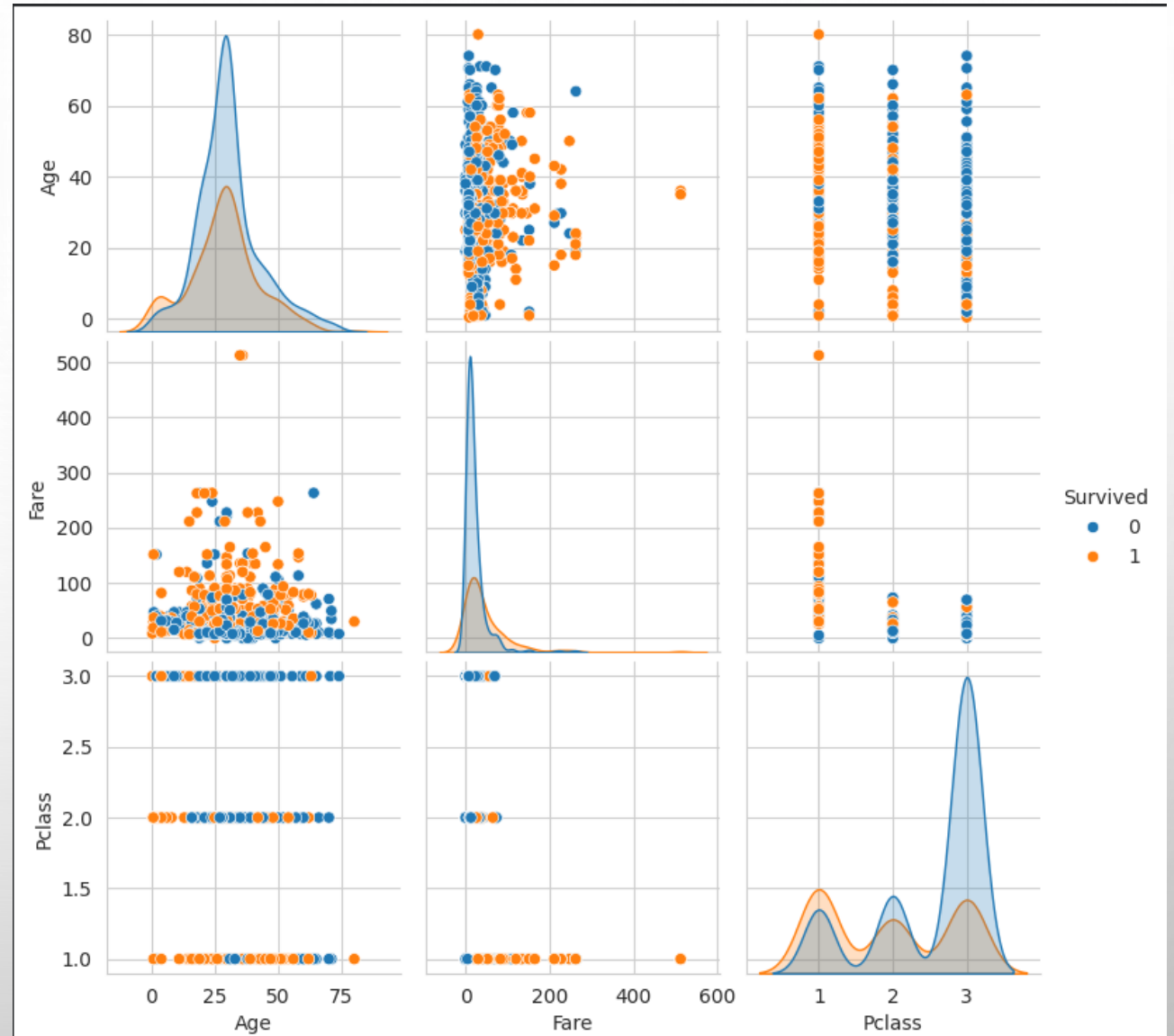
# HEATMAP

- SEX SHOWS STRONG POSITIVE CORRELATION WITH SURVIVAL ( $\sim 0.54$ ).
- PCLASS HAS MODERATE NEGATIVE CORRELATION WITH SURVIVAL ( $\sim -0.34$ ).
- FARE SHOWS POSITIVE CORRELATION WITH SURVIVAL ( $\sim 0.26$ ).
- PCLASS AND FARE ARE STRONGLY NEGATIVELY CORRELATED ( $\sim -0.55$ ).
- AGE HAS WEAK CORRELATION WITH SURVIVAL.



# SCATTER PLOTS

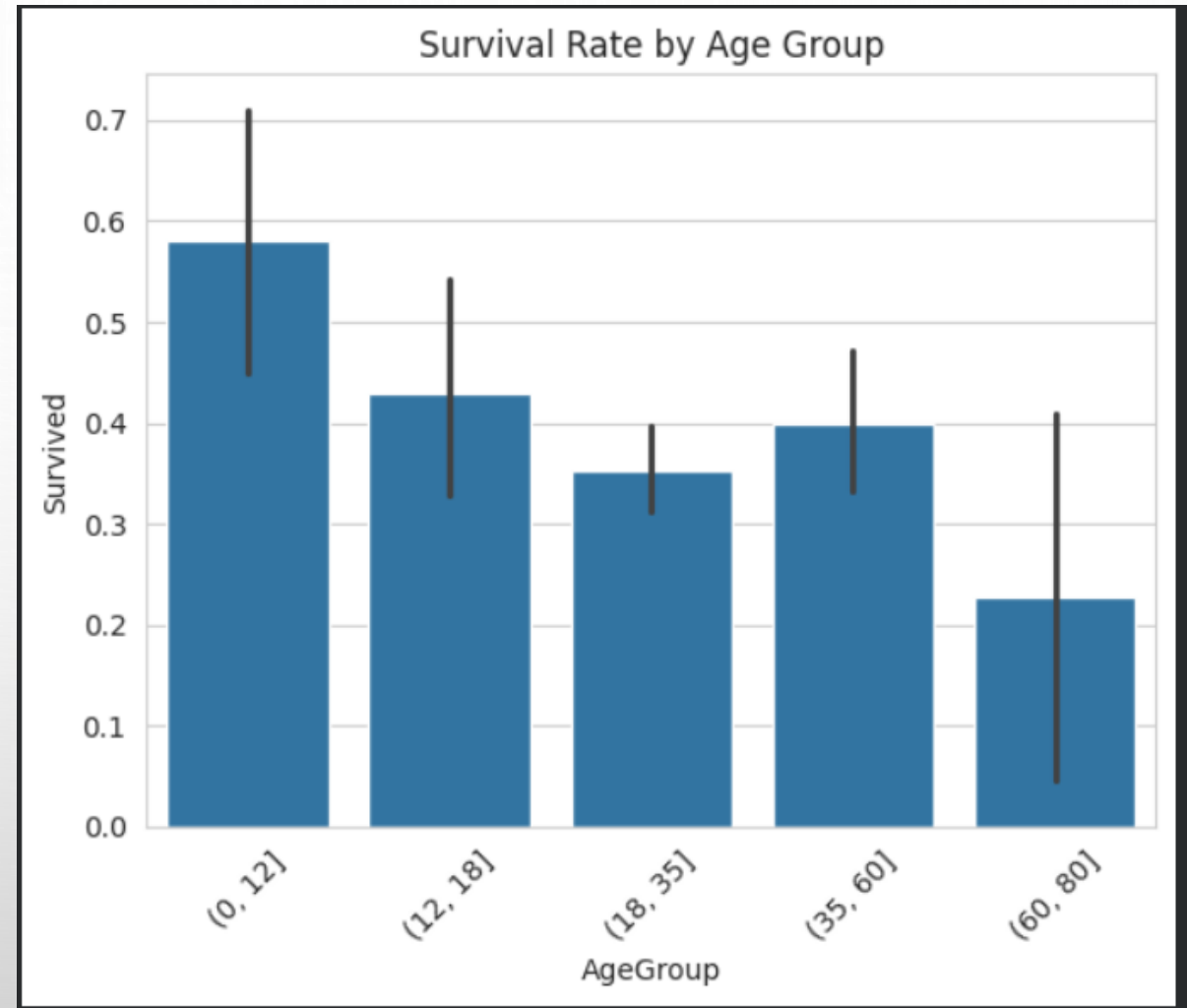
- HIGHER SURVIVAL OBSERVED IN 1ST CLASS PASSENGERS.
- PASSENGERS WHO PAID HIGHER FARE HAD BETTER SURVIVAL CHANCES.
- AGE SHOWS WEAK IMPACT ON SURVIVAL.
- FAMILY SIZE SHOWS MODERATE INFLUENCE.



# SURVIVAL RATE BY AGE GROUP

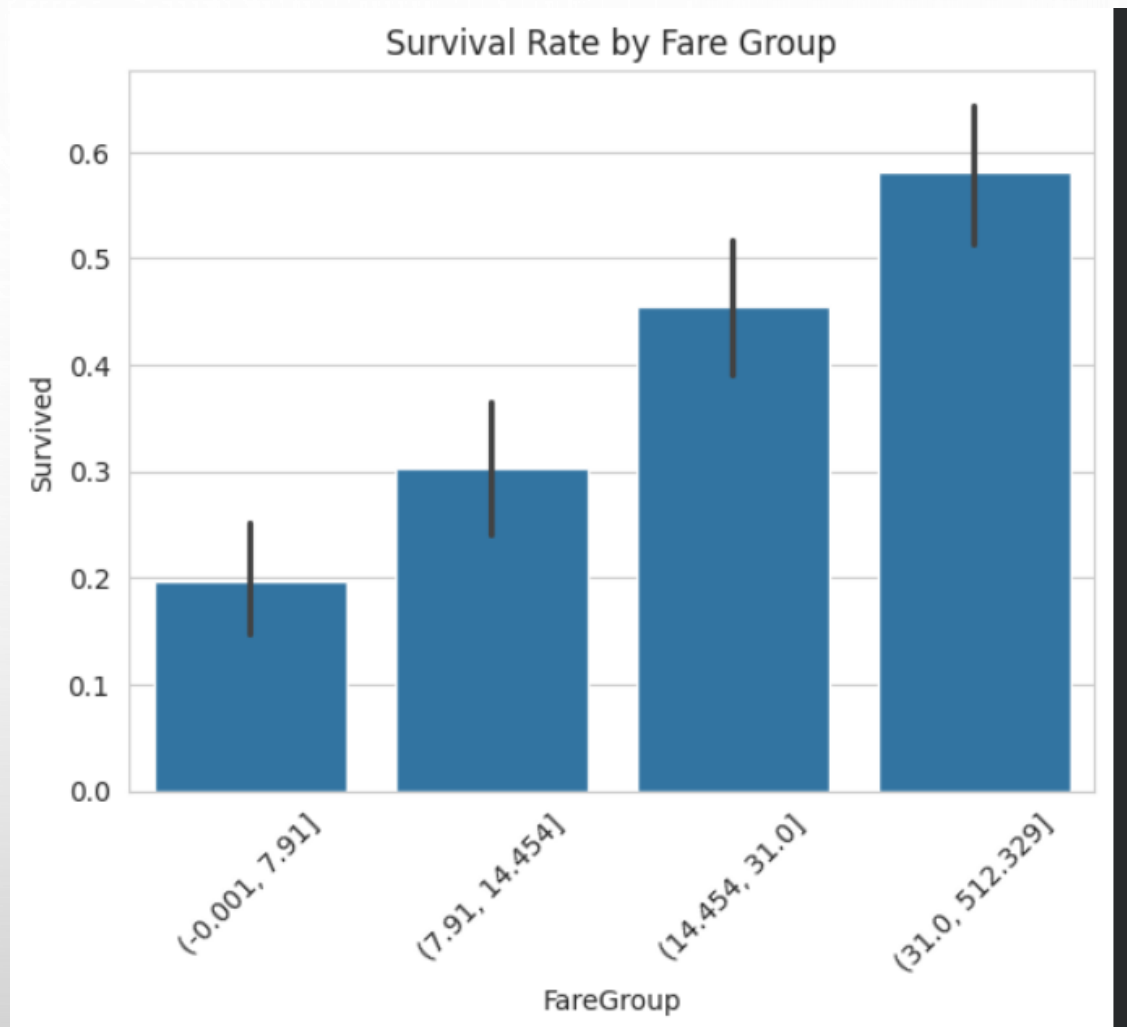
## OBSERVATIONS:

- CHILDREN (0–12 YEARS) SHOW THE HIGHEST SURVIVAL RATE.
- TEENAGERS (12–18 YEARS) HAVE MODERATE SURVIVAL CHANCES.
- ADULTS (18–35 YEARS) SHOW LOWER SURVIVAL COMPARED TO CHILDREN.
- MIDDLE-AGED PASSENGERS (35–60 YEARS) HAVE MODERATE SURVIVAL RATE.
- ELDERLY PASSENGERS (60+ YEARS) SHOW THE LOWEST SURVIVAL RATE.



# SURVIVAL RATE FOR FARE GROUP

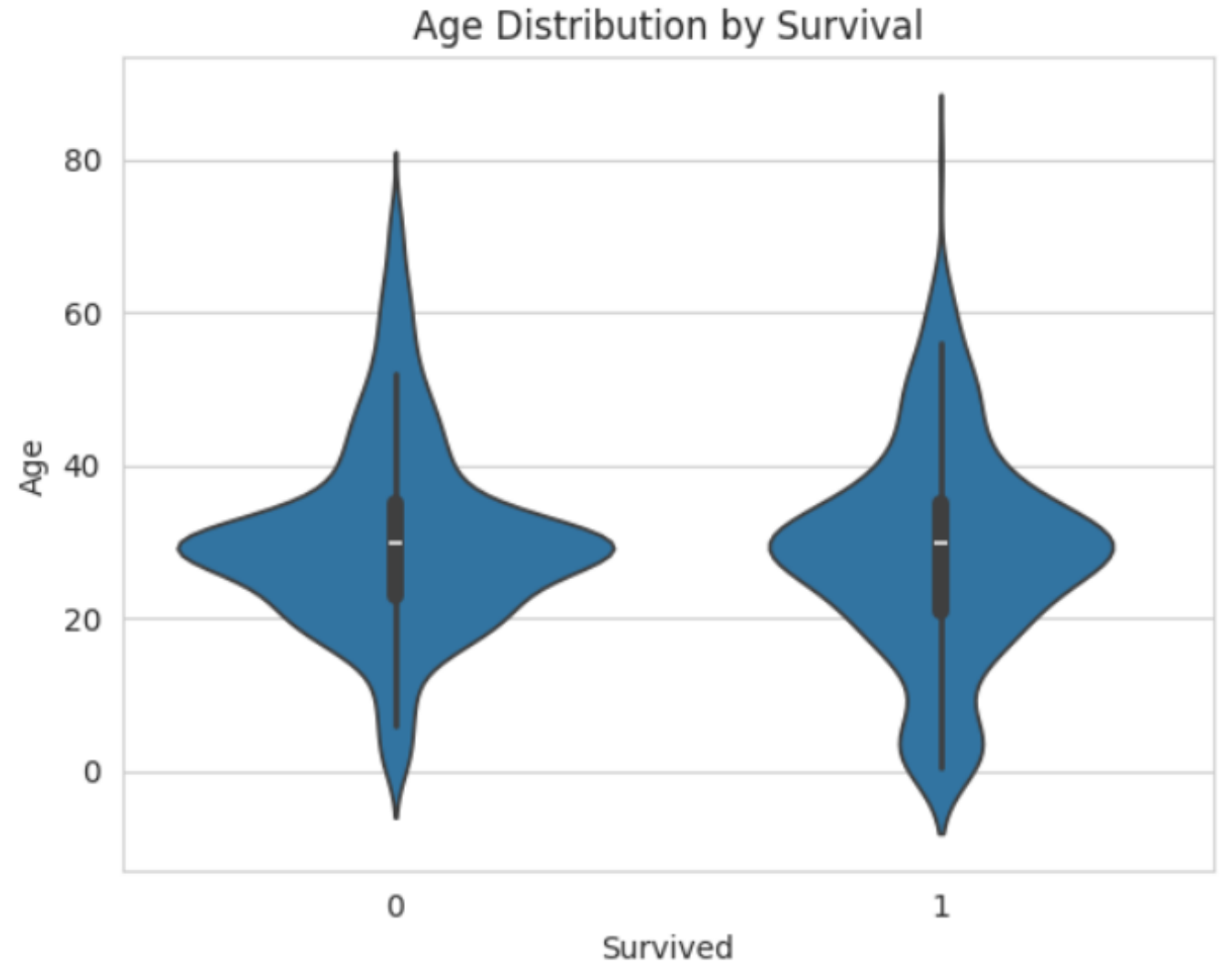
- SURVIVAL RATE INCREASES AS FARE INCREASES.
- PASSENGERS IN THE HIGHEST FARE GROUP SHOW THE HIGHEST SURVIVAL PROBABILITY.
- PASSENGERS IN THE LOWEST FARE GROUP HAVE THE LOWEST SURVIVAL RATE.
- THERE IS A CLEAR UPWARD TREND IN SURVIVAL FROM LOW FARE TO HIGH FARE GROUPS.
- FARE HAS A POSITIVE RELATIONSHIP WITH SURVIVAL.





# VIOLIN PLOT

- SURVIVORS WERE GENERALLY YOUNGER.
- HIGHER SURVIVAL RATE OBSERVED AMONG CHILDREN.
- AGE SHOWS WEAK TO MODERATE INFLUENCE ON SURVIVAL.





The image features a light gray background with a subtle gradient. In the top-left and bottom-right corners, there are several realistic water droplets of varying sizes, some overlapping. A faint, circular, textured pattern is visible in the upper center of the image.

**THANKYOU**