EDGE- Digital Skills for Students
Course: Introduction to Programming with Python
Batch: PP-06
Project Works

Task For Project: Analyze and Visualize on individual Dataset

Datasets: *Titanic: Machine Learning from Disaster*

Question for PP-06-1

1. Count the total number of passengers.
2. Identify the number of unique embarkation towns.
3. Display the first 5 rows of the dataset.
4. Calculate the survival rate of passengers.
5. Check for missing values in the dataset.
6. Plot a histogram for the distribution of passenger ages.

Question for PP-06-2

1. What percentage of passengers survived?
2. Replace missing values in age with the median age.
3. Calculate the average fare for passengers in each class.
4. Display the unique values in the embarked column.
5. Create a bar plot showing the count of passengers in each class.
6. Analyze the survival rate for male and female passengers.

Question for PP-06-3

1. Determine the range of fares paid by passengers.
2. Plot a pie chart showing the proportion of survivors and non-survivors.
3. Find the number of children (passengers under 18 years old).
4. Fill missing values in the deck column with "Unknown".
5. Identify which gender had the highest survival rate.
6. Create a boxplot showing fare distribution across classes.

Question for PP-06-4

1. Find the average age of survivors and non-survivors.
2. Replace missing values in embark_town with the most frequent value.
3. Plot a scatter plot for fare vs. age, coloring points by survival status.
4. Count the number of passengers who traveled alone.
5. Identify which embarkation point had the highest survival rate.
6. Create a bar chart comparing survival rates across gender and class.

Question for PP-06-5

1. How many passengers traveled in each class (pclass)?
2. What is the median fare paid by survivors?
3. Plot a histogram showing the fare distribution.
4. Find the survival rate for passengers who traveled alone.
5. Identify the most common age of passengers.
6. Create a heatmap to show missing values in the dataset.

Question for PP-06-6

1. Find the number of siblings/spouses onboard for passengers who survived.
2. Compare the survival rate for passengers who paid above and below the average fare.
3. Plot a count plot for embarked showing passenger counts for each embarkation point.
4. Identify the age group with the highest survival rate.
5. Create a pair plot for age, fare, and survived.
6. Determine if the fare is correlated with survival using a scatter plot.

Question for PP-06-7

1. Find the number of passengers by embarkation point.
2. Replace missing values in age with the mean age.
3. Analyze the fare distribution using a violin plot by survival status.
4. Count the number of families (passengers with non-zero sibsp or parch).
5. Find the proportion of male and female passengers in each class.
6. Plot a bar chart comparing survival rates across embarkation points.

Question for PP-06-8

1. Identify the deck with the most passengers.
2. Calculate the average age of passengers in each class.
3. Create a density plot for fare by survival status.
4. Find the number of passengers who embarked from Southampton.

5.   Identify the survival rate for passengers in the "C" deck.
6.   Generate a bar plot for the count of passengers by class and survival.

Question for PP-06-9

1.   Find the total fare paid by passengers in each class.
2.   Plot a pie chart showing the proportion of passengers who traveled alone.
3.   Determine the survival rate for passengers with no family members onboard.
4.   Create a swarm plot showing survival based on age.
5.   Count the number of children who survived.
6.   Analyze the distribution of ages among survivors using a KDE plot.

Question for PP-06-10

1.   Find the maximum fare paid by a survivor.
2.   Identify which class had the highest percentage of survivors.
3.   Create a stacked bar plot of survival by gender and class.
4.   Find the survival rate for passengers aged 60 and above.
5.   Count the number of unique ticket numbers in the dataset.
6.   Create a box plot of fare by survival status.

Question for PP-06-11

1.   Identify the embarkation point with the lowest survival rate.
2.   Plot a bar chart for the number of passengers grouped by who (man, woman,
child).
3.   Find the average age of female passengers.
4.   Determine the survival rate for passengers with a fare above the 75th percentile.
5.   Count the number of adult male survivors.
6.   Create a scatter plot of age vs. fare for each class.

Question for PP-06-12

1.   Analyze the number of passengers in each category of who.
2.   Calculate the median age of passengers in the first class.
3.   Find the survival rate for passengers in the lowest class.
4.   Replace missing embarked values with the most common value.
5.   Create a histogram of fares paid by survivors and non-survivors.
6.   Compare survival rates for passengers traveling with family and those traveling
alone.

Question for PP-06-13

1. Find the proportion of passengers in each class.
2. Count the number of passengers in each deck.
3. Plot a bar chart showing survival rates across different decks.
4. Determine the average fare for each embarkation point.
5. Calculate the survival rate for passengers with family size greater than 3.
6. Create a histogram of ages grouped by survival status.

Question for PP-06-14

1. Calculate the median age of passengers who embarked at Cherbourg.
2. Plot a pie chart for the proportion of passengers in each age group (children, young adults, adults, seniors).
3. Analyze the survival rate for passengers with sibsp = 0 and parch = 0.
4. Find the maximum fare paid by passengers who did not survive.
5. Create a scatter plot for fare vs. age with different markers for each embarkation point.
6. Count the number of passengers in each class by gender.

Question for PP-06-15

1. Find the youngest survivor and the oldest survivor.
2. Create a bar chart showing the distribution of passengers by who.
3. Analyze the survival rate for passengers with more than 2 family members onboard.
4. Replace missing deck values with the mode of the deck column.
5. Create a violin plot of fare distribution for survivors and non-survivors by class.
6. Find the total number of male passengers who embarked from Queenstown.

Question for PP-06-16

1. Determine the survival rate of passengers aged under 5.
2. Plot a count plot of embarked with survival as the hue.
3. Calculate the average fare paid by passengers aged above 50.
4. Compare survival rates for passengers with different family sizes using a bar plot.
5. Identify the proportion of passengers in each who category who survived.
6. Create a KDE plot for the age distribution of survivors.

Question for PP-06-17

1. Identify the passenger who paid the highest fare and find their survival status.
2. Calculate the average age of passengers in each gender and class combination.
3. Create a scatter plot of fare vs. survival colored by embarkation point.

4.   Compare survival rates for passengers in different who categories.
5.   Create a histogram showing the fare distribution by survival status.
6.   Analyze the survival rate for passengers traveling in large families (family size > 4).

Question for PP-06-18

1.   Find the survival rate for passengers with a fare below 10.
2.   Create a count plot for gender with survival as the hue.
3.   Determine the survival rate for passengers from each embarkation point.
4.   Analyze the age distribution of passengers in each class using a box plot.
5.   Replace missing age values with a random value sampled from the age distribution.
6.   Plot a pie chart showing the proportion of passengers in each deck.

Question for PP-06-19

1.   Calculate the survival rate of passengers with no siblings or spouses onboard.
2.   Find the average fare for survivors in the second class.
3.   Compare the age distribution for passengers who traveled alone vs. with family using a violin plot.
4.   Determine the proportion of passengers aged below 18 who survived.
5.   Analyze the correlation between fare and survived.
6.   Create a grouped bar chart comparing survival rates by gender and embarkation point.

Question for PP-06-20

1.   Identify which deck had the highest survival rate.
2.   Plot a histogram showing the distribution of family sizes.
3.   Compare the survival rates for male and female passengers in the third class.
4.   Replace missing embark_town values with "Unknown".
5.   Calculate the percentage of survivors who paid above the median fare.
6.   Create a swarm plot for fare vs. class colored by survival status.

Question for PP-06-21

1.   Analyze the survival rate for passengers with family size 1.
2.   Compare survival rates for passengers who embarked from different ports.
3.   Plot a bar chart for survival rates across age groups (children, adults, seniors).
4.   Calculate the average age for passengers who embarked at Southampton.
5.   Determine the survival rate for female passengers traveling alone.
6.   Create a heatmap showing survival rates across combinations of class and gender.

Question for PP-06-22

1. Find the total number of survivors from the first class.
2. Plot a scatter plot for fare vs. class colored by survival status.
3. Calculate the proportion of passengers who embarked from each port.
4. Analyze the survival rates for passengers in each who category.
5. Determine the average fare for passengers traveling in large families.
6. Create a density plot for the fare distribution of survivors.

Question for PP-06-23

1. Analyze the survival rate for male children.
2. Plot a count plot for the alone feature with survival as the hue.
3. Calculate the average age of passengers in each deck.
4. Find the youngest passenger in each class and determine their survival status.
5. Create a violin plot for age distribution by class and survival status.
6. Analyze the survival rate for passengers aged between 30 and 40.

Question for PP-06-24

1. Identify the passengers with the highest and lowest fares and their survival statuses.
2. Calculate the survival rate for female passengers aged above 50.
3. Create a scatter plot of fare vs. age with survival as a color gradient.
4. Compare survival rates for passengers with no family members vs. small families (family size ≤ 3).
5. Replace missing deck values with a new category "Missing".
6. Plot a pie chart for the proportion of passengers who traveled alone.

Question for PP-06-25

1. Find the survival rate for passengers who embarked from Queenstown.
2. Compare the survival rates for passengers in the lowest fare quartile.
3. Create a bar chart showing the survival rates for each age group.
4. Analyze the survival rate for adult males in the second class.
5. Create a heatmap for survival rates based on who and embarked.
6. Generate a swarm plot for fare vs. survival status grouped by class.