

# Data Science - Assignment 3

---

## Submission Requirement

- Submit **one** `.ipynb` **notebook**
- Use:
  - Code cells for implementation
  - Markdown cells for explanations and answers
- Your notebook should be well-organized and readable

## Tasks

You will use the Titanic dataset. You can download code from Kaggle: <https://www.kaggle.com/datasets/yasserh/titanic-dataset> or from course website: [https://xiaochenli-w.github.io/CSC-405-605-705-Data-Science/upload%20slides/datasets\\_used\\_pandas2.zip](https://xiaochenli-w.github.io/CSC-405-605-705-Data-Science/upload%20slides/datasets_used_pandas2.zip)

## Requirements

- Answer **ALL 5 questions**
- For each question:
  - Create 1–2 different types of visualizations
  - Use different chart types across the assignment
  - Provide clear explanations and insights
- Include code (Python + matplotlib / seaborn / plotly, etc.)
- Your submission should include:
  - At least:
    - 1 heatmap
    - 1 density/violin/box plot
    - 1 scatter plot
    - 2 basic charts (bar/stacked/pie/line/bubble/etc.)
  - Clean labels, titles, legends
  - Readable formatting

### Question 1: Survival vs Passenger Class & Gender (Code + Chart(s) + Analysis: 20 points)

Is survival rate related to passenger class? Does gender change this relationship?

Suggested attributes: `Survived`, `Pclass`, `Sex`

Write down your analysis:

- Which class had the highest survival rate?
- Did gender amplify or reduce class differences?

### Question 2: Age Distribution and Survival Patterns (Code + Chart(s) + Analysis: 20 points)

How does age affect survival? Are children or elderly passengers more likely to survive?

Suggested attributes: Age, Survived, Sex

Write down your analysis:

- Distribution differences between survivors and non-survivors
- Any noticeable trends across age groups

**Question 3: Family Size and Survival (Code + Chart(s) + Analysis: 20 points)**

Did traveling with family increase or decrease survival chances?

Suggested attributes: Sibsp, Parch, Survived

(You may create a new feature: FamilySize = Sibsp + Parch)

Write down your analysis:

- Is there an optimal family size for survival?
- Do people traveling alone have higher/lower chances?

**Question 4: Fare, Class, and Survival Relationship (Code + Chart(s) + Analysis: 20 points)**

Does ticket fare correlate with survival? How does this interact with passenger class?

Suggested attributes: Fare, Pclass, Survived

Write down your analysis:

- Are higher-paying passengers more likely to survive?
- Does fare reflect class differences clearly?

**Question 5: Embarkation Port and Survival Behavior (Code + Chart(s) + Analysis: 20 points)**

Did the port of embarkation influence survival rates?

Suggested attributes: Embarked, Survived, Pclass, Sex

Write down your analysis:

- Which port had the highest survival rate?
- Are there hidden confounding factors (class, gender)?