

*Worksheet 2 - Sep 19, 2025*

Data Practice with Titanic Dataset

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Goal:** Learn how to load and explore a real dataset in Google Colab.

**Step 1 – Upload the Dataset**

1. Open your Colab notebook.
2. On the **left sidebar**, click the **folder icon** 🗂️.
3. Click the **Upload** button ⬆️.
4. Select **mini\_titanic.csv** from your Downloads folder.
5. You should now see mini\_titanic.csv listed in the file explorer.

**Step 2 – Load the Data**

import pandas as pd

# Load the Titanic dataset

data = pd.read\_csv("mini\_titanic.csv")

# Show the first 5 rows

data.head()

**Step 3 – Exercises**

**Exercise 1 – First Look**

* Print the first 5 rows using .head().
* How many rows and columns are in the dataset?

**Exercise 2 – Survival Counts**

* How many passengers survived?
* How many did not survive?  
  (Hint: use data["Survived"].value\_counts())

**Exercise 3 – Who Survived More?**

* Compare survival between males and females.  
  (Hint: use data.groupby("Gender")["Survived"].value\_counts())

**Bonus:**

* How many kids (Age < 18) survived?
* Which class had the most survivors?