


Exploratory Data Analysis **SANDBOX**

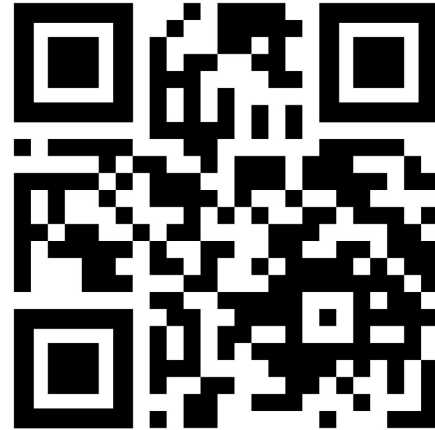


Wednesday, November 6th
The Diamond, Workroom 2
16:30 - 18:30

Before we start..



Consent Form



Accessibility Form

SOCIAL ALERT!



ESCAPE

ROOM 

THURSDAY, NOVEMBER 7th

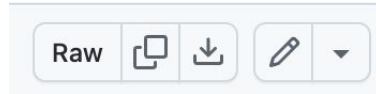
**GET TICKETS
NOW!!!**

Search For

> **Sheffield University Data Science Society > 2024-2025**

> **Sandbox Sessions > EDA > EDA.ipynb**

1. Download the file



2. Do one of the following:

Save it on the same folder you use for VSCode

OR

Upload to your Google Drive for Colab

Open DataVis Guide And VSCode Guide
from our website



[Pandas Guide](#)



[VS Code Guide](#)

<https://uosdss.wordpress.com>



Setup

**Install / Import Library,
Import Dataset**



#Challenge 1

Load Data

Summarize Your Data

how many passengers are there?



#Challenge 2

Separate the data into numerical and categorical lists





#Challenge 3

- Find the percentage of men who survived
 - Did more women or men survive?
 - How old was the youngest passenger?
- 



Challenge 4

Visualizing categorical columns (sex ,
survived)



#Challenge 5

Via visualisation or mathematical calculation
find the following:

Find out if the age column is normally
distributed.

Find out if the fates column is normally
distributed.

#Challenge 6

Visualise the correlation between all columns using sub graphs heat map Find if there is correlation between surviving and any other column

**THANK
YOU !**

We hope to see you next week :)

