

Data Analysis

TITANIC PROJECT



AND



27, August 2023



CONTENT

01

Project
Background

02

Description of
Columns and
Headers

03

Query
language and
Screenshots

04

Visualization
Screenshots

05

Insights
and
Recommendat
ions

06

Conclusion

07

Team

08

Strategy

PROJECT BACKGROUND

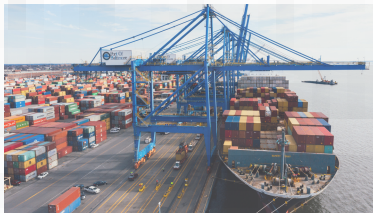
- This project is about THE TITANIC SHIP
- The Titanic is about a ship that conveyed passengers onboard whose personal details are stored in the database.
- This ship conveying passengers from different works of life, gender and family background was wrecked leaving many dead and few survivors.

The objective of the project is to extract specific data sets using SQL and visualizing them using Power BI to provide insights and recommendations.



AGE

LENGTH OF TIME
A PERSON
HAS LIVED



PORT of EMBARKATION

FROM WHICH THE
SHIP DEPARTS OR
PASSENGER
BOARDS THE SHIP.



PASSENGER ID

UNIQUE IDENTITY
OF THE PASSENGER



FARE PAID

THE TICKET PRICE



CABIN

IS A CLASS OF
ACCOMODATION ON
A PASSENGER SHIP



NAME

IS A COLLECTION OF
DATA PERTAINNING
TO AN INDIVIDUAL
TRAVELLER OR
GROUP

DESCRIPTION OF COLUMNS
& HEADERS

1. Calculate the average age of passengers who survived and those who did not

```
select AVG([Age]) AS 'Avg age of survivors of not',[Survived]
from[dbo].[train_clean (3)]
group by[Survived]
```

90 %



Results



Messages

	Avg age of survivors of not	Survived
1	30.2495446265938	0
2	28.1540058479149	1

1. Calculate the average age of passengers who survived and those who did not

```
select AVG([Age]) AS 'Avg age of survivors of not',[Survived]  
from[dbo].[train_clean (3)]  
group by[Survived]
```

90 %



Results



Messages

	Avg age of survivors of not	Survived
1	30.2495446265938	0
2	28.1540058479149	1

2. Find the top 5 most common titles (e.g., Mr., Mrs., Miss) among the passengers

```
select top 5 ([Title]), count ([PassengerId]) AS 'most common title'
from [dbo].[train_clean (3)]
group by [Title]
order by count ([PassengerId]) DESC
```

90 %



Results



Messages

	Title	most common title
1	Mr	525
2	Miss	185
3	Mrs	128
4	Master	40
5	Dr	7

3. Calculate the average fare paid by passengers in each passenger class (Pclass)

```
select AVG ([Fare]) AS 'AVG fare paid by pax', [Pclass]
from [dbo].[train_clean (3)]
group by [Pclass]
order by [Pclass] DESC
```

90 %

Results

Messages

	AVG fare paid by pax	Pclass
1	13.6755502102574	3
2	20.6621831810993	2
3	84.154687528257	1

A compilation of passengers with the highest fare.

4. Retrieve passengers with the highest and lowest fare

=

```
select*
from [dbo].[train_clean (3)]
where [Fare] = (select max ([Fare]) from [dbo].[train_clean (3)])
```

90 %

Results

Messages

	Age	Cabin	Embarked	Fare	Name	Parch	PassengerId	Pclass	Sex
1	35	NULL	C	512.329223632813	Ward, Miss. Anna	0	259	1	female
2	36	B51 B53 B55	C	512.329223632813	Cardeza, Mr. Thomas Drake Martinez	1	680	1	male
3	35	B101	C	512.329223632813	Lesurer, Mr. Gustave J	0	738	1	male

A compilation of passengers with the lowest fare.

4. Retrieve passengers with the highest and lowest fare

```
select*
from [dbo].[train_clean (3)]
where [Fare] = (select min ([Fare]) from [dbo].[train_clean (3)])
```

90 %

Results

Messages

	Age	Cabin	Embarked	Fare	Name	Parch	PassengerId	Pclass	Sex	SibSp	Survived	Ticket	Title	Family_Size
1	36	NULL	S	0	Leonard, Mr. Lionel	0	180	3	male	0	0	LINE	Mr	0
2	40	B94	S	0	Hamison, Mr. William	0	264	1	male	0	0	112059	Mr	0
3	25	NULL	S	0	Tomquist, Mr. William Henry	0	272	3	male	0	1	LINE	Mr	0
4	30	NULL	S	0	Parkes, Mr. Francis "Frank"	0	278	2	male	0	0	239853	Mr	0
5	19	NULL	S	0	Johnson, Mr. William Cahoon Jr	0	303	3	male	0	0	LINE	Mr	0
6	30	NULL	S	0	Cunningham, Mr. Alfred Fleming	0	414	2	male	0	0	239853	Mr	0
7	30	NULL	S	0	Campbell, Mr. William	0	467	2	male	0	0	239853	Mr	0
8	30	NULL	S	0	Frost, Mr. Anthony Wood "Arc...	0	482	2	male	0	0	239854	Mr	0
9	49	NULL	S	0	Johnson, Mr. Alfred	0	598	3	male	0	0	LINE	Mr	0
10	30	NULL	S	0	Parr, Mr. William Henry Marsh	0	634	1	male	0	0	112052	Mr	0
11	30	NULL	S	0	Watson, Mr. Ennis Hastings	0	675	2	male	0	0	239856	Mr	0
12	30	NULL	S	0	Knight, Mr. Robert J	0	733	2	male	0	0	239855	Mr	0
13	39	A36	S	0	Andrews, Mr. Thomas Jr	0	807	1	male	0	0	112050	Mr	0
14	30	B102	S	0	Fry, Mr. Richard	0	816	1	male	0	0	112058	Mr	0
15	38	NULL	S	0	Reuchlin, Jonkheer. John Geo...	0	823	1	male	0	0	19972	Mr	0

NO. of Rows - 15

5. Find the survival rate for male and female passengers

```
select [Sex], sum([Survived]) AS 'Total survived', count(*) AS 'Total passengers', sum([Survived])*100.0/count(*) AS 'Survival rate'
from [dbo].[train_clean (3)]
group by [Sex]
```

90 %

Results Messages

	Sex	Total survived	Total passengers	Survival rate
1	male	109	577	18.8908145580589
2	female	233	314	74.203821656051

6. Calculate the average family size (SibSp + Parch + 1) for passengers who survived and those who did not

```
=select AVG([SibSp]+[Parch]+1) AS 'Family size',[Survived]
from[dbo].[train_clean (3)]
group by[Survived]
```

100 %

	Family size	Survived
1	1	0
2	1	1

7. Retrieve the names of passengers whose age is missing (NULL)

```
= select [Age],[Name]
from [dbo].[train_clean (3)]
where [Age] IS NULL
```

90 %



Results



Messages

Age	Name
-----	------

8. Find the most common port of embarkation (Embarked) and its corresponding number of passengers

```
select [Embarked], count ([PassengerId]) AS 'Most embarked port'  
from [dbo].[train_clean (3)]  
group by [Embarked]  
order by count([PassengerId]) DESC
```

90 %



Results



Messages

	Embarked	Most embarked port
1	S	645
2	C	169
3	Q	77

9. Calculate the average age for passengers in each passenger class and gender combination

```
select AVG([Age]) AS 'AVG age of pax and gender',[Pclass],[Sex]
from[dbo].[train_clean (3)]
group by[Pclass],[Sex]
order by [Pclass]
```

90 %



Results



Messages

	AVG age of pax and gender	Pclass	Sex
1	34.5531914893617	1	female
2	39.4952459017761	1	male
3	28.5460526315789	2	female
4	30.678981481327	2	male
5	22.6666666666667	3	female
6	27.1539481267634	3	male

10. Retrieve passengers who were traveling alone (Family_Size = 1) and did not survive:

```
=select [Family_Size],[Survived]
from[dbo].[train_clean (3)]
where[Family_Size]=1 and [Survived] = 0
```

90 %

Results Messages

	Family_Size	Survived
1	1	0
2	1	0
3	1	0
4	1	0
5	1	0
6	1	0
7	1	0
8	1	0
9	1	0
10	1	0
11	1	0
12	1	0
13	1	0
14	1	0
15	1	0
16	1	0
17	1	0
18	1	0
19	1	0
20	1	0
21	1	0

VISUALIZATION VIA POWER BI

TITANIC DASHBOARD

Total Number of Passengers Onboard

891

Total Amount Paid by Passengers Onboard

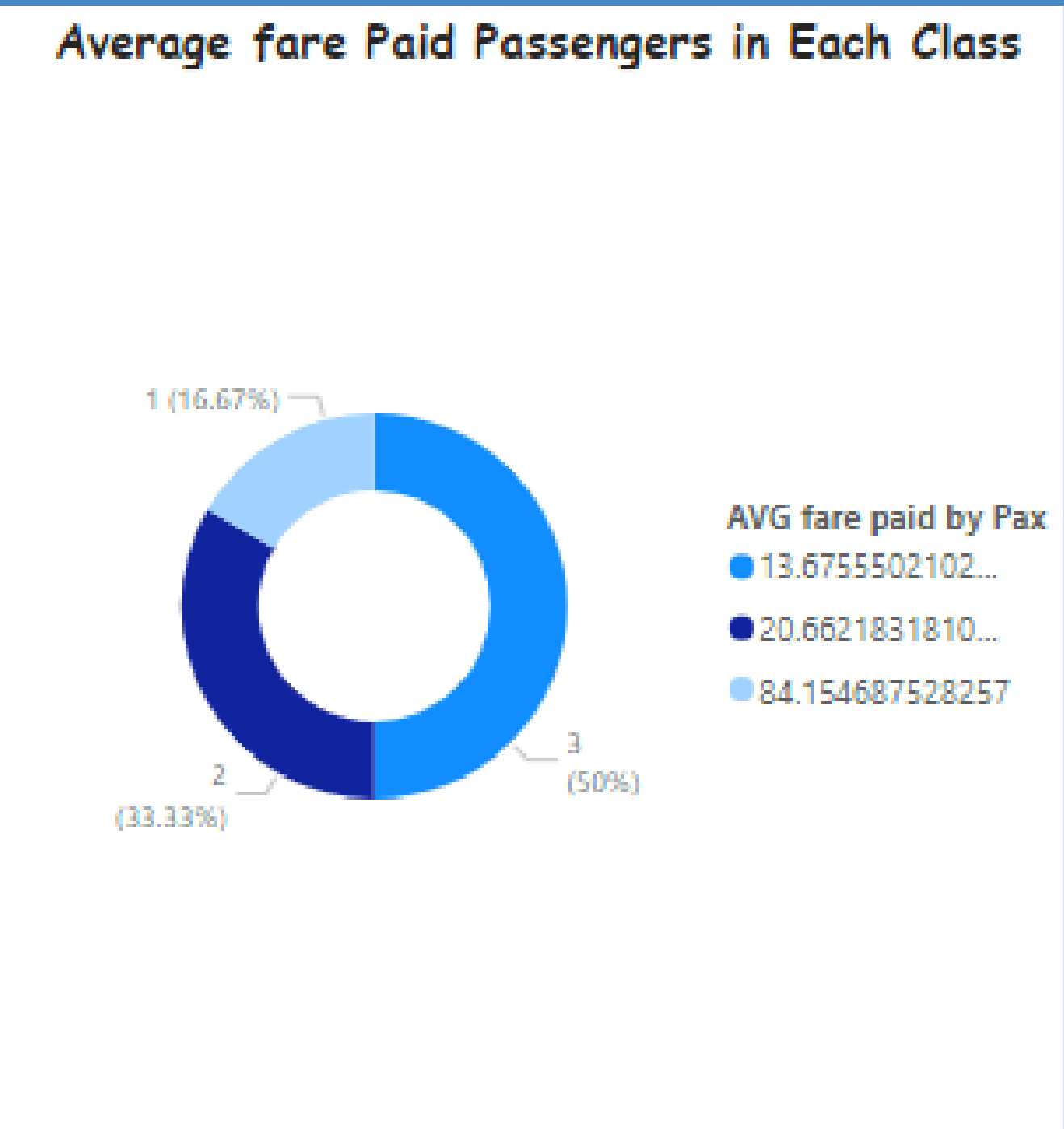
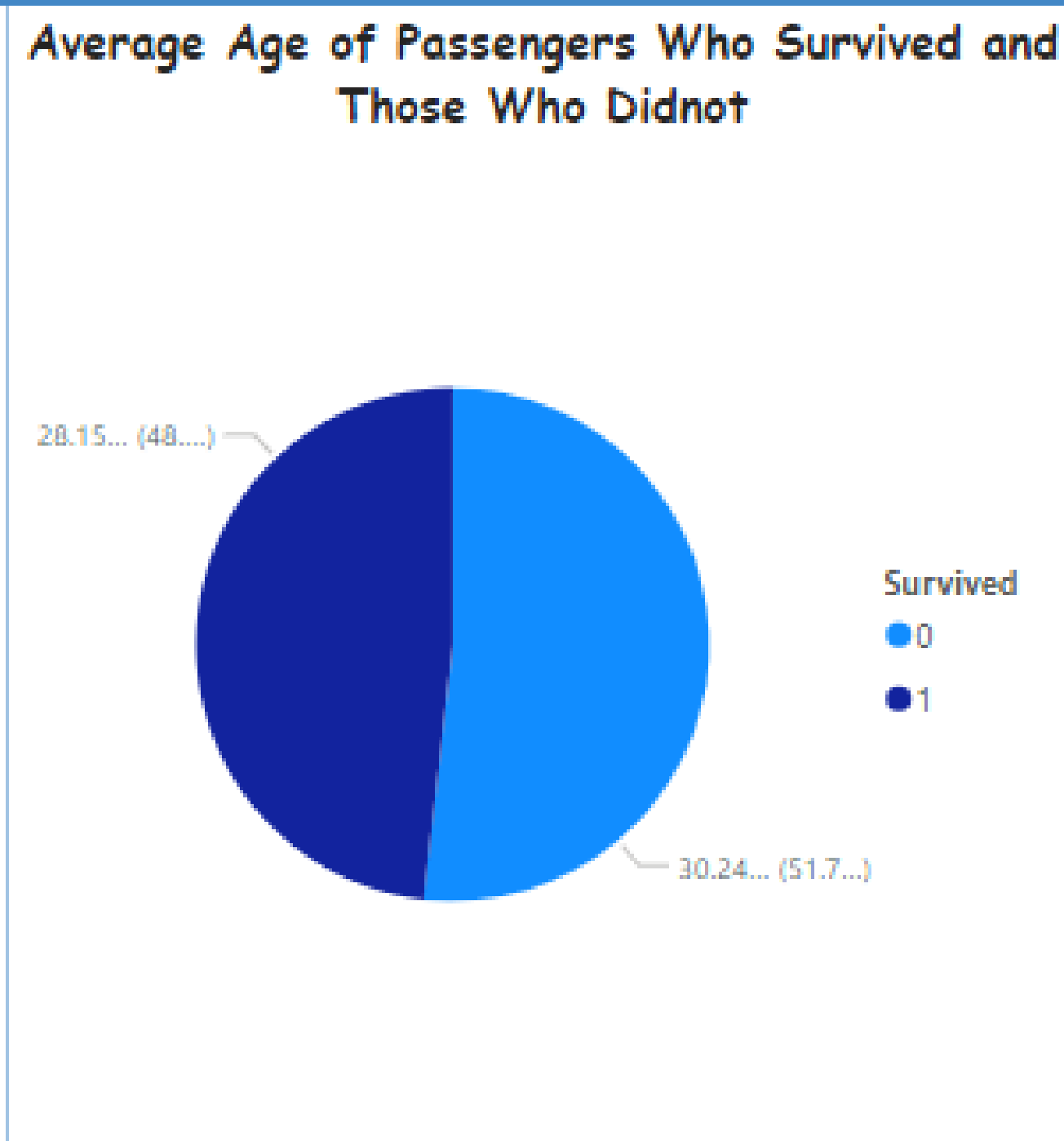
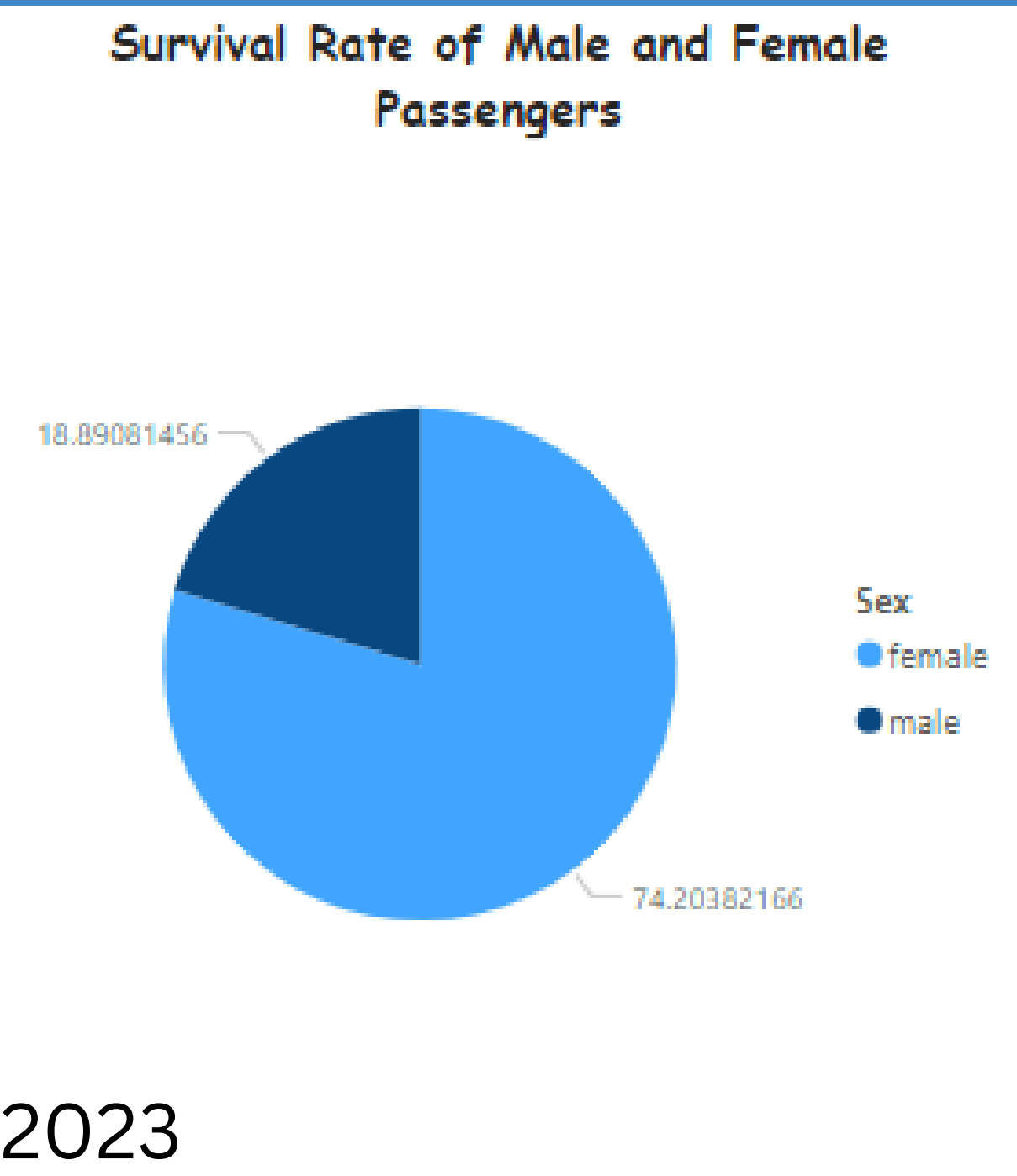
28.69K

Number of Passengers With Highest Fare

3

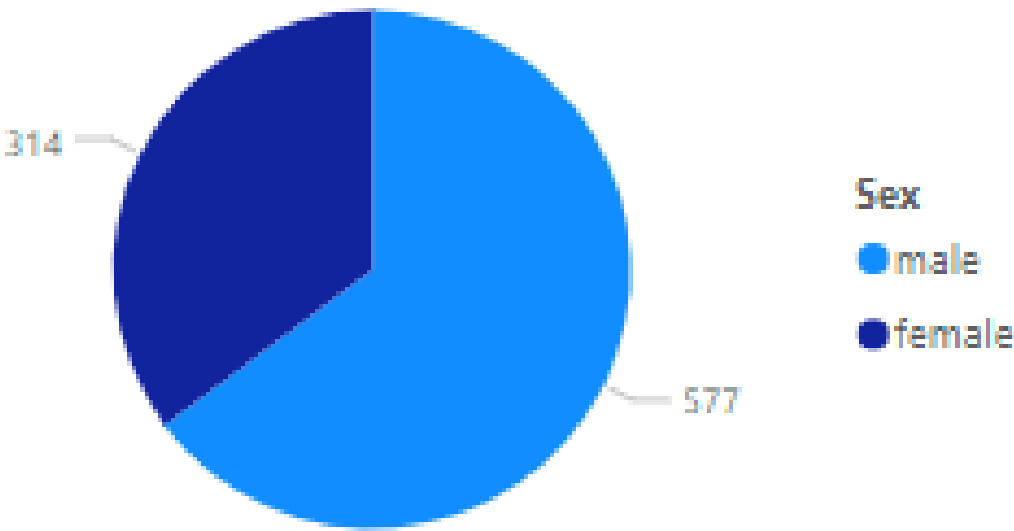
Number of Passengers With Lowest fare

15



TITANIC DASHBOARD

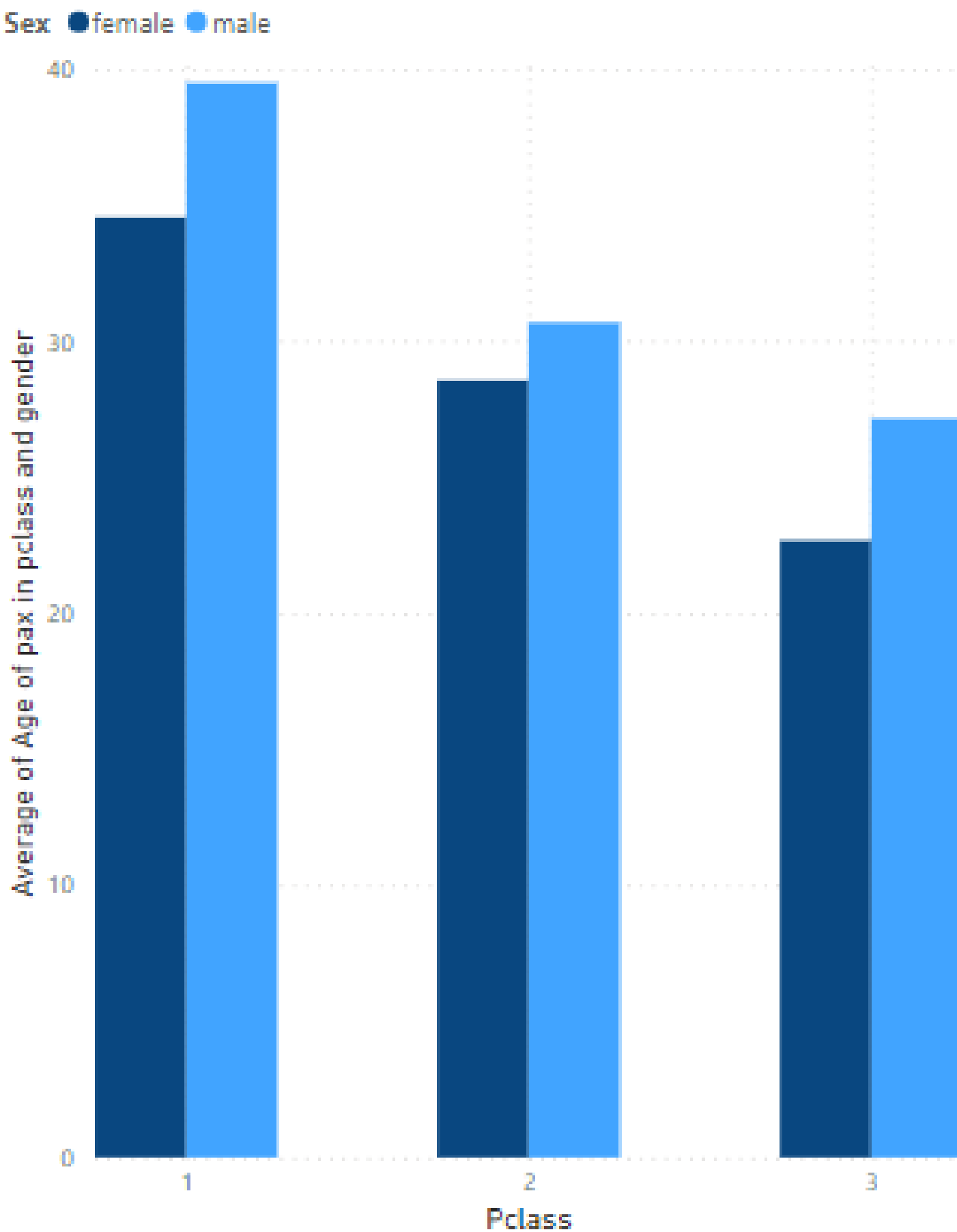
Number of Passengers by Sex



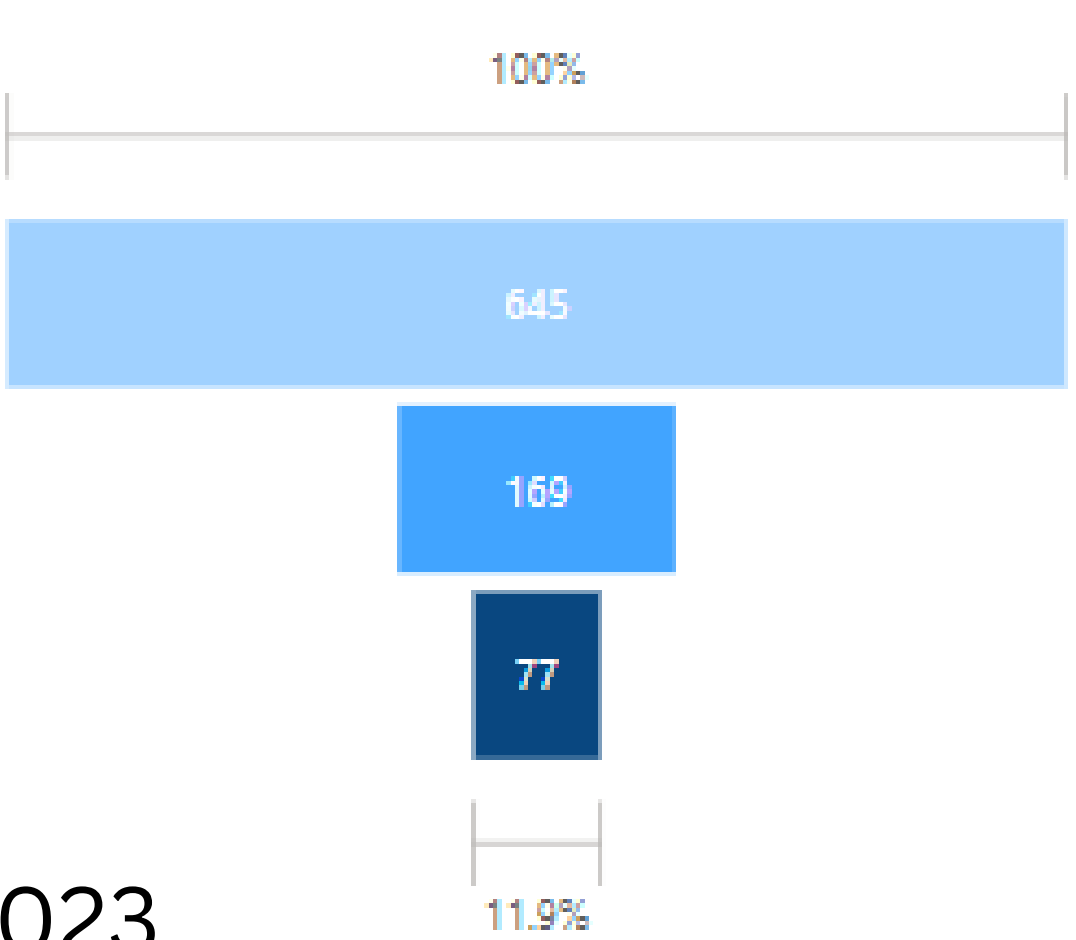
Passengers Who Travelled Alone and Didnot Survive



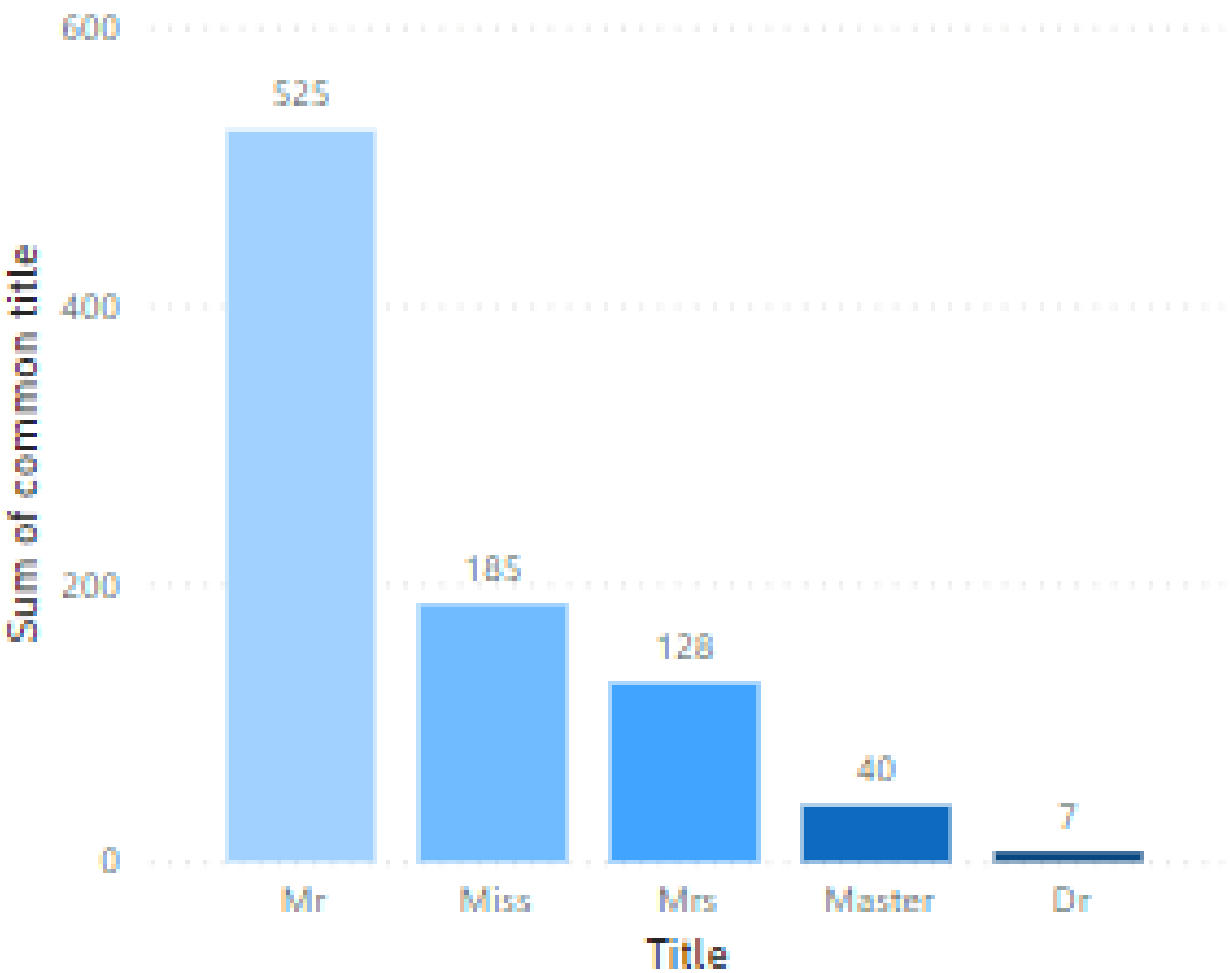
Average Age of Passenger in Each Class and Gender



Passengers and Port of Embarkation



Five Most Common Titles



INSIGHTS

01

From the analysis, sum of no of passengers for male (577) was higher than female (314).

02

At 525 Mr had the highest some of common title and was 7400.00% higher than Dr which had the sum of least common title at 7.

03

The deaths were more than the survivors with a death rate of 57.1% and survivor rate of 48.2%. The survival rate for female (79.71%) was higher than male (20.29%).

port S had the highest number of passengers embarkation with a total of 645 passenger while port C had 169 and port Q had 77 respectively.

Though PClass 1 had a low population of (16.63%) the fare of 84.5 generated was higher than PClass 3 (13.68) with the highest population of 50%.

Average average of age for passengers in PClass and gender was higher for male (32.44) than female (28.59)

04

05

06

RECOMMENDATIONS

01

To avoid possible high rate of deaths, the passengers, crew members and spare life jackets should be provided.

02

There should be constant safety briefing of passengers before boarding to create awareness of what to do incase of emergency..

03

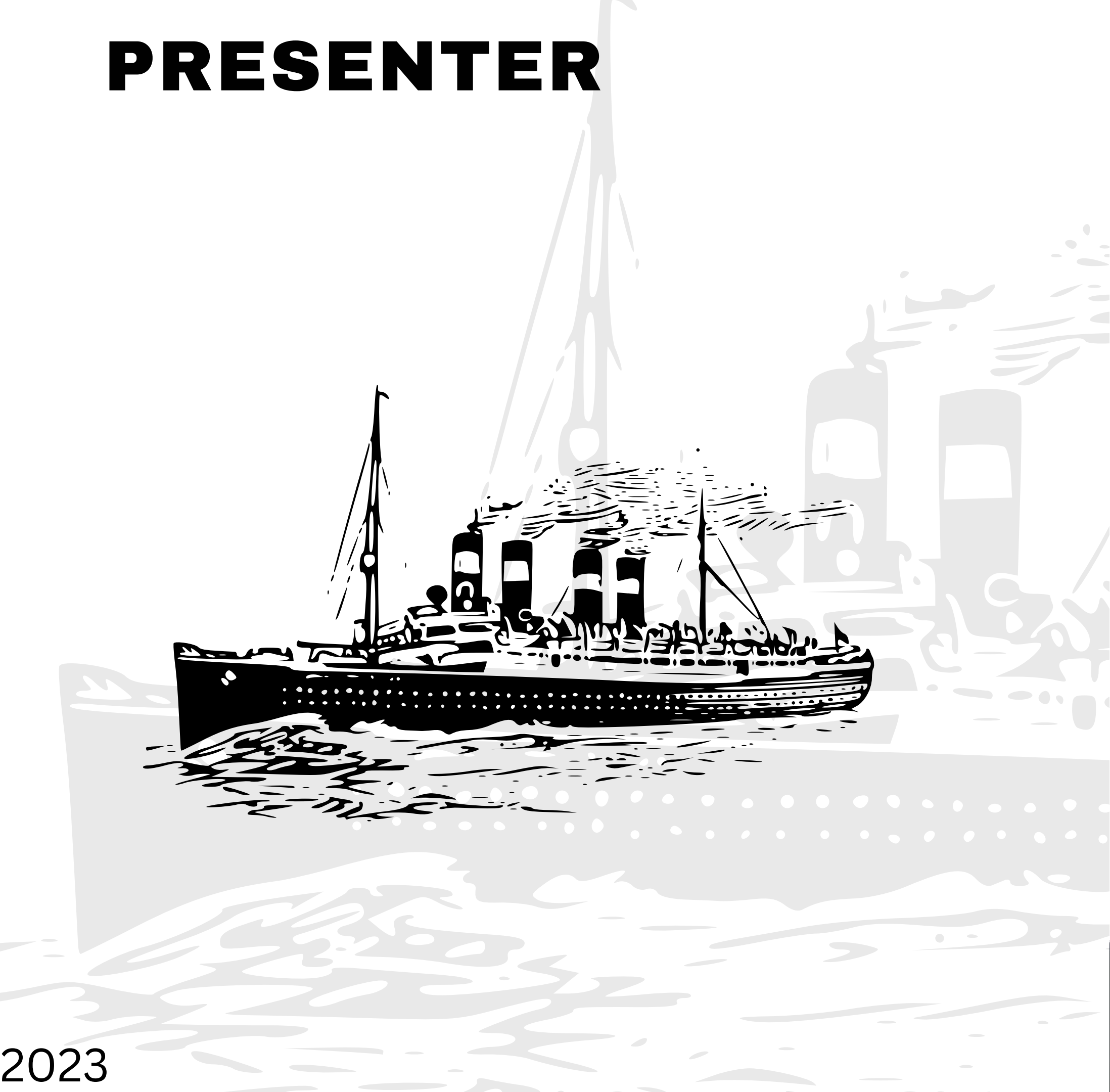
Sufficient provision of life boat to accomodate thenumber of passengers on board for safe evacuation.

04

There should be effective ways of communicating unexpected signals of dangers between the tower watchkeeper and the crew on deck.



PRESENTER



2023



TITANIC PROJECT

**THANK'S FOR
WATCHING**

Chiemeziem Okeke

2023

