```
In []: # === First cell: Imports and basic settings ===

import os
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
import missingno as msno

# Make plots look nicer
plt.rcParams['figure.figsize'] = (8, 5)
sns.set_theme(style="whitegrid")

# Paths

DATA_PATH = "data/train_eda.csv" # adjust if your file has a different name
FIG_DIR = "figures"
os.makedirs(FIG_DIR, exist_ok=True)
```

Importing Libraries

- Imported essential Python libraries for data handling (pandas , numpy).
- Set up plotting libraries (matplotlib , seaborn) for visualizations.
- Added missingno for missing data visualization.
- Configured plot style and created a folder to save generated figures.

```
In [4]: df = pd.read_csv(DATA_PATH)
print(f"Shape: {df.shape}")
df.head()
```

Shape: (891, 12)

Out[4]:		PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	I
	0	1	0	3	Braund, Mr. Owen Harris	male	22.0	1	0	A/5 21171	7.2
	1	2	1	1	Cumings, Mrs. John Bradley (Florence Briggs Th	female	38.0	1	0	PC 17599	71.2
	2	3	1	3	Heikkinen, Miss. Laina	female	26.0	0	0	STON/O2. 3101282	7.9
	3	4	1	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0	1	0	113803	53.1
	4	5	0	3	Allen, Mr. William Henry	male	35.0	0	0	373450	8.0
	4										•

Loading the Dataset

- Loaded the dataset from the data folder into a pandas DataFrame.
- Verified the file path and confirmed successful loading.
- Displayed the first 5 rows to get an initial look at the structure and values.

```
In [5]: # === Third cell: Inspect dataset info and missing values ===

# Basic info about columns and data types
df.info()

# Summary statistics for numeric and categorical columns
display(df.describe(include='all').T)

# Missing values count (sorted)
missing = df.isnull().sum().sort_values(ascending=False)
print("\nMissing values per column:")
print(missing[missing > 0])
```

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 891 entries, 0 to 890
Data columns (total 12 columns):

#	Column	Non-Null Count	Dtype			
0	PassengerId	891 non-null	int64			
1	Survived	891 non-null	int64			
2	Pclass	891 non-null	int64			
3	Name	891 non-null	object			
4	Sex	891 non-null	object			
5	Age	714 non-null	float64			
6	SibSp	891 non-null	int64			
7	Parch	891 non-null	int64			
8	Ticket	891 non-null	object			
9	Fare	891 non-null	float64			
10	Cabin	204 non-null	object			
11	Embarked	889 non-null	object			
dtypes: $float64(2)$ int64(5) object(5)						

dtypes: float64(2), int64(5), object(5)

memory usage: 83.7+ KB

	count	unique	top	freq	mean	std	min	25%	50%
PassengerId	891.0	NaN	NaN	NaN	446.0	257.353842	1.0	223.5	446.0
Survived	891.0	NaN	NaN	NaN	0.383838	0.486592	0.0	0.0	0.0
Pclass	891.0	NaN	NaN	NaN	2.308642	0.836071	1.0	2.0	3.0
Name	891	891	Braund, Mr. Owen Harris	1	NaN	NaN	NaN	NaN	NaN
Sex	891	2	male	577	NaN	NaN	NaN	NaN	NaN
Age	714.0	NaN	NaN	NaN	29.699118	14.526497	0.42	20.125	28.0
SibSp	891.0	NaN	NaN	NaN	0.523008	1.102743	0.0	0.0	0.0
Parch	891.0	NaN	NaN	NaN	0.381594	0.806057	0.0	0.0	0.0
Ticket	891	681	347082	7	NaN	NaN	NaN	NaN	NaN
Fare	891.0	NaN	NaN	NaN	32.204208	49.693429	0.0	7.9104	14.4542
Cabin	204	147	G6	4	NaN	NaN	NaN	NaN	NaN
Embarked	889	3	S	644	NaN	NaN	NaN	NaN	NaN

Missing values per column:

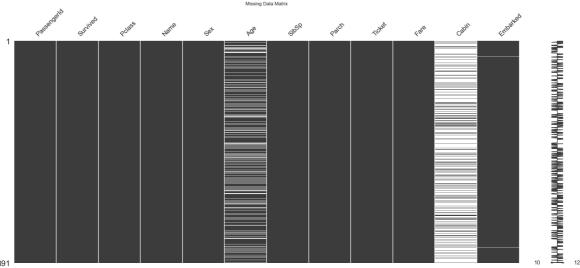
Cabin 687 Age 177 Embarked 2 dtype: int64

Inspecting Data Types and Missing Values

- .info() showed column names, data types, and non-null counts.
- .describe() provided summary statistics for numeric and categorical columns.
- Identified columns with missing values to address in later cleaning steps.

```
In [17]: # === Fourth cell: Missing values visualization ===

msno.matrix(df)
plt.title("Missing Data Matrix")
plt.savefig(f"{FIG_DIR}/missing_matrix.png", bbox_inches='tight', dpi=150)
plt.show()
```

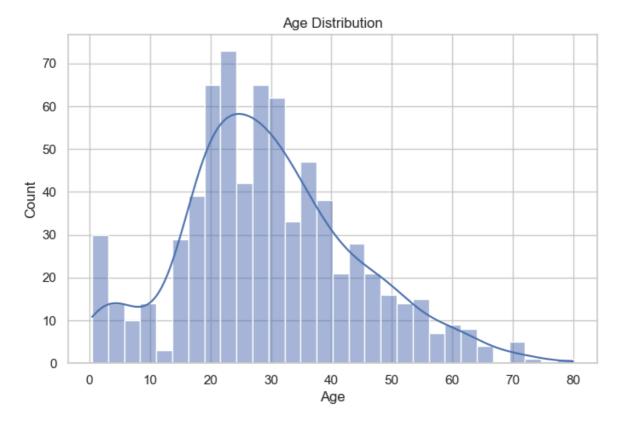


Missing Data Matrix

- Several columns have missing values, with some showing significant gaps.
- White areas indicate missing entries; solid lines mean complete data.
- Columns with heavy missingness may need imputation or removal.

```
In [7]: # === Fifth cell: Histogram for a numeric column ===

# Change 'Age' to any numeric column you want to see
sns.histplot(df['Age'], kde=True, bins=30)
plt.title("Age Distribution")
plt.xlabel("Age")
plt.ylabel("Count")
plt.savefig(f"{FIG_DIR}/hist_age.png", bbox_inches='tight', dpi=150)
plt.show()
```

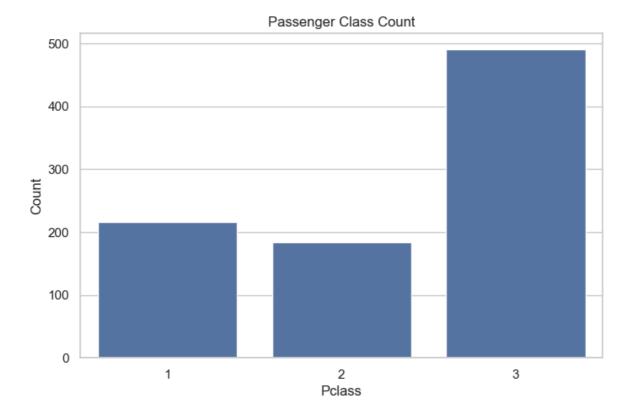


Age Distribution

- Most passengers are between 20 and 40 years old.
- Few passengers above 60 years; tail extends toward older ages.
- Slight right skew in distribution.

```
In [10]: # === Sixth cell: Countplot for a categorical column ===

# Change 'Pclass' to any categorical column in your dataset
sns.countplot(x='Pclass', data=df)
plt.title("Passenger Class Count")
plt.xlabel("Pclass")
plt.ylabel("Count")
plt.savefig(f"{FIG_DIR}/count_pclass.png", bbox_inches='tight', dpi=150)
plt.show()
```

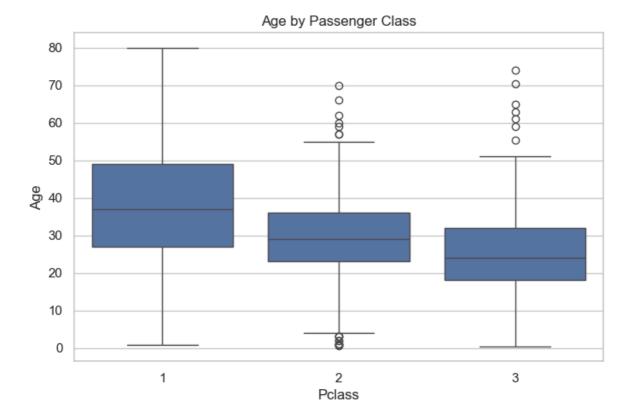


Passenger Class Count

- Most passengers traveled in 3rd class, followed by 1st and 2nd.
- Class imbalance suggests socio-economic differences in the dataset.
- Could influence survival patterns in later analysis.

```
In [12]: # === Seventh cell: Boxplot for numeric vs categorical ===

# Example: Age distribution across Passenger Class
sns.boxplot(x='Pclass', y='Age', data=df)
plt.title("Age by Passenger Class")
plt.xlabel("Pclass")
plt.ylabel("Pclass")
plt.ylabel("Age")
plt.savefig(f"{FIG_DIR}/box_age_by_pclass.png", bbox_inches='tight', dpi=150)
plt.show()
```



Age by Passenger Class

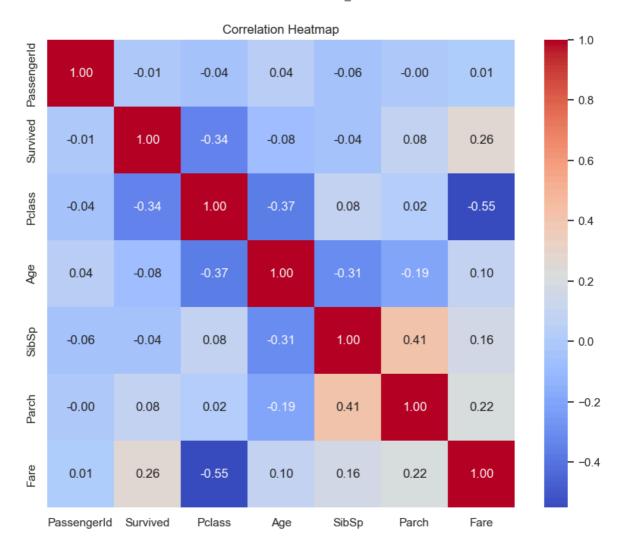
- Median age is highest in 1st class and lowest in 3rd class.
- Wider spread of ages in 1st class; 3rd class has more younger passengers.
- Several outliers appear across all classes.

```
In [13]: # === Eighth cell: Correlation heatmap for numeric columns ===

# Select only numeric columns
numeric_df = df.select_dtypes(include=[np.number])

# Compute correlations
corr = numeric_df.corr()

# Plot heatmap
plt.figure(figsize=(10, 8))
sns.heatmap(corr, annot=True, fmt=".2f", cmap='coolwarm', square=True)
plt.title("Correlation Heatmap")
plt.savefig(f"{FIG_DIR}/correlation_heatmap.png", bbox_inches='tight', dpi=150)
plt.show()
```



Correlation Heatmap

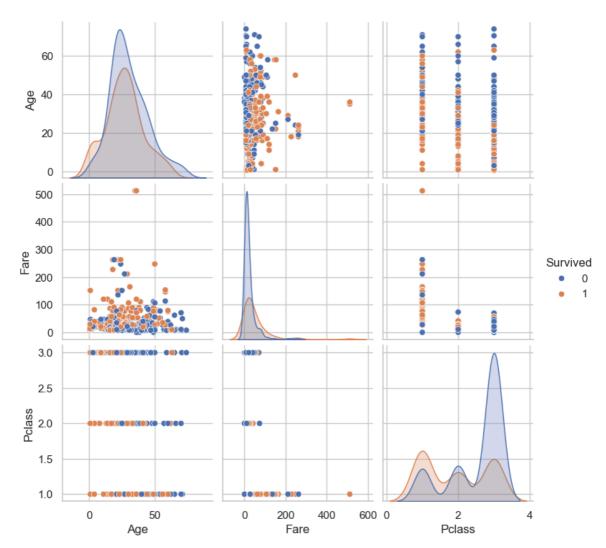
- Fare shows a moderate negative correlation with passenger class (Pclass).
- Age has low correlation with most other variables.
- No extreme multicollinearity detected between numeric columns.

```
In [14]: # === Ninth cell: Pairplot for selected numeric columns ===

# Pick up to 4 numeric columns for pairplot
cols_for_pairplot = ['Age', 'Fare', 'Pclass', 'Survived'] # change if needed

# Sample data if too large
plot_df = df[cols_for_pairplot].sample(min(500, len(df)), random_state=1)

sns.pairplot(plot_df, hue='Survived', diag_kind='kde') # remove hue if no targe
plt.savefig(f"{FIG_DIR}/pairplot_sample.png", bbox_inches='tight', dpi=150)
plt.show()
```



Pairplot

- Clear separation in Fare between higher and lower passenger classes.
- Age does not show strong separation between survival outcomes.
- Some clustering visible in Pclass vs Fare scatterplots.

```
In [16]: # === Tenth cell: Grouped bar chart for target vs categorical ===

target_col = 'Survived' # change if your target column is different

# Choose a few categorical columns except the target itself
cat_cols = df.select_dtypes(include=['object', 'category']).columns.tolist()

for col in cat_cols:
    if col != target_col:
        plt.figure()
        sns.countplot(x=target_col, hue=col, data=df)
        plt.title(f"{target_col} by {col}")
        plt.savefig(f"{FIG_DIR}/{target_col}_by_{col}.png", bbox_inches='tight',
        plt.show()
```

Name Braund, Mr. Owen Harris Allen, Mr. William Henry Moran, Mr. James McCarthy, Mr. Timothy J Palsson, Master. Gosta Leonard Saundercock, Mr. William Henry Andersson, Mr. Anders Johan Vestrom, Miss. Hulda Amanda Adolfina Rice, Master. Eugene Vander Planke, Mrs. Julius (Emelia Maria Vandemoortele) Fynney, Mr. Joseph J Palsson, Miss. Torborg Danira Emir, Mr. Farred Chehab Fortune, Mr. Charles Alexander Todoroff, Mr. Lalio Uruchurtu, Don. Manuel E Wheadon, Mr. Edward H Meyer, Mr. Edgar Joseph Holverson, Mr. Alexander Oskar Cann, Mr. Ernest Charles Vander Planke, Miss. Augusta Maria Ahlin, Mrs. Johan (Johanna Persdotter Larsson) Turpin, Mrs. William John Robert (Dorothy Ann Wonnacott) Kraeff, Mr. Theodor Rogers, Mr. William John Lennon, Mr. Denis Samaan, Mr. Youssef Arnold-Franchi, Mrs. Josef (Josefine Franchi) Panula, Master. Juha Niilo Nosworthy, Mr. Richard Cater Ostby, Mr. Engelhart Cornelius Novel, Mr. Mansouer Goodwin, Master. William Frederick Sirayanian, Mr. Orsen Harris, Mr. Henry Birkhardt Skoog, Master. Harald Stewart, Mr. Albert A Crease, Mr. Ernest James Kink, Mr. Vincenz Jenkin, Mr. Stephen Curnow Goodwin, Miss. Lillian Amy Hood, Mr. Ambrose Jr Chronopoulos, Mr. Apostolos Moen, Mr. Sigurd Hansen Staneff, Mr. Ivan Moutal, Mr. Rahamin Haim Waelens, Mr. Achille Carrau, Mr. Francisco M Ford, Mr. William Neal Slocovski, Mr. Selman Francis Celotti, Mr. Francesco Christmann, Mr. Emil Andreasson, Mr. Paul Edvin

Chaffee, Mr. Herbert FullerDean, Mr. Bertram Frank

- Coxon, Mr. Daniel
- Shorney, Mr. Charles Joseph
- Goldschmidt, Mr. George B
- Kantor, Mr. Sinai
- Petranec, Miss. Matilda
- Petroff, Mr. Pastcho ("Pentcho")
- White, Mr. Richard Frasar
- Johansson, Mr. Gustaf Joel
- Gustafsson, Mr. Anders Vilhelm
- Mionoff, Mr. Stoytcho
- Rekic, Mr. Tido
- Porter, Mr. Walter Chamberlain
- Zabour, Miss. Hileni
- Barton, Mr. David John
- Jussila, Miss. Katriina
- Attalah, Miss. Malake
- Pekoniemi, Mr. Edvard
- Connors. Mr. Patrick
- Turpin, Mr. William John Robert
- Baxter, Mr. Quigg Edmond
- Andersson, Miss. Ellis Anna Maria
- Hickman, Mr. Stanley George
- Moore, Mr. Leonard Charles
- Nasser, Mr. Nicholas
- White, Mr. Percival Wayland
- McMahon, Mr. Martin
- Ekstrom, Mr. Johan
- Drazenoic, Mr. Jozef
- Coelho, Mr. Domingos Fernandeo
- Robins, Mrs. Alexander A (Grace Charity Laury)
- Sobey, Mr. Samuel James Hayden
- Richard, Mr. Emile
- Futrelle, Mr. Jacques Heath
- Osen, Mr. Olaf Elon
- Giglio, Mr. Victor
- Boulos, Mrs. Joseph (Sultana)
- Burke, Mr. Jeremiah
- Andrew, Mr. Edgardo Samuel
- Nicholls, Mr. Joseph Charles
- Ford, Miss. Robina Maggie "Ruby"
- Navratil, Mr. Michel ("Louis M Hoffman")
- Byles, Rev. Thomas Roussel Davids
- Bateman, Rev. Robert James
- Meo, Mr. Alfonzo
- van Billiard, Mr. Austin Blyler
- Olsen, Mr. Ole Martin
- Williams, Mr. Charles Duane
- Corn, Mr. Harry
- Smiljanic, Mr. Mile
- Sage, Master. Thomas Henry
- Cribb, Mr. John Hatfield
- Bengtsson, Mr. John Viktor
- Calic, Mr. Jovo
- Panula, Master. Eino Viljami
- Skoog, Mrs. William (Anna Bernhardina Karlsson)
- Baumann, Mr. John D
- Ling, Mr. Lee

- Van der hoef, Mr. Wyckoff
- Rice, Master. Arthur
- Sivola, Mr. Antti Wilhelm
- Smith, Mr. James Clinch
- Klasen, Mr. Klas Albin
- Lefebre, Master. Henry Forbes
- Isham, Miss. Ann Elizabeth
- Hale, Mr. Reginald
- Leonard, Mr. Lionel
- Sage, Miss. Constance Gladys
- Pernot, Mr. Rene
- Asplund, Master. Clarence Gustaf Hugo
- Rood, Mr. Hugh Roscoe
- Bourke, Mr. John
- Turcin, Mr. Stjepan
- Carbines, Mr. William
- Mernagh, Mr. Robert
- Olsen, Mr. Karl Siegwart Andreas
- Yrois, Miss. Henriette ("Mrs Harbeck")
- Vande Walle, Mr. Nestor Cyriel
- Sage, Mr. Frederick
- Johanson, Mr. Jakob Alfred
- Youseff, Mr. Gerious
- Strom, Miss. Telma Matilda
- Backstrom, Mr. Karl Alfred
- Ali, Mr. Ahmed
- Perkin, Mr. John Henry
- Givard, Mr. Hans Kristensen
- Kiernan, Mr. Philip
- Jacobsohn, Mr. Sidney Samuel
- Harris, Mr. Walter
- Bracken, Mr. James H
- Green, Mr. George Henry
- Nenkoff, Mr. Christo
- Berglund, Mr. Karl Ivar Sven
- Lovell, Mr. John Hall ("Henry")
- Fahlstrom, Mr. Arne Jonas
- Lefebre, Miss. Mathilde
- Larsson, Mr. Bengt Edvin
- Sjostedt, Mr. Ernst Adolf
- Leyson, Mr. Robert William Norman
- Harknett, Miss. Alice Phoebe
- Hold, Mr. Stephen
- Pengelly, Mr. Frederick William
- Hunt, Mr. George Henry
- Zabour, Miss. Thamine
- Coleridge, Mr. Reginald Charles
- Maenpaa, Mr. Matti Alexanteri
- Attalah, Mr. Sleiman
- Minahan, Dr. William Edward
- Lindahl, Miss. Agda Thorilda Viktoria
- Carter, Rev. Ernest Courtenay
- Reed, Mr. James George
- Strom, Mrs. Wilhelm (Elna Matilda Persson)
- Stead, Mr. William Thomas
- Lobb, Mr. William Arthur
- Rosblom, Mrs. Viktor (Helena Wilhelmina)

- Smith, Mr. Thomas
- Taussig, Mr. Emil
- Harrison, Mr. William
- Henry, Miss. Delia
- Reeves, Mr. David
- Panula, Mr. Ernesti Arvid
- Cairns, Mr. Alexander
- Natsch, Mr. Charles H
- Lindblom, Miss. Augusta Charlotta
- Parkes, Mr. Francis "Frank"
- Rice, Master. Eric
- Duane, Mr. Frank
- Olsson, Mr. Nils Johan Goransson
- de Pelsmaeker, Mr. Alfons
- Smith, Mr. Richard William
- Stankovic, Mr. Ivan
- Naidenoff, Mr. Penko
- Levy, Mr. Rene Jacques
- Haas, Miss. Aloisia
- Mineff, Mr. Ivan
- Lewy, Mr. Ervin G
- Hanna, Mr. Mansour
- Allison, Miss. Helen Loraine
- Johnson, Mr. William Cahoone Jr
- Williams, Mr. Howard Hugh "Harry"
- Abelson, Mr. Samuel
- Lahtinen, Mrs. William (Anna Sylfven)
- Hendekovic, Mr. Ignjac
- Hart, Mr. Benjamin
- Moraweck, Dr. Ernest
- Dennis, Mr. Samuel
- Danoff, Mr. Yoto
- Sage, Mr. George John Jr
- Nysveen, Mr. Johan Hansen
- Partner, Mr. Austen
- Graham, Mr. George Edward
- Vander Planke, Mr. Leo Edmondus
- Denkoff, Mr. Mitto
- Pears, Mr. Thomas Clinton
- Blackwell, Mr. Stephen Weart
- Collander, Mr. Erik Gustaf
- Sedgwick, Mr. Charles Frederick Waddington
- Fox, Mr. Stanley Hubert
- Dimic, Mr. Jovan
- Odahl, Mr. Nils Martin
- Williams-Lambert, Mr. Fletcher Fellows
- Elias, Mr. Tannous
- Arnold-Franchi, Mr. Josef
- Yousif, Mr. Wazli
- Vanden Steen, Mr. Leo Peter
- Funk, Miss, Annie Clemmer
- Skoog, Mr. Wilhelm
- del Carlo, Mr. Sebastiano
- Barbara, Mrs. (Catherine David)
- Asim, Mr. Adola
- O'Brien, Mr. Thomas
- Adahl. Mr. Mauritz Nils Martin

- Wiklund, Mr. Jakob Alfred
- Beavan, Mr. William Thomas
- Ringhini, Mr. Sante
- Palsson, Miss. Stina Viola
- Widener, Mr. Harry Elkins
- Betros, Mr. Tannous
- Gustafsson, Mr. Karl Gideon
- Tikkanen, Mr. Juho
- Plotcharsky, Mr. Vasil
- Davies, Mr. Charles Henry
- Goodwin, Master. Sidney Leonard
- Sadlier, Mr. Matthew
- Gustafsson, Mr. Johan Birger
- Johansson, Mr. Erik
- Olsson, Miss. Elina
- McKane, Mr. Peter David
- Pain, Dr. Alfred
- Adams, Mr. John
- Jussila, Miss. Mari Aina
- Hakkarainen, Mr. Pekka Pietari
- Oreskovic, Miss. Marija
- Gale, Mr. Shadrach
- Widegren, Mr. Carl/Charles Peter
- Birkeland, Mr. Hans Martin Monsen
- Lefebre, Miss. Ida
- Sdycoff, Mr. Todor
- Hart, Mr. Henry
- Cunningham, Mr. Alfred Fleming
- Meek, Mrs. Thomas (Annie Louise Rowley)
- Matthews, Mr. William John
- Van Impe, Miss. Catharina
- Gheorgheff, Mr. Stanio
- Charters, Mr. David
- Zimmerman, Mr. Leo
- Danbom, Mrs. Ernst Gilbert (Anna Sigrid Maria Brogren)
- Rosblom, Mr. Viktor Richard
- Wiseman, Mr. Phillippe
- Flynn, Mr. James
- Kallio, Mr. Nikolai Erland
- Silvey, Mr. William Baird
- Ford, Miss. Doolina Margaret "Daisy"
- Fortune, Mr. Mark
- Kvillner, Mr. Johan Henrik Johannesson
- Hampe, Mr. Leon
- Petterson, Mr. Johan Emil
- West, Mr. Edwy Arthur
- Hagland, Mr. Ingvald Olai Olsen
- Foreman, Mr. Benjamin Laventall
- Peduzzi, Mr. Joseph
- Millet, Mr. Francis Davis
- O'Connor, Mr. Maurice
- Morley, Mr. William
- Gee, Mr. Arthur H
- Milling, Mr. Jacob Christian
- Maisner, Mr. Simon
- Goncalves, Mr. Manuel Estanslas
- Campbell Mr William

- Campbon, ivii. vviiiiaini Smart, Mr. John Montgomery Scanlan, Mr. James Keefe, Mr. Arthur Cacic, Mr. Luka Strandberg, Miss. Ida Sofia Clifford, Mr. George Quincy Renouf, Mr. Peter Henry Braund, Mr. Lewis Richard Karlsson, Mr. Nils August Goodwin, Master. Harold Victor Frost, Mr. Anthony Wood "Archie" Rouse, Mr. Richard Henry Lefebre, Miss. Jeannie Kent, Mr. Edward Austin Somerton, Mr. Francis William Hagland, Mr. Konrad Mathias Reiersen Windelov, Mr. Einar Molson, Mr. Harry Markland Artagaveytia, Mr. Ramon Stanley, Mr. Edward Roland Yousseff, Mr. Gerious Shellard, Mr. Frederick William Allison, Mrs. Hudson J C (Bessie Waldo Daniels) Svensson, Mr. Olof Calic, Mr. Petar Canavan, Miss. Mary O'Sullivan, Miss. Bridget Mary Laitinen, Miss. Kristina Sofia Penasco y Castellana, Mr. Victor de Satode Olsen, Mr. Henry Margido Webber, Mr. James Coleff, Mr. Satio Walker, Mr. William Anderson Ryan, Mr. Patrick Pavlovic, Mr. Stefo Vovk, Mr. Janko Lahoud, Mr. Sarkis Kassem, Mr. Fared Farrell, Mr. James Farthing, Mr. John Salonen, Mr. Johan Werner Hocking, Mr. Richard George Toufik, Mr. Nakli Elias, Mr. Joseph Jr Cacic, Miss. Marija Butt, Major. Archibald Willingham Risien, Mr. Samuel Beard Andersson, Miss. Ingeborg Constanzia Andersson, Miss. Sigrid Elisabeth Douglas, Mr. Walter Donald Nicholson, Mr. Arthur Ernest Goldsmith, Mr. Frank John Sharp, Mr. Percival James R O'Brien, Mr. Timothy Wright, Mr. George Robbins, Mr. Victor

iviorrow, ivir. i nomas kowan Sivic, Mr. Husein Norman, Mr. Robert Douglas Simmons, Mr. John Meanwell, Miss. (Marion Ogden) Davies, Mr. Alfred J Stoytcheff, Mr. Ilia Palsson, Mrs. Nils (Alma Cornelia Berglund) Doharr, Mr. Tannous Rush, Mr. Alfred George John Patchett, Mr. George Caram, Mrs. Joseph (Maria Elias) Downton, Mr. William James Ross, Mr. John Hugo Paulner, Mr. Uscher Jarvis, Mr. John Denzil Gilinski, Mr. Eliezer Murdlin, Mr. Joseph Rintamaki, Mr. Matti Elsbury, Mr. William James Bourke, Miss. Mary Chapman, Mr. John Henry Van Impe, Mr. Jean Baptiste Johnson, Mr. Alfred Boulos, Mr. Hanna Slabenoff, Mr. Petco Harrington, Mr. Charles H Torber, Mr. Ernst William Lindell, Mr. Edvard Bengtsson Karaic, Mr. Milan Andersson, Mrs. Anders Johan (Alfrida Konstantia Brogren) Jardin, Mr. Jose Neto Horgan, Mr. John Brocklebank, Mr. William Alfred Danbom, Mr. Ernst Gilbert Lobb, Mrs. William Arthur (Cordelia K Stanlick) Gavey, Mr. Lawrence Yasbeck, Mr. Antoni Hansen, Mr. Henry Damsgaard Bowen, Mr. David John "Dai" Sutton, Mr. Frederick Kirkland, Rev. Charles Leonard Bostandyeff, Mr. Guentcho O'Connell, Mr. Patrick D Lundahl, Mr. Johan Svensson Parr, Mr. William Henry Marsh Skoog, Miss. Mabel Leinonen, Mr. Antti Gustaf Collyer, Mr. Harvey Panula, Mrs. Juha (Maria Emilia Ojala) Thorneycroft, Mr. Percival Jensen, Mr. Hans Peder Skoog, Miss. Margit Elizabeth Cor, Mr. Liudevit Willey, Mr. Edward Mitkoff, Mr. Mito Kalvik, Mr. Johannes Halvorsen

8/11/25, 11:51 PM

- task5 eda Hegarty, Miss. Hanora "Nora" Hickman, Mr. Leonard Mark Radeff, Mr. Alexander Bourke, Mrs. John (Catherine) Eitemiller, Mr. George Floyd Newell, Mr. Arthur Webster Badt, Mr. Mohamed Colley, Mr. Edward Pomeroy Coleff, Mr. Peju Hickman, Mr. Lewis Butler, Mr. Reginald Fenton Rommetvedt, Mr. Knud Paust Cook, Mr. Jacob Davidson, Mr. Thornton Mitchell, Mr. Henry Michael Watson, Mr. Ennis Hastings Edvardsson, Mr. Gustaf Hjalmar Sawyer, Mr. Frederick Charles Goodwin, Mrs. Frederick (Augusta Tyler) Peters, Miss. Katie Olsvigen, Mr. Thor Anderson Goodwin, Mr. Charles Edward Brown, Mr. Thomas William Solomon Laroche, Mr. Joseph Philippe Lemercier Panula, Mr. Jaako Arnold Dakic, Mr. Branko Fischer, Mr. Eberhard Thelander Saad, Mr. Khalil Weir, Col. John Chapman, Mr. Charles Henry Kelly, Mr. James Thayer, Mr. John Borland Humblen, Mr. Adolf Mathias Nicolai Olsen Barbara, Miss. Saiide Gallagher, Mr. Martin Hansen, Mr. Henrik Juul Morley, Mr. Henry Samuel ("Mr Henry Marshall") Klaber, Mr. Herman Larsson, Mr. August Viktor Greenberg, Mr. Samuel Soholt, Mr. Peter Andreas Lauritz Andersen McEvoy, Mr. Michael Johnson, Mr. Malkolm Joackim Jensen, Mr. Svend Lauritz Gillespie, Mr. William Henry Hodges, Mr. Henry Price Oreskovic, Mr. Luka Bryhl, Mr. Kurt Arnold Gottfrid Ilmakangas, Miss. Pieta Sofia Hassan, Mr. Houssein G N Knight, Mr. Robert J Berriman, Mr. William John Troupiansky, Mr. Moses Aaron
- Nankoff, Mr. Minko

file:///C:/Users/DELL/Task5_EDA/report_task5.html

Williams, Mr. Leslie

Ivanoff, Mr. Kanio

Ford, Mrs. Edward (Margaret Ann Watson)

8/11/25, 11:51 PM

task5 eda Cavendish, Mr. Tyrell William McNamee, Mr. Neal Crosby, Capt. Edward Gifford Abbott, Mr. Rossmore Edward Marvin, Mr. Daniel Warner Connaghton, Mr. Michael Vande Velde, Mr. Johannes Joseph Jonkoff, Mr. Lalio Carlsson, Mr. August Sigfrid Bailey, Mr. Percy Andrew Theobald, Mr. Thomas Leonard Garfirth, Mr. John Nirva, Mr. lisakki Antino Aijo Eklund, Mr. Hans Linus Brewe, Dr. Arthur Jackson Mangan, Miss. Mary Moran, Mr. Daniel J Gronnestad, Mr. Daniel Danielsen Lievens, Mr. Rene Aime Jensen, Mr. Niels Peder Mack, Mrs. (Mary) Elias, Mr. Dibo Myhrman, Mr. Pehr Fabian Oliver Malkolm Tobin, Mr. Roger Kilgannon, Mr. Thomas J Long, Mr. Milton Clyde Johnston, Mr. Andrew G Ali, Mr. William Harmer, Mr. Abraham (David Lishin) Rice, Master. George Hugh Guggenheim, Mr. Benjamin Keane, Mr. Andrew "Andy" Gaskell, Mr. Alfred Sage, Miss. Stella Anna Hoyt, Mr. William Fisher Dantcheff, Mr. Ristiu Otter, Mr. Richard Ibrahim Shawah, Mr. Yousseff Van Impe, Mrs. Jean Baptiste (Rosalie Paula Govaert) Ponesell, Mr. Martin Johansson, Mr. Karl Johan Andrews, Mr. Thomas Jr Pettersson, Miss. Ellen Natalia Meyer, Mr. August Alexander, Mr. William Lester, Mr. James Slemen, Mr. Richard James Andersson, Miss. Ebba Iris Alfrida

- Tomlin, Mr. Ernest Portage Fry, Mr. Richard
- Heininen, Miss. Wendla Maria
- Mallet, Mr. Albert
- Holm, Mr. John Fredrik Alexander
- Skoog, Master. Karl Thorsten
- Reuchlin, Jonkheer. John George
- Panula, Master. Urho Abraham
- Flynn, Mr. John

Lam, Mr. Len Saad, Mr. Amin Augustsson, Mr. Albert Allum, Mr. Owen George Pasic, Mr. Jakob Sirota, Mr. Maurice Alhomaki, Mr. Ilmari Rudolf Mudd, Mr. Thomas Charles Lemberopolous, Mr. Peter L Culumovic, Mr. Jeso Abbing, Mr. Anthony Sage, Mr. Douglas Bullen Markoff, Mr. Marin Harper, Rev. John Andersson, Master. Sigvard Harald Elias Svensson, Mr. Johan Boulos, Miss. Nourelain Carter, Mrs. Ernest Courtenay (Lilian Hughes) Razi, Mr. Raihed Hansen, Mr. Claus Peter Giles, Mr. Frederick Edward Sage, Miss. Dorothy Edith "Dolly" Gill, Mr. John William Roebling, Mr. Washington Augustus II van Melkebeke, Mr. Philemon Balkic, Mr. Cerin Carlsson, Mr. Frans Olof Vander Cruyssen, Mr. Victor Gustafsson, Mr. Alfred Ossian Petroff, Mr. Nedelio Laleff, Mr. Kristo Markun, Mr. Johann Dahlberg, Miss. Gerda Ulrika Banfield, Mr. Frederick James Sutehall, Mr. Henry Jr Rice, Mrs. William (Margaret Norton) Montvila, Rev. Juozas Johnston, Miss. Catherine Helen "Carrie" Dooley, Mr. Patrick Cumings, Mrs. John Bradley (Florence Briggs Thayer) Heikkinen, Miss, Laina Futrelle, Mrs. Jacques Heath (Lily May Peel) Johnson, Mrs. Oscar W (Elisabeth Vilhelmina Berg) Nasser, Mrs. Nicholas (Adele Achem) Sandstrom, Miss. Marguerite Rut Bonnell, Miss. Elizabeth Hewlett, Mrs. (Mary D Kingcome) Williams, Mr. Charles Eugene Masselmani, Mrs. Fatima Beesley, Mr. Lawrence McGowan, Miss. Anna "Annie" Sloper, Mr. William Thompson Asplund, Mrs. Carl Oscar (Selma Augusta Emilia Johansson) O'Dwyer, Miss. Ellen "Nellie" Spencer, Mrs. William Augustus (Marie Eugenie) Glynn, Miss. Mary Agatha

Mamee Mr Hanna

Nicola-Yarred, Miss. Jamila Laroche, Miss. Simonne Marie Anne Andree Devaney, Miss. Margaret Delia O'Driscoll, Miss. Bridget Harper, Mrs. Henry Sleeper (Myna Haxtun) Faunthorpe, Mrs. Lizzie (Elizabeth Anne Wilkinson) Woolner, Mr. Hugh Rugg, Miss. Emily West, Miss. Constance Mirium Icard, Miss. Amelie Moubarek, Master, Gerios Nye, Mrs. (Elizabeth Ramell) Andersson, Miss. Erna Alexandra Bing, Mr. Lee Caldwell, Master. Alden Gates Dowdell, Miss. Elizabeth Sheerlinck, Mr. Jan Baptist McDermott, Miss. Brigdet Delia Ilett, Miss. Bertha Backstrom, Mrs. Karl Alfred (Maria Mathilda Gustafsson) Fortune, Miss. Mabel Helen Greenfield, Mr. William Bertram Doling, Mrs. John T (Ada Julia Bone) Salkjelsvik, Miss. Anna Kristine Moss, Mr. Albert Johan Moran, Miss. Bertha Webber, Miss. Susan Nicola-Yarred, Master, Elias Madsen, Mr. Fridtjof Arne Peter, Miss. Anna Weisz, Mrs. Leopold (Mathilde Francoise Pede) Newsom, Miss. Helen Monypeny Nysten, Miss. Anna Sofia Hakkarainen, Mrs. Pekka Pietari (Elin Matilda Dolck) Andersson, Mr. August Edvard ("Wennerstrom") Pears, Mrs. Thomas (Edith Wearne) Gilnagh, Miss. Katherine "Katie" Watt, Mrs. James (Elizabeth "Bessie" Inglis Milne) Goldsmith, Master. Frank John William "Frankie" Chibnall, Mrs. (Edith Martha Bowerman) Johnson, Miss. Eleanor Ileen Becker, Master. Richard F Kink-Heilmann, Miss. Luise Gretchen O'Brien, Mrs. Thomas (Johanna "Hannah" Godfrey) Romaine, Mr. Charles Hallace ("Mr C Rolmane") Pinsky, Mrs. (Rosa) Andersen-Jensen, Miss. Carla Christine Nielsine Navratil, Master. Michel M Brown, Mrs. James Joseph (Margaret Tobin) Lurette, Miss. Elise Madigan, Miss. Margaret "Maggie" Cohen, Mr. Gurshon "Gus" Albimona, Mr. Nassef Cassem Carr, Miss. Helen "Ellen" Blank, Mr. Henry Cameron, Miss. Clear Annie Nowall Miss Madalaina

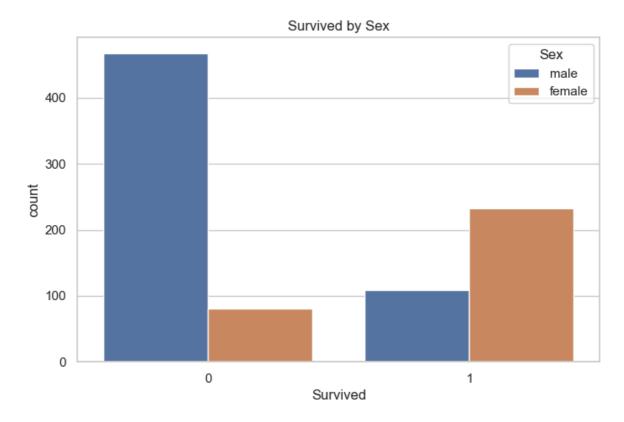
INCWCII, IVIISS. IVIAUCICIIIC Honkanen, Miss. Eliina Bazzani, Miss. Albina Sunderland, Mr. Victor Francis Hoyt, Mr. Frederick Maxfield Mellors, Mr. William John Harris, Mrs. Henry Birkhardt (Irene Wallach) Asplund, Miss. Lillian Gertrud Collyer, Miss. Marjorie "Lottie" Murphy, Miss. Katherine "Kate" Hamalainen, Mrs. William (Anna) Beckwith, Mr. Richard Leonard Touma, Mrs. Darwis (Hanne Youssef Razi) Thorne, Mrs. Gertrude Maybelle Cherry, Miss. Gladys Ward, Miss. Anna Parrish, Mrs. (Lutie Davis) Asplund, Master. Edvin Rojj Felix Persson, Mr. Ernst Ulrik Graham, Mrs. William Thompson (Edith Junkins) Bissette, Miss. Amelia Tornquist, Mr. William Henry Mellinger, Mrs. (Elizabeth Anne Maidment) Healy, Miss. Hanora "Nora" Andrews, Miss. Kornelia Theodosia Abbott, Mrs. Stanton (Rosa Hunt) Dorking, Mr. Edward Arthur de Mulder, Mr. Theodore Hosono, Mr. Masabumi Connolly, Miss. Kate Barber, Miss. Ellen "Nellie" Bishop, Mrs. Dickinson H (Helen Walton) Saalfeld, Mr. Adolphe Baxter, Mrs. James (Helene DeLaudeniere Chaput) Kelly, Miss. Anna Katherine "Annie Kate" McCoy, Mr. Bernard Keane, Miss. Nora A Allison, Master. Hudson Trevor Fleming, Miss. Margaret Penasco y Castellana, Mrs. Victor de Satode (Maria Josefa Perez de Soto y Vallejo) Francatelli, Miss. Laura Mabel Hays, Miss. Margaret Bechstein Ryerson, Miss. Emily Borie Nilsson, Miss. Helmina Josefina Kantor, Mrs. Sinai (Miriam Sternin) Wick, Miss. Mary Natalie Spedden, Mrs. Frederic Oakley (Margaretta Corning Stone) Slayter, Miss. Hilda Mary Caldwell, Mrs. Albert Francis (Sylvia Mae Harbaugh) Young, Miss. Marie Grice Ball, Mrs. (Ada E Hall) Goldsmith, Mrs. Frank John (Emily Alice Brown) Hippach, Miss. Jean Gertrude McCoy, Miss. Agnes Frauenthal, Mrs. Henry William (Clara Heinsheimer) Burns, Miss. Elizabeth Margaret Dahl, Mr. Karl Edwart

- Navratil, Master. Edmond Roger
- Fortune, Miss. Alice Elizabeth
- Brown, Miss. Amelia "Mildred"
- Smith, Miss. Marion Elsie
- Davison, Mrs. Thomas Henry (Mary E Finck)
- Coutts, Master. William Loch "William"
- Bowerman, Miss. Elsie Edith
- McGovern, Miss. Mary
- Mockler, Miss. Helen Mary "Ellie"
- Warren, Mrs. Frank Manley (Anna Sophia Atkinson)
- Moussa, Mrs. (Mantoura Boulos)
- Jermyn, Miss. Annie
- Aubart, Mme. Leontine Pauline
- Harder, Mr. George Achilles
- Meyer, Mrs. Edgar Joseph (Leila Saks)
- Landergren, Miss. Aurora Adelia
- Bidois, Miss. Rosalie
- Nakid, Miss. Maria ("Mary")
- Holverson, Mrs. Alexander Oskar (Mary Aline Towner)
- Buss, Miss. Kate
- Lehmann, Miss. Bertha
- Carter, Mr. William Ernest
- Jansson, Mr. Carl Olof
- Newell, Miss. Marjorie
- Sandstrom, Mrs. Hjalmar (Agnes Charlotta Bengtsson)
- Trout, Mrs. William H (Jessie L)
- Niskanen, Mr. Juha
- Richards, Master. William Rowe
- Minahan, Miss. Daisy E
- Sundman, Mr. Johan Julian
- Drew, Mrs. James Vivian (Lulu Thorne Christian)
- Silven, Miss. Lyyli Karoliina
- Clarke, Mrs. Charles V (Ada Maria Winfield)
- Phillips, Miss. Kate Florence ("Mrs Kate Louise Phillips Marshall")
- Pickard, Mr. Berk (Berk Trembisky)
- Bjornstrom-Steffansson, Mr. Mauritz Hakan
- Thorneycroft, Mrs. Percival (Florence Kate White)
- Louch, Mrs. Charles Alexander (Alice Adelaide Slow)
- Carter, Miss. Lucile Polk
- Richards, Mrs. Sidney (Emily Hocking)
- Hart, Mrs. Benjamin (Esther Ada Bloomfield)
- Reynaldo, Ms. Encarnacion
- Johannesen-Bratthammer, Mr. Bernt
- Dodge, Master. Washington
- Mellinger, Miss. Madeleine Violet
- Seward, Mr. Frederic Kimber
- Baclini, Miss. Marie Catherine
- Peuchen, Major. Arthur Godfrey
- Goldenberg, Mr. Samuel L
- Jalsevac, Mr. Ivan
- Kenyon, Mrs. Frederick R (Marion)
- Toomey, Miss, Ellen
- Anderson, Mr. Harry
- Baclini, Miss. Helene Barbara
- West, Mrs. Edwy Arthur (Ada Mary Worth)
- Jerwan, Mrs. Amin S (Marie Marthe Thuillard)
- Hirvonen, Miss. Hildur E

- Turkula, Mrs. (Hedwig)
- Bishop, Mr. Dickinson H
- Hoyt, Mrs. Frederick Maxfield (Jane Anne Forby)
- Coutts, Master. Eden Leslie "Neville"
- Eustis, Miss. Elizabeth Mussey
- Maioni, Miss. Roberta
- Quick, Mrs. Frederick Charles (Jane Richards)
- Bradley, Mr. George ("George Arthur Brayton")
- Lang, Mr. Fang
- Daly, Mr. Eugene Patrick
- McGough, Mr. James Robert
- Rothschild, Mrs. Martin (Elizabeth L. Barrett)
- Lemore, Mrs. (Amelia Milley)
- Angle, Mrs. William A (Florence "Mary" Agnes Hughes)
- Perreault, Miss. Anne
- Hippach, Mrs. Louis Albert (Ida Sophia Fischer)
- Ridsdale, Miss. Lucy
- Quick, Miss. Phyllis May
- Peter, Mrs. Catherine (Catherine Rizk)
- Hart, Miss. Eva Miriam
- LeRoy, Miss. Bertha
- Frolicher, Miss. Hedwig Margaritha
- Crosby, Miss. Harriet R
- Beane, Mr. Edward
- Beane, Mrs. Edward (Ethel Clarke)
- Padro y Manent, Mr. Julian
- Davies, Master. John Morgan Jr
- Thayer, Mr. John Borland Jr
- Leeni, Mr. Fahim ("Philip Zenni")
- Ohman, Miss. Velin
- Duff Gordon, Lady. (Lucille Christiana Sutherland) ("Mrs Morgan")
- Taussig, Mrs. Emil (Tillie Mandelbaum)
- de Messemaeker, Mrs. Guillaume Joseph (Emma)
- Jonsson, Mr. Carl
- Harris, Mr. George
- Appleton, Mrs. Edward Dale (Charlotte Lamson)
- Flynn, Mr. John Irwin ("Irving")
- Kelly, Miss. Mary
- Garside, Miss. Ethel
- Silvey, Mrs. William Baird (Alice Munger)
- Jussila, Mr. Eiriik
- Christy, Miss. Julie Rachel
- Thayer, Mrs. John Borland (Marian Longstreth Morris)
- Taussig, Miss. Ruth
- Frolicher-Stehli, Mr. Maxmillian
- Stephenson, Mrs. Walter Bertram (Martha Eustis)
- Leitch, Miss. Jessie Wills
- Duff Gordon, Sir. Cosmo Edmund ("Mr Morgan")
- Jacobsohn, Mrs. Sidney Samuel (Amy Frances Christy)
- Homer, Mr. Harry ("Mr E Haven")
- Daniel, Mr. Robert Williams
- Laroche, Mrs. Joseph (Juliette Marie Louise Lafargue)
- Shutes, Miss. Elizabeth W
- Murphy, Miss. Margaret Jane
- Herman, Miss. Alice
- Becker, Miss. Marion Louise
- Kimball, Mr. Edwin Nelson Jr

- Nakid, Mr. Sahid
- Longley, Miss. Gretchen Fiske
- Barkworth, Mr. Algernon Henry Wilson
- Stahelin-Maeglin, Dr. Max
- Davis, Miss. Mary
- Sagesser, Mlle. Emma
- Foo, Mr. Choong
- Baclini, Miss. Eugenie
- Harper, Mr. Henry Sleeper
- Simonius-Blumer, Col. Oberst Alfons
- Stanley, Miss. Amy Zillah Elsie
- Doling, Miss. Elsie
- O'Leary, Miss. Hanora "Norah"
- Frauenthal, Dr. Henry William
- Lindqvist, Mr. Eino William
- Taylor, Mrs. Elmer Zebley (Juliet Cummins Wright)
- Brown, Mrs. Thomas William Solomon (Elizabeth Catherine Ford)
- Wilhelms, Mr. Charles
- Turja, Miss. Anna Sofia
- Cardeza, Mr. Thomas Drake Martinez
- Hassab, Mr. Hammad
- Madill, Miss. Georgette Alexandra
- Dick, Mr. Albert Adrian
- Karun, Miss. Manca
- Lam, Mr. Ali
- Mullens, Miss. Katherine "Katie"
- Astor, Mrs. John Jacob (Madeleine Talmadge Force)
- Silverthorne, Mr. Spencer Victor
- Kelly, Mrs. Florence "Fannie"
- Calderhead, Mr. Edward Pennington
- Cleaver, Miss. Alice
- Moubarek, Master. Halim Gonios ("William George")
- Mayne, Mlle. Berthe Antonine ("Mrs de Villiers")
- Taylor, Mr. Elmer Zebley
- Endres, Miss. Caroline Louise
- Troutt, Miss. Edwina Celia "Winnie"
- Harper, Miss. Annie Jessie "Nina"
- Chambers, Mr. Norman Campbell
- Renouf, Mrs. Peter Henry (Lillian Jefferys)
- Mannion, Miss. Margareth
- Allen, Miss. Elisabeth Walton
- Lesurer, Mr. Gustave J
- Hawksford, Mr. Walter James
- Ryerson, Miss. Susan Parker "Suzette"
- Stranden, Mr. Juho
- Sinkkonen, Miss. Anna
- Wells, Miss. Joan
- Moor, Master. Meier
- Herman, Mrs. Samuel (Jane Laver)
- Hamalainen, Master. Viljo
- Rothes, the Countess. of (Lucy Noel Martha Dyer-Edwards)
- Barah, Mr. Hanna Assi
- Carter, Mrs. William Ernest (Lucile Polk)
- Hogeboom, Mrs. John C (Anna Andrews)
- Hocking, Mrs. Elizabeth (Eliza Needs)
- Emanuel, Miss. Virginia Ethel
- Robert, Mrs. Edward Scott (Elisabeth Walton McMillan)



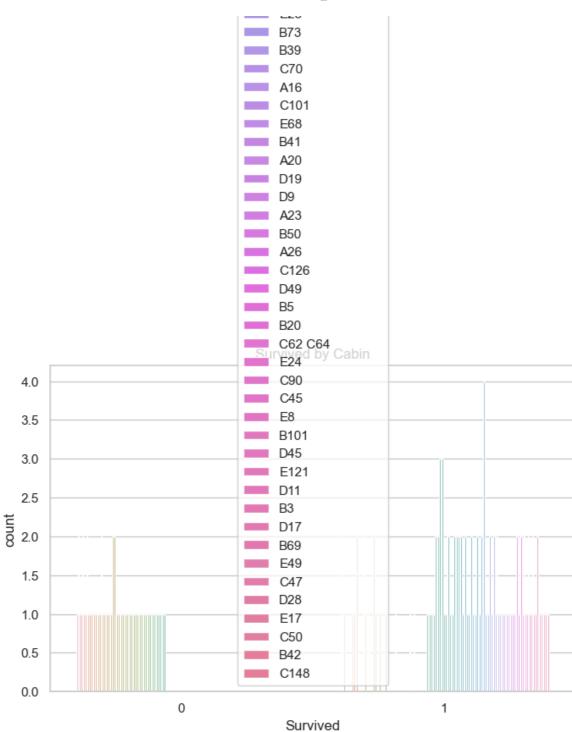


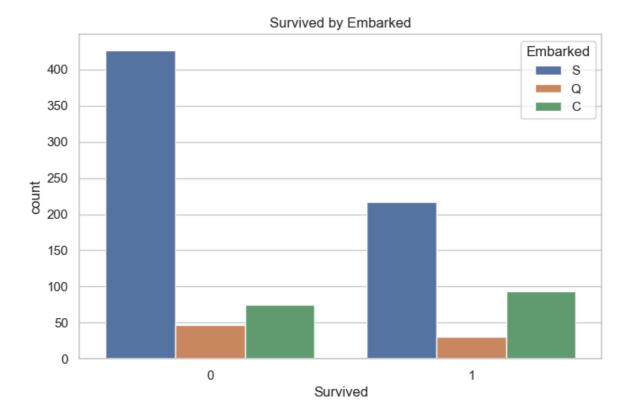


Cabin E46 C23 C25 C27 B30 C83 F G73 E31 A5 D26 C110 B58 B60 C123 B86 F2 B19 A7 C49 A32 G6 C78 C87 E67 B94 C118 A19 ___ D C22 C26 C124 C91 C2 T C128 C82 E44 C111 E38 E63 A14 B37 C30 C65 D46 C95 B38 C86 A10 D50 D48 === E58 B71 C68 F G63 C46 B22 D30

____ C77

F38 D6 B82 B84 A36 B102 A24 B51 B53 B55 C85 C103 D56 A6 B78 D33 C52 B28 F33 D10 D12 E101 F E69 D47 E33 F4 **B**4 B80 A31 D36 D15 C93 D35 B77 C125 C99 D7 B49 C106 E36 C54 B57 B59 B63 B66 C7 E34 C32 B18 E40 D37 B35 E50 B96 B98 E10 A34 C104 C92 D21 E12 D20 B79 F25





Survival by Categorical Variables

- Female passengers had a much higher survival rate than male passengers.
- 1st class passengers had higher survival rates than 2nd or 3rd.
- Some embarkation ports show differences in survival rates.

Final Summary of Findings

- The dataset contained X rows and Y columns, with some missing values in specific columns (e.g., Cabin).
- Key demographic patterns: majority passengers aged 20–40; male passengers outnumbered females.
- Socio-economic influence: 1st class passengers tended to be older and had higher survival rates.
- Survival patterns: females and younger passengers had better survival chances.
- Recommendations for next steps:
 - Create new features such as "Family Size" and "Title" from names.
 - Handle missing values for critical features using median/mode or predictive imputation.
 - Use findings to inform machine learning models.