

TITANIC

March 9, 2025

1 Titanic dataset analysis

1.0.1 Esta tarea implica la limpieza y el análisis del conjunto de datos del Titanic. El conjunto de datos está disponible en Kaggle y contiene información sobre los pasajeros del Titanic, como su edad, clase, tarifa, etc.

1.0.2 1. Import and clean the dataset

	PassengerId	Survived	Pclass	Name \
0	1	0	3	Braund
1	2	1	1	Cumings
2	3	1	3	Heikkinen
3	4	1	1	Futrelle
4	5	0	3	Allen

	Lastname	Sex	Age	SibSp	Parch \
0	Mr. Owen Harris	male	22.0	1	0
1	Mrs. John Bradley (Florence Briggs Thayer)	female	38.0	1	0
2	Miss. Laina	female	26.0	0	0
3	Mrs. Jacques Heath (Lily May Peel)	female	35.0	1	0
4	Mr. William Henry	male	35.0	0	0

	Ticket	Fare	Embarked
0	A/5 21171	7.2500	S
1	PC 17599	71.2833	C
2	STON/O2. 3101282	7.9250	S
3	113803	53.1000	S
4	373450	8.0500	S

```
<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   Lastname      891 non-null    object
5   Sex           891 non-null    object
6   Age           891 non-null    float64
7   SibSp         891 non-null    int64
8   Parch         891 non-null    int64
9   Ticket        891 non-null    object
10  Fare          891 non-null    float64
11  Embarked      891 non-null    object
dtypes: float64(2), int64(5), object(5)
memory usage: 83.7+ KB
Columns: ['PassengerId', 'Survived', 'Pclass', 'Name', 'Lastname', 'Sex',
'Age', 'SibSp', 'Parch', 'Ticket', 'Fare', 'Embarked']

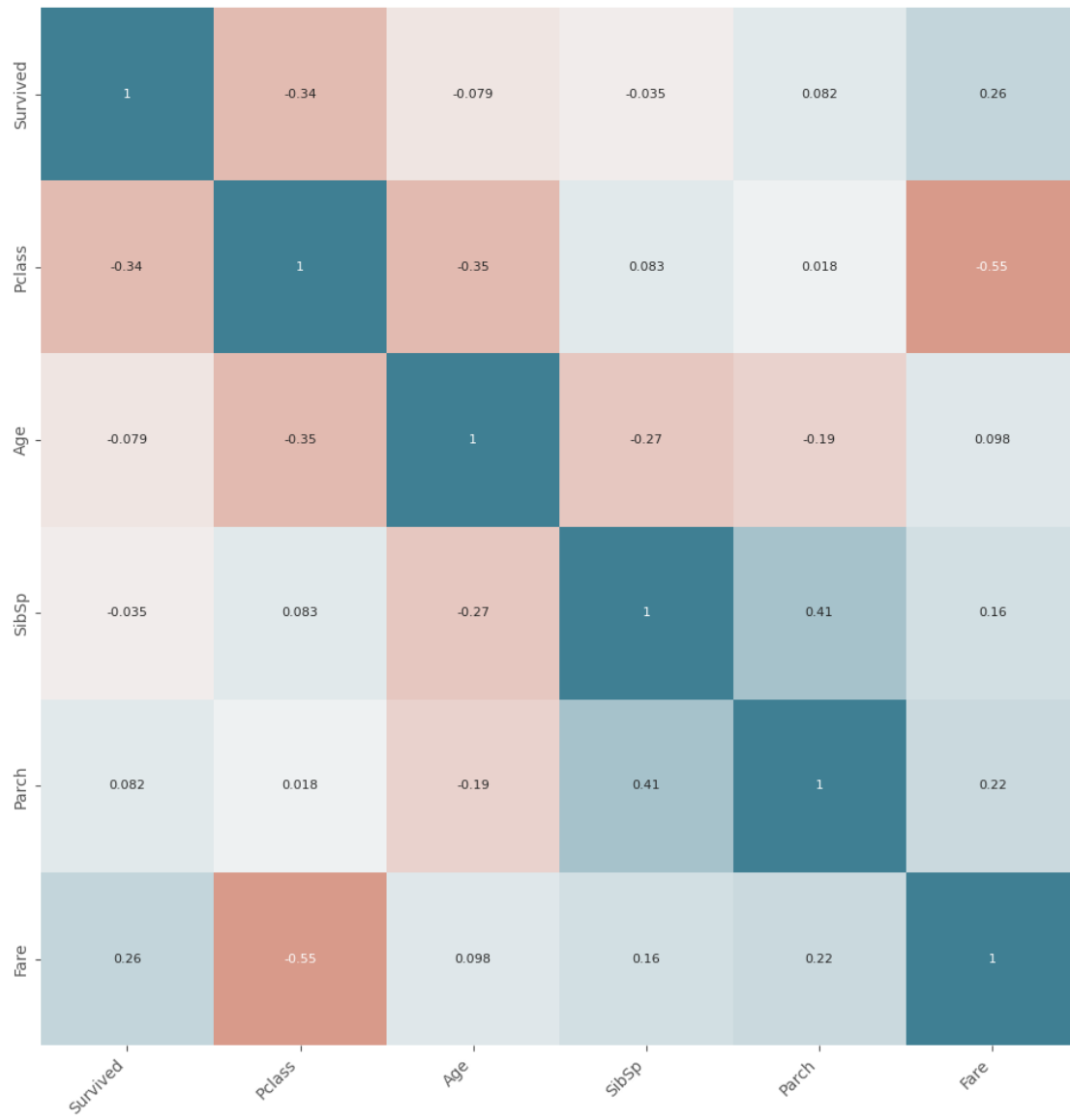
```

	Survived	Pclass	Age	SibSp	Parch	Fare
count	891.000	891.000	891.000	891.000	891.000	891.000
mean	0.384	2.309	29.385	0.523	0.382	32.204
std	0.487	0.836	13.260	1.103	0.806	49.693
min	0.000	1.000	0.420	0.000	0.000	0.000
25%	0.000	2.000	21.000	0.000	0.000	7.910
50%	0.000	3.000	30.000	0.000	0.000	14.454
75%	1.000	3.000	35.000	1.000	0.000	31.000
max	1.000	3.000	80.000	8.000	6.000	512.329

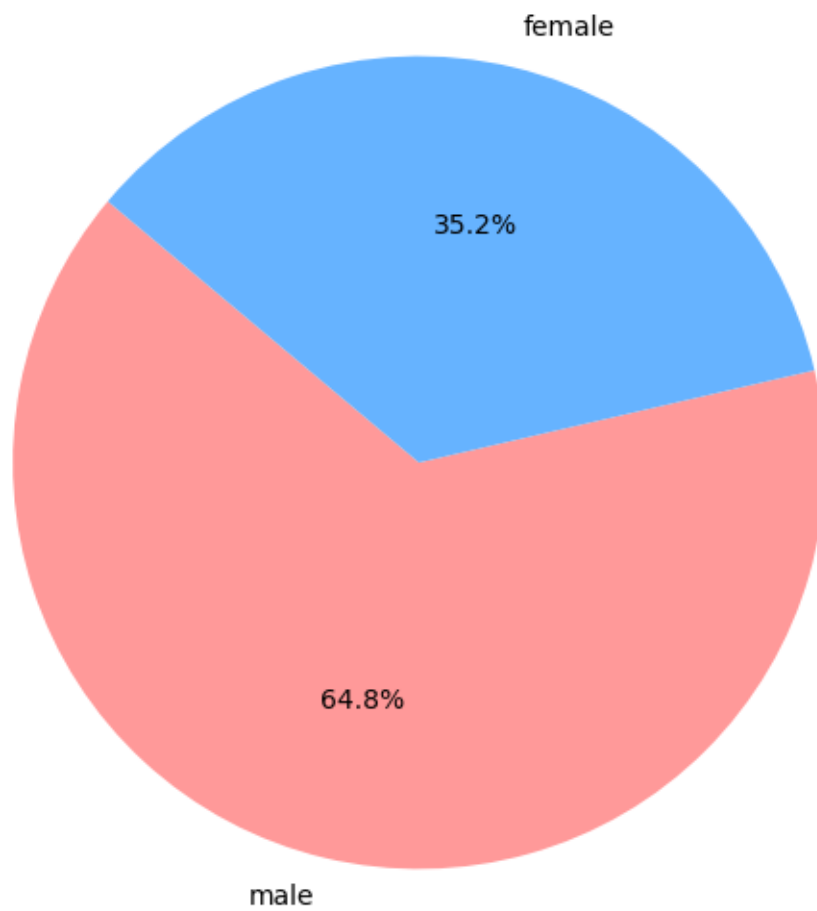
```

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 891 entries, 0 to 890
Data columns (total 6 columns):
#   Column      Non-Null Count  Dtype
---  -
0   Survived    891 non-null    int64
1   Pclass      891 non-null    int64
2   Age         891 non-null    float64
3   SibSp       891 non-null    int64
4   Parch       891 non-null    int64
5   Fare        891 non-null    float64
dtypes: float64(2), int64(4)
memory usage: 41.9 KB

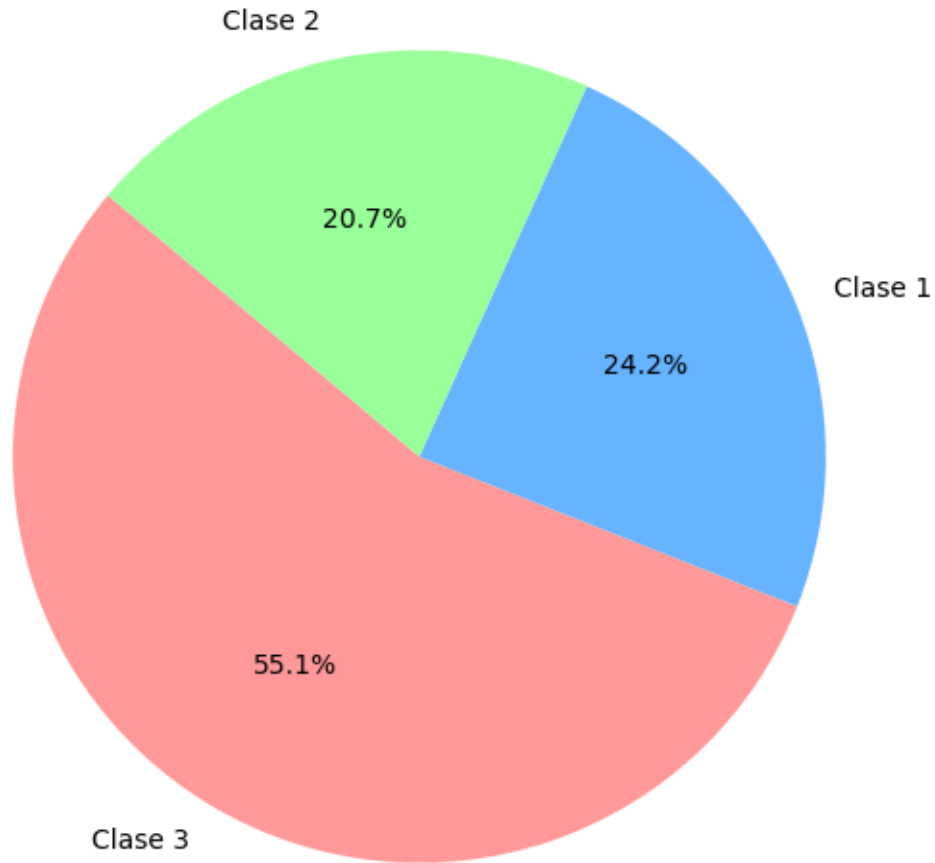
```



Distribución de pasajeros según género



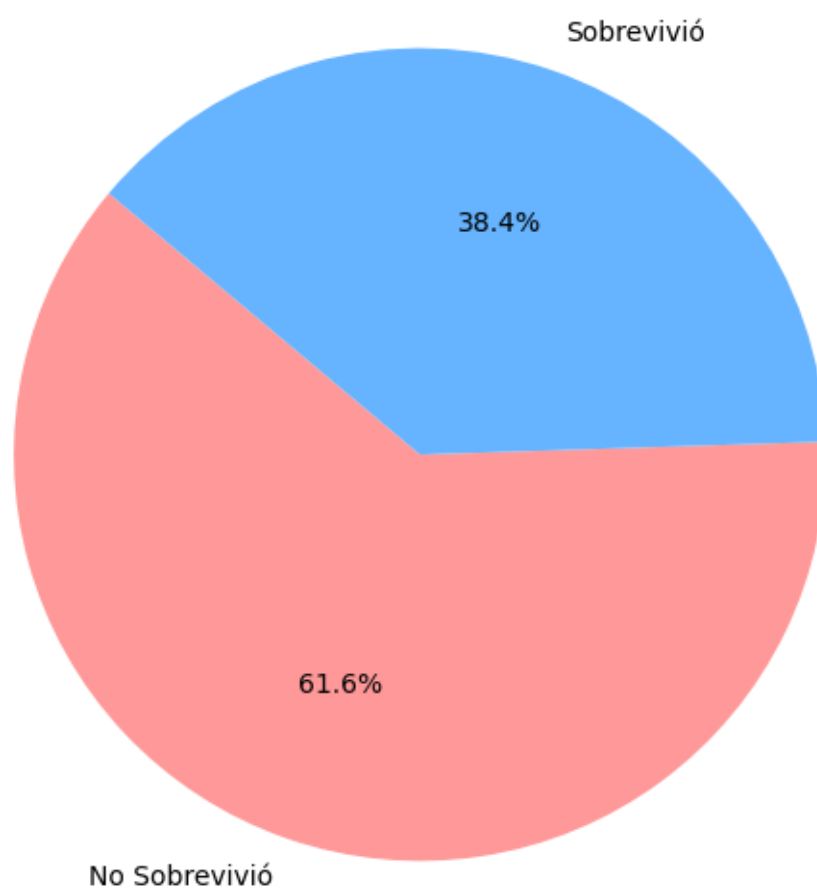
Distribución de Pasajeros según Clase



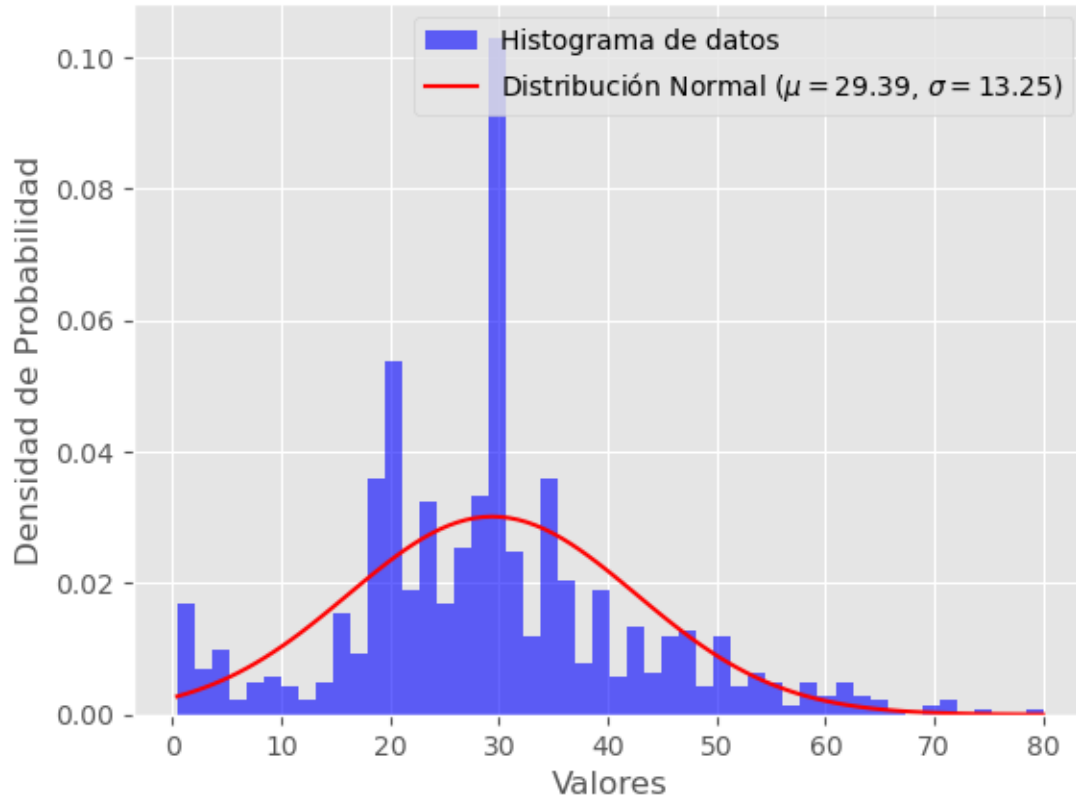
Dado que la clase, a pesar de estar determinada con un valor numérico, se trata de una variable categórica, carece de sentido analizarla por su distribución estadística. Una mejor forma de representar esa información puede ser con gráficos especializados en mostrar variables categóricas.

Como podemos observar, la gran mayoría de pasajeros se encontraban en **tercera clase**.

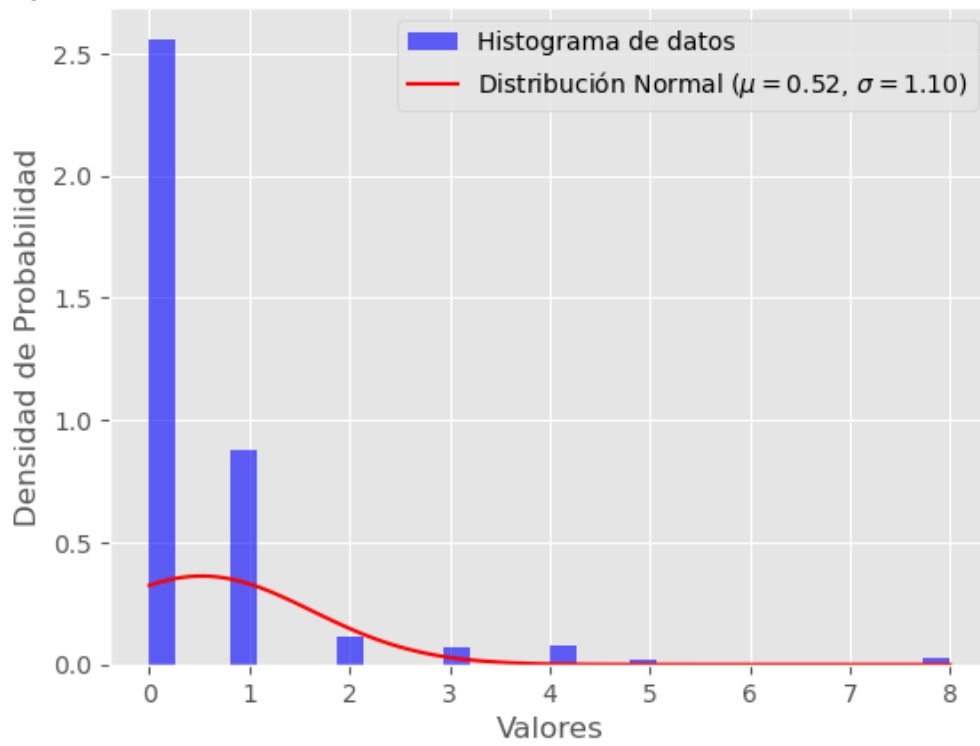
Distribución de pasajeros sobrevivientes



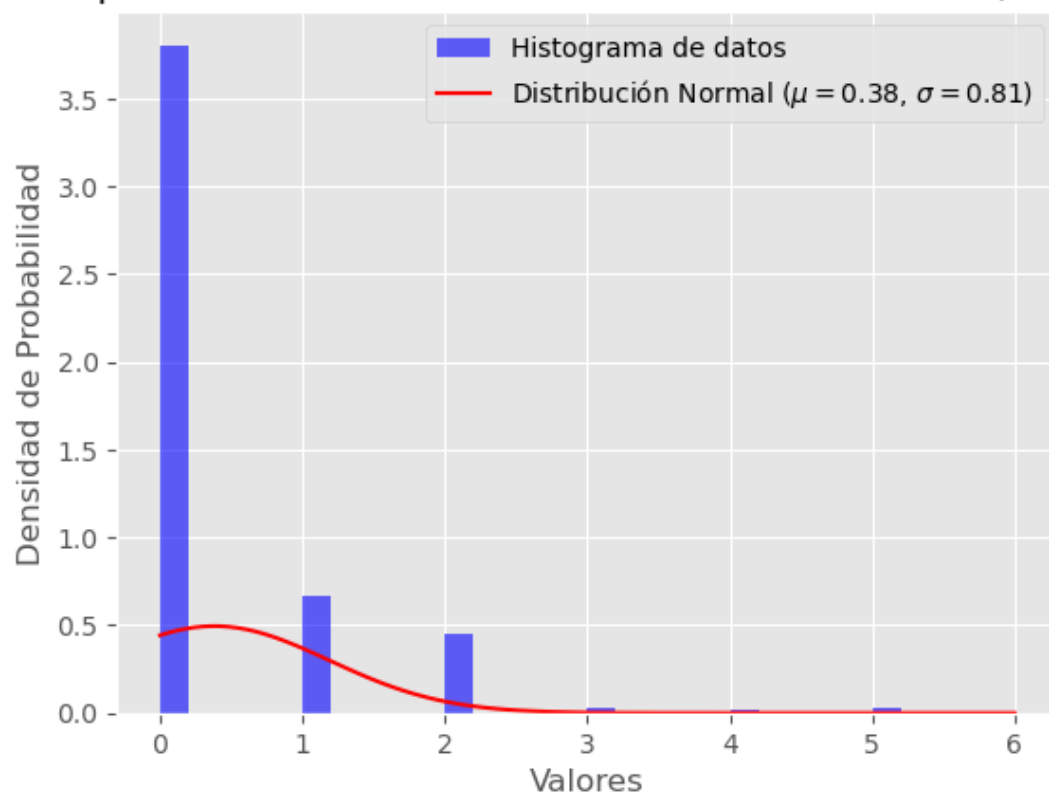
Aproximación a la Distribución Normal de Edad



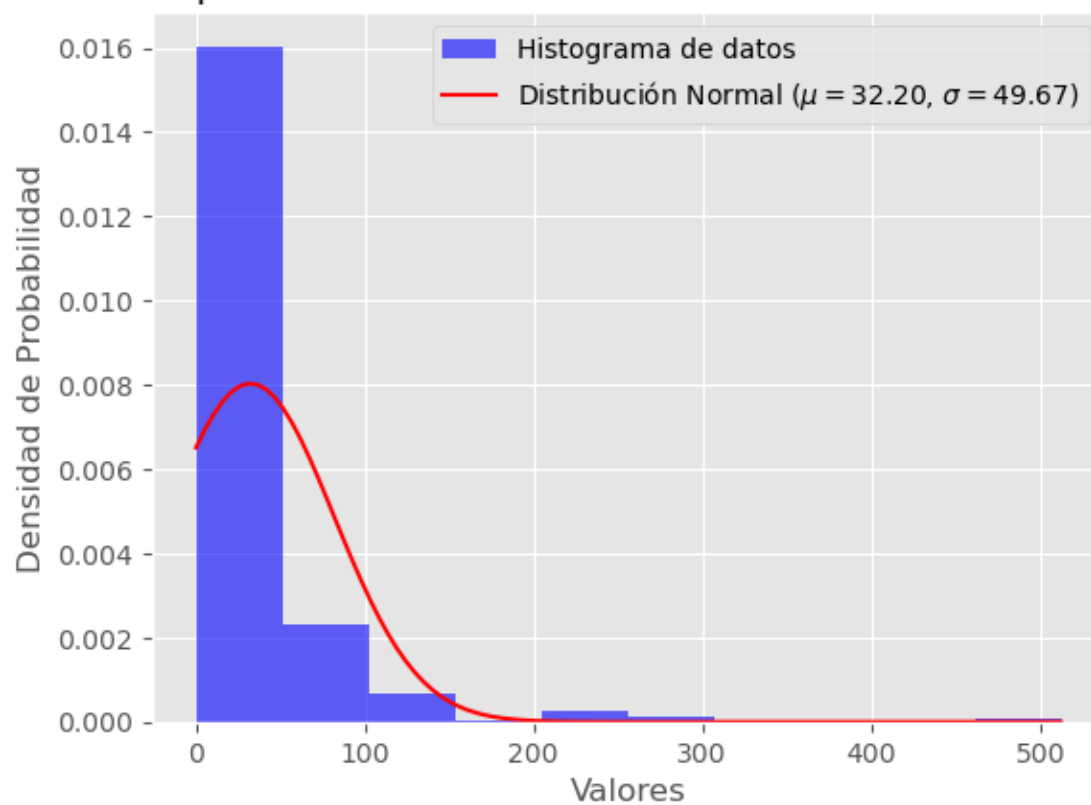
Aproximación a la Distribución Normal de Hermanos/Esposos

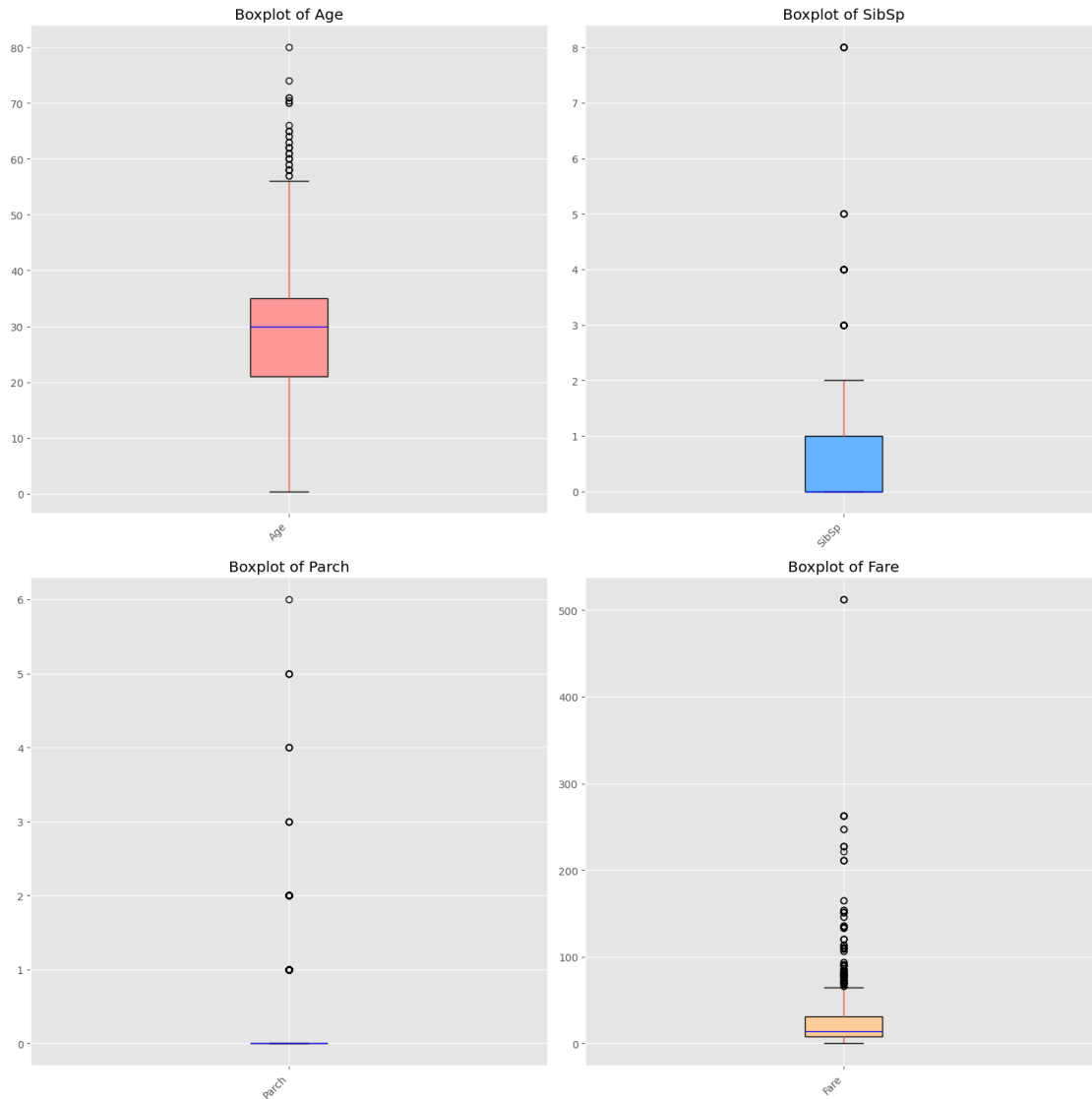


Aproximación a la Distribución Normal de Padres/Hijos



Aproximación a la Distribución Normal en Tarifa





1.1 Equipo:

- Coconi Dafne
- Cortés López
- Sánchez Erik
- Villegas Getsemaní

Ejemplo de grafico interactivo con plotly

<IPython.core.display.HTML object>

Comando para generar reporte PDF

Requirement already satisfied: nbconvert in c:\programdata\miniconda3\lib\site-packages (7.16.6)

Requirement already satisfied: beautifulsoup4 in c:\users\ville\appdata\roaming\python\python311\site-packages (from nbconvert) (4.12.3)

Requirement already satisfied: bleach!=5.0.0 in c:\users\ville\appdata\roaming\python\python311\site-packages (from bleach[css]!=5.0.0->nbconvert) (6.1.0)

Requirement already satisfied: defusedxml in c:\users\ville\appdata\roaming\python\python311\site-packages (from nbconvert) (0.7.1)

Requirement already satisfied: jinja2>=3.0 in c:\users\ville\appdata\roaming\python\python311\site-packages (from nbconvert) (3.1.4)

Requirement already satisfied: jupyter-core>=4.7 in c:\users\ville\appdata\roaming\python\python311\site-packages (from nbconvert) (5.7.2)

Requirement already satisfied: jupyterlab-pygments in c:\users\ville\appdata\roaming\python\python311\site-packages (from nbconvert) (0.3.0)

Requirement already satisfied: markupsafe>=2.0 in c:\users\ville\appdata\roaming\python\python311\site-packages (from nbconvert) (2.1.5)

Requirement already satisfied: mistune<4,>=2.0.3 in c:\users\ville\appdata\roaming\python\python311\site-packages (from nbconvert) (3.0.2)

Requirement already satisfied: nbclient>=0.5.0 in c:\users\ville\appdata\roaming\python\python311\site-packages (from nbconvert) (0.10.0)

Requirement already satisfied: nbformat>=5.7 in c:\users\ville\appdata\roaming\python\python311\site-packages (from nbconvert) (5.10.4)

Requirement already satisfied: packaging in c:\programdata\miniconda3\lib\site-packages (from nbconvert) (23.1)

Requirement already satisfied: pandocfilters>=1.4.1 in c:\users\ville\appdata\roaming\python\python311\site-packages (from nbconvert) (1.5.1)

Requirement already satisfied: pygments>=2.4.1 in c:\users\ville\appdata\roaming\python\python311\site-packages (from nbconvert) (2.18.0)

Requirement already satisfied: traitlets>=5.1 in c:\users\ville\appdata\roaming\python\python311\site-packages (from nbconvert) (5.14.3)

Requirement already satisfied: six>=1.9.0 in c:\programdata\miniconda3\lib\site-packages (from bleach!=5.0.0->bleach[css]!=5.0.0->nbconvert) (1.16.0)

Requirement already satisfied: webencodings in c:\users\ville\appdata\roaming\python\python311\site-packages (from bleach!=5.0.0->bleach[css]!=5.0.0->nbconvert) (0.5.1)

Requirement already satisfied: tinycss2<1.3,>=1.1.0 in
c:\programdata\miniconda3\lib\site-packages (from bleach[css]!=5.0.0->nbconvert)
(1.2.1)

Requirement already satisfied: platformdirs>=2.5 in
c:\programdata\miniconda3\lib\site-packages (from jupyter-core>=4.7->nbconvert)
(3.10.0)

Requirement already satisfied: pywin32>=300 in
c:\users\ville\appdata\roaming\python\python311\site-packages (from jupyter-
core>=4.7->nbconvert) (306)

Requirement already satisfied: jupyter-client>=6.1.12 in
c:\users\ville\appdata\roaming\python\python311\site-packages (from
nbclient>=0.5.0->nbconvert) (8.6.3)

Requirement already satisfied: fastjsonschema>=2.15 in
c:\users\ville\appdata\roaming\python\python311\site-packages (from
nbformat>=5.7->nbconvert) (2.20.0)

Requirement already satisfied: jsonschema>=2.6 in
c:\users\ville\appdata\roaming\python\python311\site-packages (from
nbformat>=5.7->nbconvert) (4.23.0)

Requirement already satisfied: soupsieve>1.2 in
c:\users\ville\appdata\roaming\python\python311\site-packages (from
beautifulsoup4->nbconvert) (2.6)

Requirement already satisfied: attrs>=22.2.0 in
c:\users\ville\appdata\roaming\python\python311\site-packages (from
jsonschema>=2.6->nbformat>=5.7->nbconvert) (24.2.0)

Requirement already satisfied: jsonschema-specifications>=2023.03.6 in
c:\users\ville\appdata\roaming\python\python311\site-packages (from
jsonschema>=2.6->nbformat>=5.7->nbconvert) (2023.12.1)

Requirement already satisfied: referencing>=0.28.4 in
c:\users\ville\appdata\roaming\python\python311\site-packages (from
jsonschema>=2.6->nbformat>=5.7->nbconvert) (0.35.1)

Requirement already satisfied: rpds-py>=0.7.1 in
c:\users\ville\appdata\roaming\python\python311\site-packages (from
jsonschema>=2.6->nbformat>=5.7->nbconvert) (0.20.0)

Requirement already satisfied: python-dateutil>=2.8.2 in
c:\programdata\miniconda3\lib\site-packages (from jupyter-
client>=6.1.12->nbclient>=0.5.0->nbconvert) (2.8.2)

Requirement already satisfied: pyzmq>=23.0 in
c:\users\ville\appdata\roaming\python\python311\site-packages (from jupyter-
client>=6.1.12->nbclient>=0.5.0->nbconvert) (26.2.0)

Requirement already satisfied: tornado>=6.2 in
c:\programdata\miniconda3\lib\site-packages (from jupyter-
client>=6.1.12->nbclient>=0.5.0->nbconvert) (6.3.3)

Note: you may need to restart the kernel to use updated packages.

WARNING: Skipping c:\ProgramData\miniconda3\Lib\site-packages\networkx-3.1.dist-
info due to invalid metadata entry 'name'

WARNING: Skipping c:\ProgramData\miniconda3\Lib\site-packages\networkx-3.1.dist-
info due to invalid metadata entry 'name'

WARNING: Skipping c:\ProgramData\miniconda3\Lib\site-packages\networkx-3.1.dist-info due to invalid metadata entry 'name'
WARNING: Skipping c:\ProgramData\miniconda3\Lib\site-packages\networkx-3.1.dist-info due to invalid metadata entry 'name'
WARNING: Skipping c:\ProgramData\miniconda3\Lib\site-packages\networkx-3.1.dist-info due to invalid metadata entry 'name'
WARNING: Skipping c:\ProgramData\miniconda3\Lib\site-packages\networkx-3.1.dist-info due to invalid metadata entry 'name'

Requirement already satisfied: pandoc in c:\programdata\miniconda3\lib\site-packages (2.4)

Requirement already satisfied: plumbum in c:\programdata\miniconda3\lib\site-packages (from pandoc) (1.9.0)

Requirement already satisfied: ply in c:\programdata\miniconda3\lib\site-packages (from pandoc) (3.11)

Requirement already satisfied: pywin32 in c:\users\ville\appdata\roaming\python\python311\site-packages (from plumbum->pandoc) (306)

Note: you may need to restart the kernel to use updated packages.

WARNING: Skipping c:\ProgramData\miniconda3\Lib\site-packages\networkx-3.1.dist-info due to invalid metadata entry 'name'
WARNING: Skipping c:\ProgramData\miniconda3\Lib\site-packages\networkx-3.1.dist-info due to invalid metadata entry 'name'
WARNING: Skipping c:\ProgramData\miniconda3\Lib\site-packages\networkx-3.1.dist-info due to invalid metadata entry 'name'
WARNING: Skipping c:\ProgramData\miniconda3\Lib\site-packages\networkx-3.1.dist-info due to invalid metadata entry 'name'
WARNING: Skipping c:\ProgramData\miniconda3\Lib\site-packages\networkx-3.1.dist-info due to invalid metadata entry 'name'
WARNING: Skipping c:\ProgramData\miniconda3\Lib\site-packages\networkx-3.1.dist-info due to invalid metadata entry 'name'

[NbConvertApp] Converting notebook TITANIC.ipynb to pdf
c:\ProgramData\miniconda3\share\jupyter\nbconvert\templates\latex\display_priority.j2:32: UserWarning: Your element with mimetype(s) dict_keys(['text/html']) is not able to be represented.

((*- endblock -*))

c:\ProgramData\miniconda3\share\jupyter\nbconvert\templates\latex\display_priority.j2:32: UserWarning: Your element with mimetype(s) dict_keys(['application/vnd.plotly.v1+json', 'text/html']) is not able to be represented.

((*- endblock -*))

[NbConvertApp] Support files will be in titanic_reports/TITANIC3_files\
[NbConvertApp] Making directory .\titanic_reports/TITANIC3_files\titanic_reports
[NbConvertApp] Writing 80169 bytes to notebook.tex
[NbConvertApp] Building PDF
[NbConvertApp] Running xelatex 3 times: ['xelatex', 'notebook.tex', '-quiet']
[NbConvertApp] Running bibtex 1 time: ['bibtex', 'notebook']

[NbConvertApp] WARNING | b had problems, most likely because there were no citations

[NbConvertApp] PDF successfully created

[NbConvertApp] Writing 358451 bytes to titanic_reports\TITANIC3.pdf

^C

[NbConvertApp] Converting notebook TITANIC.ipynb to pdf

c:\ProgramData\miniconda3\share\jupyter\nbconvert\templates\latex\display_priority.j2:32: UserWarning: Your element with mimetype(s) dict_keys(['text/html']) is not able to be represented.

((*- endblock -*))

c:\ProgramData\miniconda3\share\jupyter\nbconvert\templates\latex\display_priority.j2:32: UserWarning: Your element with mimetype(s) dict_keys(['application/vnd.plotly.v1+json', 'text/html']) is not able to be represented.

((*- endblock -*))

[NbConvertApp] Support files will be in TITANICr_files\

[NbConvertApp] Making directory .\TITANICr_files

[NbConvertApp] Writing 40751 bytes to notebook.tex

[NbConvertApp] Building PDF

[NbConvertApp] Running xelatex 3 times: ['xelatex', 'notebook.tex', '-quiet']

[NbConvertApp] Running bibtex 1 time: ['bibtex', 'notebook']

[NbConvertApp] WARNING | b had problems, most likely because there were no citations

[NbConvertApp] PDF successfully created

[NbConvertApp] Writing 316966 bytes to TITANICr.pdf