

統計諮詢

Yi-Ting Kuo

2025-02-28

目錄

一、敘述統計量	1
二、視覺化圖表分析	3
(一)Survived生還者	3
(二)Pclass 艙等	4
(三)Sex 性別	4
(四)Age 年紀	5
(五)SibSp 兄弟姊妹/配偶數	5
(六)Parch 父母或子女人數數	6
(七)Fare 票價	6
(八)Cabin 升級艙等與否	7
(九)Embarked 登船口	7
三、比較	8
(一)不同艙等的生還率	8
(二)男女的生還率	8
(三)不同登船口的生還率	9

一、敘述統計量

titanic

12 Variables 891 Observations

PassengerId

n	missing	distinct	Info	Mean	pMedian	Gmd	.05
891	0	891	1	446	446	297.3	45.5
.10	.25	.50	.75	.90	.95		
90.0	223.5	446.0	668.5	802.0	846.5		

lowest : 1 2 3 4 5, highest: 887 888 889 890 891

Survived

n	missing	distinct
891	0	2

Value	0	1
Frequency	549	342
Proportion	0.616	0.384

Pclass

n	missing	distinct
891	0	3

Value	1	2	3
Frequency	216	184	491
Proportion	0.242	0.207	0.551

Name

n	missing	distinct
891	0	891

lowest : Abbing, Mr. Anthony
highest: Yousseff, Mr. Gerious

Abbott, Mr. Rossmore Edward
Yrois, Miss. Henriette ("Mrs Harbeck")

Abbott, Mrs. Star
Zabour, Miss. Hil

Sex

n	missing	distinct
891	0	2

Value	female	male
Frequency	314	577
Proportion	0.352	0.648

Age

n	missing	distinct	Info	Mean	pMedian	Gmd	.05
714	177	88	0.999	29.7	29	16.21	4.00
.10	.25	.50	.75	.90	.95		
14.00	20.12	28.00	38.00	50.00	56.00		

lowest : 0.42 0.67 0.75 0.83 0.92, highest: 70 70.5 71 74 80

SibSp

n	missing	distinct	Info	Mean	pMedian	Gmd
891	0	7	0.669	0.523	0.5	0.823

Value	0	1	2	3	4	5	8
Frequency	608	209	28	16	18	5	7
Proportion	0.682	0.235	0.031	0.018	0.020	0.006	0.008

For the frequency table, variable is rounded to the nearest 0

Parch

n	missing	distinct	Info	Mean	pMedian	Gmd
891	0	7	0.556	0.3816	0	0.6259

Value	0	1	2	3	4	5	6
Frequency	678	118	80	5	4	5	1
Proportion	0.761	0.132	0.090	0.006	0.004	0.006	0.001

For the frequency table, variable is rounded to the nearest 0

Ticket

n	missing	distinct
---	---------	----------

891	0	681					
-----	---	-----	--	--	--	--	--

lowest : 110152 110413 110465 110564 110813
highest: W./C. 6608 W./C. 6609 W.E.P. 5734 W/C 14208 WE/P 5735

Fare

n	missing	distinct	Info	Mean	pMedian	Gmd	.05
891	0	248	1	32.2	19.6	36.78	7.225
.10	.25	.50	.75	.90	.95		
7.550	7.910	14.454	31.000	77.958	112.079		

lowest : 0 4.0125 5 6.2375 6.4375
highest: 227.525 247.521 262.375 263 512.329

Cabin

n	missing	distinct
891	0	2

Value No Yes
Frequency 687 204
Proportion 0.771 0.229

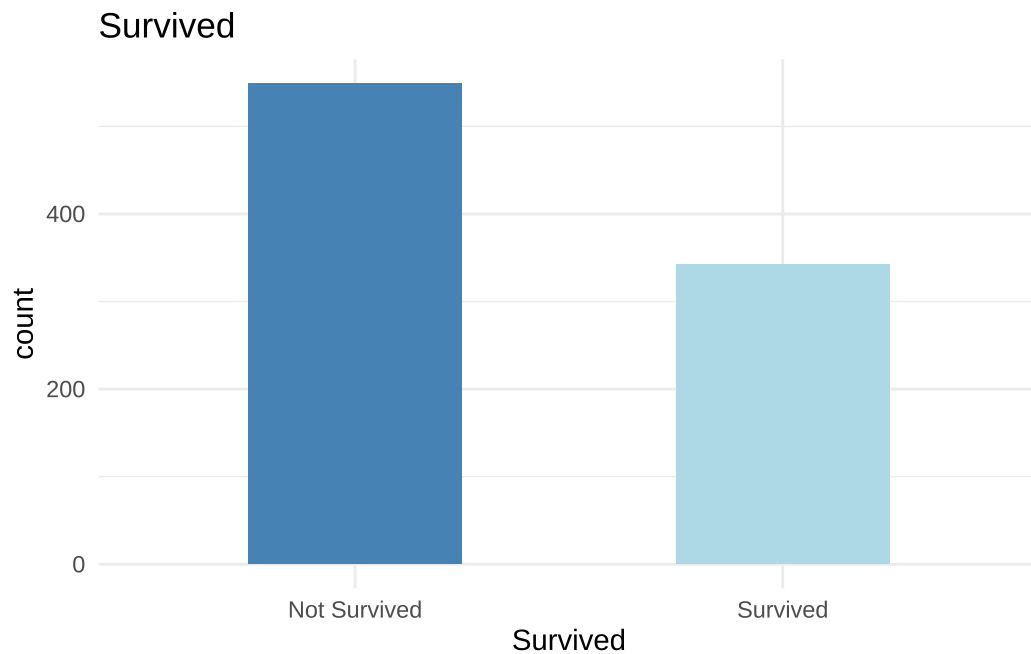
Embarked

n	missing	distinct
891	0	4

Value C Q S
Frequency 2 168 77 644
Proportion 0.002 0.189 0.086 0.723

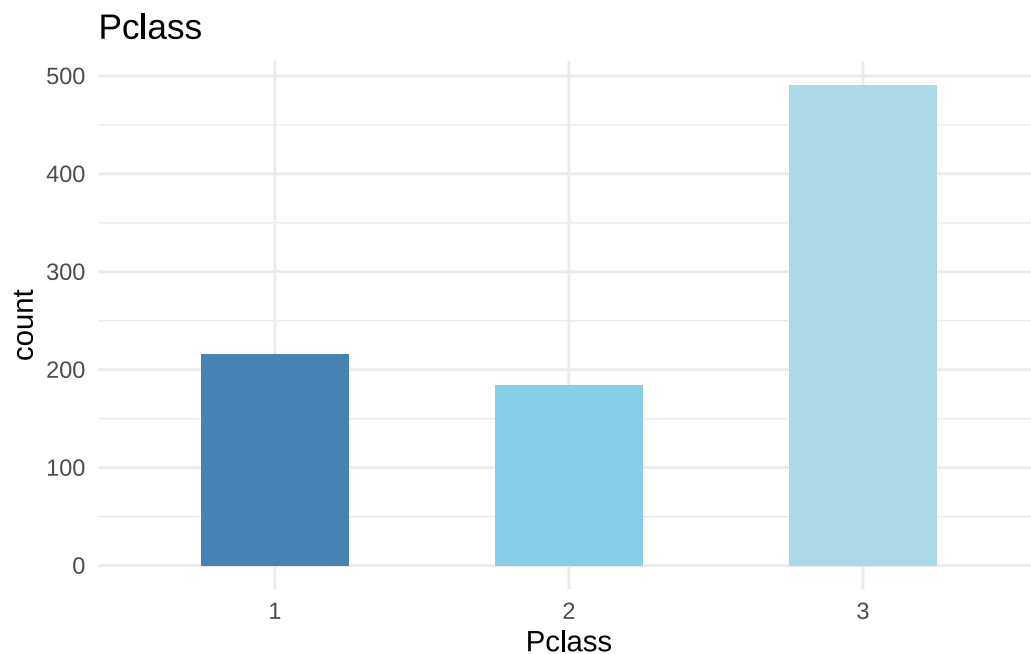
二、視覺化圖表分析

(一)Survived生還者



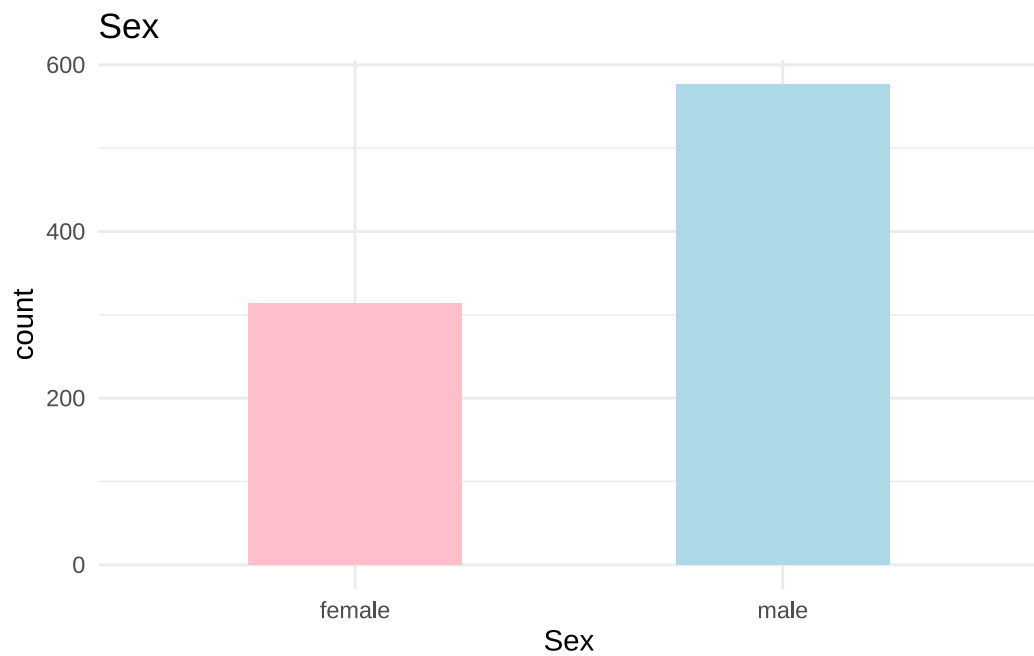
在這場事故中，未生還者的人數明顯高於生還者的人數，差距大約兩百餘人。

(二)Pclass 艙等



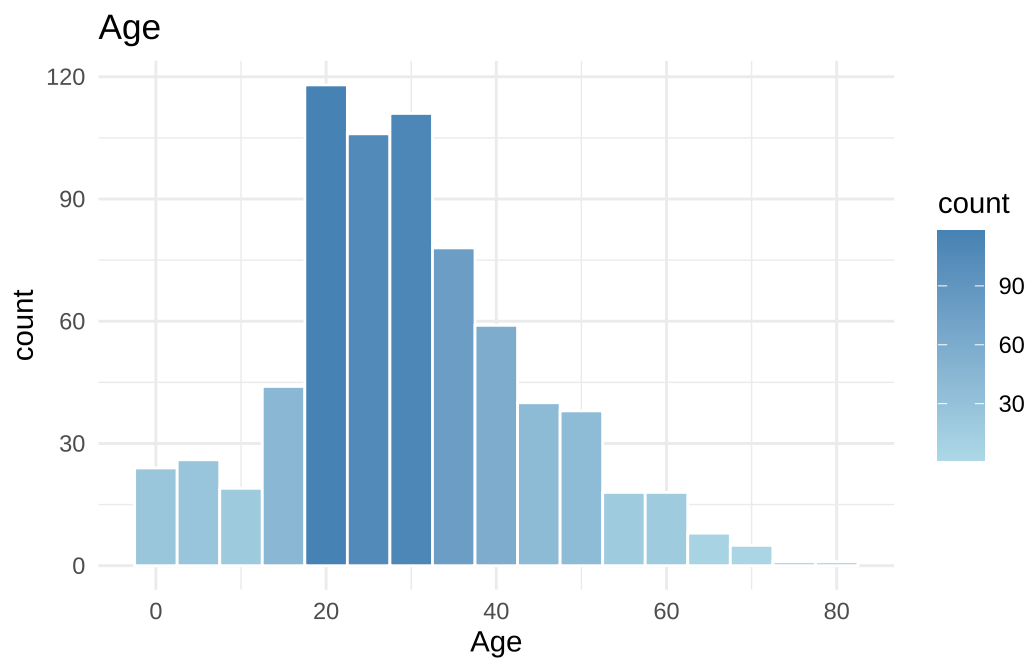
在第三個艙等的人數最多，因此判斷第三個艙等為票價較便宜的艙等，並在下一個章節中比較這三個艙等的存活率。

(三)Sex 性別



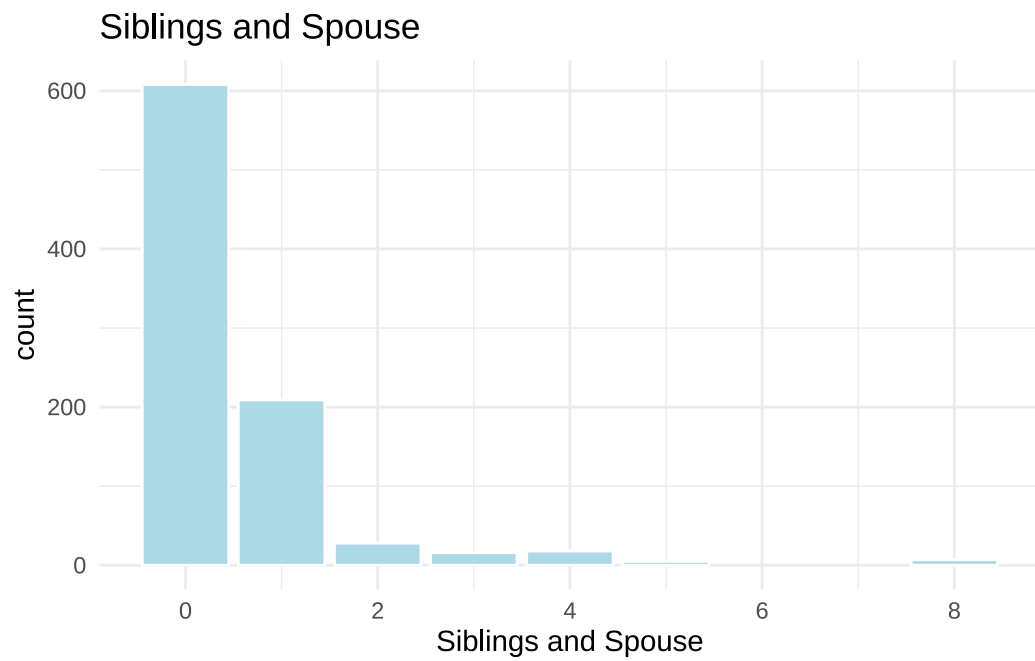
乘客中男女的比例大約為2：1。

(四)Age 年紀

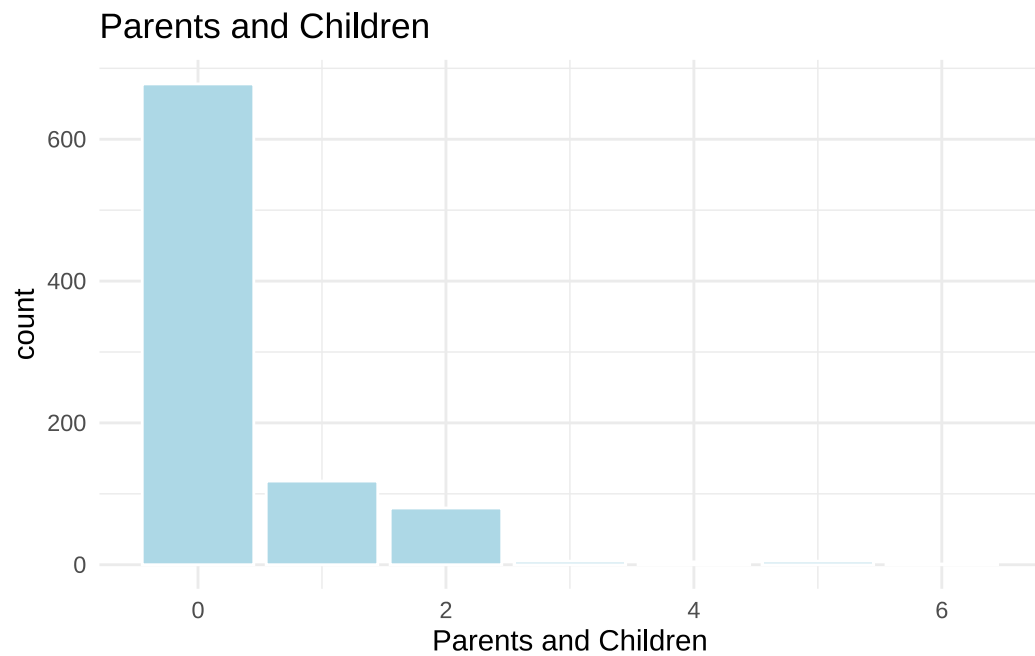


年紀分佈以20-40歲的青壯年人居多。

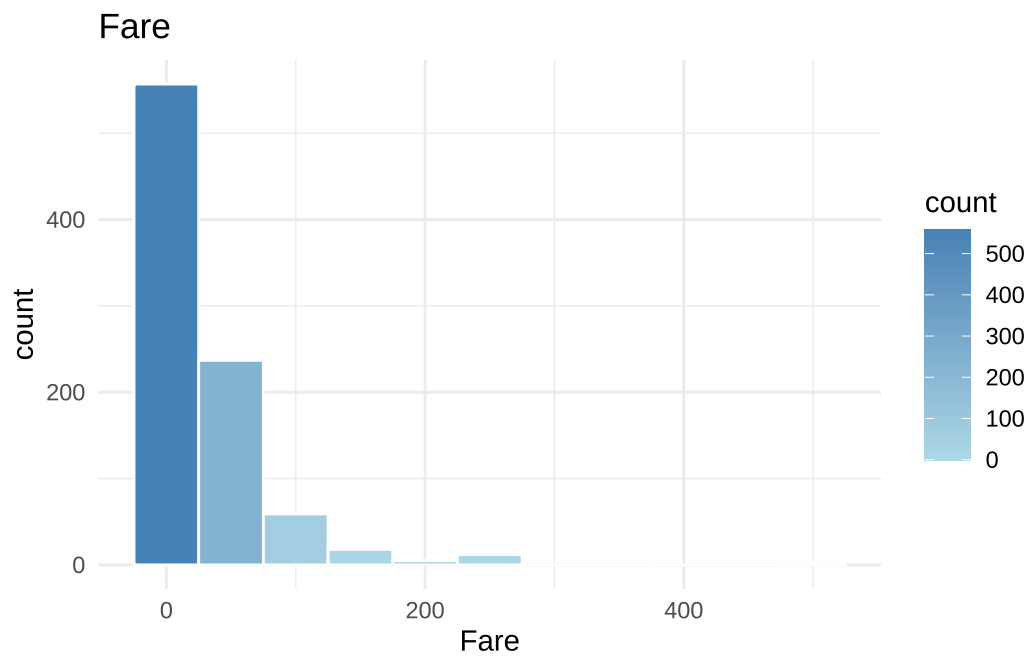
(五) SibSp 兄弟姊妹/配偶數



(六) Parch 父母或子女人數數

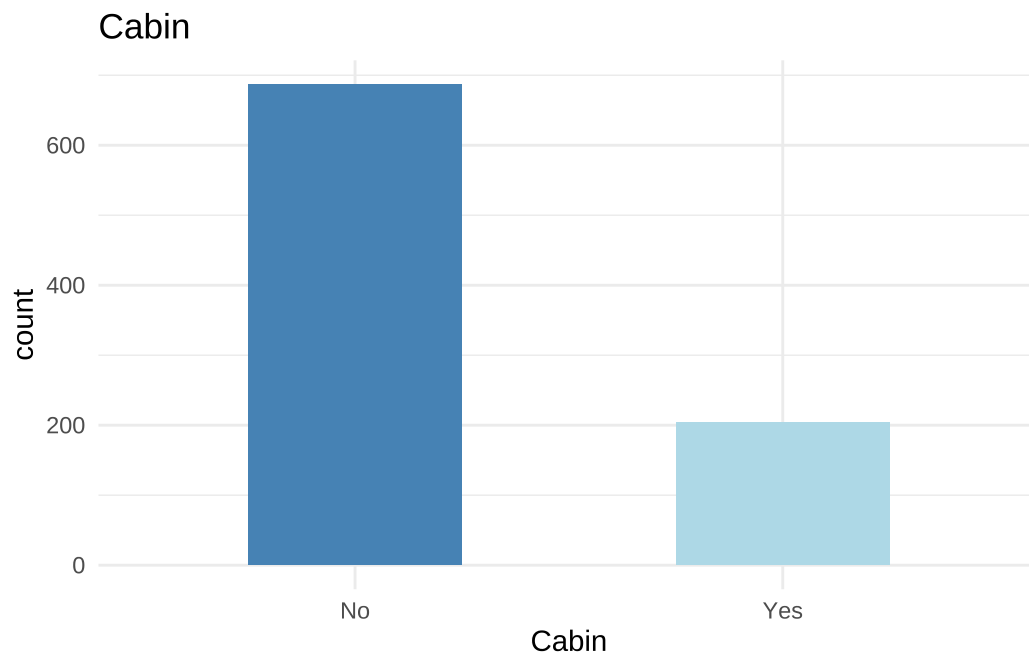


(七)Fare 票價

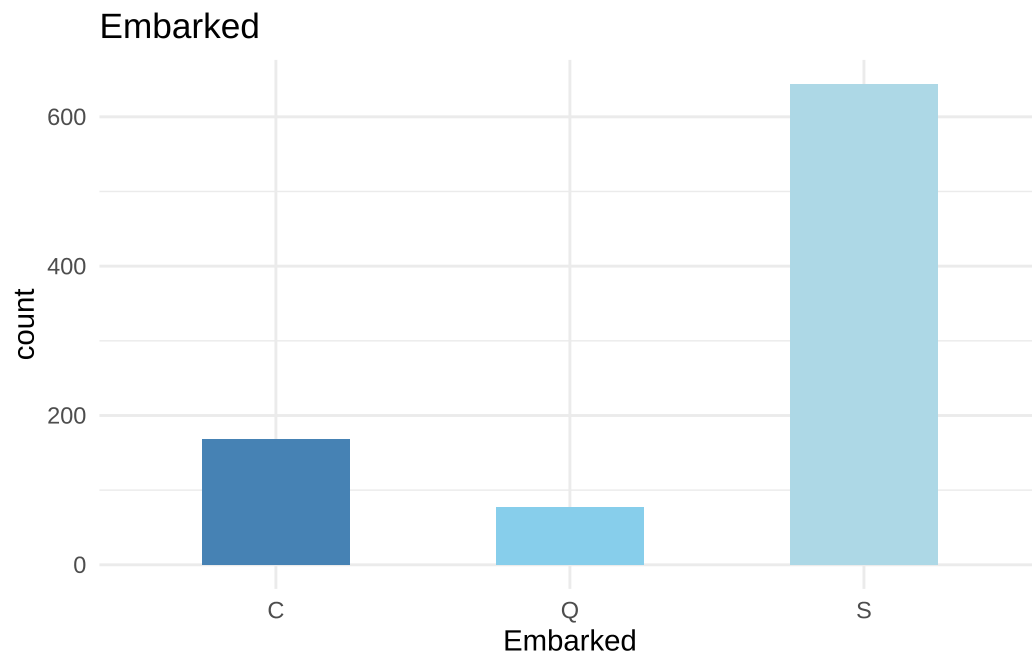


大約六成的乘客都選擇最低的票價，極少數的乘客回選擇較高的票價。

(八)Cabin 升級艙等與否

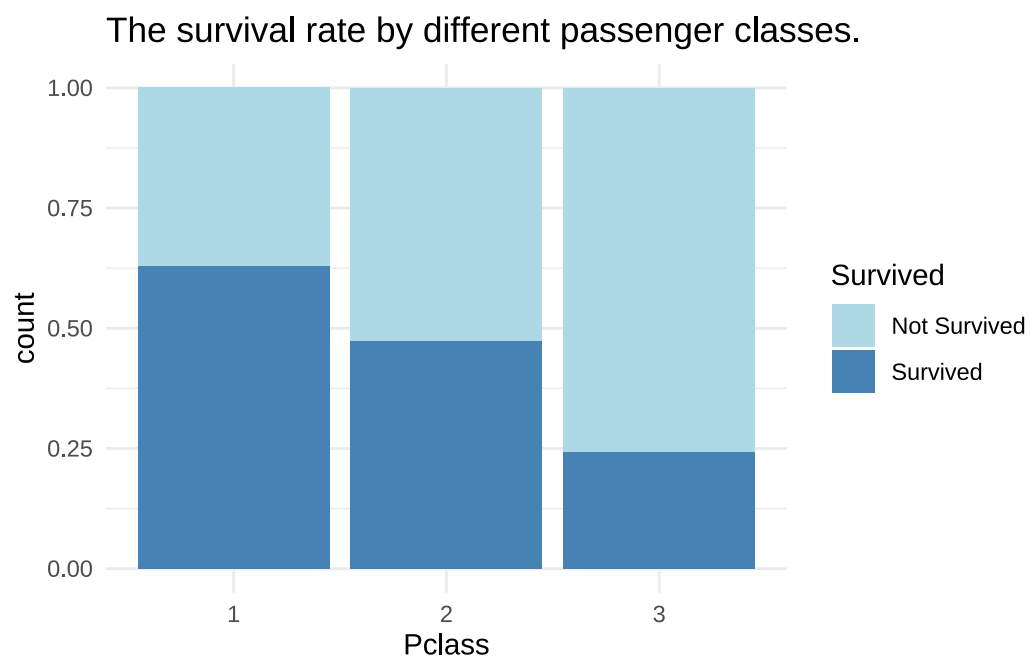


(九)Embarked 登船口



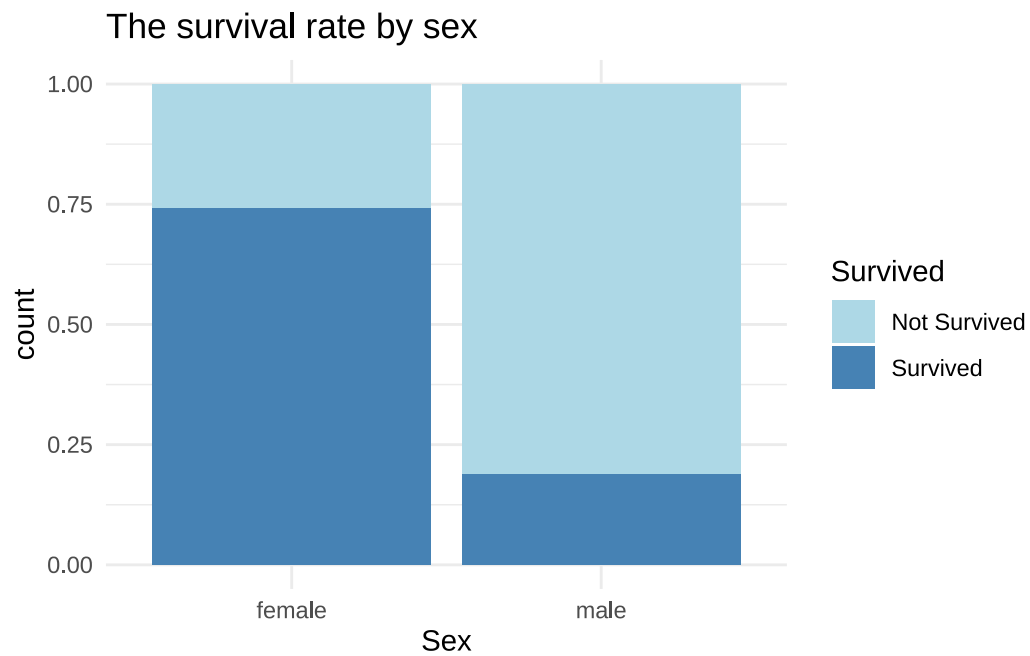
三、比較

(一)不同艙等的生還率



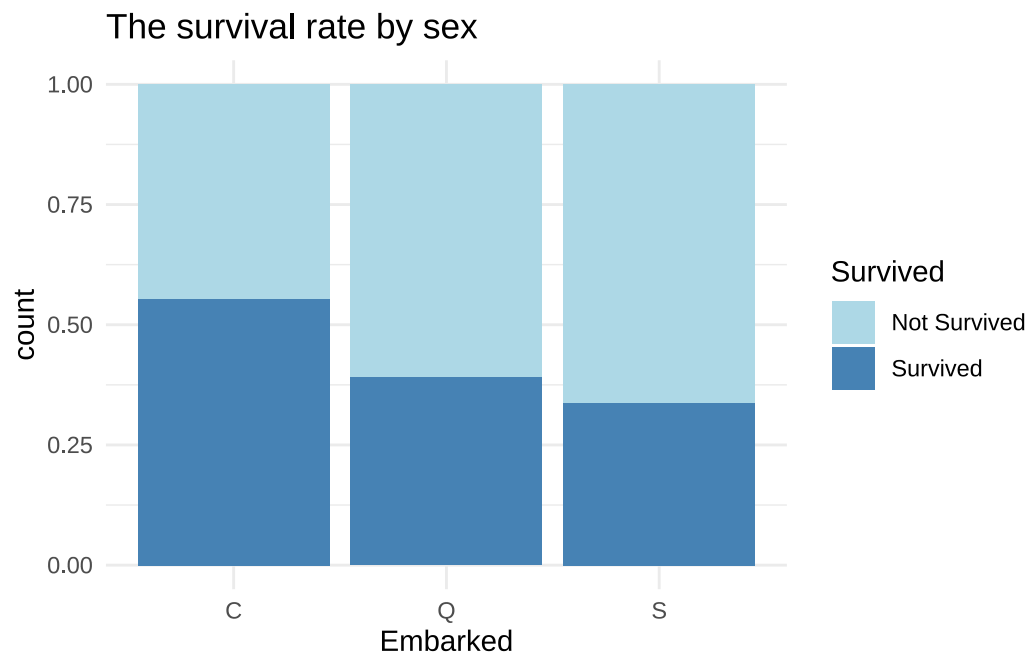
第一個艙等的生還率最高，第三個艙等的生還率最低，因此推論第一個艙等是較高級的艙等。

(二)男女的生還率



女性的生還率遠遠高於男性，或許我們可以推斷男性多為衝鋒搭救的一方，而女性為受保護的一方。

(三)不同登船口的生還率



雖然最多人從S登船口，但該登船口的生還率最低。