



รายงาน

Case Study

การวิเคราะห์ปัจจัยที่มีผลต่อการรอดชีวิตของผู้โดยสารบนเรือ Titanic โดยใช้การถดถอยโลจิสติก

เสนอ

ผศ.ดร.มายิอนิง อีสอ

ผศ.ดร.นุริน ตือเร๊ะ

ผศ.ดร.อารินดา มะอาลี

จัดทำโดย

นางสาว มัยมูน เต็งโต รหัสนักศึกษา 6520310134

MISS LIZA THEA รหัสนักศึกษา 6520310202

สาขาคณิตศาสตร์และวิทยาการคอมพิวเตอร์

คณะวิทยาศาสตร์และเทคโนโลยี

รายงานฉบับนี้เป็นส่วนหนึ่งของ

รายวิชา 747-341 ชุดวิชาการวิเคราะห์ข้อมูลและการนำเสนอข้อมูลเชิงทัศน์

ภาคเรียนที่ 1 ปีการศึกษา 2567

มหาวิทยาลัยสงขลานครินทร์ วิทยาเขตปัตตานี

Case Study

1. หัวข้อ: การวิเคราะห์ปัจจัยที่มีผลต่อการรอดชีวิตของผู้โดยสารบนเรือ Titanic โดยใช้การถดถอยโลจิสติก

2. คำถามในการวิจัย

คำถามวิจัย : ปัจจัยใดที่มีผลต่อการรอดชีวิตของผู้โดยสารบนเรือ Titanic

3. วัตถุประสงค์ของการศึกษา

เพื่อสำรวจความสัมพันธ์ของตัวแปรอิสระต่างๆ ที่มีอิทธิพลต่อผลลัพธ์ของการรอดชีวิตของผู้โดยสารบนเรือ Titanic

4. ที่มาและความสำคัญ

เหตุผลที่เลือกทำเรื่องนี้ เนื่องจากเหตุการณ์การจมของเรือ Titanic เป็นเหตุการณ์ประวัติศาสตร์ที่มีผู้เสียชีวิตจำนวนมาก และมีข้อมูลที่สามารถวิเคราะห์ได้อย่างละเอียด การศึกษานี้มีความสำคัญเพราะช่วยทำความเข้าใจว่าปัจจัยใดที่มีผลต่อการรอดชีวิตของผู้โดยสาร เช่น ชั้นผู้โดยสาร เพศ อายุ เป็นต้น ปัญหาที่สำคัญคือการขาดความเข้าใจในปัจจัยที่มีผลต่อการรอดชีวิตในสถานการณ์วิกฤต ซึ่งสามารถนำไปปรับใช้กับการจัดการภัยพิบัติหรือเหตุการณ์ฉุกเฉินในอนาคต

5. รายละเอียดของตัวแปรที่เลือกใช้

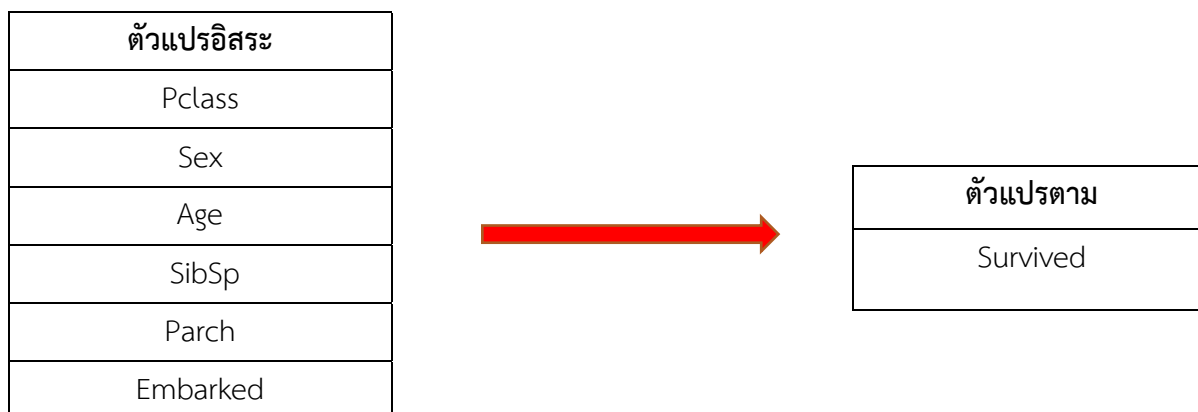
ข้อมูลทั้งหมดมี 891 รายและ 12 ตัวแปรโดยตัวแปรที่เลือกใช้มี 8 ตัวแปรตามตารางดังนี้ :

ตารางที่ 1 : คำอธิบายของตัวแปร

ตัวแปร	คำอธิบายของตัวแปร	ประเภทของตัวแปร	รายละเอียดของตัวแปร (คำอธิบายของตัวแปร)
Survived (ตัวแปรตาม)	การรอดชีวิตของผู้โดยสาร	ตัวแปรกลุ่ม	1 : รอด 0 : ไม่รอด
PassengerId	รหัสของผู้โดยสาร	ตัวแปรต่อเนื่อง	
Pclass	ชั้นผู้โดยสาร	ตัวแปรกลุ่ม	1 : ชั้น 1 2 : ชั้น 2 3 : ชั้น 3
Sex	เพศ	ตัวแปรกลุ่ม	male : ชาย female : หญิง
Age	อายุของผู้โดยสาร (ปี)	ตัวแปรต่อเนื่อง	
SibSp	จำนวนพี่น้องหรือคู่สมรสของผู้โดยสารบนเรือไททานิค	ตัวแปรต่อเนื่อง	
Parch	จำนวนพ่อแม่หรือลูกของผู้โดยสารบนเรือไททานิค	ตัวแปรต่อเนื่อง	

Embarked	ท่าเรือแห่งการลงเรือ	ตัวแปรกลุ่ม	C : Cherbourg, Q : Queenstown, S : Southampton
----------	----------------------	-------------	--

ตารางที่ 2 : ตัวแปรอิสระ 6 ตัวแปรและตัวแปรตาม 1 ตัวแปร



6. การจัดการข้อมูลและทำความสะอาดของข้อมูล

ข้อมูล Titanic-Dataset เก็บไว้ในตัวแปร dt1 โดยมีการเลือกใช้ library(epiDisplay) ,

library(readxl) ในการวิเคราะห์ข้อมูลโดยใช้ R

```
> str(dt1)
'data.frame':  891 obs. of  8 variables:
 $ PassengerId: int  1 2 3 4 5 6 7 8 9 10 ...
 $ Survived   : int  0 1 1 1 0 0 0 0 1 1 ...
 $ Pclass     : int  3 1 3 1 3 3 1 3 3 2 ...
 $ Sex        : chr  "male" "female" "female" "female" ...
 $ Age        : num  22 38 26 35 35 NA 54 2 27 14 ...
 $ SibSp      : int  1 1 0 1 0 0 0 3 0 1 ...
 $ Parch      : int  0 0 0 0 0 0 0 1 2 0 ...
 $ Embarked   : chr  "S" "C" "S" "S" ...
```

6.1 การตรวจสอบข้อมูลที่ซ้ำกัน

```
> dtl$checkId <- paste(dtl$PassengerId)
> dup <- duplicated(dtl$checkId)
> table(dup)
dup
FALSE
 891
> dtl[dup==T,]
[1] PassengerId Survived  Pclass      Sex      Age      SibSp
[7] Parch      Embarked  checkId
<0 rows> (or 0-length row.names)
> dtl <- dtl[dup==F,]
> str(dtl)
'data.frame': 891 obs. of 9 variables:
 $ PassengerId: int  1 2 3 4 5 6 7 8 9 10 ...
 $ Survived   : int  0 1 1 1 0 0 0 0 1 1 ...
 $ Pclass     : int  3 1 3 1 3 3 1 3 3 2 ...
 $ Sex        : chr  "male" "female" "female" "female" ...
 $ Age        : num  22 38 26 35 35 NA 54 2 27 14 ...
 $ SibSp      : int  1 1 0 1 0 0 0 3 0 1 ...
 $ Parch      : int  0 0 0 0 0 0 0 1 2 0 ...
 $ Embarked   : chr  "S" "C" "S" "S" ...
 $ checkId    : chr  "1" "2" "3" "4" ...
```

จากผลการวิเคราะห์พบว่าไม่มีชุดข้อมูลที่ซ้ำกัน

6.2 การตรวจสอบข้อมูลสูญหาย

ตรวจสอบข้อมูลสูญหายของแต่ละตัวแปร

```
> dtl[dtl == ""] <- NA
> missing_value <- is.na(dtl)
> missing_count <- colSums(missing_value)
> missing_count
PassengerId  Survived    Pclass      Sex      Age      SibSp
           0           0          0          0      177          0
   Parch  Embarked  checkId
           0           2          0
```

จากผลการวิเคราะห์พบว่าตัวแปรที่มีข้อมูลสูญหายคือ ตัวแปร Age และ Embarked จำนวน 177 ราย และ 2 ราย ตามลำดับ

- ทำความสะอาดของตัวแปร Age

เนื่องจากตัวแปรอายุมีข้อมูลสูญหายจำนวน 177 ราย ดังนั้นเราจะแทนข้อมูลสูญหายด้วยค่าเฉลี่ยของอายุ

```
> mean_age <- mean(dtl$Age, na.rm = TRUE)
> mean_age
[1] 29.69912
> dtl$age <- ifelse(is.na(dtl$Age), mean_age, dtl$Age)
> str(dtl)
'data.frame': 891 obs. of 10 variables:
 $ PassengerId: int 1 2 3 4 5 6 7 8 9 10 ...
 $ Survived : int 0 1 1 1 0 0 0 0 1 1 ...
 $ Pclass : int 3 1 3 1 3 3 1 3 3 2 ...
 $ Sex : chr "male" "female" "female" "female" ...
 $ Age : num 22 38 26 35 35 NA 54 2 27 14 ...
 $ SibSp : int 1 1 0 1 0 0 0 3 0 1 ...
 $ Parch : int 0 0 0 0 0 0 0 1 2 0 ...
 $ Embarked : chr "S" "C" "S" "S" ...
 $ checkId : chr "1" "2" "3" "4" ...
 $ age : num 22 38 26 35 35 ...
```

- ทำความสะอาดของตัวแปร Embarked

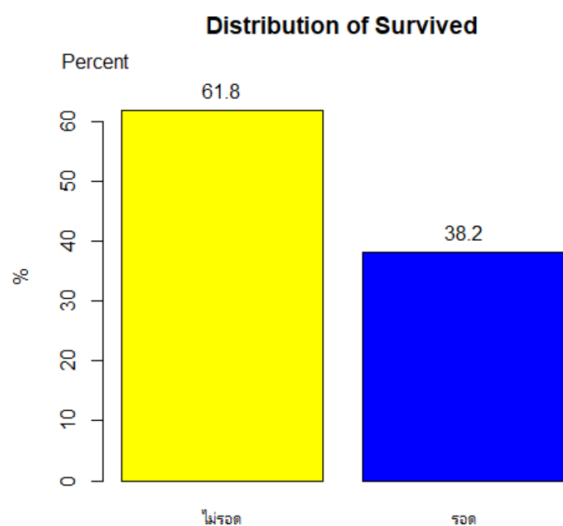
เนื่องจากข้อมูลสูญหายของตัวแปร Embarked มี 2 ราย ดังนั้นเราจะทำความสะอาดข้อมูลโดยลบข้อมูลที่สูญหายออกจากชุดข้อมูล dt1

```
> missing_embarked_count <- sum(is.na(dtl$Embarked))
> missing_embarked_count
[1] 2
> # drop 2 columns
> dtl <- dtl[!is.na(dtl$Embarked), ]
> str(dtl)
'data.frame': 889 obs. of 10 variables:
 $ PassengerId: int 1 2 3 4 5 6 7 8 9 10 ...
 $ Survived : int 0 1 1 1 0 0 0 0 1 1 ...
 $ Pclass : int 3 1 3 1 3 3 1 3 3 2 ...
 $ Sex : chr "male" "female" "female" "female" ...
 $ Age : num 22 38 26 35 35 NA 54 2 27 14 ...
 $ SibSp : int 1 1 0 1 0 0 0 3 0 1 ...
 $ Parch : int 0 0 0 0 0 0 0 1 2 0 ...
 $ Embarked : chr "S" "C" "S" "S" ...
 $ checkId : chr "1" "2" "3" "4" ...
 $ age : num 22 38 26 35 35 ...
```

7. ผลการวิเคราะห์ข้อมูลเบื้องต้น

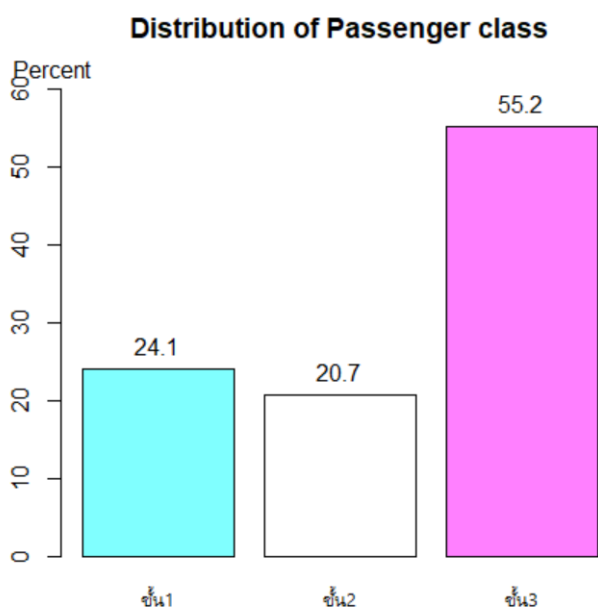
7.1. การวิเคราะห์ข้อมูล 1 ตัวแปร

Survived (ตัวแปรตาม)



จากกราฟพบว่า ผู้โดยสารที่อยู่บนเรือไททานิคมีผู้รอดชีวิต คิดเป็น 38.2% ในขณะที่มีผู้เสียชีวิตคิดเป็น 61.8%

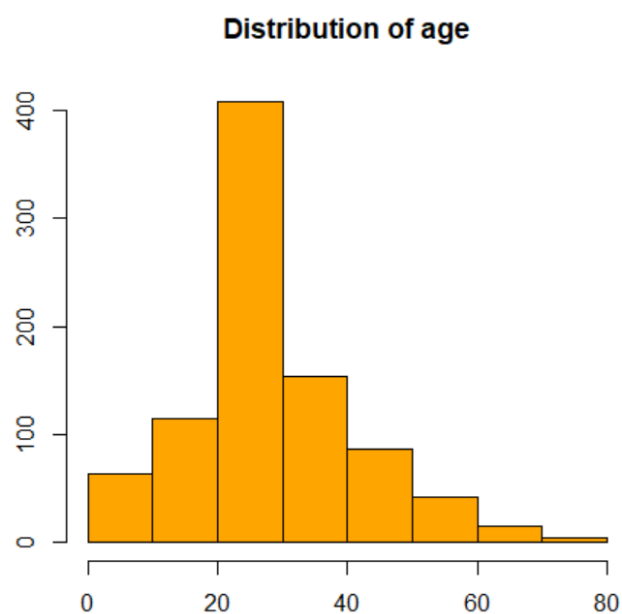
ตัวแปร Pclass



จากกราฟพบว่า ผู้โดยสารที่อยู่บนเรือไททานิคส่วนใหญ่อยู่ชั้นที่ 3 รองลงมาคือชั้นที่ 1 และ ชั้นที่ 2 คิดเป็น 55.2% , 24.1% และ 20.7% ตามลำดับ

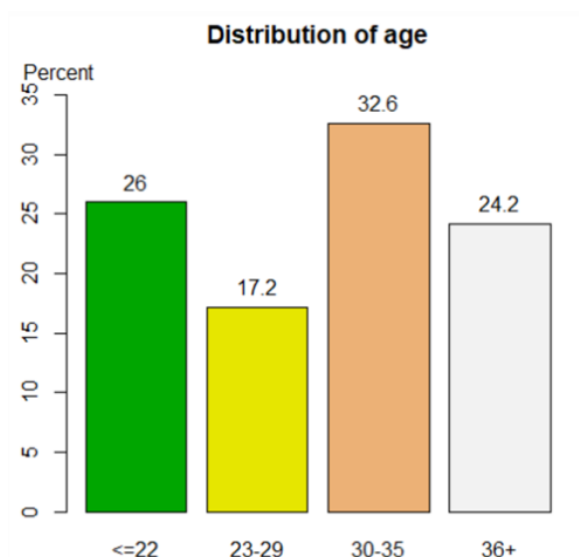
ตัวแปร age

#Before



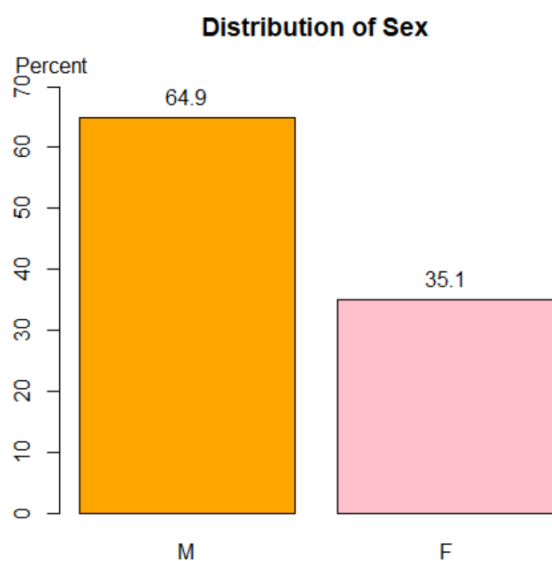
จากกราฟ Histogram พบว่า ข้อมูลไม่มีการแจกแจงแบบปกติ ดังนั้นเราจะจัดกลุ่มข้อมูลของตัวแปร age โดยใช้ quantile เป็นหลัก

#After



จากกราฟพบว่า ผู้โดยสารที่อยู่บนเรือไททานิคส่วนใหญ่มีอายุ 30 - 35 ปี และรองลงมาคือ มีอายุน้อยกว่าหรือเท่ากับ 22 ปี คิดเป็น 32.6% และ 26% ตามลำดับ

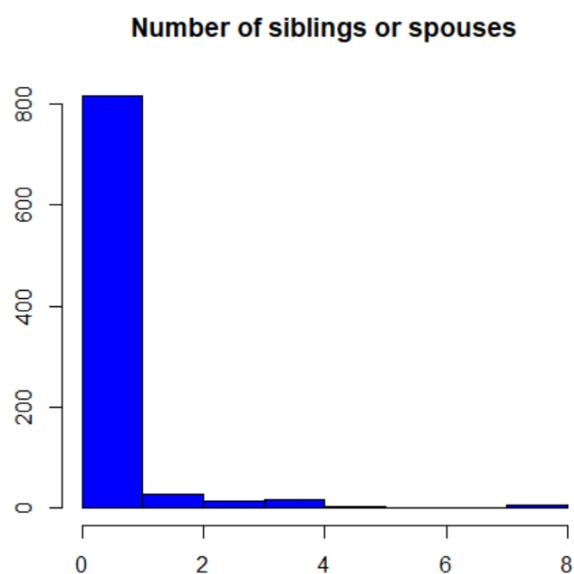
ตัวแปร Sex



จากกราฟพบว่า ผู้โดยสารที่อยู่บนเรือไททานิคส่วนใหญ่เป็นเพศชายคิดเป็น 64.9% ในขณะที่ผู้โดยสารเป็นเพศหญิงคิดเป็น 35.1%

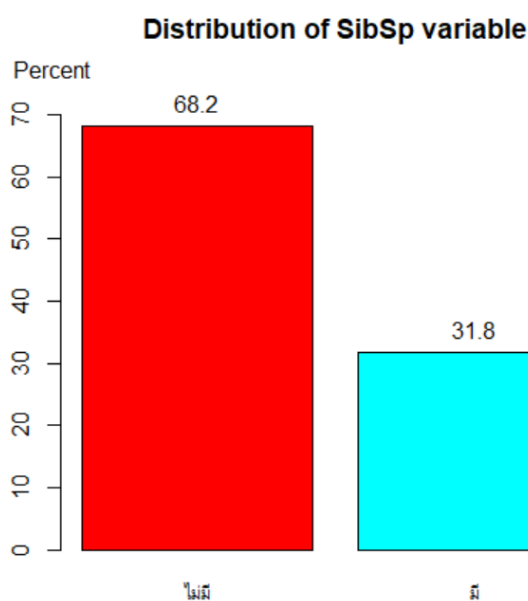
ตัวแปร SibSp

#Before



จากกราฟHistogramพบว่า ข้อมูลไม่มีการแจกแจงแบบปกติ ดังนั้นเราจะจัดกลุ่มข้อมูลของตัวแปรSibSpโดยให้ 0 = “ไม่มี” และ 1-8 = “มี” เพื่ออธิบายผลการวิเคราะห์ได้ง่ายขึ้น

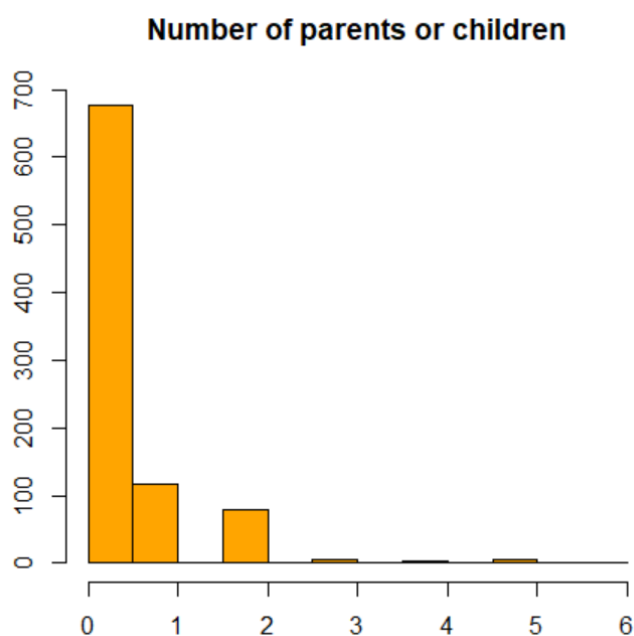
#After



จากกราฟพบว่า ผู้โดยสารที่อยู่บนเรือไททานิคส่วนใหญ่ไม่มีพี่น้องหรือคู่สมรส คิดเป็น 68.2% ในขณะที่ผู้โดยสารมีพี่น้องหรือคู่สมรส คิดเป็น 31.8%

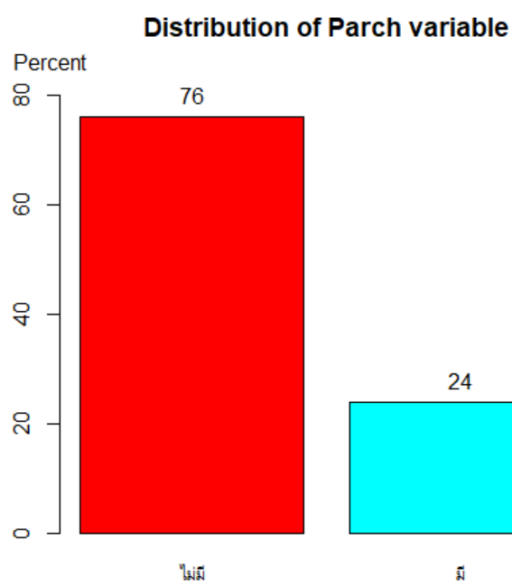
ตัวแปร Parch

#Before



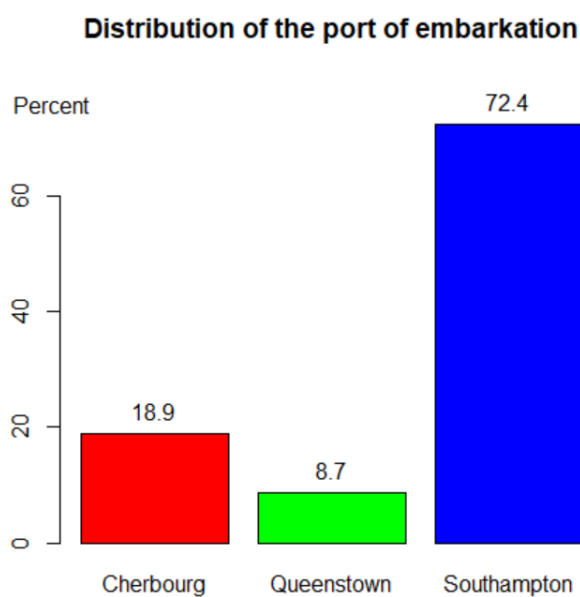
จากกราฟ Histogram พบว่าข้อมูลไม่มีการแจกแจงแบบปกติ ดังนั้นเราจะจัดกลุ่มข้อมูลของตัวแปร Parch โดยให้ 0 = “ไม่มี ” และให้ 1 - 8 = “ มี ” เพื่ออธิบายผลการวิเคราะห์ได้ง่ายขึ้น

#After



จากกราฟพบว่า ผู้โดยสารที่อยู่บนเรือไททานิคส่วนใหญ่ไม่มีพ่อแม่หรือลูก คิดเป็น 76% ในขณะที่ผู้โดยสารที่มีพ่อแม่หรือลูก คิดเป็น 24%

ตัวแปร Embarked



จากกราฟพบว่า ผู้โดยสารที่อยู่บนเรือไททานิคส่วนใหญ่ขึ้นท่าเรือ Southampton รองลงมาคือ Cherbourg และ Queenstown คิดเป็น 72.4% , 18.9% และ 8.7% ตามลำดับ

8. ขั้นตอนการวิเคราะห์ข้อมูล

8.1 การวิเคราะห์ตัวแปรที่ละคู่ระหว่างตัวแปรตามและตัวแปรอิสระ

```
> tableStack(data= dt1, c(Pclass, Sex, age, SibSp, Parch, Embarked), by = Survived)
```

ตารางที่1: ผลการวิเคราะห์ตัวแปรที่ละคู่ระหว่างตัวแปรตามและตัวแปรอิสระ

ตัวแปรอิสระ	0 : ไม่รอด	1 : รอด	χ^2 (degrees of freedom)	P-value
Pclass			100.98 (2 df)	< 0.001
ชั้น1	80 (14.6)	134 (39.4)		
ชั้น2	97 (17.7)	87 (25.6)		
ชั้น3	372 (67.8)	119 (35)		
Sex			260.76 (1 df)	< 0.001
M	468 (85.2)	109 (32.1)		
F	81 (14.8)	231 (67.8)		
age			2.66 (3 df)	0.447
<=22	133 (24.2)	98 (28.8)		
23-29	95 (17.3)	58 (17.1)		
30-35	187 (34.1)	103 (30.3)		
36+	134 (24.4)	81 (23.8)		
SibSp			12.4 (1 df)	< 0.001
ไม่มี	398 (72.5)	208 (61.2)		
มี	151 (27.5)	132 (38.8)		
Parch			19.82 (1 df)	< 0.001
ไม่มี	445 (81.1)	231 (67.9)		
มี	104 (18.9)	109 (32.1)		
Embarked			26.49 (2 df)	< 0.001
Cherbourg	75 (13.7)	93 (27.4)		
Queenstown	47 (8.6)	30 (8.8)		
Southampton	427 (77.8)	217 (63.8)		

จากผลการวิเคราะห์ตัวแปรตามและตัวแปรอิสระจะเห็นได้ว่าชั้นผู้โดยสาร (Pclass) , เพศ (Sex) , จำนวนพี่น้องหรือคู่สมรสของผู้โดยสารบนเรือไททานิค (SibSp) , จำนวนพ่อแม่หรือลูกของผู้โดยสารบนเรือไททานิค (Parch) , และท่าเรือ (Embarked) มีความสัมพันธ์กับการรอดชีวิตของผู้โดยสารบนเรือ (Survived) อย่างมีนัยสำคัญทางสถิติ

9. การวิเคราะห์หลายตัวแปร

วิเคราะห์ความสัมพันธ์ระหว่างตัวแปรอิสระหลายตัวแปรกับตัวแปรตามด้วยสถิติการถดถอยโลจิสติกพหุคูณ (Multiple Logistic Regression)

ข้อตกลง (Assumption) :

- ค่าความคลาดเคลื่อน (error) เป็นอิสระต่อกัน
- ไม่มี multicollinearity ระหว่างตัวแปรอิสระ

การวิเคราะห์หาความสัมพันธ์ระหว่างตัวแปรตามกับตัวแปรอิสระหลายตัวพร้อมกัน โดยนำตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามอย่างมีนัยสำคัญทางสถิติทุกตัว จากผลการวิเคราะห์สองตัวแปรในข้อที่ 7 เข้าไปในตัวแบบตั้งต้น (Initial model)

9.1 การสร้างตัวแบบตั้งต้น (Initial model)

ใช้คำสั่ง glm() โดยกำหนด link function เป็น binomial

```
# สร้างตัวแบบ
> mod <- glm(formula = Survived ~ Pclass + age + Sex + SibSp + Parch + Embarked,
              family = binomial, data = dt1)
> summary(mod)
```

ตารางที่1: ผลการวิเคราะห์การสร้างตัวแบบตั้งต้น

Call:

```
glm(formula = Survived ~ Pclass + age + Sex + SibSp + Parch + Embarked
, family = binomial, data = dt1)
```

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	0.64394	0.31770	2.027	0.042672 *
Pclass ^{ชั้น} 2	-0.87133	0.26450	-3.294	0.000987 ***
Pclass ^{ชั้น} 3	-2.18887	0.25256	-8.667	< 2e-16 ***
age23-29	-0.33591	0.28390	-1.183	0.236737
age30-35	-0.39135	0.24842	-1.575	0.115170
age36+	-0.95531	0.27510	-3.473	0.000516 ***
SexF	2.58417	0.19447	13.288	< 2e-16 ***
SibSp ^{มี}	-0.07991	0.20435	-0.391	0.695753
Parch ^{มี}	0.09250	0.22866	0.405	0.685808
EmbarkedQueenstown	-0.03139	0.37870	-0.083	0.933946
EmbarkedSouthampton	-0.54762	0.23217	-2.359	0.018337 *

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

Null deviance: 1182.82 on 888 degrees of freedom

Residual deviance: 805.31 on 878 degrees of freedom

AIC: 827.31

Number of Fisher Scoring iterations: 5

จากผลการวิเคราะห์พบว่าตัวแปรอิสระที่มีผลกับตัวแปรตามคือ ^{ชั้น}ผู้โดยสาร (Pclass) , เพศ (Sex) , อายุของผู้โดยสาร (age) และท่าเรือ (Embarked)

ตัวแบบข้างต้นไม่เป็นตัวแบบที่ดีที่สุดเพราะว่าตัวแบบต้องทำนายของตัวแปรตามอย่างแม่นยำและตัวแบบควรมีความสามารถในการตีความให้เข้าใจง่าย

9.2. การคัดเลือกตัวแปรไว้ในตัวแบบ

9.2.1. การคัดเลือกตัวแบบด้วยวิธีถอยหลัง (Backward selection)

ทำได้ง่ายด้วยคำสั่ง step() ระบุตัวเลือก direction = “ backward ”

```
> bw <- step(mod, direction = "backward")
```

ตารางที่1: ผลการวิเคราะห์การคัดเลือกตัวแบบด้วยวิธีถอยหลัง

```
> bw <- step(mod, direction = "backward")
Start: AIC=827.31
Survived ~ Pclass + Sex + SibSp + Parch + Embarked + age
```

	Df	Deviance	AIC
- SibSp	1	805.46	825.46
- Parch	1	805.47	825.47
<none>		805.31	827.31
- Embarked	2	812.38	830.38
- age	3	818.02	834.02
- Pclass	2	896.90	914.90
- Sex	1	1022.66	1042.66

```
Step: AIC=825.46
Survived ~ Pclass + Sex + Parch + Embarked + age
```

	Df	Deviance	AIC
- Parch	1	805.54	823.54
<none>		805.46	825.46
- Embarked	2	812.63	828.63
- age	3	818.15	832.15
- Pclass	2	896.90	912.90
- Sex	1	1023.77	1041.77

```
Step: AIC=823.54
Survived ~ Pclass + Sex + Embarked + age
```

	Df	Deviance	AIC
<none>		805.54	823.54
- Embarked	2	812.70	826.70
- age	3	818.54	830.54
- Pclass	2	896.96	910.96
- Sex	1	1038.78	1054.78

จากผลการวิเคราะห์พบว่า ตัวแปรอิสระที่ควรตัดออกจากตัวแบบคือจำนวนพี่น้องหรือคู่สมรสของผู้โดยสารบนเรือไททานิค (SibSp) , จำนวนพ่อแม่หรือลูกของผู้โดยสารบนเรือ (Parch) เนื่องจากค่า AIC ของตัวแปร SibSp , Parch มีค่าน้อยกว่าค่า AIC ตั้งต้น

10. การสร้างตัวแบบสุดท้าย (Final model)

การสร้างตัวแบบสุดท้าย (Final model) เพื่อนำไปสู่สรุปผลการศึกษา

```
> m2 <- glm(data=dt1, Survived ~ Embarked + age + Pclass + Sex, family=binomial)
> summary(m2)
```

ตารางที่1: ผลการวิเคราะห์การสร้างตัวแบบสุดท้าย

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	-3.0506	0.2717	-11.227	< 2e-16 ***
EmbarkedCherbourg	0.5535	0.2318	2.387	0.016965 *
EmbarkedQueenstown	0.5112	0.3311	1.544	0.122558
age<=22	0.9629	0.2730	3.528	0.000419 ***
age23-29	0.6155	0.2886	2.132	0.032973 *
age30-35	0.5572	0.2542	2.192	0.028409 *
Pclass ^{ชั้น} 1	2.1847	0.2524	8.657	< 2e-16 ***
Pclass ^{ชั้น} 2	1.3158	0.2341	5.169	1.92e-08 ***
SexF	2.5892	0.1878	13.786	< 2e-16 ***

จากผลการวิเคราะห์ เราจะสร้างสมการของตัวแบบ

$$\log \text{it}(p_i) = -3.0506 + (0.5535)(\text{EmbarkedCherbourg}) + (0.9629)(\text{age} \leq 22) + (0.6155)(\text{age} 23-29) \\ + (0.5572)(\text{age} 30-35) + (2.1847)(\text{Pclass}^{\text{ชั้น}}1) + (1.3158)(\text{Pclass}^{\text{ชั้น}}2) + (2.5892)(\text{SexF})$$

11. การแปลผลจากตัวแบบสุดท้ายที่เลือก

โดยใช้คำสั่ง `logistic.display ()`

```
> logistic.display(m2)
```

ตารางที่1: การแปลผลจากตัวแบบโลจิสติกที่เลือกในการทำนายการรอดชีวิต: รอด vs ไม่รอด

Logistic regression prediction Survived : รอด vs ไม่รอด				
	crude OR (95%CI)	adj. OR (95%CI)	P(Wald's test)	P(LR-test)
Embarked:				0.028
Southampton (ref.)		1	--	
Cherbourg	2.44 (1.73,3.45)	1.74 (1.1, 2.74)	0.017	
Queenstown	1.26 (0.77, 2.04)	1.67 (0.87, 3.19)	0.123	
age:				0.005
36+ (ref.)		1	--	
<=22	1.22 (0.83,1.78)	2.62 (1.53,4.47)	< 0.001	
23-29	1.01 (0.66,1.55)	1.85 (1.05,3.26)	0.033	
30-35	0.91 (0.63, 1.31)	1.75 (1.06,2.87)	0.028	
Pclass:				< 0.001
ชั้น3 (ref.)		1	--	
ชั้น1	5.24 (3.71,7.4)	8.89 (5.42,14.58)	< 0.001	
ชั้น2	2.8 (1.96,4)	3.73 (2.36,5.9)	< 0.001	
Sex:				< 0.001
M (ref.)		1	--	
F	12.24 (8.82, 17)	13.32 (9.22, 19.25)	< 0.001	
Log-likelihood = -402.7701				
No. of observations = 889				
AIC value = 823.5402				

จากผลการวิเคราะห์พบว่า เมื่อเทียบกับผู้โดยสารบนเรือที่ขึ้นท่าเรือ Southampton ผู้โดยสารที่ขึ้นท่าเรือ Cherbourg จะมีโอกาสรอดชีวิตมากกว่า คิดเป็น **1.74** เท่า

ในขณะที่เทียบกับผู้โดยสารบนเรือที่มีอายุน้อยกว่าหรือเท่ากับ 36 ปี ผู้โดยสารที่มีอายุน้อยกว่าหรือเท่ากับ 22 ปี 23-29 ปี และ 30-35ปี มีโอกาสที่จะรอดชีวิตมากกว่า คิดเป็น **2.62** เท่า **1.85** เท่า และ **1.75** เท่า ตามลำดับ

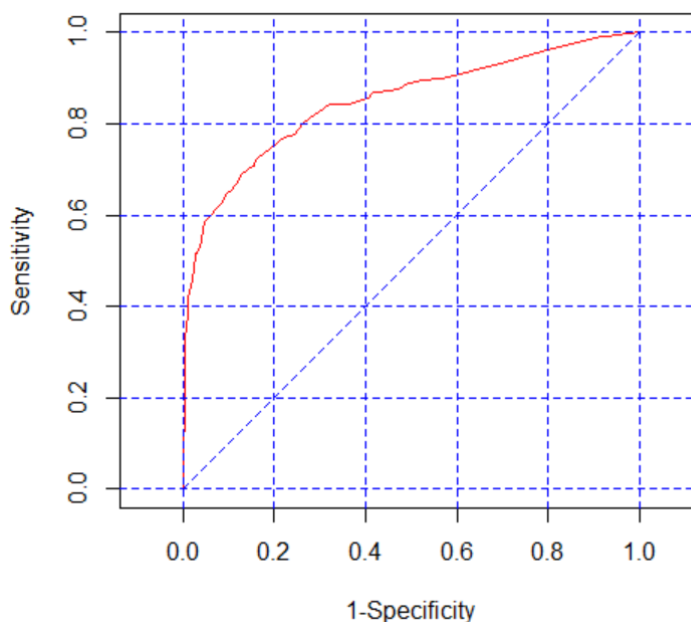
ต่อมาเมื่อเทียบกับผู้โดยสารที่อยู่ชั้นที่ 3 ผู้โดยสารที่อยู่ชั้นที่ 1 และ ชั้นที่ 2 บนเรือไททานิคจะมีโอกาสรอดชีวิตมากกว่า คิดเป็น 8.89 เท่า และ 3.73 เท่า ตามลำดับ

สุดท้ายเมื่อเทียบกับผู้โดยสารบนเรือเพศชาย ผู้โดยสารเพศหญิงที่อยู่บนเรือไททานิคมีโอกาสรอดชีวิตมากกว่า คิดเป็น 13.32 เท่า

12. ทดสอบประสิทธิภาพของตัวแบบ

การวินิจฉัยตัวแบบ

```
> ro <- lroc(m2, graph=TRUE)
> attributes(ro)
> ro$auc
```



ค่า $AUC = 0.85$ แปลผลว่าคือตัวแบบมีความกลมกลืนดี

ในกรณีนี้ค่า $AUC = 0.85$ หมายความว่า ตัวแบบมีความแม่นยำดีมากในการจำแนกกลุ่ม มีความกลมกลืนในการทำนายผลของตัวแปรตอบสนอง (Survived) ค่อนข้างสูง

เอกสารอ้างอิง

แหล่งข้อมูลมาจาก : <https://www.kaggle.com/datasets/lovishbansal123/titanic-dataset>

วรรณ นานา พิมาน แพง. "อัตรา การ รอดชีวิต และ ปัจจัย ที่มี ผล ต่อ การ รอดชีวิต จาก การ ช่วย พ้น คับ ซึพ ใน โรง พยาบาล ศรีสะเกษ." *วารสาร การ แพทย์ โรง พยาบาล ศรีสะเกษ สุรินทร์ บุรีรัมย์* 38.2 (2023): 369-382.

<https://he02.tci-thaijo.org/index.php/MJSSBH/article/view/264889>

Jampates, C., C. Yottasurodom, and C. Kasemsarn. "อัตรา รอดชีวิต ใน ระยะ ยาว ภายหลัง การ ผ่าตัด เลาะ เยื่อ หุ้ม หัวใจ ใน ผู้ ป่วย โรค ถู หุ้ม หัวใจ อักเสบ เรื้อรัง ใน สถาบัน โรค ทรวงอก." *Journal of The Department of Medical Services* 45.2 (2020): 141-148.

<https://he02.tci-thaijo.org/index.php/JDMS/article/download/245544/166893/852646>

Piyapadungkit, Suriya. "การ รอดชีวิต ของ ผู้ ป่วย ภายหลัง ผ่าตัด ภาวะ เลือด ออก ใน สมอง แปร ตาม ระยะ Midline Shift ใน ภาพถ่าย เอ็กซเรย์ คอมพิวเตอร์ และ ความ สำคัญ ของ ICU ภายหลัง ผ่าตัด โรง พยาบาล พระ นั้ ง เกล้า จังหวัด นนทบุรี." *Journal of Health Science of Thailand* 29.4 (2020): 660-669.

<https://thaidj.org/index.php/JHS/article/view/9310>

ปิย รัตน์ ลิ ม ป วิทยา กุล, et al. "การ ศึกษา ความ ชุก และ วิธี การ รักษา ที่มี ผล ต่อ อัตรา การ รอดชีวิต ใน ผู้ ป่วย มะเร็ง ปอด ชนิด ไม่ใช้ เซลล์ ขนาด เล็ก ระยะ เริ่ม ต้น ใน โรง พยาบาล พระมงกุฎเกล้า." *Royal Thai Army Medical Journal* 72.2 (2019): 109-120.

<https://he02.tci-thaijo.org/index.php/rtamedj/article/view/199395>

Vichapat, Voralak. "ปัจจัย ที่ เกี่ยว กับ การ เสีย ชีวิต และ อัตรา การ รอดชีวิต ใน ผู้ ป่วย โรค มะเร็ง ปอด ชนิด เซลล์ ขนาด ไม่ เล็ก ระยะ แพร่ กระจาย ที่ ได้ รับ การ วินิจฉัย และ รักษา ที่ โรง พยาบาล สระบุรี." *Journal of The Department of Medical Services* 46.1 (2021): 182-192.

<https://he02.tci-thaijo.org/index.php/JDMS/article/view/251800>