。 ~ a	~	al
ลาดับท	รหัส นศ.	ข-สกุล

## ใบงาน รายวิชา 747-341 Data analytics and data visualization module

## week 5-6: inferential statistics

## 1.ข้อมูลและรายละเอียดข้อมูล

ข้อมูลน้ำหนักของทารกแรกคลอด ที่เก็บรวบรวม จากศูนย์การแพทย์ Baystate (Baystate Medical Center) ในเมือง Springfield รัฐแมสซาซูเซตส์ จากหญิงตั้งครรภ์จำนวน 189 ราย ข้อมูลนี้ถูกเก็บในชื่อ "Bweight.csv" รายละเอียดข้อมูลแสดงดังตารางที่ 1 การวิเคราะห์ข้อมูล 1 ตัวแปร โดยใช้สถิติเบื้องต้น

ตารางที่ 1 รายละเอีบดข้อมูล

ตัวแปร	คำอธิบายตัวแปร	คำอธิบายค่าตัวเลข
low	น้ำหนักแรกคลอดต่ำกว่า 2,500 กรัม	0 คือ >= 2,500 , 1 คือ < 2,500
age	อายุของมารดาเป็นปี	
lwt	น้ำหนักของมารดาที่มีประจำเดือนครั้งสุดท้าย (ปอนด์)	
race	เชื้อชาติของมารดา	1=ผิวขาว, 2=ผิวดำ, 3=อื่นๆ
smoke	การสูบบุหรี่ขณะตั้งครรภ์	0= ไม่สูบ, 1= สูบ
ptl	จำนวนของการคลอดก่อนกำหนด	
ht	ประวัติการเป็นโรคความดันโลหิตสูง	0=ไม่มี, 1=มี
ui	การกลั้นปัสสาวะไม่อยู่	0=ไม่มี, 1=มี
ftv	จำนวนครั้งในการพบแพทย์ของไตรมาสแรก	
bwt	น้ำหนักแรกคลอดเป็นกรัม ( <b>ตัวแปรตาม</b> )	

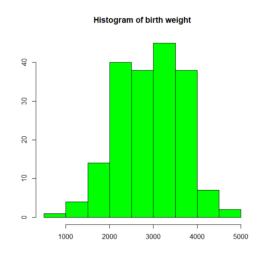
						າ
92	9,	92 6 92		•	т И т	9
จากข้อม	<u>เลข้าง</u>	าต้ามไห้เ	าปได้	ตอบคำถา	มตัด โป	٩Ĺ
1 111 11 110	10.1	4 7 1 70 0 7 1	POF 1.	VIII	1471000	•

1.	จากข้อมูล	"Bweight.csv"	ตัวแปรทั้งหมดมีกี่ตัว
	ข	U	

<sup>2.</sup> จงระบุชื่อตัวแปรต่อเนื่อง คือ ......

3. จงระบุชื่อตัวแปรกลุ่ม คือ ......

4. กราฟฮีสโตแกรมแสดงการแจกแจงของน้ำหนักของทารกแรกเกิด

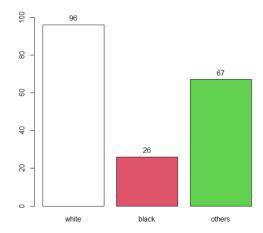


> summ(wt\$bwt)
obs. mean median s.d. min. max.
189 2944.587 2977 729.214 709 4990

จงอธิบายกราฟพร้อมทั้งแสดงค่าสถิติเบื้องต้น

5. กราฟแท่งแสดงการแจกแจงของตัวแปรสีผิวของแม่





จงอธิบายกราฟโดยใช้สถิติเบื้องต้นร้อยละ

		1		l l
0	~	a 2		പ്പ
2	าดำเ	ที่รหิส	นศ.	୩୭-ଷ୍ଧର
61	ил	VI VI 61	wii.	п п – еи ї еи

## การวิเคราะห์ข้อมูลรายคู่ 2 ตัวแปร

6. ผลการวิเคราะห์ข้อมูลแสดงความสัมพันธ์ระหว่างอัตราการเต้นของชีพจรของ ก่อนวิ่งและหลังวิ่งแตกต่างกันหรือไม่

```
> t.test(pul$PULSE2_NR,pul$PULSE1_NR,paired=TRUE,var.equal=TRUE)

Paired t-test

data: pul$PULSE2_NR and pul$PULSE1_NR

t = -0.70662, df = 30, p-value = 0.4853

alternative hypothesis: true mean difference is not equal to 0

95 percent confidence interval:
-1.6313753 0.7926656

sample estimates:
mean difference
-0.4193548
```

```
จากผลการวิเคราะห์ข้อมูลข้างต้น จงแสดงลำดับการทดสอบสมมตฐาน
```

7. ผลการวิเคราะห์ข้อมูลแสดงความสัมพันธ์ระหว่างการสูบบุหรี่ระหว่างตั้งครรภ์กับน้ำหนักของทารกแรกคลอด

```
> t.test(wt$bwt~wt$smoke1, var.equal=T)

Two Sample t-test

data: wt$bwt by wt$smoke1

t = 2.6529, df = 187, p-value = 0.008667

alternative hypothesis: true difference
95 percent confidence interval:
    72.75612 494.79735

sample estimates:
    mean in group no mean in group yes
    3055.696 2771.919
```

8	ลำดับที่	รหัส นศ	ชื่อ-สกุล
จากผล	การวิเครา	าะห์ข้อมูลข้างต้น จงแสดงลำด่	<b>้</b> บการทดสอบสมมตฐาน
8. ผลก	ารวิเคราะ	ห์ข้อมูลแสดงความสัมพันธ์ระ	หว่างชนสีผิวของมารดากับน้ำหนักของทารกแรกคลอด
> 0	neway	.test(wt\$bwt~w	t\$race1, var.equal=T)
		One-way analys:	is of means
		t\$bwt and wt\$ra	ace1 , denom df = 186, p-value = 0.008336
-	1.51	237 Hull GI = 2	, denom dr = 100, p varde = 0.000550
จากผล	การวิเครา	ะห์ข้อมูลข้างต้น จงแสดงลำด์	์ ทำการทดสุดบุสมมติฐาน

	ลำดับที่	รหัส นศ			ชื่อ-สกุล.			
ตัวแบ	บการถดถอยเชิง	แส้น						
9. ผลf	าารวิเคราะห์ข้อมูล	แสดงความสัม	พันธ์ระหว่างจำ	นวนการคลอดเ	า่อนกำหนดข	องมารดากับ	บน้ำหนักของท	ารกแรก
ନରପମ								
> n	simple 1 n1 <- lm(b summary(m1	wt~ptl1,						
Cal lm(	l: (formula =	bwt ~ p	tl1, dat	a = wt)				
	siduals: Min 992.49 -4	1Q I 95.40	Median -8.49	3Q 586.51	Ma 1976.5	x 1		
(In	efficients ntercept) .11-3	Estimate 3013.49	56	.59 53.	252 <	2e-16		
Res Mul F-s	gnif. code sidual sta tiple R-s statistic: กการวิเคราะห์ข้อมู	ndard er: quared: 9.34 o	ror: 713 0.04757 n 1 and	.6 on 18 , Adju 187 DF,	7 degre sted R- p-valu	es of :	freedom d: 0.04	

ลาดิบทรหิส นศขอ-สกุลขอ-สกุล	
·	
10. ผลการวิเคราะห์ข้อมูลแสดงความสัมพันธ์ระหว่างสีผิวของมารดากับน้ำหนักของทารกแรกคลอด	
> m2 <- lm(bwt~race1, data=wt) > summary(m2)	
Call: lm(formula = bwt ~ race1, data = wt)	
Residuals: Min 1Q Median 3Q Max	
-2096.28 -502.72 -12.72 526.28 1887.28	
Coefficients: Estimate Std. Error t value Pr(> t )	
(Intercept) 3102.72 72.92 42.548 < 2e-16 ***	
race1black -383.03 157.96 -2.425 0.01627 *	
racelothers -297.44 113.74 -2.615 0.00965 **	
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ''	1
Residual standard error: 714.5 on 186 degrees of freedom Multiple R-squared: 0.05017, Adjusted R-squared: 0.03996 F-statistic: 4.913 on 2 and 186 DF, p-value: 0.008336	
จากผลการวิเคราะห์ข้อมูลข้างต้น จงแสดงลำดับการทดสอบสมมตฐาน	
	••••

ลาดับทีร	หล นศ		ชอ-สกุล		
11. ผลการวิเคราะห์ข้อมูลแ	สดงความสัมพันธ์ระ	หว่างตัวแปรอิสระ 2 เ	ทั่วแปร (จำนวนการ	คลอดก่อนกำเ	หนดและสี่ผิวของ
มารดา)กับน้ำหนักของทารก	แรกคลอด(ตัวแปรต	าม)			
<pre>&gt; m3 &lt;- lm(bwt &gt; summary(m3)</pre>	~ptl1+race1	l, data=wt)			
Call:					
lm(formula = b	wt ~ ptl1 +	race1, dat	a = wt)		
Residuals:					
		an 3Q			
-2143.07 -430	.63 -32.0	534.93	1825.93		
Coefficients:					
Es (Intercept) 3		Error t va			
ptl11-3 -					
race1black -	379.66	154.63 -2	455 0.01	500 *	
racelothers -	283.44	111.44 -2	544 0.01	179 *	
Signif. codes:	0 ****	0.001 '**'	0.01 '*' 0	.05 '.'	0.1 ' ' 1
Residual stand Multiple R-squ F-statistic: 6	ared: 0.09	9477, Adjı	ısted R-sq	uared:	0.08009
			_		
จากผลการวิเคราะห์ข้อมูลข้	้างต้น จงแสดงลำดับ	เการทดสอบสมมตฐา	น		