**ใบงานที่ 4**

1. จากข้อมูลตารางที่ 1 ตารางที่ 2 และคำอธิบายข้อมูลในตารางที่ 3 ให้นักศึกษาอธิบายการทำความสะอาดข้อมูล และการผสานข้อมูลตารางที่ 1 และตารางที่ 2 เข้าด้วยกัน พร้อมเขียนคำสั่งในโปรแกรม R

**ตารางที่ 1** ข้อมูลบุคคล



**ตารางที่ 2** ข้อมูลการศึกษาและการทำงาน



**ตารางที่ 3** คำอธิบายข้อมูลการศึกษาและการทำงาน



คำอธิบายการทำความสะอาดข้อมูล และการผสานข้อมูล

ตารางที่1 ข้อมูลบุคคล ทำความสะอาด และผสานข้อมูลจากตารางที่ต่างกันเข้าด้วยกัน ของตัวแปร titnme โดยที่ เด็กชาย = 1, เด็กหญิง = 2, นาย = 3, นางสาว = 4 และค่าว่าง เป็นNA

ตารางที่2 ข้อมูลการศึกษาและการทำงานให้เปลี่ยนจากภาษาเป็นตัวเลข และค่าว่างให้เป็น NA เขียนคำสั่งในโปรแกรม R (Command R)

dt$titnme <- ifelse(titnme==1,”เด็กชาย”,

ifelse(titnme==2,”เด็กหญิง”,

ifelse(titnme==3,”นาย”,”นางสาว”)))

dt$age <- ifelse(dt$age==””,NA,dt$age)

dt$gender <- ifelse(is.na(gender), titnme, gender)

dt$a9 <- ifelse(dt$a9==”ได้”,1,0)

dt$a10 <- ifelse(dt$a10==”ได้”,1,0)

dt$a11 <- ifelse(dt$a11 == “”,NA, dt$a11)

dt$a12 <- ifelse(dt$a12 == “”,NA, dt$a12)

dt$a13 <- ifelse(dt$a13 == “”,NA, dt$a13)

dt$a14 <- ifelse(dt$a14 == “”,NA, dt$a14)

2. จากข้อมูลตารางที่ 4 ข้อมูลปริมาณน้ำฝน และตารางที่ 5 ข้อมูลอุณหภูมิ ให้นักศึกษาอธิบายการผสานข้อมูล พร้อมเขียนคำสั่งในโปรแกรม R

**ตารางที่ 3** ข้อมูลปริมาณน้ำฝน



**ตารางที่ 3** ข้อมูลอุณหภูมิ



คำอธิบายการทำความสะอาดข้อมูล และการผสานข้อมูล

ตารางที่ 4 ข้อมูลปริมาณน้ำฝน ตัวแปร ปริมาณฝน 24 ชม. (มม.) เปลี่ยนจาก ไม่มีฝน = 0

และตารางที่ 5 ข้อมูลอุณหภูมิ จัดการค่าข้อมูลจังหวัดให้อยู่ในประเภทเดียวกัน คือภาคใต้ ภาคกลาง ภาคเหนือ ภาคอีสาน และอุณหภูมิสูงสุด ไม่มี = NA อุณหภูมิต่ำสุด ที่ยังไม่ได้รับรายงาน = NA

ภาคอีสานเขียนคำสั่งในโปรแกรม R (Command R)

d$ปริมาณฝน 24 ชม. (มม.) <- ifelse(d$ปริมาณฝน 24 ชม. (มม.)==”ไม่มีฝน”,0 , d$ปริมาณฝน 24 ชม. (มม.))

d$อุณหภูมิ (°C) สูงสุด <- ifelse(d$อุณหภูมิ (°C) สูงสุด == “”,NA, d$อุณหภูมิ (°C) สูงสุด)

d$อุณหภูมิ (°C) ต่ำสุด <- ifelse(d$อุณหภูมิ (°C) ต่ำสุด== “ยังไม่ได้รับรายงาน”,NA, d$อุณหภูมิ (°C) ต่ำสุด)