

เอกสารวิชาการ

เรื่องที่ 1

คู่มือการวิเคราะห์ข้อมูลระบาดวิทยาทางสัตวแพทย์ด้วยโปรแกรม R
สำหรับเจ้าหน้าที่ปศุสัตว์

โดย

อรพรรณ อัจคำภา

เลขทะเบียนวิชาการ	64(2)-0116(4)-161
สถานที่ดำเนินการ	สำนักงานปศุสัตว์เขต 4
ระยะเวลาดำเนินการ	กุมภาพันธ์ 2564 - กันยายน 2564
การเผยแพร่	เว็บไซต์สำนักงานปศุสัตว์เขต 4

คำนำ

ในปัจจุบันงานระบาดวิทยาทางสัตวแพทย์มีบทบาทสำคัญกับการปฏิบัติงานของเจ้าหน้าที่ของกรมปศุสัตว์มากยิ่งขึ้น เนื่องมาจากกรมปศุสัตว์ได้จัดทำระบบเฝ้าระวังโรคที่สำคัญในปศุสัตว์ เช่น โรคปากและเท้าเปื่อย โรคล้มปี่ สกีน โรคไขหวัดนก โรคระบาดในสุกร โรคพิษสุนัขบ้า เป็นต้น เพื่อให้เจ้าหน้าที่ทั้งในระดับอำเภอและระดับจังหวัดรายงานโรคในฐานข้อมูลต่างๆ อย่างมีระบบและรายละเอียดเพิ่มมากขึ้น และกรมปศุสัตว์ยังดำเนินการฝึกอบรมเจ้าหน้าที่ปศุสัตว์เพื่อเตรียมความพร้อมให้เจ้าหน้าที่ปศุสัตว์ให้มีทักษะในการนำข้อมูลต่างๆ ข้างต้น มาทำการวิเคราะห์เพื่อนำข้อมูลในพื้นที่ไปใช้เป็นแนวทางในการวางแผนหรือดำเนินงานในการป้องกันและควบคุมโรคให้มีประสิทธิภาพมากยิ่งขึ้น เพื่อที่จะใช้ในการแก้ไขปัญหาให้แก่ผู้เลี้ยงปศุสัตว์ในแต่ละพื้นที่ได้อย่างแท้จริง ดังนั้นการจัดการข้อมูล การวิเคราะห์ข้อมูล และการนำเสนอข้อมูลทางด้านระบาดวิทยาทางสัตวแพทย์ ถือเป็นหนึ่งในทักษะที่เจ้าหน้าที่ปศุสัตว์จะต้องมีการพัฒนาอย่างต่อเนื่อง ซึ่งคู่มือการวิเคราะห์ข้อมูลระบาดวิทยาทางสัตวแพทย์ด้วยโปรแกรมอาร์ (R) สำหรับเจ้าหน้าที่ปศุสัตว์เล่มนี้ เป็นคู่มือที่อธิบายถึงวิธีการจัดการข้อมูล วิเคราะห์ข้อมูล และนำเสนอข้อมูลที่เกี่ยวข้องกับงานระบาดวิทยาทางสัตวโดยใช้โปรแกรมอาร์ (R) ขึ้นพื้นฐานที่เจ้าหน้าที่ปศุสัตว์สามารถนำมาใช้เป็นแนวทางในการปฏิบัติงานได้เป็นอย่างดี

ผู้เขียนหวังเป็นอย่างยิ่งว่าคู่มือเล่มนี้จะเป็นประโยชน์ต่อเจ้าหน้าที่ปศุสัตว์ทั้งในระดับอำเภอ ระดับจังหวัด ระดับเขต ในการเพิ่มพูนทักษะความรู้เรื่องการวิเคราะห์ข้อมูลระบาดวิทยาทางสัตวแพทย์ด้วยโปรแกรมอาร์ (R) และสามารถใช้เป็นแนวทางในการเรียนรู้ เพิ่มทักษะ และประยุกต์ใช้ในการปฏิบัติงานให้มีประสิทธิภาพมากยิ่งขึ้นต่อไป

สพ.ญ.ดร.อรพรรณ อัจคำภา

สารบัญ

	หน้า
คำนำ	1
สารบัญ	2
สารบัญภาพ	3
สารบัญตาราง	4
บทที่ 1 ความรู้เบื้องต้นเกี่ยวกับโปรแกรม R และ RStudio	5
บทที่ 2 การดาวน์โหลดและติดตั้งโปรแกรม R และ Rstudio	8
บทที่ 3 การใช้งานโปรแกรม R เบื้องต้น	14
บทที่ 4 การนำเข้าข้อมูล	27
บทที่ 5 การจัดการข้อมูล	35
บทที่ 6 การวิเคราะห์ข้อมูลสถิติพื้นฐานสำหรับงานระบาดวิทยาทางสัตวแพทย์	47
บทที่ 7 การบันทึกข้อมูลและส่งออกผลการวิเคราะห์ข้อมูล	64
บทที่ 8 ข้อเสนอแนะและข้อควรระวังของการใช้โปรแกรม	69
เอกสารอ้างอิง	70
ภาคผนวก ก ข้อมูลสำหรับการวิเคราะห์เชิงพรรณนาและเชิงอนุมาน	71
ภาคผนวก ข ข้อมูลสำหรับการวิเคราะห์เชิงเวลา	76
ภาคผนวก ค ข้อมูลสำหรับการวิเคราะห์เชิงพื้นที่	79

สารบัญรูปภาพ

	หน้า
รูปที่ 1.1 ความสามารถหลักของโปรแกรม R	6
รูปที่ 1.2 ตัวอย่างการนำเสนอข้อมูลจากการวิเคราะห์ข้อมูลด้วยโปรแกรม R	7
รูปที่ 2.1 เว็บไซต์สำหรับดาวน์โหลดโปรแกรม R	8
รูปที่ 2.2 หน้าต่าง CRAN mirror ของประเทศไทยสำหรับใช้ดาวน์โหลด	9
รูปที่ 2.3 ระบบปฏิบัติการที่ต้องการทำการติดตั้ง ได้แก่ Windows, Mac OS หรือ Linux	10
รูปที่ 2.4 การติดตั้งโปรแกรม R สำหรับระบบปฏิบัติการ Windows	10
รูปที่ 2.5 หน้าต่างสำหรับติดตั้งโปรแกรม R ของระบบปฏิบัติการ Windows	11
รูปที่ 2.6 หน้าต่างสำหรับติดตั้งโปรแกรม R ของระบบปฏิบัติการ Mac OS	12
รูปที่ 2.7 เว็บไซต์สำหรับดาวน์โหลดโปรแกรม RStudio	13
รูปที่ 2.8 หน้าต่างสำหรับดาวน์โหลดโปรแกรม RStudio	13
รูปที่ 3.1 ส่วนประกอบหน้าต่างการทำงานของโปรแกรม RStudio	15
รูปที่ 3.2 เมื่อต้องการให้คำสั่งที่ต้องการทำงาน ให้คลิกเลือกคำสั่งนั้นๆ หรือเอาเคอร์เซอร์วางไว้ในบรรทัดคำสั่งที่ต้องการ แล้วกดเลือกที่คำว่า “Run” เพื่อสั่งให้โปรแกรมวิเคราะห์และประมวลผล	15
รูปที่ 3.3 การตั้งค่าหน้าต่างแสดงผลของโปรแกรม R	16
รูปที่ 3.4 การติดตั้งชุดคำสั่งผ่านระบบอินเทอร์เน็ต	24
รูปที่ 3.5 การค้นหาคำอธิบายหรือตัวอย่างคำสั่งของชุดคำสั่งที่ต้องการ ผ่านหน้าต่าง Help	26
รูปที่ 4.1 ชื่อตัวแปรที่ถูกนำเข้ามาในโปรแกรม R เรียบร้อยแล้ว	30
รูปที่ 5.1 ลักษณะการจัดการข้อมูลด้วยชุดคำสั่ง dplyr	36
รูปที่ 7.1 การบันทึกคำสั่งที่สร้างขึ้นให้เป็นแบบสกุ่ล .R	64
รูปที่ 7.2 การบันทึกกราฟที่ได้จากการวิเคราะห์เป็นรูปภาพ	66
รูปที่ 7.3 การบันทึกรายงานคำสั่งและผลการวิเคราะห์โดยใช้ Compie report	67
รูปที่ 7.4 ตัวอย่างเอกสารรายงานในรูปแบบ Microsoft word จาก Compie report	68

สารบัญตาราง

	หน้า
ตารางที่ 3.1 โครงสร้างของข้อมูลในโปรแกรม R	17
ตารางที่ 3.2 ชนิดของข้อมูลในโปรแกรม R	20
ตารางที่ 3.3 ค่าการกระทำ ฟังก์ชันทางคณิตศาสตร์ และสัญลักษณ์ทางตรรกศาสตร์ อย่างง่ายที่ใช้ในภาษา R	21
ตารางที่ 3.4 ชุดคำสั่ง (Packages) เบื้องต้นที่ใช้บ่อยในการจัดการข้อมูล วิเคราะห์ข้อมูล และแสดงผลข้อมูลในงานระบาดวิทยาทางสัตวแพทย์	25
ตารางที่ 4.1 คำสั่งสำหรับนำเข้าไฟล์ข้อมูลในระบบปฏิบัติการ Windows และ Mac OS	29
ตารางที่ 5.1 สัญลักษณ์ที่ใช้ในการจัดรูปแบบวันที่ที่ใช้ในโปรแกรม R	46
ตารางที่ 6.1 ลักษณะของข้อมูลและวิธีการใช้สถิติเชิงพรรณนา	48

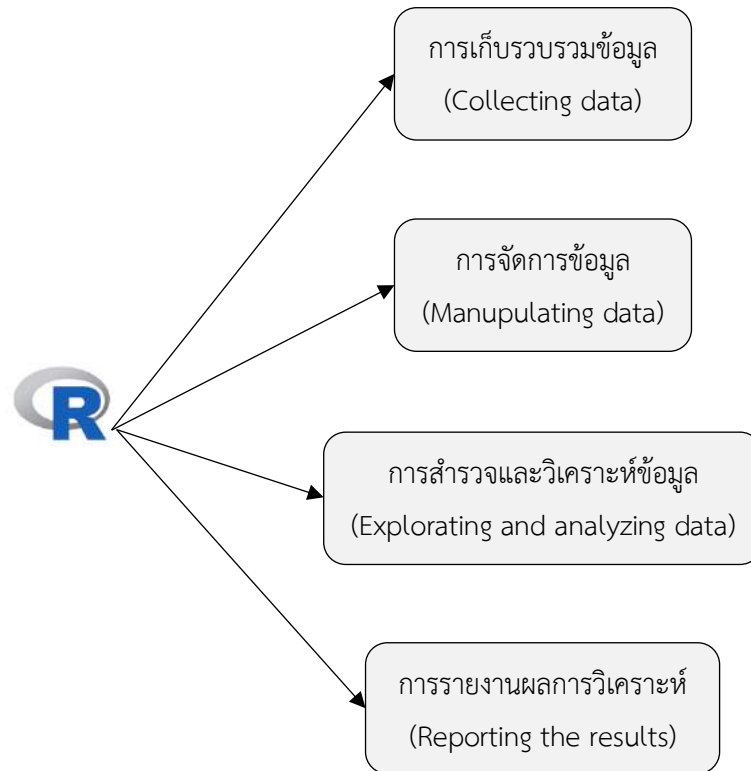
บทที่ 1

ความรู้เบื้องต้นเกี่ยวกับโปรแกรม R และ RStudio

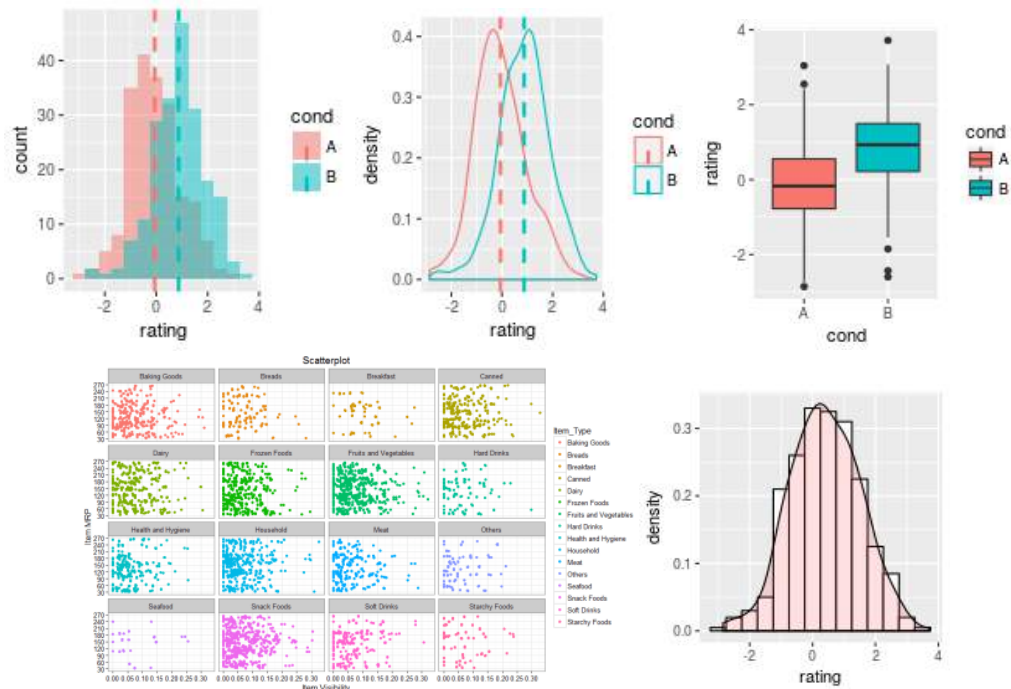
โปรแกรม R เป็นโปรแกรมคอมพิวเตอร์ที่ใช้สำหรับจัดการข้อมูล วิเคราะห์ข้อมูล และแสดงผลข้อมูลทางสถิติ ถูกสร้างและเผยแพร่ครั้งแรกเมื่อปี 1993 โดย Ross Ihaka และ Robert Gentleman ภาควิชาสถิติ มหาวิทยาลัยโอ๊คแลนด์ ประเทศนิวซีแลนด์ ในปัจจุบันมีทีมพัฒนาโปรแกรมที่เป็นนักสถิติ นักวิจัย อาจารย์ และผู้ใช้งานจากทั่วโลก ซึ่งโปรแกรม R เป็นซอฟต์แวร์ประเภท Open source ที่ทุกคนสามารถนำมาใช้ฟรีโดยไม่เสียค่าใช้จ่าย และผู้ใช้ไม่ต้องกังวลกับเรื่องของการละเมิดลิขสิทธิ์เหมือนกับโปรแกรมสำเร็จรูปทางสถิติอื่นๆ ผลวิเคราะห์สถิติด้วยโปรแกรม R นั้นได้รับการยอมรับจากนานาชาติอย่างกว้างขวาง โดยเฉพาะในวงการนักสถิติ (Statistician) นักวิเคราะห์ข้อมูล (Data analyst) และนักวิทยาการข้อมูล (Data scientist) เป็นโปรแกรมที่ดีมากสำหรับการใช้ในการเรียนรู้ทางสถิติ เนื่องจากสามารถทำให้ผู้ใช้เกิดความเข้าใจในกระบวนการทางสถิติได้ดียิ่งขึ้น อีกทั้งยังเป็นโปรแกรมที่มีความยืดหยุ่นในการวิเคราะห์ทางสถิติ จึงทำให้ผู้ใช้สามารถขยายกระบวนการวิเคราะห์ออกไปได้ตามความต้องการ และยังสามารถทำการบันทึกคำสั่งในการวิเคราะห์ได้อย่างง่ายดาย รวมถึงสามารถใช้ร่วมกับโปรแกรมการสร้างเอกสารอื่นๆ และสามารถใช้ได้ทั้งกับระบบปฏิบัติการ Windows, Mac OS และ Linux อีกด้วย ซึ่งข้อมูลและรายละเอียดต่างๆ เกี่ยวกับโปรแกรม R สามารถดูเพิ่มเติมได้ที่เว็บไซต์ <http://www.r-project.org>

ความสามารถหลักของโปรแกรม R แบ่งออกเป็น 4 ส่วน (รูปที่ 1.1) ได้แก่ 1) การเก็บรวบรวมข้อมูล (Collecting data) โปรแกรม R มีความสามารถในการนำเข้าข้อมูลจากแหล่งต่างๆ ทั้งข้อมูลที่อยู่ในรูปแบบ Flat File, Excel File, Database, Website รวมทั้งไฟล์ข้อมูลจากโปรแกรมวิเคราะห์ข้อมูลทางสถิติต่าง เช่น SPSS, SAS, STATA และ MiniTab เป็นต้น 2) การจัดการข้อมูล (Manipulating data) โปรแกรม R มีความสามารถในการจัดการข้อมูลเพื่อจัดระเบียบข้อมูลดิบที่นำเข้า และสร้างชุดข้อมูลใหม่ที่มีความพร้อมต่อการนำไปวิเคราะห์ข้อมูลตามที่ต้องการ 3) การสำรวจและวิเคราะห์ข้อมูล (Exploring and analyzing data) โดยโปรแกรม R มีฟังก์ชันที่รองรับการวิเคราะห์ข้อมูลด้วยโมเดลทางสถิติที่หลากหลายเทียบเท่า (หรืออาจมากกว่า) โปรแกรมวิเคราะห์ข้อมูลทางสถิติชั้นนำ เช่น SPSS, SAS, Minitab และ Stata โดยในปัจจุบันโปรแกรม R มีชุดคำสั่ง (Packages) มากกว่า 10,000 ชุดคำสั่ง และ 4) การรายงานผลการวิเคราะห์ (Reporting the results) โดยสามารถในการสร้างทัศนภาพข้อมูล (Visualizing data) ที่สามารถสร้างได้อย่างหลากหลายตามความต้องการของผู้วิจัย และมีความสวยงาม เป็นมืออาชีพ (รูปที่ 1.2) และสามารถการสร้างรายงาน (Report) ในหลากหลาย

รูปแบบทั้งในรูปแบบของ Webpage (HTML), PDF document, Word document และ Presentation slide เป็นต้น และยังสามารถนำเสนอผลการวิเคราะห์ได้อย่างหลากหลายโดยเฉพาะการนำเสนอทัศนภาพเชิงปฏิสัมพันธ์ (Interactive data visualization) ได้อีกด้วย



รูปที่ 1.1 ความสามารถหลักของโปรแกรม R



รูปที่ 1.2 ตัวอย่างการนำเสนอข้อมูลจากการวิเคราะห์ข้อมูลด้วยโปรแกรม R

ส่วน RStudio เป็นโปรแกรมที่ถูกพัฒนาขึ้นเพื่อเป็นเครื่องมือช่วยในการทำงานด้านการวิเคราะห์ข้อมูลด้วยภาษา R เรียกว่า Intergrated development environment (IDE) ซึ่งโปรแกรมนี้นี้จะคล้ายกับ Text editor ที่มีการเอื้อให้เหมาะกับการใช้งานและการพัฒนาโปรแกรมให้สะดวกมากขึ้น มีความยืดหยุ่นสูง ทำให้ผู้ใช้งานสามารถทำงานได้ง่าย รวดเร็ว และมีประสิทธิภาพมากขึ้น ผู้ใช้สามารถศึกษาคุณสมบัติต่างๆ ของโปรแกรม RStudio ได้เพิ่มเติมจากเว็บไซต์ <https://www.rstudio.com/products/rstudio/features/> ซึ่งในปัจจุบันการเรียนรู้ชุดคำสั่งต่างๆ รวมถึงการถามตอบปัญหาที่พบในการใช้งาน ผู้ใช้สามารถเข้าไปที่เว็บไซต์ <https://community.rstudio.com/> เพื่อแลกเปลี่ยนองค์ความรู้ต่างๆ ได้

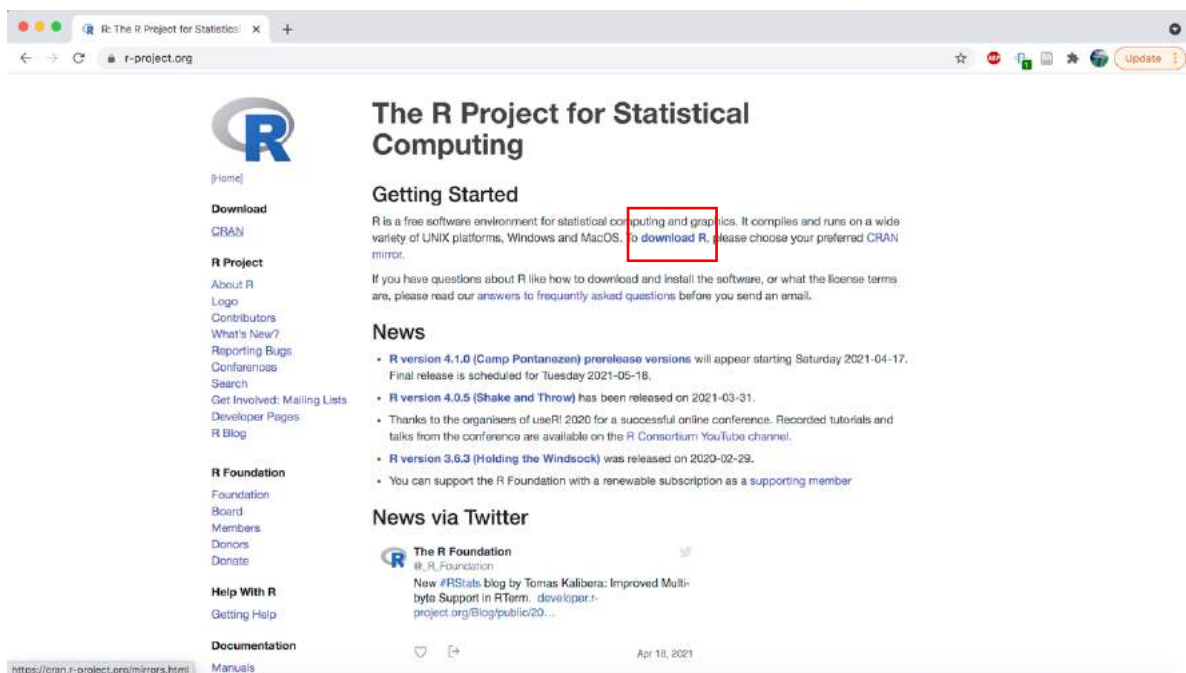
บทที่ 2

การดาวน์โหลดและติดตั้งโปรแกรม R และ RStudio

การติดตั้งโปรแกรม R ในครั้งแรกจำเป็นต้องติดตั้งทั้งโปรแกรม R และ RStudio เนื่องจาก RStudio เป็นโปรแกรมหนึ่งที่จะช่วยให้การทำงานในโปรแกรม R มีความสะดวกและง่ายขึ้น โดยที่โปรแกรม R จะเป็นเครื่องมือในการประมวลผลเบื้องหลัง ซึ่งในการเรียกใช้งานโปรแกรม R เราจะเปิดใช้งานเฉพาะโปรแกรม RStudio เท่านั้น

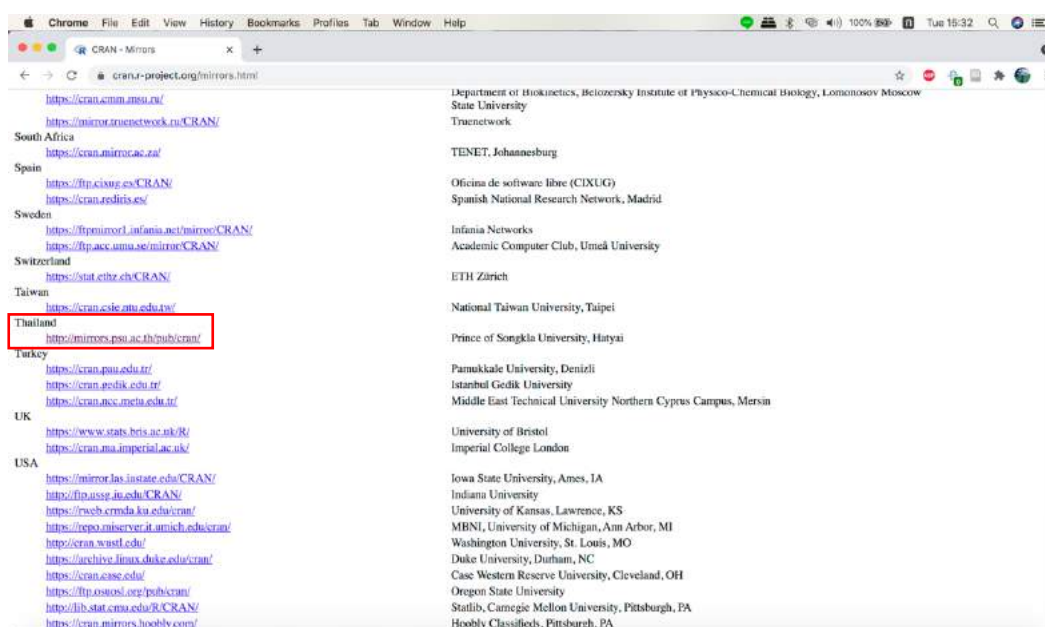
โปรแกรม R สามารถเลือกดาวน์โหลดโปรแกรมที่เหมาะสมกับระบบปฏิบัติการของเครื่องคอมพิวเตอร์ โดยสามารถเลือกได้ 3 ระบบปฏิบัติการ ได้แก่ Windows, Mac OS และ Linux จากนั้นดำเนินการติดตั้งโปรแกรมตามคำแนะนำของตัวช่วยติดตั้ง โดยในคู่มือฉบับนี้จะอธิบายวิธีการติดตั้งสำหรับระบบปฏิบัติการ Windows และ Mac OS

การติดตั้งโปรแกรม R ทำได้โดยดาวน์โหลดโปรแกรม R จาก <https://www.r-project.org/> แล้วเลือกคำว่า download R (รูปที่ 2.1)



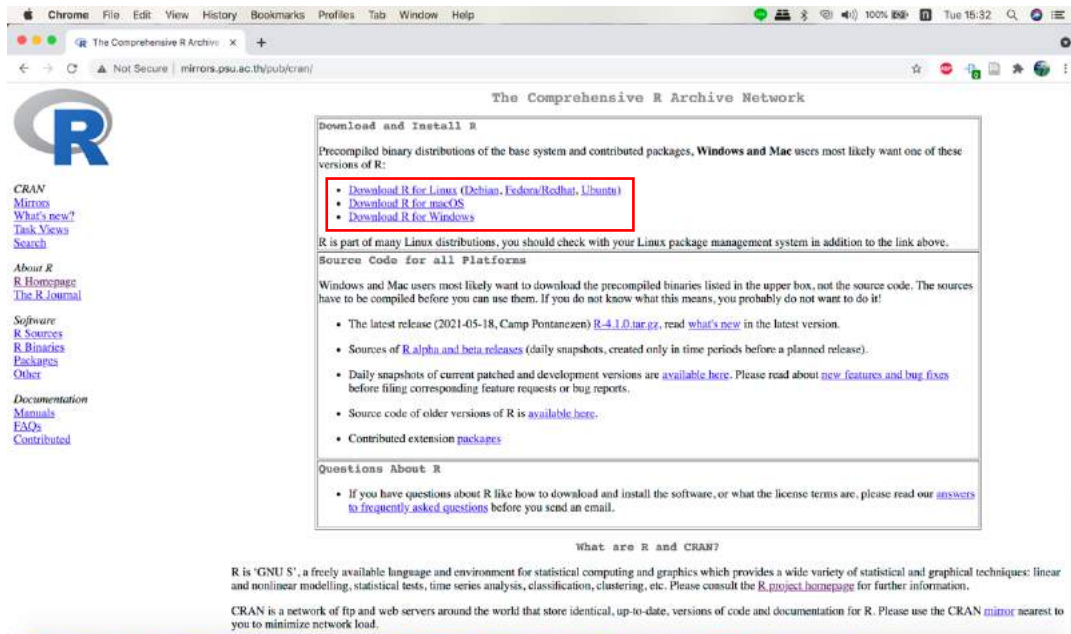
รูปที่ 2.1 เว็บไซต์สำหรับดาวน์โหลดโปรแกรม R

ขั้นตอนที่ 2 หน้าต่างของ CRAN mirror จะปรากฏขึ้น ซึ่งเราสามารถเลือกประเทศที่เราต้องการดาวน์โหลด โดยเลือกประเทศใดก็ได้ เพราะเป็นที่อยู่ของการดาวน์โหลด แต่หากเลือกประเทศใกล้เคียงเรา หรือเลือกประเทศไทย ก็จะดาวน์โหลดได้เร็วกว่าเลือกจากที่อื่น โดย CRAN mirror ของประเทศไทย คือมหาวิทยาลัยสงขลานครินทร์ หรือ Prince of Songkla University, Hatyai (รูปที่ 2.2)



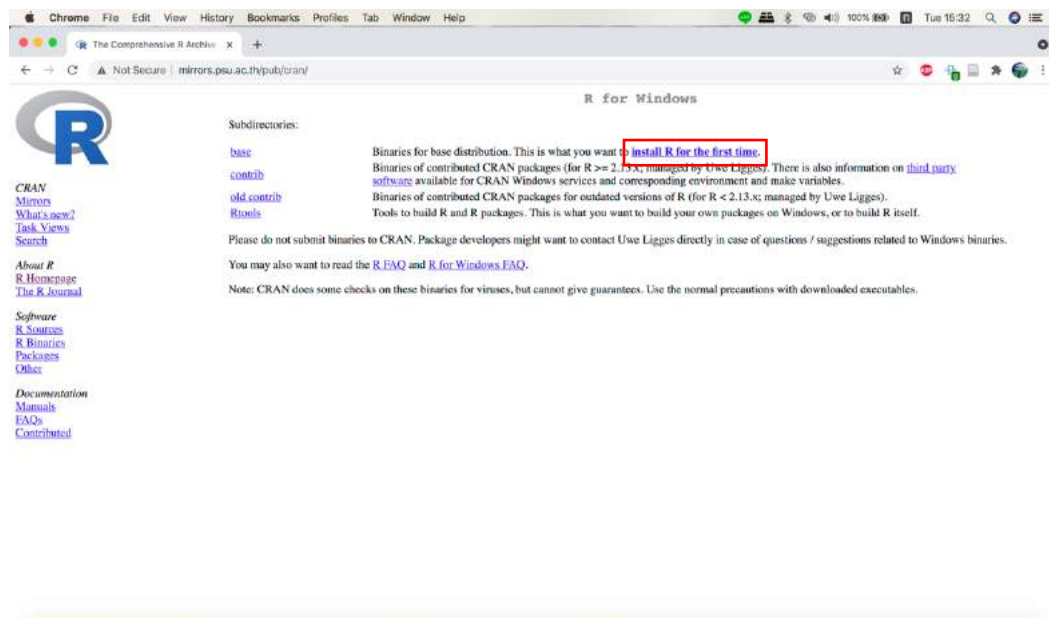
รูปที่ 2.2 หน้าต่าง CRANN mirror ของประเทศไทยสำหรับใช้ดาวน์โหลด

ขั้นตอนที่ 3 ให้ทำการเลือกระบบปฏิบัติการคอมพิวเตอร์ให้เหมาะสมกับระบบคอมพิวเตอร์ของผู้ใช้ เช่น Windows, Mac OS และ Linux (รูปที่ 2.3)



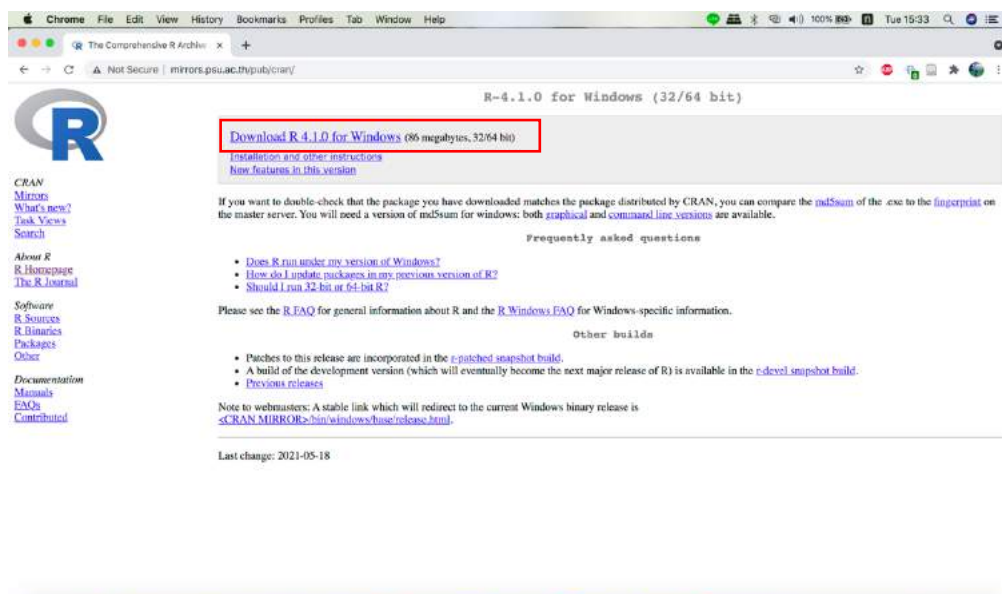
รูปที่ 2.3 ระบบปฏิบัติการที่ต้องการทำการติดตั้ง ได้แก่ Windows, Mac OS หรือ Linux

ในระบบปฏิบัติการ Windows ให้เลือก Download R for Windows จากหน้าต่างดังกล่าว แล้วทำการเลือก install R for the first time ดังรูปที่ 2.4



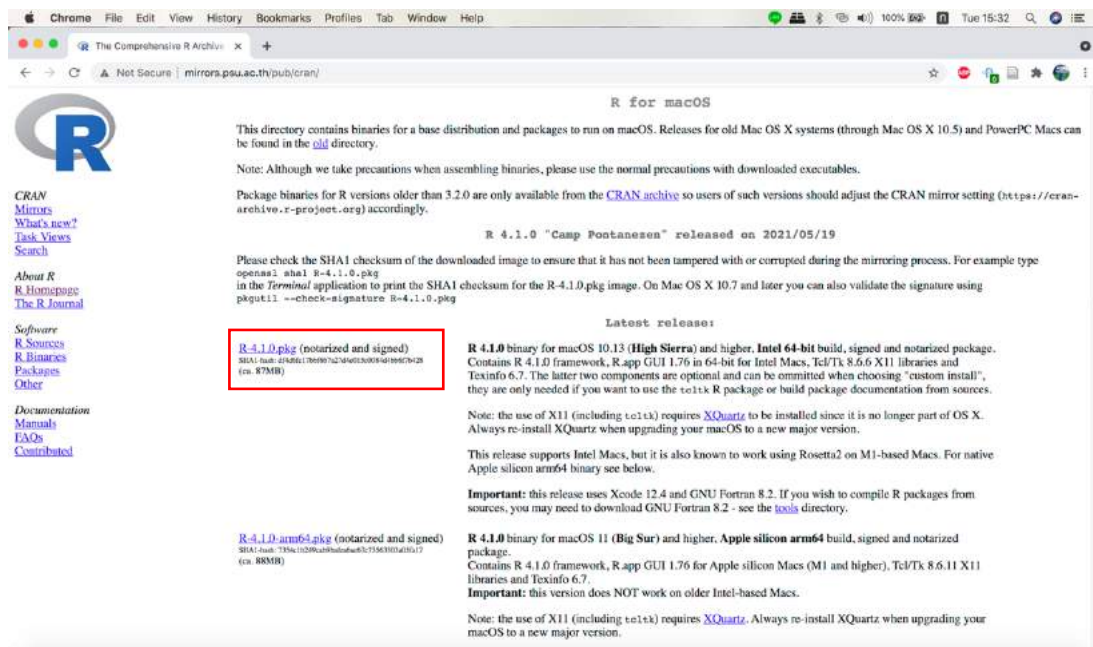
รูปที่ 2.4 การติดตั้งโปรแกรม R สำหรับระบบปฏิบัติการ Windows

จากนั้นจะมีหน้าต่างให้ดาวน์โหลดโปรแกรม R ที่เป็นเวอร์ชันล่าสุด ซึ่งในขณะนี้จะเป็นเวอร์ชัน 4.1.0 โดยสามารถเข้าไปดาวน์โหลดได้ที่ <http://mirrors.psu.ac.th/pub/cran/> (รูปที่ 2.5) เมื่อดาวน์โหลดมา ก็จะได้ไฟล์ R-x.x-xxx.exe แล้วแต่รุ่นและระบบปฏิบัติการ เช่น R-4.1.0-win32.exe โดยจะเก็บอยู่ที่โฟลเดอร์ Download จากนั้นทำการ Double click ไฟล์ดังกล่าว แล้วระบบจะแสดงหน้าจอการติดตั้งโปรแกรมโดยให้กด Next ไปเรื่อยๆ และกดปุ่ม Finish เมื่อทำการติดตั้งเรียบร้อยแล้ว



รูปที่ 2.5 หน้าต่างสำหรับติดตั้งโปรแกรม R ของระบบปฏิบัติการ Windows

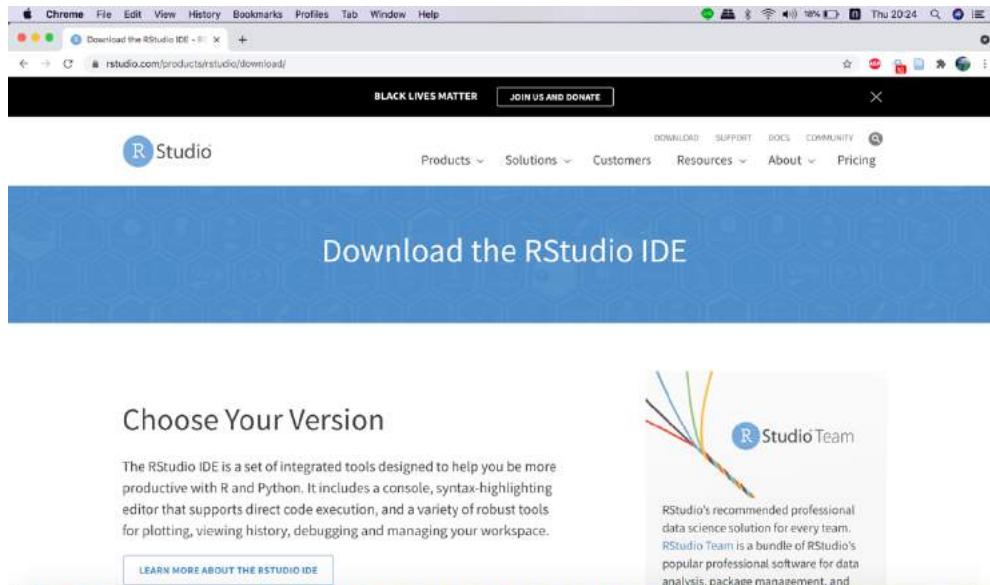
ส่วนระบบปฏิบัติการ Mac OS ให้ทำการเลือกเวอร์ชันล่าสุดคือ R-4.1.0.pkg เพื่อดาวน์โหลดโปรแกรม R (รูปที่ 2.6) เมื่อดาวน์โหลดมา ก็จะได้ไฟล์ R-4.1.0.pkg exe โดยจะเก็บอยู่ที่โฟลเดอร์ Download จากนั้นทำการ Double click ไฟล์ดังกล่าว แล้วระบบจะแสดงหน้าจอการติดตั้งโปรแกรมโดยให้กด Next ไปเรื่อยๆ และกดปุ่ม Finish เมื่อทำการติดตั้งเสร็จสิ้น



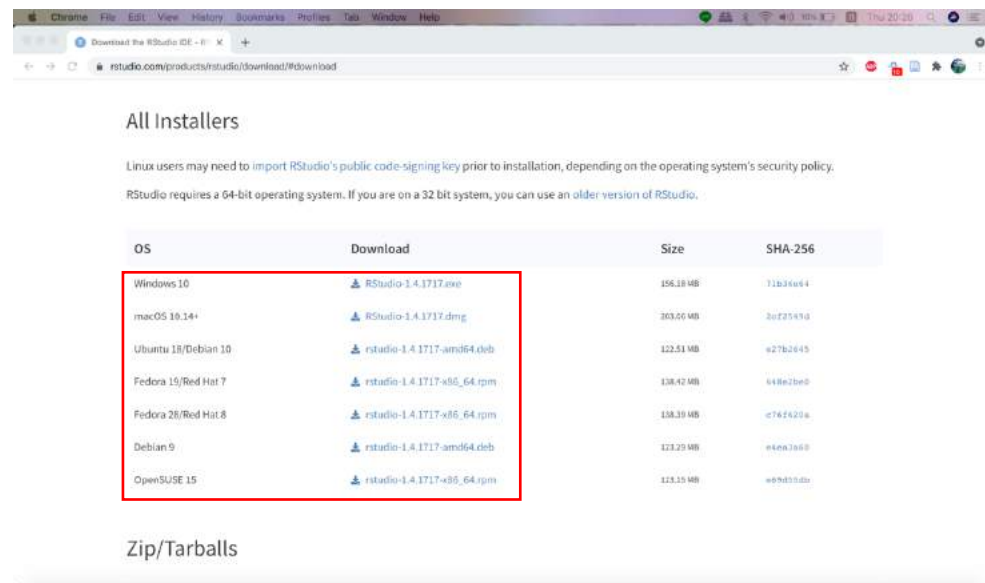
รูปที่ 2.6 หน้าต่างสำหรับติดตั้งโปรแกรม R ของระบบปฏิบัติการ Mac OS

ในการดาวน์โหลดและติดตั้งโปรแกรม RStudio จะต้องติดตั้งโปรแกรม R ก่อนเสมอ ถ้าหากไม่ทำการติดตั้ง R นั้นจะทำให้ RStudio ไม่สามารถทำงานได้ โดยการดาวน์โหลดโปรแกรม RStudio สามารถดาวน์โหลดได้จากเว็บไซต์ <https://www.rstudio.com/> (รูปที่ 2.7) แล้วเลื่อนหน้าจอลงมาด้านล่างที่ All installers แล้วเลือกดาวน์โหลดโปรแกรมที่เหมาะสมกับระบบปฏิบัติการของเครื่องคอมพิวเตอร์ โดยมีระบบปฏิบัติการต่างๆ ให้เลือกใช้ เช่น Windows, Mac OS หรือ Ubuntu/Debian เป็นต้น (รูปที่ 2.8) จากนั้นดำเนินการติดตั้งโปรแกรม

เมื่อดาวน์โหลดมาโปรแกรม RStudio ก็จะได้ไฟล์ RStudio-1.4.1717.exe สำหรับระบบปฏิบัติการ Windows หรือไฟล์ RStudio_1.4.171.dmg สำหรับระบบปฏิบัติการ Mac OS โดยจะเก็บอยู่ที่โฟลเดอร์ Download เมื่อทำการ Double click ไฟล์ดังกล่าว แล้วระบบจะแสดงหน้าจอการติดตั้งโปรแกรมโดยให้กด Next ไปเรื่อยๆ และกดปุ่ม Finish เมื่อทำการติดตั้งเสร็จสิ้น ซึ่งเมื่อทำการติดตั้งโปรแกรม RStudio เสร็จเรียบร้อยแล้วสามารถเรียกใช้โปรแกรมผ่าน Start menu สำหรับระบบปฏิบัติการ Windows หรือเรียกใช้ผ่านทาง Applications สำหรับระบบปฏิบัติการ Mac OS



รูปที่ 2.7 เว็บไซต์สำหรับดาวน์โหลดโปรแกรม RStudio



รูปที่ 2.8 หน้าต่างสำหรับดาวน์โหลดโปรแกรม RStudio

บทที่ 3

การใช้งานโปรแกรม R เบื้องต้น

3.1 การเริ่มใช้งานโปรแกรม RStudio

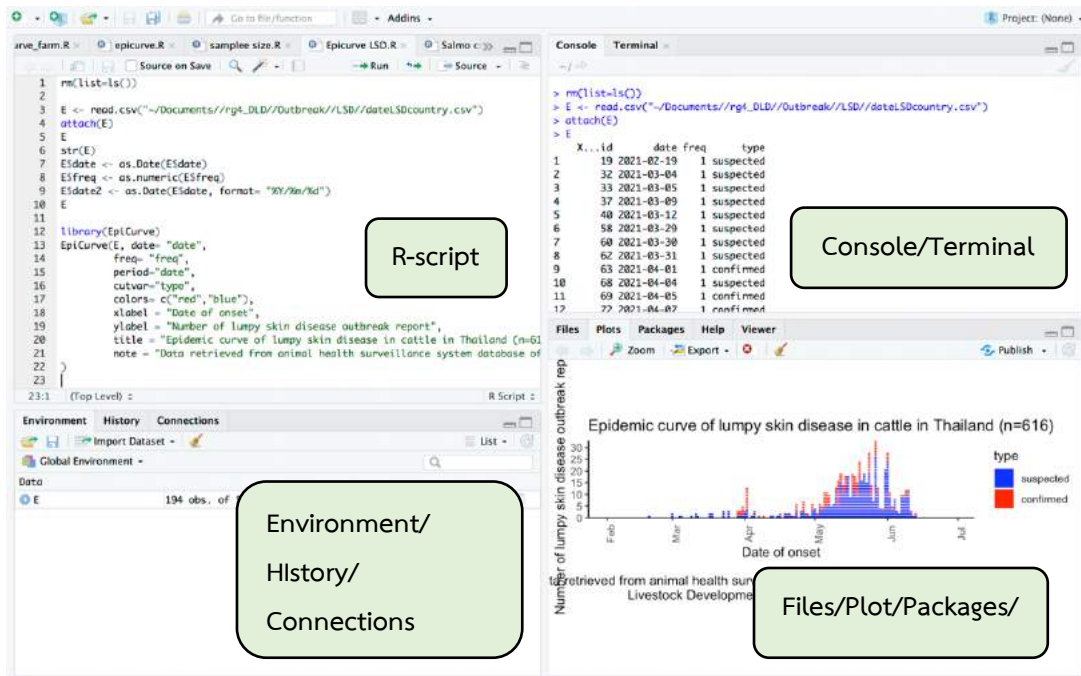
เริ่มต้นการใช้งานโดยเลือกที่โปรแกรม RStudio โดยหน้าต่างของโปรแกรมนี้อาจแบ่งออกเป็น 4 ส่วน (รูปที่ 3.1) คือ

1) R-script เป็นหน้าต่างที่ใช้สำหรับเขียนคำสั่งภาษา R ซึ่งการเรียกใช้หน้าต่างนี้สามารถทำได้โดยการคลิกเลือกที่เมนู File > New Script สำหรับระบบปฏิบัติการ Windows หรือ File > New File > R Script สำหรับระบบปฏิบัติการ Mac OS โดยเมื่อเขียนคำสั่งแล้วให้ทำการ Highlight คำสั่งนั้นๆ แล้วกด Run หรือกด Control+Enter เพื่อสั่งการให้โปรแกรมประมวลผลของคำสั่งนั้นๆ (รูปที่ 3.2)

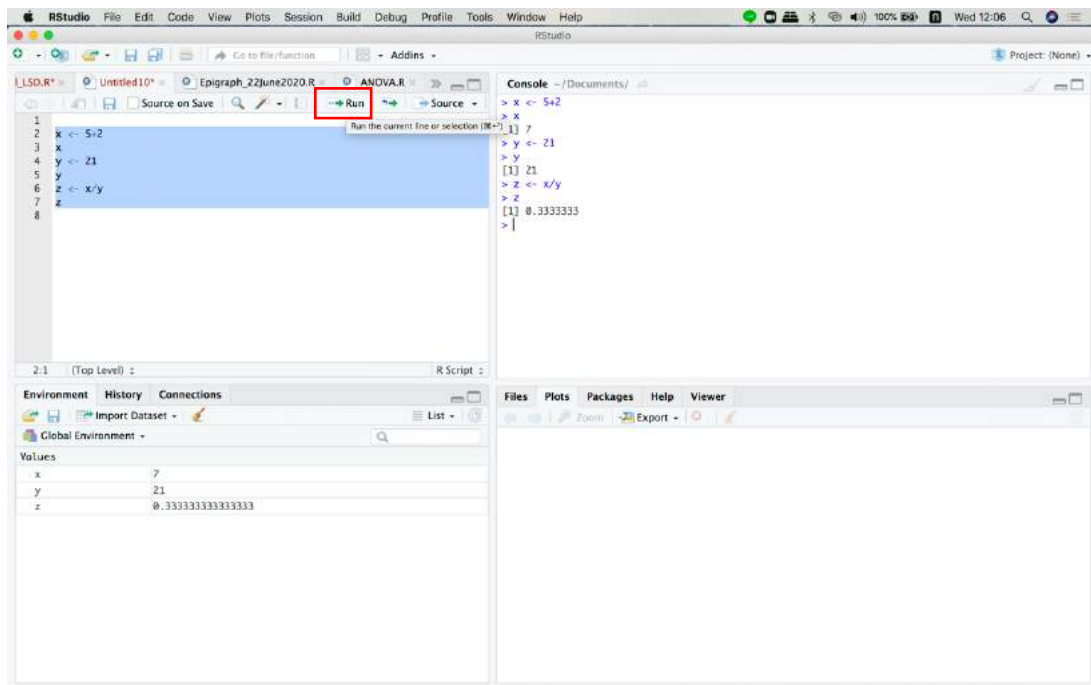
2) Console/Terminal ซึ่ง Console จะเป็นส่วนที่รับคำสั่งด้วย Command line และแสดงผลของคำสั่ง ส่วน Terminal เป็นส่วนที่ทำหน้าที่รับคำสั่งและทำหน้าที่ประมวลผลบนระบบปฏิบัติการ Dos

3) Environment/History/Connections โดย Environment เป็นส่วนที่แสดงรายละเอียดค่าหรือข้อมูลที่ได้กำหนดและเก็บไว้ในโปรแกรม ส่วน History เป็นส่วนที่แสดงประวัติการใช้งานคำสั่ง ซึ่งผู้ใช้สามารถที่จะเรียกดูหรือเรียกใช้คำสั่งเดิมได้ และ Connection เป็นส่วนที่ทำหน้าที่ติดต่อกับแหล่งข้อมูลหรือฐานข้อมูลภายนอกต่างๆ

4) Files/Plots/Packages/Help/Viewer โดย Files เป็นส่วนที่ทำหน้าที่แสดงข้อมูลปัจจุบันที่กำลังทำงานอยู่ ในขณะที่ Plots เป็นส่วนที่จะแสดงกราฟลักษณะต่างๆ ส่วน Packages เป็นส่วนที่แสดงรายการชุดคำสั่งที่ผู้ใช้สามารถเรียกใช้งาน โดยจะต้องทำการติดตั้งชุดคำสั่งต่างๆ ที่ต้องการใช้ในครั้งแรกและอัปเดตชุดคำสั่งได้เช่นเดียวกัน Help เป็นส่วนที่แสดงข้อมูลหัวข้อและรายละเอียดที่ผู้ใช้งานต้องการเพิ่มเติมในการใช้งาน และ Viewer เป็นส่วนที่แสดงผลเว็บไซต์บนหน้าจอ RStudio

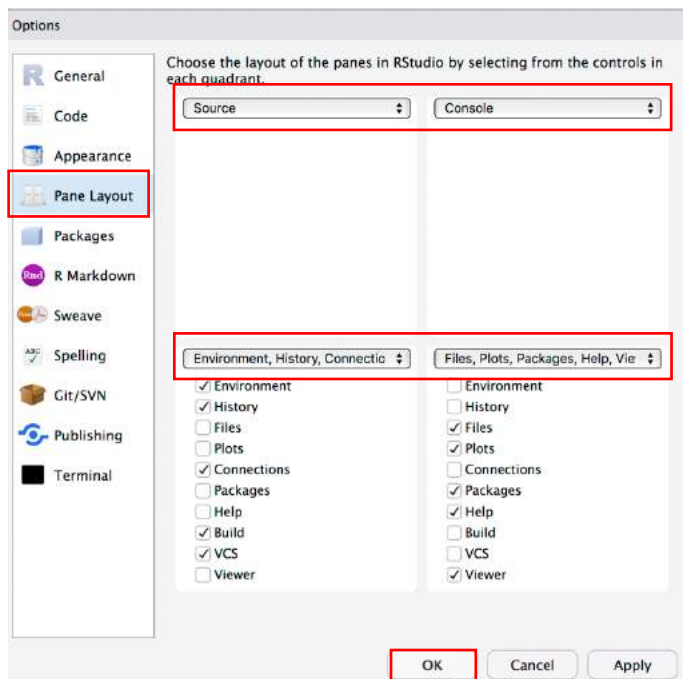


รูปที่ 3.1 ส่วนประกอบหน้าต่างการทำงานของโปรแกรม RStudio



รูปที่ 3.2 เมื่อต้องการให้คำสั่งที่ต้องการทำงาน ให้คลิกเลือกคำสั่งนั้นๆ หรือเอาเคอร์เซอร์วางไว้ในบรรทัดคำสั่งที่ต้องการ แล้วกดเลือกที่คำว่า “Run” เพื่อสั่งให้โปรแกรมวิเคราะห์และประมวลผล

ในส่วนของการตั้งค่าการแสดงผลหน้าจอในรูป 3.1 นั้นสามารถทำได้ โดยเลือกเมนู Tools ที่โปรแกรม R แล้วเลือก Global options จากนั้นจะมีหน้าต่างต่างเปิดขึ้นมาให้เลือก Pane Layout แล้วปรับค่าการแสดงผลของหน้าต่าง ดังรูปที่ 3.3



รูปที่ 3.3 การตั้งค่าหน้าต่างแสดงผลของโปรแกรม R

นอกจากนี้ผู้ใช้จะต้องทำการกำหนดพื้นที่ทำงาน (Working directory) สำหรับไฟล์ ข้อมูล หรือผลการวิเคราะห์ในรูปแบบกราฟต่างๆ ที่ต้องการนำมาใช้หรือบันทึก ก่อนเริ่มใช้งาน R โดยมีคำสั่งที่สำคัญ ดังนี้

```
getwd()
```

ใช้สำหรับตรวจสอบว่าในขณะนี้เราทำงานอยู่ที่พื้นที่ทำงานใด

```
setwd()
```

ใช้สำหรับกำหนดพื้นที่ทำงาน

- ระบบปฏิบัติการ Windows เช่น `setwd("d:FMD")`
- ระบบปฏิบัติการ Mac OS เช่น `setwd("~/Documents/FMD")`

```
dir()
```

ใช้สำหรับเรียกดูไฟล์ทั้งหมดที่จัดเก็บอยู่ในพื้นที่ทำงาน

```
Sys.setlocale(locale="Thai")
```

ใช้สำหรับตั้งค่าให้ R อ่านภาษาไทยจากข้อมูล

3.2 โครงสร้างข้อมูลในโปรแกรม R

ในการใช้งานโปรแกรม R โครงสร้างของข้อมูลที่ใช้จะมีทั้งหมด 4 แบบ คือ เวกเตอร์ (Vector) เมทริกซ์ (Matrix) อาร์เรย์ (Array) และ กรอบข้อมูล (Data frame) โดยมีรายละเอียดดังนี้

ตารางที่ 3.1 โครงสร้างของข้อมูลในโปรแกรม R

ข้อมูล	ความหมาย
เวกเตอร์ (Vector)	เป็นโครงสร้างข้อมูลพื้นฐาน
เมทริกซ์ (Matrix)	เป็นเวกเตอร์หลายๆ อัน มาประกอบกัน
อาร์เรย์ (Array)	เป็นกลุ่มของข้อมูลที่เกี่ยวข้องกัน มีจำนวนแน่นอน เป็นข้อมูลประเภทเดียวกัน
กรอบข้อมูล (Data frame)	เป็นข้อมูลที่ประกอบด้วยแถวและคอลัมน์

ตัวอย่างการใช้คำสั่งในการกำหนดแบบของข้อมูล ในโปรแกรม R เช่น

```
# การสร้างเวกเตอร์
# การสร้างเวกเตอร์ตัวแปรที่เป็นข้อความ
species <- c("dog", "cat", "dog", "chicken", "dog")
species
# ผล
[1] "dog" "cat" "dog" "chicken" "dog"
# การสร้างเวกเตอร์ตัวแปรที่เป็นตัวเลข
population <- c(11, 25, 36, 14, 9)
population
# ผล
[1] 11 25 36 14 9
# การสร้างเมทริกซ์ โดยเป็นการรวมตัวแปร species และ population และกำหนดชื่อออกเป็น farm
farm <- cbind(species,population)
farm
      species population
[1,] "dog"      "11"
```

```

# การสร้างเมทริกซ์ โดยเป็นการรวมตัวแปร species และ population และกำหนดชื่อออบเจ็คเป็น farm
farm <- cbind(species,population)

farm

# ผล

  species population
[1,] "dog"      "11"
[2,] "cat"      "25"
[3,] "dog"      "36"
[4,] "chicken" "14"
[5,] "dog"      "9"

# การสร้างกรอบข้อมูล โดยกำหนดชื่อออบเจ็คเป็น A
A <- data.frame(farm)

A

# ผล

  species population
1   dog         11
2   cat         25
3   dog         36
4 chicken         14
5   dog          9

```

หมายเหตุ: ในการใช้คำสั่งต่างๆ ในภาษา R ชื่อวัตถุหรือตัวแปร อักษรตัวเล็กหรือตัวใหญ่ จะแตกต่างกัน และสามารถใช้อักษร # เพื่อเขียนข้อความต่างๆ โดยที่ไม่ถือเป็นคำสั่ง

โดยส่วนใหญ่ข้อมูลทางระบาดวิทยาสัตว์แพทย์ จะมีโครงสร้างของข้อมูลเป็นแบบกรอบข้อมูล (Data frame) ซึ่งเป็นรูปแบบคล้ายกับตารางของโปรแกรม Excel ที่เมื่อนำเข้ามาในโปรแกรม R จะถูกเรียกว่า กรอบข้อมูล ซึ่งโดยทั่วไปเราสามารถที่จะสร้างกรอบข้อมูลขึ้นมาได้ด้วย เช่น

```
# การสร้างกรอบข้อมูล โดยกำหนดชื่อออบเจ็คเป็น id, dog, cat
id <- c(1, 2, 3, 4, 5)
dog <- c(20, 30, 15, 40, 9)
cat <- c(15, 22, 13, 25, 7)
data.frame(id, dog, cat)

# ผล
  id dog cat
1  1  20  15
2  2  30  22
3  3  15  13
4  4  40  25
5  5   9   7
```

3.3 ชนิดของข้อมูลในโปรแกรม R

ชนิดของข้อมูลในโปรแกรม R จะแบ่งออกเป็น 4 ประเภท คือ ข้อความ (Character) ตัวเลขที่มีจุดทศนิยม (Numeric) จำนวนเต็มที่ไม่มีจุดทศนิยม (Integer) และตรรกะ (Logical) โดยมีรายละเอียดดังตารางที่ 3.2 ซึ่งในการวิเคราะห์ข้อมูลจะต้องมีการกำหนดชนิดของข้อมูลให้เหมาะสมกับการวิเคราะห์นั้นด้วย

ตารางที่ 3.2 ชนิดของข้อมูลในโปรแกรม R

ประเภทข้อมูล	ตัวอย่างข้อมูล	ตัวอย่างคำสั่ง
ข้อความ (Character): ตัวอักษรต่างๆ	“Yes”, “No”, “Y”, “N”, “disease”, “non-disease”	A <- “cattle” class (A) [1] “character”
ตัวเลข (Numeric): ตัวเลขที่มีจุดทศนิยม	3.58, 20.0, 55.5	A <- 150 class (A) [1] “numeric”
จำนวนเต็ม (Integer): ตัวเลขที่ไม่มีจุดทศนิยม	2K, 6K, 8K	A <- 12K class (A) [1] “integer”
ตรรกะ (Logical)	TRUE (T) FALSE (F)	A <- TRUE class (A) [1] “logical”

3.4 คำสั่งพื้นฐานในการใช้งานภาษา R

3.4.1 การคำนวณอย่างง่าย

โปรแกรม R มีความสามารถในการคำนวณทางคณิตศาสตร์อย่างง่าย โดยใช้สัญลักษณ์พื้นฐานในการคำนวณ หรือสามารถใช้งานฟังก์ชันทางคณิตศาสตร์ สัญลักษณ์ทางตรรกศาสตร์ โดยใช้ร่วมกับค่าการกระทำต่างๆ ของภาษา R (ตารางที่ 3.3) ดังนี้

ตารางที่ 3.3 ค่าการกระทำ ฟังก์ชันทางคณิตศาสตร์ และสัญลักษณ์ทางตรรกศาสตร์ อย่างง่ายที่ใช้ในภาษา R

หมวดฟังก์ชัน	ความหมาย	สัญลักษณ์ที่ใช้ในภาษา R
ค่าการกระทำ	เป็นการนำค่าหรือผลจากการคำนวณหรือการทำงานของคำสั่งที่ปลาย	<-
	ลูกศร (ด้านขวา) ไปใส่ในชื่อวัตถุที่เขียนไว้ที่อยู่ด้านหัวลูกศร (ด้านซ้าย)	
	เป็นการนำค่าหรือผลจากการคำนวณหรือการทำงานของคำสั่งที่ปลาย	->
	ลูกศร (ด้านซ้าย) ไปใส่ในชื่อ	
	วัตถุที่เขียนไว้ที่อยู่ด้านหัวลูกศร (ด้านขวา)	
	มีความหมายเหมือนกับ <-	=
ฟังก์ชันทางคณิตศาสตร์	บวก	+
	ลบ	-
	คูณ	*
	หาร	/
	รากที่สอง	sqrt()
	ค่าลอการิทึม	log()
	ค่าเอกโพเนนเชียล	exp()
สัญลักษณ์ทางตรรกศาสตร์	การเท่ากัน	==
	การไม่เท่ากัน	!=
	น้อยกว่า	<
	มากกว่า	>
	น้อยกว่าหรือเท่ากับ	<=
	มากกว่าหรือเท่ากับ	>=
	และ	&
	หรือ	

ตัวอย่างการคำนวณค่าทางคณิตศาสตร์อย่างง่ายร่วมกับค่าการกระทำต่างๆ ในโปรแกรม R เช่น

```
# กำหนดให้ออบเจ็กต์ x เป็น 5+2
x <- 5+2
x
# ผล
[1] 7
# กำหนดให้ออบเจ็กต์ y เป็น 21
y <- 21
y
# ผล
[1] 21
# กำหนดให้ออบเจ็กต์ z เป็นค่า x หารด้วยค่า y
z <- x/y
z
# ผล
[1] 0.3333333
```

ตัวอย่างการใช้สัญลักษณ์ทางตรรกศาสตร์และค่าการกระทำต่างๆ โปรแกรม R เช่น

```
# กำหนดให้ออบเจ็กต์ x เป็น 10, y เป็น 20 และ z เป็น 10
x <- 10
y <- 20
z <- 10
# ค่า x เท่ากับค่า y
x==y
# ผล เป็นเท็จ
[1] FALSE
# ค่า x เท่ากับค่า z
x==z
# ผล เป็นจริง
[1] TRUE
```

3.4.2 คำสั่งอื่นๆ ที่สำคัญในการใช้งานในโปรแกรม R

`str()`

ใช้สำหรับแสดงโครงสร้างของข้อมูล

`class()`

ใช้สำหรับแสดงชนิดของข้อมูล

`names()`

ใช้สำหรับแสดงชื่อตัวแปรของข้อมูล

`length()`

ใช้สำหรับแสดงจำนวนของข้อมูล

`cbind()`

ใช้สำหรับรวมข้อมูลโดยคอลัมน์

`rbind()`

ใช้สำหรับรวมข้อมูลโดยแถว

`rm()`

ใช้สำหรับลบวัตถุหรือตัวแปร

`cut()`

ใช้สำหรับเปลี่ยนข้อมูลตัวเลขให้เป็นกลุ่มช่วง

`glimpse ()`

ใช้สำหรับแสดงข้อมูลในแนวคอลัมน์

`summary()`

ใช้สำหรับสรุปข้อมูลด้วยสถิติอย่างง่าย เช่น min, max, mean, median

`help()`

ใช้สำหรับแสดงข้อมูลการใช้งานตามหัวข้อที่ต้องการ

`print()`

ใช้สำหรับแสดงค่าของตัวแปรบนจอภาพ

`c ()`

ใช้สำหรับนำข้อมูลมาต่อกันเป็นเวกเตอร์ เช่น ชื่อตัวแปร <- c (ค่าข้อมูล 1, ค่าข้อมูล 2, ค่าข้อมูล 3,...)

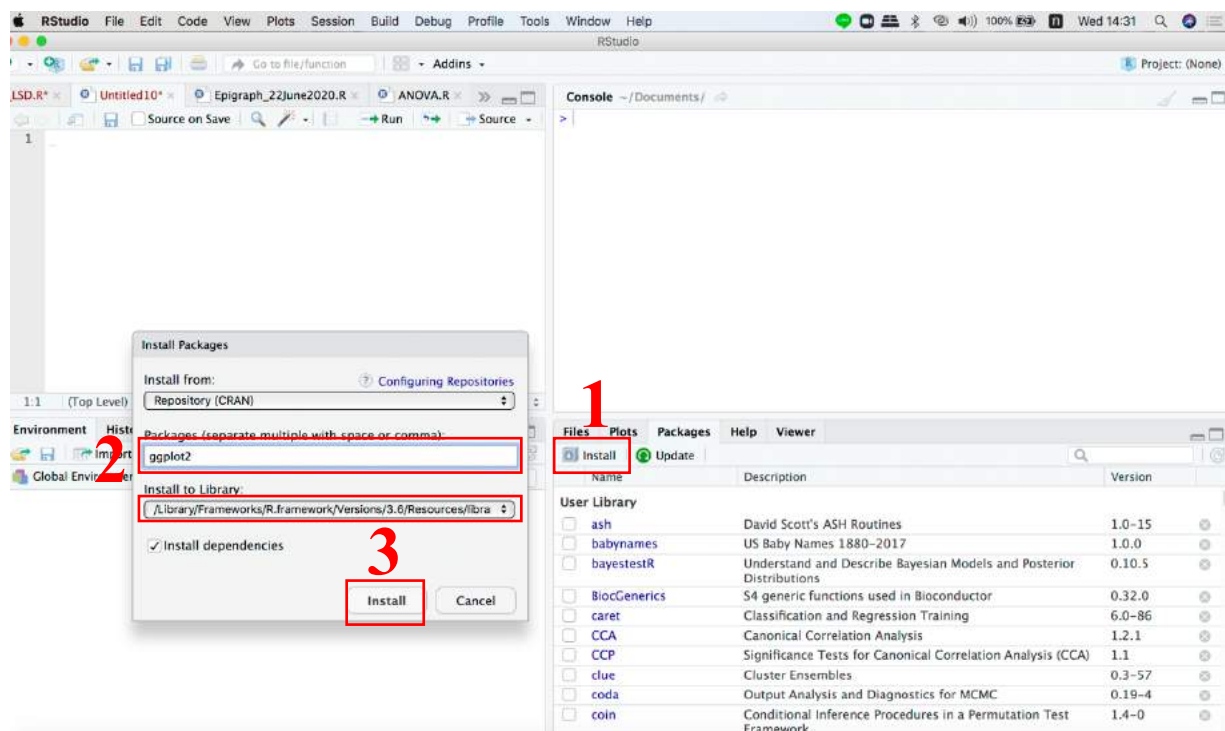
หรือ A <- c(25, 18, 34, 9, 23)

3.5 การติดตั้งชุดคำสั่งที่สำคัญสำหรับงานระบาดวิทยาทางสัตวแพทย์

ในส่วนของการติดตั้งชุดคำสั่งในโปรแกรม R พบว่าในปัจจุบันโปรแกรม R มีชุดคำสั่งมากกว่า 10,000 ชุด ที่ผู้ใช้สามารถติดตั้งและเรียกใช้ตามความเหมาะสมของการวิเคราะห์ข้อมูล ซึ่งในครั้งแรกที่ใช้งานให้ผู้ใช้ทำการติดตั้งชุดคำสั่งที่ต้องการด้วยคำสั่ง `install.packages()` หรือทำการติดตั้งชุดคำสั่งผ่านอินเทอร์เน็ต โดยไปที่เมนู Tool เลือก Install packages จากนั้นเลือก Respository (CRAN) (รูปที่ 3.4) และให้พิมพ์ชื่อชุดคำสั่งที่ต้องการติดตั้ง เมื่อต้องการเรียกใช้งานชุดคำสั่งนั้นๆ ให้ใช้คำสั่ง `library()` ทุกครั้งที่ใช้งาน เช่น

```
install.packages("ggplot2")
```

```
library(ggplot2)
```

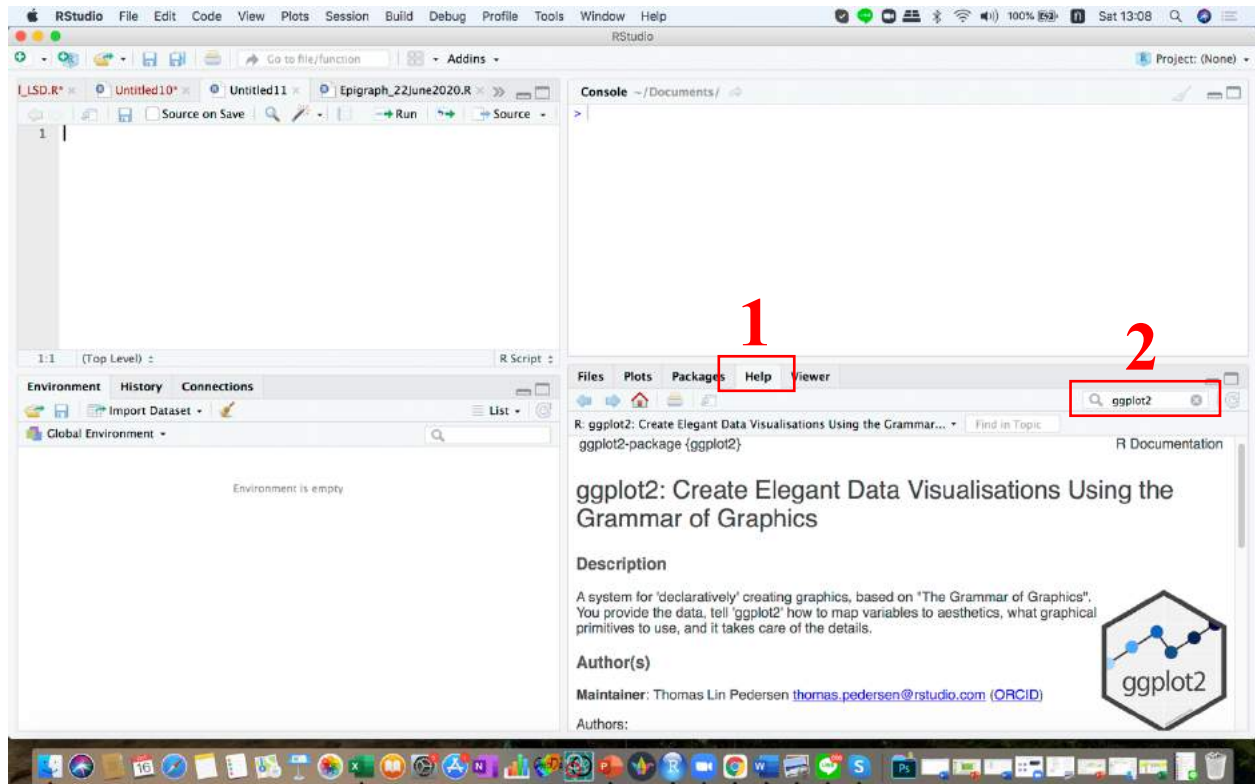


รูปที่ 3.4 การติดตั้งชุดคำสั่งผ่านระบบอินเทอร์เน็ต

ชุดคำสั่งที่ใช้สำหรับงานระบาดวิทยาทางสัตวแพทย์ สามารถประยุกต์การใช้งานจากชุดคำสั่งต่างๆ ซึ่งขึ้นอยู่กับวัตถุประสงค์ที่ผู้ใช้งานต้องการ โดยในที่นี้จะยกตัวอย่างชุดคำสั่งเบื้องต้นอย่างง่ายเกี่ยวกับการจัดการข้อมูล วิเคราะห์ข้อมูลทางระบาดวิทยา และการแสดงผลข้อมูล (ตารางที่ 3.4) โดยลักษณะของคำสั่ง (Code/Script) ของแต่ละชุดคำสั่งจะมีความแตกต่างกัน ผู้ใช้สามารถเรียนรู้และค้นหาเพิ่มเติมได้ในเว็บไซต์ต่างๆ หรือสามารถค้นหาจากหน้าต่าง Help โดยสามารถพิมพ์ชื่อชุดคำสั่งหรือคำสั่งที่ต้องการทำการค้นหาได้โดยตรง (รูปที่ 3.5) หรือพิมพ์ ?function.name หรือ help(function.name) ใน console ก็ได้เช่นเดียวกัน

ตารางที่ 3.4 ชุดคำสั่ง (Packages) เบื้องต้นที่ใช้บ่อยในการจัดการข้อมูล วิเคราะห์ข้อมูล และแสดงผลข้อมูลในงานระบาดวิทยาทางสัตวแพทย์

วัตถุประสงค์	ชื่อชุดคำสั่ง	การนำไปใช้
การจัดการข้อมูล	readxl	นำเข้าข้อมูล
	readr	นำเข้าข้อมูลขนาดใหญ่ให้รวดเร็วขึ้น
	dplyr	ตรวจสอบข้อมูลและแปลงข้อมูลให้อยู่ในรูปแบบที่ต้องการ
	purrr	ตรวจสอบข้อมูลและแปลงข้อมูลให้อยู่ในรูปแบบที่ต้องการ
	tidyverse	ตรวจสอบข้อมูลและแปลงข้อมูลให้อยู่ในรูปแบบที่ต้องการ
การวิเคราะห์ข้อมูล	Hmisc	วิเคราะห์ข้อมูลอย่างง่าย
	DataExplorer	วิเคราะห์ข้อมูลอย่างง่าย
	psych	วิเคราะห์ข้อมูลสถิติเชิงพรรณนา
	janitor	วิเคราะห์ข้อมูลสถิติเชิงพรรณนา
	epiR	วิเคราะห์ข้อมูลสถิติเชิงพรรณนาและเชิงอนุมาน
	epitools	วิเคราะห์ข้อมูลสถิติเชิงพรรณนาและเชิงอนุมาน
	gmodels	วิเคราะห์ข้อมูลสถิติเชิงพรรณนาและเชิงอนุมาน
	epiDisplay	วิเคราะห์ข้อมูลสถิติเชิงพรรณนาและเชิงอนุมาน
การแสดงผล	ploty	ทำกราฟอย่างง่าย
	ggplot2	ทำกราฟในรูปแบบที่ง่ายและซับซ้อน
	ggiraph	สร้างกราฟแบบ Interactive plot
	sf	ทำแผนที่



รูปที่ 3.5 การค้นหาคำอธิบายหรือตัวอย่างคำสั่งของชุดคำสั่งที่ต้องการ ผ่านหน้าต่าง Help

บทที่ 4

การนำเข้าข้อมูล

การนำเข้าไฟล์ข้อมูลมาใช้ในโปรแกรม R ในเบื้องต้น จะต้องทำการล้างข้อมูลที่เคยทำการวิเคราะห์ก่อนหน้านี้ แล้วกำหนดพื้นที่ทำงานเพื่อให้โปรแกรม R สามารถเข้าถึงได้ จากนั้นให้ทำการติดตั้งชุดคำสั่ง (Packages) ที่จะใช้ในการวิเคราะห์ และทำการนำเข้าข้อมูล (Import data) มาในโปรแกรม R โดยวิธีการนำเข้าข้อมูลสามารถทำได้หลายวิธีขึ้นอยู่กับชนิดข้อมูลที่มี เช่น Text (.txt), Comma separated values (.csv), Excel (.xls, .xlsx), SPSS (.sav), STATA (.dta) และ SAS (.xpt) เป็นต้น โดยในกรณีที่ข้อมูลมีขนาดเล็ก ผู้ใช้สามารถที่จะคัดลอกนำข้อมูลที่บันทึกไว้ในรูปแบบตารางมาวางในหน้าต่าง R-script ที่โปรแกรม R ได้โดยตรง หรือถ้าหากข้อมูลมีขนาดใหญ่ ผู้ใช้สามารถนำเข้าไฟล์ข้อมูลในรูปแบบต่างๆ ข้างต้นเข้ามาใช้งาน โดยจะต้องใช้คำสั่งที่เหมาะสมกับไฟล์แต่ละประเภท ซึ่งก่อนที่จะนำเข้าข้อมูลใหม่มาในโปรแกรม R จะต้องทำการล้างข้อมูล และกำหนดพื้นที่ทำงานก่อนทุกครั้ง โดยผู้ใช้งานจะต้องเตรียมไฟล์ข้อมูลที่จะใช้ในการวิเคราะห์เก็บไว้ในโฟลเดอร์เดียวกันกับพื้นที่ทำงานเสมอ

```
rm(list=ls())
```

ใช้สำหรับล้างข้อมูลเดิมที่ใช้งานในโปรแกรมก่อนหน้านี้

```
setwd ()
```

ใช้สำหรับกำหนดพื้นที่ทำงาน ซึ่งระบบปฏิบัติการ Windows และ Mac OS จะใช้คำสั่งที่แตกต่างกันและขึ้นอยู่กับแหล่งที่ผู้ใช้เก็บข้อมูล:

- ระบบปฏิบัติการ Windows: ให้ทำการตั้งโฟลเดอร์ชื่อ “R” ที่ไดรฟ์ D จากนั้นใช้คำสั่ง `setwd ("D:/R")`
- ระบบปฏิบัติการ Mac OS: ให้ทำการตั้งโฟลเดอร์ชื่อ “R” ที่ Documents จากนั้นใช้คำสั่ง `setwd ("~/Documents/R")`

นอกจากนี้ก่อนการนำเข้าข้อมูลเพื่อมาทำการวิเคราะห์ในโปรแกรม R ผู้ใช้จะต้องตรวจสอบข้อมูลที่จะนำมาใช้เบื้องต้นก่อน โดยมีข้อคำนึงดังนี้ 1) ถ้าทำงานกับข้อมูลแบบตาราง ข้อมูลแถวแรกจะต้องเป็นชื่อของคอลัมน์ 2) ควรหลีกเลี่ยงชื่อหรือค่าที่มีช่องว่าง โดยให้ใช้ “_” หรือ “.” แทนช่องว่าง เช่น `owner_name` หรือ `owner.name` หรือ `ownername` เป็นต้น 3) เขียนชื่อคอลัมน์ให้สั้น กระชับ และเข้าใจง่าย 4) หลีกเลี่ยงการใช้

ค่าที่มีสัญลักษณ์ต่างๆ เช่น ?, \$, %, ^, &, *, (,), -, #, ?, , , <, >, /, |, \, [,] , {, และ }; 4) ลบข้อความอื่นๆ ที่ไม่เกี่ยวข้องกับข้อมูลที่ต้องการวิเคราะห์ออก และ 5) ใช้คำว่า “NA” แทนค่าที่หายไป (Missing values) เป็นต้น

4.1 การนำเข้าข้อมูลขนาดเล็ก

ผู้ใช้ทั้งในระบบปฏิบัติการแบบ Windows และ Mac OS สามารถใช้คำสั่งนี้ได้โดยตรง ซึ่งอักขรแบบหน้า หมายถึง ข้อมูลที่ผู้ใช้มีอยู่ในรูปแบบ Data frame ที่มีทั้งแถวและคอลัมน์ โดยผู้ใช้สามารถคัดลอกข้อมูลที่เตรียมไว้มาวางแทนอักขรแบบหน้าได้เลย ในขณะที่กรูปร่างคำสั่งอื่นๆ ไว้ ซึ่งในที่นี้ได้ใช้อักขรแบบหน้าเพื่อใช้เป็นตัวอย่างเท่านั้น โดยในที่นี้ได้กำหนดชื่อออบเจ็คเป็น A

```
A <- read.table(header=T, text="
      Type      Proportion
      Dog       75.00
      Cat       83.33
      Pig       40.74
      Poultry   98.05
      ")
attach (A)
```

4.2 การนำเข้าข้อมูลขนาดใหญ่

การนำเข้าข้อมูลขนาดใหญ่ ในคู่มือเล่มนี้จะเน้นการนำเข้าข้อมูลที่เจ้าหน้าที่ปศุสัตว์มักจะทำการบันทึกในรูปแบบของ “.txt.” หรือ “.csv” หรือ “.xls” หรือ “.xlsx” ซึ่งในที่นี้จะอธิบายโดยแบ่งออกเป็น 2 วิธีดังนี้

4.2.1 ในกรณีที่ต้องการเลือกไฟล์ด้วยตัวเองทุกครั้ง

ไฟล์ที่สามารถใช้ในกรณีนี้จะต้องเป็นไฟล์นามสกุล “.csv” ผู้ใช้สามารถใช้คำสั่งด้านล่างนี้ เมื่อส่งคำสั่งปฏิบัติการคำสั่งแล้วจะมีหน้าต่างไฟล์เปิดขึ้นมา จากนั้นผู้ใช้อีกก็สามารถเลือกไปยังไฟล์ข้อมูลที่ต้องการนำเข้ามาวิเคราะห์ได้ ซึ่งในที่นี้จะกำหนดชื่อออบเจ็คเป็น LSD

```
LSD <- read.csv (file.choose(), header = TRUE)
```

4.2.2 ในกรณีที่ต้องการลิงค์ไปยังไฟล์เดอร์ที่เก็บข้อมูลหรือตามที่กำหนดพื้นที่ทำงานโดยตรง

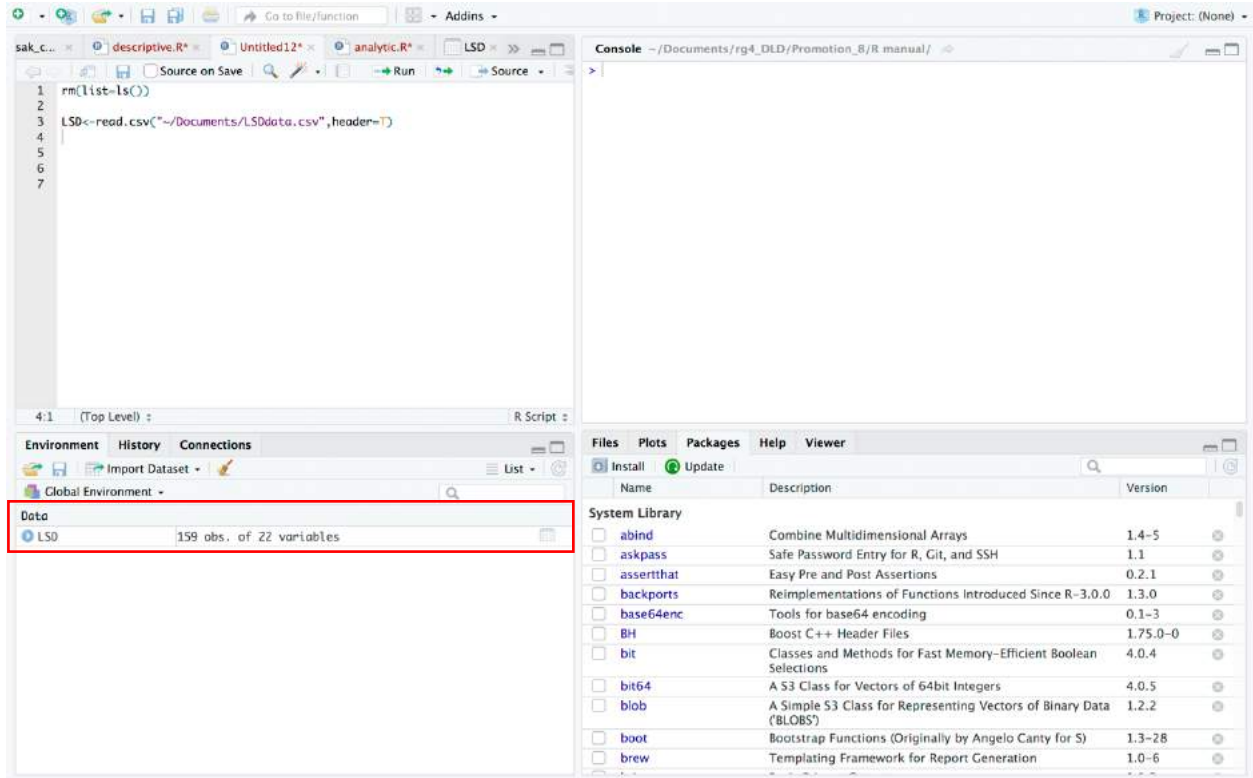
ผู้ใช้สามารถใช้คำสั่งเพื่อนำเข้าไฟล์ข้อมูลในรูปแบบต่างๆ แยกตามประเภทไฟล์ และตามระบบปฏิบัติการ ซึ่งในที่นี้จะกำหนดชื่อออบเจ็คเป็น LSD ดังแสดงในตารางที่ 4.1

ตารางที่ 4.1 คำสั่งสำหรับนำเข้าไฟล์ข้อมูลในระบบปฏิบัติการ Windows และ Mac OS

ชนิดไฟล์	ระบบปฏิบัติการ Windows	ระบบปฏิบัติการ Mac OS
.txt	LSD <- read.table ("D:/R/LSDdata.txt", header=TRUE)	LSD <- read.table ("~/Documents/R/LSDdata.txt ", header=TRUE)
.csv	LSD <- read.csv ("D:/R/LSDdata.csv", header=TRUE)	LSD <- read.csv ("~/Documents/R /LSDdata.csv", header=TRUE)
.xls หรือ .xlsx	library(readxl) # xls files LSD <- read_xls("LSDdata.xls") # xlsx files LSD <- read_xlsx("LSDdata.xlsx") หมายเหตุ: กรณีนี้ผู้ใช้จะต้องกำหนดพื้นที่ทำงานก่อน พร้อมทำการติดตั้งและเรียกใช้ชุดคำสั่ง โดยใช้ install.packages ("readxl") ในครั้งแรกที่ใช้งานและเรียกใช้ชุดคำสั่งนี้ด้วย library(readxl) ทุกครั้งที่จะนำเข้าข้อมูล	

หมายเหตุ : คำสั่ง sep="\\t" เป็นการระบุว่าชุดข้อมูล LSDdata.txt มีการคั่นแต่ละคอลัมน์ด้วย tab แต่หากชุดข้อมูลใดคั่นด้วยเครื่องหมายจุลภาค (,) ก็ต้องระบุเป็น sep=" , " , “

เมื่อนำเข้าข้อมูลเข้ามาในโปรแกรม R เรียบร้อยแล้ว จะปรากฏชื่อตัวแปร จำนวนแถว และจำนวนคอลัมน์ ที่หน้าต่าง Environment ดังรูปที่ 4.1



รูปที่ 4.1 ชื่อตัวแปรที่ถูกนำเข้ามาในโปรแกรม R เรียบร้อยแล้ว

4.3 การตรวจสอบข้อมูล

สำหรับการตรวจสอบหรือสำรวจข้อมูลที่น่าเข้ามา ถือว่าเป็นกระบวนการที่สำคัญมากของการดำเนินการตรวจสอบเบื้องต้นในการวิเคราะห์ข้อมูล ผู้ใช้สามารถใช้คำสั่ง head() หรือ tail() เพื่อแสดงข้อมูล 6 แถวแรกของชุดข้อมูลที่ระบุชื่อในวงเล็บ หรือ แสดงข้อมูล 6 แถวหลังของชุดข้อมูลที่ระบุชื่อในวงเล็บ หรือคำสั่ง View() เพื่อแสดงตารางของข้อมูลในหน้าต่างใหม่ ซึ่งข้อมูลตัวอย่างชื่อ LSD จะมีข้อมูลทั้งหมด 159 ข้อมูล เมื่อใช้คำสั่ง head(LSD) หรือ tail(LSD) หรือ View(LSD) จะแสดงผลดังนี้

head(LSD)

```
> head(LSD)
  X...id District lat lon Status AgeFarmer Sex Jobs Study HusbandryTime F_timefarm
1 1 San Kamphaeng 18.78596 99.19050 Case 47 Female Owner Secondary 20 more5
2 2 San Kamphaeng 18.73208 99.12086 Control 64 Male Owner Secondary 30 more5
3 3 San Kamphaeng 18.70780 99.16715 Case 61 Male Owner Secondary 29 more5
4 4 Mae On 18.78231 99.26404 Control 57 Male Owner Secondary 22 more5
5 5 Mae On 18.78160 99.25924 Control 65 Male Owner Primary 15 more5
6 6 Mae On 18.79565 99.24526 Control 65 Female Owner Primary 20 more5
  totalcattle casenumber Farmtype F_Herdsize F_pregnant F_calfinfarm F_mixbreed F_OthersSpecies F_Sharebull
1 15 3 more5 more5 0 1 0 1 0
2 13 NA more5 more5 1 0 0 0 0
3 12 2 more5 more5 0 0 0 1 0
4 9 NA more5 more5 1 1 0 0 0
5 8 NA more5 more5 1 0 0 0 1
6 7 NA more5 more5 0 0 0 1 0
  F.ManyInseminator F_insect
1 1 1
2 0 1
3 0 1
4 0 1
5 0 1
6 0 1
```

tail(LSD)

```
> tail(LSD)
  X...id District lat lon Status AgeFarmer Sex Jobs Study HusbandryTime F_timefarm
154 154 San Kamphaeng 18.73754 99.11965 Control 56 Female Owner Secondary 5 more5
155 155 San Kamphaeng 18.73117 99.12720 Case 51 Male Owner Primary 8 more5
156 156 San Kamphaeng 18.73113 99.13886 Control 82 Female Owner Primary 10 more5
157 157 San Kamphaeng 18.73193 99.12054 Control 48 Male Owner Secondary 5 more5
158 158 San Kamphaeng 18.72749 99.13254 Control 66 Female Owner Primary 5 more5
159 159 San Kamphaeng 18.73105 99.11763 Case 39 Female Owner Secondary 5 more5
  totalcattle casenumber Farmtype F_Herdsize F_pregnant F_calfinfarm F_mixbreed F_OthersSpecies F_Sharebull
154 2 NA more5 1to5 1 0 0 0 1
155 2 4 more5 1to5 0 0 0 1 0
156 1 NA more5 1to5 1 0 0 1 0
157 1 NA more5 1to5 0 0 0 1 0
158 1 NA more5 1to5 0 0 1 1 0
159 1 1 more5 1to5 1 0 0 1 0
  F.ManyInseminator F_insect
154 0 1
155 0 0
156 0 1
157 0 1
158 0 1
159 1 0
```


View(LSD)

X_id	District	lat	lon	Status	AgeFarmer	Sex	Jobs	Study	HusbandryTime	F_timefarm	totalcattle	casenumber	Farmtype	F_Herdsiz
1	San Kamphaeng	18.78596	99.19050	Case	47	Female	Owner	Secondary	20	more5	15	3	large	more5
2	San Kamphaeng	18.73208	99.12086	Control	64	Male	Owner	Secondary	30	more5	13	N/A	large	more5
3	San Kamphaeng	18.70780	99.16715	Case	61	Male	Owner	Secondary	29	more5	12	2	large	more5
4	Mae On	18.78231	99.26404	Control	57	Male	Owner	Secondary	22	more5	9	N/A	large	more5
5	Mae On	18.78160	99.25924	Control	65	Male	Owner	Primary	15	more5	8	N/A	large	more5
6	Mae On	18.79565	99.24526	Control	65	Female	Owner	Primary	20	more5	7	N/A	large	more5
7	Mae On	18.78122	99.25905	Case	65	Male	Owner	Primary	20	more5	7	1	large	more5
8	Mae On	18.78371	99.25979	Case	47	Male	Owner	Primary	17	more5	7	1	large	more5
9	Mae On	18.74941	99.24760	Control	56	Male	Owner	Primary	30	more5	6	N/A	large	more5
10	Mae On	18.78665	99.24816	Control	64	Male	Owner	Primary	26	more5	6	N/A	large	more5
11	Mae On	18.76248	99.25978	Case	66	Male	Owner	Primary	20	more5	6	2	large	more5
12	Mae On	18.78565	99.24882	Control	69	Male	Owner	Primary	20	more5	6	N/A	large	more5
13	Mae On	18.77003	99.25782	Case	45	Male	Owner	Secondary	20	more5	6	1	large	more5
14	Mae On	18.81444	99.25803	Control	64	Male	Owner	Primary	20	more5	6	N/A	large	more5
15	Mae On	18.78590	99.26496	Case	50	Male	Owner	Primary	15	more5	6	1	large	more5
16	Mae On	18.78089	99.26451	Control	57	Male	Owner	Diploma	15	more5	6	N/A	large	more5
17	Mae On	18.79700	99.25849	Control	57	Male	Owner	Primary	15	more5	6	N/A	large	more5
18	Mae On	18.81292	99.25456	Case	51	Female	Owner	Primary	20	more5	5	2	small	1to5
19	Mae On	18.63932	99.28559	Control	52	Male	Owner	Tertiary	16	more5	5	N/A	small	1to5
20	Mae On	18.76043	99.26214	Case	51	Male	Owner	Primary	20	more5	4	1	small	1to5
21	Mae On	18.60416	99.26463	Control	63	Male	Owner	Secondary	30	more5	4	N/A	small	1to5
22	Mae On	18.58249	99.26370	Case	54	Male	Owner	Secondary	25	more5	4	2	small	1to5
23	San Sai	18.88444	99.06861	Case	55	Female	Owner	Primary	25	more5	4	3	small	1to5
24	San Sai	18.88003	99.08250	Case	60	Male	Owner	Diploma	20	more5	4	4	small	1to5
25	San Sai	19.04547	98.96394	Control	53	Female	Owner	Primary	20	more5	4	N/A	small	1to5

หรืออาจใช้คำสั่ง `names()` ซึ่งใช้สำหรับแสดงชื่อตัวแปรของชุดข้อมูลที่ระบุชื่อในวงเล็บ และใช้เพื่อการเปลี่ยนชื่อตัวแปร ซึ่งในตัวอย่างนี้ข้อมูล LSD ประกอบไปด้วย 22 คอลัมน์

names(LSD)

```
> names(LSD)
[1] "X...id"      "District"    "lat"         "lon"
[5] "Status"     "AgeFarmer"  "Sex"         "Jobs"
[9] "Study"      "HusbandryTime" "F_timefarm"  "totalcattle"
[13] "casenumber" "Farmtype"    "F_Herdsiz"   "F_pregnant"
[17] "F_calfinfarm" "F_mixbreed"  "F_OthersSpicies" "F_Sharebull"
[21] "F.ManyInseminator" "F_insect"
> |
```

ส่วนคำสั่ง `str()` จะใช้เพื่อให้แสดงรายละเอียดของชุดข้อมูลที่ระบุชื่อในวงเล็บ โดยระบุจำนวนข้อมูลจำนวนตัวแปร และแสดงรายละเอียดของตัวแปรทุกตัว ประกอบด้วยชื่อตัวแปร ชนิดของตัวแปร และแสดงค่าของตัวแปร

```
str(LSD)
```

```
> str(LSD)
'data.frame': 159 obs. of 22 variables:
 $ X...id      : int  1 2 3 4 5 6 7 8 9 10 ...
 $ District    : chr   "San Kamphaeng" "San Kamphaeng" "San Kamphaeng" "Mae On" ...
 $ lat         : num   18.8 18.7 18.7 18.8 18.8 ...
 $ lon         : num   99.2 99.1 99.2 99.3 99.3 ...
 $ Status      : chr   "Case" "Control" "Case" "Control" ...
 $ AgeFarmer   : int   47 64 61 57 65 65 65 47 56 64 ...
 $ Sex         : chr   "Female" "Male" "Male" "Male" ...
 $ Jobs        : chr   "Owner" "Owner" "Owner" "Owner" ...
 $ Study       : chr   "Secondary" "Secondary" "Secondary" "Secondary" ...
 $ HusbandryTime : int   20 30 29 22 15 20 20 17 30 26 ...
 $ F_timefarm  : chr   "more5" "more5" "more5" "more5" ...
 $ totalcattle : int   15 13 12 9 8 7 7 7 6 6 ...
 $ casenumber  : int    3 NA 2 NA NA NA 1 1 NA NA ...
 $ Farmtype    : chr   "large" "large" "large" "large" ...
 $ F_Herdsiz   : chr   "more5" "more5" "more5" "more5" ...
 $ F_pregnant  : int    0 1 0 1 1 0 1 0 1 0 ...
 $ F_calfinfarm : int    1 0 0 1 0 0 1 1 1 1 ...
 $ F_mixbreed  : int    0 0 0 0 0 0 0 0 0 0 ...
 $ F_OthersSpicies : int    1 0 1 0 0 1 0 1 1 0 ...
 $ F_Sharebull : int    0 0 0 0 1 0 1 0 0 0 ...
 $ F.ManyInseminator: int    1 0 0 0 0 0 0 0 0 0 ...
 $ F_insect    : int    1 1 1 1 1 1 1 0 1 1 ...
>
```

หรืออีกทางเลือกหนึ่งคือ ผู้ใช้สามารถเลือกใช้ชุดคำสั่ง `tidyverse` ซึ่งเป็นที่นิยมเนื่องจากใช้งานง่ายและครอบคลุมตั้งแต่การนำเข้าข้อมูล, ดัดแปลงรูปแบบข้อมูล, เลือกเฉพาะบางข้อมูลหรือรวมกลุ่มข้อมูล เป็นต้น ซึ่งในที่นี้ก็สามารถนำมาใช้เพื่อตรวจสอบข้อมูลที่นำเข้ามาโดยใช้คำสั่งดังต่อไปนี้

```
install.packages("tidyverse")
```

```
library(tidyverse)
```

```
LSD %>% glimpse()
```

*ข้อมูลนี้ประกอบด้วยค่าสังเกต (จำนวนแถว) 159 และ 22 ตัวแปร (จำนวนคอลัมน์)

```

> library(tidyverse)
> LSD %>% glimpse()
Rows: 159
Columns: 22
$ X...id      <int> 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 2~
$ District    <chr> "San Kamphaeng", "San Kamphaeng", "San Kamphaeng", "Mae On", "Mae On", "Mae On", "Mae On", "Mae ~
$ lat         <dbl> 18.78596, 18.73208, 18.70780, 18.78231, 18.78160, 18.79565, 18.78122, 18.78371, 18.74941, 18.786~
$ lon         <dbl> 99.19050, 99.12086, 99.16715, 99.26404, 99.25924, 99.24526, 99.25905, 99.25979, 99.24760, 99.248~
$ Status      <chr> "Case", "Control", "Case", "Control", "Control", "Control", "Case", "Case", "Control", "Control"~
$ AgeFarmer   <int> 47, 64, 61, 57, 65, 65, 65, 47, 56, 64, 66, 69, 45, 64, 50, 57, 57, 51, 52, 51, 63, 54, 55, 60, ~
$ Sex         <chr> "Female", "Male", "Male", "Male", "Male", "Female", "Male", "Male", "Male", "Male", "Male", "Mal~
$ Jobs        <chr> "Owner", "Owner", "Owner", "Owner", "Owner", "Owner", "Owner", "Owner", "Owner", "Owner", "Owner", "Owner~
$ Study       <chr> "Secondary", "Secondary", "Secondary", "Secondary", "Secondary", "Primary", "Primary", "Primary", "Primary", ~
$ HusbandryTime <int> 20, 30, 29, 22, 15, 20, 20, 17, 30, 26, 20, 20, 20, 20, 15, 15, 15, 20, 16, 20, 30, 25, 25, 20, ~
$ F_timefarm  <chr> "more5", "more5", "more5", "more5", "more5", "more5", "more5", "more5", "more5", "more5", "more5", "more5~
$ totalcattle <int> 15, 13, 12, 9, 8, 7, 7, 7, 6, 6, 6, 6, 6, 6, 6, 5, 5, 4, 4, 4, 4, 4, 4, 4, 4, 4, ~
$ casenumber  <int> 3, NA, 2, NA, NA, NA, 1, 1, NA, NA, 2, NA, 1, NA, 1, NA, NA, 2, NA, 1, NA, 2, 3, 4, NA, NA, 1, 1~
$ Farmlandtype <chr> "large", "large", "large", "large", "large", "large", "large", "large", "large", "large", "large", "large~
$ F_Herdsize  <chr> "more5", "more5", "more5", "more5", "more5", "more5", "more5", "more5", "more5", "more5", "more5", "more5~
$ F_pregnant  <int> 0, 1, 0, 1, 1, 0, 1, 0, 1, 0, 1, 1, 1, 0, 1, 0, 0, 1, 1, 1, 1, 1, 0, 1, 1, 1, 0, 0, 1, 1, 0, ~
$ F_calfinfarm <int> 1, 0, 0, 1, 0, 0, 1, 1, 1, 1, 1, 1, 0, 1, 0, 0, 0, 0, 1, 1, 1, 0, 0, 1, 1, 1, 0, 0, 1, 1, 0, ~
$ F_mixbreed  <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 1, 0, 1, 0, 0, 0, 1, 0, 0, 0, 0, ~
$ F_OthersSpecies <int> 1, 0, 1, 0, 0, 1, 0, 1, 1, 0, 1, 1, 1, 1, 0, 0, 1, 1, 1, 1, 1, 1, 1, 1, 0, 0, 1, 0, 0, 1, 1, ~
$ F_Sharebull <int> 0, 0, 0, 0, 1, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 1, 0, 0, 1, 0, ~
$ F.ManyInseminator <int> 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 1, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, ~
$ F_insect    <int> 1, 1, 1, 1, 1, 1, 1, 0, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 0, 1, ~
>

```

บทที่ 5

การจัดการข้อมูล

งานระบาดวิทยาทางสัตวแพทย์ของกรมปศุสัตว์โดยทั่วไปมักจะเกี่ยวข้องกับงานสอบสวนโรคระบาดในสัตว์และงานเฝ้าระวังโรคในสัตว์ รวมไปถึงจนถึงการศึกษาวิจัยต่างๆ ที่เกี่ยวข้อง ซึ่งข้อมูลที่ถูกบันทึกลงในฐานข้อมูลต่างๆ นั้นจะเรียกว่า ข้อมูลดิบ (Raw data) คือ ข้อมูลที่ไม่ได้ผ่านขั้นตอนก่อนกระบวนการประมวลผล (Pre-processing) มาจากแหล่งข้อมูลได้หลายรูปแบบ ซึ่งมักจะมีปัญหาด้านคุณภาพของข้อมูล (Data quality) เช่น การมีข้อมูลไม่สอดคล้องกัน ข้อมูลซ้ำ ข้อมูลขาดหาย รูปแบบของข้อมูลไม่ตรงกัน หรือข้อมูลผิดปกติ ทำให้ยากต่อการวิเคราะห์ในขั้นตอนต่อไป ซึ่ง Data wrangling เป็นกระบวนการเปลี่ยนข้อมูลดิบให้อยู่ในรูปแบบที่ง่ายต่อการใช้งาน สามารถทำได้หลายวิธีโดยขึ้นอยู่กับลักษณะข้อมูล และเป้าหมายในการใช้งาน ซึ่งก่อนที่เราจะนำข้อมูลไปใช้งาน ผู้ใช้จึงต้องทำการจัดการข้อมูลเพื่อให้ข้อมูลจะได้อยู่ในรูปแบบที่เหมาะสม เข้าใจง่าย และสามารถนำไปใช้ให้เกิดประโยชน์มากที่สุด ซึ่งในบทนี้จะอธิบายถึงการจัดการข้อมูลในวิธีการต่างๆ เพื่อเตรียมข้อมูลให้เป็นรูปแบบเดียวกัน ซึ่งการจัดการข้อมูลเบื้องต้นจากชุดคำสั่ง dplyr โดยในชุดคำสั่งนี้จะมีคำสั่งที่นิยมใช้ในการจัดการข้อมูล ดังรูปที่ 5.1 ได้แก่

`select()`

ใช้สำหรับเลือกคอลัมน์ที่ต้องการ

`filter()`

ใช้สำหรับฟิลเตอร์แถวที่เราต้องการด้วยเงื่อนไขต่างๆ โดยการสร้างเงื่อนไขสามารถสร้างได้มากกว่าหนึ่งข้อ โดยใช้ & (AND) และ | (OR) operators มาช่วย

`arrange()`

ใช้สำหรับเรียงข้อมูล จากมากไปหาน้อย (Descending order) หรือจากน้อยไปหามาก (Ascending order)

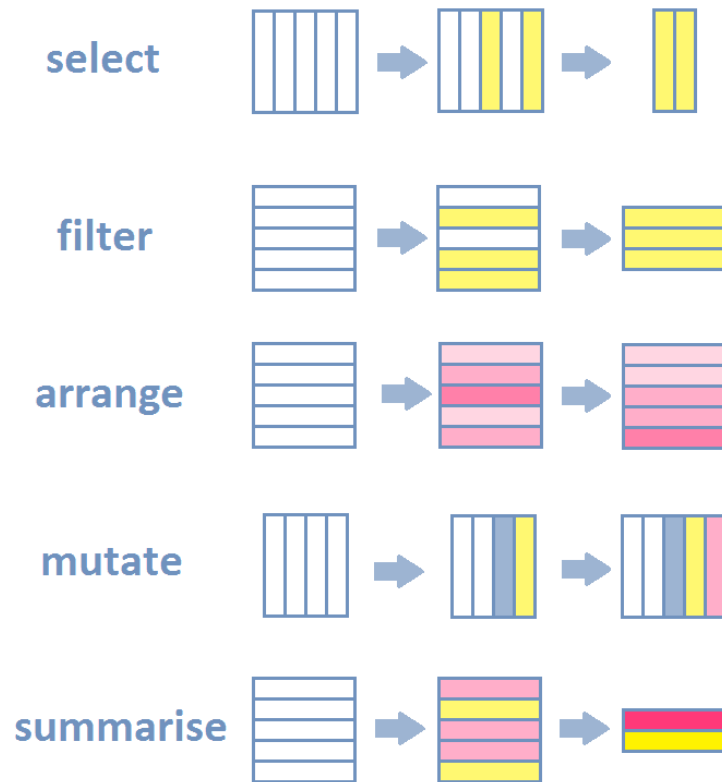
`mutate()`

ใช้สำหรับสร้างคอลัมน์หรือตัวแปรใหม่

`summarise()`

ใช้สำหรับสรุปผลสถิติของคอลัมน์ที่เราเลือก

หมายเหตุ ในชุดคำสั่งนี้จะใช้สัญลักษณ์ pipe operator (`%>%`) เพื่อเขียนเชื่อมคำสั่งเข้าด้วยกัน



รูปที่ 5.1 ลักษณะการจัดการข้อมูลด้วยชุดคำสั่ง dplyr

ตัวอย่างการใช้คำสั่ง มีดังนี้

```
install.packages("dplyr")
library(dplyr)
# กรณีที่ต้องการเลือกคอลัมน์ที่ต้องการ โดยกำหนดให้ออบเจ็กต์ LSD_con มีเฉพาะตัวแปรแบบต่อเนื่อง
# คือ AgeFarmer, HusbandryTime, totalcattle
LSD_con <- LSD %>%
  dplyr::select(AgeFarmer, HusbandryTime, totalcattle)
LSD_con
```

	AgeFarmer	HusbandryTime	totalcattle
1	47	20	15
2	64	30	13
3	61	29	12
4	57	22	9
5	65	15	8
6	65	20	7
7	65	20	7
8	47	17	7
9	56	30	6
10	64	26	6
11	66	20	6
12	69	20	6
13	45	20	6
14	64	20	6
15	50	15	6
16	57	15	6
17	57	15	6
18	51	20	5
19	52	16	5
20	51	20	4
21	63	30	4
22	54	25	4
23	55	25	4
24	60	20	4
25	53	20	4

Showing 1 to 26 of 159 entries

กรณีที่ต้องการเลือกคอลัมน์ที่ต้องการ โดยกำหนดให้ออบเจ็ค LSD_cat มีเฉพาะตัวแปรแบ่งกลุ่มคือ Status, Farmtype , F_pregnant, F_calfinfarm, F_mixbreed, F_OthersSpicies,F_Sharebull, F_Sharebull, F.ManyInseminator, F_insect

```
LSD_cat <- LSD %>%
```

```
  dplyr::select(Status, Farmtype , F_pregnant, F_calfinfarm, F_mixbreed,
                F_OthersSpicies,F_Sharebull, F_Sharebull, F.ManyInseminator,
                F_insect)
```

```
LSD_cat
```

หรือสามารถเลือกคอลัมน์เป็นตัวเลขลำดับที่ของคอลัมน์ คือ คอลัมน์ลำดับที่ 5, 14-22

```
LSD_cat <- LSD_cat <- LSD %>%
```

```
  dplyr::select(5, 14:22)
```

```
LSD_cat
```

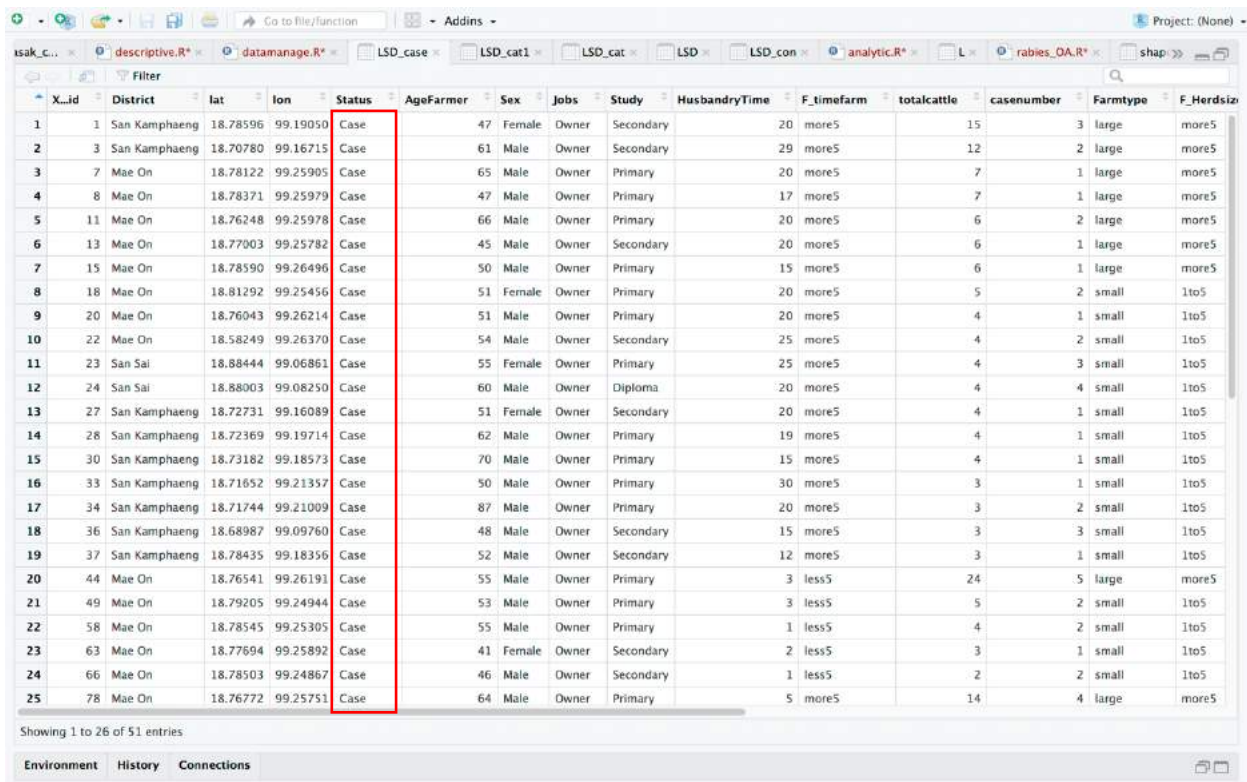
Filter										
	Status	Farmtype	F_pregnant	F_calfinfarm	F_mixbreed	F_OthersSpicies	F_Sharebull	F.ManyInseminator	F_insect	
1	Case	large	0	1	0	1	0	1	1	
2	Control	large	1	0	0	0	0	0	0	1
3	Case	large	0	0	0	1	0	0	0	1
4	Control	large	1	1	0	0	0	0	0	1
5	Control	large	1	0	0	0	1	1	0	1
6	Control	large	0	0	0	1	0	0	0	1
7	Case	large	1	1	0	0	1	0	0	1
8	Case	large	0	1	0	1	0	0	0	0
9	Control	large	1	1	0	1	0	1	0	1
10	Control	large	0	1	0	0	0	0	0	1
11	Case	large	1	1	0	1	0	0	0	1
12	Control	large	1	0	0	1	0	0	0	1
13	Case	large	1	1	0	1	0	0	0	1
14	Control	large	0	0	0	1	0	0	0	1
15	Case	large	1	0	1	0	0	0	0	1
16	Control	large	0	0	0	0	0	0	0	1
17	Control	large	0	0	0	1	0	0	0	1
18	Case	small	1	1	0	1	0	1	1	1
19	Control	small	1	1	0	1	0	0	0	1
20	Case	small	1	1	1	1	0	0	0	1
21	Control	small	1	0	0	0	0	1	1	1
22	Case	small	1	0	1	1	0	0	0	1
23	Case	small	0	1	1	1	1	0	0	1
24	Case	small	1	1	0	1	0	0	0	1
25	Control	small	1	1	0	1	0	0	0	1

Showing 1 to 26 of 159 entries

กรณีที่ต้องการเลือกแถวที่มีค่าหรือข้อความที่ต้องการ โดยกำหนดให้ออบเจ็ค LSD_case มีแถวที่มีข้อความเท่ากับ Case

```
LSD_case <- filter(LSD, Status == "Case")
```

```
LSD_case
```



The screenshot shows the RStudio interface with a data frame named `LSD_case` loaded. The data frame has 26 rows and 14 columns. The columns are: `X_id`, `District`, `lat`, `lon`, `Status`, `AgeFarmer`, `Sex`, `Jobs`, `Study`, `HusbandryTime`, `F_timefarm`, `totalcattle`, `casenumber`, `Farmtype`, and `F_Herdsiz`. The `Status` column is highlighted in red, indicating that all rows in this filtered dataset have a status of "Case".

X_id	District	lat	lon	Status	AgeFarmer	Sex	Jobs	Study	HusbandryTime	F_timefarm	totalcattle	casenumber	Farmtype	F_Herdsiz
1	San Kamphaeng	18.78596	99.19050	Case	47	Female	Owner	Secondary	20	more5	15	3	large	more5
2	San Kamphaeng	18.70780	99.16715	Case	61	Male	Owner	Secondary	29	more5	12	2	large	more5
3	Mae On	18.78122	99.25905	Case	65	Male	Owner	Primary	20	more5	7	1	large	more5
4	Mae On	18.78371	99.25979	Case	47	Male	Owner	Primary	17	more5	7	1	large	more5
5	Mae On	18.76248	99.25978	Case	66	Male	Owner	Primary	20	more5	6	2	large	more5
6	Mae On	18.77003	99.25782	Case	45	Male	Owner	Secondary	20	more5	6	1	large	more5
7	Mae On	18.78590	99.26496	Case	50	Male	Owner	Primary	15	more5	6	1	large	more5
8	Mae On	18.81292	99.25456	Case	51	Female	Owner	Primary	20	more5	5	2	small	1to5
9	Mae On	18.76043	99.26214	Case	51	Male	Owner	Primary	20	more5	4	1	small	1to5
10	Mae On	18.58249	99.26370	Case	54	Male	Owner	Secondary	25	more5	4	2	small	1to5
11	San Sai	18.88444	99.06861	Case	55	Female	Owner	Primary	25	more5	4	3	small	1to5
12	San Sai	18.88003	99.08250	Case	60	Male	Owner	Diploma	20	more5	4	4	small	1to5
13	San Kamphaeng	18.72731	99.16089	Case	51	Female	Owner	Secondary	20	more5	4	1	small	1to5
14	San Kamphaeng	18.72369	99.19714	Case	62	Male	Owner	Primary	19	more5	4	1	small	1to5
15	San Kamphaeng	18.73182	99.18573	Case	70	Male	Owner	Primary	15	more5	4	1	small	1to5
16	San Kamphaeng	18.71652	99.21357	Case	50	Male	Owner	Primary	30	more5	3	1	small	1to5
17	San Kamphaeng	18.71744	99.21009	Case	87	Male	Owner	Primary	20	more5	3	2	small	1to5
18	San Kamphaeng	18.68987	99.09760	Case	48	Male	Owner	Secondary	15	more5	3	3	small	1to5
19	San Kamphaeng	18.78435	99.18356	Case	52	Male	Owner	Secondary	12	more5	3	1	small	1to5
20	Mae On	18.76541	99.26191	Case	55	Male	Owner	Primary	3	less5	24	5	large	more5
21	Mae On	18.79205	99.24944	Case	53	Male	Owner	Primary	3	less5	5	2	small	1to5
22	Mae On	18.78545	99.25305	Case	55	Male	Owner	Primary	1	less5	4	2	small	1to5
23	Mae On	18.77694	99.25892	Case	41	Female	Owner	Secondary	2	less5	3	1	small	1to5
24	Mae On	18.78503	99.24867	Case	46	Male	Owner	Secondary	1	less5	2	2	small	1to5
25	Mae On	18.76772	99.25751	Case	64	Male	Owner	Primary	5	more5	14	4	large	more5

Showing 1 to 26 of 51 entries

Environment History Connections

กรณีที่ต้องการเรียงลำดับแถว ในที่นี้คือตัวแปร AgeFarmer

```
LSD_agefarmer <- LSD %>% arrange(AgeFarmer)
```

```
LSD_agefarmer
```

The screenshot shows the RStudio interface with a data frame loaded. The 'AgeFarmer' column is highlighted with a red box. The data frame contains 159 entries, showing columns for District, lat, lon, Status, AgeFarmer, Sex, Jobs, Study, HusbandryTime, F_timefarm, totalcattle, casenumber, Farmtype, F_Herdsize, and F_preg.

	District	lat	lon	Status	AgeFarmer	Sex	Jobs	Study	HusbandryTime	F_timefarm	totalcattle	casenumber	Farmtype	F_Herdsize	F_preg
29	San Sai	18.97857	98.96994	Control	32	Male	Owner	Diploma	6	more5		3	NA	small	1to5
30	Mae On	18.78508	99.24890	Control	36	Male	Owner	Diploma	1	less5		4	NA	small	1to5
15	San Sai	19.07528	98.94000	Control	36	Male	Owner	Tertiary	6	more5		4	NA	small	1to5
23	Mae On	18.76621	99.26342	Control	37	Female	Owner	Diploma	2	less5		2	NA	small	1to5
32	Mae On	18.76257	99.26373	Case	37	Male	Owner	Secondary	6	more5		8	3	large	more5
21	Mae On	18.64382	99.28535	Control	38	Male	Owner	Secondary	1	less5		2	NA	small	1to5
17	San Sai	18.97262	98.97960	Control	38	Male	Owner	Diploma	5	more5		4	NA	small	1to5
17	Mae On	18.79620	99.24649	Control	39	Female	Owner	Secondary	3	less5		7	NA	large	more5
19	San Kamphaeng	18.73105	99.11763	Case	39	Female	Owner	Secondary	5	more5		1	1	small	1to5
36	Mae On	18.76621	99.25627	Case	40	Female	Owner	Secondary	5	more5		9	2	large	more5
14	San Sai	18.86222	99.07417	Control	40	Female	Owner	Primary	5	more5		3	NA	small	1to5
33	Mae On	18.77694	99.25892	Case	41	Female	Owner	Secondary	2	less5		3	1	small	1to5
39	Mae On	18.78528	99.24786	Control	42	Male	Owner	Primary	2	less5		2	NA	small	1to5
26	Mae On	18.78704	99.26428	Control	42	Male	Owner	Primary	3	less5		1	NA	small	1to5
10	San Sai	18.96745	98.98286	Case	42	Male	Owner	Secondary	7	more5		4	1	small	1to5
10	San Kamphaeng	18.71897	99.20720	Control	42	Female	Owner	Secondary	5	more5		2	NA	small	1to5
12	Mae On	18.74767	99.25663	Control	44	Female	Owner	Primary	2	less5		5	NA	small	1to5
36	Mae On	18.79221	99.24513	Control	44	Male	Owner	Primary	4	less5		4	NA	small	1to5
13	Mae On	18.77003	99.25782	Case	45	Male	Owner	Secondary	20	more5		6	1	large	more5
16	Mae On	18.78619	99.26555	Control	45	Male	Owner	Secondary	4	less5		8	NA	large	more5
31	Mae On	18.75399	99.25951	Control	45	Male	Owner	Primary	3	less5		3	NA	small	1to5
22	Mae On	18.78388	99.26102	Control	45	Male	Owner	Diploma	1	less5		2	NA	small	1to5
18	San Kamphaeng	18.73451	99.23511	Control	45	Male	Owner	Secondary	10	more5		2	NA	small	1to5
36	Mae On	18.78503	99.24867	Case	46	Male	Owner	Secondary	1	less5		2	2	small	1to5
34	Mae On	18.76665	99.25637	Case	46	Female	Owner	Primary	5	more5		5	3	small	1to5

Showing 1 to 26 of 159 entries

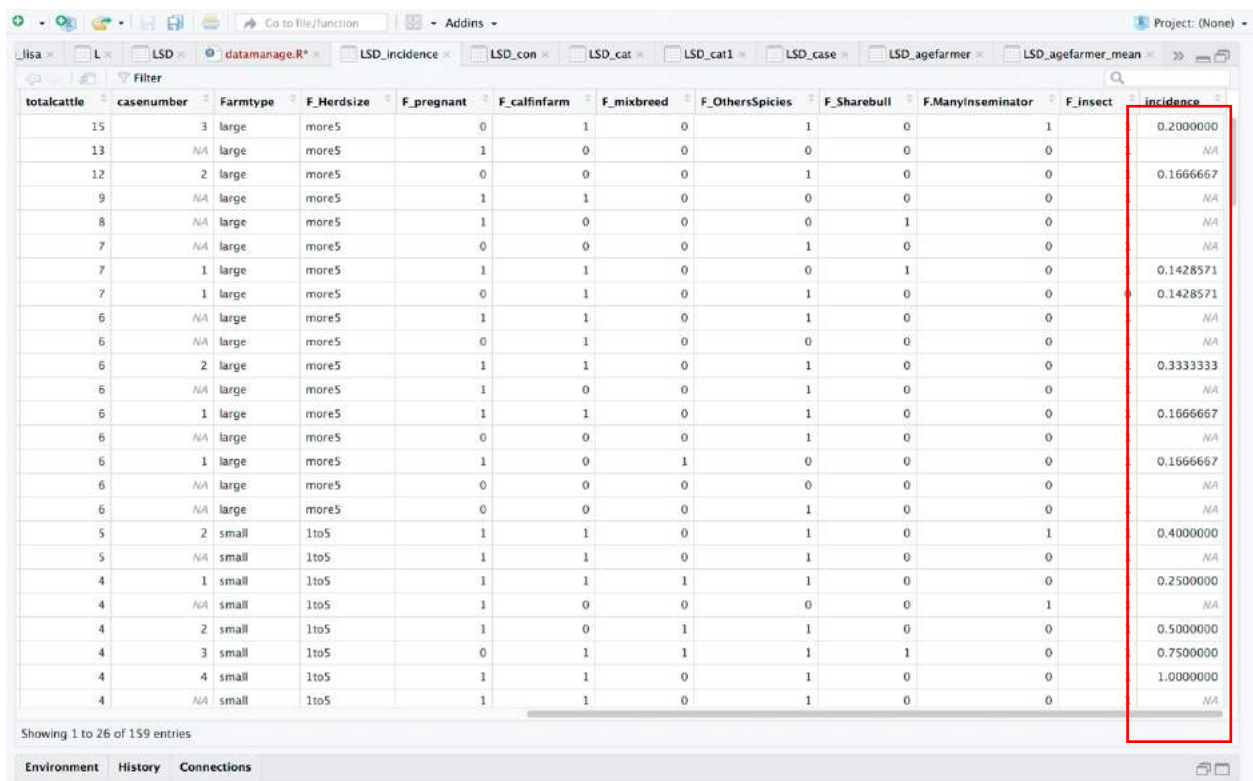
Environment History Connections

กรณีที่ต้องการสร้างคอลัมน์ใหม่ โดยกำหนดให้คอลัมน์ใหม่ชื่อ incidence เกิดจากจำนวนของ casenumber หารด้วย จำนวนของ totalcattle ในออบเจ็คชื่อ LSD_incidence

```
LSD_incidence <- LSD %>%
```

```
  mutate(incidence = casenumber/totalcattle)
```

```
LSD_incidence
```



The screenshot shows the RStudio interface with the 'LSD_incidence' data frame loaded. The table has 12 columns: totalcattle, casenumber, Farmtype, F_Herdsize, F_pregnant, F_calfinfirm, F_mixbreed, F_OthersSpices, F_Sharebull, F.Manyinseminator, F_insect, and incidence. The 'incidence' column is highlighted with a red box. The table shows 159 entries, with the first 26 entries visible in the screenshot.

totalcattle	casenumber	Farmtype	F_Herdsize	F_pregnant	F_calfinfirm	F_mixbreed	F_OthersSpices	F_Sharebull	F.Manyinseminator	F_insect	incidence
15	3	large	more5	0	1	0	1	0	1	1	0.2000000
13	N/A	large	more5	1	0	0	0	0	0	0	N/A
12	2	large	more5	0	0	0	1	0	0	0	0.1666667
9	N/A	large	more5	1	1	0	0	0	0	0	N/A
8	N/A	large	more5	1	0	0	0	1	0	0	N/A
7	N/A	large	more5	0	0	0	1	0	0	0	N/A
7	1	large	more5	1	1	0	0	1	0	0	0.1428571
7	1	large	more5	0	1	0	1	0	0	0	0.1428571
6	N/A	large	more5	1	1	0	1	0	0	0	N/A
6	N/A	large	more5	0	1	0	0	0	0	0	N/A
6	2	large	more5	1	1	0	1	0	0	0	0.3333333
6	N/A	large	more5	1	0	0	1	0	0	0	N/A
6	1	large	more5	1	1	0	1	0	0	0	0.1666667
6	N/A	large	more5	0	0	0	1	0	0	0	N/A
6	1	large	more5	1	0	1	0	0	0	0	0.1666667
6	N/A	large	more5	0	0	0	0	0	0	0	N/A
6	N/A	large	more5	0	0	0	1	0	0	0	N/A
5	2	small	1to5	1	1	0	1	0	1	1	0.4000000
5	N/A	small	1to5	1	1	0	1	0	0	0	N/A
4	1	small	1to5	1	1	1	1	0	0	0	0.2500000
4	N/A	small	1to5	1	0	0	0	0	0	1	N/A
4	2	small	1to5	1	0	1	1	0	0	0	0.5000000
4	3	small	1to5	0	1	1	1	1	1	0	0.7500000
4	4	small	1to5	1	1	0	1	0	0	0	1.0000000
4	N/A	small	1to5	1	1	0	1	0	0	0	N/A

```
# กรณีที่ต้องการสรุปค่าการวิเคราะห์ทางสถิติอย่างง่าย
LSD %>%
  summarise(mean_AgeFarmer= mean(AgeFarmer),
  sd_AgeFarmer = sd(AgeFarmer),
  n = n())
# ผล
  mean_AgeFarmer sd_AgeFarmer
1      55.90566      9.516669
  n
1 159
```

ในกรณีที่ต้องการตรวจสอบการซ้ำซ้อนของข้อมูล สามารถใช้ชุดคำสั่ง tidyverse โดยใช้คำสั่ง duplicated() โดยคำสั่งนี้จะแสดงข้อมูลแถวที่มีข้อมูลซ้ำกัน ซึ่งเมื่อพบข้อมูลซ้ำให้ลบข้อมูลนั้นออกด้วยคำสั่ง

```
duplicated (LSD)
# จากนั้นทำการลบข้อมูลที่ซ้ำออกด้วยคำสั่ง
LSD [!duplicated (LSD), ] -> LSD_new
# หรือ
unique (LSD) -> LSD_new
```

ในกรณีที่ข้อมูลมีค่า Missing value อยู่ แล้วต้องการลบข้อมูลแถวนั้นออกสามารถใช้คำสั่ง

```
LSD_new <- na.omit (LSD)
```

นอกจากนี้การเปลี่ยนลักษณะการเรียงตัวของข้อมูลจาก wide form เป็น long form ด้วยคำสั่ง gather() ซึ่งอยู่ในชุดคำสั่ง tidyร สามารถทำได้ดังนี้

```
# นำเข้าข้อมูลชุดใหม่ที่มีชื่ออบเจ็กต์ว่า olddata_wide
rm(list=ls())

olddata_wide<- read.table(header = T, text="
  id species control groupA groupB
1  dog    7.9  12.3  10.7
2  cat    6.3  10.6  11.1
3  cat    9.5  13.1  13.8
4  dog   11.5  13.4  12.9
")

attach(olddata_wide)

olddata_wide

# เปลี่ยนลักษณะการเรียงตัวของข้อมูลจาก wide form เป็น long form
library(tidyร)

data_long <- gather(olddata_wide, group, RBC, control:groupB, factor_key=TRUE)

data_long
```

```

> olddata_wide
  id species control groupA groupB
1  1    dog    7.9   12.3   10.7
2  2    cat    6.3   10.6   11.1
3  3    cat    9.5   13.1   13.8
4  4    dog   11.5   13.4   12.9
> library(tidyr)
> data_long <- gather(olddata_wide, group, RBC, control:groupB, factor_key=TRUE)
> data_long
  id species  group RBC
1  1    dog control  7.9
2  2    cat control  6.3
3  3    cat control  9.5
4  4    dog control 11.5
5  1    dog groupA 12.3
6  2    cat groupA 10.6
7  3    cat groupA 13.1
8  4    dog groupA 13.4
9  1    dog groupB 10.7
10 2    cat groupB 11.1
11 3    cat groupB 13.8
12 4    dog groupB 12.9
> |

```

ในกรณีที่ต้องการเปลี่ยนชื่อตัวแปรหรือชื่อคอลัมน์ สามารถทำได้โดยใช้คำสั่ง `names()`

```

# เปลี่ยนตัวแปรคอลัมน์ที่ 1 เป็น species
names(data_long)[2] <- "animal"
data_long

```

```

> names(data_long)[2] <- "animal"
> data_long
  id animal  group RBC
1  1    dog control  7.9
2  2    cat control  6.3
3  3    cat control  9.5
4  4    dog control 11.5
5  1    dog groupA 12.3
6  2    cat groupA 10.6
7  3    cat groupA 13.1
8  4    dog groupA 13.4
9  1    dog groupB 10.7
10 2    cat groupB 11.1
11 3    cat groupB 13.8
12 4    dog groupB 12.9
> |

```

การจัดรูปแบบของข้อมูลเพื่อทำการวิเคราะห์ในโปรแกรม R เราสามารถตรวจสอบ class หรือ type ของข้อมูลเราได้ด้วยคำสั่ง `is.numeric()`, `is.logical()`, `is.character()` ในโปรแกรม R ยังมีอีกหลายคำสั่งที่ใช้ในการตรวจสอบชนิดของข้อมูล โดยพิมพ์ `is.` ลงไปใน console แล้วโปรแกรม R จะแนะนำคำสั่งทั้งหมดที่ขึ้นต้นด้วย `is.` ขึ้นมาให้เราเลือกใช้งาน

```
# กรณีที่ต้องการตรวจสอบชนิดของข้อมูล
is.numeric(data_long$RBC)

# ผล เป็นจริง คือ ตัวแปร RBC เป็นข้อมูลแบบตัวเลข
[1] TRUE
```

นอกจากนี้เราสามารถเปลี่ยน class หรือ type ด้วยคำสั่ง `as.numeric()`, `as.logical()`, `as.character()` ฯลฯ เช่น ในการวิเคราะห์หาปัจจัยเสี่ยงของโรคทุกครั้ง จะต้องทำการกำหนดให้แต่ละปัจจัยเป็นชนิด factor ตามตัวอย่างคำสั่งต่อไปนี้

```
# กำหนดตัวแปรแบบแบ่งกลุ่มให้เป็น factor
animals <- c("dog", "cat", "cat", "cat", "dog")
as.factor(animals)

# ผล
[1] dog cat cat cat dog
Levels: cat dog

# กำหนดตัวแปรแบบแบ่งกลุ่มที่มีลำดับให้เป็น factor
farmer_income <- c("high", "medium", "low", "low", "medium")
factor(farmer_income, levels = c("low", "medium", "high"), ordered = TRUE)

# ผล
[1] high medium low low medium
Levels: low < medium < high
```

ส่วนการตั้งค่าของรูปแบบวันที่ (Date values) ในการทำงานในโปรแกรม R ควรจะต้องกำหนดรูปแบบของวันที่เป็นให้เหมาะสมในแต่ละชุดคำสั่ง ซึ่งในโปรแกรม R จะใช้รูปแบบวันที่ที่ใช้สัญลักษณ์ที่เฉพาะ (ตารางที่ 5.1) โดยการตั้งค่ารูปแบบวันที่สามารถทำได้โดยใช้คำสั่ง `as.Date()` ดังตัวอย่างต่อไปนี้

```
# จัดรูปแบบวันที่
dates <- c("20nov1980", "20nov1991", "20nov1993", "10sep1993")
dt <- as.Date(dates, "%d%b%Y")
dt
# ผล
[1] "1980-11-20" "1991-11-20" "1993-11-20" "1993-09-10"
```

ตารางที่ 5.1 สัญลักษณ์ที่ใช้ในการจัดรูปแบบวันที่ที่ใช้ในโปรแกรม R

สัญลักษณ์	ความหมาย	ตัวอย่าง
%d	เลขวันที่	0-31
%a	ตัวย่อวัน	Mon, Tue,....
%A	วัน	Monday, Tuesday,....
%m	เลขเดือน	00-12
%b	ตัวย่อเดือน	Jan, Feb,.....
%B	เดือน	January, February,.....
%y	ตัวย่อปี ค.ศ.	21
%Y	ปี	2021

บทที่ 6

การวิเคราะห์ข้อมูลทางสถิติพื้นฐานสำหรับงานระบาดวิทยาทางสัตวแพทย์

โปรแกรม R สามารถนำมาใช้ในการวิเคราะห์ข้อมูลทางสถิติได้หลากหลายรูปแบบ ขึ้นอยู่กับวัตถุประสงค์และลักษณะข้อมูลที่มี โดยในบทนี้จะกล่าวถึงการวิเคราะห์ข้อมูลทางสถิติพื้นฐานสำหรับงานระบาดวิทยาทางสัตวแพทย์เกี่ยวกับระบาดวิทยาเชิงพรรณนา (Descriptive statistics) เกี่ยวกับเจ้าของสัตว์/ตัวสัตว์ (Person/Animal) เวลา (Time) และสถานที่ (Place) และระบาดวิทยาเชิงอนุมาน (Inferential statistics) ด้วยชุดคำสั่งและคำสั่งอย่างง่าย โดยจะมีการแสดงผลของการวิเคราะห์ในรูปแบบต่างๆ รวมด้วย

6.1 ระบาดวิทยาเชิงพรรณนา

สถิติเชิงพรรณนา เป็นสถิติที่ใช้อธิบาย บรรยาย หรือสรุป ลักษณะของข้อมูลกลุ่มตัวอย่างที่ทำการศึกษา ซึ่งไม่สามารถนำไปอ้างอิงลักษณะประชากรอื่นๆ ได้ สามารถอธิบายได้ด้วยความถี่ (Frequency) ร้อยละ (Percentage) อัตรา(Rate) อัตราส่วน (Ratio) สัดส่วน (Proportion) ค่าเฉลี่ย (Mean) ส่วนเบี่ยงเบนมาตรฐาน (Standard deviation) ค่ามัธยฐาน (Median) ค่าพิสัย (Range) และค่าฐานนิยม (Mode) เป็นต้น ซึ่งข้อมูลที่นำมาใช้สำหรับการวิเคราะห์เชิงพรรณนานั้นจะต้องมีการระบุชนิดของข้อมูลหรือตัวแปรว่าเป็นชนิดใด ดังตัวอย่างแสดงในตารางที่ 6.1 โดยตัวอย่างในการวิเคราะห์จะใช้ข้อมูลที่ชื่อ LSDdata.csv (ภาคผนวก ก) จากนั้นเมื่อนำเข้าข้อมูลมาในโปรแกรม R จะถูกตั้งชื่อใหม่เป็น LSD ดังนั้นในตัวอย่างต่อไปนี้จะใช้ LSD เป็นชื่อออบเจกต์ในการสั่งคำสั่งต่างๆ ให้ทำงานต่อไป

ตารางที่ 6.1 ลักษณะของข้อมูลและวิธีการใช้สถิติเชิงพรรณนา

ลักษณะของข้อมูล	วิธีการใช้สถิติเชิงพรรณนา	ตัวอย่างข้อมูล
ข้อมูลแบบแบ่งกลุ่ม (Categorical data): ข้อมูล ที่แสดงลักษณะที่แตกต่างกัน	ค่าความถี่และร้อยละ	ชนิดสัตว์ (สุนัข/แมว/โค) พันธุ์สัตว์ (พื้นเมือง/ต่างประเทศ) เพศ (ผู้/เมีย) สถานะภาพการเกิดโรค (เป็น/ไม่เป็นโรค) ชนิดแหล่งน้ำ (บาดาล/ประปา/สาธารณะ) ชนิดอาหาร (หยาบ/ขึ้น) การเคลื่อนย้าย (มี/ไม่มี) ฯลฯ
ข้อมูลแบบต่อเนื่อง (Continuous data): ข้อมูลที่ เป็นเลขที่มีค่าได้ทุกค่าในช่วง ที่กำหนดและมีความหมาย	กรณีข้อมูลมีการแจกแจงแบบปกติ: ค่าเฉลี่ย (ส่วนเบี่ยงเบนมาตรฐาน) กรณีข้อมูลมีการแจกแจงแบบไม่ปกติ: ค่ามัธยฐาน (ค่าสูงสุด-ต่ำสุด) หรือ ค่ามัธยฐาน (ค่าควอไทล์ที่ 1-ค่าควอไทล์ที่ 3)	อายุสัตว์ น้ำหนักสัตว์ ความสูง ระยะเวลา การเลี้ยง อัตราการเต้นของหัวใจ อัตรา การหายใจ ระดับน้ำตาลในเลือด ค่าเลือด ฯลฯ
ข้อมูลแบบไม่ต่อเนื่อง (Discrete data): ข้อมูลที่เป็น เลขจำนวนเต็มที่มีความหมาย	ร้อยละและอัตราส่วน	จำนวนสัตว์ จำนวนการตั้งท้อง จำนวน สัตว์ป่วย จำนวนสัตว์ที่แสดงอาการ

6.1.1 สถิติเชิงพรรณนาเกี่ยวกับเจ้าของสัตว์/ตัวสัตว์

ในการอธิบายเกี่ยวกับเจ้าของสัตว์ (เช่น เพศ อายุ อาชีพ ตำแหน่ง ระดับการศึกษา ระยะเวลาการเลี้ยงสัตว์ ฯลฯ) หรือตัวสัตว์ (เช่น เพศ พันธุ์ อายุ ชนิดสัตว์ จำนวนสัตว์ ประวัติการได้รับวัคซีน ประวัติการนำสัตว์ใหม่เข้ามาในฝูง ฯลฯ) นั้นจะมีประเภทของข้อมูลเป็นทั้งแบบแบ่งกลุ่ม แบบต่อเนื่อง และแบบไม่ต่อเนื่อง (ดังแสดงในตารางที่ 6.1) ซึ่งสามารถทำการวิเคราะห์ข้อมูลได้ดังนี้

6.1.1.1 การวิเคราะห์ข้อมูลแบบแบ่งกลุ่ม

การวิเคราะห์ข้อมูลเชิงพรรณนาของตัวแปรแบบแบ่งกลุ่ม สามารถใช้ค่าความถี่หรือร้อยละ โดยสามารถวิเคราะห์ค่าสิ่งที่ละตัวแปร โดยใช้ชุดคำสั่ง janitor ซึ่งในที่นี้จะใช้ตัวแปรที่มีชื่อว่า Farmtype ซึ่งแบ่งออกเป็น 2 กลุ่มคือ ฟาร์มขนาดเล็ก (small farm) และฟาร์มขนาดใหญ่ (large farm) ดังตัวอย่างต่อไปนี้

```
library(janitor)
# กรณีที่ต้องการแสดงความถี่เพียง 1 ตัวแปร คือ Farmtype
LSD %>% tabyl(Farmtype)
# ผล
Farmtype  n          percent
large     43          0.2704403
small    116          0.7295597
# กรณีที่ต้องการแสดงความถี่และร้อยละของตัวแปร 1 ตัวแปร คือ Farmtype
LSD %>%
  tabyl(Farmtype) %>%
  adorn_pct_formatting()
# ผล
Farmtype  n          percent
large     43          27.0%
small    116          73.0%
```

หรือถ้าต้องการทำเป็นตารางไขว้ (Cross-tabulation table) ที่แสดงความถี่และร้อยละระหว่าง 2 ตัวแปร โดยในตัวอย่างนี้จะแสดงความถี่และร้อยละของตัวแปรที่มีชื่อว่า Farmtype และ Status สามารถทำได้โดย

```
# กรณีที่ต้องการแสดงความถี่ระหว่าง 2 ตัวแปร คือ Farmtype และ Status
LSD %>% tabyl(Farmtype, Status)
# ผล
Farmtype  Case  Control
large     22    21
small     29    87
```

```
# กรณีที่ต้องการแสดงร้อยละระหว่าง 2 ตัวแปร คือ Farmtype และ Status
```

```
LSD %>%
```

```
tabyl(Farmtype, Status) %>%
```

```
adorn_totals(where = "row") %>%
```

```
adorn_percentages(denominator = "row") %>%
```

```
adorn_pct_formatting(digits = 2)
```

```
# ผล
```

Farmtype	Case	Control
large	51.16%	48.84%
small	25.00%	75.00%
Total	32.08%	67.92%

```
# กรณีที่ต้องการแสดงความถี่และร้อยละระหว่าง 2 ตัวแปร คือ Farmtype และ Status
```

```
LSD %>%
```

```
tabyl(Farmtype, Status) %>%
```

```
adorn_totals(where = "row") %>%
```

```
adorn_percentages(denominator = "col") %>%
```

```
adorn_pct_formatting() %>%
```

```
adorn_ns(position = "front") %>%
```

```
adorn_title(
```

```
row_name = "Farm type",
```

```
col_name = "LSD status")
```

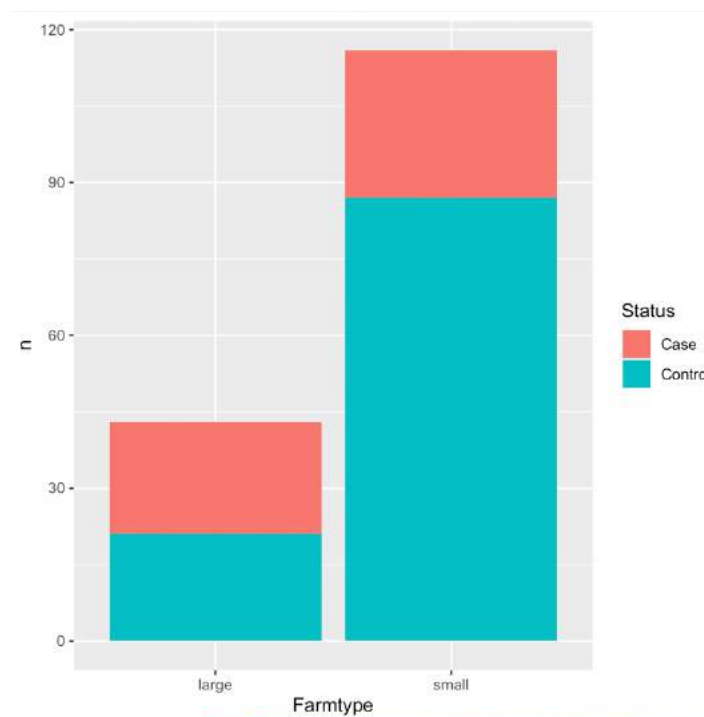
```
# ผล
```

LSD status

Farm type	Case	Control
large	22 (43.1%)	21 (19.4%)
small	29 (56.9%)	87 (80.6%)
Total	51 (100.0%)	108 (100.0%)

ในกรณีที่ต้องการสร้างกราฟ Bar plot ของตัวแปรแบบต่อเนื่อง 2 ตัวแปร สามารถทำได้โดยใช้คำสั่ง

```
# ตัวแปรแบบต่อเนื่อง คือ AgeFarmer และตัวแปรแบบแบ่งกลุ่ม คือ Status
library(ggplot2)
library(dplyr)
LSD %>%
count(Status, Farmtype) %>%
ggplot()+
geom_col(
mapping = aes(
x = Farmtype,
fill = Status,
y = n))
```



6.1.1.2 การวิเคราะห์ข้อมูลแบบต่อเนื่อง

ในที่นี้จะทำการวิเคราะห์ข้อมูลแบบต่อเนื่องสามารถทำได้หลายวิธี โดยใช้ตัวอย่างคำสั่งดังนี้

```
# กรณีที่ต้องการวิเคราะห์ทีละ 1 ตัวแปร คือ AgeFarmer
summary(LSD$AgeFarmer)

# ผล

  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 32.00  50.00   56.00   55.91  63.50   87.00

# กรณีที่ต้องการวิเคราะห์หลายตัวแปร คือ AgeFarmer, HusbandryTime, totalcattle
library(rstatix)
LSD %>%
  get_summary_stats(
    AgeFarmer, HusbandryTime, totalcattle,
    type = "common")

# ผล

A tibble: 3 x 10
  variable      n  min  max median  iqr mean  sd  se  ci
  <chr>      <dbl> <dbl> <dbl>  <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>
1 AgeFarmer   159   32   87    56 13.5 55.9  9.52 0.755 1.49
2 HusbandryTime 159    1   30    8  10  9.97  7.33 0.582 1.15
3 totalcattle  159    1   24    4   3  4.62  3.10 0.246 0.486
```

กรณีที่ต้องการวิเคราะห์ตัวแปรแบบต่อเนื่องกับตัวแปรแบบแบ่งกลุ่ม คือ AgeFarmer กับ Status

tapply(LSD\$AgeFarmer, LSD\$Status, summary)

ผล

\$Case

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
37.0	50.0	55.0	55.8	64.0	87.0

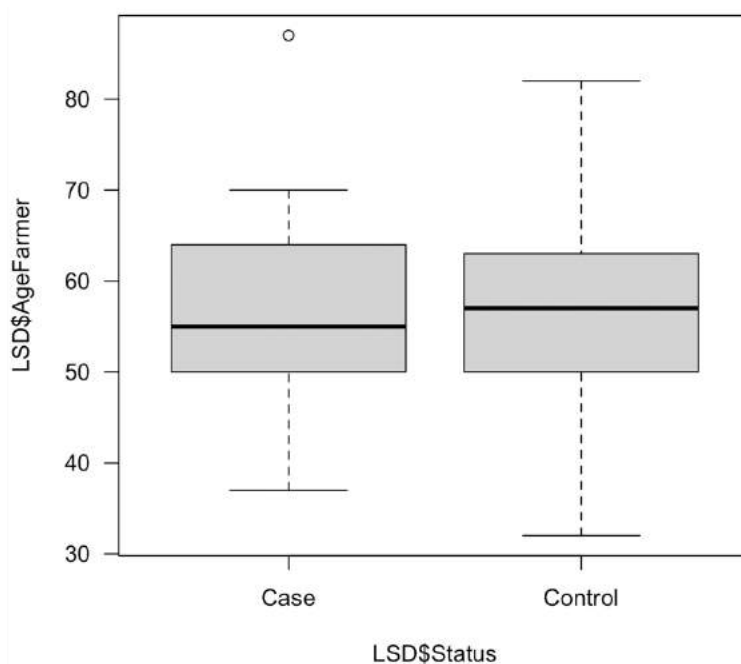
\$Control

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
32.00	50.00	57.00	55.95	63.00	82.00

ในกรณีที่ต้องการสร้างกราฟ Box plot ของตัวแปรแบบต่อเนื่องกับตัวแปรแบบแบ่งกลุ่ม สามารถทำได้โดยใช้คำสั่ง

ตัวแปรแบบต่อเนื่อง คือ AgeFarmer และตัวแปรแบบแบ่งกลุ่ม คือ Status

boxplot(LSD\$AgeFarmer~LSD\$Status, las=1)



6.1.2 สถิติเชิงพรรณนาเกี่ยวกับเวลา

ในการสอบสวนโรคระบาดในสัตว์ การแสดงความถี่เชิงเวลาสามารถทำได้โดยใช้กราฟระบาด (Epidemic curve) มาใช้ในการอธิบายการกระจายตัวของการเกิดโรคในกรณีนั้นๆ โดยในเบื้องต้นให้ผู้จัดทำเตรียมข้อมูลในรูปแบบตาราง โดยรูปแบบของวันที่ให้กำหนดให้เป็น ปี-เดือน-วัน และทำการเรียงลำดับข้อมูลวันที่ให้ถูกต้อง โดยเติมวันที่ให้ครบถ้วน (ภาคผนวก ข) ดังตัวอย่างต่อไปนี้

date	count	Status
2021-08-01	0	NA
2021-08-02	0	NA
2021-08-03	0	NA
2021-08-04	0	NA
2021-08-05	0	NA
2021-08-06	1	Confirmed
2021-08-07	2	Suspected
2021-08-07	3	Confirmed
.....		
.....		

หมายเหตุ: date หมายถึง วันที่เกิดโรคหรือรายงานโรค, count หมายถึง จำนวนฟาร์มที่พบโรค, Status หมายถึง สถานะของฟาร์ม, NA หมายถึง ไม่ทราบสถานะของฟาร์ม, Confirmed หมายถึง ฟาร์มยืนยันโรค, Suspected หมายถึง ฟาร์มสงสัยโรค

เนื่องจากในตัวอย่างนี้มีข้อมูลปริมาณน้อย จึงสามารถนำเข้าข้อมูลมาสู่โปรแกรม R ได้โดยตรง ซึ่งในที่นี้ได้ทำการกำหนดชื่อออบเจ็ค คือ E โดยใช้คำสั่ง

```
rm(list=ls())

E <- read.table(header = T, text="

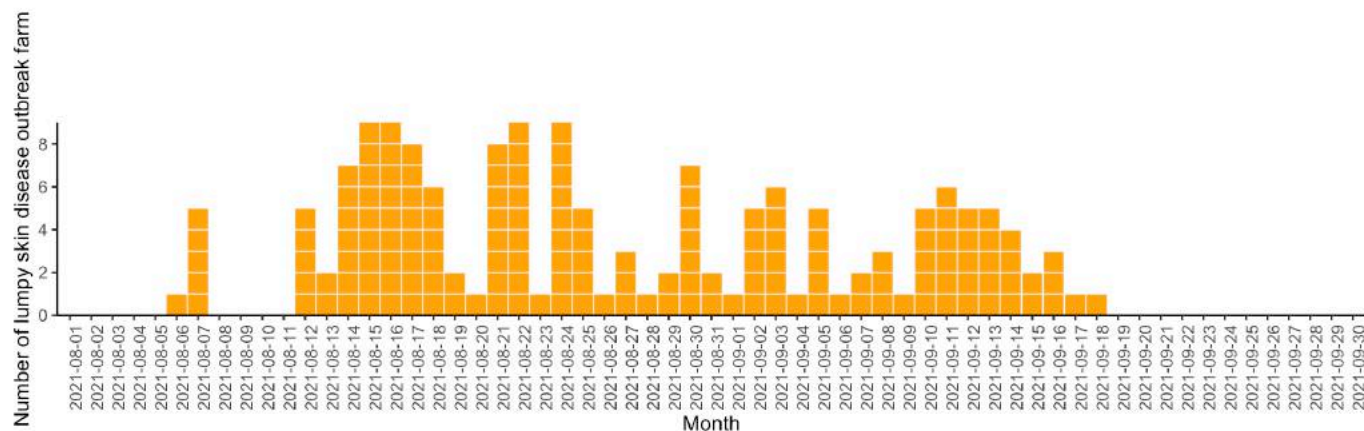
    date                                count Status
    2021-08-01                          0    NA
    2021-08-02                          0    NA
    2021-08-03                          0    NA
    2021-08-04                          0    NA
    2021-08-05                          0    NA
    2021-08-06                          1  Confirmed
    .....                               ....
    .....                               ....
    2021-09-29                          0    NA
    2021-09-30                          0    NA

    ")

attach(E)
```

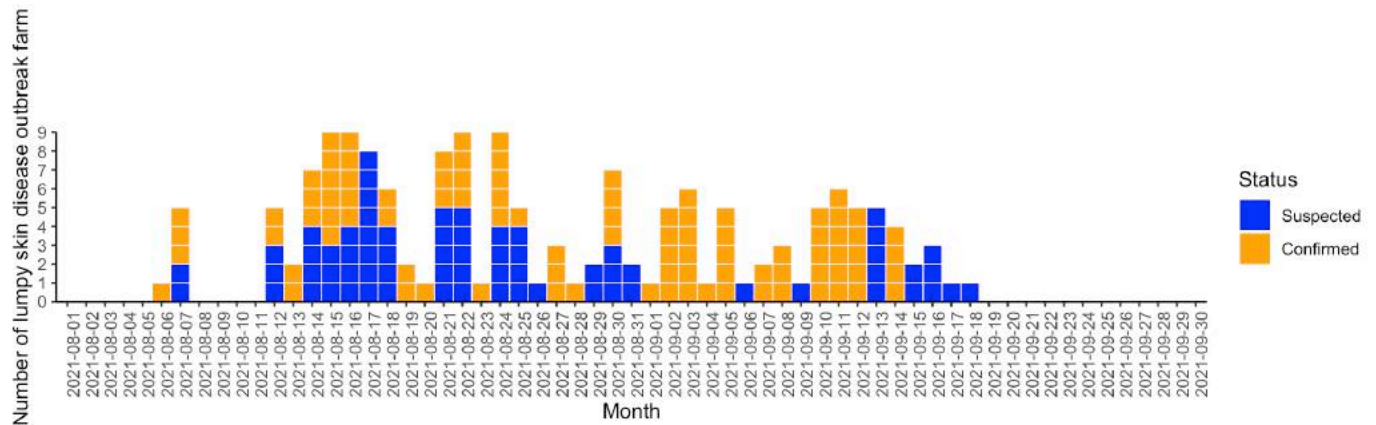
จากนั้นเรียกใช้งานชุดคำสั่งที่ใช้ในการทำกราฟระบาด คือ EpiCurve โดยทำการติดตั้งและเรียกใช้คำสั่งดังต่อไปนี้

```
# กรณีที่ต้องการกราฟระบาดเพียง 1 สี
install.packages("EpiCurve")
library(EpiCurve)
EpiCurve(E, date= "date",
  freq= "count",
  period="date",
  colors= "orange",
  xlabel = "Month",
  ylabel = "Number of lumpy skin disease outbreak farm"
)
```

หรือถ้าต้องการแยกสีตามกลุ่มของข้อมูล ซึ่งในตัวอย่างข้อมูลนี้ใช้ตัวแปร Status ซึ่งในตัวแปรนี้จะแยกเป็น Confirmed และ Suspected farms โดยสามารถใช้คำสั่งดังต่อไปนี้

```
EpiCurve(E, date= "date",
  freq= "count",
  period="date",
  cutvar="Status",
  colors= c("orange", "blue"),
  xlabel = "Month",
  ylabel = "Number of lumpy skin disease outbreak farm"
)
# กำหนดชนิดของตัวแปรแบบแบ่งกลุ่มที่ต้องการแยกสีที่ cutvar= "..."
```



6.1.3 สถิติเชิงพรรณนาเกี่ยวกับพื้นที่

ในการอธิบายลักษณะของการกระจายตัวเชิงพื้นที่ของการระบาดของโรคในสัตว์ สามารถแสดงจุดเกิดโรคโดยใช้พิกัดฟาร์มแบบละติจูดและลองจิจูด นำมาแสดงจุดที่เกิดโรคในโปรแกรม R ได้เช่นเดียวกัน ผู้ใช้สามารถสร้างแผนที่ในโปรแกรม R โดยเตรียมข้อมูล เช่น ละติจูด/ลองจิจูดของฟาร์ม อำเภอหรือตำบลที่เกิดโรค และสถานะของฟาร์ม (เกิดโรค/ไม่เกิดโรค) เป็นต้น โดยจะต้องเตรียมไฟล์ข้อมูลในสกุล .txt (ภาคผนวก ค) ดังตัวอย่างของข้อมูลต่อไปนี้

id	District	Latitude	Longitude	Status
DKSK201	San Kamphaeng	18.78595621	99.1905035	Confirmed case
DKSK202	San Kamphaeng	18.73208333	99.12086111	Confirmed case
DKSK52	San Kamphaeng	18.70780073	99.16714982	Confirmed case
DMFM100	Mae On	18.78231489	99.26403812	Confirmed case
DMFM16	Mae On	18.78159858	99.25923567	Suspected case
DMFM21	Mae On	18.7956532	99.24526001	Suspected case
DMFM24	Mae On	18.78121926	99.25904531	Suspected case
DMFM29	Sansai	18.78371267	99.25978926	Suspected case
.....				

ในการสร้างแผนที่การระบาดของโรคจะใช้ชุดคำสั่งที่สำคัญ คือ sf และ ggplot2 โดยจะต้องทำการติดตั้งและเรียกใช้ชุดคำสั่งดังนี้

```
install.packages("sf")  
install.packages("ggplot2")  
library(sf)  
library(ggplot2)
```

ขอบเขตของแผนที่สามารถทำได้ทั้งในระดับตำบล/อำเภอ/จังหวัด/ประเทศ ซึ่งในที่นี้เป็นตัวอย่างขอบเขตระดับอำเภอที่ได้คัดเลือกพื้นที่ที่ต้องการสร้างแผนที่มาแล้วในโปรแกรม QGIS จากนั้นนำเข้าไฟล์แผนที่ (shape file) ของพื้นที่ที่ต้องการทำแผนที่ซึ่งเป็นไฟล์สกุล .shp โดยในที่นี้ได้ทำการนำเข้าไฟล์แผนที่ชื่อ 3district.shp และกำหนดชื่อออบเจกต์ของไฟล์แผนที่เป็น M โดยใช้คำสั่ง

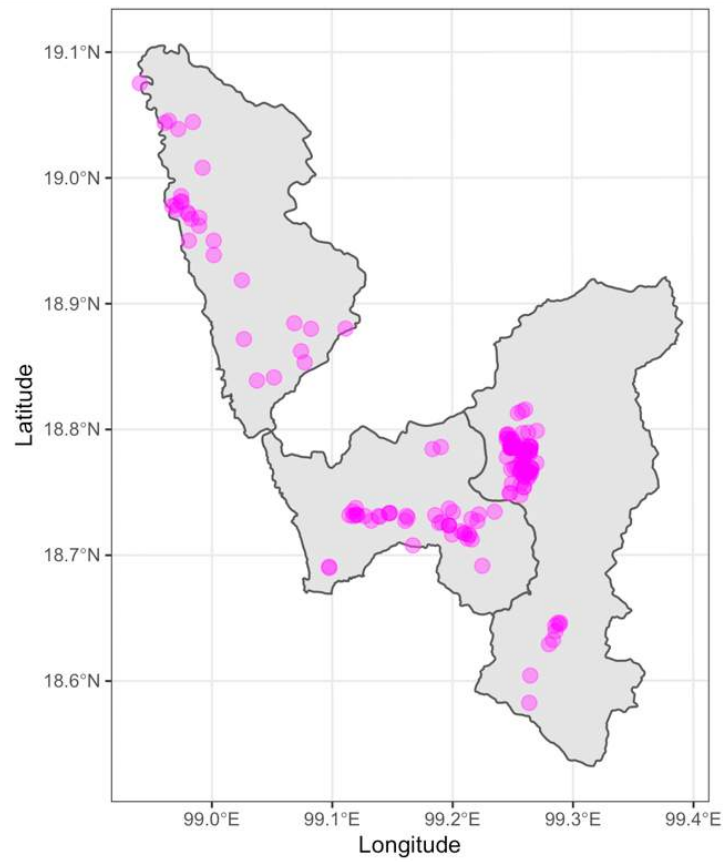
```
M <- st_read("3 district.shp")  
attach(M)
```

จากนั้นใช้คำสั่งเพื่อนำเข้าข้อมูลของการเกิดโรคที่เป็นสกุล .txt ข้างต้น ด้วยคำสั่ง

```
locate <- read.table ("~/Documents/lsd.txt", header=T, sep="\t")  
loc2 <- locate %>% st_as_sf(coords = c("Longitude", "Latitude"), crs = 4326)
```

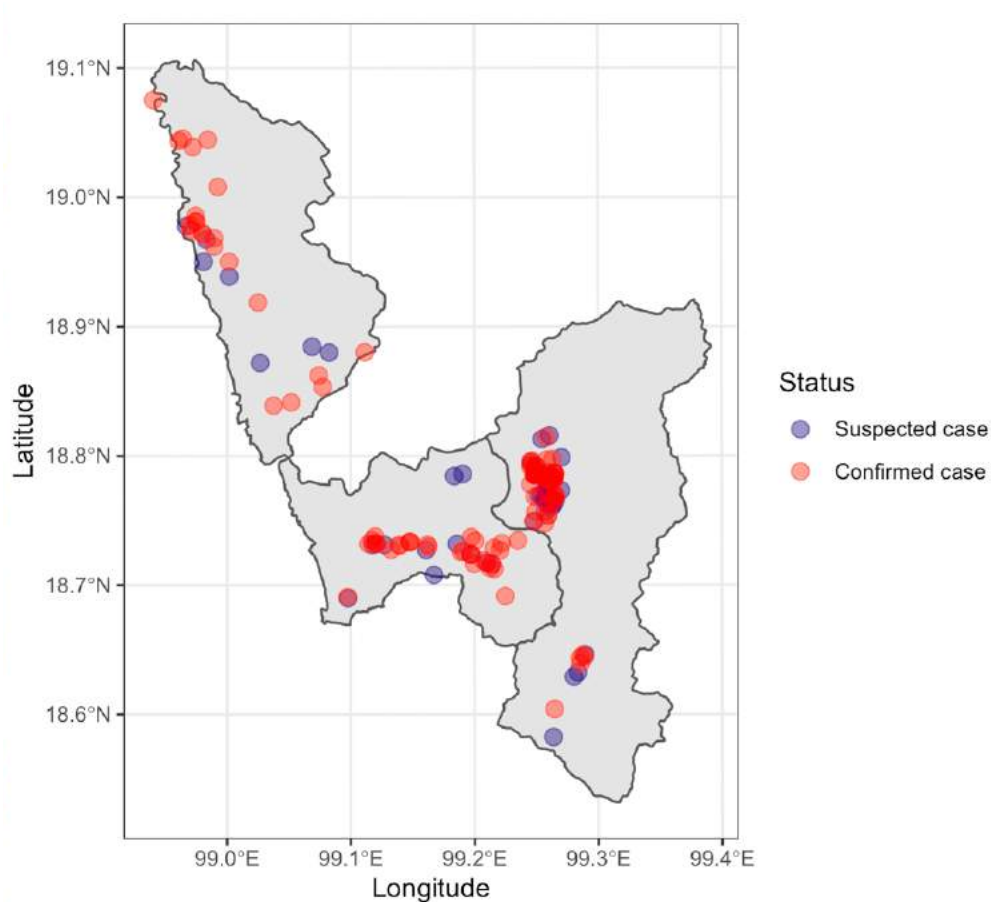
จากนั้นทำการสร้างแผนที่ ด้วยคำสั่ง

```
# กรณีที่ต้องการแสดงจำนวนฟาร์มในพื้นที่เกิดโรคทั้งหมดเป็นสีเดียวกัน  
library(ggplot2)  
ggplot(m)+  
  geom_sf() +  
  geom_sf(data= loc2, color="magenta1", size=3,alpha=0.5) +  
  theme_bw() +  
  labs(  
    x = "Longitude",  
    y = "Latitude"  
  )
```



กรณีที่ต้องการแสดงตำแหน่งของฟาร์มในพื้นที่เกิดโรค โดยแยกตามสถานภาพการเกิดโรค คือ ตัวแปรชื่อ “Status”

```
ggplot(m)+
  geom_sf() +
  geom_sf(data= loc2, aes(color=Status), size=3,alpha=0.5) +
  scale_color_manual(values=c("darkblue","red"),
    labels=c("Suspected case", "Confirmed case"))+
  theme_bw()+
  labs(
    x = "Longitude",
    y = "Latitude"
  )
```



6.2 ระบาดวิทยาเชิงอนุมาน

ในการศึกษาปัจจัยเสี่ยงจากการเกิดโรคระบาดในสัตว์ สามารถทำการทดสอบหาความสัมพันธ์ระหว่างตัวแปรแบบแบ่งกลุ่ม 2 ตัวแปร โดยขึ้นอยู่กับวิธีการออกแบบวิธีการศึกษาที่ใช้ ซึ่งในที่นี้จะอธิบายการวิเคราะห์ 2 วิธี คือ การทดสอบไคสแควร์ (Chi-squared test) และการหาอัตราส่วนออดส์ (Odds ratio)

6.2.1 การทดสอบไคสแควร์

ในการทดสอบหาความสัมพันธ์ระหว่างตัวแปรกลุ่ม 2 ตัวแปร โดยสามารถใช้สถิติทดสอบ Chi-squared หรือ Fisher's exact โดยสถิติทดสอบ Chi-squared จะใช้เมื่อมีขนาดตัวอย่างใหญ่เพียงพอ หรือสามารถพิจารณาจากค่าที่คาดหวัง (Expected value) ไม่ควรต่ำกว่า 5 ซึ่งถ้าหากค่าที่คาดหวังน้อยกว่า 5 หรือน้อยกว่าร้อยละ 20 ของจำนวนเซลล์ทั้งหมดในตารางแบบไขว้ จะต้องทำการวิเคราะห์ด้วยสถิติทดสอบ Fisher's exact สำหรับคำสั่งที่ใช้ในการหาค่าความถี่คาดหวัง คือ `chisq.test()$expected` ส่วนคำสั่ง `chisq.test()` ใช้ในการวิเคราะห์ Chi-squared และ คำสั่ง `fisher.test()` จะใช้ในการวิเคราะห์ Fisher's exact

ในที่นี้เป็นตัวอย่างการวิเคราะห์ว่าขนาดของฟาร์ม (Farmtype) กับการเกิดโรค LSD (Status) ในประชากรที่ทำการศึกษามีความสัมพันธ์กันหรือไม่ โดยสามารถตั้งสมมุติฐานได้ดังนี้

H_0 : ขนาดของฟาร์มกับการเกิดโรค LSD ในประชากรที่ทำการศึกษาไม่มีความสัมพันธ์กัน

H_a : ขนาดของฟาร์มกับการเกิดโรค LSD ในประชากรที่ทำการศึกษามีความสัมพันธ์กัน

โดยใช้คำสั่งดังนี้

```
# หาความสัมพันธ์ระหว่างตัวแปร 2 ตัวแปร คือ Farmtype และ Status ด้วยสถิติทดสอบ Chi-squared
# ตรวจสอบค่าความถี่ของแต่ละตัวแปร
table(LSD$Farmtype, LSD$Status)

# ผล

      Case Control
large    22     21
small    29     87

# ทดสอบสถิติ Chi-squared
chisq.test(LSD$Farmtype, LSD$Status, correct=FALSE)

# ผล

      Pearson's Chi-squared test

data:  LSD$Farmtype and LSD$Status
X-squared = 9.8559, df = 1, p-value = 0.001693
```

จากผลการทดสอบ พบว่าค่าความถี่ที่คาดหวังมากกว่า 5 ทั้ง 4 เซลล์ จึงสามารถทำการวิเคราะห์ด้วยสถิติทดสอบ Chi-squared ได้ค่า p-value = 0.001 จึงทำการปฏิเสธ H_0 สรุปได้ว่า ขนาดของฟาร์มกับการเกิดโรค LSD ในประชากรที่ทำการศึกษามีความสัมพันธ์กันอย่างมีนัยสำคัญทางสถิติ ที่ระดับนัยสำคัญ 0.05

6.2.2 การหาอัตราส่วนออดส์

การหาอัตราส่วนออดส์ (Odds ratio: OR) คำนวณมาจากอัตราส่วนความน่าจะเป็น 2 ค่า คือ ความน่าจะเป็นของการเกิดเหตุการณ์ที่สนใจ เช่น ฟาร์มที่เป็นโรค (Case) ต่อความน่าจะเป็นของการเกิดเหตุการณ์ที่ไม่สนใจ เช่น ฟาร์มที่ไม่เป็นโรค (Control)

โดยในตัวอย่างนี้จะหาความสัมพันธ์ระหว่างการมีโคตั้งท้องในฟาร์ม (F_pregnant: 1=yes, 0=no) กับการเกิดโรคล้มปี่ สกิน LSD (Status: Case, Control) จึงทำการทดสอบความสัมพันธ์ ด้วยคำสั่งดังต่อไปนี้

```
# หาความสัมพันธ์ระหว่างตัวแปร 2 ตัวแปร คือ F_pregnant และ Status ด้วยค่า Odds ratio
```

```
# กำหนดตัวแปรที่จะทำการวิเคราะห์ให้เป็นปัจจัย
```

```
LSD$Status <- as.factor(LSD$Status)
```

```
LSD$Farmltype <- as.factor(LSD$F_pregnant)
```

```
# กำหนดการเรียงลำดับของปัจจัย
```

```
LSD$Status <- relevel (LSD$Status, ref="Case")
```

```
LSD$Farmltype <- relevel (LSD$F_pregnant, ref="0")
```

```
# แสดงค่าความถี่และร้อยละในตารางไขว้
```

```
library(epiDisplay)
```

```
tabpct(LSD$F_pregnant, LSD$Status, percent="row")
```

```
# ผล
```

```
Row percent
```

```
    LSD$Status
```

```
LSD$F_pregnant   Case Control Total
```

```
      0      17      35      52
```

```
      (32.7) (67.3) (100)
```

```
      1      34      73     107
```

```
      (31.8) (68.2) (100)
```

```
# คำนวณค่าอัตราส่วนออดส์
```

```
cc(LSD$F_pregnant , LSD$Status)
```

ผล

LSD\$Status

LSD\$F_pregnant Case Control Total

0 17 35 52

1 34 73 107

Total 51 108 159

OR = 1.04

95% CI = 0.51, 2.12

Chi-squared = 0.01, 1 d.f., P value = 0.908

Fisher's exact test (2-sided) P value = 1

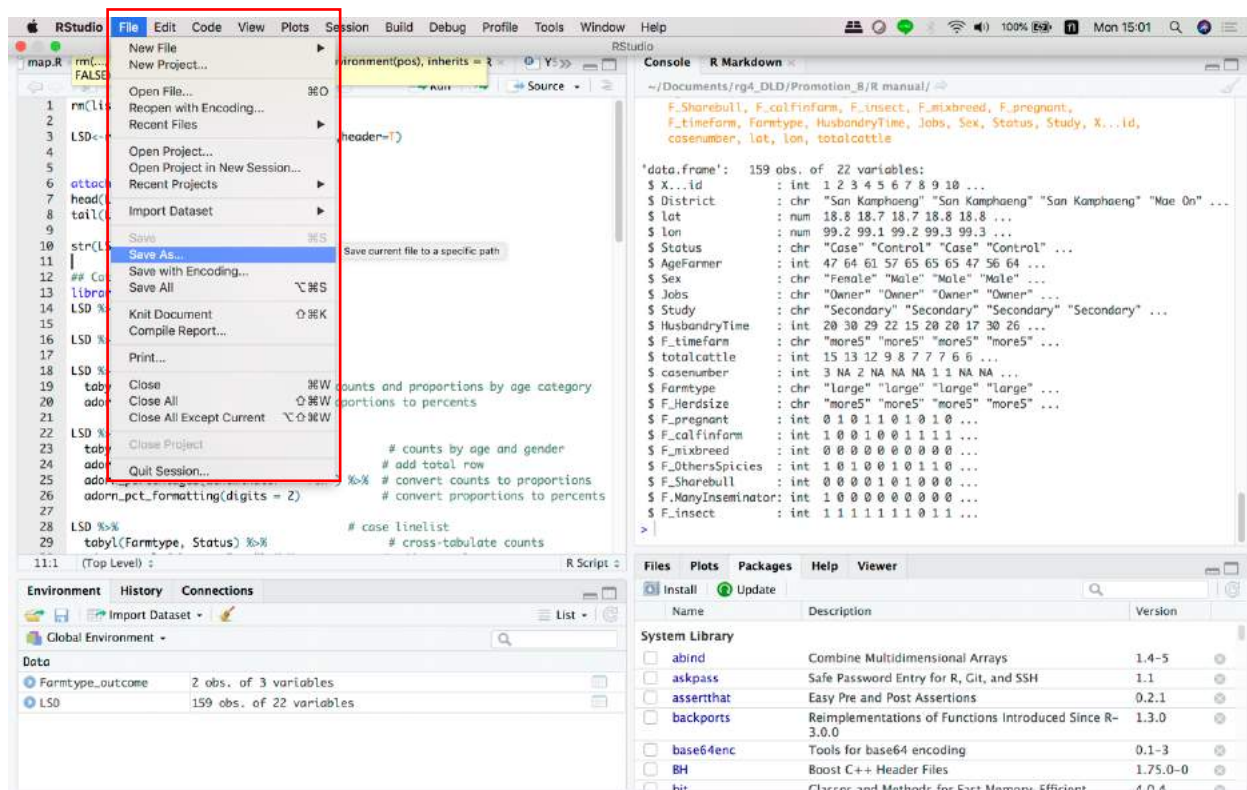
จากผลการคำนวณค่า OR ด้วยคำสั่ง cc() พบว่า ค่า OR มีค่าเท่ากับ 1.04 และค่าช่วงความเชื่อมั่นเท่ากับ 0.51-2.12 จึงสามารถสรุปได้ว่า ฟาร์มที่มีโคตั้งท้องอยู่ในฟาร์มมีโอกาสเกิดโรค LSD เป็น 1.04 เท่าเมื่อเทียบกับฟาร์มที่ไม่มีโคตั้งท้องอยู่ในฟาร์ม ในขณะที่พบว่าฟาร์มที่มีโคตั้งท้องอยู่ในฟาร์มและการเกิดโรค LSD ไม่มีความสัมพันธ์กันอย่างมีนัยสำคัญทางสถิติที่ระดับนัยสำคัญ 0.05

บทที่ 7

การบันทึกข้อมูลและส่งออกผลการวิเคราะห์ข้อมูล

ในการบันทึกข้อมูลและส่งออกผลการวิเคราะห์ข้อมูล ผู้ใช้สามารถส่งออกคำสั่งที่ใช้ ไฟล์ข้อมูล กราฟ หรือรายงาน ได้หลากหลายรูปแบบขึ้นอยู่กับความต้องการของผู้ใช้เป็นหลัก ซึ่งในที่นี้จะแนะนำวิธีการอย่างง่าย และจำเป็นสำหรับข้อมูลและผลการวิเคราะห์ที่ได้จากการวิเคราะห์ข้อมูลจากการสอบสวนโรคและเฝ้าระวังโรคระบาดในสัตว์

เมื่อผู้ใช้ได้ทำการวิเคราะห์ข้อมูลต่างๆ เรียบร้อยแล้ว ให้ทำการบันทึกคำสั่งต่างๆ ที่สร้างขึ้น โดยจะเป็นนามสกุล .R ซึ่งจะสามารถเปิดได้ในโปรแกรม R โดยเลือก File แล้วเลือก Save as จากนั้นไฟล์ดังกล่าวจะถูกเบรบทไปยังโฟลเดอร์ที่ได้กำหนดพื้นที่ทำงานโดยอัตโนมัติ หรือสามารถเลือกโฟลเดอร์ที่ต้องการบันทึกเองก็ได้เช่นเดียวกัน ดังรูปที่ 7.1



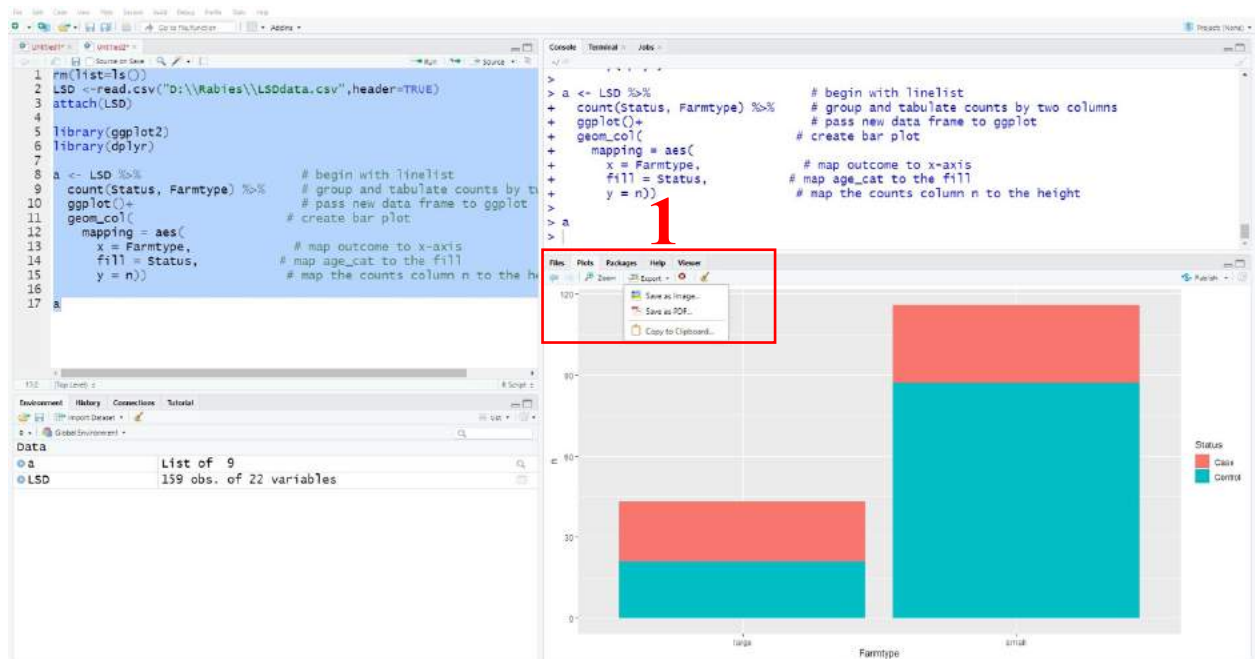
รูปที่ 7.1 การบันทึกคำสั่งที่สร้างขึ้นให้เป็นแบบสกุล .R

ส่วนข้อมูลตารางของออบเจ็คต่างๆ ที่ถูกสร้างขึ้นนั้นสามารถที่จะทำการส่งออกเป็นรูปแบบไฟล์ในสกุลแบบต่างๆ เช่น .txt .csv .xlsx .xls เป็นต้น โดยในที่นี้จะเป็นอย่างการบันทึกตารางที่สร้างขึ้นในรูปแบบสกุล .csv โดยใช้คำสั่ง

```
# กรณีที่ต้องการบันทึกข้อมูลตารางในโปรแกรม R เป็นสกุล .csv ชื่อ LSD_incidence.csv
# การบันทึกข้อมูลตารางของระบบปฏิบัติการ Windows
write.table(LSD_incidence, file = "D:/R/LSD_incidence.csv",
col.names = TRUE, row.names = FALSE, sep = "\t")

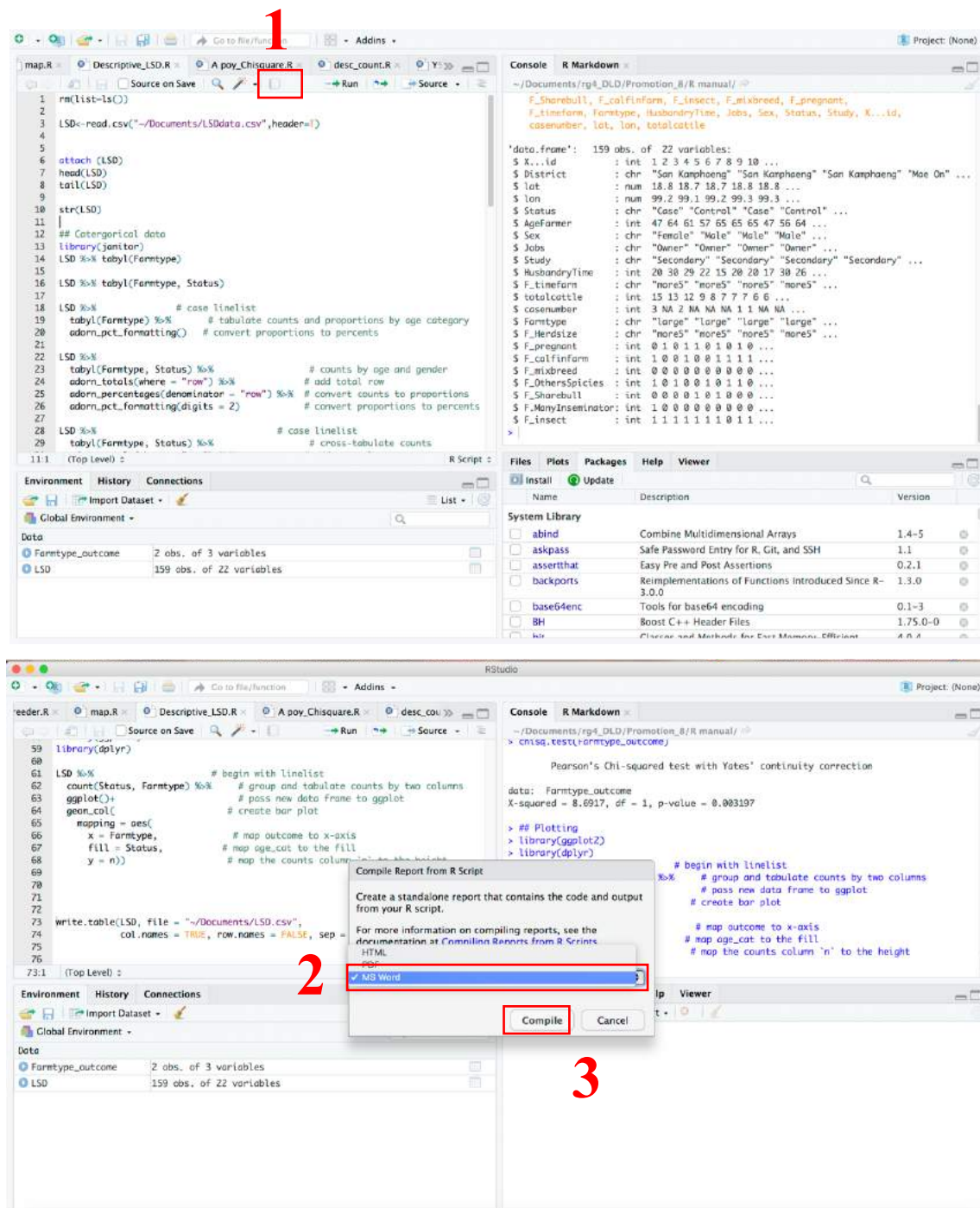
# การบันทึกข้อมูลตารางของระบบปฏิบัติการ Mac OS
write.table(LSD_incidence, file = "~/Documents/R/LSD_incidence.csv",
col.names = TRUE, row.names = FALSE, sep = "\t")
```

ส่วนการบันทึกกราฟชนิดต่างๆ ที่ถูกสร้างขึ้นในโปรแกรม R สามารถทำได้หลายวิธี แต่ในที่นี้จะเป็นการบันทึกกราฟ เป็นสกุล .jpg สามารถทำได้โดย เลือกคำว่า Export ที่หน้าต่าง Plot แล้วเลือกเป็น Save as Image จากนั้นจะมีหน้าต่างปรากฏขึ้นมา (1) ให้ทำการเลือกรูปแบบสกุลของรูปภาพที่ต้องการบันทึกที่ Image format พร้อมเลือกโฟลเดอร์ที่ต้องการจัดเก็บรูปที่ Directory (2) แล้วเลือกขนาดของรูปภาพที่ต้องการ (3) แล้วทำการเลือก Save เพื่อบันทึกรูปภาพที่ต้องการ (4) ดังรูปที่ 7.2



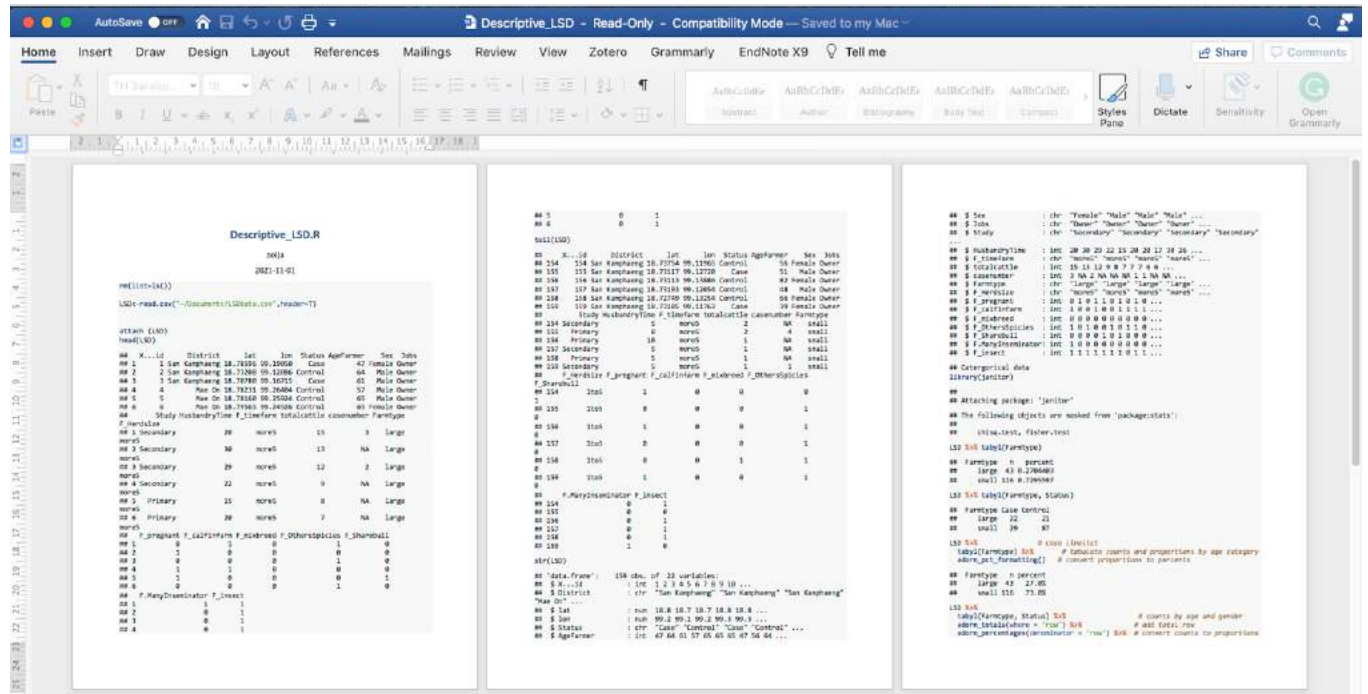
รูปที่ 7.2 การบันทึกกราฟที่ได้จากการวิเคราะห์เป็นรูปภาพ

นอกจากนี้ผู้ใช้งานยังสามารถส่งออกรายงานในรูปแบบเอกสาร Microsoft Word หรือ PDF หรือ HTML ได้เช่นเดียวกัน โดยในรายงานจะประกอบด้วยคำสั่งและผลการวิเคราะห์ โดยใช้ Compile report ซึ่งในตัวอย่างนี้จะทำการบันทึกรายงานในรูปแบบ Microsoft Word (รูปที่ 7.3)



รูปที่ 7.3 การบันทึกรายงานคำสั่งและผลการวิเคราะห์โดยใช้ Compile report

หลังจากที่ทำการบันทึกเรียบร้อยแล้ว เอกสารรายงานในรูปแบบ Microsoft word นั้นจะปรากฏขึ้นมาโดยอัตโนมัติ ตัวอย่างไฟล์ดังแสดงในรูปที่ 7.4



รูปที่ 7.4 ตัวอย่างเอกสารรายงานในรูปแบบ Microsoft word จาก Compile report

บทที่ 8

ข้อแนะนำหรือข้อควรระวังของการใช้โปรแกรม

ในบทนี้จะขอสรุปข้อแนะนำและข้อควรระวังในการใช้โปรแกรม R เบื้องต้น สำหรับการนำมาใช้ในงานระบาดวิทยาทางสัตวแพทย์ ดังนี้

1. ในการติดตั้งโปรแกรม ผู้ใช้จะต้องเลือกระบบปฏิบัติการที่เหมาะสมที่ใช้งาน และในการเริ่มใช้งานในครั้งแรกจะต้องติดตั้งทั้งโปรแกรม R และ RStudio จากนั้นเมื่อต้องการใช้งานในแต่ละครั้งให้เรียกใช้งานโปรแกรม RStudio
2. การติดตั้งชุดคำสั่งในการใช้งาน ให้ติดตั้งชุดคำสั่งที่ต้องการใช้เฉพาะในครั้งแรกที่นำมาใช้งาน โดยจะต้องต่ออินเทอร์เน็ตทุกครั้งที่ทำกรดาวน์โหลดและติดตั้งชุดคำสั่งนั้นๆ จากนั้นจะเป็นการเรียกใช้งานชุดคำสั่งแต่ละชุดเมื่อต้องการใช้งาน
3. การนำเข้าและส่งออกข้อมูล คำสั่งของแต่ละระบบปฏิบัติการ เช่น Windows หรือ Mac OS จะมีความแตกต่างกัน ผู้ใช้จะต้องเลือกใช้กันอย่างถูกต้องเหมาะสม
4. การเตรียมข้อมูลสำหรับการใช้ในการวิเคราะห์ การจัดการข้อมูลเพื่อเตรียมสำหรับนำไปใช้วิเคราะห์ในโปรแกรม ควรเตรียมข้อมูลในรูปแบบ long form (หน้า 43) เนื่องจากการวิเคราะห์ทางสถิติส่วนมากต้องการรูปแบบนี้ นอกจากนี้ไม่ควรรวมข้อมูลอื่นๆ ที่ไม่เกี่ยวข้องกับการวิเคราะห์ในไฟล์ดังกล่าวเนื่องจากบางครั้งผู้วิเคราะห์ข้อมูลอาจจะเกิดความสับสน
5. การจัดตัวแปร ในการวิเคราะห์ปัจจัยเสี่ยงต่างๆ ผู้ใช้จะต้องทำการกำหนดให้ตัวแปรแบบแบ่งกลุ่มเป็นปัจจัย (as.factor) ก่อนทุกครั้ง (หน้า 45)
6. ผู้ใช้จะต้องจัดเตรียมรูปแบบของวันที่ ให้เป็น ปี-เดือน-วัน เพื่อนำมาใช้กับชุดคำสั่ง EpiCurves (หน้า 54)
7. เมื่อคำสั่งต่างๆ ที่ใช้เกิด error ให้ตรวจสอบคำสั่ง ชื่อออบเจ็ค และชื่อตัวแปรต่างๆ ให้ถูกต้อง เพราะชื่อตัวแปรต้องเป็นเป็นอักษรตัวเล็กหรือตัวใหญ่ให้ตรงกับชื่อเป็นชื่อคอลัมน์หรือออบเจ็ค

เอกสารอ้างอิง

- วีระศักดิ์ จงสู่วิวัฒน์วงศ์. 2563. การวิเคราะห์ข้อมูลสุขภาพในระบบสาธารณสุขไทยโดยใช้โปรแกรม R สถาบันวิจัยและพัฒนาสุขภาพภาคใต้ คณะแพทยศาสตร์ มหาวิทยาลัยสงขลานครินทร์. ภูเก็ต มีเดีย, จังหวัดสงขลา.
- Batra, N., Spina, A., Blomquist, P., Campbell, F., Laurenson-Schafer, H., Florence, I., Fischer, N., Ndiaye, A., Coyer, L., Polonsky, J., Izawa, Y., Bailey, C., Molling, D., Berry, I., Buajitti, E., Mousset, M., Hollis, S., Lin, W. 2021. The Epidemiologist R Handbook: R for applied epidemiology and public health. Available online at: <https://epirhandbook.com/en/>
- Chang, W. 2013. R graphics cookbook. O'Reilly Media, Inc., CA. 397 pp.
- Kabacoff, R.I. 2015. R in action: Data analysis and graphic with R (2nd ed.). Manning publications, Shelter Island, NY. 447 pp.
- Porta, M.S., International Epidemiology Association. 2014. A dictionary of epidemiology (6thed.). Oxford University Press. UK.
- R Development Core Team. 2005. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL: <http://www.R-project.org>.
- Venables, W.N., Smith, D.M., the R Core Team. 2017. An introduction to R: Note on R: A programming environment for data analysis and graphics version 3.4.1. Available online at: <https://docs.google.com/viewer?a=v&pid=sites&srcid=ZGVmYXVsdGRvbWFpbm9xzdGF0aXN0aWNzdWZmcGt8Z3g6MmNiYWl3ODVLZTE1NzA4OA>.

ภาคผนวก ก

ข้อมูลสำหรับการวิเคราะห์เชิงพรรณนาและเชิงอนุมาน

id	Status	AgeFarmer	HusbandryTime	totalcattle	casenumber	Farmtype	F_pregnant	F_calfinfarm	F_mixbreed	F_OthersSpicies	F_Sharebull	F-ManyInseminator	F_insect
1	Case	47	20	15	3	large	0	1	0	1	0	1	1
2	Control	64	30	13		large	1	0	0	0	0	0	1
3	Case	61	29	12	2	large	0	0	0	1	0	0	1
4	Control	57	22	9		large	1	1	0	0	0	0	1
5	Control	65	15	8		large	1	0	0	0	1	0	1
6	Control	65	20	7		large	0	0	0	1	0	0	1
7	Case	65	20	7	1	large	1	1	0	0	1	0	1
8	Case	47	17	7	1	large	0	1	0	1	0	0	0
9	Control	56	30	6		large	1	1	0	1	0	0	1
10	Control	64	26	6		large	0	1	0	0	0	0	1
11	Case	66	20	6	2	large	1	1	0	1	0	0	1
12	Control	69	20	6		large	1	0	0	1	0	0	1
13	Case	45	20	6	1	large	1	1	0	1	0	0	1
14	Control	64	20	6		large	0	0	0	1	0	0	1
15	Case	50	15	6	1	large	1	0	1	0	0	0	1
16	Control	57	15	6		large	0	0	0	0	0	0	1
17	Control	57	15	6		large	0	0	0	1	0	0	1
18	Case	51	20	5	2	small	1	1	0	1	0	1	1
19	Control	52	16	5		small	1	1	0	1	0	0	1
20	Case	51	20	4	1	small	1	1	1	1	0	0	1
21	Control	63	30	4		small	1	0	0	0	0	1	1
22	Case	54	25	4	2	small	1	0	1	1	0	0	1
23	Case	55	25	4	3	small	0	1	1	1	1	0	1
24	Case	60	20	4	4	small	1	1	0	1	0	0	1
25	Control	53	20	4		small	1	1	0	1	0	0	1
26	Control	55	20	4		small	1	0	0	0	0	0	1
27	Case	51	20	4	1	small	1	0	1	0	0	0	1
28	Case	62	19	4	1	small	0	1	0	1	1	1	1
29	Control	60	18	4		small	0	1	0	0	0	0	1
30	Case	70	15	4	1	small	1	0	0	0	0	0	1
31	Control	78	20	4		small	1	1	0	1	1	0	0
32	Control	64	30	3		small	0	0	0	1	0	0	1
33	Case	50	30	3	1	small	0	0	1	0	0	0	1
34	Case	87	20	3	2	small	1	0	0	0	0	0	1
35	Control	58	20	3		small	1	1	0	0	0	0	1
36	Case	48	15	3	3	small	1	1	0	1	0	0	1

37	Case	52	12	3	1	small	0	1	0	1	0	0	1
38	Control	77	20	3		small	1	0	0	0	1	0	0
39	Control	61	20	3		small	1	1	0	1	0	0	0
40	Control	62	20	2		small	1	0	0	1	0	0	1
41	Control	60	20	2		small	1	0	0	1	0	0	1
42	Control	55	15	2		small	0	0	0	1	0	0	1
43	Control	55	15	2		small	0	0	0	0	0	0	1
44	Case	55	3	24	5	large	1	1	0	1	0	0	1
45	Control	62	4	9		large	1	0	0	0	0	0	0
46	Control	45	4	8		large	1	0	0	0	0	0	0
47	Control	39	3	7		large	1	1	0	1	1	0	1
48	Control	50	4	6		large	1	0	0	0	0	0	0
49	Case	53	3	5	2	small	1	0	0	1	0	0	1
50	Control	50	3	5		small	1	1	0	1	1	0	1
51	Control	52	3	5		small	1	1	0	0	1	0	1
52	Control	44	2	5		small	0	1	0	1	0	1	1
53	Control	54	2	5		small	1	1	0	1	0	0	1
54	Control	59	2	5		small	0	1	0	1	0	1	1
55	Control	58	4	4		small	0	0	1	0	0	0	1
56	Control	44	4	4		small	1	0	0	1	0	0	1
57	Control	52	2	4		small	0	0	0	1	0	0	1
58	Case	55	1	4	2	small	0	0	0	1	0	0	1
59	Control	47	1	4		small	1	1	0	1	0	0	1
60	Control	36	1	4		small	1	0	0	1	0	0	1
61	Control	45	3	3		small	0	0	0	1	0	0	1
62	Control	60	2	3		small	0	0	1	1	0	1	1
63	Case	41	2	3	1	small	1	1	0	1	0	0	1
64	Control	58	1	3		small	0	0	0	1	0	0	1
65	Control	60	1	3		small	1	1	0	1	0	0	1
66	Case	46	1	2	2	small	0	1	0	1	0	1	1
67	Control	55	4	2		small	0	0	0	0	0	0	1
68	Control	49	3	2		small	0	0	0	1	0	1	1
69	Control	42	2	2		small	1	0	0	1	0	0	1
70	Control	50	2	2		small	1	1	0	0	0	1	1
71	Control	38	1	2		small	0	0	1	1	0	1	1
72	Control	45	1	2		small	0	0	0	1	0	0	1
73	Control	37	2	2		small	0	0	0	1	1	0	0
74	Control	48	4	1		small	1	0	1	0	0	0	1
75	Control	51	3	1		small	1	0	0	1	0	0	1
76	Control	42	3	1		small	0	0	0	1	0	0	1

77	Control	57	1	1		small	0	0	0	1	0	0	1
78	Case	64	5	14	4	large	1	1	0	1	0	0	1
79	Control	58	8	12		large	1	1	0	1	0	0	1
80	Control	51	10	10		large	1	1	0	1	0	1	1
81	Case	68	10	10	2	large	0	1	0	1	0	1	1
82	Case	70	10	10	3	large	1	1	0	1	0	1	1
83	Case	65	10	9	5	large	1	1	1	0	0	1	1
84	Case	64	10	9	4	large	1	1	0	1	1	1	1
85	Case	52	5	9	6	large	1	1	0	1	0	0	1
86	Case	40	5	9	2	large	1	0	0	1	0	1	1
87	Control	47	10	9		large	1	0	0	1	0	1	0
88	Control	64	6	9		large	1	0	0	1	0	1	0
89	Case	56	10	8	3	large	1	0	1	1	0	0	1
90	Case	60	10	8	2	large	1	1	0	1	0	0	1
91	Case	56	10	8	1	large	0	1	0	1	0	0	1
92	Case	37	6	8	3	large	0	1	0	1	0	0	1
93	Control	52	10	7		large	1	0	1	0	0	0	1
94	Case	50	10	7	4	large	0	1	0	1	1	0	1
95	Case	54	8	6	2	large	0	0	0	0	0	0	1
96	Control	58	8	6		large	1	0	0	1	0	0	1
97	Control	55	5	6		large	0	0	0	1	0	0	1
98	Case	67	9	6	1	large	1	0	0	1	0	0	0
99	Control	71	10	5		small	1	0	1	1	0	0	1
100	Control	63	10	5		small	1	0	0	0	0	1	1
101	Case	57	7	5	1	small	1	0	0	1	0	0	1
102	Case	65	6	5	1	small	1	0	0	1	0	0	1
103	Case	60	5	5	2	small	1	1	0	1	0	0	1
104	Case	46	5	5	3	small	1	1	0	0	0	1	1
105	Control	53	5	5		small	1	1	0	1	0	0	1
106	Control	50	10	5		small	1	1	0	0	0	0	0
107	Control	65	10	4		small	1	1	0	1	0	0	1
108	Case	57	10	4	1	small	0	1	0	1	1	0	1
109	Case	66	10	4	1	small	1	0	0	0	0	0	1
110	Control	53	10	4		small	1	0	0	0	0	0	1
111	Control	60	8	4		small	1	0	0	1	0	0	1
112	Control	66	8	4		small	1	0	0	1	0	0	1
113	Control	70	8	4		small	1	0	0	0	0	0	1
114	Control	64	7	4		small	1	0	1	1	0	0	1
115	Control	36	6	4		small	0	0	0	0	0	0	1
116	Control	55	5	4		small	1	1	0	1	1	1	1

117	Control	38	5	4		small	0	1	0	1	0	0	1
118	Case	50	5	4	1	small	0	1	0	1	0	1	1
119	Control	64	5	4		small	1	0	1	0	0	0	1
120	Case	42	7	4	1	small	1	1	1	0	1	0	1
121	Control	67	10	4		small	1	0	0	1	0	0	0
122	Control	57	5	4		small	1	0	0	1	0	0	0
123	Control	46	10	3		small	1	1	0	1	0	0	1
124	Case	56	10	3	5	small	1	0	0	0	0	0	1
125	Control	69	10	3		small	1	0	0	1	0	1	1
126	Control	65	8	3		small	1	0	0	1	0	0	1
127	Case	65	7	3	3	small	1	0	0	0	0	0	1
128	Control	65	7	3		small	1	0	0	1	0	0	1
129	Control	32	6	3		small	0	1	0	1	0	1	1
130	Control	61	6	3		small	1	0	0	1	0	0	1
131	Case	67	6	3	2	small	1	0	0	0	0	0	1
132	Control	62	5	3		small	1	0	0	1	1	0	1
133	Control	60	5	3		small	0	0	1	0	0	0	1
134	Control	40	5	3		small	1	0	0	1	0	0	1
135	Control	69	10	3		small	1	0	0	0	0	0	0
136	Control	48	6	3		small	0	0	0	0	0	0	0
137	Control	57	5	3		small	1	0	0	0	0	0	0
138	Control	45	10	2		small	1	0	0	1	0	0	1
139	Control	64	10	2		small	0	1	0	1	0	0	1
140	Control	67	10	2		small	1	0	0	1	0	0	1
141	Control	57	10	2		small	1	0	1	0	0	1	1
142	Control	52	10	2		small	0	1	0	1	0	0	1
143	Control	61	10	2		small	1	0	0	0	1	0	1
144	Control	52	10	2		small	1	0	0	0	0	0	1
145	Control	59	10	2		small	1	0	0	1	0	0	1
146	Control	54	10	2		small	1	0	0	1	0	1	1
147	Control	53	7	2		small	1	0	0	1	0	0	1
148	Control	57	6	2		small	0	0	0	1	0	0	1
149	Control	69	5	2		small	1	0	0	1	0	1	1
150	Control	42	5	2		small	0	0	0	1	0	1	1
151	Control	57	5	2		small	1	0	0	0	1	0	1
152	Control	52	5	2		small	1	0	0	0	0	0	1
153	Control	59	5	2		small	1	0	0	1	0	0	1
154	Control	56	5	2		small	1	0	0	0	1	0	1
155	Case	51	8	2	4	small	0	0	0	1	0	0	0
156	Control	82	10	1		small	1	0	0	1	0	0	1

157	Control	48	5	1		small	0	0	0	1	0	0	1
158	Control	66	5	1		small	0	0	1	1	0	0	1
159	Case	39	5	1	1	small	1	0	0	1	0	1	0

ภาคผนวก ข
ข้อมูลสำหรับการวิเคราะห์เชิงเวลา

date	count	Status
2021-08-01	0	NA
2021-08-02	0	NA
2021-08-03	0	NA
2021-08-04	0	NA
2021-08-05	0	NA
2021-08-06	1	Confirmed
2021-08-07	2	Suspected
2021-08-07	3	Confirmed
2021-08-08	0	NA
2021-08-09	0	NA
2021-08-10	0	NA
2021-08-11	0	NA
2021-08-12	2	Confirmed
2021-08-12	3	Suspected
2021-08-13	2	Confirmed
2021-08-14	4	Suspected
2021-08-14	3	Confirmed
2021-08-15	6	Confirmed
2021-08-15	3	Suspected
2021-08-16	5	Confirmed
2021-08-16	4	Suspected
2021-08-17	8	Suspected
2021-08-18	4	Suspected
2021-08-18	2	Confirmed
2021-08-19	2	Confirmed
2021-08-20	1	Confirmed

2021-08-21	5	Suspected
2021-08-21	3	Confirmed
2021-08-22	5	Suspected
2021-08-22	4	Confirmed
2021-08-23	1	Confirmed
2021-08-24	4	Suspected
2021-08-24	5	Confirmed
2021-08-25	4	Suspected
2021-08-25	1	Confirmed
2021-08-26	1	Suspected
2021-08-27	3	Confirmed
2021-08-28	1	Confirmed
2021-08-29	2	Suspected
2021-08-30	4	Confirmed
2021-08-30	3	Suspected
2021-08-31	2	Suspected
2021-09-01	1	Confirmed
2021-09-02	5	Confirmed
2021-09-03	6	Confirmed
2021-09-04	1	Confirmed
2021-09-05	5	Confirmed
2021-09-06	1	Suspected
2021-09-07	2	Confirmed
2021-09-08	3	Confirmed
2021-09-09	1	Suspected
2021-09-10	5	Confirmed
2021-09-11	6	Confirmed
2021-09-12	5	Confirmed
2021-09-13	5	Suspected
2021-09-14	4	Confirmed

2021-09-15	2	Suspected
2021-09-16	3	Suspected
2021-09-17	1	Suspected
2021-09-18	1	Suspected
2021-09-19	0	NA
2021-09-20	0	NA
2021-09-21	0	NA
2021-09-22	0	NA
2021-09-23	0	NA
2021-09-24	0	NA
2021-09-25	0	NA
2021-09-26	0	NA
2021-09-27	0	NA
2021-09-28	0	NA
2021-09-29	0	NA
2021-09-30	0	NA

ภาคผนวก ค
ข้อมูลสำหรับการวิเคราะห์เชิงพื้นที่

id	District	Latitude	Longitude	Status
1	San Kamphaeng	18.7859562	99.1905035	Case
2	San Kamphaeng	18.7320833	99.1208611	Control
3	San Kamphaeng	18.7078007	99.1671498	Case
4	Mae On	18.7823149	99.2640381	Control
5	Mae On	18.7815986	99.2592357	Control
6	Mae On	18.7956532	99.24526	Control
7	Mae On	18.7812193	99.2590453	case
8	Mae On	18.7837127	99.2597893	Case
9	Mae On	18.7494125	99.2476029	Control
10	Mae On	18.7866477	99.2481601	Control
11	Mae On	18.7624828	99.2597757	case
12	Mae On	18.7856527	99.2488229	Control
13	Mae On	18.7700321	99.2578232	Case
14	Mae On	18.8144437	99.2580332	Control
15	Mae On	18.7859016	99.2649642	Case
16	Mae On	18.7808862	99.2645103	Control
17	Mae On	18.7970001	99.2584906	Control
18	Mae On	18.812921	99.2545573	Case
19	Mae On	18.6393221	99.2855939	Control
20	Mae On	18.7604279	99.2621445	Case
21	Mae On	18.6041643	99.2646257	Control
22	Mae On	18.582487	99.2637004	Case
23	San Sai	18.8844448	99.0686084	Case
24	San Sai	18.8800278	99.0824998	case
25	San Sai	19.0454726	98.9639446	Control
26	San Kamphaeng	18.691478	99.22481	Control
27	San Kamphaeng	18.7273055	99.1608884	case
28	San Kamphaeng	18.723694	99.197139	Case
29	San Kamphaeng	18.7313039	99.1619271	Control
30	San Kamphaeng	18.7318241	99.1857282	Case
31	San Kamphaeng	18.7256027	99.1891078	Control
32	San Kamphaeng	18.7270683	99.220773	Control
33	San Kamphaeng	18.7165204	99.2135695	case
34	San Kamphaeng	18.7174371	99.2100896	Case
35	San Kamphaeng	18.7332051	99.1470926	Control
36	San Kamphaeng	18.6898733	99.0975974	case

37	San Kamphaeng	18.7843454	99.1835557	Case
38	San Kamphaeng	18.7305446	99.1397007	Control
39	San Kamphaeng	18.7297222	99.1627778	Control
40	San Kamphaeng	18.7322939	99.119326	Control
41	San Kamphaeng	18.734844	99.117127	Control
42	Mae On	18.7835297	99.2613831	Control
43	Mae On	18.7834436	99.26478	Control
44	Mae On	18.7654081	99.2619055	Case
45	Mae On	18.7920048	99.2498951	Control
46	Mae On	18.78619	99.2655529	Control
47	Mae On	18.7962029	99.2464945	Control
48	Mae On	18.7904329	99.2493803	Control
49	Mae On	18.7920506	99.2494396	Case
50	Mae On	18.7857174	99.2544121	Control
51	Mae On	18.7568843	99.2494923	Control
52	Mae On	18.7476651	99.2566319	Control
53	Mae On	18.7895123	99.2484585	Control
54	Mae On	18.7812193	99.2590453	Control
55	Mae On	18.7841434	99.2555009	Control
56	Mae On	18.79221	99.2451317	Control
57	Mae On	18.7624195	99.2597946	Control
58	Mae On	18.7854482	99.2530452	Case
59	Mae On	18.7858592	99.2498385	Control
60	Mae On	18.7850832	99.248898	Control
61	Mae On	18.7539876	99.2595066	Control
62	Mae On	18.7853982	99.2498758	Control
63	Mae On	18.7769355	99.2589154	case
64	Mae On	18.7855095	99.2544877	Control
65	Mae On	18.7731332	99.2570499	Control
66	Mae On	18.7850292	99.2486702	Case
67	Mae On	18.7824032	99.2634826	Control
68	Mae On	18.782212	99.2600432	Control
69	Mae On	18.7852834	99.247864	Control
70	Mae On	18.786985	99.2593008	Control
71	Mae On	18.6438233	99.2853455	Control
72	Mae On	18.7838827	99.2610231	Control
73	Mae On	18.7662104	99.2634153	Control
74	Mae On	18.7690903	99.265659	Control
75	Mae On	18.7824322	99.2641996	Control
76	Mae On	18.7870413	99.2642827	Control

77	Mae On	18.7974365	99.2631032	Control
78	Mae On	18.7677189	99.2575066	Case
79	Mae On	18.7815359	99.2588465	Control
80	Mae On	18.7683333	99.2488889	Control
81	Mae On	18.8157325	99.2606261	case
82	Mae On	18.7491227	99.2480673	case
83	Mae On	18.6325015	99.2835632	Case
84	Mae On	18.7732596	99.2699068	Case
85	Mae On	18.769263	99.2518256	Case
86	Mae On	18.7662112	99.2562709	Case
87	Mae On	18.7537256	99.2594682	Control
88	Mae On	18.7694797	99.2650618	Control
89	Mae On	18.7567206	99.2569491	Case
90	Mae On	18.7663538	99.264241	case
91	Mae On	18.7947754	99.2460938	Case
92	Mae On	18.7625677	99.2637322	Case
93	Mae On	18.7862545	99.2646705	Control
94	Mae On	18.7693865	99.2540839	Case
95	Mae On	18.7759061	99.2582402	Case
96	Mae On	18.7836011	99.2620285	Control
97	Mae On	18.7335207	99.1479656	Control
98	Mae On	18.7710334	99.2591911	Case
99	Mae On	18.7654407	99.2644673	Control
100	Mae On	18.7927686	99.2464136	Control
101	Mae On	18.7987276	99.2703364	Case
102	Mae On	18.7654571	99.2656153	Case
103	Mae On	18.7773062	99.2588401	case
104	Mae On	18.766654	99.2563664	Case
105	Mae On	18.7778576	99.2453861	Control
106	Mae On	18.7685566	99.2659807	Control
107	Mae On	18.6460433	99.2875773	Control
108	Mae On	18.646167	99.2893884	case
109	Mae On	18.6290088	99.2802109	Case
110	Mae On	18.6445229	99.2888547	Control
111	San Sai	18.8413851	99.0516663	Control
112	San Sai	18.9855971	98.9743775	Control
113	San Sai	18.9683329	98.9894443	Control
114	San Sai	19.0437973	98.9608068	Control
115	San Sai	19.0752779	98.9400005	Control
116	San Sai	18.8801945	99.1112499	Control

117	San Sai	18.9726196	98.9796043	Control
118	San Sai	18.9776879	98.9669117	Case
119	San Sai	18.9619378	98.9892756	Control
120	San Sai	18.9674504	98.9828632	case
121	San Sai	18.9186547	99.0249665	Control
122	San Sai	18.8533333	99.077222	Control
123	San Sai	19.0445419	98.9841668	Control
124	San Sai	18.8719165	99.0266116	case
125	San Sai	18.9811235	98.9749767	Control
126	San Sai	18.8388886	99.0375004	Control
127	San Sai	18.9386949	99.0015557	Case
128	San Sai	18.9811948	98.9740276	Control
129	San Sai	18.9785741	98.9699421	Control
130	San Sai	18.9747784	98.9704177	Control
131	San Sai	18.9502152	98.9808038	Case
132	San Sai	19.0081206	98.9922847	Control
133	San Sai	18.971255	98.9804879	Control
134	San Sai	18.862222	99.074167	Control
135	San Sai	19.038889	98.971944	Control
136	San Sai	18.950162	99.0014628	Control
137	San Kamphaeng	18.7262694	99.1911005	Control
138	San Kamphaeng	18.7345071	99.2351068	Control
139	San Kamphaeng	18.7237498	99.1971116	Control
140	San Kamphaeng	18.7125001	99.2161112	Control
141	San Kamphaeng	18.7370995	99.1972883	Control
142	San Kamphaeng	18.734213	99.2004913	Control
143	San Kamphaeng	18.7288275	99.2157572	Control
144	San Kamphaeng	18.7163892	99.1997219	Control
145	San Kamphaeng	18.7158337	99.2141662	Control
146	San Kamphaeng	18.7335388	99.1479372	Control
147	San Kamphaeng	18.7237498	99.1971116	Control
148	San Kamphaeng	18.7323992	99.2220131	Control
149	San Kamphaeng	18.7172839	99.2097195	Control
150	San Kamphaeng	18.7189678	99.2071984	Control
151	San Kamphaeng	18.7318362	99.1142791	Control
152	San Kamphaeng	18.6909489	99.0974463	Control
153	San Kamphaeng	18.7132047	99.212503	Control
154	San Kamphaeng	18.7375357	99.1196522	Control
155	San Kamphaeng	18.731174	99.127199	Case
156	San Kamphaeng	18.7311327	99.1388569	Control

157	San Kamphaeng	18.7319317	99.1205399	Control
158	San Kamphaeng	18.727495	99.1325364	Control
159	San Kamphaeng	18.7310478	99.1176271	Case