



รายงานวิจัยฉบับสมบูรณ์

โครงการ

การพัฒนาการใช้โปรแกรม R ในการวิเคราะห์ข้อมูล
R Program Development for Research Utilization

โดย

หัชชา ศรีปลั่ง และคณะ

เดือน ปี ที่เสร็จโครงการตามสัญญา

ตุลาคม 2549

คำนำ

ตามสัญญาเลขที่ RDG4730207 ซึ่งสำนักงานกองทุนสนับสนุนการวิจัย (สกว.) ได้ทำกับมหาวิทยาลัยสงขลานครินทร์ โดย นายประเสริฐ ชิตพงศ์ อธิการบดีมหาวิทยาลัยสงขลานครินทร์ (ในขณะนั้น) เป็นฝ่ายที่ 1 ในฐานะ “ผู้รับทุน” และข้าพเจ้า นายหัชชา ศรีปลั่ง สังกัดหน่วยระบาดวิทยา คณะแพทยศาสตร์ มหาวิทยาลัยสงขลานครินทร์ ในฐานะ “หัวหน้าโครงการวิจัยผู้รับทุน” เป็นฝ่ายที่ 2 เมื่อวันที่ 1 เดือน พฤศจิกายน พ.ศ. 2546 โดยตกลงทำสัญญาในโครงการ “การพัฒนาการใช้โปรแกรม R ในการวิเคราะห์ข้อมูล” ในชื่อภาษาอังกฤษว่า “R Program Development for Research Utilization” และมีคณาจารย์และบุคลากรในมหาวิทยาลัยสงขลานครินทร์เป็นผู้ร่วมวิจัยด้วยอีกจำนวนหนึ่ง

งานดังกล่าวได้รับการสนับสนุนและช่วยเหลืออย่างดีจากผู้เกี่ยวข้อง ทั้งในส่วนของสำนักงานกองทุนสนับสนุนการวิจัย (สกว.) มูลนิธิสาธารณสุขแห่งชาติ (มสช.) ในฐานะผู้ทำการแทนผู้ให้ทุน และจากหน่วยงานต่างๆ ในมหาวิทยาลัยสงขลานครินทร์ที่เกี่ยวข้อง บัดนี้งานวิจัยตามโครงการดังกล่าวได้ดำเนินการสิ้นสุดลงแล้วตามแผนงานที่ได้เสนอไว้อย่างดี จึงได้สรุปการดำเนินการทั้งหมดมา ณ ที่นี้

เนื่องจากงานวิจัยนี้มีส่วนที่เป็นนวัตกรรม ประดิษฐ์กรรมด้วยส่วนหนึ่ง คือการสร้างสภาพแวดล้อม R-ICE ซึ่งเป็นสภาพแวดล้อมในการวิเคราะห์ข้อมูลทางสถิติ และยังมีอีกส่วนหนึ่งซึ่งเป็นการเผยแพร่การใช้งานโปรแกรม R และสภาพแวดล้อมที่ประดิษฐ์ขึ้นตามโครงการนี้ ด้วยการนำองค์ความรู้ที่ได้ศึกษาและสร้างขึ้นนี้เผยแพร่แก่สาธารณะในรูปแบบการอบรมเชิงปฏิบัติการ การสร้างเว็บไซต์ การเรียนทางสื่ออินเทอร์เน็ต และการตีพิมพ์หนังสือแก่ประชาชน คณะผู้วิจัยจึงหวังเป็นอย่างยิ่งว่างานวิจัยนี้ได้เป็นประโยชน์ต่อการพัฒนางานวิจัยของประเทศในด้านต่างๆ ในอนาคตต่อไป และช่วยปลูกฝัง เน้นย้ำเจตคติในการเคารพสิทธิในทรัพย์สินทางปัญญาของซอฟต์แวร์ทางการค้า และเจตคติที่จะไม่ใช้ซอฟต์แวร์ที่ได้มาโดยผิดกฎหมาย ในหมู่ประชาชนชาวไทยโดยทั่วไปอีกด้วย

นายหัชชา ศรีปลั่ง
(หัวหน้าโครงการวิจัยผู้รับทุน)

สารบัญ

	หน้า
บทสรุปสำหรับผู้บริหาร	1
บทคัดย่อ	5
Abstract	7
เนื้อหางานวิจัย	9
หลักการและเหตุผล	9
วัตถุประสงค์ของการวิจัย	11
หน่วยงานที่รับผิดชอบ	11
คณะผู้ร่วมวิจัย	11
วิธีการดำเนินงาน	12
ผลการดำเนินงาน	14
1. การพัฒนาฟังก์ชันในการจัดการและวิเคราะห์ข้อมูลทางสถิติ	15
2. การทำความเข้าใจการใช้งานเป็นภาษาอังกฤษและภาษาไทย	26
3. การทดสอบฟังก์ชัน	30
4. การพัฒนารายการเมนูกราฟิก	30
5. การทำแพ็คเกจเพื่อเผยแพร่แก่สาธารณะ	31
6. การหาสภาพแวดล้อมเสริมในการวิเคราะห์ข้อมูลทางสถิติ	62
7. การสอนในหลักสูตรระดับบัณฑิตวิทยาลัย และหลักสูตรสถิติ	63
8. การจัดทำคู่มือการใช้งานโปรแกรม R	63
9. การจัดประชุมอบรมเชิงปฏิบัติการ	64
10. การสร้างกลุ่มผู้ใช้ R	67
11. การจัดทำเว็บไซต์	67
12. การจัดการเรียนการสอนทางอินเทอร์เน็ต	73
13. การอบรมและเข้าร่วมประชุมในต่างประเทศ	75
14. การจัดทำ mirror site ของ CRAN	76
15. การจัดทำหนังสือเพื่อเผยแพร่แก่ประชาชน	76
วิจารณ์และสรุป	77
ภาคผนวก	79
ภาคผนวก ก.	79
ภาคผนวก ข.	80
ภาคผนวก ค.	81
ภาคผนวก ง.	83
ภาคผนวก จ.	86

บทสรุปสำหรับผู้บริหาร

หลักการและเหตุผล

ปัจจุบันมีการใช้ซอฟต์แวร์ทางการค้าอย่างผิดกฎหมายกันมากในประเทศไทย โดยเฉพาะซอฟต์แวร์ทางสถิติ ซึ่งเป็นส่วนสำคัญในการสร้างองค์ความรู้ในสาขาต่างๆ ปัจจุบันมีซอฟต์แวร์ทางสถิติที่มีลิขสิทธิ์แบบ GPL คือใช้งานได้ แจกจ่ายได้ และนำไปปรับปรุงพัฒนาได้โดยไม่เสียค่าใช้จ่าย คือโปรแกรม **R** ซึ่งโปรแกรมนี้มีสมรรถนะสูงมาก และได้รับการยอมรับในวงวิชาการสถิติทั่วโลก แต่การใช้งานยังใช้ได้ค่อนข้างยาก เนื่องจากโปรแกรมและคู่มือต่างๆ เน้นไปทางด้านการเป็นภาษาคอมพิวเตอร์ที่มีสมรรถนะสูง ทำให้ฟังก์ชันบางด้านยังมีน้อย และส่วนติดต่อกับผู้ใช้ไม่มีส่วนที่เป็นกราฟิกและรายการเมนูต่างๆ ทำให้ผู้ใช้ที่เป็นนักสถิติและผู้ที่ไม่มีความรู้พื้นฐานด้านการเขียนโปรแกรมคอมพิวเตอร์ใช้งานได้ยาก

วัตถุประสงค์

โครงการวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนา เผยแพร่ และสนับสนุนให้มีการใช้โปรแกรม **R** ซอฟต์แวร์ทางสถิติที่ไม่ต้องเสียค่าลิขสิทธิ์ ในการวิเคราะห์ผลงานวิจัยทั้งทางด้านวิทยาศาสตร์สุขภาพ และระเบียบวิธีทางสถิติ

หน่วยงานที่รับผิดชอบ

หน่วยระบาดวิทยา คณะแพทยศาสตร์ และภาควิชาคณิตศาสตร์ คณะวิทยาศาสตร์
มหาวิทยาลัยสงขลานครินทร์ อำเภอหาดใหญ่ จังหวัดสงขลา

คณะผู้ร่วมวิจัย

รศ. นพ. หัซซา ศรีปลั่ง หน่วยระบาดวิทยา คณะแพทยศาสตร์ หัวหน้าโครงการ

วิธีดำเนินการ

การดำเนินการแบ่งได้เป็นสามกลุ่มกิจกรรมคือ 1. การพัฒนาการใช้โปรแกรม **R** ให้ใช้ได้ง่ายขึ้น 2. พัฒนาส่วนติดต่อกับผู้ใช้เป็นภาษาไทย และ 3. การจัดกิจกรรมเผยแพร่การใช้งานโปรแกรม **R** และสิ่งประดิษฐ์ที่สร้างขึ้นในกิจกรรมสองกลุ่มที่กล่าวข้างต้น ในการวิเคราะห์ทางสถิติ

การพัฒนาการใช้โปรแกรม **R** ให้ใช้ได้ง่ายขึ้น ประกอบด้วยการสร้างฟังก์ชันเพิ่มเติม โดยเฉพาะในด้านการจัดการข้อมูล และการแสดงผลการทำงานที่เข้าใจได้ง่าย ซึ่งเป็นจุดอ่อนที่สำคัญของ **R** และทำชุดฟังก์ชันเหล่านั้นเป็นแพ็คเกจเพื่อเผยแพร่แก่ผู้ต้องการใช้ และการพัฒนาส่วนติดต่อกับผู้ใช้ให้เป็นแบบรายการเมนูและหน้าต่างกราฟิกเช่นเดียวกับโปรแกรมทางการค้าอื่นๆ

ในด้านการพัฒนาส่วนติดต่อกับผู้ใช้เป็นภาษาไทยนั้น ประกอบด้วยการทำความเข้าใจแบบ on-line และแบบเป็นแฟ้มสำหรับฟังก์ชันและแพ็คเกจที่สร้างขึ้นนั้นเป็นภาษาไทย และการทำรายการเมนูและหน้าต่างกราฟิกต่างๆ ให้เป็นภาษาไทย

ในด้านการจัดกิจกรรมเผยแพร่การใช้งานโปรแกรม **R** ได้จัดการประชุมเชิงปฏิบัติการให้แก่ผู้ใช้กลุ่มต่างๆ จัดการเรียนการสอนผ่านทางอินเทอร์เน็ต และการตีพิมพ์หนังสือเผยแพร่แก่ประชาชนและนักวิชาการ และยังคงจัดทำเว็บไซต์เพื่อการเรียนรู้ แลกเปลี่ยนความคิดเห็น และการจัดกิจกรรมต่างๆ ที่เกี่ยวข้องกับ **R**

ผลการดำเนินงาน

1. การพัฒนาฟังก์ชันในการจัดการและวิเคราะห์ข้อมูลทางสถิติ

คณะผู้วิจัยได้สร้างแพ็คเกจชุดฟังก์ชันด้านการจัดการแฟ้มและข้อมูล ชื่อ ice มี 50 ฟังก์ชันที่ได้สร้างขึ้นใหม่ แพ็คเกจ epid ซึ่งเป็นชุดฟังก์ชันพื้นฐานด้านระบาดวิทยา มีฟังก์ชัน 9 ฟังก์ชัน

2. การทำความช่วยเหลือการใช้งานเป็นภาษาอังกฤษและภาษาไทย

ความช่วยเหลือของฟังก์ชันและแพ็คเกจที่สร้างขึ้นเหล่านี้มีทั้งที่เป็นภาษาอังกฤษและภาษาไทย และในอนาคตอาจแปลเป็นภาษาอื่นๆ ได้อีกด้วย จากต้นฉบับที่ทำเป็นภาษาอังกฤษ

3. การทดสอบฟังก์ชัน

ได้ใช้กระบวนการสามขั้นตอนในการทดสอบฟังก์ชัน คือ 1. การทดสอบด้วยการทำงานตัวอย่าง 2. การทดสอบด้วยผู้ร่วมวิจัยท่านอื่น และ 3. การแจ้งความผิดพลาดในการทำงาน ทั้งในระหว่างการอบรมและการที่ผู้นำไปใช้งานจริง

4. การพัฒนารายการเมนูกราฟิก

ได้พัฒนารายการเมนูกราฟิกโดยอาศัยแพ็คเกจ tcltk ภายใน R เอง เพื่อให้ทำงานร่วมกับ R ได้อย่างสมบูรณ์ และทำงานได้บนทุกระบบปฏิบัติการ ได้สร้างสภาพแวดล้อม ICE ซึ่งเป็นสภาพแวดล้อมที่เป็นรายการเมนูและหน้าต่างกราฟิก สามารถแสดงได้ทั้งภาษาไทยและภาษาอังกฤษ

5. การทำแพ็คเกจเพื่อเผยแพร่แก่สาธารณะ

สภาพแวดล้อม ICE ประกอบด้วยแพ็คเกจหลักคือ ice.main และแพ็คเกจประกอบอื่นๆ ซึ่งเมื่อทำงานภายใต้โปรแกรม R จะเรียกเป็นสภาพแวดล้อม Integrated Computing Environment for R (R-ICE)

6. การหาสภาพแวดล้อมเสริมในการวิเคราะห์ข้อมูลทางสถิติ

เพื่อสนับสนุนการทำงานวิเคราะห์ข้อมูลอย่างครบวงจร จึงจำเป็นต้องจัดหาโปรแกรมเสริมการทำงานที่เป็นโปรแกรมฟรีด้วยเช่นกัน ที่สำคัญคือ text editor สำหรับระบบปฏิบัติการต่างๆ ซึ่งจะต้องรู้จักชื่อฟังก์ชันต่างๆ ของ R และสามารถแจ้งความผิดพลาดในการเขียนบรรทัดคำสั่งได้ ที่แนะนำในระบบวินโดวส์ได้แก่โปรแกรม Tinn-R และ Crimson Editor สำหรับระบบปฏิบัติการแมคอินทอช OSX แนะนำโปรแกรม TextWrangler และระบบลินุกซ์แนะนำโปรแกรม Blue Fish

7. การสอนในหลักสูตรระดับวิทยานานาชาติและหลักสูตรสถิติ

ได้จัดการเรียนการสอนรายวิชา Statistics I และ Statistics II ให้แก่นักศึกษาหลักสูตรระดับวิทยานานาชาติ ระดับบัณฑิตศึกษา ด้วยโปรแกรม R มาตั้งแต่ พ.ศ. 2547 และทางคณะวิทยาศาสตร์ได้ใช้โปรแกรม R ในการสอนหลักสูตรสถิติแก่นักศึกษาระดับปริญญาตรีด้วย

8. การจัดทำคู่มือการใช้งานโปรแกรม R

ได้จัดทำคู่มือการใช้งานโปรแกรม R ขึ้นสองเล่ม ทั้งนี้เพื่อให้สอดคล้องกับการจัดการอบรมเชิงปฏิบัติการที่มีสองสาย คือสายวิทยาศาสตร์สุขภาพ ซึ่งเน้นการวิเคราะห์ข้อมูลทางระบาดวิทยา และสายวิทยาศาสตร์ทั่วไป

9. การจัดประชุมอบรมเชิงปฏิบัติการ

ในการจัดกิจกรรมเพื่อเผยแพร่การใช้งานโปรแกรม **R** นั้น ได้จัดการประชุมเชิงปฏิบัติการวิเคราะห์ทางสถิติด้วยโปรแกรม **R** ขึ้น รวมทั้งสองสายได้ 13 ครั้ง ให้แก่ผู้สนใจทั้งภายในและภายนอกมหาวิทยาลัยสงขลานครินทร์ รวมถึงการจัดนอกสถานที่ด้วย

10. การสร้างกลุ่มผู้ใช้

ได้ใช้เว็บไซต์ทั่วไปสำหรับโปรแกรม **R** สำหรับคนไทย เป็นที่รวมผู้ใช้โปรแกรม **R** และแพ็คเกจต่างๆ ที่เป็นคนไทย (Thai UseR) ซึ่งมีกิจกรรมในการตอบข้อซักถาม แลกเปลี่ยนความคิดเห็น แจ้งข่าว และกิจกรรมต่างๆ ที่เกี่ยวข้องกับการใช้โปรแกรม **R** ที่ <http://medipe.psu.ac.th/R/>

11. การจัดทำเว็บไซต์

ได้จัดทำเว็บไซต์ขึ้น 4 เว็บไซต์ มีเว็บไซต์สำหรับผู้สนใจโปรแกรม **R** โดยทั่วไป ซึ่งเป็นที่พบปะสนทนาของกลุ่มผู้ใช้ **R** เว็บไซต์สำหรับหลักสูตรทางอินเทอร์เน็ต เว็บไซต์สำหรับสภาพแวดล้อม **R-ICE** และเว็บบล็อกเผยแพร่การเขียนฟังก์ชันในโปรแกรม **R**

12. การจัดการเรียนการสอนทางอินเทอร์เน็ต

จัดทำเว็บไซต์ <http://medipe.psu.ac.th/ices/> โดยใช้ชื่อว่า ICES โดยย่อมาจาก Internet Courses in Epidemiology and Statistics หลักสูตรทางสถิติที่เกี่ยวข้องกับ **R** เป็นหลักสูตรรหัส ICS02 การวิเคราะห์ทางสถิติด้วยโปรแกรม **R** และ ICS03 การวิเคราะห์ทางสถิติด้วยโปรแกรม **R-ICE**

13. การอบรมและเข้าร่วมประชุมในต่างประเทศ

หัวหน้าโครงการได้เดินทางไปอบรมเชิงปฏิบัติการเรื่อง Analysis of Time Trends ซึ่งใช้โปรแกรม **R** ในการวิเคราะห์ทางสถิติ ที่ International Agency for Research on Cancer เมือง Lyon ประเทศฝรั่งเศส และคณะผู้วิจัยส่วนหนึ่งได้ไปร่วมประชุม UseR! 2006 Conference ที่กรุงเวียนนา และได้รายงานสภาพแวดล้อม **R-ICE** เป็น oral presentation ในที่ประชุมนี้ด้วย

14. การจัดทำ mirror site ของ CRAN

ได้จัดทำ mirror ของ CRAN เพื่อให้ดาวน์โหลดโปรแกรมได้รวดเร็วยิ่งขึ้น แต่เนื่องจากข้อมูลมีขนาดใหญ่มาก การ update จึงทำได้ไม่บ่อยนัก

15. การจัดทำหนังสือเพื่อเผยแพร่แก่ประชาชน

ส่วนการตีพิมพ์หนังสือเพื่อส่งเสริมการใช้โปรแกรม **R** นั้น ได้ทำหนังสือขึ้นสองเล่ม คือ “การวิเคราะห์ทางสถิติด้วย **R-ICE**” ซึ่งมีกลุ่มเป้าหมายเป็นประชาชนทั่วไปที่สนใจการวิเคราะห์ข้อมูลทางสถิติ และ “คู่มือการใช้โปรแกรม **R** การใช้คำสั่งหรือฟังก์ชันพื้นฐาน” ซึ่งมีกลุ่มเป้าหมายเป็นนักวิชาการในสถาบันวิชาการต่างๆ ขณะนี้ทั้งสองเล่มกำลังอยู่ในระหว่างการจัดพิมพ์

วิจารณ์และสรุป

คณะผู้วิจัยได้เรียนรู้ในด้านเทคนิคต่างๆ มากมายในการเขียนฟังก์ชัน การทำแพ็คเกจ และการแปลภาษาตามวิธีการของโปรแกรม **R** และได้สร้างเครือข่ายกับนักพัฒนาโปรแกรม **R** ในระดับนานาชาติโดยการเข้าร่วมประชุม และนำเสนอผลงานเรื่องสภาพแวดล้อม **R-ICE** ที่พัฒนาขึ้นนี้ในการประชุม UseR! 2006 ที่กรุงเวียนนา

ในการเผยแพร่การใช้งานโปรแกรม **R** นั้น การจัดการประชุมเชิงปฏิบัติมีประสิทธิภาพในการเรียนรู้สูงสุด แต่ทำได้จำนวนจำกัด เนื่องจากต้องใช้ทรัพยากรมาก การจัดการเรียนการสอนทางอินเทอร์เน็ตไม่ได้ผลดีเท่าที่ควร เนื่องจากผู้เรียนอาจไม่อยู่ในสภาพที่พร้อมจะเรียน เช่น ไม่สามารถติดตั้งโปรแกรมในเครื่องที่ใช้เรียนได้

และหากมีปัญหาเล็กน้อยก็มักจะเลิกเรียนได้ง่าย คิดว่าการทำหนังสือคู่มือน่าจะเผยแพร่ได้ในวงกว้างและ
ผู้ใช้สามารถทำตามขั้นตอนและตัวอย่างในหนังสือได้ และไม่มีข้อจำกัดเรื่องเวลาและสถานที่ในการเรียนรู้
นอกจากนี้ การสร้างเว็บไซต์เพื่อเป็นที่รวมเสมือนของผู้ใช้ **R** ในประเทศไทยน่าจะนำไปสู่การสร้างเครือข่ายผู้ใช้
R ที่เป็นคนไทยได้อย่างต่อเนื่องไปในอนาคตระยะยาว

ทั้งนี้คณะผู้วิจัยได้ใช้โปรแกรม **R** ในการทำงานวิเคราะห์ทางสถิติ และการเรียนการสอนในหลักสูตร
ต่างๆ จึงจะยังคงทำงานนี้ต่อไปแม้จะสิ้นสุดโครงการวิจัยนี้แล้วก็ตาม โดยมีความคาดหวังว่าจะสามารถผลักดัน
ให้มีการใช้โปรแกรม **R** ในประเทศไทย ทั้งนี้เพราะเรามั่นใจอย่างสูงว่า **R** เป็นซอฟต์แวร์วิเคราะห์ทางสถิติที่มี
ประสิทธิภาพสูง เป็นที่ยอมรับของนักสถิติผู้เชี่ยวชาญทั่วโลก และยังใช้และแจกจ่ายได้ฟรี โดยหวังว่าจะสามารถ
ปลูกฝังเจตคติแก่สังคมไทยในการเคารพกฎหมายทรัพย์สินทางปัญญา และการไม่ใช้ซอฟต์แวร์ที่ได้มาโดยผิด
กฎหมายอีกด้วย

บทคัดย่อ

ปัจจุบันการวิเคราะห์ข้อมูลทางสถิติ ต้องอาศัยโปรแกรมสำเร็จรูปที่สามารถรองรับข้อมูลขนาดใหญ่ และระเบียบวิธีทางสถิติที่ซับซ้อนได้ แต่ซอฟต์แวร์ที่มีประสิทธิภาพดังกล่าวส่วนใหญ่เป็นซอฟต์แวร์ทางการค้าที่มีราคาสูง แต่ในขณะนี้ มีผู้พัฒนาโปรแกรม **R** ขึ้นมา โดยเป็นโปรแกรมที่มีสมรรถนะสูง เป็นที่นิยมในหมู่นักวิชาการและนักสถิติระดับโลก และคุณสมบัติที่โดดเด่นของมันคือเป็นซอฟต์แวร์ที่มีลิขสิทธิ์แบบ GPL คือใช้งาน แจกจ่าย และปรับปรุงพัฒนาได้ฟรี โดยไม่ต้องเสียค่าใช้จ่าย แต่ซอฟต์แวร์นี้ยังเป็นที่รู้จักและมีผู้ใช้เป็นนัยมากในประเทศไทย และยังมีข้อจำกัดที่สำคัญในเรื่องเอกสารประกอบการใช้งาน เนื่องจากตัวโปรแกรมเองเน้นการเป็นภาษาคอมพิวเตอร์ ทำให้ผู้ใช้ทั่วไปเรียนรู้ได้ยาก

วัตถุประสงค์

โครงการนี้จึงมีวัตถุประสงค์เพื่อพัฒนา เผยแพร่ และสนับสนุนให้มีการใช้โปรแกรม **R** ซึ่งเป็นซอฟต์แวร์ทางสถิติที่ไม่ต้องเสียค่าลิขสิทธิ์ ในการวิเคราะห์ผลงานวิจัยทั้งทางด้านวิทยาศาสตร์สุขภาพและระเบียบวิธีทางสถิติ

การดำเนินงาน

การดำเนินงานเพื่อให้บรรลุวัตถุประสงค์ข้างต้น จึงมีสามกลุ่มกิจกรรมคือ 1. การพัฒนาการใช้โปรแกรม **R** ให้ใช้ได้ง่ายขึ้น 2. พัฒนาส่วนติดต่อกับผู้ใช้เป็นภาษาไทย และ 3. การจัดกิจกรรมเผยแพร่การใช้งานโปรแกรม **R** และสิ่งประดิษฐ์ที่สร้างขึ้นในกิจกรรมสองกลุ่มที่กล่าวข้างต้น ในการวิเคราะห์ทางสถิติ การพัฒนาการใช้โปรแกรม **R** ให้ใช้ได้ง่ายขึ้น ประกอบด้วยการสร้างฟังก์ชันเพิ่มเติม โดยเฉพาะในด้านการจัดการข้อมูล และการแสดงผลการทำงานที่เข้าใจได้ง่าย ซึ่งเป็นจุดอ่อนที่สำคัญของ **R** และทำชุดฟังก์ชันเหล่านั้นเป็นแพ็คเกจเพื่อเผยแพร่แก่ผู้ต้องการใช้ และการพัฒนาส่วนติดต่อกับผู้ใช้ให้เป็นแบบรายการเมนูและหน้าต่างกราฟิก เช่นเดียวกับโปรแกรมทางการค้าอื่นๆ ในด้านการพัฒนาส่วนติดต่อกับผู้ใช้เป็นภาษาไทยนั้น ประกอบด้วยการทำความช่วยเหลือแบบ on line และแบบเป็นแฟ้มสำหรับฟังก์ชันและแพ็คเกจที่สร้างขึ้นนั้นเป็นภาษาไทย และการทำรายการเมนูและหน้าต่างกราฟิกต่างๆ ให้เป็นภาษาไทย ในด้านการจัดกิจกรรมเผยแพร่การใช้งานโปรแกรม **R** ได้จัดการประชุมเชิงปฏิบัติการให้แก่ผู้ใช้กลุ่มต่างๆ จัดการเรียนการสอนผ่านทางอินเทอร์เน็ต และการตีพิมพ์หนังสือเผยแพร่แก่ประชาชนและนักวิชาการ และยังคงจัดทำเว็บไซต์เพื่อการเรียนรู้ แลกเปลี่ยนความคิดเห็น และการจัดกิจกรรมต่างๆ ที่เกี่ยวข้องกับ **R**

ผลการดำเนินงาน

ผลการดำเนินงานในกลุ่มกิจกรรมการพัฒนาการใช้โปรแกรม **R** ให้ใช้ได้ง่ายขึ้น ประกอบด้วยการพัฒนาชุดฟังก์ชันด้านการจัดการแฟ้มและข้อมูล ชื่อ ice มีฟังก์ชันที่คณะวิจัยได้สร้างขึ้น 50 ฟังก์ชัน แพ็คเกจ epid ซึ่งเป็นชุดฟังก์ชันพื้นฐานด้านระบาดวิทยา มีฟังก์ชันที่คณะวิจัยได้สร้างขึ้น 9 ฟังก์ชัน ความช่วยเหลือของฟังก์ชันที่สร้างขึ้นเหล่านี้มีทั้งที่เป็นภาษาอังกฤษและภาษาไทย และได้สร้างสภาพแวดล้อม **ICE** ซึ่งเป็นสภาพแวดล้อมที่เป็นรายการเมนูและหน้าต่างกราฟิก สามารถแสดงได้ทั้งภาษาไทยและภาษาอังกฤษ ประกอบด้วยแพ็คเกจหลักคือ ice.main และแพ็คเกจประกอบอื่นๆ ซึ่งเมื่อทำงานภายใต้โปรแกรม **R** จะเรียกเป็นสภาพแวดล้อม Integrated Computing Environment for **R** (**R-ICE**) ในการจัดกิจกรรมเพื่อ

เผยแพร่การใช้งานโปรแกรม **R** นั้น ได้จัดการประชุมเชิงปฏิบัติการวิเคราะห์ทางสถิติด้วยโปรแกรม **R** ขึ้น 13 ครั้ง ให้แก่ผู้สนใจทั้งภายในและภายนอกมหาวิทยาลัยสงขลานครินทร์ รวมถึงการจัดนอกสถานที่ด้วย ได้จัดทำเว็บไซต์ขึ้น 4 เว็บไซต์ มีเว็บไซต์สำหรับผู้สนใจโปรแกรม **R** โดยทั่วไป ซึ่งเป็นที่พบปะสนทนาของกลุ่มผู้ใช้ **R** เว็บไซต์สำหรับผู้สนใจหลักสูตรทางอินเทอร์เน็ต เว็บไซต์สำหรับสภาพแวดล้อม **R-ICE** และเว็บบล็อกเผยแพร่การเขียนฟังก์ชันในโปรแกรม **R** และยังสามารถจัดทำ mirror ของ CRAN เพื่อให้ดาวน์โหลดโปรแกรมได้รวดเร็วยิ่งขึ้น ส่วนการตีพิมพ์หนังสือเพื่อส่งเสริมการใช้โปรแกรม **R** นั้น ได้ทำหนังสือขึ้นสองเล่ม คือ “การวิเคราะห์ทางสถิติด้วย **R-ICE**” ซึ่งมีกลุ่มเป้าหมายเป็นประชาชนทั่วไปที่สนใจการวิเคราะห์ข้อมูลทางสถิติ และ “คู่มือการใช้โปรแกรม **R** การใช้คำสั่งหรือฟังก์ชันพื้นฐาน” ซึ่งมีกลุ่มเป้าหมายเป็นนักวิชาการในสถาบันวิชาการต่างๆ ขณะนี้ทั้งสองเล่มกำลังอยู่ในระหว่างการดำเนินการจัดพิมพ์

วิจารณ์และสรุป

ในการดำเนินโครงการนี้ คณะผู้วิจัยได้เรียนรู้ในด้านเทคนิคต่างๆ มากมายในการเขียนฟังก์ชัน การทำแพ็คเกจ และการแปลภาษาตามวิธีการของโปรแกรม **R** และได้สร้างเครือข่ายกับนักพัฒนาโปรแกรม **R** ในระดับนานาชาติโดยการเข้าร่วมประชุม และนำเสนอผลงานเรื่องสภาพแวดล้อม **R-ICE** ที่พัฒนาขึ้นนี้ในการประชุม UseR! 2006 ที่กรุงเวียนนา ในการเผยแพร่การใช้งานโปรแกรม **R** นั้น การจัดการประชุมเชิงปฏิบัติมีประสิทธิภาพในการเรียนรู้สูงสุด แต่ทำได้จำนวนจำกัด เนื่องจากต้องใช้ทรัพยากรมาก การจัดการเรียนการสอนทางอินเทอร์เน็ตไม่ได้ผลดีเท่าที่ควร เนื่องจากผู้เรียนอาจไม่อยู่ในสภาพที่พร้อมจะเรียน เช่นไม่สามารถติดตั้งโปรแกรมในเครื่องที่ใช้เรียนได้ และหากมีปัญหาเล็กน้อยก็มักจะเลิกเรียนได้ง่าย คิดว่าการทำหนังสือคู่มือ น่าจะเผยแพร่ได้ในวงกว้างและผู้ใช้สามารถทำตามขั้นตอนและตัวอย่างในหนังสือได้ และไม่มีข้อจำกัดเรื่องเวลา และสถานที่ในการเรียนรู้ นอกจากนี้ การสร้างเว็บไซต์เพื่อเป็นที่รวมเสมือนของผู้ใช้ **R** ในประเทศไทยน่าจะนำไปสู่การสร้างเครือข่ายผู้ใช้ **R** ที่เป็นคนไทยได้อย่างต่อเนื่องไปในอนาคตระยะยาว

ทั้งนี้คณะผู้วิจัยได้ใช้โปรแกรม **R** ในการทำงานวิเคราะห์ทางสถิติ และการเรียนการสอนในหลักสูตรต่างๆ จึงจะยังคงทำงานนี้ต่อไปแม้จะสิ้นสุดโครงการวิจัยนี้แล้วก็ตาม โดยมีความคาดหวังว่าจะสามารถผลักดันให้มีการใช้โปรแกรม **R** ในประเทศไทย ทั้งนี้เพราะเรามั่นใจอย่างสูงว่า **R** เป็นซอฟต์แวร์วิเคราะห์ทางสถิติที่มีประสิทธิภาพสูง เป็นที่ยอมรับของนักสถิติผู้เชี่ยวชาญทั่วโลก และยังใช้และแจกจ่ายได้ฟรี โดยหวังว่าจะสามารถปลูกฝังเจตคติแก่สังคมไทยในการเคารพกฎหมายทรัพย์สินทางปัญญา และการไม่ใช้ซอฟต์แวร์ที่ได้มาโดยผิดกฎหมายอีกด้วย

Abstract

Nowadays, statistical analysis requires softwares that are capable of large dataset handling and sophisticated statistical methodologies, while most of them are high cost commercial softwares. **R**, a high capacity open source software for statistical analysis, is recently developed and rapidly accepted by world class statisticians and academic professions. The most distinguishing characteristic of **R** is its GPL (GNU General Public License) term where the software is open for free use, distribution, and modification. However, it is rarely known and used in Thailand. One of the important limitations is its strength as a computer language environment, thus, most of the documentations are available for programmers but not for statisticians and general users.

Objectives:

This project is aiming at development, distribution, and facilitation of using **R** as a free software for statistical analysis of research projects both in the fields of health science and statistical methodologies.

Methods and plans:

To achieve the above objectives, three groups of activities were planned: 1. modification of **R** for ease of use, 2. development of user interfaces in Thai, and 3. promotion of **R** and the products of the two previous activities for statistical analysis. Activities in development of **R** for ease of use included creation of new functions especially in data management and output display which were important weakness of **R**. Functions were bundled in a small set of packages that could be easily distributed to users. Graphical user interfaces (GUI) like those existing in other well known commercial statistical softwares was built. User interfaces in Thai included on-line and off-line help for those functions and packages and drop down menus and dialog windows in Thai. Activities for promotion of **R** in statistical analysis comprised workshops for different targeted users, internet-based courses, and web-sites for learning, discussion, and other activities related to **R**.

Results:

To modify **R** for ease of use, two packages are created. A package emphasizing on file and data management called “ice” comprises 50 newly created functions and “epid” package contains 9 basic epidemiological functions. Those functions have both English and Thai on-line and off-line help files. A GUI environment called “**ICE**” which has both English and Thai displays is also created. It contains one core package, “ice.main”, and a set of associated packages. When it is running under **R**, Integrated Computing Environment for **R** (**R-ICE**) is the name for such an environment. To promote using **R** in statistical analysis, 13 workshops were held for university staff and students and those from other institutes. Four web sites have been set up for **R** users (UseR). One is a meeting place and discussion forum for Thai UseRs, one for internet-based courses, and one for **R-ICE** users. The last site is a web blog for **R** programming. A mirror site of CRAN is set up at a server in the Epidemiology Unit, Prince of Songkla University to facilitate rapid download of **R** software and packages within the university and the country. Two books are on the process of publishing. “Statistical Analysis with **R-ICE**” is targeted at general users

who are interested in statistical analysis. “Manual for Basic **R** Functions” is targeting at academic personnels.

Discussion and conclusion:

We have learned a lot of valuable **R** programming, package building, and internationalization techniques during the project. We have a network with **R** developers and **R** core team in UseR! 2006 conference in Vienna where we had an opportunity to present the **R-ICE** environment. In promoting the use of **R**, there are evidences that a workshop with computer practice is the most successful environment for learning and using **R** but it consumes a lot of resources while only a number of participates can be enrolled at a time. Internet-based courses is not a highly fruitful environment since participants may not be ready all the time of the courses and their problems may not be promptly solved and learning failure is ensured. Books and manuals can be widely distributed to target groups and they can follow the contents and practice examples at their own paces and places. Furthermore, web sites as virtual meeting places might lead to a last long network of Thai UseRs.

The research team is now using **R** in statistical analysis and teaching courses, thus, we still continue working on **R** even the project has come to an end. We hope we could continue promoting use of **R** in Thailand since we are best convinced that **R** is a free and highly efficient statistical software that is accepted by statistician communities worldwide and we also wish we could convince Thai communities an attitude of respecting intellectual property rights and not using any illegal softwares.

เนื้อหางานวิจัย

หลักการและเหตุผล

ความเป็นมาและความสำคัญ

รัฐบาลไทยได้ประกาศนโยบายรับรองทรัพย์สินทางปัญญา โดยเฉพาะอย่างยิ่งซอฟต์แวร์คอมพิวเตอร์ โดยมีประกาศให้หน่วยราชการต่างๆ ใช้ซอฟต์แวร์ที่ถูกกฎหมายเท่านั้น อย่างไรก็ตาม ในทางปฏิบัติ หน่วยราชการต่างๆ ไม่มีงบประมาณในการจัดซื้อซอฟต์แวร์มากเพียงพอ หากจัดซื้อจริงคาดว่าจะค่าใช้จ่ายต่างๆ จะสูงขึ้นอย่างมาก มีความจำเป็นทางยุทธศาสตร์ที่จะต้องแก้ไขปัญหานี้ในระดับประเทศ

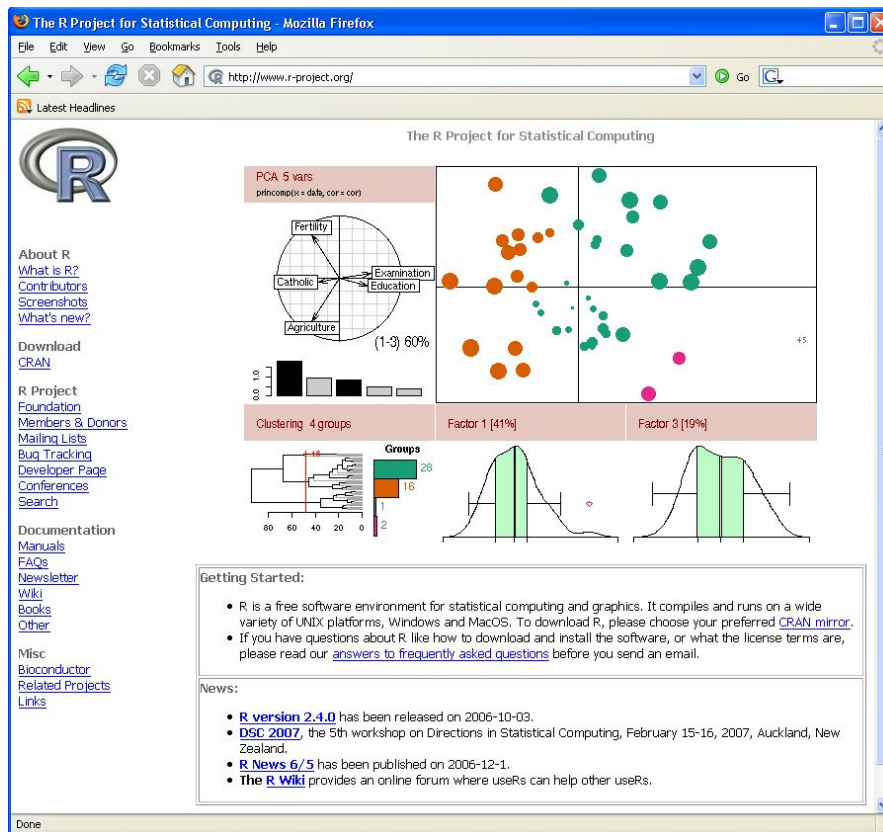
ปัจจุบันการวิเคราะห์ข้อมูลทางสถิติ ต้องอาศัยโปรแกรมสำเร็จรูป (software) ที่สามารถรองรับข้อมูลที่มีขนาดใหญ่ และระเบียบวิธีทางสถิติที่ซับซ้อน อย่างไรก็ตามซอฟต์แวร์ที่มีประสิทธิภาพดังกล่าวถูกผลิตขึ้นเพื่อการค้า และมีราคาแพง ประกอบกับหน่วยงานต่างๆ ไม่มีงบประมาณสำหรับซื้อซอฟต์แวร์เหล่านั้น ทำให้ต้องใช้ซอฟต์แวร์อย่างผิดกฎหมาย ขณะที่ซอฟต์แวร์ทางสถิติ เป็นส่วนสำคัญในการสร้างองค์ความรู้ในสาขาต่างๆ ขึ้นในประเทศไทย การพัฒนาการใช้ซอฟต์แวร์ที่ไม่ใช่เพื่อการค้าสำหรับวิเคราะห์ข้อมูลให้ชำนาญมากขึ้น เป็นสิ่งจำเป็นและในลำดับต่อไปให้แพร่หลายขึ้น เพื่อให้หลุดพ้นจากปัญหาเชิงการค้าให้ได้มากที่สุด

ศักยภาพที่มีอยู่และสามารถพัฒนาเพิ่มเติม

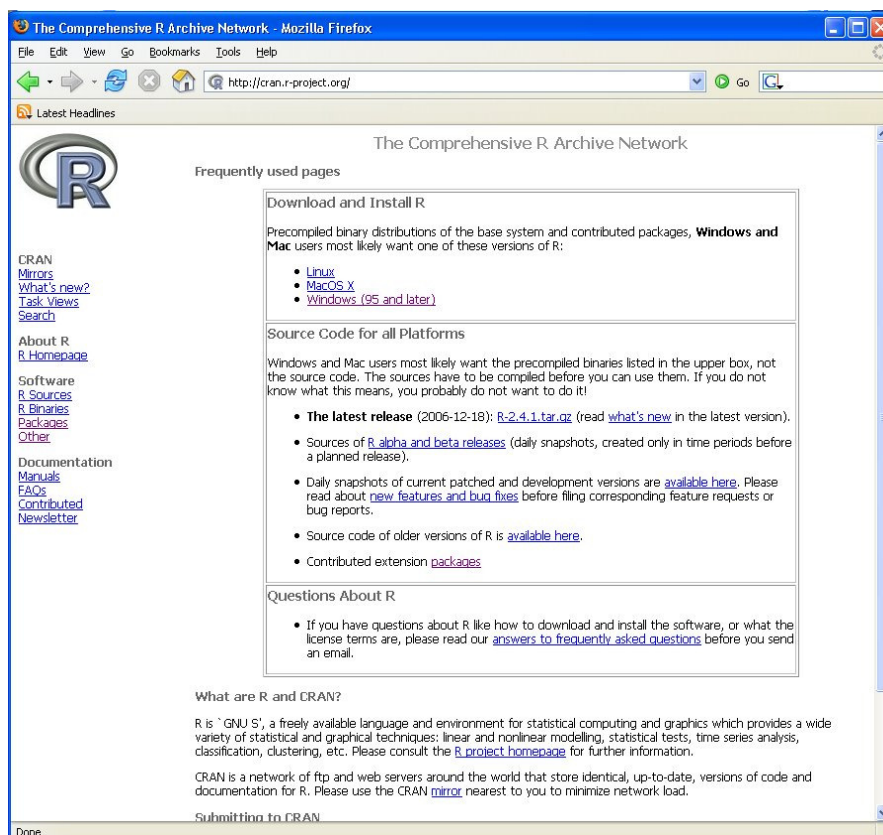
กิจกรรมการพัฒนาการใช้ซอฟต์แวร์ เป็นสิ่งสำคัญควบคู่ไปกับการพัฒนาความรู้ความเข้าใจในการวิเคราะห์ข้อมูล โครงการที่คณะผู้วิจัยนำเสนอนี้จึงเป็นทั้งการลงทุนเพื่อให้ได้ความรู้จากข้อมูล พัฒนาศักยภาพให้มีความสามารถในการวิเคราะห์ข้อมูล และเป็นการทดลองใช้ซอฟต์แวร์ที่นำเสนอให้นักวิจัยใช้ได้ฟรีและไม่ผิดกฎหมาย ที่เราสามารถดาวน์โหลดผ่านอินเทอร์เน็ตได้

โปรแกรมสำเร็จรูป **R** เป็นซอฟต์แวร์แบบ open source ที่นำเสนอให้ฟรี ไม่มีลิขสิทธิ์ โดย The R Development Core Team ซึ่งผู้เขียนโปรแกรมนี้เป็นผู้เขียนโปรแกรม S Plus ที่ได้รับรางวัลโปรแกรมดีเด่นในปี พ.ศ. 2540, 2542, 2544 และ 2545 โปรแกรม **R** มีมูลนิธิ R Software Foundation เป็นผู้สนับสนุน โดยมีเว็บไซต์ที่ <http://www.r-project.org> ดังรูปที่ 1 มีผลงานตีพิมพ์ที่ใช้โปรแกรมนี้ในการวิเคราะห์ข้อมูลเพิ่มมากขึ้นทุกวันในวารสารวิชาการด้าน methodology โปรแกรมนี้จึงเป็นทางเลือกหนึ่งที่สามารถนำมาพัฒนาเพื่อใช้ในการเรียนการสอนได้

เนื่องจากโปรแกรม **R** มีองค์กร Comprehensive R Archive Network (CRAN) ดูแลในการพัฒนาโปรแกรมและแพ็คเกจ โดยมีเว็บไซต์อยู่ที่ <http://cran.r-project.org> ดังรูปที่ 2 โครงการนี้จึงไม่มีวัตถุประสงค์ในการพัฒนาในส่วนของโปรแกรม แต่จะพัฒนา library เพื่อใช้งานเฉพาะด้าน ในที่นี้คือในด้านระบาดวิทยา และสถิติทั่วไป ซึ่งการพัฒนาในส่วนของ library ทำได้โดยไม่มีข้อจำกัดทางลิขสิทธิ์ และหากคำสั่งที่สร้างขึ้นโดยคนไทย คำสั่งใดเป็นคำสั่งที่ยังไม่มีผู้ใดทำไว้ก่อน และมีความโดดเด่นก็สามารถส่งไปรวมไว้ใน library ที่เป็นส่วนกลาง เพื่อแจกจ่ายแก่ผู้ใช้ทั่วโลกได้



รูปที่ 1 หน้าเว็บหลักของ R ที่ <http://www.r-project.org/>



รูปที่ 2 หน้าเว็บดาวน์โหลดโปรแกรม R สำหรับระบบปฏิบัติการต่างๆ จาก CRAN ที่ <http://cran.r-project.org/>

แม้ว่าโปรแกรม **R** จะมีข้อดีดังที่กล่าวมาแล้ว แต่เนื่องจากโปรแกรม **R** เป็นโปรแกรมที่แจกจ่ายให้ใช้งานได้ฟรี และเปิดโอกาสให้นักสถิติและนักโปรแกรมคอมพิวเตอร์ทั้งหลายสามารถพัฒนาทั้งตัวโปรแกรมและแพ็คเกจได้อย่างอิสระ แต่การที่มีอิสระมากเช่นนี้ก็เป็นข้อจำกัดที่ทำให้คำสั่งต่างๆ ในแพ็คเกจที่แตกต่างกันที่พัฒนาจากบุคคลต่างๆ ขาดเอกภาพในด้านวิธีใช้งานและการแสดงผลการวิเคราะห์ การทำงานกับคำสั่ง หนึ่งอาจแตกต่างไปจากคำสั่งอื่นๆ ที่ทำงานใกล้เคียงกัน และผลที่แสดงก็มีรูปแบบต่างกัน ทำให้ผู้ใช้สับสนและเรียกใช้งานคำสั่งผิดพลาดได้ง่าย ทำให้จำเป็นต้องเรียกดูข้อความช่วยเหลือบ่อยๆ และนอกจากนี้ ข้อความช่วยเหลือของคำสั่งจำนวนมากจะเขียนในเชิงลึกทางด้านสถิติขั้นสูงหรือทางด้านเทคนิคคอมพิวเตอร์ ซึ่งเข้าใจได้ยากมาก และผู้ใช้ทั่วไปจะไม่สามารถเข้าใจได้เลย

วัตถุประสงค์ของการวิจัย

โครงการวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนา เผยแพร่ และสนับสนุนให้มีการใช้โปรแกรม **R** ซอฟต์แวร์ทางสถิติที่ไม่ต้องเสียค่าลิขสิทธิ์ ในการวิเคราะห์ผลงานวิจัยทั้งทางด้านวิทยาศาสตร์สุขภาพ และระเบียบวิธีทางสถิติ

หน่วยงานที่รับผิดชอบ

หน่วยระบาดวิทยา คณะแพทยศาสตร์ และภาควิชาคณิตศาสตร์ คณะวิทยาศาสตร์
มหาวิทยาลัยสงขลานครินทร์ อำเภอหาดใหญ่ จังหวัดสงขลา

คณะผู้ร่วมวิจัย

ที่ปรึกษาโครงการ

1. ศ. นพ. วีระศักดิ์ จิงสู่วินวงศ์ หน่วยระบาดวิทยา คณะแพทยศาสตร์
2. อ. อลัน กีเตอร์ หน่วยระบาดวิทยา คณะแพทยศาสตร์

ผู้วิจัยหลัก

1. รศ. นพ. หัซซา ศรีปลั่ง หน่วยระบาดวิทยา คณะแพทยศาสตร์
(สังกัดภาควิชาพยาธิวิทยา คณะแพทยศาสตร์)
2. ดร. นราทิพย์ จันสกุล ภาควิชาคณิตศาสตร์และสถิติ คณะวิทยาศาสตร์

ผู้วิจัยร่วม

1. นายเอ็ดเวิร์ด แมคเนล หน่วยระบาดวิทยา คณะแพทยศาสตร์
2. นส. อภิรดี แซ่ลิ่ม หน่วยระบาดวิทยา คณะแพทยศาสตร์
(ปัจจุบันเป็นอาจารย์ในคณะวิทยาศาสตร์และเทคโนโลยี)
3. ผศ. กัลยา นิตเรืองจรัส ภาควิชาคณิตศาสตร์และสถิติ คณะวิทยาศาสตร์
4. นายกิตติศักดิ์ ชูมาลี หน่วยระบาดวิทยา คณะแพทยศาสตร์
(สังกัดภาควิชาเวชศาสตร์ชุมชน คณะแพทยศาสตร์)

วิธีการดำเนินงาน

คณะผู้วิจัยได้เสนอแผนการดำเนินโครงการเมื่อเริ่มดำเนินโครงการ ประกอบไปด้วย 3 ระยะ คือ

แผนการดำเนินงานวิจัยที่เสนอ

ระยะที่ 1 ระยะเวลาในการเตรียมการ 1 ปี

1. พัฒนาฟังก์ชัน (คำสั่ง) แพคเกจ และสื่อการเรียนการสอนเบื้องต้น ด้วยโปรแกรม **R** สำหรับการใช้งานด้านวิทยาศาสตร์สุขภาพ เพื่อให้ง่ายแก่ผู้ที่ต้องการเรียนรู้
2. ทดลองการใช้โปรแกรม **R** ในระยะเริ่มต้นสำหรับผู้สนใจในประเทศไทย โดยการ
 - 2.1 แปลความช่วยเหลือของคำสั่งในแพคเกจที่มีอยู่แล้วเป็นภาษาไทย เพื่อความสะดวกแก่ผู้ใช้
 - 2.2 เขียนคู่มือการใช้งานและคู่มือคำสั่งเป็นภาษาไทย
3. สร้างกลุ่มผู้เขียนคำสั่งใน **R** และกลุ่มผู้ใช้ **R**
4. เชิญวิทยากรผู้พัฒนาโปรแกรมจากนานาชาติมาจัดอบรมในมหาวิทยาลัยและสถาบันวิจัยในประเทศไทย เพื่อส่งเสริมการใช้และพัฒนาซอฟต์แวร์ให้เหมาะสมกับประเทศไทย นอกจากนี้ความรู้เรื่องการใช้โปรแกรม **R** ในประเทศไทยยังมีน้อย จึงต้องอาศัยผู้เชี่ยวชาญให้คำแนะนำ
5. สร้างความเข้มแข็งให้กับโครงการฝึกอบรมเชิงปฏิบัติการในการใช้โปรแกรม **R** และเปิดเป็นหลักสูตรสอนให้แก่นักศึกษาทั้งระดับปริญญาตรี ปริญญาโท และปริญญาเอก และผู้สนใจ

ระยะที่ 2 ระยะเวลา 1 ปี

1. จัดประชุมเชิงปฏิบัติการในประเทศไทย โดยวิทยากรไทย โดยจัดอบรมในสถาบันต่าง ๆ ในประเทศไทยโดยร่วมมือกับนักวิชาการประจำสถาบันอย่างน้อย 5 แห่ง ให้แก่ผู้เข้าร่วมจากสถาบันการศึกษา และสถาบันต่าง ๆ ในแต่ละท้องถิ่น ๆ ทางหน่วยระดับวิทยา คณะแพทยศาสตร์ จะจัดให้แก่ผู้เข้าร่วมจากหน่วยงานสาธารณสุข แพทย์และพยาบาล ส่วนภาควิชาคณิตศาสตร์ คณะวิทยาศาสตร์ จัดอบรมให้แก่คณาจารย์ทางด้านสถิติ หรือผู้สนใจจากต่างสถาบันที่เกี่ยวข้องกับงานด้านสถิติ
2. ทำการว่าจ้างผู้เชี่ยวชาญเพื่อสร้างเว็บไซต์สำหรับการลงทะเบียนเรียนโปรแกรม **R** ผ่านทางอินเทอร์เน็ต และผู้ใช้สามารถดาวน์โหลดโปรแกรมและแพคเกจได้ฟรี
3. ทดลองให้มีการใช้โปรแกรมนี้นี้กับเจ้าหน้าที่สาธารณสุข และสถาบันอื่น ๆ และทดลองให้มีการสอนผ่านทางอินเทอร์เน็ต
4. สร้างความช่วยเหลือแบบ on-line เป็นภาษาไทย
5. จัดทำคู่มือทั้งภาษาไทยและอังกฤษ ในการใช้งานโปรแกรม **R** ให้แก่ผู้สนใจ
6. หลังจากพัฒนาการเขียนคำสั่งในโปรแกรม **R** ได้ระยะหนึ่ง ผู้วิจัยหลักจะศึกษาดูงานด้านการใช้โปรแกรม **R** ในต่างประเทศ เพื่อศึกษาว่าต่างประเทศพัฒนาอะไรไปบ้าง มีอะไรที่ประเทศไทยจะต้องทำต่อไปในอนาคต และต่างประเทศได้พัฒนาโปรแกรมไปมาอย่างน้อยเพียงใด เพื่อเรียนรู้คำสั่งใหม่ๆ ซึ่งกลุ่มผู้เขียนโปรแกรม **R** จากต่างประเทศมีการพัฒนาโปรแกรมตลอดเวลา จึงจำเป็นที่ผู้เขียนโปรแกรมในประเทศไทยจะต้องตามให้ทัน

ระยะที่ 3 ระยะเวลา 1 ปี

1. จัดทำหนังสือเพื่อเป็นบทเรียนให้แก่นักศึกษาและผู้สนใจ

2. จัดทำบทเรียนใน Internet ให้ดีขึ้น เพื่อให้ผู้ที่สนใจสามารถเข้ามาลงทะเบียนทางอินเทอร์เน็ตโดยไม่ต้องมีปัญหาในเรื่องของลิขสิทธิ์โปรแกรม
3. จัดทำศูนย์ mirror และเว็บบอร์ดเพื่อตอบปัญหาให้แก่ผู้สนใจ
4. จัดทำหนังสือคู่มือการใช้เผยแพร่สู่สาธารณะ ที่หาซื้อได้จากร้านขายหนังสือทั่วไป

อย่างไรก็ตาม ในระหว่างการดำเนินงาน ได้ปรับเปลี่ยนแผนการทำงานไปบ้าง เนื่องจากเหตุผลบางประการ ซึ่งจะได้แจกแจงรายละเอียดต่อไป โดยสรุป แผนการดำเนินงานที่แท้จริงมีดังนี้

แผนการดำเนินงานวิจัยที่ได้ทำจริง

ระยะที่ 1 ระยะเวลาในการเตรียมการ 1 ปี

1. พัฒนาฟังก์ชัน (คำสั่ง) แพลกเจ และสื่อการเรียนการสอนเบื้องต้น ด้วยโปรแกรม **R** สำหรับการใช้งานด้านวิทยาศาสตร์สุขภาพ เพื่อให้ง่ายแก่ผู้ที่ต้องการเรียนรู้
2. ทดลองการใช้โปรแกรม **R** ในระยะเริ่มต้นสำหรับผู้สนใจในประเทศไทย โดยการ
 - 2.1 แปลความช่วยเหลือของคำสั่งในแพลกเจที่มีอยู่แล้วเป็นภาษาไทย เพื่อความสะดวกแก่ผู้ใช้
 - 2.2 เขียนคู่มือการใช้งานและคู่มือคำสั่งเป็นภาษาไทย
 (ส่วนนี้ไม่ได้ทำ เนื่องจากพบว่าฟังก์ชันที่ได้พัฒนาขึ้นในข้อ 1 สามารถครอบคลุมการทำงานส่วนใหญ่ในการวิเคราะห์ข้อมูลทางสถิติได้ โดยเกือบไม่ต้องใช้คำสั่งพื้นฐานของ **R** จึงได้ตัดสินใจไม่แปลความช่วยเหลือคำสั่งพื้นฐานของ **R** ซึ่งมีจำนวนมาก และส่วนใหญ่ (เกือบทั้งหมด) ผู้ใช้โดยทั่วไปไม่ได้ใช้)
3. สร้างกลุ่มผู้เขียนคำสั่งใน **R** และกลุ่มผู้ใช้ **R**
 (การสร้างกลุ่มผู้เขียนคำสั่ง **R** ไม่สามารถทำได้ง่ายนัก เนื่องจากผู้ที่เขียนคำสั่งได้ ต้องมีประสบการณ์ในการใช้งานโปรแกรม **R** มานานพอสมควร จนมีประสบการณ์มากพอ และมีโจทย์ที่จะนำไปสร้างเป็นฟังก์ชันและแพลกเจได้ ดังนั้นกลุ่มผู้ใช้ **R** จึงยังไม่ได้เกิดขึ้น แต่เมื่อสิ้นสุดโครงการ พบว่ามีผู้ใช้งานหนึ่งที่มีศักยภาพถึงขั้นที่จะเขียนได้)
4. เชิญวิทยากรผู้พัฒนาโปรแกรมจากนานาชาติมาจัดอบรมในมหาวิทยาลัยและสถาบันวิจัยในประเทศไทย เพื่อส่งเสริมการใช้และพัฒนาซอฟต์แวร์ให้เหมาะสมกับประเทศไทย นอกจากนี้ความรู้เรื่องการใช้โปรแกรม **R** ในประเทศไทยยังมีน้อย จึงต้องอาศัยผู้เชี่ยวชาญให้คำแนะนำ (เนื่องจากเมื่อได้เรียนรู้การเขียนฟังก์ชันและแพลกเจไประยะหนึ่งแล้ว พบว่าส่วนใหญ่สามารถเรียนรู้เองได้ จึงเปลี่ยนการตัดสินใจไม่เชิญผู้เชี่ยวชาญจากต่างประเทศ แต่คิดว่าการไปประชุม *UseR conference* น่าจะทำให้ได้ความรู้และรู้จักผู้เชี่ยวชาญจำนวนมาก ซึ่งสามารถปรึกษาทาง email ได้เมื่อติดขัด จะเป็นการดีกว่า)
5. สร้างความเข้มแข็งให้กับโครงการฝึกอบรมเชิงปฏิบัติการในการใช้โปรแกรม **R** และเปิดเป็นหลักสูตรสอนให้แก่นักศึกษาทั้งระดับปริญญาตรี ปริญญาโท และปริญญาเอก และผู้สนใจ

ระยะที่ 2 ระยะเวลา 1 ปี

1. จัดประชุมเชิงปฏิบัติการในประเทศไทย โดยวิทยากรไทย โดยจัดอบรมในสถาบันต่าง ๆ ในประเทศไทยโดยร่วมมือกับนักวิชาการประจำสถาบันอย่างน้อย 5 แห่ง ให้แก่ผู้เข้าร่วมจากสถาบันการศึกษา และสถาบันต่าง ๆ ในแต่ละท้องถิ่น ๆ ทางหน่วยระดับวิทยา คณะ

แพทยศาสตร์ จะจัดให้แก่ผู้เข้าอบรมจากหน่วยงานสาธารณสุข แพทย์และพยาบาล ส่วน
ภาควิชาคณิตศาสตร์ คณะวิทยาศาสตร์ จัดอบรมให้แก่อาจารย์ทางด้านสถิติ หรือผู้สนใจจาก
ต่างสถาบันที่เกี่ยวข้องกับงานด้านสถิติ

2. ทำการว่าจ้างผู้เชี่ยวชาญเพื่อสร้างเว็บไซต์สำหรับการลงทะเบียนเรียนโปรแกรม **R** ผ่านทาง
อินเทอร์เน็ต และผู้ใช้งานสามารถดาวน์โหลดโปรแกรมและแพ็คเกจได้ฟรี
(จากการศึกษาโปรแกรมสร้างเว็บไซต์ต่างๆ พบว่ามีโปรแกรมสำเร็จรูปแบบ *open source*
จำนวนหนึ่ง ทั้งทางด้านการจัดทำเว็บไซต์ เว็บบอร์ด และจัดการเรียนการสอนทางไกล จึงเลือกใช้
โปรแกรมเหล่านั้นมาใช้แทนการว่าจ้าง และใช้ลูกจ้างโครงการมาช่วยจัดการดูแลเว็บไซต์แทน)
3. ทดลองให้มีการใช้โปรแกรมนี้กับเจ้าหน้าที่สาธารณสุข และสถาบันอื่น ๆ และทดลองให้มีการ
สอนผ่านทางอินเทอร์เน็ต
4. สร้างความช่วยเหลือแบบ on-line เป็นภาษาไทย
5. จัดทำคู่มือทั้งภาษาไทยและอังกฤษ ในการใช้งานโปรแกรม **R** ให้แก่ผู้สนใจ
6. หลังจากพัฒนาการเขียนคำสั่งในโปรแกรม **R** ได้ระยะหนึ่ง ผู้วิจัยหลักจะศึกษาดูงานด้านการใช้
โปรแกรม **R** ในต่างประเทศ เพื่อศึกษาว่าทางต่างประเทศพัฒนาอะไรไปบ้าง มีอะไรที่ประเทศ
ไทยจะต้องทำต่อไปในอนาคต และต่างประเทศได้พัฒนาโปรแกรมไปมาน้อยเพียงใด เพื่อเรียนรู้
คำสั่งใหม่ๆ ซึ่งกลุ่มผู้เขียนโปรแกรม **R** จากต่างประเทศมีการพัฒนาโปรแกรมตลอดเวลา จึง
จำเป็นที่ผู้เขียนโปรแกรมในประเทศไทยจะต้องตามให้ทัน

ระยะที่ 3 ระยะเวลา 1 ปี

1. จัดทำหนังสือเพื่อเป็นบทเรียนให้แก่นักศึกษาและผู้สนใจ
2. จัดทำบทเรียนในอินเทอร์เน็ตให้ดีขึ้น เพื่อให้ผู้ที่สนใจสามารถเข้ามาลงทะเบียนทางอินเทอร์เน็ต
ได้โดยไม่ต้องมีปัญหาในเรื่องของลิขสิทธิ์โปรแกรม
3. จัดทำศูนย์ mirror และเว็บบอร์ดเพื่อตอบปัญหาให้แก่ผู้สนใจ
4. จัดทำหนังสือคู่มือการใช้เผยแพร่สู่สาธารณะ ที่หาซื้อได้จากร้านขายหนังสือทั่วไป

สรุปได้เป็นสามกลุ่มกิจกรรมหลักๆ คือ 1. การพัฒนาการใช้โปรแกรม **R** ให้ใช้ได้ง่ายขึ้น 2. พัฒนา
ส่วนติดต่อกับผู้ใช้เป็นภาษาไทย และ 3. การจัดกิจกรรมเผยแพร่การใช้งานโปรแกรม **R** และสิ่งประดิษฐ์ที่สร้าง
ขึ้นในกิจกรรมสองกลุ่มที่กล่าวข้างต้น ในการวิเคราะห์ทางสถิติ

ผลการดำเนินงาน

ผลการดำเนินงานตามโครงการ จะแยกบรรยายเป็นหัวข้อ เพื่อให้สอดคล้องกับรายการที่เขียนใน
แผนการดำเนินงานดังนี้

1. การพัฒนาฟังก์ชันในการจัดการและวิเคราะห์ข้อมูลทางสถิติ
2. การทำความช่วยเหลือการใช้งานเป็นภาษาอังกฤษและภาษาไทย
3. การทดสอบฟังก์ชัน
4. การพัฒนารายการเมนูกราฟิก
5. การทำแพ็คเกจเพื่อเผยแพร่ผลงานแก่สาธารณะ
6. การหาสภาพแวดล้อมเสริมในการวิเคราะห์ทางสถิติ

7. การสอนในหลักสูตรระดับมหาวิทยาลัย และหลักสูตรสถิติ
8. การจัดทำคู่มือการใช้งานโปรแกรม R
9. การจัดประชุมเชิงปฏิบัติการ
10. การสร้างกลุ่มผู้ใช้ R
11. การจัดทำเว็บไซต์
12. การจัดการเรียนการสอนทางอินเทอร์เน็ต
13. การอบรมและเข้าร่วมประชุมในต่างประเทศ
14. การจัดทำ mirror site ของ CRAN
15. การจัดทำหนังสือเพื่อเผยแพร่แก่ประชาชน

ซึ่งมีรายละเอียดดังต่อไปนี้

1. การพัฒนาฟังก์ชันในการจัดการและวิเคราะห์ข้อมูลทางสถิติ

เมื่อเริ่มใช้ R ใหม่ๆ ผู้ใช้ระดับนักสถิติหรือผู้ใช้ที่ไม่มีพื้นฐานในการเขียนโปรแกรมคอมพิวเตอร์มักพบปัญหาว่าโปรแกรม R นั้นใช้งานยากกว่าโปรแกรมวิเคราะห์ข้อมูลทางสถิติตัวอื่นๆ ทั้งนี้เนื่องจากโปรแกรม R เป็นโปรแกรมภาษาคอมพิวเตอร์ ซึ่งออกแบบมาให้มีความสามารถทางด้านการคำนวณทางสถิติ และด้านกราฟิกอย่างสูง ฟังก์ชันใน R ส่วนใหญ่จึงเป็นฟังก์ชันที่เน้นทางการเขียนโปรแกรมคอมพิวเตอร์ แต่มีฟังก์ชันที่ทำงานติดต่อกับผู้ใช้ระดับนักสถิติหรือผู้ใช้ที่ไม่มีพื้นฐานในการเขียนคอมพิวเตอร์โดยทั่วไป ฟังก์ชันส่วนที่ขาดไปมากคือส่วนที่เกี่ยวข้องกับการจัดการ เปลี่ยนค่าข้อมูล อย่างง่ายๆ ฟังก์ชันที่มีอยู่ค่อนข้างซับซ้อนไปในเชิงการเขียนโปรแกรมคอมพิวเตอร์ แม้จะมีเพิ่มความช่วยเหลือที่เรียกดูได้ทันที แต่คำอธิบายฟังก์ชันหลายฟังก์ชันนั้นอ่านแล้วไม่สามารถทำความเข้าใจได้ เพราะเขียนโดยนักเขียนโปรแกรมและเขียนให้นักเขียนโปรแกรมใช้ จึงเป็นที่น่าเสียดายที่ทำให้ผู้เริ่มใช้งานหนึ่งเลิกใช้โปรแกรม R ไป

เนื่องจากจุดนี้เป็นจุดอ่อนที่สำคัญที่ทำให้ผู้ใช้ทั่วไปไม่ใช้โปรแกรม R คณะวิจัยจึงตั้งใจจะแก้ไขปรับปรุงให้มีฟังก์ชันในการจัดการข้อมูลที่ใช้ได้ง่าย มีการแสดงผลของฟังก์ชันที่อ่านเข้าใจง่ายสำหรับนักสถิติ และผู้ใช้ไม่จำเป็นต้องเรียนรู้ด้านการเขียนโปรแกรมระดับลึกก็สามารถใช้โปรแกรมในการวิเคราะห์ทางสถิติได้

เมื่อตั้งวัตถุประสงค์ไว้เช่นนี้แล้ว ในเบื้องต้นจึงเลือกที่จะสร้างฟังก์ชันเพิ่มเติมให้มีชื่อและการแสดงผลคล้ายกับคำสั่งในโปรแกรม Stata ซึ่งคณะนักวิจัยคุ้นเคยอยู่ แต่เป็นโปรแกรมที่มีลิขสิทธิ์ทางการค้า ไม่สามารถแจกจ่ายให้แก่บุคคลทั่วไปได้ฟรี คณะผู้วิจัยหลีกเลี่ยงที่จะทำฟังก์ชันให้เหมือน Stata เกินไป เพียงแต่ทำให้คล้ายกันเท่านั้น เพื่อให้ผู้ใช้โดยเฉพาะนักศึกษาในหลักสูตรระดับมหาวิทยาลัยที่เคยใช้ Stata อยู่เปลี่ยนมาใช้โปรแกรม R ได้ง่าย

คณะผู้วิจัยไม่ใช้โปรแกรม SPSS เป็นต้นแบบเนื่องจากผู้ใช้ทั่วไปใช้โปรแกรม SPSS อยู่แล้ว ดังนั้นถ้าใช้ SPSS เป็นต้นแบบ ผู้ใช้จะรู้สึกว่าโปรแกรม R สู้ SPSS ไม่ได้ และจะหันกลับไปใช้ SPSS ซึ่งหาซื้ออย่างผิดกฎหมายลิขสิทธิ์ได้ง่ายตามตลาดทั่วไป นั่นคือโปรแกรม R จะต้องมีความเอกลักษณ์ของตัวเอง ใช้งานได้ง่ายระดับหนึ่ง อย่างน้อยคล้ายๆ กับ Stata ซึ่งเป็นโปรแกรมที่ต้องเขียนบรรทัดคำสั่งเหมือนกัน และในที่สุด หากจะสามารถพัฒนาโปรแกรมให้มีส่วนติดต่อกับผู้ใช้แบบกราฟิก เช่น รายการเมนู และหน้าต่างโต้ตอบกับผู้ใช้ได้ ก็จะเป็นการดีที่สุด

จึงได้วางกลุ่มฟังก์ชันที่ต้องพัฒนาเพิ่มไว้ดังนี้

- การบริหารจัดการข้อมูล
- การสรุปข้อมูล
- และการวิเคราะห์ทางสถิติพื้นฐาน

แต่ในที่สุดแล้วได้พัฒนากลุ่มฟังก์ชันเพิ่มเติมอีกหลายอย่าง ดังที่จะกล่าวในเรื่องแพ็คเกจ EpiCan ต่อไป

ด้วยความคิดดังกล่าว จึงได้เริ่มทำงานโดยทำรายการฟังก์ชันของ Stata ที่ต้องการให้มีอยู่ใน R ด้วย ซึ่งมีรายการฟังก์ชันมากมาย เช่น replace recode และอื่นๆ แต่พบว่าบางชื่อมีฟังก์ชันชื่อเดียวกันอยู่แล้วใน R แต่ทำงานต่างกัน หรือคล้ายกันแต่ให้ผลลัพธ์ที่ต่างกัน ในกรณีเช่นนี้จึงต้องเปลี่ยนชื่อฟังก์ชันบางตัวที่มีชื่อซ้ำกัน แต่ทำงานต่างกัน เพื่อป้องกันการสับสน

เมื่อได้รายการฟังก์ชันแล้ว จึงเริ่มสร้างฟังก์ชันเหล่านั้น โดยส่วนใหญ่แล้ว หัวหน้าโครงการ (รศ. นพ. หัซซา ศรีปลั่ง) จะเป็นคนเขียนฟังก์ชันเอง เนื่องจากการเขียนฟังก์ชันจำเป็นต้องอาศัยประสบการณ์ในการเขียนหลายปี และเคยเรียนการเขียนฟังก์ชันด้วยภาษาคอมพิวเตอร์ตามมาตรฐานมาก่อน เช่น ภาษา Pascal หรือภาษา C มิฉะนั้นจะเขียนออกมาอย่างไม่มีโครงสร้าง สับสน วกวน เยิ่นเย้อ และทำให้แก้ไขในภายหลังได้ยาก แต่เนื่องจากหัวหน้าโครงการฯ มีได้ศึกษาและมีประสบการณ์ในการเขียนโปรแกรมคอมพิวเตอร์ภาษามาตรฐานมาหลายปี จึงพอจะทำงานนี้ได้รวดเร็วกว่าผู้ร่วมวิจัยคนอื่น นอกจากนี้ Edward McNeil ก็ยังช่วยเขียนบางฟังก์ชันให้ด้วย และ นส. อภิรดี แซ่ลิ่มทำหน้าที่ทดสอบฟังก์ชันที่เขียนขึ้นเหล่านั้นเป็นหลัก

ในส่วนของผู้ร่วมวิจัยทางภาควิชาคณิตศาสตร์จะให้นเน้นด้านการจัดอบรมการใช้งานโปรแกรม R พื้นฐาน ซึ่ง ดร. นราทิพย์ จันสกุล ได้ทำวิทยานิพนธ์คุณสมบัติโดยใช้โปรแกรมนี้มาแล้ว จึงเป็นเรื่องไม่ยากที่จะทำงานนี้ โดยจะต้องเขียนหลักสูตรและเอกสารประกอบการอบรมให้เข้าใจและใช้ได้ง่าย ซึ่งจะได้กล่าวในส่วนนี้ในหัวข้อต่อไป

การเขียนฟังก์ชันใน R

ฟังก์ชันหรือคำสั่งใน R คือวัตถุชนิดหนึ่งที่ใช้สำหรับการทำงานต่างๆ ในการออกคำสั่งให้ฟังก์ชันทำงาน ทำได้โดยการเขียนชื่อฟังก์ชันนั้น ตามด้วยวงเล็บ ซึ่งอาจจะมีอาร์กิวเมนต์อยู่ภายในหรือไม่มีก็ได้

การเขียนฟังก์ชันใน R นั้นทำได้โดยเริ่มต้นจากการระบุชื่อฟังก์ชันดังนี้

```
myfun <- function () {}
```

ซึ่งขณะนี้จะได้ฟังก์ชัน myfun ที่ไม่ได้ทำงานอะไร และเมื่อกำหนดเช่นนี้แล้ว ก็สามารถเรียกใช้งานได้โดยพิมพ์ myfun() ที่ R console ซึ่งจะได้ผลลัพธ์เป็น NULL คือไม่มีอะไรเกิดขึ้น ดังนี้

```
> myfun()
NULL
```

เมื่อต้องการให้ฟังก์ชันทำงาน ให้ใส่คำสั่งการทำงานของฟังก์ชันไว้ภายในวงเล็บปีกกา ดังนี้

```
myfun <- function ()
{
  cat("This is my first function.\n")
}
```

ตอนนี้ฟังก์ชัน `myfun()` จะทำงานพิมพ์ประโยคว่า `This is my first function.` บนหน้าจอ **R console** ข้อความ `\n` เป็นรหัสบอกให้ขึ้นบรรทัดใหม่ เมื่อกำหนดฟังก์ชันเช่นนี้แล้ว เมื่อเรียกใช้ฟังก์ชันบนหน้าจอ **R console** จะได้ผลดังนี้

```
> myfun()
This is my first function.
```

การเขียนชุดของบรรทัดคำสั่งเพื่อสร้างฟังก์ชันเช่นนี้ ควรเขียนเป็นแฟ้มสคริปต์ดีกว่าการเขียนโดยตรงบน **R console** ในแฟ้มสคริปต์เดียวควรเขียนฟังก์ชันเดียวเท่านั้น หรือชุดของฟังก์ชันที่ทำงานร่วมกันหรือทำงานคล้ายกันในเงื่อนไขที่แตกต่างกัน และตั้งชื่อแฟ้มสคริปต์นั้นเป็นชื่อของฟังก์ชันนั้นๆ การเรียกฟังก์ชันนั้นเข้ามาในหน่วยความจำของ **R** ทำได้โดยการใช้คำสั่ง `source` แล้วใส่ชื่อแฟ้มที่เก็บฟังก์ชันนั้นในวงเล็บเป็นอาร์กิวเมนต์ ดังนี้

```
> source("myfun.R")
> myfun()
This is my first function.
```

การส่งผ่านอาร์กิวเมนต์ใน **R** นั้น ทำโดยส่งผ่านอาร์กิวเมนต์ในวงเล็บของฟังก์ชัน ซึ่งอาร์กิวเมนต์อาจมีหลายตัวก็ได้ โดยแยกกันด้วยเครื่องหมายจุลภาค โดยลำดับของอาร์กิวเมนต์มีความสำคัญ คือจะระบุลำดับของอาร์กิวเมนต์เรียงกันไปตามลำดับ แต่หากต้องการสลับที่อาร์กิวเมนต์ก็ทำได้โดยระบุชื่อของอาร์กิวเมนต์นั้นๆ ตามด้วยเครื่องหมาย = เช่น `x = "This is my text."` โดยที่ `x` คือชื่อของอาร์กิวเมนต์ที่ระบุไว้ในขั้นตอนการเขียนฟังก์ชัน ดังตัวอย่างการระบุการส่งผ่านอาร์กิวเมนต์ `x` ให้กับฟังก์ชัน `myfun()` ดังนี้

```
myfun <- function (x) {
  cat(x, "\n")
}
```

จากนั้นเมื่อเรียก `source("myfun.R")` ใหม่อีกครั้ง `myfun` เดิมจะถูกแทนที่ด้วยฟังก์ชันที่แก้ไขใหม่แล้วนั้น จากนั้นก็จะสามารถส่งผ่านข้อความใน `myfun` ได้ดังนี้

```
> source("myfun.R")
> myfun("Here it is.")
Here it is.
>
```

ฟังก์ชันใน **R** สามารถรับอาร์กิวเมนต์ชนิดต่างๆ กันได้ โดยไม่จำกัด และไม่ต้องระบุชนิดเหมือนภาษาโปรแกรมอื่นบางภาษา ดังตัวอย่างการส่งผ่านอาร์กิวเมนต์ชนิดอื่นๆ ให้กับ `myfun` ต่อไปนี้

```
> myfun(2*5)
10
> x <- c(1,2,3,4,5,6)
> myfun(x)
1 2 3 4 5 6
```

เราสามารถระบุการทำงานของฟังก์ชันให้แตกต่างกันตามคลาสของอาร์กิวเมนต์ตัวแรกที่ส่งผ่านเข้าไปได้ โดยใช้ฟังก์ชัน `UseMethod` ประกอบ ดังตัวอย่าง

```
myfun <- function (x)
```

```

UseMethod("myfun")

myfun.character <- function (x)
{
  cat("The STRING is:", x, "\n")
}

myfun.numeric <- function (x)
{
  cat("The NUMBER is:", x, "\n")
}

myfun.default <- function (x)
{
  cat("The OBJECT is:", x, "\n")
}

```

ในที่นี้หากส่งผ่านอาร์กิวเมนต์ชนิด `character` ให้กับฟังก์ชัน `myfun` ฟังก์ชัน `myfun.character` จะถูกเรียกใช้ หากส่งผ่านอาร์กิวเมนต์ชนิด `numeric` ให้ ฟังก์ชัน `myfun.numeric` จะถูกเรียกใช้ หากส่งผ่านอาร์กิวเมนต์ชนิดอื่นนอกเหนือจากนี้ ฟังก์ชัน `myfun.default` จะถูกเรียกใช้ เมื่อเรียก `source("myfun.R")` ใหม่อีกครั้ง `myfun` จะทำงานต่างๆ กันตามแต่ละคลาสของอาร์กิวเมนต์ที่ส่งผ่านเข้าไปดังนี้

```

> myfun("Here it is.")
The STRING is: Here it is.
> myfun(3*4)
The NUMBER is: 12

```

เราสามารถระบุชื่อฟังก์ชันหลายชื่อให้ทำงานอย่างเดียวกันได้ สิ่งนี้มีประโยชน์ในการเรียกฟังก์ชันด้วยชื่อย่อ เพื่อให้พิมพ์คำสั่งเพียงสั้นๆ แต่วิธีนี้ไม่เป็นวิธีที่ **R** นิยมใช้กัน แต่เป็นที่นิยมใช้ในโปรแกรม **Stata** การทำงานเช่นนี้ทำได้โดยผ่านการเรียกใช้คำสั่ง `UseMethod` ดังนี้

```

myfun <- function (x)
UseMethod("myfun")

myf <- function (x)
UseMethod("myfun")

my <- function (x)
UseMethod("myfun")

myfun.character <- function (x)
{
  cat("The STRING is:", x, "\n")
}
...

```

นั่นคือเรากำหนดให้ชื่อ `myfun`, `myf` และ `my` เป็นฟังก์ชันที่ไปเรียกชุดวิธีการทำงานเดียวกัน คือ `myfun.character`, `myfun.numeric` และ `myfun.default` ดังนั้นเมื่อเราเรียก `myfun` หรือ `myf` หรือ `my` ก็จะทำงานเช่นเดียวกันทั้งหมด

การส่งค่ากลับให้กับฟังก์ชันใดๆ ทำโดยการใช้คำสั่ง `return` ดังตัวอย่าง

```

myfun <- function (x)
UseMethod("myfun")

myfun.character <- function (x)

```

```

{
  result <- paste("The STRING is:", x, "\n")
  cat(result)
  return(result)
}

myfun.numeric <- function (x)
{
  result <- paste("The NUMBER is:", x, "\n")
  cat(result)
  return(result)
}

myfun.default <- function (x)
{
  result <- paste("The OBJECT is:", x, "\n")
  cat(result)
  return(result)
}

```

และเมื่อเรียกใช้ฟังก์ชันนี้ ก็จะสามารถส่งค่าให้กับวัตถุได้ดังนี้

```

> x <- myfun("Here it is.")
The STRING is: Here it is.
> x
The STRING is: Here it is.
> y <- myfun(3*4)
The NUMBER is: 12
> y
The NUMBER is: 12

```

นอกจากนี้ยังมีวิธีการเรียกใช้ข้อมูลและชื่อตัวแปรในกรอบข้อมูล นำมาใช้ทำงานภายในฟังก์ชัน ซึ่งเป็นสิ่งที่ผู้วิจัยได้ศึกษาเรียนรู้ด้วยตัวเอง ซึ่งการเรียกตัวแปรในกรอบข้อมูลที่ค้นพบนี้ สามารถทำให้เรียกชื่อตัวแปร หรือเรียกเป็นเลขลำดับของตัวแปรมาใช้งานได้ และยังมีวิธีการทำงานอีกมากมายซึ่งเป็นรายละเอียดในทางเทคนิคและจะไม่นำมากล่าวในที่นี้ ทั้งนี้ที่กล่าวมาข้างต้นเพียงเพื่อให้เห็นพอดเป็นตัวอย่างการเขียนฟังก์ชันเท่านั้น

ตัวอย่างการเขียนฟังก์ชัน

ต่อไปนี้เป็นตัวอย่างของฟังก์ชันใน ice ที่เขียนขึ้นซึ่งมีจำนวนมากมาย จะยกตัวอย่างเพียงสองฟังก์ชันให้เห็นเท่านั้นดังนี้

ฟังก์ชัน `list.rep`

ใช้ตรวจสอบหมายเลข record number ที่ซ้ำกันในกรอบข้อมูลหนึ่งๆ

```

### List replicated records in a vector
list.rep <- function (data, x, n=2)
UseMethod ("list.rep")

list.rep.default <- function (x, n=2)
{
  rep <- table(x)[table(x)>=n]
  if (length(rep)==0)
    cat("No replication found.\n")
  else {
    list.rep <- data.frame(rep)
    return(list.rep)
  }
}

```

```
list.rep.data.frame <- function (data, x, n=2)
{
  nl <- as.list(1:ncol(data))
  names(nl) <- names(data)
  x <- eval(substitute(x),envir=data)
  rep <- table(x)[table(x)>=n]
  if (length(rep)==0)
    cat("No replication found.\n")
  else {
    list.rep <- data.frame(rep)
    return(list.rep)
  }
}
```

ฟังก์ชัน summarize

ใช้สำหรับสรุปวัตถุนิตต่างๆ ในรูปแบบอย่างย่อ โดยแสดงผลคล้ายกับคำสั่ง summarize ของโปรแกรม Stata

```
### Summarize variables in a data set
summarize <- function (object, subset, select, group, pctl =
  get(".ec.pctl",.GlobalEnv), frame=get(".ec.frame",.GlobalEnv))
UseMethod ("summarize")

summ <- function (object, subset, select, group, pctl=get(".ec.pctl",.GlobalEnv),
  frame=get(".ec.frame",.GlobalEnv))
UseMethod ("summarize")

summarize.default <- function (object, subset=NULL, group=NULL,
  pctl=get(".ec.pctl",.GlobalEnv), frame=get(".ec.frame",.GlobalEnv))
{
  if (is.list(object)) {
    vecx <- deparse(substitute(object))
    vec.str <- strsplit(vecx,"")
    vec.num <- length(vec.str[[1]])
    xx <- vec.str[[1]]
    if (vec.num>1) {
      xx <- substring(xx,2)
      xx[1] <- substring(xx[1],5)
      xx[vec.num] <- strsplit(xx[vec.num],"")
    }
    if (!missing(subset) || !missing(group) || !missing(pctl))
      cat("subset, group, and pctl not allowed\n")
    else {
      aa <- NULL
      for (i in 1:length(object)) {
        if (is.factor(object[i][[1]]))
          aa <- rbind(aa,s.factor(object[i][[1]]))
        else if (is.numeric(object[i][[1]]))
          aa <- rbind(aa,summarize.continuous(object[i][[1]]))
      }
      row.names(aa) <- xx
      if (frame) print.frame(aa,type="data")
      else print.noquote(aa)
    }
  }
  if (!missing(subset)) objectx <- subset(object,subset)
  else objectx <- object
  if (typeof(object)=="double" || typeof(object)=="integer") {
    if (!missing(group)) {
      cat(deparse(substitute(object)))
      cat(" stratified by", deparse(substitute(group)), "\n")
      aa <- NULL
      for (i in 1:length(table(group)))
        aa <- rbind(aa,summarize.continuous(objectx[group==names(table(
          group)[i])],pctl))
      row.names(aa) <- names(table(group))
    }
    else {
      aa <- summarize.continuous(objectx,pctl)
      row.names(aa) <- deparse(substitute(object))
    }
  }
}
```

```

        if (frame) print.frame(aa,type="character")
        else print.noquote(aa)
    }
}

summarize.POSIXt <- function (object, frame=get(".ec.frame",.GlobalEnv))
{
  a <- rep("",5)
  dim(a) <- c(1,5)
  xx <- summary(object)
  a[1, ] <- c(length(na.omit(as.character(object))),
    length(as.character(object))-length(na.omit(as.character(object))),
    format.Date(xx[1]),
    format.Date(xx[3]),
    format.Date(xx[6]))
  colnames(a) <- c("Obs.", "NA", "Min.", "Median", "Max.")
  rownames(a) <- " "
  print.noquote(a)
}

summarize.Date <- function (object, frame=get(".ec.frame",.GlobalEnv))
{
  a <- rep("",5)
  dim(a) <- c(1,5)
  xx <- summary(object)
  a[1, ] <- c(length(na.omit(as.character(object))),
    length(as.character(object))-length(na.omit(as.character(object))),
    format.Date(xx[1]),
    format.Date(xx[3]),
    format.Date(xx[6]))
  colnames(a) <- c("Obs.", "NA", "Min.", "Median", "Max.")
  rownames(a) <- " "
  print.noquote(a)
}

summarize.Time <- function (object, frame=get(".ec.frame",.GlobalEnv))
{
  a <- rep("",5)
  dim(a) <- c(1,5)
  xx <- summary(object)
  a[1, ] <- c(length(na.omit(as.character(object))),
    length(as.character(object))-length(na.omit(as.character(object))),
    format.Date(xx[1]),
    format.Date(xx[3]),
    format.Date(xx[6]))
  colnames(a) <- c("Obs.", "NA", "Min.", "Median", "Max.")
  rownames(a) <- " "
  print.noquote(a)
}

summarize.continuous <- function (object, pctl=get(".ec.pctl",.GlobalEnv),
  frame=get(".ec.frame",.GlobalEnv))
{
  a <- rep("",7)
  dim(a) <- c(1,7)
  a[1, ] <- round(c(length(na.omit(object)),
    length(object)-length(na.omit(object)),
    mean(na.omit(object)),
    ifelse(is.na(mean(na.omit(object))), NA, sd(na.omit(object))),
    min(na.omit(object)),
    quantile(na.omit(object),0.5),
    max(na.omit(object)),2)
  colnames(a) <- c("Obs.", "NA", "Mean", "S.D.", "Min.", "Median", "Max.")
  if (pctl) {
    a <- rep("",16)
    dim(a) <- c(1,16)
    a[1, ] <- round(c(length(na.omit(object)),
      length(object)-length(na.omit(object)),
      mean(object),
      sd(object,na.rm=TRUE),
      min(object), quantile(object,
        probs=c(1,2,5,10,25,50,75,90,98,99, NA)/100)[1:10],
      max(object)), 3)
    colnames(a) <- c("Obs.", "NA", "Mean", "S.D.", "Min.",
      "1pct", "2pct", "5pct", "10pct", "25pct",
      "Median", "75pct", "90pct", "98pct", "99pct", "Max.")
  }
}

```



```

    rownames(a) <- " "
    return(a)
}

s.factor <- function (object,group)
{
  a <- rep("",7)
  dim(a) <- c(1,7)
  a[1, ] <- round(c(length(na.omit(object)),
    length(object)-length(na.omit(object)),
    mean(as.numeric(na.omit(object))),
    sd(as.numeric(na.omit(object))),
    min(as.numeric(na.omit(object))),
    quantile(as.numeric(na.omit(object)),0.5),
    max(as.numeric(na.omit(object)))), 3)

  colnames(a) <- c("Obs.", "NA", "Mean", "S.D.", "Min.", "Median", "Max.")
  if (!missing(group))
    rownames(a) <- names(table(group,deparse.level=1))[1]
  else rownames(a) <- " "
  return(a)
}

summarize.factor <- function (object, group=NULL, frame=get(".ec.frame",.GlobalEnv))
{
  if (!missing(group)) {
    cat(deparse(substitute(object)),sep="")
    cat(" stratified by", deparse(substitute(group)),"\n")
    aa <- NULL
    for (i in 1:length(table(group)))
      aa <- rbind(s.factor(object[group==names(table(group))[i]]),aa)
    row.names(aa) <- names(table(group))
  }
  else {
    aa <- s.factor(object)
    row.names(aa) <- deparse(substitute(object))
  }
  if (frame) print.frame(aa,type="character")
  else print.noquote(aa)
}

s.data <- function (data)
{
  a <- rep("", (dim(data)[2]) * 8)
  dim(a) <- c(dim(data)[2], 8)
  colnames(a) <- c("Variable", "Obs.", "NA", "Mean", "S.D.",
    "Min.", "Median", "Max.")
  a[, 1] <- attr(data, "names")
  rownames(a) <- 1:nrow(a)
  for (i in 1:(dim(data)[2])) {
    if ((typeof(data[i][1, ]) == "character") ||
      is.na(mean(na.omit(as.numeric(data[[i]]))))) {
      a[i, 2] <- as.character(length(data[[i]]))
      a[i, 3] <- as.character(length(data[[i]])-length(na.omit(data[[i]])))
      a[i, 4:8] <- ""
    }
    else if (class(data[[i]])[1] == "Date") {
      xx <- summary(data[[i]])
      a[i, 2] <- as.character(length(na.omit(unclass(data[[i]]))))
      a[i, 3] <- as.character(length(unclass(data[[i]])
        -length(na.omit(unclass(data[[i]]))))
      a[i, 4:5] <- "-"
      a[i, 6:8] <- format.Date(c(xx[1],xx[3],xx[6]))
    }
    else if (class(data[[i]])[1] == "POSIXt" | class(data[[i]])[1] ==
      "POSIXct") {
      xx <- summary(data[[i]])
      a[i, 2] <- as.character(length(na.omit(as.character(data[[i]]))))
      a[i, 3] <- as.character(length(as.character(data[[i]])
        -length(na.omit(as.character(data[[i]]))))
      a[i, 4:5] <- "-"
      a[i, 6:8] <- format.Date(c(xx[1],xx[3],xx[6]))
    }
    else if (class(data[[i]]) == "Time" ) {
      xx <- summary(data[[i]])
      a[i, 2] <- as.character(length(na.omit(as.character(data[[i]]))))
      a[i, 3] <- as.character(length(as.character(data[[i]]))

```

```

        -length(na.omit(as.character(data[[i]])))
        a[i, 4:5] <- "-"
        a[i, 6:8] <- as.POSIXct(c(xx[1],xx[3],xx[6]))
    }
    else if (is.integer(data[[i]]) || is.real(data[[i]]) {
        a[i, 4:8] <- round(c(mean(na.omit(as.numeric(data[[i]]))),
            sd(na.omit(as.numeric(data[[i]]))),
            min(na.omit(as.numeric(data[[i]]))),
            quantile(na.omit(as.numeric(data[[i]])), 0.5),
            max(na.omit(as.numeric(data[[i]]))), 2)
        a[i, 2] <- as.character(length(na.omit(as.numeric(data[[i]]))))
        a[i, 3] <- as.character(length(data[[i]])-length(na.omit(data[[i]])))
    }
    else if (is.null(class(data[[i]])) {
        a[i, 4:8] <- round(c(mean(na.omit(as.numeric(data[[i]]))),
            sd(na.omit(as.numeric(data[[i]]))),
            min(na.omit(as.numeric(data[[i]]))),
            quantile(na.omit(as.numeric(data[[i]])), 0.5),
            max(na.omit(as.numeric(data[[i]]))), 2)
        a[i, 2] <- as.character(length(na.omit(data[[i]])))
        a[i, 3] <- as.character(length(data[[i]])-length(na.omit(data[[i]])))
    }
    else if (is.factor(data[i][2, ])) {
        a[i, 4:8] <- round(c(mean(as.numeric(na.omit(data[i][, ]))),
            sd(as.numeric(na.omit(data[i][, ]))),
            min(as.numeric(na.omit(data[i][, ]))),
            median(as.numeric(na.omit(data[i][, ]))),
            max(as.numeric(na.omit(data[i][, ]))), 3)
        a[i, 2] <- as.character(length(na.omit(data[[i]])))
        a[i, 3] <- as.character(length(data[[i]])-length(na.omit(data[[i]])))
    }
}
return(a)
}

summarize.data.frame <- function (data, subset=NULL, select=NULL, group=NULL,
    pctl=get(".ec.pctl",.GlobalEnv), frame=get(".ec.frame",.GlobalEnv))
{
    if (!missing(subset)) {
        subset <- eval(substitute(subset),data)
        data <- subset(data,subset)
    }
    if (!missing(group)) {
        groupx <- eval(substitute(group),data)
        xx <- table(groupx,deparse.level=1)
        for (i in 1:length(xx)) {
            if (missing(select)) {
                z <- attr(xx,"dimnames")[[1]][i]
                if (i>1) cat("\n")
                cat("group:", deparse(substitute(group)), "=", z, "\n")
                aa <- subset(data,groupx==z)
                bb <- s.data(aa)
                aa <- subset.data.frame(bb,select=-1)
                row.names(aa) <- bb[,1]
            }
            else {
                nl <- as.list(1:ncol(data))
                names(nl) <- names(data)
                vars <- eval(substitute(select), nl)
                z <- attr(xx,"dimnames")[[1]][i]
                zz <- data[vars[1]][[1]]
                if (length(vars)>1)
                    for (j in 2:length(vars))
                        zz <- data.frame(zz,data[vars[j]][[1]])
                else zz <- data.frame(zz)
                names(zz) <- names(data[vars])
                if (i>1) cat("\n")
                cat("group:", deparse(substitute(group)), "=", z, "\n")
                aa <- subset(zz,groupx==z)
                bb <- s.data(aa)
                aa <- subset.data.frame(bb,select=-1)
                row.names(aa) <- bb[,1]
            }
        }
        if (frame) print.frame(aa,type="data")
        else print.noquote(aa, digits = 3)
    }
}

```

```

}
else {
  if (missing(select)) {
    bb <- s.data(data)
    aa <- subset.data.frame(bb,select=-1)
    row.names(aa) <- bb[,1]
  }
  else {
    nl <- as.list(1:ncol(data))
    names(nl) <- names(data)
    vars <- eval(substitute(select), nl)
    xx <- data[vars[1]][[1]]
    if (length(vars)>1)
      for (j in 2:length(vars)) {
        xx <- data.frame(xx,data[vars[j]][[1]])
      }
    else xx <- data.frame(xx)
    names(xx) <- names(data[vars])
    bb <- s.data(xx)
    aa <- subset.data.frame(bb,select=-1)
    row.names(aa) <- bb[,1]
  }
  if (frame) print.frame(aa,type="data")
  else print.noquote(aa, digits = 3)
}
}

```

รายการฟังก์ชันที่พัฒนาขึ้นในโครงการนี้

แพ็คเกจ ice

ฟังก์ชันต่างๆ ในแพ็คเกจนี้แบ่งเป็นกลุ่มๆ ตามการใช้งานได้เป็น 6 กลุ่ม คือ file management, data management, data summary, statistical test, statistical modeling, และ memory and system management ซึ่งเป็นการทำงานโดยทั่วไปในส่วนของจัดการแฟ้มข้อมูล ข้อมูล การวิเคราะห์ทางสถิติ การทำโมเดลทางสถิติ และอื่นๆ ไม่ได้เกี่ยวข้องกับทางระบาดวิทยาโดยตรง ซึ่งในส่วนนี้มีจำนวนมากและครอบคลุมกระบวนการทำงานในการจัดการแฟ้มและข้อมูลไปจนถึงการวิเคราะห์ทางสถิติ จึงพบว่าการจัดการสอนการใช้งานโปรแกรม R นั้น นอกจากการทำงานพื้นฐานจริงๆ แล้วก็แทบจะไม่จำเป็นต้องสอนฟังก์ชันอื่นของ R อีก ทำให้ไม่มีความจำเป็นที่จะต้องแปลความช่วยเหลือฟังก์ชันของ R อื่นๆ เป็นภาษาไทย เนื่องจากฟังก์ชันที่พัฒนาขึ้นทั้งหมดนี้มีความช่วยเหลือทั้งในภาษาอังกฤษและภาษาไทยอยู่แล้ว

รายการฟังก์ชันที่พัฒนาขึ้นในแพ็คเกจ ice มีดังนี้

File magement

- create
- do
- fread
- fwrite
- logg
- output

Data management

- as.na
- change
- cjoin
- defector
- expand
- label
- pack
- recode.na
- recode

- relabel
- rename
- shift
- sort.data
- to.character
- to.factor
- to.numeric
- to.vector
- Data summary
 - describe
 - display
 - ftab
 - list.rep
 - match.table
 - summarize
 - tab
 - xtab
- Statistical test
 - anova.test
 - ci
 - lr.test
 - shapiro.test
 - st.test
 - wilcoxon.test
- Statistical modeling
 - logistic
 - logit
- Memory and system management
 - clean
 - ec.data
 - ec.frame
 - ec.options
 - exit
 - frame.line
 - get.ec.option
 - li
 - pause
 - print.frame
 - print.mtable

แพ็คเกจ epid

ฟังก์ชันในแพ็คเกจนี้เป็นฟังก์ชันที่เกี่ยวข้องกับการคำนวณทางระบาดวิทยาพื้นฐาน หลายฟังก์ชันพัฒนา
เลียนแบบคำสั่งที่มีในโปรแกรม Stata เช่น cc, cs เป็นต้น รายการฟังก์ชันมีดังนี้

- cc
- cs
- epidemic
- ir
- mcc
- mhor

- randomise
- spower
- ssize

2. การทำความช่วยเหลือการใช้งานเป็นภาษาอังกฤษและภาษาไทย

เมื่อมีการทำแพ็คเกจสำหรับโปรแกรม **R** นั้น แม้จะไม่ใช่ข้อกำหนดตายตัว แต่ส่วนใหญ่ก็จะต้องทำความช่วยเหลือการใช้งานฟังก์ชันต่างๆ ด้วย ทั้งนี้เพื่อให้ผู้อื่นเข้าใจ และใช้ฟังก์ชันเหล่านั้นได้ โดยส่วนใหญ่ก็จะมีตัวอย่างการใช้งานประกอบด้วย

การทำความช่วยเหลือใน **R** นั้น ทำได้โดยการสร้างแฟ้มชนิด R documentation (Rd) ขึ้น ซึ่งมี tag ที่ต้องเขียนเสมอคือ

```
\name{}
\alias{}
\title{}
\description{}
\usage{}
```

ส่วน tag นอกจากนั้นจะเขียนหรือไม่เขียนก็ได้ และยังมี tag ที่ช่วยเน้นการแสดงข้อความอื่นๆ อีก เช่น `\bold{}`, `\code{}`, `\emph{}` เป็นต้น

ในที่นี้จะกล่าวถึงตัวอย่างแฟ้มความช่วยเหลือสำหรับแพ็คเกจ `ice.dataman` ซึ่งเขียนไว้ในแฟ้มชื่อ `ice.dataman.Rd` ดังนี้

ice.dataman.Rd

เป็นแฟ้มอธิบายการทำงานของแพ็คเกจและฟังก์ชัน ในที่นี้คือแพ็คเกจ `ice.dataman` เขียนในรูปแบบของ R documentation format (Rd) โดยทำเป็นภาษาอังกฤษ และจะมีแฟ้มคู่กันที่แปลเป็นภาษาไทย บรรจุไว้ในแพ็คเกจ `ice.help.th` ซึ่งเป็นนวัตกรรมที่ทางคณะผู้วิจัยได้ทำให้ เพื่อความสะดวกในการสร้างแฟ้มช่วยเหลือเป็นหลายภาษา แต่ในที่นี้จะได้แสดงข้อความภายในแฟ้มทั้งในภาษาอังกฤษและภาษาไทยให้เห็นเทียบกันดังนี้

ภาษาอังกฤษ

```
\name{ice.dataman}
\alias{ice.dataman}
\title{Data Manager}
\description{
Collection of functions for data management in \code{base} and
\code{ice} packages, including vector and data frame management,
changing codes and values in variables in a data frame, and factor
labelling. It is designed for \bold{ICE} environment.
}

\usage{
ice.dataman()
}

\details{
Functions or a few command lines supporting data frame management
include creating a blank data frame, duplicating, renaming, and
subsetting a data frame. Renaming of variables in a data frame is
also available. A data frame can be transformed by row and column
conditions. Values or codes within variables in a data frame can be
```

changed, recoded, and NA's can be easily handled. Vector of character and numeric types and factors can be interchanged easily.

`{ice.dataman}` menu is also included in `{ice.datamain}` and `{ice.fullmain}` menus for convenience of use. If `{ice.dataman}` is used as a floating menu (not including in the `{main}` menu), the command line it creates can be recorded into a script file by pressing `{Record}` button on the `{main}` menu bar.

```
\note{
The project is funded by the Health Research Network, Thailand
Research Fund, and the Thailand National Health Foundation.

\emph{Web site:} \url{http://www.r-ice.org}
}

\author{
Hutcha Sriplung, Epidemiology Unit, Prince of Songkla University, Hat
Yai, Songkhla, 90110, Thailand.

\emph{e-mail:} \email{hutcha.s@psu.ac.th}
}

\keyword{misc}
```

ภาษาไทย

```
\name{ice.dataman}
\alias{ice.dataman}
\title{Data Manager}
\description{
ที่รวมฟังก์ชันสำหรับการจัดการข้อมูลในแพ็คเกจ {base} และ {ice} ประกอบด้วย
การจัดการเวกเตอร์และกรอบข้อมูล เปลี่ยนรหัสและค่าในตัวแปรในกรอบข้อมูล และการให้ป้ายแฟกเตอร์ ออกแบบ
มาเพื่อสภาพแวดล้อม {ICE}
}

\usage{
ice.dataman()
}

\details{
ฟังก์ชันหรือชุดบรรทัดคำสั่งเล็กๆ ที่สนับสนุนการจัดการกรอบข้อมูล ประกอบด้วยการสร้างกรอบข้อมูลเปล่า
การทำซ้ำ การเปลี่ยนชื่อ และการทำเซตย่อยกรอบข้อมูล และยังสามารถเปลี่ยนชื่อตัวแปรในกรอบข้อมูล
เปลี่ยนแปลงกรอบข้อมูลตามเงื่อนไขด้านแถวและสดมภ์ และยังสามารถเปลี่ยนค่าหรือรหัสค่าและรหัส และ
จัดการค่า NA ในตัวแปรภายในกรอบข้อมูลได้อย่างง่ายดาย และยังสามารถเปลี่ยนแปลงเวกเตอร์ชนิดข้อความ
และตัวเลข และแฟกเตอร์ไปมาได้ง่าย
```

รายการเมนู `{ice.dataman}` ยังเป็นส่วนประกอบในรายการเมนู `{ice.datamain}` และ `{ice.fullmain}` เพื่อความสะดวกในการทำงาน ถ้าใช้งาน `{ice.dataman}` แบบ

```

รายการเมนูลอย (ไม่รวมอยู่ในเมนู \code{main}) บรรทัดคำสั่งที่ถูกสร้างขึ้นสามารถบันทึกในแฟ้มสคริปต์
ได้โดยการกดปุ่ม \code{บันทึก} บนรายการเมนู \code{main}
}

\note{
โครงการนี้ได้รับทุนจากเครือข่ายวิจัยสุขภาพ สำนักงานกองทุนสนับสนุนการวิจัย (สกว.) และ มูลนิธิ
สาธารณสุขแห่งชาติ (มสช.)

\emph{Web site:} \url{http://www.r-ice.org}
}

\author{
หัชชา ศรีปลั่ง หน่วยระบาดวิทยา คณะแพทยศาสตร์ มหาวิทยาลัยสงขลานครินทร์ อ. หาดใหญ่ จ. สงขลา
90110

\emph{e-mail:} \email{hutcha.s@psu.ac.th}
}

\keyword{misc}

```

ในการ compile แฟ้มความช่วยเหลือ จะทำในระหว่างการ compile แพคเกจ จึงจะนำไปกล่าวในหัวข้อการ compile แพคเกจข้างล่างเพียงทีเดียว

ในกระบวนการ compile แพคเกจนั้น ถ้าในแพคเกจมีแฟ้มความช่วยเหลือในรูปแบบ Rd ดังกล่าวอยู่ด้วยแล้ว การ compile จะสร้างแฟ้มความช่วยเหลือแบบต่างๆ ทั้งแบบ on-line และแฟ้มชนิด pdf มาด้วย ซึ่งความช่วยเหลือแบบ online นั้น มีทั้งชนิดที่เป็นแฟ้ม text แฟ้ม html ที่เปิดได้โดยโปรแกรม web browser ทั่วไป แบบ chtml ของระบบปฏิบัติการวินโดวส์ และแบบ dvi สำหรับระบบยูนิกซ์ โดยที่คำปรีายบนระบบปฏิบัติการต่างๆ กันนั้นตั้งไว้ไม่เหมือนกัน เมื่อเรียกคำสั่ง help จากหน้าจอ R console จากระบบวินโดวส์ จะเรียกความช่วยเหลือแบบ chtml มาแสดง บนระบบแมคอินทอช OSX จะเรียกแฟ้ม html มาแสดงโดยใช้ browser ของ RGUI เอง และบนลินุกซ์นั้น คำปรีายเป็น text แต่หากต้องการดูความช่วยเหลือแบบ html ก็ทำได้โดยใส่ตัวเลือก html=TRUE ในการเรียก help ได้ด้วย

ข้อจำกัดของแฟ้มความช่วยเหลือแบบ pdf คือไม่สามารถสร้างแฟ้มความช่วยเหลือเป็นภาษาไทยได้โดยตรงในการ compile แพคเกจ เนื่องจาก script ในการแปลระหว่างการ compile ไม่สามารถจัดการภาษาไทยได้อย่างถูกต้อง จะต้องสร้างเป็นภาษาอังกฤษเท่านั้น ซึ่งจะมีหน้าตาแฟ้มความช่วยเหลือดังนี้

Package ‘ice.dataman’

January 3, 2007

Version 1.1-1

Date 2007-01-03

Title ICE Data Manager

Author Hutcha Sriplung

Maintainer Hutchia Sriplung <hutchia.s@psu.ac.th>

Depends R (>= 2.3.0), ice

Description Functions for data management

License GPL Version 2 or later

URL <http://www.r-ice.org>

R topics documented:

ice.dataman 1
Index 3

ice.dataman *Data Manager*

Description

Collection of functions for data management in `base` and `ice` packages, including vector and data frame management, changing codes and values in variables in a data frame, and factor labelling. It is designed for **ICE** environment.

Usage

```
ice.dataman()
```

Details

Functions or a few command lines supporting data frame management include creating a blank data frame, duplicating, renaming, and subsetting a data frame. Renaming of variables in a data frame is also available. A data frame can be transformed by row and column conditions. Values or codes within variables in a data frame can be changed, recoded, and NA's can be easily handled. Vector of character and numeric types and factors can be interchanged easily.

`ice.dataman` menu is also included in `ice.datamain` and `ice.fullmain` menues for convenience of use. If `ice.dataman` is used as a floating menu (not including in the main menu), the command line it creates can be recorded into a script file by pressing `Record` button on the `main` menu bar.

Note

The project is funded by the Health Research Network, Thailand Research Fund, and the Thailand National Health Foundation.

Web site: <http://www.r-ice.org>

Author(s)

Hutchia Sriplung, Epidemiology Unit, Prince of Songkla University, Hat Yai, Songkhla, 90110, Thailand.

e-mail: (hutchia.s@psu.ac.th)

Index

*Topic **misc**

ice.dataman, 1

ice.dataman, 1

หากแพ็คเกจมีหลายฟังก์ชันอยู่ด้วยกัน เช่นในแพ็คเกจ ice ความช่วยเหลือของทุกฟังก์ชันจะรวมอยู่ในแฟ้ม pdf เดียวกันนี้ โดยมี index อยู่ท้ายแฟ้ม เป็นตัวบอกว่าความช่วยเหลือของแต่ละฟังก์ชันอยู่ที่หน้าไหนบ้าง

3. การทดสอบฟังก์ชัน

ฟังก์ชันต่างๆ ที่ได้สร้างขึ้นนั้น จะผ่านกระบวนการทดสอบเพื่อหาข้อผิดพลาด บกพร่อง และข้อจำกัดโดยขั้นตอนดังนี้

ในการทำเพิ่มความช่วยเหลือนั้น ส่วนใหญ่จะต้องเขียนตัวอย่างการใช้งานให้ครอบคลุมในกรณีต่างๆ ซึ่งในกระบวนการเขียนตัวอย่างนี้เอง ผู้เขียนฟังก์ชันได้ทดสอบไปรอบหนึ่งแล้วว่าฟังก์ชันสามารถทำงานในเงื่อนไขตัวอย่างได้ถูกต้อง

ขั้นตอนที่สอง ฟังก์ชันจะถูกทดสอบโดยผู้ร่วมวิจัยอีกหนึ่งท่าน โดยเฉพาะ นส. อภริดี ซึ่งมีหน้าที่ในหน่วยระบาดวิทยาในการให้คำปรึกษางานวิจัยแก่บุคลากรในคณะแพทย์ ทำให้พบเงื่อนไขที่แปลกแตกต่างไปหลายๆ รูปแบบ จึงเหมาะในการทดสอบการใช้งานฟังก์ชันซึ่งอาจมีบางครั้งที่ผลไม่ได้ตามที่คาดหวัง ซึ่งผู้พัฒนาฟังก์ชันจะได้นำกลับมาปรับปรุงแก้ไขอีกครั้งหนึ่ง

ขั้นตอนที่สาม เมื่อเผยแพร่ฟังก์ชันไปแล้วในรูปแบบแพ็คเกจ ผู้ใช้ยังสามารถแจ้งการทำงานที่ผิดพลาดหรือความคาดหวังที่ต้องการเพิ่มเติมจากฟังก์ชันนั้นๆ มาให้ทราบ ซึ่งก็จะทำการปรับปรุงแก้ไขในแพ็คเกจรุ่นต่อไป

ในการจัดอบรมเชิงปฏิบัติการแต่ละครั้งนั้น คณะผู้วิจัยได้พบข้อบกพร่องหรือสิ่งที่ควรเพิ่มเติมในฟังก์ชันต่างๆ ซึ่งเป็นประโยชน์มากในการพัฒนาฟังก์ชัน

นอกจากนี้ ในระหว่างการค้าเนินโครงการ **R** ได้พัฒนาโปรแกรมรุ่นใหม่อย่างต่อเนื่อง จากรุ่น 2.1.x มาจนถึงรุ่น 2.4.1 ในเดือนธันวาคม 2549 ซึ่งในแต่ละรุ่นที่ปรับปรุงนั้น ฟังก์ชันพื้นฐานที่ใช้ในการเขียนฟังก์ชันในแพ็คเกจมีการพัฒนาเปลี่ยนแปลงการทำงานหรือการแสดงผลไปบ้าง ทำให้ต้องมีการทดสอบและปรับปรุงฟังก์ชันตามรุ่นของโปรแกรม **R** ที่เปลี่ยนไปตลอดเวลา

4. การพัฒนารายการเมนูกราฟิก

หลังจากได้ศึกษาแพ็คเกจที่มีรายการเมนูกราฟิกต่างๆ แบบ GUI ภายใต้อุปกรณ์ **R** มากขึ้น พบว่ามีผู้พัฒนาแพ็คเกจกราฟิกที่สามารถนำมาใช้เขียน R-GUI ได้อยู่จำนวนหนึ่ง เช่น tcltk, RGtk, wxPython และยังมีผู้พัฒนา GUI โดยใช้แพ็คเกจเหล่านี้บ้างแล้ว เช่น R Commander (RCmdr) ซึ่งพัฒนาโดย John Fox ใช้แพ็คเกจ tcltk แพ็คเกจ Poor Man GUI (PGM) ใช้แพ็คเกจ RGtk2 แพ็คเกจ RKward (อาร์คัวร์ด) ซึ่งใช้ได้เฉพาะภาษา Java พัฒนาโดยทีมจากมหาวิทยาลัย Augsburg ประเทศเยอรมันนี

ในช่วงแรกบางแพ็คเกจใช้ได้หรือติดตั้งได้ง่ายกับระบบปฏิบัติการบางระบบ เช่น PMG เฉพาะระบบ Linux หรือ JGR กับระบบ Macintosh เท่านั้น อย่างไรก็ตาม ในปัจจุบันเกือบทุกแพ็คเกจสามารถใช้งานระบบปฏิบัติการได้ทั้ง Windows, UNIX และ Macintosh ได้

การนำ GUI มาใช้ในระยะนั้น ผู้วิจัยยังพบว่ามีข้อจำกัดหลายประการ ดังนี้ 1. GUI เป็นภาษาอังกฤษเท่านั้น ไม่สามารถทำเป็นภาษาอื่นได้ 2. GUI ส่วนใหญ่นับพื้นที่การพิมพ์คำสั่ง มีรายการเมนูจำกัด หรือบางอันไม่มีรายการเมนูเลย 3. การปรับปรุง เปลี่ยนแปลง แก้ไขให้เหมาะกับการทำงานที่แตกต่างกันทำได้หรือทำได้ยาก เนื่องจากเป็นโครงสร้าง GUI รวมขนาดใหญ่ และแต่ละส่วนผูกพันกันเอง จนมีแต่ทีมที่พัฒนาเท่านั้นที่

จะแก้ไขได้ ประกอบกับ R ที่ใช้ยังเป็นรุ่น 2.1.x ถึง 2.2.x ซึ่งยังไม่สนับสนุน internationalization คือยังไม่สามารถใช้กับภาษาอื่นได้ นอกจากภาษาอังกฤษ จนเมื่อมีการพัฒนามาจนถึงรุ่น 2.3.x แล้ว จึงสามารถแสดงผลเป็นภาษาอื่นๆ ได้ ทำให้ยังไม่ตัดสินใจที่จะนำแพ็คเกจใดแพ็คเกจหนึ่งมาใช้ในโครงการ

จนเมื่อผู้วิจัยได้รู้จักแพ็คเกจ GraspR ซึ่งเป็น GUI ที่พัฒนาโดยใช้แพ็คเกจ tcltk ซึ่งเกี่ยวกับการทำ generalized regression analysis and spatial prediction for R และได้ลองใช้งานดู จึงพบว่าแพ็คเกจที่แยกหน้าต่างออกจากกันเป็นหลายหน้าต่าง ทำให้สามารถพัฒนาแยกกันได้ จึงเกิดความคิดที่จะพัฒนาระบบ R-GUI ที่มีแนวคิดหลักของการทำงานดังแบบ COMP และได้ตั้งชื่อว่า Integrated Computing Environment for R (R-ICE อ่านว่า ไรซ์)

แนวคิดของ COMP มีดังต่อไปนี้

Customizability
Open environment
Modularity
Platform independence

ในส่วนของ customizability หมายความว่า R-ICE จะต้องสามารถใช้กับภาษาเขียนต่างๆ ได้ (ภาษามนุษย์ ไม่ใช่ภาษาคอมพิวเตอร์) และผู้ใช้สามารถเลือกส่วนประกอบของระบบให้เข้ากับการใช้งานของแต่ละคนได้ และยังสามารถสร้าง GUI ใหม่เองได้ถ้าต้องการทำ

Open environment หมายถึง ผู้ใช้สามารถแจกจ่ายและปรับปรุงเปลี่ยนแปลง source code ภายใต้ลิขสิทธิ์แบบ GPL ได้ และเปิดให้นักพัฒนาที่จะสามารถสร้างโมดูล (แพ็คเกจ) ของตัวเองที่ทำงานร่วมกับสภาพแวดล้อม R-ICE และแจกจ่ายให้กับผู้ใช้คนอื่นๆ ได้

ความหมายของ modularity คือระบบนี้แบ่งเป็นโมดูล (แพ็คเกจ) หลายโมดูล โดยแต่ละโมดูลอาจขึ้นต่อกันอยู่บ้างหรือส่วนใหญ่จะไม่ขึ้นต่อกันเลย และสามารถนำมาประกอบใช้งานร่วมกันได้แบบ plug and play โดยไม่รบกวนการทำงานของโมดูลอื่น

และสุดท้าย platform independence คือ R-ICE จะต้องสามารถใช้ได้กับคอมพิวเตอร์ทุกระบบหลัก โดยเป็นแพ็คเกจของ R ที่จะสามารถติดตั้งได้เหมือนแพ็คเกจปกติทั่วไปของ R นั่นเอง ทั้งนี้เป็นไปได้เนื่องจาก ICE เขียนโดยใช้แพ็คเกจ tcltk ของ R นั่นเอง ดังนั้นหากจะมีข้อจำกัดใดๆ ก็เป็นข้อจำกัดที่เกิดขึ้นจากแพ็คเกจ tcltk เอง

5. การทำแพ็คเกจเพื่อเผยแพร่ผลงานแก่สาธารณะ

ก. แพ็คเกจ EpiCan

ในช่วงแรกของโครงการ ได้ทำแพ็คเกจรวมฟังก์ชันต่าง ตั้งชื่อว่า EpiCan โดยแบ่งกลุ่มฟังก์ชันดังนี้

1. general
2. file manager
3. data manager
4. data summary
5. statistical tests
6. statistical models
7. graphics
8. epidemiology
9. miscellaneous

โดยทั้งหมดไม่มี graphic user interface (GUI) เป็นแต่ฟังก์ชันที่ต้องเขียนแบบ command line เท่านั้น ได้ปรับปรุง แก้ไข เพิ่มเติมมาโดยลำดับ จากรุ่น 1.1.1 จนถึงรุ่น 1.3-7 แพคเกจ EpiCan ได้ถูกนำมาใช้ในการเรียนการสอนของนักศึกษาในหลักสูตรระดับวิทยาลัยนานาชาติ และในการอบรมให้แก่แพทย์ประจำบ้านและแพทย์ใช้ทุนของคณะแพทยศาสตร์ มหาวิทยาลัยสงขลานครินทร์ด้วย จุดเด่นของแพคเกจนี้คือมีฟังก์ชันเกี่ยวกับการจัดการแฟ้มและข้อมูลได้สะดวกกว่าฟังก์ชันพื้นฐานของ **R** เอง ทำให้ผู้ใช้สามารถจัดการนำเข้าแฟ้มข้อมูลและจัดการตัวแปรต่างๆ ได้สะดวกขึ้นมาก และนอกจากนี้ฟังก์ชันการวิเคราะห์ทางสถิติบางอย่างที่ได้เปลี่ยนแปลงการแสดงผลให้ดูง่าย คล้ายกับการแสดงผลของโปรแกรม Stata ทำให้ผู้ใช้ตีความผลลัพธ์ของการทดสอบต่างๆ ได้ง่ายกว่าผลลัพธ์ของฟังก์ชันในแพคเกจ Stats ของ **R**

จะเห็นได้ว่าการเรียกชื่อรุ่นของ EpiCan มีความแตกต่างไปจากเดิม คือเดิมใช้ 1.1.1 แต่ต่อมาในรุ่น 1.3-6 ได้เปลี่ยนมาใช้สัญลักษณ์ - ระหว่างเลขรุ่นตำแหน่งที่สองกับสาม ทั้งนี้เนื่องจากในระยะแรกยังไม่ทราบวิธีการเขียนเลขรุ่นของแพคเกจ จนในระยะหลังได้เรียนรู้มากขึ้น จึงทราบว่าวิธีการของ **R** นั้น จะใช้สัญลักษณ์ - ระหว่างเลขรุ่นตำแหน่งที่สองกับสาม นี่เป็นตัวอย่างให้เห็นว่าผู้วิจัยได้เรียนรู้เพิ่มเติมเกี่ยวกับโปรแกรม **R** มากขึ้นเป็นระยะๆ ในระหว่างการทำโครงการนี้ และนอกจากนี้ยังมีอีกหลายอย่างที่ได้ประสบการณ์เพิ่มมากขึ้นจากการลองผิดลองถูก เนื่องจากเอกสารประกอบสำหรับการเขียนฟังก์ชันและแพคเกจใน **R** มีน้อยมาก และเขียนสั้นมากจนมักจะอ่านไม่เข้าใจ เนื่องจากงานเช่นนี้เป็นงานที่น้อยคนจะมีโอกาสทำถึงขั้นนี้

ข. แพคเกจ ICE

ในระบบ **R-ICE** นั้น จะประกอบด้วยสองส่วนใหญ่ คือ ส่วนประกอบต่างๆ ของ **R** มาตรฐาน กับส่วนที่เป็น **ICE** ที่พัฒนาขึ้น ซึ่งส่วนใหญ่เป็นส่วนของ GUI ซึ่งการสร้างทำได้โดยการแยกระบบ EpiCan ออกเป็นชิ้นส่วนเอกเทศไม่ขึ้นแก่กันสองส่วน คือ ice ซึ่งทำงานเกี่ยวกับการจัดการแฟ้มข้อมูล การจัดการกรอบข้อมูลและตัวแปร การสรุปข้อมูล การสร้างกราฟ และการทดสอบทางสถิติอย่างง่าย และอีกส่วนคือ epid ซึ่งทำงานเกี่ยวกับระบาดวิทยา และสร้างส่วน GUI เพิ่มเติม คือแพคเกจที่มีคำว่า ice นำหน้า รวมทั้งหมดแล้วในระยะแรกมีอยู่ 10 แพคเกจ เรียงตามตัวอักษรคือ

1. epid
2. ice
3. ice.commands
4. ice.dataman
5. ice.epid
6. ice.graph
7. ice.main
8. ice.objects
9. ice.statis
10. ice.summary

และจัดเป็นกลุ่มได้ดังนี้

1. Global module
 - a. ice
2. Main module
 - a. ice.main
3. Associated modules
 - a. ice.commands
 - b. ice.dataman

- c. ice.graph
 - d. ice.objects
 - e. ice.summary
 - f. ice.statis
4. Extended modules
- a. epid
 - b. ice.epid

Main module เป็นแกนกลางของ **ICE GUI** ทำหน้าที่บริหารจัดการให้แพ็คเกจ ice และแพ็คเกจอื่นๆ ทำงานร่วมกันเป็นระบบเดียว จัดการเปิดหรือสร้างแฟ้มสคริปต์สำหรับระบบ **R-ICE** ทั้งหมด ทำงานแฟ้มสคริปต์และส่งผลบนหน้าจอหรือบันทึกลงแฟ้มก็ได้ และยังรับหน้าที่จัดการอ่านเขียนแฟ้มข้อมูล และจัดการวัตถุต่างๆ ในหน่วยความจำ รวมทั้งมีปุ่มบันทึกบรรทัดคำสั่งที่ถูกสร้างขึ้นจากรายการเมนูของทุกแพ็คเกจ ในระบบ **ICE** ลงสู่แฟ้มสคริปต์เดียวกันด้วย ในอนาคตส่วนนี้อาจมีแพ็คเกจเพิ่มเติมอีกก็ได้ แต่โดยพื้นฐานแล้วควรเป็นการเพิ่มฟังก์ชันให้กับ ice โดยไม่มีการสร้างแพ็คเกจเพิ่มเติม แต่อาจเป็นไปได้ที่จะแบ่งแพ็คเกจ ice เป็นส่วนย่อยมากกว่าหนึ่งแพ็คเกจก็ได้

Associated modules เป็นแพ็คเกจกราฟิก ทำหน้าที่จำเพาะตามชื่อ ice.commands ทำหน้าที่เป็น editor ขนาดเล็ก ที่สามารถเขียนบรรทัดคำสั่งลงแฟ้มสคริปต์ที่สร้างขึ้นหรือเปิดโดย ice.main, ice.objects ทำหน้าที่เรียกดูวัตถุประเภทต่างๆ ในหน่วยความจำในขณะหนึ่งๆ ice.dataman ทำหน้าที่เกี่ยวกับการจัดการกรอบข้อมูล ice.graph ทำหน้าที่เกี่ยวกับการวาดกราฟพื้นฐาน เป็นต้น ในส่วนนี้อาจมีการรวมโมดูลต่างๆ เข้าด้วยกันเป็นชุด ให้ผู้ใช้ได้เลือกใช้ตามความชอบหรือให้เข้ากับงานที่ต้องทำเป็นประจำก็ได้

Extended modules เป็นแพ็คเกจอื่นๆ ซึ่งอาจเป็นที่รวมฟังก์ชัน หรือแพ็คเกจ GUI ก็ได้ ที่ทำงานพิเศษในเรื่องใดเรื่องหนึ่ง ไม่ใช่งานพื้นฐานทั่วไป ซึ่งส่วนนี้สามารถสร้างขยายได้ไม่จำกัด โดยอาจพัฒนาไปใช้ทางด้านระบาดวิทยา ชีวสถิติ วิศวกรรม สังคมศาสตร์ หรืองานในสาขาอื่นใดก็ได้ แต่ทั้งหมดสามารถทำงานร่วมกับแพ็คเกจอื่นๆ ในระบบ **R-ICE** ได้โดยผ่านแพ็คเกจ ice.main นั่นเอง ในส่วนนี้มีทางพัฒนาไปได้อย่างกว้างขวางมากไม่มีที่สิ้นสุด

สภาพแวดล้อม **R-ICE** ที่กล่าวมานี้ ได้สรุปนำเสนอในที่ประชุม UseR! 2006 ที่กรุงเวียนนา ประเทศออสเตรีย เมื่อเดือนมิถุนายน 2549 และได้เปิดเว็บไซต์ขึ้นที่ <http://www.r-ice.org> โดยผู้สนใจสามารถดาวน์โหลดแพ็คเกจต่างๆ ได้ฟรี และพบว่ามีชาวต่างชาติเข้ามาดาวน์โหลดไปใช้งานเช่นกัน

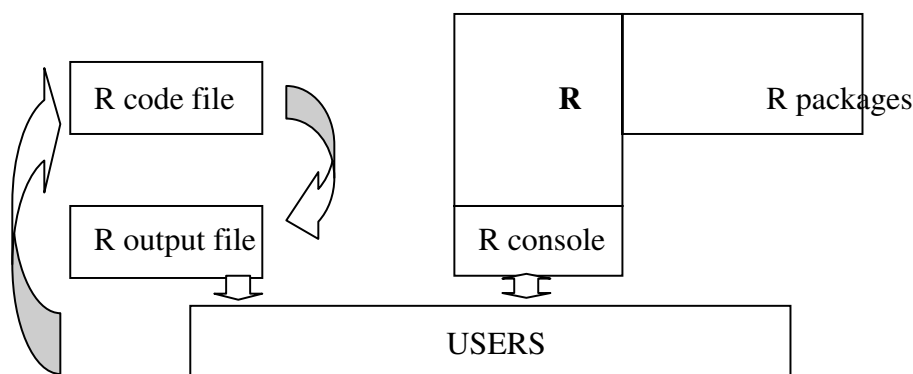
สุดท้ายแล้ว ขณะนี้ (มกราคม 2550) มีแพ็คเกจต่างๆ 13 แพ็คเกจ เรียงตามตัวอักษรภาษาอังกฤษดังรายการนี้ (แพ็คเกจต่างๆ สร้างขึ้นในระยะเวลาต่างๆ กัน)

1. epid
2. ice
3. ice.commands
4. ice.datamain
5. ice.dataman
6. ice.epid
7. ice.fullmain
8. ice.graph
9. ice.help
10. ice.main
11. ice.objects
12. ice.statis
13. ice.summary

แพ็คเกจ ice.datamain เป็นการรวม ice.main กับ ice.dataman และแพ็คเกจ ice.fullmain เป็นการรวม ice.main, ice.dataman, ice.summary, ice.statis และ ice.graph เข้าด้วยกัน เพื่อให้เป็นระบบเมนูรวมเพื่อจะได้ติดตั้งได้ง่ายด้วยคำสั่งเรียกใช้งานเพียงคำสั่งเดียว

สรุปแนวคิดการทำงานของ **R-ICE** เป็นแผนภาพจะทำให้เกิดความเข้าใจกระบวนการทำงานในส่วนของ **ICE** ได้ง่ายขึ้น จึงจะสรุปให้เห็นเป็นขั้นตอนดังนี้

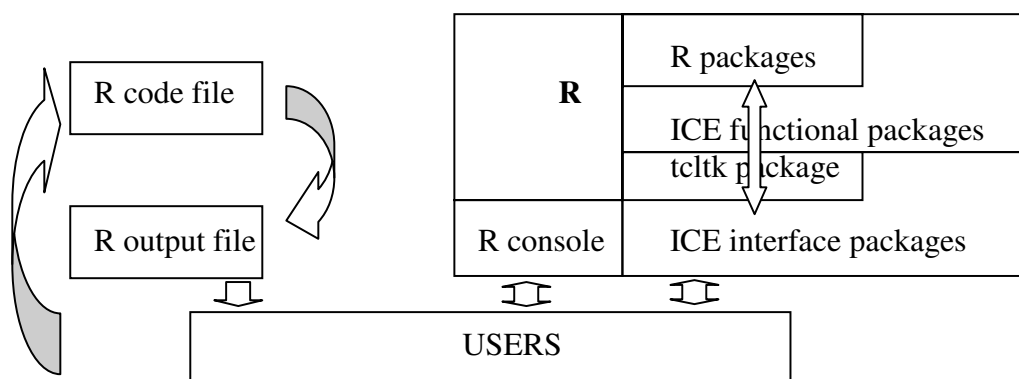
ในการทำงานกับ **R** ตามปกตินั้น จะมีแผนภาพการทำงานดังนี้



รูปที่ 3 กระบวนการทำงานวิเคราะห์ทางสถิติด้วยโปรแกรม **R** ตามปกติ

นั่นคือผู้ใช้ทำงานกับ **R** ผ่านทาง **R console** และ **R** อาจทำงานกับแพ็คเกจต่างๆ ที่ผู้ใช้ต้องการเรียกใช้งานเพิ่มเติม อาจทำงานโดยตรง หรือทำงานผ่านการเขียนเพิ่ม **R code** ก็ได้ นี่เป็นวิธีทำงานตามปกติที่ผู้ใช้ส่วนใหญ่ใช้งานอยู่ แม้แต่ผู้ใช้ EpiCan ก็ทำงานเช่นนี้เหมือนกัน

แต่ในสภาพแวดล้อม **R-ICE** จะมีส่วนเพิ่มเติมดังนี้



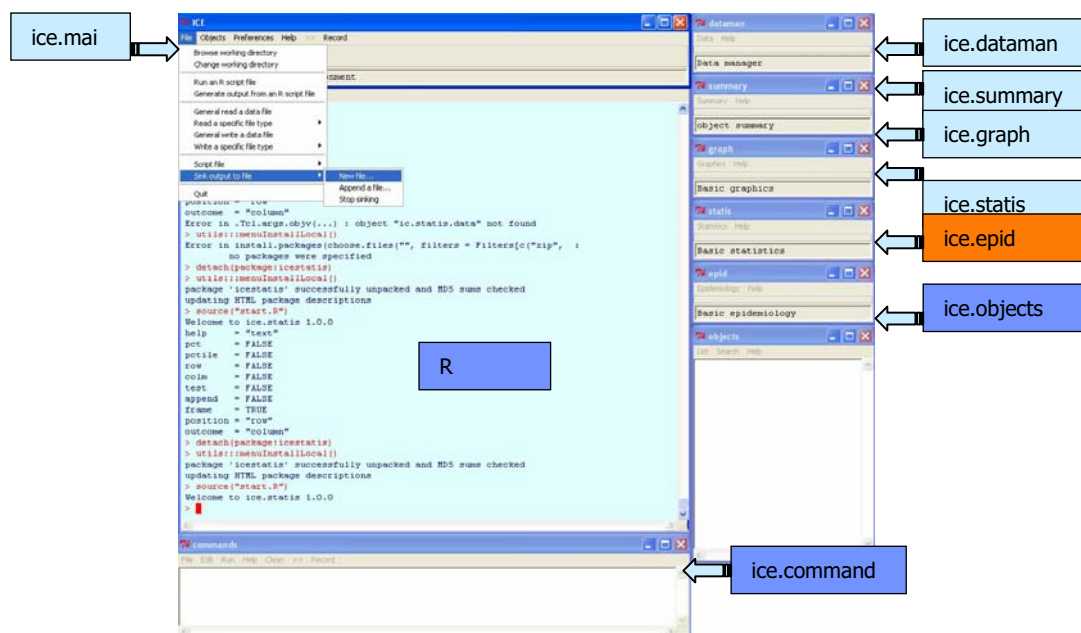
รูปที่ 4 กระบวนการทำงานวิเคราะห์ทางสถิติด้วยสภาพแวดล้อม **R-ICE**

จะเห็นว่ามีส่วนของ **ICE interface packages** มาทำงานเป็นตัวติดต่อกับผู้ใช้ควบคู่ไปกับ **R console** คือมีส่วนของรายการเมนูและหน้าต่างได้ตอบให้ใช้งานได้สะดวกขึ้น และ **ICE interface packages** นี้จะทำงานผ่าน **tcltk package** ในการติดต่อกับ **R** และยังสามารถเรียกใช้ **R packages** และ **ICE functional packages** ต่างๆ ได้อีกด้วย **ICE interface packages** ทำงานร่วมกับ **R console** โดยการส่งบรรทัดคำสั่งที่สร้างจากการที่ผู้ใช้เลือกรายการเมนูและกรอกตัวเลือกต่างๆ เสร็จแล้วไปทำงานกับ **R** ผ่าน **R console** อีกต่อหนึ่ง โดยผลการทำงานจะแสดงบน **R console** นั้นเอง

จึงเห็นได้ว่าในสภาพแวดล้อมนี้ **ICE** กับ **R** ทำงานร่วมกันตลอดเวลา โดยที่ **ICE** ไม่ได้ซ่อนเร้น **R** ไว้เบื้องหลังแต่อย่างใด ซึ่งต่างจากแพ็คเกจบางแพ็คเกจที่เป็น R-GUI เช่นกัน แต่ **R** และ R console กลับไม่มีบทบาทอะไรเลย คือถูก GUI ทำงานรับหน้าไว้ทั้งหมดทั้งการรับคำสั่งและแสดงผล จึงเรียกสภาพแวดล้อมนี้เป็น **R-ICE** คือทั้งสองส่วนทำงานร่วมกัน

ค. รูปร่างหน้าตาของ **R-ICE**

ในที่นี้จะได้แสดงหน้าตาของหน้าจอ **R-ICE** ให้เห็น เพื่อจะได้เข้าใจลักษณะการทำงานได้ชัดเจนขึ้น ดังรูปที่ 5



รูปที่ 5 หน้าจอการทำงานของสภาพแวดล้อม **R-ICE** บนระบบปฏิบัติการวินโดวส์

ในรูปที่ 5 จะเห็นว่าหน้าจอ R console ยังคงเป็นศูนย์กลางของการทำงานในสภาพแวดล้อม **R-ICE** ซึ่งจะต่างจากแพ็คเกจ R Commander ซึ่ง R console ไม่มีความจำเป็นแต่อย่างใด เพราะ R Commander จะมีหน้าจอแสดงผลการทำงานของตัวเองอยู่แล้ว ด้านบนสุดเป็นรายการเมนูของแพ็คเกจ ice.main ด้านขวามือเป็นรายการเมนูของแพ็คเกจต่างๆ ที่เป็นส่วนประกอบของ **ICE** และมีหน้าต่าง ice.objects และ ice.commands อยู่ด้านล่าง

นั่นคือเมื่อออกคำสั่งรายการเมนูจาก ice.main หรือรายการเมนูอื่นๆ ของ **ICE** เมนูต่างๆ เหล่านี้จะส่งคำสั่งไปแสดงที่ R console และส่งคำสั่งไปทำงานที่ **R** เมื่อได้ผลลัพธ์การทำงาน **R** จะแสดงผลบนหน้าจอ R console จะเห็นว่าโดยวิธีนี้ผู้วิจัยไม่จำเป็นต้องเขียนหน้าต่างเพื่อรับผลของการทำงานของ **R** แต่อย่างใด และผู้ใช้สามารถสลับการทำงานไปที่ R console ได้ตามใจ

ส่วนคำสั่งที่รายการเมนูต่างๆ ของ **ICE** เรียกทำงานนั้น อาจเป็นฟังก์ชันพื้นฐานของ **R** เอง ฟังก์ชันในแพ็คเกจ ice หรือฟังก์ชันในแพ็คเกจอื่นๆ ก็ได้

ตัวอย่างแฟ้มต่างๆ ใน ICE interface packages

เนื่องจากฟังก์ชันในแพ็คเกจต่างๆ มีเป็นจำนวนมาก ในที่นี้จะกล่าวเป็นตัวอย่างเฉพาะรายการฟังก์ชันในแพ็คเกจ ice.dataman ซึ่งถูกจัดอยู่ในแฟ้มต่างๆ และฟังก์ชันใน ice หรือ base (ของ R) ดังนี้

ชื่อแฟ้ม (ซึ่งตรงกับชื่อฟังก์ชันด้วย)	ฟังก์ชันที่ถูกเรียก	จากแพ็คเกจ
- dataman.change.R	change	ice
- dataman.create.data.R	create	ice
- dataman.duplicate.data.R	<-	base
- dataman.help.R	help, example	base
- dataman.label.R	label	ice
- dataman.recode.na.R	recode.na	ice
- dataman.recode.R	recode	ice
- dataman.relabel.R	relabel	ice
- dataman.rename.data.R	<-, rm	base
- dataman.rename.variable.R	rename	ice
- dataman.sort.data.R	sort.data	ice
- dataman.subset.data.R	subset	base
- dataman.to.character.R	to.character	ice
- dataman.to.factor.R	to.factor	ice
- dataman.to.numeric.R	to.numeric	ice
- dataman.transform.R	transform	base
- ice.dataman.R	-	-

จะเห็นว่าแพ็คเกจนี้ทำงานเกี่ยวกับการจัดการข้อมูลและกรอข้อมูล ซึ่งฟังก์ชันรายการเมนูจะไปเรียกใช้ฟังก์ชันทำงานในแพ็คเกจอื่นอีกทอดหนึ่ง ซึ่งในกรณีนี้ส่วนใหญ่เรียกจากแพ็คเกจ ice แต่ก็มีบางรายการเรียกใช้ฟังก์ชันในแพ็คเกจ base ของ R โดยตรง

โครงสร้างโฟลเดอร์ของแพ็คเกจ

ในการศึกษาการสร้างและ compile แพ็คเกจให้ใช้งานได้ตั้งแต่เริ่มโครงการนั้น ผู้วิจัยได้ศึกษาด้วยตนเองเป็นหลัก โดยการอ่านเอกสารของ R บทความในอินเทอร์เน็ต และที่สำคัญคือการศึกษาจากแพ็คเกจที่มีคนสร้างมาแล้ว แต่เนื่องจาก R เป็นโปรแกรม open source ซึ่งมีทุนในการจัดทำเอกสารและตัวอย่างเหล่านี้จำกัด ขึ้นอยู่กับการอุทิศเวลาส่วนตัวของผู้พัฒนาแต่ละคนจะเขียนขึ้น จึงหาบทความอ้างอิงให้เรียนรู้ได้ยาก จึงต้องใช้เวลาในการเรียนรู้ค่อนข้างนาน แต่เนื่องจากหัวหน้าโครงการ (รศ. นพ. หัซซา ศรีปลั่ง) มีพื้นฐานการเขียนโปรแกรมคอมพิวเตอร์ในภาษาอื่น เช่น BASIC, Pascal, C, dBase อยู่พอสมควร จึงใช้พื้นฐานเหล่านั้นมาประยุกต์ใช้ จึงสามารถเขียนฟังก์ชันและทำแพ็คเกจได้ในที่สุด

สรุปโครงสร้างโฟลเดอร์เพื่อการทำแพ็คเกจดังนี้

Package_name

- DESCRIPTION
- NAMESPACE
- R (R function folder)
 - o Function files (*function_name.R*)
- man (manual folder)
 - o Documentation files (*function_name.Rd*)

- po (language translation folder)
 - o *R-package_name.pot*
 - o *R-th.po*
 - o (Translation files for other languages)
- inst
 - o po (language translation folder)
 - th (Thai)
 - LC_MESSAGES (locale messages)
 - o *R-package_name.mo* (binary language translation file)
 - (Binary translation folder/files for other languages)
 - o (Other folder/files)

ในส่วนของการแปลภาษาอังกฤษเป็นภาษาอื่นๆ (internationalization) ซึ่งเป็นการทำงานของโฟลเดอร์ po และ inst เป็นส่วนที่ได้จากคำแนะนำของ Uwe Ligges ซึ่งเป็นหนึ่งใน **R** core team ที่ได้สนทนาคู่ในงานประชุมนานาชาติ UseR 2006 ซึ่งหัวหน้าโครงการฯ (รศ. นพ. หัซซา ศรีปลั่ง) ศ. นพ. วีระศักดิ์ จงสู่วิวัฒนวงศ์ และ Edward McNeil ได้มีโอกาสเข้าร่วมประชุมด้วย ซึ่งได้ใช้โอกาสนี้รายงานโครงการนี้ด้วย เมื่อเดือนมิถุนายน 2549 ที่กรุงเวียนนา ประเทศออสเตรีย แต่เนื่องจากเอกสารที่แนะนำการสร้างโฟลเดอร์และแฟ้มต่างๆ ในการทำ internationalization นั้นเขียนไว้สั้นมาก และไม่มีตัวอย่างการสร้างเป็นขั้นๆ จึงต้องศึกษาเพิ่มเติมโดยการแกะรอยการทำงานของแพ็คเกจ R Commander (Rcmdr) ของ John Fox (<http://socserv.mcmaster.ca/jfox/Misc/Rcmdr/>) ซึ่งเป็นแพ็คเกจ R-GUI อีกแพ็คเกจหนึ่ง ซึ่งมีส่วนของ internationalization เช่นกัน แต่แพ็คเกจนี้มีแนวคิดการทำงานต่างกับ **R-ICE**

เพื่อให้เห็นการทำงานของแฟ้มต่างๆ จะได้ยกตัวอย่างแฟ้มในแพ็คเกจ ice.dataman ที่เป็นส่วนหนึ่งของสภาพแวดล้อม **ICE** ที่สร้างขึ้นดังต่อไปนี้

ตัวอย่างแฟ้มต่างๆ ในแพ็คเกจ ice.dataman รุ่น 1.1-1

DESCRIPTION

เป็นแฟ้มอธิบายแพ็คเกจ ประกอบด้วยชื่อ รุ่น วันที่สร้าง และข้อมูลทางเทคนิคอื่นๆ

```
Package: ice.dataman
Version: 1.1-1
Date: 2007-01-03
Title: ICE Data Manager
Author: Hutcha Sriplung
Maintainer: Hutcha Sriplung <hutcha.s@psu.ac.th>
Depends: R (>= 2.3.0), ice
Description: Functions for data management
License: GPL Version 2 or later
URL: http://www.r-ice.org
```

NAMESPACE

เป็นแฟ้มบอกชื่อฟังก์ชันที่จะส่งผ่านออกจากแพ็คเกจ เพื่อให้แพ็คเกจติดต่อกับผู้ใช้งาน **R** ได้

```
export(ice.dataman, dataman.assign,
ic.dataman.data.change.vector, ic.dataman.data.change.data,
```



```

ic.dataman.data.create.data,
ic.dataman.data.duplicate.data,
ic.help.functions.create, ic.help.functions.subset,
ic.help.functions.rename, ic.help.functions.change,
ic.help.functions.recode, ic.help.functions.as.na,
ic.help.functions.recode.na, ic.help.functions.to.character,
ic.help.functions.to.numeric, ic.help.functions.to.factor,
ic.help.functions.label, ic.help.functions.relabel,
ic.help.functions.sort.data,
ic.help.examples.create, ic.help.examples.subset,
ic.help.examples.rename, ic.help.examples.change,
ic.help.examples.recode, ic.help.examples.as.na,
ic.help.examples.recode.na, ic.help.examples.to.character,
ic.help.examples.to.numeric, ic.help.examples.to.factor,
ic.help.examples.label, ic.help.examples.relabel,
ic.help.examples.sort.data,
ic.help.dataman.about.text, ic.help.dataman.about.html,
ic.help.dataman.about.chm, ic.help.dataman.package,
ic.dataman.data.label.vector, ic.dataman.data.label.data,
ic.dataman.data.recode.na.vector, ic.dataman.data.recode.na.data,
ic.dataman.data.recode.vector, ic.dataman.data.recode.data,
ic.dataman.data.relabel.vector, ic.dataman.data.relabel.data,
ic.dataman.data.rename.data,
ic.dataman.data.rename.variable,
ic.dataman.data.sort,
ic.dataman.data.subset.data,
ic.dataman.data.to.char.vector, ic.dataman.data.to.char.data,
ic.dataman.data.to.fact.vector, ic.dataman.data.to.fact.data,
ic.dataman.data.to.num.vector, ic.dataman.data.to.num.data,
ic.dataman.data.transform)

```

ice.dataman.R

เป็นแฟ้มสร้างรายการเมนูของแพคเกจ ice.dataman

```

# Define constants.
dataman.assign <- function ()
{
  assign(".ic.dataman.width",
    ifelse(exists(".ic.dataman.width",envir=GlobalEnv),
      get(".ic.dataman.width",.GlobalEnv),20),
    envir=.GlobalEnv)
  assign(".ic.dataman.font",
    ifelse(exists(".ic.dataman.font",envir=.GlobalEnv),
      get(".ic.dataman.font",.GlobalEnv),
      ifelse(.Platform$OS.type=="windows","courier 10","courier 11")),
    envir=.GlobalEnv)
  assign(".ic.dataman.prompt",
    ifelse(exists(".ic.dataman.prompt",envir=.GlobalEnv),
      get(".ic.dataman.prompt",.GlobalEnv),"dataman>>"),
    envir=.GlobalEnv)
}

## ice.dataman
"ice.dataman" <- function ()
{
  require(tcltk)
  require(ice)
  dataman.assign()

# Initialize window object.
w.dataman <- tkoplevel()
tkwm.title(w.dataman,"dataman")
dataman.statusbar <- tkframe(w.dataman)
dataman.statusbar.label <- tklablel(dataman.statusbar, text =
  gettext("Data manager", domain="R-ice.dataman"),

```

```

font=get(".ic.dataman.font",.GlobalEnv), width=get(".ic.dataman.width",
.GlobalEnv)+2,relief = "sunken", anchor = "w")
tkpack(dataman.statusbar.label, side = "left", padx = 2, expand = "yes",
fill = "both")
tkpack(dataman.statusbar, side = "bottom", fill = "x", pady = 2)

# Set dataman menus.
top.menu <- tkmenu(w.dataman, tearoff=FALSE)
tkconfigure(w.dataman, menu = top.menu)
dataman.data.menu <- tkmenu(top.menu, tearoff=FALSE)
dataman.help.menu <- tkmenu(top.menu, tearoff=FALSE)
ic.dataman.data <- tkmenu(top.menu, tearoff=FALSE)
ic.dataman.data.change <- tkmenu(top.menu, tearoff=FALSE)
ic.dataman.data.recode <- tkmenu(top.menu, tearoff=FALSE)
ic.dataman.data.recode.na <- tkmenu(top.menu, tearoff=FALSE)
ic.dataman.data.to.char <- tkmenu(top.menu, tearoff=FALSE)
ic.dataman.data.to.num <- tkmenu(top.menu, tearoff=FALSE)
ic.dataman.data.to.fact <- tkmenu(top.menu, tearoff=FALSE)
ic.dataman.data.label <- tkmenu(top.menu, tearoff=FALSE)
ic.dataman.data.relabel <- tkmenu(top.menu, tearoff=FALSE)
ic.help.functions <- tkmenu(top.menu, tearoff=FALSE)
ic.help.examples <- tkmenu(top.menu, tearoff=FALSE)
ic.help.dataman.about <- tkmenu(top.menu, tearoff=FALSE)

# Set top level menus.
tkadd(top.menu, "cascade", label=gettext("Data",domain="R-ice.dataman"),
menu=dataman.data.menu)
tkadd(top.menu, "cascade", label=gettext("Help",domain="R-ice.dataman"),
menu=dataman.help.menu)

# Set components of Data menu.
tkadd(dataman.data.menu, "command", label=gettext("Create a blank data frame",
domain="R-ice.dataman"), command=ic.dataman.data.create.data)
tkadd(dataman.data.menu, "command", label=gettext("Duplicate a data frame",
domain="R-ice.dataman"), command=ic.dataman.data.duplicate.data)
tkadd(dataman.data.menu, "command", label=gettext("Rename a data frame",
domain="R-ice.dataman"), command=ic.dataman.data.rename.data)
tkadd(dataman.data.menu, "command", label=gettext("Subset a data frame",
domain="R-ice.dataman"), command=ic.dataman.data.subset.data)
tkadd(dataman.data.menu, "separator")
tkadd(dataman.data.menu, "command", label=gettext("Rename variables",
domain="R-ice.dataman"), command=ic.dataman.data.rename.variable)
tkadd(dataman.data.menu, "command", label=gettext("Transform a data frame",
domain="R-ice.dataman"), command=ic.dataman.data.transform)
tkadd(dataman.data.menu, "cascade", label=gettext("Change values",
domain="R-ice.dataman"), menu=ic.dataman.data.change)
tkadd(ic.dataman.data.change, "command", label=gettext("in a vector",
domain="R-ice.dataman"), command=ic.dataman.data.change.vector)
tkadd(ic.dataman.data.change, "command", label=gettext("within a data frame",
domain="R-ice.dataman"), command=ic.dataman.data.change.data)
tkadd(dataman.data.menu, "cascade", label=gettext("Recode values",
domain="R-ice.dataman"), menu=ic.dataman.data.recode)
tkadd(ic.dataman.data.recode, "command", label=gettext("in a vector",
domain="R-ice.dataman"), command=ic.dataman.data.recode.vector)
tkadd(ic.dataman.data.recode, "command", label=gettext("within a data frame",
domain="R-ice.dataman"), command=ic.dataman.data.recode.data)
tkadd(dataman.data.menu, "cascade", label=gettext("Recode NA",
domain="R-ice.dataman"), menu=ic.dataman.data.recode.na)
tkadd(ic.dataman.data.recode.na, "command", label=gettext("in a vector",
domain="R-ice.dataman"), command=ic.dataman.data.recode.na.vector)
tkadd(ic.dataman.data.recode.na, "command", label=gettext(
"within a data frame",domain="R-ice.dataman"), command=
ic.dataman.data.recode.na.data)
tkadd(dataman.data.menu, "separator")
tkadd(dataman.data.menu, "cascade", label=gettext("Factor to character",
domain="R-ice.dataman"), menu=ic.dataman.data.to.char)
tkadd(ic.dataman.data.to.char, "command", label=gettext("factor",
domain="R-ice.dataman"), command=ic.dataman.data.to.char.vector)
tkadd(ic.dataman.data.to.char, "command", label=gettext("within a data frame",
domain="R-ice.dataman"), command=ic.dataman.data.to.char.data)
tkadd(dataman.data.menu, "cascade", label=gettext("Factor to numeric",
domain="R-ice.dataman"), menu=ic.dataman.data.to.num)
tkadd(ic.dataman.data.to.num, "command", label=gettext("factor",domain=
"R-ice.dataman"), command=ic.dataman.data.to.num.vector)
tkadd(ic.dataman.data.to.num, "command", label=gettext("within a data frame",
domain="R-ice.dataman"), command=ic.dataman.data.to.num.data)
tkadd(dataman.data.menu, "cascade", label=gettext("Numeric to factor",

```

```

domain="R-ice.dataman"), menu=ic.dataman.data.to.fact)
  tkadd(ic.dataman.data.to.fact, "command", label=gettext("numeric vector",
domain="R-ice.dataman"), command=ic.dataman.data.to.fact.vector)
  tkadd(ic.dataman.data.to.fact, "command", label=gettext("within a data frame",
domain="R-ice.dataman"), command=ic.dataman.data.to.fact.data)
  tkadd(dataman.data.menu, "cascade", label=gettext("Label factors",
domain="R-ice.dataman"), menu=ic.dataman.data.label)
  tkadd(ic.dataman.data.label, "command", label=gettext("in a factor",
domain="R-ice.dataman"), command=ic.dataman.data.label.vector)
  tkadd(ic.dataman.data.label, "command", label=gettext("within a data frame",
domain="R-ice.dataman"), command=ic.dataman.data.label.data)
  tkadd(dataman.data.menu, "cascade", label=gettext("Relabel factors",
domain="R-ice.dataman"), menu=ic.dataman.data.relabel)
  tkadd(ic.dataman.data.relabel, "command", label=gettext("in a factor",
domain="R-ice.dataman"), command=ic.dataman.data.relabel.vector)
  tkadd(ic.dataman.data.relabel, "command", label=gettext("within a data frame",
domain="R-ice.dataman"), command=ic.dataman.data.relabel.data)
  tkadd(dataman.data.menu, "separator")
  tkadd(dataman.data.menu, "command", label=gettext("Sort a data frame",
domain="R-ice.dataman"), command=ic.dataman.data.sort)

# Set components of Help menu.
  tkadd(dataman.help.menu, "cascade", label=gettext("Functions",
domain="R-ice.dataman"), menu=ic.help.functions)
  tkadd(ic.help.functions, "command", label="create", command=
ic.help.functions.create)
  tkadd(ic.help.functions, "command", label="subset", command=
ic.help.functions.subset)
  tkadd(ic.help.functions, "command", label="rename", command=
ic.help.functions.rename)
  tkadd(ic.help.functions, "command", label="change", command=
ic.help.functions.change)
  tkadd(ic.help.functions, "command", label="recode", command=
ic.help.functions.recode)
  tkadd(ic.help.functions, "command", label="as.na", command=
ic.help.functions.as.na)
  tkadd(ic.help.functions, "command", label="recode.na", command=
ic.help.functions.recode.na)
  tkadd(ic.help.functions, "command", label="to.character", command=
ic.help.functions.to.character)
  tkadd(ic.help.functions, "command", label="to.numeric", command=
ic.help.functions.to.numeric)
  tkadd(ic.help.functions, "command", label="to.factor", command=
ic.help.functions.to.factor)
  tkadd(ic.help.functions, "command", label="label", command=
ic.help.functions.label)
  tkadd(ic.help.functions, "command", label="relabel", command=
ic.help.functions.relabel)
  tkadd(ic.help.functions, "command", label="sort.data", command=
ic.help.functions.sort.data)
  tkadd(dataman.help.menu, "cascade", label=gettext("Examples",
domain="R-ice.dataman"), menu=ic.help.examples)
  tkadd(ic.help.examples, "command", label="create", command=
ic.help.examples.create)
  tkadd(ic.help.examples, "command", label="subset", command=
ic.help.examples.subset)
  tkadd(ic.help.examples, "command", label="rename", command=
ic.help.examples.rename)
  tkadd(ic.help.examples, "command", label="change", command=
ic.help.examples.change)
  tkadd(ic.help.examples, "command", label="recode", command=
ic.help.examples.recode)
  tkadd(ic.help.examples, "command", label="as.na", command=
ic.help.examples.as.na)
  tkadd(ic.help.examples, "command", label="recode.na", command=
ic.help.examples.recode.na)
  tkadd(ic.help.examples, "command", label="to.character", command=
ic.help.examples.to.character)
  tkadd(ic.help.examples, "command", label="to.numeric", command=
ic.help.examples.to.numeric)
  tkadd(ic.help.examples, "command", label="to.factor", command=
ic.help.examples.to.factor)
  tkadd(ic.help.examples, "command", label="label", command=
ic.help.examples.label)
  tkadd(ic.help.examples, "command", label="relabel", command=
ic.help.examples.relabel)
  tkadd(ic.help.examples, "command", label="sort.data", command=

```

```

ic.help.examples.sort.data)
  tkadd(dataman.help.menu, "separator")
  tkadd(dataman.help.menu, "cascade", label=gettext("About dataman window",
domain="R-ice.dataman"), menu=ic.help.dataman.about)
  tkadd(ic.help.dataman.about, "command", label="text", command=
ic.help.dataman.about.text)
  tkadd(ic.help.dataman.about, "command", label="html", command=
ic.help.dataman.about.html)
  if (.Platform$OS.type == "windows")
  tkadd(ic.help.dataman.about, "command", label="chm", command=
ic.help.dataman.about.chm)
  tkadd(dataman.help.menu, "command", label=gettext("About dataman package",
domain="R-ice.dataman"), command=ic.help.dataman.package)

assign("dataman", w.dataman, pos = 1)
frame1 <- tkframe(w.dataman, relief = "groove", borderwidth = 2)
dataman.backgrd.label.dataman <- tklabel(frame1, text = "",
anchor = "w", font = "arial 1 bold italic", justify = "left")
tkpack(frame1, dataman.backgrd.label.dataman, side = "top", fill = "x")
}

```

dataman.change.R

เป็นแฟ้มสร้างหน้าต่างรับค่า argument ต่างของฟังก์ชัน `change` ซึ่งเป็นฟังก์ชันในแฟ้ม `ice` เพื่อส่งไปทำงานใน **R** โดยสุดท้ายจะส่งบรรทัดคำสั่งให้ **R** ในรูป

```
ReturnVal <- change(data, variable, condition, value, append,
variable names)
```

โดยส่งบรรทัดคำสั่งไปทำงานใน **R** แล้วแสดงบรรทัดคำสั่งนั้นบนหน้าจอ **R console** ด้วย

```

## Functions in data menu.
ic.dataman.data.change.vector <- function ()
{
  objs <- as.vector(eval(ls(pos = 1)))
  lst <- NA
  for (i in 1:length(objs)) {
    com.line <- paste("class(",objs[i],")",sep = " ")
    cls <- eval.parent(parse(text = com.line)[[1]],n = 2)[1]
    if (cls == "numeric" | cls == "integer" | cls == "character" |
        cls == "double" | cls == "factor") lst <- c(lst,objs[i])
  }
  lst <- lst[2:length(lst)]
  if (is.na(lst[1])) {
    tkmessageBox(message=gettext("No vector in memory.", domain=
      "R-ice.dataman"))
    return()
  }
  dlg <- tktoplevel()
  tkwm.deiconify(dlg)
  tkgrab.set(dlg)
  tkfocus(dlg)
  tkwm.title(dlg,"change")
  tkgrid(tklabel(dlg,text = "          "))
  tkgrid(tklabel(dlg,text = gettext("Change values in a vector",
    domain="R-ice.dataman"))))
  tkgrid(tklabel(dlg,text = gettext("Select a vector", domain=
    "R-ice.dataman"))))
  scr <- tkscrollbar(dlg,repeatinterval=5,command=function(...)
    tkyview(var.lst,...))
  var.lst <- tklistbox(dlg,selectmode = "single", height = 5,
    font = .ic.dataman.font,exportselection = FALSE,width = 33,
    yscrollcommand=function(...)tkset(scr,...))
  tkgrid(var.lst,scr,columnspan=2)
  tkgrid.configure(scr,rowspan=4,sticky="nsw")
  for (i in 1:length(lst))
    tkinsert(var.lst,"end",lst[i])
}

```

```

tkselection.set(var.lst, 0) # The first is 0.
textEntryVarTcl <- tclVar(paste(""))
textEntryWidget <- tkentry(dlg,width = paste(20),
  textvariable = textEntryVarTcl)
tkgrid(tklabel(dlg,text = "      "))

tkgrid(tklabel(dlg,text = gettext("Return output in vector",
  domain="R-ice.dataman")),textEntryWidget)
tkgrid(tklabel(dlg,text = "      "))

onOK <- function ()
{
  dlg2 <- tkoplevel()
  tkwm.deiconify(dlg2)
  tkgrab.set(dlg2)
  tkfocus(dlg2)
  tkwm.title(dlg2,gettext("Change a value", domain="R-ice.dataman"))
  tkgrid(tklabel(dlg2,text = "      "))
  tkgrid(tklabel(dlg2,text = gettext("Condition and value",
    domain="R-ice.dataman"))))

  textEntryVarTcl1 <- tclVar(paste(""))
  textEntryWidget1 <- tkentry(dlg2,width = paste(25),
    textvariable = textEntryVarTcl1)
  textEntryVarTcl2 <- tclVar(paste(""))
  textEntryWidget2 <- tkentry(dlg2,width = paste(25),
    textvariable = textEntryVarTcl2)
  tkgrid(tklabel(dlg2,text = gettext("Condition (such as v1>1)",
    domain="R-ice.dataman")),textEntryWidget1)
  tkgrid(tklabel(dlg2,text = gettext("Value", domain="R-ice.dataman")),
    textEntryWidget2)
  tkgrid(tklabel(dlg2,text = "      "))

  onOK2 <- function ()
  {
    var.selected <- lst[as.numeric(tkcurselection(var.lst))+1]
    ReturnVal <- tclvalue(textEntryVarTcl)
    if (ReturnVal=="")
      ReturnVal <- data.selected

    condi <- tclvalue(textEntryVarTcl1)
    value <- tclvalue(textEntryVarTcl2)

    com.line <- paste(ReturnVal," <- change(",var.selected," ", " ,condi," ", " ,
      value," )",sep="")
    cat(.ic.dataman.prompt,"")
    print(parse(text=com.line)[[1]])
    void <- eval.parent(parse(text=com.line)[[1]])
    assign(".ec.last.com.line",com.line,env=.GlobalEnv)
    tkgrab.release(dlg2)
    tkdestroy(dlg2)
    tkfocus(dlg)
  }

  onCancel2 <- function ()
  {
    tkgrab.release(dlg2)
    tkdestroy(dlg2)
    tkfocus(dlg)
  }

  OK2.but <- tkbutton(dlg2, text=paste(" ",gettext("OK", domain=
    "R-ice.dataman")," "), command=onOK2)
  Cancel2.but <- tkbutton(dlg2, text=paste(" ",gettext("Cancel", domain=
    "R-ice.dataman")," "), command=onCancel2)
  tkgrid(OK2.but,Cancel2.but)
  tkgrid(tklabel(dlg2,text="      "))

  tkfocus(dlg2)
  tkbind(dlg2,"<Destroy>",
    function() {tkgrab.release(dlg2)})
  tkwait.window(dlg2)

  tkgrab.release(dlg)
  tkdestroy(dlg)
}

```

```

onCancel <- function ()
{
  tkgrab.release(dlg)
  tkdestroy(dlg)
}

OK.but <- tkbutton(dlg, text=paste(" ",gettext("OK", domain=
  "R-ice.dataman")," "), command=onOK)
Cancel.but <- tkbutton(dlg, text=paste(" ",gettext("Cancel", domain=
  "R-ice.dataman")," "), command=onCancel)
tkgrid(OK.but,Cancel.but)
tkgrid(tklabel(dlg,text="  "))

tkfocus(dlg)
tkbind(dlg,"<Destroy>",
  function() {tkgrab.release(dlg)})
tkwait.window(dlg)
}

ic.dataman.data.change.data <- function ()
{
  objs <- as.vector(eval(ls(pos = 1)))
  lst <- NA
  for (i in 1:length(objs)) {
    com.line <- paste("class(",objs[i],")",sep = "")
    cls <- eval.parent(parse(text = com.line)[[1]],n = 2)[1]
    if (cls == "data.frame") lst <- c(lst,objs[i])
  }
  lst <- lst[2:length(lst)]
  if (is.na(lst[1])) {
    tkmessageBox(message=gettext("No data frame in memory.", domain=
      "R-ice.dataman"))
    return()
  }
  dlg <- tkoplevel()
  tkwm.deiconify(dlg)
  tkgrab.set(dlg)
  tkfocus(dlg)
  tkwm.title(dlg,"change")
  tkgrid(tklabel(dlg,text = "  "))
  tkgrid(tklabel(dlg,text = gettext("Select a data frame:", domain=
    "R-ice.dataman"))))
  scr <- tkscrollbar(dlg,repeatinterval=5,command=function(...)
    tkyview(data.lst,...))
  data.lst <- tklistbox(dlg,selectmode = "single", height = 5,
    font = .ic.dataman.font,exportselection = FALSE,width = 33,
    yscrollcommand=function(...)tkset(scr,...))
  tkgrid(data.lst,scr,columnspan=2)
  tkgrid.configure(scr,rowspan=4,sticky="nsw")
  for (i in 1:length(lst))
    tkinsert(data.lst,"end",lst[i])
  tkselection.set(data.lst, 0) # The first is 0.

  textEntryVarTcl <- tclVar(paste(""))
  textEntryWidget <- tkentry(dlg,width = paste(20),
    textvariable = textEntryVarTcl)
  tkgrid(tklabel(dlg,text = "  "))
  tkgrid(tklabel(dlg,text = gettext("Return output in data frame",
    domain="R-ice.dataman")),textEntryWidget)
  tkgrid(tklabel(dlg,text = "  "))

  onOK <- function ()
  {
    data.selected <- lst[as.numeric(tkcurselection(data.lst))+1]
    com.line <- paste("colnames(",data.selected,"",sep="")
    lst <- eval.parent(parse(text = com.line)[[1]],n = 2)
    ReturnVal <- tclvalue(textEntryVarTcl)
    if (ReturnVal=="")
      ReturnVal <- data.selected

    dlgVar <- tkoplevel()
    tkwm.deiconify(dlgVar)
    tkgrab.set(dlgVar)
    tkfocus(dlgVar)
    tkwm.title(dlgVar,"change")
    tkgrid(tklabel(dlgVar,text = "  "))
  }
}

```

```

tkgrid(tklabel(dlgVar,text = gettext("Select variables in the list",
domain="R-ice.dataman"))))
scr <- tkscrollbar(dlgVar,repeatinterval=5,command=function(...)
tkyview(var.lst,...))
var.lst <- tklistbox(dlgVar,selectmode = "multiple", height = 5,
font = .ic.dataman.font,exportselection = FALSE,width = 33,
yscrollcommand=function(...)tkset(scr,...))
tkgrid(var.lst,scr,columnspan=2)
tkgrid.configure(scr,rowspan=4,sticky="nsw")
for (i in 1:length(lst))
tkinsert(var.lst,"end",paste(ifelse(i<10," ",""),i," ",lst[i],sep=""))
tkselection.set(var.lst, -1) # The first is 0, no initial selection = -1
tkgrid(var.lst,columnspan = 2)

textEntryVarTcl2 <- tclVar(paste(""))
textEntryWidget2 <- tkentry(dlgVar,width = paste(23),
textvariable = textEntryVarTcl2)
textEntryVarTcl3 <- tclVar(paste(""))
textEntryWidget3 <- tkentry(dlgVar,width = paste(23),
textvariable = textEntryVarTcl3)
textEntryVarTcl4 <- tclVar(paste(""))
textEntryWidget4 <- tkentry(dlgVar,width = paste(23),
textvariable = textEntryVarTcl4)
tkgrid(tklabel(dlgVar,text = " "))
tkgrid(tklabel(dlgVar,text = gettext("Condition (such as v1>1)",
domain="R-ice.dataman")),textEntryWidget2)
tkgrid(tklabel(dlgVar,text = gettext("Value", domain="R-ice.dataman")),
textEntryWidget3)
tkgrid(tklabel(dlgVar,text = " "))
append1 <- tkradiobutton(dlgVar)
append2 <- tkradiobutton(dlgVar)
appendValue <- tclVar(ifelse(ifelse(exists(".ec.append",env=.GlobalEnv),
get(".ec.append",.GlobalEnv),FALSE),"TRUE","FALSE"))
tkconfigure(append1,variable=appendValue,value="TRUE")
tkconfigure(append2,variable=appendValue,value="FALSE")
tkgrid(tklabel(dlgVar,text = " "),tklabel(dlgVar,text = "TRUE"),
tklabel(dlgVar,text = "FALSE"))
tkgrid(tklabel(dlgVar,text = gettext("Append new variables", domain=
"R-ice.dataman")), append1, append2)
tkgrid(tklabel(dlgVar,text = gettext("New names", domain=
"R-ice.dataman")),textEntryWidget4)
tkgrid(tklabel(dlgVar,text = " "))

onOK2 <- function ()
{
var.selected.lst <- lst[as.numeric(tkcurselection(var.lst))+1]
var.selected <- paste("c(",var.selected.lst[1],sep="")
if (length(var.selected.lst)>1)
for (i in 2:length(var.selected.lst))
var.selected <- paste(var.selected,var.selected.lst[i],sep=",")
var.selected <- paste(var.selected,")",sep="")
var.names <- tclvalue(textEntryVarTcl4)
if (substr(var.names,1,1)!="\\")
if (length(strsplit(var.names,"")[[1]])==1)
var.names <- paste("\\",var.names,"\\",sep="")
else var.names <- as.character(strsplit(var.names,""))
else var.names <- paste("c(",as.character(var.names),")")

condi <- tclvalue(textEntryVarTcl2)
value <- tclvalue(textEntryVarTcl3)
appendVal <- as.character(tclvalue(appendValue))
if (appendVal=="TRUE")
com.line <- paste(ReturnVal," <- change(",data.selected," ",
var.selected," ",condi," ",value," append = ",appendVal,
", var.names = ",var.names,")",sep="")
else
com.line <- paste(ReturnVal," <- change(",data.selected," ",
var.selected," ",condi," ",value,")",sep="")
cat(.ic.dataman.prompt,"")
print(parse(text = com.line)[[1]])
void <- eval.parent(parse(text = com.line)[[1]])
assign(".ec.last.com.line",com.line,env=.GlobalEnv)
tkgrab.release(dlgVar)
tkdestroy(dlgVar)
tkfocus(dlg)
}

```

```

onCancel2 <- function ()
{
  tkgrab.release(dlgVar)
  tkdestroy(dlgVar)
  tkfocus(dlg)
}

OK2.but <- tkbutton(dlgVar, text=paste(" ",gettext("OK", domain=
  "R-ice.dataman")), " "), command=onOK2)
Cancel2.but <- tkbutton(dlgVar, text=paste(" ",gettext("Cancel", domain=
  "R-ice.dataman")), " "), command=onCancel2)
tkgrid(OK2.but, Cancel2.but)
tkgrid(tklabel(dlgVar, text="  "))

tkfocus(dlgVar)
tkbind(dlgVar, "<Destroy>", function() {tkgrab.release(dlgVar)})
tkbind(textEntryWidget, "<Return>", onOK2)
tkwait.window(dlgVar)

tkgrab.release(dlg)
tkdestroy(dlg)
}

onCancel <- function ()
{
  tkgrab.release(dlg)
  tkdestroy(dlg)
}

OK.but <- tkbutton(dlg, text=paste(" ",gettext("OK", domain=
  "R-ice.dataman")), " "), command=onOK)
Cancel.but <- tkbutton(dlg, text=paste(" ",gettext("Cancel", domain=
  "R-ice.dataman")), " "), command=onCancel)
tkgrid(OK.but, Cancel.but)
tkgrid(tklabel(dlg, text="  "))

tkfocus(dlg)
tkbind(dlg, "<Destroy>", function() {tkgrab.release(dlg)})
tkbind(textEntryWidget, "<Return>", onOK)
tkwait.window(dlg)
}

```

dataman.create.data.R

เป็นแฟ้มสร้างหน้าต่างรับชื่อกรอบข้อมูลที่ใช้ต้องการจะสร้าง แล้วนำไปสร้างข้อมูลเปล่าโดยใช้ editor ภายในของ **R** ในรูป

```
.dat.name <- create()
```

โดยส่งบรรทัดคำสั่งไปทำงานใน **R** แล้วแสดงบรรทัดคำสั่งนั้นบนหน้าจอ **R** console ด้วย

```

## Functions in data menu.
ic.dataman.data.create.data <- function ()
{
  dlg <- tktoplevel()
  tkwm.deiconify(dlg)
  tkgrab.set(dlg)
  tkfocus(dlg)
  tkwm.title(dlg, "create")
  textEntryVarTcl <- tclVar(paste("new.dat"))
  textEntryWidget <- tkentry(dlg, width = paste(30),
    textvariable = textEntryVarTcl)
  tkgrid(tklabel(dlg, text = "  "))
  tkgrid(tklabel(dlg, text = gettext("Return output in data frame", domain=
    "R-ice.dataman")), textEntryWidget)
  tkgrid(tklabel(dlg, text = "  "))
}

```



```

onOK <- function ()
{
  .dat.name <- tclvalue(textEntryVarTcl)
  com.line <- paste(.dat.name,"<- create()")
  cat(.ic.dataman.prompt,"")
  print(parse(text=com.line)[[1]])
  void <- eval.parent(parse(text=com.line)[[1]])
  assign(".ec.last.com.line",com.line,env=.GlobalEnv)
  tkgrab.release(dlg)
  tkdestroy(dlg)
}

onCancel <- function ()
{
  tkgrab.release(dlg)
  tkdestroy(dlg)
}

OK.but <- tkbutton(dlg, text=paste(" ",gettext("OK", domain=
  "R-ice.dataman")," "), command=onOK)
Cancel.but <- tkbutton(dlg, text=paste(" ",gettext("Cancel", domain=
  "R-ice.dataman")," "), command=onCancel)
tkgrid(OK.but,Cancel.but)
tkgrid(tklabel(dlg,text="  "))

tkfocus(dlg)
tkbind(dlg,"<Destroy>", function() {tkgrab.release(dlg)})
tkbind(textEntryWidget,"<Return>",onOK)
tkwait.window(dlg)
}

```

ในที่นี้จะไม่แสดงเนื้อหาในแฟ้มคำสั่งทุกแฟ้ม เนื่องจากมีความยาวมาก หากรวมกันทั้งหมดในระบบ **ICE** จะมีหลายหมื่นบรรทัด แต่ได้รวบรวมแฟ้มคำสั่งในรุ่นต่างๆ ไว้ในแผ่นซีดีส่งมาพร้อมรายงานนี้แล้วในภาคผนวก ก.

การแปลเป็นภาษาไทย

ในส่วนของการแปลเป็นภาษาไทย จะได้กล่าวถึงในสองส่วน คือการแปลรายการเมนูและข้อความในหน้าต่างโต้ตอบกับผู้ใช้เป็นภาษาไทย กับการแปลแฟ้มช่วยเหลือเป็นภาษาไทย ดังได้กล่าวแล้วในหัวข้อข้างต้น ในที่นี้จะได้แสดงตัวอย่างแฟ้มที่อยู่ในแพคเกจ ice.dataman เท่านั้น

ก. การแปลรายการเมนูและข้อความในหน้าต่างโต้ตอบกับผู้ใช้

แฟ้มที่เกี่ยวข้องกับการแปลในส่วนนี้มีสามแฟ้มคือ R-ice.dataman.pot, R-th.po และ R-ice.dataman.mo ซึ่งในกระบวนการสร้างแฟ้มเหล่านี้จะต้องทำงานต่อเนื่องกันเป็นทอดๆ โดยเริ่มต้นจากการสร้างแฟ้ม R-ice.dataman.pot (po template) ขึ้นก่อน

R-ice.dataman.pot

แฟ้ม R-ice.dataman.pot นี้อยู่ในโฟลเดอร์ ../ice.dataman/po/ (ในภาษา R ให้เขียนเครื่องหมายแยก folder เป็น / เพื่อให้เป็นกลางกับทุกระบบปฏิบัติการ และไม่ซ้ำซ้อนกับเครื่องหมาย \ ที่ใช้ระบุอักขระพิเศษ เช่น \t ที่หมายความว่า tab นั่นเอง) เป็นการรวบรวมวลีภาษาอังกฤษที่จะต้องแปลเป็นภาษาอื่น โดยจะใส่ไว้ใน msgid

```
msgid ""
```

```

msgstr ""
"Project-Id-Version: ice.dataman_1.1-1\n"
"Report-Msgid-Bugs-To: bugs@r-project.org\n"
"POT-Creation-Date: 2006-05-12 08:33\n"
"PO-Revision-Date: YEAR-MO-DA HO:MI+ZONE\n"
"Last-Translator: FULL NAME <EMAIL@ADDRESS>\n"
"Language-Team: LANGUAGE <LL@li.org>\n"
"MIME-Version: 1.0\n"
"Content-Type: text/plain; charset=CHARSET\n"
"Content-Transfer-Encoding: 8bit\n"
"Plural-Forms: 0\n"

# ice.dataman
msgid "Data manager"
msgstr ""

msgid "Data"
msgstr ""

msgid "Help"
msgstr ""

msgid "Create a blank data frame"
msgstr ""

msgid "Duplicate a data frame"
msgstr ""

msgid "Rename a data frame"
msgstr ""

msgid "Subset a data frame"
msgstr ""

msgid "Rename variables"
msgstr ""

msgid "Transform a data frame"
msgstr ""

msgid "Change values"
msgstr ""

msgid "in a vector"
msgstr ""

msgid "within a data frame"
msgstr ""

msgid "Recode values"
msgstr ""

msgid "Recode NA"
msgstr ""

msgid "Factor to character"
msgstr ""

msgid "factor"
msgstr ""

```

```

msgid "Factor to numeric"
msgstr ""

msgid "Numeric to factor"
msgstr ""

msgid "numeric vector"
msgstr ""

msgid "Label factors"
msgstr ""

msgid "in a factor"
msgstr ""

msgid "Relabel factors"
msgstr ""

msgid "Sort a data frame"
msgstr ""

msgid "Functions"
msgstr ""

msgid "Examples"
msgstr ""

msgid "About dataman window"
msgstr ""

msgid "About dataman package"
msgstr ""

# dataman.change
msgid "No vector in memory."
msgstr ""

msgid "Change values in a vector"
msgstr ""

msgid "Select a vector"
msgstr ""

msgid "Return output in vector"
msgstr ""

msgid "Change a value"
msgstr ""

msgid "Condition and value"
msgstr ""

msgid "Condition (such as v1>1)"
msgstr ""

msgid "Value"
msgstr ""

msgid "OK"

```

```

msgstr ""

msgid "Cancel"
msgstr ""

msgid "Select a data frame:"
msgstr ""

msgid "Return output in data frame"
msgstr ""

msgid "Select variables in the list"
msgstr ""

msgid "Append new variables"
msgstr ""

msgid "New names"
msgstr ""

# dataman.duplicate.data
msgid "Duplicate"
msgstr ""

msgid "New data frame name"
msgstr ""

# dataman.label
msgid "No vector/factor in memory."
msgstr ""

msgid "Label a factor"
msgstr ""

msgid "Ordered"
msgstr ""

msgid "Labels (in form 'l=lab')"
msgstr ""

msgid "Separator"
msgstr ""

msgid "No vector/factor in the data frame."
msgstr ""

# dataman.recode.na
msgid "Recode values in a vector"
msgstr ""

msgid "Recode NA to a value"
msgstr ""

msgid "New value for NA"
msgstr ""

# dataman.recode
msgid "Old and new value(s)"
msgstr ""

msgid "Old value(s)"

```

```

msgstr ""

msgid "New value(s)"
msgstr ""

# dataman.relabel
msgid "Relevel a factor"
msgstr ""

msgid "Reference level or level order"
msgstr ""

msgid "No factor in the data frame."
msgstr ""

# dataman.rename.variable
msgid "Rename Variable(s) in a Data Frame"
msgstr ""

msgid "Variables"
msgstr ""

msgid "Select variable(s)"
msgstr ""

msgid "New variable name(s)"
msgstr ""

# dataman.sort.data
msgid "Select the first sort order"
msgstr ""

msgid "Select the second sort order"
msgstr ""

msgid "Select the third sort order"
msgstr ""

msgid "Decreasing"
msgstr ""

# dataman.subset.data
msgid "Variables and Condition"
msgstr ""

msgid "or in the entry box below."
msgstr ""

msgid "Subset condition"
msgstr ""

msgid "',' = separate ':' = from:to"
msgstr ""

# dataman.to.character
msgid "Change a factor to"
msgstr ""

msgid "a character vector"
msgstr ""

```

```

# dataman.to.factor
msgid "Change a numeric vector to"
msgstr ""

msgid "a factor"
msgstr ""

# dataman.to.numeric
msgid "a numeric vector"
msgstr ""

msgid "Labels (in quotation)"
msgstr ""

msgid "No numeric vector in the data frame."
msgstr ""

# dataman.transform
msgid "List of variables"
msgstr ""

msgid "New name"
msgstr ""

msgid "Returned value"
msgstr ""

```

แฟ้ม *R-th.po*

แฟ้ม *R-th.po* อยู่ในโฟลเดอร์ `../ice.dataman/po/` นี่เป็นการรวบรวมวลีภาษาอังกฤษที่จะต้องแปลเป็นภาษาอื่น โดยจะใส่ไว้ใน `msgid` และวลีที่แปลเป็นภาษาไทยใน `msgstr` นั่นคือการแปลวลีในแฟ้ม *R-ice.dataman.pot* ที่กล่าวมาข้างต้นนั่นเอง โดยวิธีนี้ทำให้สามารถแปลเป็นภาษาเขียนต่างๆ ได้ทั่วโลก เช่น *R-fr.po* เป็นการแปลเป็นภาษาฝรั่งเศส *R-de.po* เป็นการแปลเป็นภาษาเยอรมัน *R-jp.po* เป็นการแปลเป็นภาษาญี่ปุ่น เป็นต้น

ในที่นี้จะเห็นว่าเราต้องกำหนดให้ character set เป็นรหัส unicode UTF-8 และกำหนด plural form เป็น 0 คือไม่มีพจน์ เนื่องจากภาษาไทยใช้คำเดียวกันไม่ว่าจะเป็นเอกพจน์หรือพหูพจน์

```

msgid ""
msgstr ""
"Project-Id-Version: ice.dataman_1.1-1\n"
"POT-Creation-Date: 2006-05-12 08:33\n"
"PO-Revision-Date: 2007-01-05 20:03+0700\n"
"Last-Translator: hutcha sriplung <hutcha.s@psu.ac.th>\n"
"Language-Team: \n"
"MIME-Version: 1.0\n"
"Content-Type: text/plain; charset=UTF-8\n"
"Content-Transfer-Encoding: 8bit\n"
"Plural-Forms: 0\n"

# ice.dataman
msgid "Data manager"
msgstr "จัดการข้อมูล"

```

```

msgid "Data"
msgstr "ข้อมูล"

msgid "Help"
msgstr "ช่วยเหลือ"

msgid "Create a blank data frame"
msgstr "สร้างกรอบข้อมูลเปล่า"

msgid "Duplicate a data frame"
msgstr "ทำซ้ำกรอบข้อมูล"

msgid "Rename a data frame"
msgstr "เปลี่ยนชื่อกรอบข้อมูล"

msgid "Subset a data frame"
msgstr "ทำห้ต้อยกรอบข้อมูล"

msgid "Rename variables"
msgstr "เปลี่ยนชื่อตัวแปร"

msgid "Transform a data frame"
msgstr "เปลี่ยนแปลงกรอบข้อมูล"

msgid "Change values"
msgstr "เปลี่ยนค่า"

msgid "in a vector"
msgstr "ในเวกเตอร์"

msgid "within a data frame"
msgstr "ภายในกรอบข้อมูล"

msgid "Recode values"
msgstr "เปลี่ยนรหัสค่า"

msgid "Recode NA"
msgstr "เปลี่ยนรหัส NA"

msgid "Factor to character"
msgstr "แฟกเตอร์เป็นข้อความ"

msgid "factor"
msgstr "แฟกเตอร์"

msgid "Factor to numeric"
msgstr "แฟกเตอร์เป็นตัวเลข"

msgid "Numeric to factor"

```

```

msgstr "ตัวเลขเป็นแฟกเตอร์"

msgid "numeric vector"
msgstr "เวกเตอร์ตัวเลข"

msgid "Label factors"
msgstr "ให้ป้ายชื่อแฟกเตอร์"

msgid "in a factor"
msgstr "ในแฟกเตอร์"

msgid "Relabel factors"
msgstr "เปลี่ยนป้ายชื่อแฟกเตอร์"

msgid "Sort a data frame"
msgstr "เรียงลำดับกรอบข้อมูล"

msgid "Functions"
msgstr "ฟังก์ชัน"

msgid "Examples"
msgstr "ตัวอย่าง"

msgid "About dataman window"
msgstr "เกี่ยวกับหน้าต่าง dataman"

msgid "About dataman package"
msgstr "เกี่ยวกับแพ็คเกจ dataman"

# dataman.change
msgid "No vector in memory."
msgstr "ไม่มีเวกเตอร์ในหน่วยความจำ"

msgid "Change values in a vector"
msgstr "เปลี่ยนค่าในเวกเตอร์"

msgid "Select a vector"
msgstr "เลือกเวกเตอร์"

msgid "Return output in vector"
msgstr "ส่งค่ากลับในเวกเตอร์"

msgid "Change a value"
msgstr "เปลี่ยนค่า"

msgid "Condition and value"
msgstr "เงื่อนไขและค่า"

msgid "Condition (such as v1>1)"
msgstr "เงื่อนไข (ตัวอย่างเช่น v1>1)"

```



```

msgid "Value"
msgstr "ค่า"

msgid "OK"
msgstr "ตกลง"

msgid "Cancel"
msgstr "ยกเลิก"

msgid "Select a data frame:"
msgstr "เลือกกรอบข้อมูล:"

msgid "Return output in data frame"
msgstr "ส่งผลกลับในกรอบข้อมูล"

msgid "Select variables in the list"
msgstr "เลือกตัวแปรในรายการ"

msgid "Append new variables"
msgstr "ต่อท้ายตัวแปรใหม่"

msgid "New names"
msgstr "ชื่อใหม่"

# dataman.duplicate.data
msgid "Duplicate"
msgstr "Duplicate"

msgid "New data frame name"
msgstr "ชื่อกรอบข้อมูลใหม่"

# dataman.label
msgid "No vector/factor in memory."
msgstr "ไม่มีเวกเตอร์/แฟกเตอร์ในหน่วยความจำ"

msgid "Label a factor"
msgstr "ให้ป้ายชื่อแฟกเตอร์"

msgid "Ordered"
msgstr "มีลำดับ"

msgid "Labels (in form '1=lab')"
msgstr "ป้ายชื่อ (ในรูปแบบ '1=lab') "

msgid "Separator"
msgstr "ตัวคั่นแยก"

msgid "No vector/factor in the data frame."
msgstr "ไม่มีเวกเตอร์/แฟกเตอร์ในกรอบข้อมูล"

```

```

# dataman.recode.na
msgid "Recode values in a vector"
msgstr "เปลี่ยนรหัสค่าในเวกเตอร์"

msgid "Recode NA to a value"
msgstr "เปลี่ยนรหัส NA เป็นค่า"

msgid "New value for NA"
msgstr "ค่าใหม่สำหรับ NA"

# dataman.recode
msgid "Old and new value(s) "
msgstr "ค่าเก่าและใหม่"

msgid "Old value(s) "
msgstr "ค่าเก่า"

msgid "New value(s) "
msgstr "ค่าใหม่"

# dataman.relabel
msgid "Relevel a factor"
msgstr "เปลี่ยนระดับในแฟกเตอร์"

msgid "Reference level or level order"
msgstr "ระดับอ้างอิงหรือลำดับของระดับ"

msgid "No factor in the data frame."
msgstr "ไม่มีแฟกเตอร์ในกรอบข้อมูล"

# dataman.rename.variable
msgid "Rename Variable(s) in a Data Frame"
msgstr "เปลี่ยนชื่อตัวแปรในกรอบข้อมูล"

msgid "Variables"
msgstr "ตัวแปร"

msgid "Select variable(s) "
msgstr "เลือกตัวแปร"

msgid "New variable name(s) "
msgstr "ชื่อตัวแปรใหม่"

# dataman.sort.data
msgid "Select the first sort order"
msgstr "เลือกตัวเรียงลำดับขั้นที่หนึ่ง"

msgid "Select the second sort order"
msgstr "เลือกตัวเรียงลำดับขั้นที่สอง"

msgid "Select the third sort order"

```

```

msgstr "เลือกตัวเรียงลำดับขั้นที่สาม"

msgid "Decreasing"
msgstr "ลดลง"

# dataman.subset.data
msgid "Variables and Condition"
msgstr "ตัวแปรและเงื่อนไข"

msgid "or in the entry box below."
msgstr "หรือในช่องพิมพ์ข้างล่าง"

msgid "Subset condition"
msgstr "เงื่อนไขเซตย่อย"

msgid "',' = separate ':' = from:to"
msgstr "',' = คั่นแยก ':' = จาก:ถึง"

# dataman.to.character
msgid "Change a factor to"
msgstr "เปลี่ยนแฟกเตอร์เป็น"

msgid "a character vector"
msgstr "เวกเตอร์ข้อความ"

# dataman.to.factor
msgid "Change a numeric vector to"
msgstr "เปลี่ยนเวกเตอร์ตัวเลขเป็น"

msgid "a factor"
msgstr "แฟกเตอร์"

# dataman.to.numeric
msgid "a numeric vector"
msgstr "เวกเตอร์ตัวเลข"

msgid "Labels (in quotation)"
msgstr "ป้ายชื่อ (ในเครื่องหมายคำพูด)"

msgid "No numeric vector in the data frame."
msgstr "ไม่มีเวกเตอร์ตัวเลขในกรอบข้อมูล"

# dataman.transform
msgid "List of variables"
msgstr "รายการตัวแปร"

msgid "New name"
msgstr "ชื่อใหม่"

msgid "Returned value"

```

msgstr "ค่าส่งกลับ"

แฟ้ม *R-ice.dataman.mo*

เป็นอีกแฟ้มหนึ่งที่มีความสำคัญมาก เกิดจากการ compile แฟ้ม *R-th.po* ข้างต้น แต่ไม่สามารถแสดงให้เห็นในที่นี้เนื่องจากเป็นแฟ้ม binary (แฟ้มที่แปลเป็นภาษาเครื่องแล้ว) แฟ้มนี้จะต้องใส่ไว้ในโฟลเดอร์ *../ice.dataman/inst/po/th/LC_MESSAGES/*

การ compile แฟ้ม *R-th.po*

การ compile แฟ้ม *R-th.po* นั้น ผู้วิจัยเลือกทำบนระบบปฏิบัติการวินโดวส์ โดยใช้โปรแกรม poEdit ช่วย โปรแกรมนี้เริ่มใช้ทำงานได้ตั้งแต่การสร้างแฟ้ม *R-th.po* จากแฟ้ม *R-package_name.pot* ซึ่งสร้างไว้ในโฟลเดอร์ *../package_name/po/* โดยโปรแกรมจะตรวจสอบว่ามีวลีที่เก็บมาซ้ำหรือไม่ ที่จริงในขั้นตอนนี้นี้ยังมีโปรแกรมที่ใช้ช่วยงานโดยค้นหาวลีที่ต้องการแปลที่เขียนอยู่ในฟังก์ชัน *gettext()* ในแฟ้มฟังก์ชัน (ที่มีนามสกุล *R*) แต่เนื่องจากผู้วิจัยแปลงแฟ้มเก่าซึ่งยังไม่ได้ทำกระบวนการ internationalization โดยเติม *gettext()* เข้าไป ในขณะเดิมก็จึงได้เก็บวลีไปด้วยพร้อมกัน จึงไม่เสียเวลาเพิ่มขึ้นมากนัก

เมื่อตรวจสอบแล้วว่าไม่มีวลีซ้ำกัน และได้แฟ้ม *R-th.po* แล้ว ก็แปลวลีเหล่านั้นเป็นภาษาไทย จากนั้นเปิดแฟ้ม *R-th.po* แล้วเลือกรายการบันทึก และเลือกภาษาเป็นไทย โปรแกรม poEdit จะสร้างแฟ้ม *R-th.mo* ขึ้น จากนั้นเปลี่ยนชื่อแฟ้มนี้เป็น *R-package_name.mo* แล้วย้ายไปไว้ในโฟลเดอร์ *../package_name/inst/po/th/LC_MESSAGES/* เป็นอันเสร็จขั้นตอนการ compile แฟ้ม *R-th.po* แต่เพียงเท่านี้

สำหรับทุกแพ็คเกจ จะมีแฟ้มที่เกี่ยวข้องกับการแปลรายการเมนู และข้อความในหน้าต่างโต้ตอบในลักษณะเช่นนี้ทั้งสิ้น การ compile ก็ทำเช่นเดียวกันทั้งสิ้น

ข. การแปลแฟ้มความช่วยเหลือ

ในกระบวนการเขียนแฟ้มความช่วยเหลือนั้น เริ่มต้นจากการเขียนแฟ้มความช่วยเหลือเป็นภาษาอังกฤษในรูปแบบ R documentation file (Rd) ขึ้นก่อน สำหรับแพ็คเกจ *ice.dataman* แฟ้มนี้จะต้องใส่ไว้ในโฟลเดอร์ *../ice.dataman/man* รูปแบบ Rd นั้น จะมี tag ที่จะต้องเขียนเสมอคือ

```
\name{}  
\alias{}  
\title{}  
\description{}  
\usage{}
```

ส่วน tag นอกจากนั้นจะเขียนหรือไม่เขียนก็ได้ และยังมี tag ที่ช่วยเน้นการแสดงผลข้อความอื่นๆ อีก เช่น *\bold{}*, *\code{}*, *\emph{}* เป็นต้น

ice.dataman.Rd

เป็นแฟ้มอธิบายการทำงานของแพ็คเกจและฟังก์ชัน ในที่นี้คือแพ็คเกจ *ice.dataman* เขียนในรูปแบบของ R documentation format (Rd) โดยทำเป็นภาษาอังกฤษ และจะมีแฟ้มคู่กันที่แปลเป็นภาษาไทย

บรรจุไว้ในแพ็คเกจ ice.help.th ซึ่งเป็นนวัตกรรมที่ทางคณะผู้วิจัยได้ทำให้ เพื่อความสะดวกในการสร้างแฟ้มช่วยเหลือเป็นหลายภาษา แต่ในที่นี้จะได้แสดงข้อความภายในแฟ้มทั้งในภาษาอังกฤษและภาษาไทยให้เห็นเทียบกันดังนี้

ภาษาอังกฤษ

```
\name{ice.dataman}
\alias{ice.dataman}
\title{Data Manager}
\description{
Collection of functions for data management in \code{base} and
\code{ice} packages, including vector and data frame management,
changing codes and values in variables in a data frame, and factor
labelling. It is designed for \bold{ICE} environment.
}

\usage{
ice.dataman()
}

\details{
Functions or a few command lines supporting data frame management
include creating a blank data frame, duplicating, renaming, and
subsetting a data frame. Renaming of variables in a data frame is
also available. A data frame can be transformed by row and column
conditions. Values or codes within variables in a data frame can be
changed, recoded, and NA's can be easily handled. Vector of character
and numeric types and factors can be interchanged easily.

\code{ice.dataman} menu is also included in \code{ice.datamain} and
\code{ice.fullmain} menus for convenience of use. If
\code{ice.dataman} is used as a floating menu (not including in the
\code{main} menu), the command line it creates can be recorded into a
script file by pressing \code{Record} button on the \code{main} menu
bar.
}

\note{
The project is funded by the Health Research Network, Thailand
Research Fund, and the Thailand National Health Foundation.

\emph{Web site:} \url{http://www.r-ice.org}
}

\author{
Hutcha Sriplung, Epidemiology Unit, Prince of Songkla University, Hat
Yai, Songkhla, 90110, Thailand.

\emph{e-mail:} \email{hutcha.s@psu.ac.th}
}

\keyword{misc}
```

ภาษาไทย

```
\name{ice.dataman}
```

```

\alias{ice.dataman}
\title{Data Manager}
\description{
ที่รวมฟังก์ชันสำหรับการจัดการข้อมูลในแพ็คเกจ \code{base} และ \code{ice} ประกอบด้วย
การจัดการเวกเตอร์และกรอบข้อมูล เปลี่ยนรหัสและค่าในตัวแปรในกรอบข้อมูล และการให้ป้ายแฟกเตอร์ ออกแบบ
มาเพื่อสภาพแวดล้อม \b{ICE}
}

\usage{
ice.dataman()
}

\details{
ฟังก์ชันหรือชุดบรรทัดคำสั่งเล็กๆ ที่สนับสนุนการจัดการกรอบข้อมูล ประกอบด้วยการสร้างกรอบข้อมูลเปล่า
การทำซ้ำ การเปลี่ยนชื่อ และการทำเซตย่อยกรอบข้อมูล และยังสามารถเปลี่ยนชื่อตัวแปรในกรอบข้อมูล
เปลี่ยนแปลงกรอบข้อมูลตามเงื่อนไขด้านแถวและสัณภูมิ และยังสามารถเปลี่ยนค่าหรือรหัสค่าและรหัส
จัดการค่า NA ในตัวแปรภายในกรอบข้อมูลได้อย่างง่ายดาย และยังสามารถเปลี่ยนแปลงเวกเตอร์ชนิดข้อความ
และตัวเลข และแฟกเตอร์ไปมาได้ง่าย

รายการเมนู \code{ice.dataman} ยังเป็นส่วนประกอบในรายการเมนู \code{ice.datamain}
และ \code{ice.fullmain} เพื่อความสะดวกในการใช้งาน ถ้าใช้งาน \code{ice.dataman} แบบ
รายการเมนูลอย (ไม่รวมอยู่ในเมนู \code{main}) บรรทัดคำสั่งที่ถูกสร้างขึ้นสามารถบันทึกในแฟ้มสคริปต์
ได้โดยการกดปุ่ม \code{บันทึก} บนรายการเมนู \code{main}
}

\note{
โครงการนี้ได้รับทุนจากเครือข่ายวิจัยสุขภาพ สำนักงานกองทุนสนับสนุนการวิจัย (สกว.) และ มูลนิธิ
สาธารณสุขแห่งชาติ (มสช.)
}

\emph{Web site:} \url{http://www.r-ice.org}
}

\author{
หัชชา ศรีปลั่ง หน่วยระบาดวิทยา คณะแพทยศาสตร์ มหาวิทยาลัยสงขลานครินทร์ อ. หาดใหญ่ จ. สงขลา
90110
}

\emph{e-mail:} \email{hutcha.s@psu.ac.th}
}

\keyword{misc}

```

ในการ `compile` เพิ่มความช่วยเหลือ จะทำในระหว่างการ `compile` แพ็คเกจ จึงจะนำไปกล่าวใน
หัวข้อการ `compile` แพ็คเกจข้างล่างเพียงทีเดียว

ขั้นตอนการ compile แพคเกจเพื่อใช้งานในคอมพิวเตอร์ระบบปฏิบัติการต่างๆ

เมื่อสร้างไฟล์เดอร์แพคเกจเสร็จเรียบร้อยแล้ว พร้อมกับ compile แปลภาษาเสร็จแล้ว จึงถึงขั้นตอนการ compile แพคเกจทั้งหมดต่อไป แต่เนื่องจากเครื่องคอมพิวเตอร์มีหลายระบบ และ R สามารถใช้ได้กับระบบปฏิบัติการหลักทั้งวินโดวส์ ยูนิกซ์ และแมคอินทอช OSX การ compile แพคเกจจึงต้องสนับสนุนระบบปฏิบัติการหลักทั้งสามด้วย ซึ่งทั้งหมดสามารถทำได้บนระบบปฏิบัติการใดก็ได้ ไม่จำกัดว่าเมื่อ compile บนระบบปฏิบัติการใดแล้วก็จะใช้ได้เฉพาะระบบปฏิบัติการนั้น แต่อย่างไรก็ตาม เมื่อทำแพคเกจใดแพคเกจหนึ่งขึ้นมาขึ้นนั้น จำเป็นต้องทดสอบการทำงานบนระบบปฏิบัติการทั้งสามระบบ เพื่อให้แน่ใจว่าไม่มีข้อจำกัด หรือมีการให้ผลที่ผิดเพี้ยนกันไป

แพคเกจที่แจกจ่ายกับผู้ใช้ทุกระบบร่วมกันได้แก่แพคเกจ `package_name.tar.gz` โดยการติดตั้งจะยุ่งยากสำหรับระบบปฏิบัติการวินโดวส์ เนื่องจากถ้าจะติดตั้งจากแพคเกจนี้ จะต้องติดตั้งในหน้าต่าง command เท่านั้น แต่สำหรับผู้ที่ใช้ระบบยูนิกซ์จะไม่เป็นการยุ่งยากแต่อย่างใด เนื่องจากจะต้องใช้งาน R จากหน้าจอ terminal อยู่แล้ว เพียงออกคำสั่ง

R CMD INSTALL `package_name.tar.gz`

ก็จะสามารถติดตั้งแพคเกจนั้นได้เลย เพียงแต่มีข้อจำกัดว่าผู้ติดตั้งจะต้องมีสิทธิ์การใช้งานเป็น root หรือ administrator สำหรับผู้ใช้เครื่องแมคอินทอช OSX ก็สามารถเรียกหน้าต่าง terminal มาติดตั้งโดยวิธีเดียวกันนี้ได้ หรือจะใช้เมนู Package install แล้วเลือก package source ก็จะได้เช่นเดียวกัน ส่วนผู้ที่ใช้ระบบวินโดวส์มักจะนิยมใช้แพคเกจ `package_name.zip` มากกว่า เพราะสามารถติดตั้งจากรายการเมนูบน RGUI ได้

ในการ compile แพคเกจนั้น สำหรับระบบปฏิบัติการวินโดวส์ จำเป็นต้องติดตั้งโปรแกรมเพิ่มเติมอีกสามโปรแกรม คือ tools.zip เป็นชุดโปรแกรมเล็กๆ ที่ทำงานคำสั่งสคริปต์ของ UNIX ในระบบปฏิบัติการวินโดวส์ ซึ่งสามารถดาวน์โหลดได้จากเว็บไซต์ Building R for Windows ของ Duncan Murdoch ซึ่งเป็นหนึ่งใน R core team ที่ <http://www.murdoch-sutherland.com/Rtools/>, HTML Help Workshop จาก Microsoft ที่ <http://www.microsoft.com/downloads/> เพื่อใช้สร้างแพคเกจช่วยเหลือชนิด chhtml ซึ่งใช้เฉพาะในระบบปฏิบัติการวินโดวส์ และส่วนที่สำคัญมากคือ ActivePerl ซึ่งจะเปิดตัวแปล Perl script ที่โปรแกรม R CMD จะเรียกใช้ในระหว่างการ compile ซึ่งดาวน์โหลดได้ที่ <http://www.activestate.com/Products/ActivePerl/Download.html> โปรแกรมทั้งหมดนี้ดาวน์โหลดได้ฟรีโดยไม่เสียค่าใช้จ่าย

อย่างไรก็ตาม การศึกษาขั้นตอนเหล่านี้ทำได้ค่อนข้างยาก เนื่องจากไม่มีเขียนไว้ในเอกสารทางการของ R แต่อย่างใด ผู้วิจัยได้ศึกษาจากเว็บไซต์ Building R for Windows ของ Duncan Murdoch นั่นเอง

ส่วนการ compile บนเครื่องระบบลินุกซ์และแมคอินทอชนั้น สามารถทำขั้นตอนต่อไปนี้ได้เลย โดยไม่ต้องติดตั้งโปรแกรมใดๆ เพิ่มเติม เนื่องจากเป็นระบบ UNIX อยู่แล้ว (ขั้นตอนการ compile แพคเกจของ R ทำตามระบบปฏิบัติการ UNIX)

กระบวนการ compile แพคเกจสำหรับ *R*

สำหรับระบบปฏิบัติการวินโดวส์ เมื่อ unzip tools.zip แล้ว ให้ย้ายทุกโปรแกรมในโฟลเดอร์ย่อย `..\tools\bin` ไปไว้ที่ `../R/R-2.x.x/bin/` เมื่อติดตั้งโปรแกรมทุกอย่างตามที่กล่าวข้างต้นหมดแล้ว ให้คัดลอกโฟลเดอร์ที่ได้เตรียมไว้แล้วนั้น ไปไว้ที่ `../R/R-2.x.x/bin/` แล้วเปิด command window ขึ้น แล้วย้าย directory ไปอยู่ที่เดียวกันนั้น

บนระบบลินุกซ์ควร login เป็น root แล้วคัดลอกโฟลเดอร์แพคเกจไปไว้ที่ home แล้วเปิด terminal window

ผู้ใช้เครื่องระบบแมคอินทอช ให้นำโฟลเดอร์แพคเกจไปไว้ที่ home เช่นกัน แล้วเปิด terminal window เช่นเดียวกับผู้ใช้ลินุกซ์

เมื่อเตรียมการเสร็จเรียบร้อยแล้วขั้นตอนการทำงานจะเหมือนกันบนทุกระบบปฏิบัติการ โดยพิมพ์คำสั่ง

R CMD check *package_name*

เพื่อตรวจสอบว่าแพคเกจที่เตรียมไว้มีข้อผิดพลาดใดๆ หรือไม่ หากมีรายงานข้อผิดพลาด ให้แก้ไขตามนั้นให้เสร็จก่อนทำขั้นตอนต่อไป เมื่อไม่มีข้อผิดพลาดแล้ว จึงพิมพ์คำสั่ง

R CMD build *package_name*

ซึ่งจะได้เพิ่ม *package_name_version.tar.gz* คือเพิ่มบีบอัด source code ที่จะให้ติดตั้งแพคเกจบนระบบลินุกซ์และแมคอินทอช ส่วนการทำเพิ่มติดตั้งสำหรับวินโดวส์ ต้องออกคำสั่งเพิ่มอีกหนึ่งบรรทัดคือ

R CMD build --binary *package_name*

ซึ่งจะได้เพิ่ม *package_name_version.zip*

อย่างไรก็ตามผู้วิจัยพบว่าวิธีนี้มีรายงานข้อผิดพลาดขึ้นเสมอ แม้ที่จริงจะไม่มีข้อผิดพลาดในการสร้างแพคเกจก็ตาม แต่เป็นข้อผิดพลาดในการอ้างอิงตัวแปรใน Perl script เอง ซึ่งเราไม่ควรเข้าไปแก้ไข และนอกจากนี้ ถ้าทำบนระบบแมคอินทอช จะไม่ได้เพิ่มดังกล่าว แต่จะได้เพิ่ม binary สำหรับแมคอินทอชแทน และหาก compile บนลินุกซ์หรือแมคอินทอช จะไม่มีเพิ่มช่วยเหลือแบบ chtml เกิดขึ้น จึงควรทำบนวินโดวส์จะดีกว่า ผู้วิจัยพบว่าการใช้โปรแกรม zip ทั่วไป เช่น WinZip ในการสร้างแพคเกจนี้ จะเป็นทางออกที่ดีกว่า ผู้วิจัยแนะนำให้ใช้โปรแกรม 7-zip.exe ซึ่งเป็นโปรแกรม zip ที่เป็น open source จะดีที่สุด เนื่องจากไม่ต้องละเมิดลิขสิทธิ์ในการใช้งานแต่อย่างใด ดาวน์โหลดได้ที่ <http://www.7-zip.org/>

การ zip โฟลเดอร์เองนี้ จะต้องใช้โฟลเดอร์ *package_name* ภายในโฟลเดอร์ *package_name.Rcheck* ที่สร้างขึ้นจากการออกคำสั่ง R CMD check *package_name* ห้ามใช้โฟลเดอร์ *package_name* ที่เป็นต้นฉบับเป็นอันขาด โดยการ zip โฟลเดอร์นั้น จะต้อง zip โดยไม่เก็บข้อมูล path เข้าไปด้วย เมื่อได้เพิ่ม *package_name.zip* แล้ว ให้เปลี่ยนชื่อโดยเติมรุ่นเข้าไปด้วย เป็น *package_name_version.zip* ก็จะได้เพิ่มที่ติดตั้งบนระบบวินโดวส์ได้ตามต้องการ

ลิขสิทธิ์ของแพ็คเกจ **ICE**

เนื่องจาก **R** เป็นซอฟต์แวร์แบบ open source ที่มีลิขสิทธิ์แบบ GNU general public license รุ่นที่ 2 (GPL 2) แพ็คเกจต่างๆ ของ **R** ก็จะใช้ลิขสิทธิ์นี้ตามไปด้วย นั่นคือแพ็คเกจต่างๆ ในระบบ **R-ICE** จะอนุญาตให้ผู้ใดก็ได้ใช้งานได้โดยไม่ต้องซื้อหรือขออนุญาต สามารถทำซ้ำ และแจกจ่าย หรือจำหน่ายได้ โดยไม่เอาค่าใด ผู้พัฒนาซอฟต์แวร์ที่เป็น open source เช่นนี้ อย่างไรก็ตามผู้พัฒนาซอฟต์แวร์สามารถรับบริจาคค่าจ้าง หรือได้ผลประโยชน์ทางการเงินโดยการขายคู่มือการใช้งาน การอบรม หรือกิจกรรมอื่นๆ ที่เกี่ยวข้องได้ ซึ่งผู้พัฒนาซอฟต์แวร์จำนวนไม่น้อยก็ใช้วิธีนี้

การเผยแพร่แพ็คเกจ **ICE**สู่สาธารณะ

สำหรับการเผยแพร่แพ็คเกจ **ICE** นั้น ผู้วิจัยได้สร้างเว็บไซต์ชื่อ **R-ICE** ขึ้น ที่ <http://www.r-ice.org/> ซึ่งเป็นที่รวบรวมแพ็คเกจต่างๆ ในระบบ **ICE** โดยผู้ใช้งานสามารถดาวน์โหลดไปใช้งานได้ฟรี และยังมีการเปิดให้ถามคำถาม สนทนา และเปลี่ยน

ความคิดเกี่ยวกับการใช้งานอีกด้วย โดยแพ็คเกจต่างๆ บนเว็บไซต์นี้เป็นรุ่นที่ทันสมัยอยู่เสมอ แต่รุ่นเก่าๆ ก็ยังได้เก็บไว้ด้วยเช่นกัน แต่ส่วนใหญ่ไม่จำเป็นต้องใช้เนื่องจากรุ่นใหม่ได้แก้ไขผิดพลาดของรุ่นเก่า และฟังก์ชันในรุ่นใหม่ก็ใช้กับแฟมิลีที่เขียนขึ้นจากฟังก์ชันในรุ่นเก่าได้

6. การหาสภาพแวดล้อมเสริมในการวิเคราะห์ข้อมูลทางสถิติ

ในการทำงานวิเคราะห์ข้อมูลทางสถิตินั้น จำเป็นต้องใช้ซอฟต์แวร์ร่วมในการทำงานที่สำคัญอีกตัวหนึ่งคือโปรแกรม text editor ซึ่งจะทำงานร่วมกับแฟมิลี **R** code และ **R** output ดังแสดงด้านซ้ายมือในรูปที่ 3 และรูปที่ 4 ข้างต้น โปรแกรม text editor ที่เหมาะสมในการใช้งานร่วมกับโปรแกรม **R** นั้นจะต้องเป็นซอฟต์แวร์ที่ใช้งานได้ฟรีโดยไม่ต้องเสียค่าใช้จ่าย ดาวน์โหลดได้ง่ายทางอินเทอร์เน็ต สามารถตรวจสอบการเปิดและปิดวงเล็บต่างๆ ได้ สามารถเปลี่ยนสีฟังก์ชันและคำสงวนของ **R** ได้ และหากมีความสามารถเสริมอื่นๆ ได้ก็จะเป็นการดี

ในการหาโปรแกรม text editor ที่เหมาะสมในแต่ละระบบปฏิบัติการนั้นนับว่าใช้เวลามากในการทดสอบ เนื่องจากในรายการ text editor ที่สนับสนุน syntax ของ **R** นั้นมีอยู่จำนวนหนึ่ง แต่ผู้วิจัยจำเป็นต้องหาตัวที่เหมาะสมที่สุดในการใช้งานร่วมกับ **R** โดยจะต้องติดตั้ง และใช้งานง่ายที่สุด และหากสามารถใช้ได้ในทุกระบบปฏิบัติการก็จะเป็นการดี

อย่างไรก็ตาม เมื่อได้ทดลองทุกโปรแกรมบนทุกระบบปฏิบัติการหลักคือ วินโดวส์ แมคอินทอช OSX และลินุกซ์แล้ว พบว่าไม่มีโปรแกรมใดที่ดีที่สุดบนทุกระบบปฏิบัติการ และได้ข้อสรุปดังนี้

ระบบปฏิบัติการวินโดวส์ โปรแกรมที่น่าใช้มากที่สุดมีสองตัว คือ Tinn-R ซึ่งดาวน์โหลดได้ที่ <https://sourceforge.net/projects/tinn-r> และ Crimson Editor ซึ่งดาวน์โหลดได้ที่ <http://crimsoneditor.com> โปรแกรมทั้งสองมีคุณสมบัติข้างต้นทุกประการ แต่ Tinn-R มีความสามารถพิเศษในการส่งบรรทัดคำสั่งไปทำงานกับ **R** ได้โดยตรง และ Crimson Editor มีความสามารถพิเศษในการตรวจสอบว่าแฟมิลีในฮาร์ดดิสก์กับที่อยู่ในหน่วยความจำไม่เหมือนกัน (มีการแก้ไขโดยโปรแกรมอื่น เช่น **ICE**) และเตือนให้อ่านข้อมูลใหม่มาแสดงแทนข้อมูลเดิมได้ ดังนั้นในการทำงานกับสภาพแวดล้อม **ICE** ซึ่งมีการเขียนแฟมิลีโดยโมดูล ice.main ด้วย โปรแกรม Crimson Editor ก็ดูจะเหมาะสมมากกว่า Tinn-R แต่หากใช้งานโปรแกรม **R** โดยไม่ใช่ **ICE** โปรแกรม Tinn-R ก็จะมีคามเหมาะสมมากกว่า

สำหรับระบบปฏิบัติการแมคอินทอช OSX โปรแกรมที่มีคุณสมบัติดังกล่าว และยังสามารถพิเศษในการเปลี่ยนข้อความที่แสดงให้ตรงกับข้อมูลที่เปลี่ยนแปลงในฮาร์ดดิสก์คือ โปรแกรม TextWrangler

สำหรับระบบปฏิบัติการลินุกซ์ โปรแกรมที่มีคุณสมบัติดังกล่าวคือโปรแกรม Blue fish ซึ่งมักจะติดตั้งมากับระบบปฏิบัติการอยู่แล้ว หรือจะดาวน์โหลดรุ่นล่าสุดจากอินเทอร์เน็ตได้ฟรีก็ได้ แต่โปรแกรมนี้ไม่สามารถเปลี่ยนข้อความที่แสดงตามข้อมูลในฮาร์ดดิสก์ที่เปลี่ยนไปได้ ผู้ใช้ต้องจัดการ reload ด้วยตัวเอง

ในปัจจุบัน นักศึกษาในหน่วยระบาดวิทยา ได้ใช้โปรแกรม Tinn-R ช่วยในการเขียนแฟ้มสคริปต์ R ทั้งสิ้น และความรู้เกี่ยวกับโปรแกรมเสริมเหล่านี้ ได้มีเขียนไว้ในหนังสือคู่มือการใช้โปรแกรม R-ICE ทั้งหมดแล้ว เพื่อให้ผู้ใช้งานในระบบปฏิบัติการต่างๆ ได้เลือกใช้และติดตั้งประกอบการทำงานในสภาพแวดล้อม R-ICE

7. การสอนในหลักสูตรระบาดวิทยานานาชาติ และหลักสูตรสถิติ

หลักสูตรระบาดวิทยานานาชาติ ของหน่วยระบาดวิทยา มีรายวิชา Statistics I และ Statistics II ซึ่งทั้งสองรายวิชานี้ได้ใช้โปรแกรม R ในการเรียนการสอนทั้งหมดมาตั้งแต่ปีการศึกษา 2547 ทั้งนี้ในระยะแรก นักศึกษายังมีปัญหาในการทำความเข้าใจอยู่บ้าง เนื่องจากนำไปเปรียบเทียบกับโปรแกรม Stata ซึ่งใช้ในการเรียนการสอนรุ่นก่อนๆ หน้านั้น

ฟังก์ชันที่สอนในหลักสูตรนี้เน้นฟังก์ชันพื้นฐานของ R เป็นหลัก และแนะนำให้ผู้รู้จักฟังก์ชันในแพ็คเกจ EpiCan และ ice, epid บ้าง ในการจัดการข้อมูล และการเรียนรู้หลักการทางระบาดวิทยา ซึ่งทำได้ง่ายกว่ามาก

อย่างไรก็ตาม ในระยะแรกนั้น การจัดการเรียนการสอนยังไม่เป็นระบบเท่าที่ควร จึงค่อยๆ เปลี่ยนแปลงเนื้อหาที่สอน โดยค่อยๆ สอนการใช้ฟังก์ชัน และความรู้เพิ่มขึ้นทีละน้อยตามลำดับ ในระยะหลัง นักศึกษาไม่มีการเปรียบเทียบกับโปรแกรมอื่น และสามารถเรียนรู้ได้อย่างดี และนักศึกษาในรุ่นหลังสามารถใช้โปรแกรม R ได้อย่างมีประสิทธิภาพอย่างเห็นได้ชัด

และในการจัดการเรียนการสอนของสาขาวิชาสถิติ ในภาควิชาคณิตศาสตร์ คณะวิทยาศาสตร์ ก็ได้ใช้โปรแกรม R ในการเรียนการสอนในระดับปริญญาตรีด้วย และในปี พศ. 2549 เมื่อ นส. อภิรดี หนึ่งในคณะผู้วิจัยได้ย้ายไปเป็นอาจารย์ในคณะวิทยาศาสตร์และเทคโนโลยี วิทยาเขตปัตตานี ก็ได้นำโปรแกรม R ไปสอนนักศึกษาที่นั่นด้วย ทำให้ขณะนี้มีการใช้โปรแกรม R ในการเรียนการสอนระดับต่างๆ ในมหาวิทยาลัยสงขลานครินทร์กว้างขวางขึ้น แต่ยังต้องใช้เวลาต่อไปอีกระยะหนึ่ง เมื่อนักศึกษาระดับบัณฑิตศึกษาจบไปมากขึ้น ซึ่งส่วนหนึ่งเป็นอาจารย์ในมหาวิทยาลัยสงขลานครินทร์เอง จึงจะเห็นการใช้โปรแกรม R ในการเรียนการสอนเป็นส่วนใหญ่ในมหาวิทยาลัยได้

8. การจัดทำคู่มือการใช้งานโปรแกรม R

ในการจัดประชุมอบรมเชิงปฏิบัติการที่จะกล่าวในหัวข้อต่อไป จำเป็นต้องจัดทำคู่มือการใช้งานโปรแกรม R ประกอบการอบรมด้วยตามแผนงานที่เสนอไว้ การอบรมจะแบ่งเป็นสองสาย คือสายวิทยาศาสตร์สุขภาพ เน้นการวิเคราะห์ทางสถิติโดยอิงการวิจัยเชิงระบาดวิทยา และสายวิทยาศาสตร์ทั่วไป เน้นความรู้ในการใช้โปรแกรม R ในการวิเคราะห์ทางสถิติโดยทั่วไป ใช้ได้กับทุกสาขา ในสายแรกนั้นผู้วิจัยในหน่วยระบาดวิทยารับผิดชอบ และอีกสายหนึ่งนั้นผู้วิจัยในภาควิชาคณิตศาสตร์เป็นผู้รับผิดชอบ แม้ว่าเนื้อหาโปรแกรม R จะคล้ายกัน แต่การ

สอนการวิเคราะห์ทางสถิติจะแตกต่างกันไปบ้าง ตามจุดเน้นที่แตกต่างกัน จึงได้จัดทำคู่มือขึ้นเป็นสองเล่มที่มีเนื้อหาส่วนหนึ่งคล้ายกัน แต่ในส่วนของการวิเคราะห์ข้อมูลมีการเน้นการทดสอบทางสถิติที่แตกต่างกัน

เนื้อหาของคู่มือแตกต่างไปจากคู่มือของ **R** ในฉบับที่เป็นภาษาอังกฤษ เนื่องจากในฉบับภาษาอังกฤษเขียนในเชิงแนะนำ **R** ในลักษณะภาษาคอมพิวเตอร์ มากกว่าการใช้งานในการวิเคราะห์ทางสถิติ ทำให้ผู้ใช้ทางสถิติอ่านเข้าใจได้ยาก และไม่เห็นภาพการนำไปใช้งานในการวิเคราะห์ข้อมูล ซึ่งเป็นจุดอ่อนที่สำคัญของคู่มือการใช้งานโปรแกรม **R** ที่ไม่มีฉบับที่เน้นการวิเคราะห์ข้อมูลทางสถิติ แต่ได้มีผู้เขียนหนังสือเป็นภาษาอังกฤษขึ้นมาเพื่อการวิเคราะห์ข้อมูลทางสถิติบ้างเช่นกัน แต่เนื่องจากผู้วิจัยเห็นว่าอาจจะไม่เหมาะกับกลุ่มผู้ใช้คนไทยทั่วไปที่อ่านภาษาอังกฤษไม่แตกฉาน จึงได้เขียนคู่มือขึ้นเป็นภาษาไทย

คู่มือการใช้โปรแกรม **R** สำหรับการวิเคราะห์ข้อมูลทางสถิติและการเขียนฟังก์ชันด้วยโปรแกรม **R** ที่ปรับปรุงล่าสุด ที่ใช้ในการอบรมบุคลากรสายวิทยาศาสตร์สุขภาพซึ่งเน้นทางด้านระบาดวิทยา แบ่งเป็นบทต่างๆ ดังนี้

- โปรแกรม **R**
- library ice
- การสำรวจข้อมูล (data exploration)
- การจัดการข้อมูล (data manipulation)
- สถิติที่ใช้ในการวิเคราะห์ข้อมูลแบบกลุ่ม
- สถิติที่ใช้ในการวิเคราะห์ข้อมูลชนิดต่อเนื่อง
- การเขียนฟังก์ชัน

คู่มือการใช้โปรแกรม **R** การใช้คำสั่งหรือฟังก์ชันพื้นฐาน ที่ปรับปรุงล่าสุด ที่ใช้ในการอบรมบุคลากรสายวิทยาศาสตร์ทั่วไป แบ่งเป็นบทต่างๆ ดังนี้

- โปรแกรม **R**
- การจัดการกับข้อมูล
- การสรุปข้อมูล
- การสร้างกราฟเบื้องต้น
- การแจกแจงความน่าจะเป็น probability distribution
- การทดสอบสมมติฐานและประมาณค่าพารามิเตอร์ของประชากร
- การทดสอบไคสแควร์ (chi-squared test)
- การวิเคราะห์ความแปรปรวนทางเดียว (one way analysis of variance)
- การวิเคราะห์การถดถอย (regression analysis)

9. การจัดประชุมอบรมเชิงปฏิบัติการ

ในการการจัดประชุมอบรมเชิงปฏิบัติการการใช้โปรแกรม **R** ในการวิเคราะห์ทางสถิตินั้น ได้เริ่มจัดให้แก่แพทย์ใช้ทุน และแพทย์ประจำบ้านในคณะแพทยศาสตร์ มหาวิทยาลัยสงขลานครินทร์ เป็นกลุ่มแรก ซึ่งแต่ละครั้งจำเป็นต้องจัดเป็น 2 รุ่น เนื่องจากมีจำนวนผู้เข้าอบรมมาก ไม่สามารถดูแลได้ทั่วถึงในครั้งเดียวได้

ตามแผนงานที่เสนอไว้ การอบรมจะแบ่งเป็นสองสาย คือสายวิทยาศาสตร์สุขภาพ เน้นการวิเคราะห์ทางสถิติโดยอิงการวิจัยระบาดวิทยา และสายวิทยาศาสตร์ทั่วไป เน้นความรู้ในการใช้โปรแกรม **R** ในการวิเคราะห์ทางสถิติโดยทั่วไป ใช้ได้กับทุกสาขา ในสายแรกนั้นผู้วิจัยในหน่วยระบาดวิทยารับผิดชอบ และอีกสาย

หนึ่งนั้นผู้วิจัยในภาควิชาคณิตศาสตร์เป็นผู้รับผิดชอบ แม้ว่าเนื้อหาโปรแกรม **R** จะคล้ายกัน แต่การสอนการวิเคราะห์ทางสถิติจะแตกต่างกันไปบ้าง ตามจุดเน้นที่แตกต่างกัน

ในการอบรมครั้งแรกๆ นั้น เป็นการทดสอบวิธีการสอนและเอกสารประกอบการสอนไปด้วย จึงเน้นการจัดภายในมหาวิทยาลัยสงขลานครินทร์เป็นหลัก และเมื่อมีประสบการณ์มากพอแล้ว จึงได้จัดให้แก่สถาบันอื่นๆ ด้วย และนอกจากนี้ แม้จะสิ้นสุดโครงการแล้ว ยังมีสถาบันที่ติดต่อขอให้จัดอบรมให้อีกด้วย

สรุปการจัดอบรมตลอดโครงการดังนี้

ครั้งที่	เวลาที่จัด	กลุ่มเป้าหมาย	จำนวน
1	20-24 ธันวาคม 2547	แพทย์ใช้ทุนและแพทย์ประจำบ้าน คณะแพทยศาสตร์	30 คน
2	10-14 มกราคม 2548	แพทย์ใช้ทุนและแพทย์ประจำบ้าน คณะแพทยศาสตร์	30 คน
3	18-20 พฤษภาคม 2548	บุคลากรในมหาวิทยาลัยสงขลานครินทร์ และบุคคลทั่วไป	30 คน
4	มิถุนายน 2548	ศิษย์เก่าหลักสูตรระบาควิทยานานาชาติ	30 คน
5	15-17 มิถุนายน 2548	บุคลากรจากทะเบียนมะเร็งทั่วประเทศ	15 คน
6	สิงหาคม 2548	อาจารย์และนักศึกษามัธยมศึกษา มหาวิทยาลัยสงขลานครินทร์ วิทยาเขตปัตตานี	30 คน
7	20 และ 27 ธันวาคม 2548	นักศึกษาปริญญาตรี มหาวิทยาลัยสงขลานครินทร์	70 คน
8	10-10 มกราคม 2549	บุคลากรจากทะเบียนมะเร็งทั่วประเทศ	15 คน
9	24-25 มีนาคม 2549	อาจารย์ บุคลากร และนักศึกษามัธยมศึกษา มหาวิทยาลัยสงขลานครินทร์	40 คน
10	26-26 เมษายน 2549	แพทย์ใช้ทุนและแพทย์ประจำบ้าน คณะแพทยศาสตร์	30 คน
11	1-3 พฤษภาคม 2549	แพทย์ใช้ทุนและแพทย์ประจำบ้าน คณะแพทยศาสตร์	30 คน
12	22-23 กรกฎาคม 2549	บุคคลทั่วไปในหน่วยงานภายนอก นักวิจัยและอาจารย์ในสังกัดมหาวิทยาลัยเทคโนโลยีราชมงคล ทุกสาขาวิชา	40 คน
13	25-27 ธันวาคม 2549	อาจารย์ บุคลากร และนักศึกษามัธยมศึกษา มหาวิทยาลัยขอนแก่น	37 คน

การจัดอบรมการวิเคราะห์ทางสถิติด้วยโปรแกรม **R** ครั้งสุดท้ายจัดระหว่างวันที่ 25-27 ธันวาคม 2549 ที่ศูนย์คอมพิวเตอร์ คณะมนุษยศาสตร์และสังคมศาสตร์ มหาวิทยาลัยขอนแก่น การจัดครั้งนี้ทางคณะแพทยศาสตร์ มหาวิทยาลัยขอนแก่น เป็นผู้ร่วมในการจัดการ แต่ได้ใช้สถานที่ของคณะมนุษยศาสตร์ฯ เนื่องจากมีห้องคอมพิวเตอร์ที่สามารถรองรับผู้เข้าร่วมอบรมได้มาก โดยมีผู้เข้ารับการอบรม 37 คน เป็นอาจารย์และ

บุคลากรในมหาวิทยาลัย 33 คน และนักศึกษาระดับบัณฑิตศึกษา 4 คน ส่วนใหญ่มีพื้นฐานด้านสถิติและคอมพิวเตอร์อยู่แล้ว

การอบรมครั้งต่างๆ นั้น ได้ปรับปรุงเปลี่ยนแปลงระยะเวลาในการสอน ลำดับการสอน และเนื้อหาไปตามความเหมาะสมตลอดเวลา ในที่นี้จึงขอนำเสนอตารางการอบรมครั้งล่าสุดให้เห็นเป็นตัวอย่างดังนี้

วัน/เดือน/ปี	รายการ	วิทยากร
25 ธันวาคม 2549		
8.45 – 9.00	ลงทะเบียน	
9.00 – 10.30	แนะนำโปรแกรม R	รศ.นพ.หัชชา ศรีปลั่ง
—	ติดตั้งโปรแกรม R-2.4.0	
—	การใช้รายการช่วย	
—	วัตถุพื้นฐานใน R	
10.30 – 10.45	พักรับประทานอาหารว่าง	
10.45 – 12.00	- ติดตั้งโปรแกรม Tinn-R	รศ.นพ.หัชชา ศรีปลั่ง
—	ติดตั้งแพ็คเกจ ice และ สภาพแวดล้อม R-ICE	
—	การนำเข้าข้อมูลเข้าและการจัดการข้อมูล	
12.00-13.00	พักรับประทานอาหารกลางวัน	
13.00 – 14.45	การสำรวจข้อมูลก่อนการวิเคราะห์ทางสถิติ	Mr. Edward McNeil
—	กราฟ	
—	คำนวณความถี่ / ตาราง	
—	คำนวณค่าเฉลี่ย ค่าเบี่ยงเบนมาตรฐาน การแจกแจงข้อมูล	
14.45 – 15.00	พักรับประทานอาหารว่าง	
15.00 – 16.30	การจัดการข้อมูลก่อนการวิเคราะห์ทางสถิติ	น.ส. อภิรดี แซ่ลิ่ม
-	การเลือกเพียงบางส่วน	
—	การเก็บคำสั่งไว้เป็นไฟล์	
—	การสร้างตัวแปรใหม่	
—	การให้คำอธิบายตัวแปร	
—	การให้คำอธิบายค่าของตัวแปร	
—	การจัดกลุ่มเป็นช่วงค่า	
—	การจัดการกับข้อมูลสูญหาย	
26 ธันวาคม 2549		
9.00-10.00	การวิเคราะห์ข้อมูลทางสถิติสำหรับตัวแปรแบบกลุ่ม	น.ส. อภิรดี แซ่ลิ่ม
—	การทดสอบไคสแควร์	
10.00-10.15	พักรับประทานอาหารว่าง	
10.15- 12.00	การวิเคราะห์ข้อมูลทางสถิติสำหรับตัวแปรแบบกลุ่ม	น.ส. อภิรดี แซ่ลิ่ม
—	การสร้างโมเดลทางสถิติ	

12.00-13.15	พักรับประทานอาหารกลางวัน	
13.15 -14.45	การวิเคราะห์ข้อมูลทางสถิติสำหรับตัวแปรต่อเนื่อง	Mr. Edward McNeil
	— การทดสอบ t-test	
	— การทดสอบ ANOVA	
14.45 – 15.00	พักรับประทานอาหารว่าง	
15.00 – 16.30	การวิเคราะห์ข้อมูลทางสถิติสำหรับตัวแปรต่อเนื่อง	Mr. Edward McNeil
	— การสร้างโมเดลทางสถิติ	
27 ธันวาคม 2549	สำหรับผู้พัฒนาโปรแกรม	
9.00-10.00	การเขียนฟังก์ชันทางสถิติในโปรแกรม R	รศ.นพ.หัชชา ศรีปลั่ง
10.00-10.15	พักรับประทานอาหารว่าง	
10.15- 12.00	การสร้าง package สำหรับโปรแกรม R	รศ.นพ.หัชชา ศรีปลั่ง
12.00-13.15	พักรับประทานอาหารกลางวัน	

จากประสบการณ์ที่ได้จัดการประชุมเชิงปฏิบัติมาหลายครั้ง พบว่าวิธีนี้มีประสิทธิภาพในการเรียนรู้สูงสุด แต่ทำได้จำนวนจำกัด เนื่องจากต้องใช้ทรัพยากรมาก

10 . การสร้างกลุ่มผู้ใช้ R

ในการสร้างกลุ่มผู้ใช้ R ในประเทศไทยนั้น นับเป็นสิ่งที่ยังทำได้ยาก และยังไม่สำเร็จ กลยุทธ์ที่ใช้คือการจัดทำหนังสือเผยแพร่การใช้งาน R และ R-ICE เพื่อสร้างผู้ใช้โปรแกรม R และสภาพแวดล้อม R-ICE ให้มีจำนวนมากเสียก่อน และการจัดทำเว็บไซต์สำหรับผู้ใช้โปรแกรม R ก็จะเป็นที่รวมของผู้ใช้ R ในการเรียนรู้ สอบถาม และสนทนา รวมถึงการเขียนฟังก์ชันใน R แต่เนื่องจากหนังสือที่จะเผยแพร่แก่ประชาชนได้ทำต้นฉบับเสร็จแล้ว กำลังอยู่ในกระบวนการตีพิมพ์ ทำให้จำนวนผู้ใช้ R และ R-ICE ยังมีจำนวนไม่มากนัก สมาชิกของเว็บไซต์จึงยังมีไม่มากนัก ทำให้การจัดกิจกรรมต่างๆ ยังอยู่ในวงจำกัด และยังมีปริมาณคำถามหรือการสนทนาจำนวนน้อย

ในการจัดการอบรมเชิงปฏิบัติการต่างๆ นั้น คณะผู้วิจัยได้แนะนำให้ผู้เข้ารับการอบรมรู้จักเว็บไซต์ที่ได้ทำขึ้นทุกครั้ง แต่การจับอบรมแต่ละครั้งสามารถรับผู้เข้าอบรมได้ในจำนวนจำกัด ขึ้นกับจำนวนของเครื่องคอมพิวเตอร์ที่มี และจำนวนผู้ช่วยวิทยากรในการดูแลผู้เข้าอบรม ทำให้จำนวนผู้ใช้ R เพิ่มขึ้นได้ช้ากว่าการเผยแพร่หนังสือมาก

คณะผู้วิจัยหวังว่าหลังจากหนังสือที่จะได้จัดทำทั้งสองเล่มออกวางจำหน่ายแล้ว จำนวนผู้ใช้ R จะเพิ่มขึ้นมาก และเว็บไซต์ก็จะสามารถจัดกิจกรรมได้อย่างชัดเจนมากขึ้น

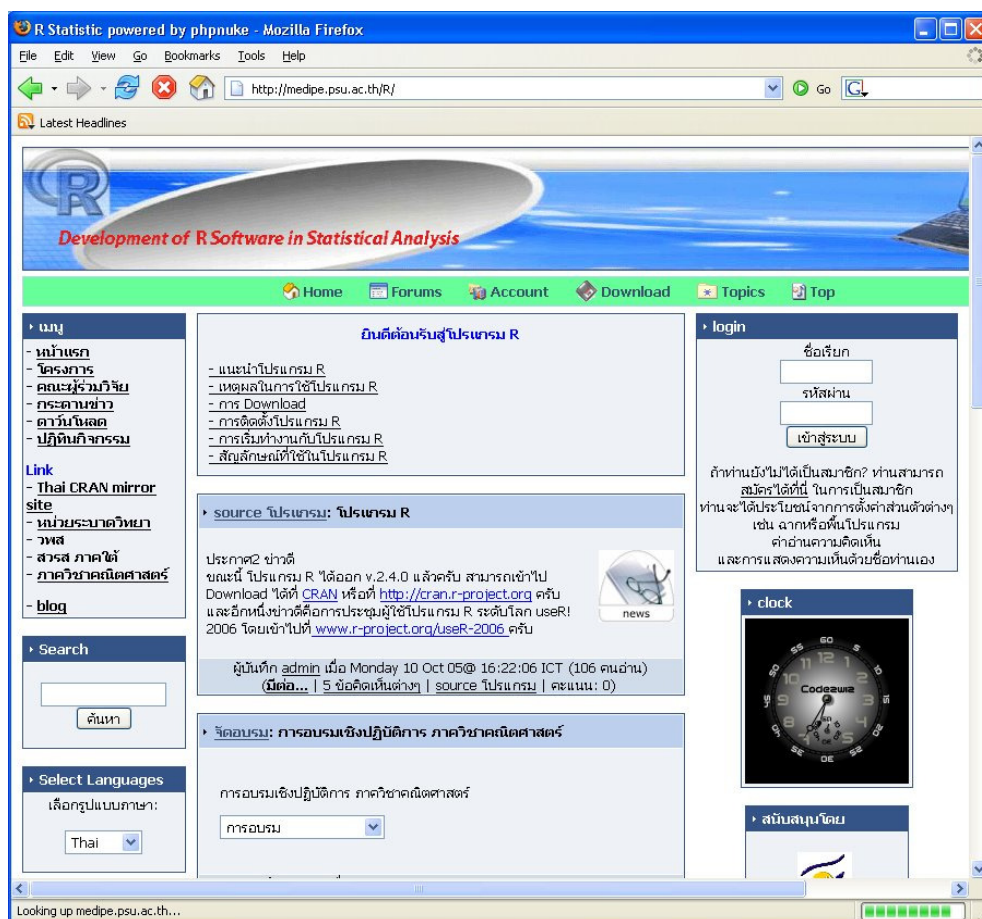
11 . การจัดทำเว็บไซต์

คณะผู้วิจัยได้จัดทำเว็บไซต์ขึ้นจำนวนหนึ่ง เพื่อกิจกรรมที่แตกต่างกันในการศึกษาเรียนรู้ การสร้างกลุ่มผู้ใช้ และสำหรับผู้พัฒนา ดังนี้

เว็บไซต์ทั่วไปสำหรับโปรแกรม R สำหรับคนไทย ดังรูปที่ 6 มีวัตถุประสงค์เพื่อเป็นศูนย์รวมผู้ใช้โปรแกรม R และแพ็คเกจที่เกี่ยวข้องต่างๆ ที่เป็นคนไทย (Thai UseR) เนื้อหาเกี่ยวกับเรื่องทั่วไปที่ควรรู้เกี่ยวกับโปรแกรม R และมีเว็บบอร์ดเพื่อสอบถามหรือสนทนาแลกเปลี่ยนความรู้เกี่ยวกับโปรแกรม R และมี

link ไปสู่หลักสูตรการอบรมที่จัดขึ้นต่างๆ และหลักสูตรการเรียนทางอินเทอร์เน็ต โดยข้อความทั้งหมดเป็นภาษาไทย และมี link ไปสู่เว็บไซต์หลักของ R และ CRAN ด้วย

เว็บนี้ใช้โปรแกรม PHP Nuke ช่วยในการทำ โปรแกรมนี้มีตัวช่วยในการทำงานอยู่มาก ทำให้ดูแลได้ง่าย และปรับปรุงได้โดยไม่ต้องเสียค่าใช้จ่าย เว็บไซต์นี้ใช้เครื่อง server ของหน่วยระบบควิทยา (ซึ่งซื้อโดยเงินของโครงการนี้) ความเร็วในการเข้าถึงค่อนข้างมีปัญหา เนื่องจากการติดต่ออินเทอร์เน็ตระหว่างมหาวิทยาลัยสงขลานครินทร์กับภายนอกมักจะมีปัญหาอยู่เป็นประจำ



รูปที่ 6 เว็บไซต์ทั่วไปสำหรับโปรแกรม R สำหรับคนไทย ที่ <http://medipe.psu.ac.th/R/>

อีกเว็บไซต์หนึ่งเป็นเว็บสำหรับ R-ICE ดังรูปที่ 7 มีวัตถุประสงค์เพื่อแจกจ่าย สอบถาม และแลกเปลี่ยนความคิดเห็นเกี่ยวกับสภาพแวดล้อมและแพ็คเกจต่างๆ ของ R-ICE เว็บนี้ทำเป็นภาษาอังกฤษ เนื่องจากมีเป้าหมายเป็นผู้ใช้จากประเทศต่างๆ ซึ่งที่ผ่านมามีชาวต่างชาติจำนวนหนึ่งได้มาดาวน์โหลดไปทดลองใช้ด้วย

ในเว็บนี้จะเผยแพร่แพ็คเกจต่างๆ ปรับปรุงทุกครั้งที่เปลี่ยนรุ่นของแพ็คเกจต่างๆ มีหน้าข่าวติดตามความก้าวหน้าของการพัฒนาแพ็คเกจ มีคู่มือการติดตั้งและใช้งาน ICE ด้วย โดยเขียนเป็นภาษาอังกฤษ แต่ยังไม่เสร็จสมบูรณ์ นอกจากนี้ยังมีเว็บบอร์ดด้วยเช่นกัน เว็บนี้พัฒนาโดยเขียนขึ้นเอง ให้มีหน้าตาคล้ายหน้าเว็บของ CRAN เพื่อให้ผู้ใช้ที่รู้จักโปรแกรม R อยู่แล้วคุ้นเคยและใช้ได้ง่าย