

บทที่ 1

สถิติเชิงพรรณนา

สถิติเบื้องต้น
โปรแกรม R เบื้องต้น

88520159

PROBABILITY AND STATISTICS FOR COMPUTING

สถิติเบื้องต้น

สถิติ (statistics)

- ยุคแห่งเทคโนโลยีสารสนเทศ ข้อมูลเป็นสิ่งที่มีความสำคัญ
- หน่วยงานต่างๆ มีการเก็บข้อมูลขององค์กรเป็นฐานข้อมูล เพื่อที่จะนำข้อมูลต่างๆ นี้มาใช้ในการวิเคราะห์และสรุปผล เพื่อให้เกิดสารสนเทศที่เป็นประโยชน์ต่อองค์กร
 - ✓ สถิติจำนวนประชาชนที่มาใช้สิทธิเลือกตั้งในแต่ละจังหวัด
 - ✓ สถิติผู้ป่วยที่เป็นไข้หวัดในแต่ละชุมชน
 - ✓ สถิติผู้สอบเข้าศึกษาต่อได้ในแต่ละคณะ

ตัวอย่าง สถิติ

- ตัวอย่าง สถิติการว่างงานของประชากรไทย

ระดับการศึกษาที่สำเร็จ	จำนวน (พันคน)	ร้อยละ
1. ไม่มีการศึกษา	43.89	1.00
2. ต่ำกว่าประถมศึกษา	232.56	5.13
3. ประถมศึกษา	631.26	14.06
4. มัธยมศึกษาตอนต้น	949.87	20.78
5. มัธยมศึกษาตอนปลาย	909.24	20.05
สายสามัญ	634.56	14.19
สายอาชีวศึกษา	262.37	5.79
สายวิชาการศึกษา	3.29	0.07
6. อุดมศึกษา	1755.57	38.66
สายวิชาการ	1154.17	25.60
สายวิชาชีพ	368.53	8.13
สายวิชาการศึกษา	232.86	4.93
7. การศึกษาอื่น ๆ	4.69	0.11
8. ไม่ทราบ	9.78	0.20
รวม	4536.82	100.00



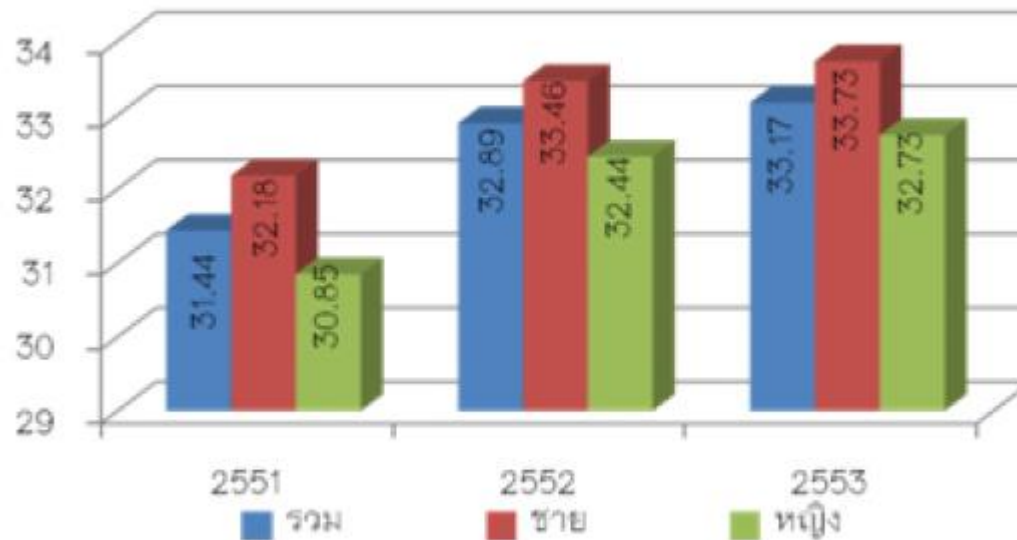
เราได้อะไรจากตารางนี้ ?

ที่มา: การสำรวจภาวะการทำงานของประชากร พ.ศ.
2559 สำนักงานสถิติแห่งชาติกระทรวงเทคโนโลยี
สารสนเทศและการสื่อสาร

ตัวอย่าง สถิติ

- ตัวอย่าง คะแนนสุขภาพจิตเฉลี่ยของผู้สูงอายุพ.ศ. 2553

คะแนนสุขภาพจิตเฉลี่ยของผู้สูงอายุ จำแนกตามเพศ
พ.ศ. 2551 - 2553



ที่มา: โครงการสำรวจสุขภาพจิตคนไทย พ.ศ. 2551 - 2553



เราได้อะไรจากตารางนี้ ?

ตัวอย่าง สถิติ

- ตัวอย่าง สถิติการอ่านหนังสือของคนไทย



ตัวอย่าง สถิติ

- ตัวอย่าง สถิติการอ่านหนังสือของคนไทย



ตัวอย่าง สถิติ

- ตัวอย่าง สถิติการอ่านหนังสือของคนไทย

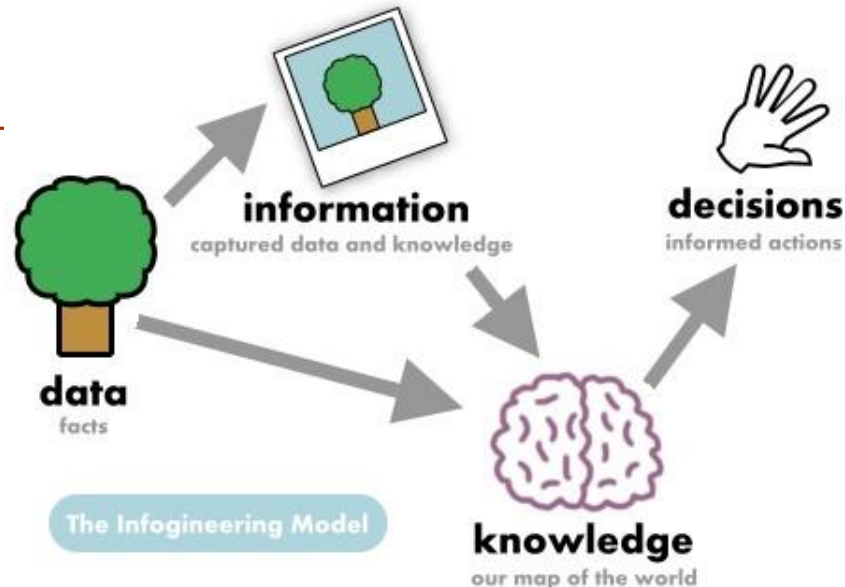


ที่มา : การสำรวจการอ่านของประชากร พ.ศ. 2556-2558 ศูนย์สารสนเทศยุทธศาสตร์ภาครัฐ สำนักงานสถิติแห่งชาติ

ความหมายของสถิติ (Defenition of statistics)

ข้อมูลสถิติ คือ ข้อเท็จจริงที่เก็บรวบรวมมาจากแหล่งต่างๆ ซึ่งอาจจะเป็นตัวเลขหรือตัวอักษร

สถิติศาสตร์ คือ วิธีการหรือกระบวนการที่ใช้ในการจัดการข้อมูลสถิติที่ได้ข้อสรุปที่เป็นประโยชน์ต่อการนำไปใช้ ช่วยในการจัดการข้อมูลเพื่อให้ได้มาซึ่งสารสนเทศ (information) ที่เป็นประโยชน์จนได้ความรู้ (knowledge) เพื่อนำไปใช้ในการตัดสินใจต่อไป



บทบาทความสำคัญของสถิติ

- ในการจัดการเกี่ยวกับปัญหาหรือคำถาม ซึ่งมักจะเป็นกรณีของสถานการณ์ที่ไม่ทราบว่าจะเกิดอะไรขึ้นในอนาคต หรือสถานการณ์ที่ไม่แน่นอน
 - ผลการเลือกตั้งที่จะเกิดขึ้นในอีก 3 เดือนข้างหน้า
 - ประสิทธิภาพของยาชนิดใหม่ที่เพิ่งคิดค้นได้
 - การเตรียมการเพื่อป้องกันน้ำท่วม
- สถิติจะเข้ามามีส่วนที่ช่วยในการตอบปัญหาหรือคำถามเหล่านี้โดยอาศัยสถิติศาสตร์เป็นเครื่องมือ
- ดังนั้นสถิติจึงเข้ามามีบทบาทที่สำคัญในทุกด้านไม่ว่าจะเป็นในชีวิตประจำวันหรือด้านวิชาชีพต่างๆ

บทบาทของสถิติในชีวิตประจำวัน

บุบอรร้อย นากะลิ้ง

3.9 ★ 525 รีวิว ๕๕๕ เปิดอยู่ 2018

อาหารทะเล, อาหารไทย

เมนูเด็ด: ปลากระพงทอดน้ำปลา, ปลากระพงทอดน้ำปลา, ปลาหมึกผัดไข่เค็ม



La Baguette นากะลิ้ง

3.8 ★ 445 รีวิว ๕๕๕ เปิดอยู่

#1 จาก 211 ร้านคาเฟ่ ในจังหวัดชลบุรี

เมนูเด็ด: Almond Croissant, Chocolate Cake, วาฟเฟิล

รับส่วนลด 10% เมื่อชำระด้วยบัตรเครดิตธนาคารออมสิน

รับส่วนลด 10% สำหรับค่าอาหารและเครื่องดื่ม (ยกเว้นสินค้าฝากขาย)

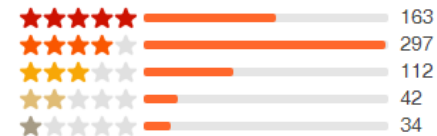


525 รีวิว 648 เรตติ้ง

62% ของรีวิวบอกว่าร้านนี้เหมาะกับครอบครัวหรือเด็ก

59% เพื่อนฝูง

40% คู่รัก



445 รีวิว 516 เรตติ้ง

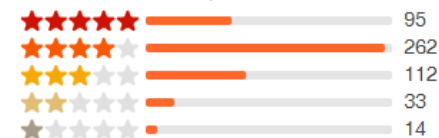
68% ของรีวิวบอกว่าร้านนี้เหมาะกับคู่รัก

66% เพื่อนฝูง

54% ครอบครัวหรือเด็ก

อันดับ #1 จาก 211

ร้านคาเฟ่ ในจังหวัดชลบุรี



บทบาทของสถิติในชีวิตประจำวัน

 ส่วนลด 74% Mall SKMEI แท้100% นาฬิกา Bluetooth น้ําก้าว... ฿1,498 ฿377 ♡ 249 ★★★★★ (27)	 ส่วนลด 60% Mall Hot! ถูกสุดๆ นาฬิกา Casio Baby-G รุ่น... ฿4,500 ฿1,799 ♡ 745 ★★★★★ (155)	 นาฬิกาโทรศัพท์ Smart Watch Phone รุ่น DZ09 ... ฿599 ♡ 72 ★★★★★ (5)	 ส่วนลด 86% Smart Watch Phone รุ่น กล้องนาฬิกา บูลทูธ ใส่ซิม... ฿2,045 ฿266 ♡ 725 ★★★★★ (489)	 เด็กสี HD โทรศัพท์นาฬิกา โทรศัพท์มือถือหน้าจอสัมผัส... ฿280 ♡ 239 ★★★★★ (63)
 ส่วนลด 50% ฤดูใบไม้ผลิหญิงสั้นและฤดูร้อน 2018 คลีนลูกใหม่ขอ... ฿344 ฿172 ♡ 398 ★★★★★ (15)	 ส่วนลด 39% ผู้หญิง เสื้อเชิ้ต สไลด์ เสื้อชีฟองใส่สบายแขนสั้น... ฿363 ฿218 ♡ 331 ★★★★★ (41)	 EASON SHOP (GU5566) แพนหนุ่มเกาหลีสีสั้น ๆ ก่อน... ฿159 ♡ 145 ★★★★★ (6)	 ✓ ร้านแนะนำ ส่วนลด 28% เสื้อยืดแขนสั้น ดอกกลม แต่งระบายรอบอกไปถึงแขนเสื้อ... ฿350 ฿250 ♡ 447 ★★★★★ (56)	 ขายส่ง 2018 ฤดูร้อนหลวมผู้หญิงขนาดใหญ่ของแฟชั่นเกาหลี... ฿158 ♡ 162 ★★★★★ (34)

บทบาทของสถิติต่อวิชาชีพต่างๆ

สถิติต่อวิชาชีพทางการแพทย์



สถิติต่อวิชาชีพทางธุรกิจ

ชนิดของข้อมูล (Type of data)

- ข้อมูลอาจจะได้มาจากการสังเกต การรวบรวม การวัดข้อมูลเป็นได้ทั้งข้อมูลตัวเลขหรือสัญลักษณ์ใด ๆ ที่สำคัญจะต้องมีความเป็นจริง
- ตัวอย่างของข้อมูล เช่น คะแนนสอบ ชื่อนักเรียน เพศ อายุ เป็นต้น
- การเก็บค่าของข้อมูลเรามักนิยมตั้งชื่อตัวแปร เพื่อแสดงหรือ กำกับ ความหมายของข้อมูล

ตัวแปร หมายถึง ลักษณะบางอย่างที่เราสนใจ ค่าของตัวแปร อาจอยู่ในรูปข้อความ หรือตัวเลขก็ได้

ข้อมูล หมายถึง ข้อเท็จจริงที่แสดงลักษณะหรือเรื่องราวเกี่ยวกับสิ่งต่างๆ เช่น คน สัตว์สิ่งของ สถานที่ซึ่งถือเป็นค่าของตัวแปร

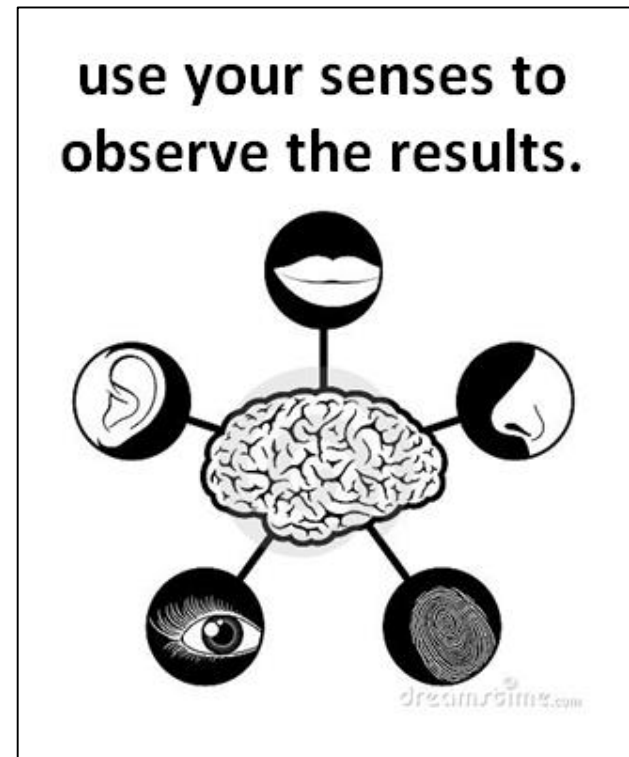
ชนิดของข้อมูล (Type of data)

เมื่อจำแนกตามลักษณะของข้อมูล สามารถแบ่งออกได้เป็น 2 ชนิดคือ

1. ข้อมูลเชิงคุณภาพ (Qualitative Data) หมายถึง ข้อมูลที่ไม่สามารถบอกได้ว่ามีค่ามากหรือน้อย แต่จะสามารถบอกได้ว่าดีหรือไม่ดีหรือบอกลักษณะความเป็นกลุ่มของข้อมูล

เช่น เพศ ศาสนา สีมม คุณภาพสินค้า

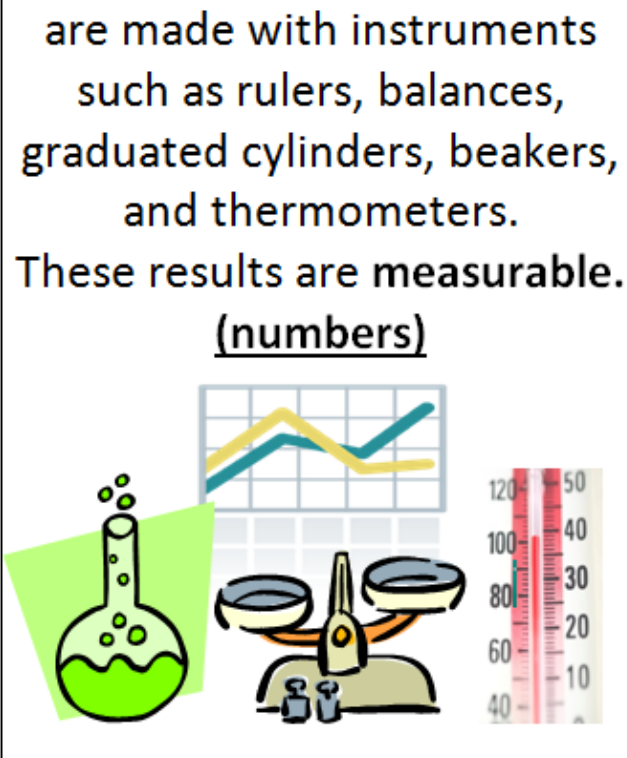
ความพึงพอใจ อันดับนางงาม ลำดับที่สอบ



ชนิดของข้อมูล (Type of data)

2. ข้อมูลเชิงปริมาณ (Quantitative Data) หมายถึง ข้อมูลที่สามารถวัดค่าได้ว่ามีค่ามากหรือน้อย ซึ่งสามารถวัดค่าออกมาเป็นตัวเลขได้

เช่น คะแนนสอบ อุณหภูมิส่วนสูง น้ำหนัก ปริมาณต่างๆ ฯลฯ



ชนิดของข้อมูล (Type of data)

จากข้อมูลในตารางนี้ ถือว่าเป็นข้อมูลเชิงปริมาณหรือคุณภาพ?

ชื่อรุ่น	ราคา
3310	2,150
3235	3,400
6830	4,000
7729	5,500
8990	6,800

ตารางแสดงราคาโทรศัพท์มือถือรุ่นต่างๆ

ชนิดของข้อมูล (Type of data)

หากจำแนกชนิดของข้อมูลให้ละเอียดเมื่อพิจารณาจากค่าที่เป็นไปได้ของข้อมูล จะสามารถจำแนกข้อมูลออกเป็น 4 มาตรา ดังนี้

1. มาตรานามบัญญัติ (Nominal scale) เป็นมาตรวัดข้อมูลที่มีค่าเป็นเชิงคุณภาพ มีลักษณะเป็นกลุ่มที่ไม่สามารถเปรียบเทียบหรือบอกถึงความแตกต่างได้เช่น เพศ (ชาย-หญิง) อาชีพ (ครู-หมอ-วิศวกร ฯลฯ) กลุ่มเลือด (O-A-B-AB) เป็นต้น

ชนิดของข้อมูล (Type of data)

2. มาตราเรียงอันดับ (Ordinal scale) เป็นมาตรวัดที่ค่าของข้อมูลเป็นข้อความ หรือตัวเลข ที่สามารถจัดลำดับความสำคัญหรือเปรียบเทียบกันได้ แต่ไม่สามารถบอกถึงขนาดของความแตกต่างได้ เช่น ระดับความคิดเห็น (มากที่สุด-มาก-ปานกลาง-น้อย-น้อยที่สุด) เราสามารถกล่าวได้ว่าระดับความคิดเห็นมากที่สุดมีค่าสูงกว่าปานกลาง แต่ไม่สามารถกล่าวได้ว่ามากกว่ากันอยู่เท่าใด

ชนิดของข้อมูล (Type of data)

3. มาตราอันตรภาค (Interval scale) เป็นมาตราที่ค่าของข้อมูลเป็นเชิงปริมาณ หรือตัวเลขที่วัดได้จริงมีระยะห่างหรือความแตกต่างของค่าที่เท่ากัน ตัวเลขที่วัดได้บอกถึงความแตกต่างและเปรียบเทียบกันได้แต่ค่าของเลขศูนย์(0) ไม่ใช่ศูนย์แท้จริง เช่น อุณหภูมิซึ่งเป็นค่าที่ใช้วัดค่าเฉลี่ยของพลังงานจลน์ของอนุภาคในสสารใดๆ ซึ่งสอดคล้องกับความร้อนหรือเย็นของสสารนั้น

4. มาตราอัตราส่วน (Ratio scale) เป็นมาตราวัดที่สูงที่สุด คือนอกจากสามารถแบ่งกลุ่มได้จัดอันดับได้มีช่วงห่างของข้อมูลเท่าๆ กันแล้ว ยังเป็นข้อมูลที่มีศูนย์แท้จริง และสามารถนำค่าของข้อมูล มาบวก ลบ คูณ หาร กันได้เช่น น้ำหนัก ส่วนสูง ระยะทาง รายได้ ฯลฯ

ชนิดของข้อมูล (Type of data)

จงจำแนกข้อมูลเหล่านี้ว่าเป็นข้อมูลที่อยู่ในมาตราใด

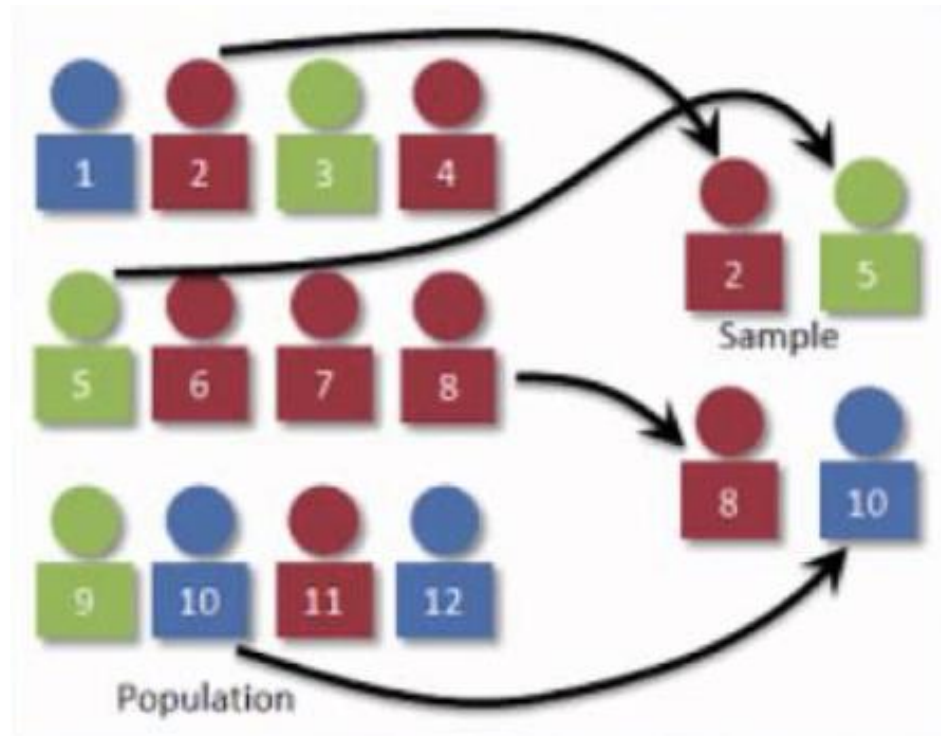
- | | |
|----------------------------|---|
| 1. ระดับการศึกษา | 7. ศาสนา |
| 2. สี | 8. ปริมาณสุทธิของนมสดในขวด |
| 3. เกรดในรายวิชา (A-F) | 9. รหัสประจำตัวนิสิต |
| 4. เกรดเฉลี่ย (GPA) | 10. ทักษะคติ (เห็นด้วยมาก เห็นด้วย ไม่เห็นด้วย) |
| 5. ผลการรักษา (หาย ไม่หาย) | 11. คะแนน IQ |
| 6. ราคาน้ำอัดลม | 12. จำนวนนิสิตในห้องเรียน |

ประชากรและตัวอย่าง (Population and Sample)

ประชากร (Population) หมายถึง หน่วยต่างๆ ทุกหน่วยที่เราสนใจศึกษาที่ให้ข้อมูลที่เราต้องการ เช่น การศึกษารายได้เฉลี่ยของข้าราชการในจังหวัดชลบุรี ดังนั้นประชากรคือ ข้าราชการทุกคนในจังหวัดชลบุรี

ตัวอย่าง (Sample) หมายถึง หน่วยบางหน่วยในประชากรที่ต้องการศึกษา เช่น การศึกษารายได้เฉลี่ยของข้าราชการในจังหวัดชลบุรี ตัวอย่างคือ ตัวแทนข้าราชการ 50 คน ที่ถูกสอบถามเกี่ยวกับรายได้

ประชากรและตัวอย่าง (Population and sample)



ประชากรและตัวอย่าง (Population and Sample)

พารามิเตอร์ คือค่าตัวเลขที่ถูกคำนวณมาจากทุกหน่วยของประชากร เป็นค่าที่อธิบายถึงลักษณะของประชากร สัญลักษณ์ที่ใช้แทนค่าพารามิเตอร์โดยส่วนใหญ่จะใช้อักษรกรีก เช่น ค่าเฉลี่ยประชากร (μ)

ค่าสถิติ คือค่าตัวเลขที่ถูกคำนวณมาจากข้อมูลตัวอย่าง ซึ่งเป็นค่าที่อธิบายลักษณะของตัวอย่าง สัญลักษณ์ที่ใช้แทนค่าสถิติจะใช้เป็นตัวอักษรภาษาอังกฤษ เช่น ค่าเฉลี่ยตัวอย่าง ($\bar{X}$)

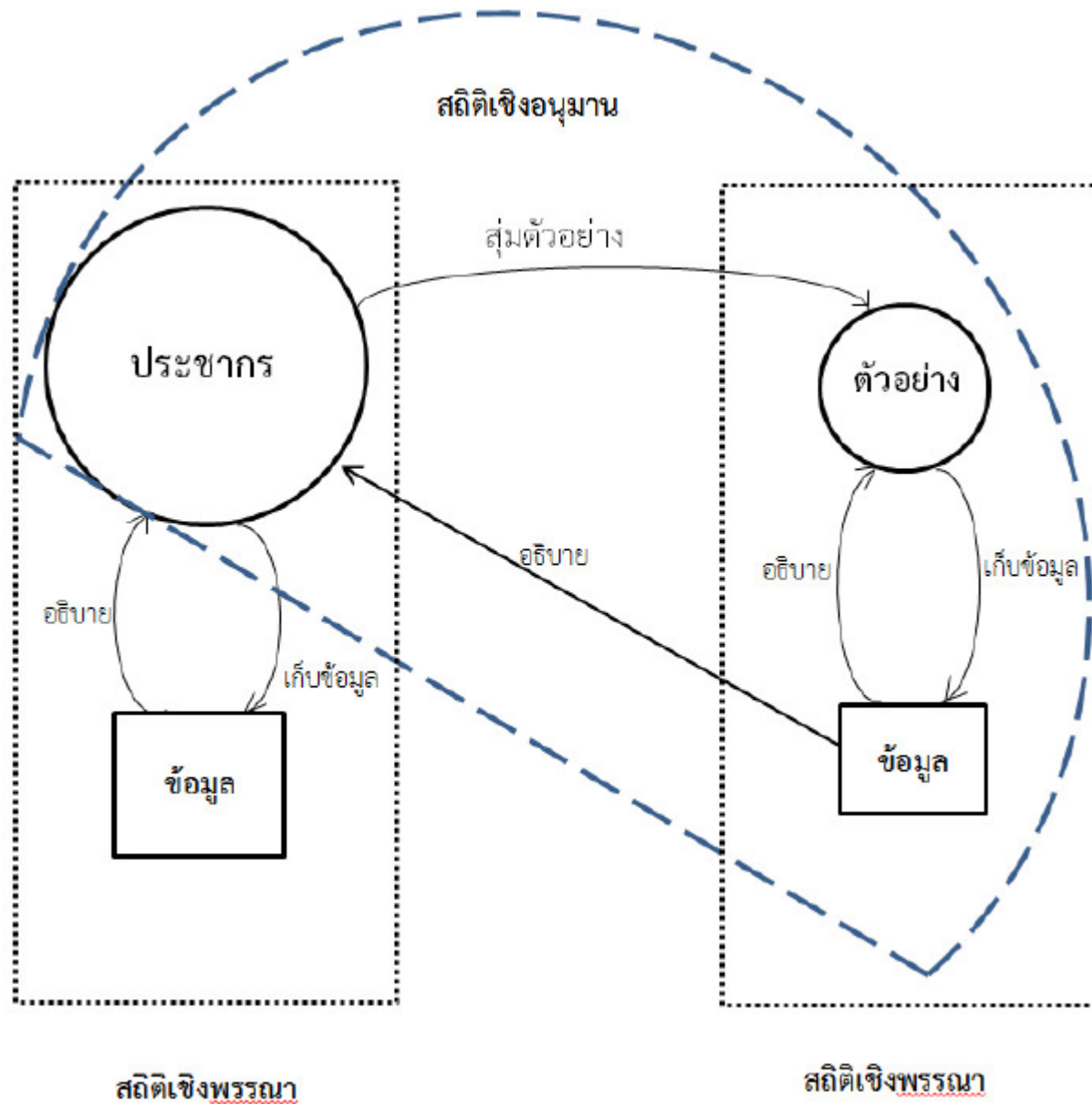
ความหมาย	สัญลักษณ์	
	พารามิเตอร์	ค่าสถิติ
ค่าเฉลี่ย	μ	$\bar{x}$
ค่าเบี่ยงเบนมาตรฐาน	σ	s
ค่าความแปรปรวน	σ^2	s^2
ค่าสัดส่วน	p	$\hat{p}$

ประเภทของสถิติศาสตร์ (Type of statistics)

สถิติศาสตร์แบ่งเป็น 2 ประเภท

- **สถิติเชิงพรรณนา (Descriptive statistics)** คือ วิธีการทางสถิติที่ใช้ในการพรรณาลักษณะของข้อมูลที่เก็บรวบรวมมาชุดหนึ่งซึ่งอาจเป็นข้อมูลตัวอย่างหรือประชากรก็ได้ในรูปแบบของ ตาราง กราฟ หรือข้อความ แต่ไม่สามารถคาดคะเนนอกเหนือข้อมูลชุดที่มีอยู่ได้
- **สถิติเชิงอนุมาน (Inferential statistics)** คือ วิธีการทางสถิติที่อาศัยทฤษฎีความน่าจะเป็น ในการอนุมานลักษณะของประชากรโดยอาศัยข้อมูลตัวอย่าง เช่น การศึกษาความคิดเห็นของนิสิตที่มีต่อการกำหนดให้สวมหมวกกันน็อกในการขับขี่รถจักรยานยนต์

สถิติเชิงพรรณนาและสถิติเชิงอนุมาน



ระเบียบวิธีการทางสถิติ (Statistical process)

1. การเก็บรวบรวมข้อมูล (Data collection) เป็นการรวบรวมข้อมูลจากแหล่งข้อมูลปฐมภูมิหรือทุติยภูมิ
2. การนำเสนอข้อมูล (Data presentation) เป็นการจัดทำข้อมูลที่รวบรวมได้ให้อยู่ในรูปแบบที่กะทัดรัด เช่น ตาราง กราฟ แผนภูมิ ข้อความ เป็นต้น
3. การวิเคราะห์ข้อมูล (Data analysis) ต้อง ใช้สูตรทางสถิติต่างๆ หรือ ใช้การอ้างอิงทางสถิติ เช่น การวิเคราะห์ แนวโน้มเข้าสู่ส่วนกลาง การวัดการกระจาย การทดสอบสมมติฐาน การประมาณค่า เป็นต้น
4. การตีความหมายและสรุปผล (Interpretation and conclusion) เป็นขั้นตอนของการนำผลการวิเคราะห์มาอธิบายให้บุคคลทั่วไปเข้าใจ

การเก็บรวบรวมข้อมูล (Data collections)

1. การสัมภาษณ์ (interview)
2. การสอบถามทางไปรษณีย์ (mail)
3. การสอบถามทางโทรศัพท์ (telephone)
4. การสังเกต (observation)
5. การทดลอง (experiment)

โปรแกรม R เบื้องต้น

โปรแกรม R

- R เป็นภาษาคอมพิวเตอร์ภาษาหนึ่ง
- นิยมในการใช้คำนวณคำนวณสถิติพื้นฐานสำหรับงานทางธุรกิจ
- มีฟังก์ชันทางสถิติที่ช่วยในการประมวลผลทาง Data Sciences และ Big Data
- ใช้ได้ฟรีบนเครื่องคอมพิวเตอร์แบบต่างๆ ทั้ง Windows, Mac OS, หรือ Linux

การติดตั้งโปรแกรม R

- โปรแกรม R สามารถดาวน์โหลดได้จาก <http://cran.r-project.org>

Download and Install R

Precompiled binary distributions of the base system for these versions of R:

- [Download R for Linux](#)
- [Download R for \(Mac\) OS X](#)
- [Download R for Windows](#)

R is part of many Linux distributions, you should use those above.

Subdirectories:

- [base](#)
- [contrib](#)
- [old contrib](#)
- [Rtools](#)

3 [Download R 3.5.1 for Windows](#) (62 megabytes, 32/64 bit)

[Installation and other instructions](#)

[New features in this version](#)

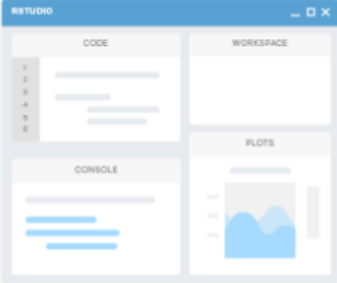
- เมื่อติดตั้งเสร็จเรียบร้อยแล้ว เราจะใช้งานโปรแกรม R ผ่าน โปรแกรม RStudio ซึ่งจะต้องติดตั้งโปรแกรม RStudio ต่อไป

การติดตั้งโปรแกรม R Studio

- เรียกมาใช้งานผ่าน R studio
- สามารถใช้งาน R ในลักษณะที่มีหน้าจอต่างๆ ให้พร้อมสำหรับการแก้ไขชุดคำสั่งต่าง ๆ หน้าจอแสดงผล หน้าจอแสดงตัวแปรต่างๆ ฯลฯ
- สามารถติดตั้ง R studio เพิ่มเติมได้โดยไปที่เว็บไซต์ <http://rstudio.com> ดาวน์โหลดโปรแกรม R studio desktop ที่เป็น open source edition มาติดตั้งการใช้ Rstudio

การติดตั้งโปรแกรม R Studio

R Studio



RStudio

RStudio makes R easier to use. It includes a code editor, debugging & visualization tools.

[Download](#) [Learn More](#)

1

RStudio Desktop
Open Source License

FREE

DOWNLOAD

[Learn More](#)

2

Installers for Supported Platforms

Installers

[RStudio 1.1.453 - Windows Vista/7/8/10](#)

[RStudio 1.1.453 - Mac OS X 10.6+ \(64-bit\)](#)

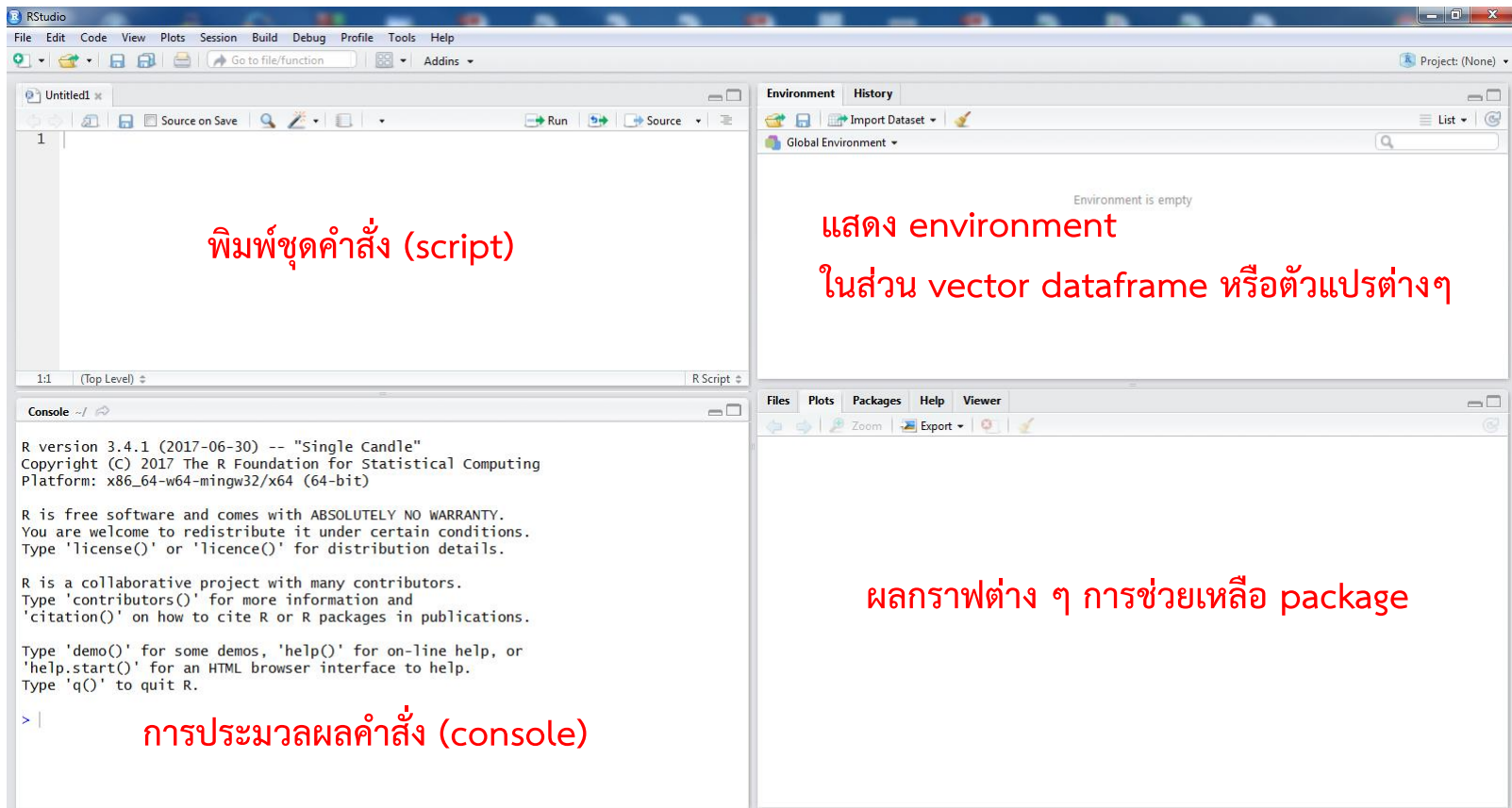
[RStudio 1.1.453 - Ubuntu 12.04-15.10/Debian 8 \(32-bit\)](#)

[RStudio 1.1.453 - Ubuntu 12.04-15.10/Debian 8 \(64-bit\)](#)

3

การติดตั้งโปรแกรม R Studio

- ได้โปรแกรม Rstudio ที่มีสัญลักษณ์ 
- เมื่อเรียกใช้งานโปรแกรม RStudio แล้วจะแสดง



การใช้งานโปรแกรม R เบื้องต้น

- การใช้งานประมวลผลคำสั่งจะใช้ส่วน console
- พิมพ์คำสั่งต่างๆ หลัง เครื่องหมาย `>` R จะประมวลผลแล้วแสดงผลลัพธ์

```
> 3+2
```

```
[1] 5
```

- สามารถเก็บค่าที่ได้ไว้ในตัวแปรแบบที่ใช้ในภาษาคอมพิวเตอร์ได้

```
> x=2+3
```

```
> x
```

```
[1] 5
```

การใช้งานโปรแกรม R เบื้องต้น

- การใช้งานประมวลผลคำสั่งจะใช้ส่วน console
- พิมพ์คำสั่งต่างๆ หลัง เครื่องหมาย > R จะประมวลผลแล้วแสดงผลลัพธ์

```
> 3+2  
[1] 5
```

- สามารถเก็บค่าที่ได้ไว้ในตัวแปรแบบที่ใช้ในภาษาคอมพิวเตอร์ได้

```
> x=2+3  
  
> x  
[1] 5
```

หรือ

```
> x <- 2+3  
  
> x  
[1] 5
```

การใช้งานโปรแกรม R เบื้องต้น

- คำสั่งช่วยเหลือ help หากมีข้อสงสัยการใช้งานคำสั่งใดให้พิมพ์ `> help(...)` หรือ เครื่องหมาย ?

```
> help("getwd")
```

```
> ?(getwd)
```

```
> getwd()
```

- คำสั่ง `getwd` เป็นคำสั่งเพื่อดูว่า working directory ณ ปัจจุบันคืออะไร
- คำสั่ง `setwd` ในการกำหนด working directory ที่ต้องการทำงาน เช่น

```
> setwd("c:/Desktop/RFiles")
```

การใช้งานโปรแกรม R เบื้องต้น

- การสร้างข้อมูล โดยใช้คำสั่ง c

```
> score = c(10,20,30,40)
```

```
> score
```

```
[1] 10 20 30 40
```

- การเลิกใช้โปรแกรม R ใช้คำสั่ง

```
> q()
```

ชนิดของข้อมูลใน R

- ข้อมูล **vector** คือ ข้อมูลเวกเตอร์คือตัวแปรหนึ่งตัวสามารถเก็บค่าได้หลายค่าใน 1 มิติในการคำนวณสามารถนำค่าทุกค่ามาคำนวณพร้อมกันได้

```
> a <- c(1,2,5.3,6,-2,4) # numeric vector
```

```
> a
```

```
[1] 1.0 2.0 5.3 6.0 -2.0 4.0
```

```
> b <- c("one","two","three") # character vector
```

```
> b
```

```
[1] "one" "two" "three"
```

```
> c <- c(TRUE,TRUE,TRUE,FALSE,TRUE,FALSE) #logical vector
```

```
> c
```

```
[1] TRUE TRUE TRUE FALSE TRUE FALSE
```

ชนิดของข้อมูลใน R

- **ข้อมูล matrices** คือตัวแปรหนึ่งตัวสามารถเก็บค่าได้หลายค่าใน 2 มิติซึ่งค่าของข้อมูลควรจะอยู่ในประเภทเดียวกันคือ (ตัวเลข ตัวอักษร หรืออื่น ๆ) และจำนวนข้อมูลในแนวแถวหรือคอลัมน์ที่เท่ากัน

รูปแบบของการกำหนดตัวแปร matrices

```
mymatrix <- matrix(vector, nrow=r, ncol=c, byrow=FALSE,  
dimnames=list(char_vector_rownames, char_vector_colnames))
```

byrow=TRUE แสดงว่าตัวแปร matrix จะถูกจัดเรียงในแนวแถว

byrow=FALSE แสดงว่าตัวแปร matrix จะถูกจัดเรียงในแนวคอลัมน์(default)

dimnames จะเป็นการกำหนดชื่อในแนวแถวและคอลัมน์ซึ่งอาจจะกำหนดหรือไม่กำหนดก็ได้

ชนิดของข้อมูลใน R

```
> y<-matrix(1:20, nrow=5,ncol=4)
> y
[,1] [,2] [,3] [,4]
[1,] 1 6 11 16
[2,] 2 7 12 17
[3,] 3 8 13 18
[4,] 4 9 14 19
[5,] 5 10 15 20
> cells <- c(1,26,24,68)
> rnames <- c("R1", "R2")
> cnames <- c("C1", "C2")
> mymatrix <- matrix(cells, nrow=2, ncol=2, byrow=TRUE,
+ dimnames=list(rnames, cnames))
> mymatrix
  C1 C2
R1 1 26
R2 24 68
```

ชนิดของข้อมูลใน R

- ข้อมูล **array** จะคล้ายกับ matrix แต่สามารถมีได้มากกว่า 2 มิติ สามารถดูรายละเอียดใน ?array
- ข้อมูล **data frame** จะมีลักษณะทั่วไปกว่าข้อมูล matrix คือในแต่ละคอลัมน์จะสามารถเป็นข้อมูลต่างชนิดกันได้ เช่น บางคอลัมน์เป็นตัวเลข บางคอลัมน์เป็นตัวอักษร

```
> d <- c(1,2,3,4)
> e <- c("red", "white", "red", NA)
> f <- c(TRUE,TRUE,TRUE,FALSE)
> mydata <- data.frame(d,e,f)
> mydata
  d e f
1 1 red TRUE
2 2 white TRUE
3 3 red TRUE
4 4 <NA> FALSE
```

การตั้งชื่อตัวแปร

- สามารถใช้ตัวอักษร ตัวเลข หรือ เครื่องหมายจุด (.) ได้
- ห้ามขึ้นต้นด้วย ตัวเลข หรือจุด (.)
- การใช้ตัวอักษรตัวใหญ่และตัวเล็กนั้นมีความหมายแตกต่างกัน

> x=2

> n=25

> a.really.long.number = 123456789

> AReallySmallNumber = 0.0000000001

การติดตั้ง packages

- package เป็นชุดคำสั่งหรือฟังก์ชันต่างๆ ที่มีผู้เขียนเพิ่มเติม
- บาง packages จะถูกติดตั้งไว้ตั้งแต่ตอนติดตั้งโปรแกรม R แต่บาง packages ผู้ใช้ต้องติดตั้งเอง
- โดยการ install package ที่ต้องการด้วยคำสั่ง `install.package(" ")` โปรแกรมจะเชื่อมต่ออินเทอร์เน็ตไปที่ <http://cran.r-project.org> เพื่อดาวน์โหลด package ต่าง ๆ ที่ระบุ
- ถ้าอยากรู้ว่าใน R มีฟังก์ชันอะไรบ้าง ไปที่

[https://cran.r-](https://cran.r-project.org/web/packages/available_packages_by_name.html)

[project.org/web/packages/available_packages_by_name.html](https://cran.r-project.org/web/packages/available_packages_by_name.html)

การติดตั้ง packages

- ตัวอย่าง การติดตั้ง packages UsingR
- เมื่อติดตั้ง package เรียบร้อยแล้ว เราสามารถ load package โดยคำสั่ง `library( )`

```
> install.package("UsingR")  
> library(UsingR)
```

- package ใดที่ติดตั้งภายหลังการลงโปรแกรม R เมื่อปิดโปรแกรมแล้ว package จะถูกปิดไปด้วย เมื่อเปิดโปรแกรมใหม่อีกครั้งแล้วต้องการใช้ package UsingR ต้องโหลด package นั้นใหม่ด้วยคำสั่ง `library()` เท่านั้น โดยไม่ต้อง install