

ศุภชัย สมพานิช

ผู้เขียนหนังสือ Bestseller ด้าน Programming

# เรียนรู้ Data Science ด้วย ภาษา R

หนึ่งในภาษายอดนิยมที่ใช้ตั้งแต่วิเคราะห์ข้อมูล (Data Analysis) ไปจนถึง Machine Learning



มหาวิทยาลัยเทคโนโลยีราชมงคลพระนคร  
ห้องสมุดทเวศร์



401016924



 **DOWNLOAD**  
provision.co.th/R

# INTRODUCTION

ในยุคปัจจุบัน ข้อมูล ข่าวสารกลายเป็นสิ่งที่มีค่ามีความหมายมากมาย จึงถือ  
เป็นอีก 1 แขนงที่ในทุกๆ สายอาชีพต้องมีสกิลนี้ติดตัวไว้ เนื้อหาในหนังสือ  
เล่มนี้ เป็นการนำเสนอการใช้งานภาษา R เป็นภาษาที่ถูกสร้างขึ้นมาเพื่อทำงาน  
กับข้อมูลโดยเฉพาะ

องค์ประกอบหนึ่งของภาษา R นั่นคือ การนำ Package ต่างๆ ที่มีความสำคัญใน  
ระดับต้นๆ เข้ามาช่วยการทำงานกับข้อมูลให้มีประสิทธิภาพดียิ่งขึ้น

หากคุณผู้อ่านท่านใดต้องการแนะนำ, ข้อติติง, เสนอแนะ รวมไปถึงข้อผิดพลาด  
ต่างๆ สามารถแจ้งผู้เขียนได้ 4 ช่องทาง คือ

1. แฟนเพจ Thaivb.NET ที่ <https://www.facebook.com/thaivb.net/>
2. กลุ่มชื่อว่า Thaivb.NET ที่ <https://www.facebook.com/groups/Thaivb.NET>
3. ช่อง YouTube ชื่อว่า Thaivb.NET
4. แฟนเพจ Thaivb.NET ที่ <https://www.blockdit.com/thaivb.net>

ศุภชัย สมพานิช

มกราคม 2564



**DOWNLOAD**

[provision.co.th/R](https://provision.co.th/R)

โครงสร้างข้อมูลแบบ Vector .....	56
การอ่านข้อมูลจากเวกเตอร์.....	61
การวนลูปอ่านค่าในเวกเตอร์.....	68
การทำงานกับข้อมูลเวกเตอร์หลายชุด.....	70
การคำนวณยอดรวมทั้งหมด ด้วยฟังก์ชัน sum() .....	72
การคำนวณสัดส่วนเปอร์เซ็นต์ ของยอดขาย .....	73
การอ่านข้อมูลจากเวกเตอร์หลายชุด แบบมีเงื่อนไข.....	76
การนำเวกเตอร์มารวมกัน .....	78
กรณีเวกเตอร์เลขจำนวนเต็ม กับเลขจำนวนเต็ม.....	78
กรณีเวกเตอร์เลขจำนวนเต็ม กับเลขทศนิยม.....	79
กรณีเวกเตอร์ข้อความกับตัวเลข.....	80
การรวมเวกเตอร์ด้วยฟังก์ชัน paste().....	82
การเปรียบเทียบข้อมูลในเวกเตอร์ เท่ากันหรือไม่ด้วยเครื่องหมาย == .....	83
การบวกกันของเวกเตอร์ตัวเลข.....	84
พื้นฐานการหาค่าสูงสุด, ต่ำสุด, ผลรวม และค่าเฉลี่ยจากเวกเตอร์ตัวเลข .....	86

การสร้างกรอบข้อมูลจากเวกเตอร์.....	90
การอ่านข้อมูลจาก Data Frame .....	92
พื้นฐานการอ่านข้อมูลใน Data Frame แบบมีเงื่อนไข.....	94
การอ่านข้อมูลจาก Data Frame แบบมีเงื่อนไขด้วยฟังก์ชัน subset().....	96
พื้นฐานการแก้ไขข้อมูลใน Data Frame.....	98
การเพิ่มหรือลบคอลัมน์ใน Data Frame..	100
การเพิ่มหรือลบแถวใน Data Frame .....	104
การทำงานกับแถวข้อมูล ที่มีค่าผิดพลาด NA.....	105
การสลับตำแหน่งคอลัมน์.....	106
การเปลี่ยนชื่อคอลัมน์ใน Data Frame ด้วยฟังก์ชัน names().....	107
การทำงานกับ Data Frame แบบถาวรด้วยฟังก์ชัน attach().....	108
การเรียงลำดับข้อมูลใน Data Frame ด้วยฟังก์ชัน order().....	109
การทำงานหลายอย่างกับ Data Frame ด้วยฟังก์ชัน with().....	111

# Contents

## CHAPTER 05 การสร้างฟังก์ชัน (Function) ขึ้นมาใช้งาน

พื้นฐานการสร้างฟังก์ชัน (Function) ใน R.....	114
การสร้างฟังก์ชันที่มีพารามิเตอร์หลายตัว.....	116
การสร้างฟังก์ชันรับค่าจากผู้ใช้งานโดยตรง.....	118
การสร้างฟังก์ชันสำหรับคำนวณพื้นที่.....	119
การสุ่มข้อมูลด้วยฟังก์ชัน sample().....	120
การสุ่มข้อมูลแบบถ่วงน้ำหนัก.....	124

## CHAPTER 06 การ Import และ Export ข้อมูลจากไฟล์ Excel, CSV และไฟล์ข้อความ

การ Import ข้อมูลจากไฟล์ข้อความ (Text File).....	127
การ Import ข้อมูลจากไฟล์ CSV.....	131
การ Import ข้อมูลจากไฟล์ Excel.....	133
การ Export ข้อมูลจาก Data Frame เป็นไฟล์ CSV.....	135

## CHAPTER 07 พื้นฐานการใช้ R ร่วมกับข้อมูลในระบบฐานข้อมูล SQL Server

การดาวน์โหลดและติดตั้งฐานข้อมูล SQL Server.....	141
การติดตั้งตัวจัดการฐานข้อมูล Microsoft SQL Server Management Studio (SSMS).....	146
การเชื่อมต่อฐานข้อมูล SQL Server ด้วย SSMS.....	148
การกำหนดให้ SSMS สามารถแก้ไขโครงสร้างฐานข้อมูลได้.....	148
การสร้างฐานข้อมูลใหม่ด้วย SSMS.....	149
การเพิ่ม/แก้ไขข้อมูลในตาราง.....	152
การดูข้อมูลที่อยู่ในตาราง.....	153
การดูโครงสร้างตาราง.....	153
การกำหนดให้ฐานข้อมูล SQL Server เข้ารหัสภาษาไทย.....	154
การดาวน์โหลดและติดตั้ง ODBC Driver สำหรับฐานข้อมูล SQL Server.....	155
การดาวน์โหลดและติดตั้ง RODBC สำหรับภาษา R.....	156
การเชื่อมต่อเข้ากับฐานข้อมูล SQL Server ด้วย R.....	157



CHAPTER

## 08 การทำงานกับข้อมูลในฐานข้อมูล SQL Server

พื้นฐานการคิวรีข้อมูลจากฐานข้อมูล SQL Server แบบ 1 ตาราง.....	159
การทำงานกับข้อมูลภาษาไทยจากฐานข้อมูล SQL Server ด้วยวิธี Export ข้อมูลเป็น Excel.....	161
การทำงานกับข้อมูลตั้งแต่ 2 ตารางขึ้นไป.....	166
การสร้าง View ในฐานข้อมูล SQL Server สำหรับข้อมูลตั้งแต่ 2 ตารางขึ้นไป.....	169

CHAPTER

## 09 ภาษา R กับข้อมูลในฐานข้อมูล MongoDB

การดาวน์โหลดและติดตั้งฐานข้อมูล MongoDB.....	177
การเชื่อมต่อกับฐานข้อมูล MongoDB.....	178
การสร้างฐานข้อมูลและคอลเล็กชันจัดเก็บข้อมูล.....	179
การปรับมุมมองข้อมูล.....	182
การดาวน์โหลดและติดตั้ง mongolite.....	183
การเพิ่มข้อมูลในฐานข้อมูล MongoDB ด้วยภาษา R.....	184
การแสดงข้อมูลจากฐานข้อมูล MongoDB ด้วยฟังก์ชัน View().....	185

การเพิ่มข้อมูลที่มีโครงสร้างแตกต่างกัน.....	186
การแก้ไขข้อมูลที่มีโครงสร้างเหมือนกัน.....	187
การแก้ไขข้อมูลที่มีโครงสร้างแตกต่างกัน.....	189
การลบข้อมูล.....	190

CHAPTER

## 10 สถิติเบื้องต้นกับภาษา R

ทำความเข้าใจกับ “ประชากร” (Population) และ “ตัวอย่าง” (Sample).....	192
ค่ากลางของข้อมูล.....	193
ค่าเฉลี่ย (Mean).....	193
ค่าเฉลี่ยเลขคณิต (Arithmetic Mean) แบบไม่แจกแจงความถี่.....	193
การหาค่าเฉลี่ยเลขคณิตกรณีมีข้อมูลผิดพลาด (NA).....	195
ค่าเฉลี่ยเลขคณิตกับค่าผิดปกติ (Outliers).....	197
ค่าเฉลี่ยถ่วงน้ำหนัก (Weighted Mean) แบบไม่แจกแจงความถี่.....	199
ค่าเฉลี่ยเรขาคณิต (Geometric Mean) แบบไม่แจกแจงความถี่.....	201

# Contents

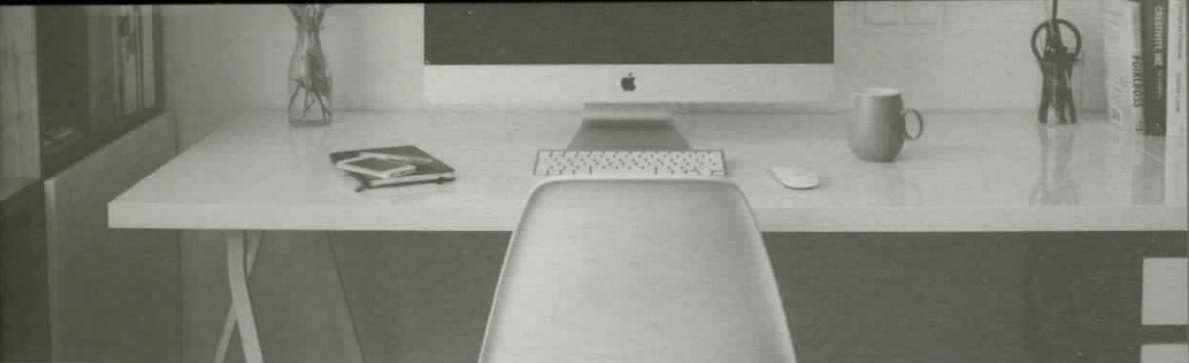
การคำนวณค่าเฉลี่ยเรขาคณิต โดยอาศัย psych.....	206
ค่าเฉลี่ยฮาร์โมนิก (Harmonic Mean) แบบไม่แจกแจงความถี่ของข้อมูล .....	208
มัธยฐาน (Median) .....	210
การหาค่ากลางมัธยฐาน (Median) แบบไม่แจกแจงความถี่ของข้อมูล .....	210
การหาค่าฐานนิยม (Mode) แบบไม่แจกแจงความถี่.....	213

## CHAPTER 11 การหาค่าแห่งของข้อมูล (Measures Of Variability)

ควอร์ไทล์ (Quartile), เดไซล์ (Decile) และ เปอร์เซ็นไทล์ (Percentile).....	217
ควอร์ไทล์ (Quartile).....	217
เดไซล์ (Decile).....	218
เปอร์เซ็นไทล์ (Percentile).....	218
วิธีการหาควอร์ไทล์, เดไซล์ และเปอร์เซ็นไทล์แบบไม่แจกแจงความถี่.....	219
การคำนวณควอร์ไทล์, เดไซล์ และ เปอร์เซ็นไทล์ด้วยฟังก์ชัน <code>quantile()</code> .....	223

## CHAPTER 12 การวัดกระจายของข้อมูล (Measure of Dispersion)

ปัญหาของค่ากลางของข้อมูล.....	226
การวัดการกระจายสัมบูรณ์ (Absolute Variation).....	228
พิสัย (Range) แบบไม่แจกแจงความถี่.....	228
ส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation) แบบไม่แจกแจงความถี่.....	229
ส่วนเบี่ยงเบนเฉลี่ย (Mean Deviation) แบบไม่แจกแจงความถี่.....	232
ส่วนเบี่ยงเบนควอร์ไทล์ (Quartile Deviation) แบบไม่แจกแจงความถี่.....	235
การวัดการกระจายสัมพัทธ์ (Relative Variation).....	237
สัมประสิทธิ์ของพิสัย แบบไม่แจกแจงความถี่.....	237
สัมประสิทธิ์ของการแปรผัน แบบไม่แจกแจงความถี่.....	239
สัมประสิทธิ์ของส่วนเบี่ยงเบนเฉลี่ย แบบไม่แจกแจงความถี่.....	240
สัมประสิทธิ์ของส่วนเบี่ยงเบนควอร์ไทล์ แบบไม่แจกแจงความถี่.....	242
การแปลความหมายของ การวัดการกระจายสัมพัทธ์.....	244



CHAPTER

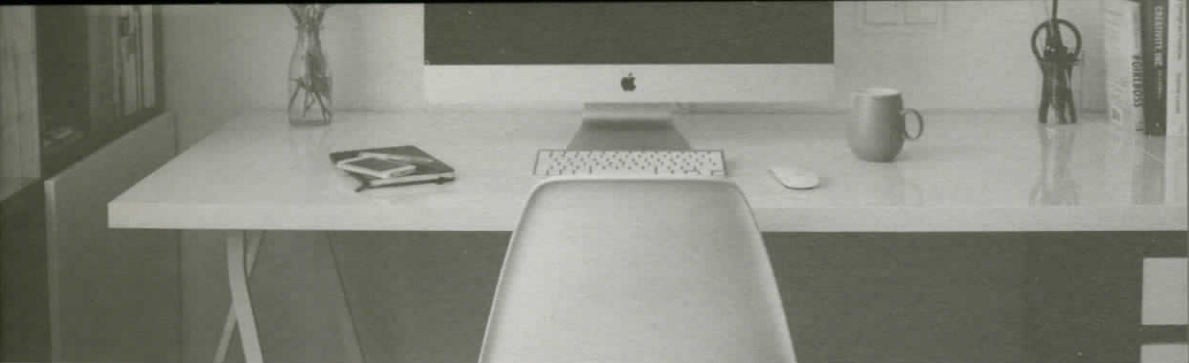
# 13 ▶ เมทริกซ์ (Matrix)

โครงสร้างข้อมูลแบบเมทริกซ์ (Matrix).....	248
การแปลงเวกเตอร์เป็นเมทริกซ์ ด้วยฟังก์ชัน dim().....	249
พื้นฐานการจัดการข้อมูลในเมทริกซ์ ด้วยฟังก์ชัน sweep().....	251
การทำงานกับข้อมูลในเมทริกซ์ แนวแถวหรือคอลัมน์.....	252
การหาผลรวมคอลัมน์ ด้วยฟังก์ชัน colSums().....	255
การหาผลรวมแถว ด้วยฟังก์ชัน rowSums().....	256
การหาค่าเฉลี่ยเลขคณิตตาม แนวคอลัมน์และแถวด้วยฟังก์ชัน colMeans() กับ rowMeans().....	257
การเพิ่มคอลัมน์ใหม่ในเมทริกซ์ ด้วยฟังก์ชัน cbind().....	258
การเพิ่มแถวใหม่ในเมทริกซ์ ด้วยฟังก์ชัน rbind().....	260
การหาค่ามัธยฐาน (Median) ในเมทริกซ์.....	261

CHAPTER

# 14 ▶ การทำงานกับ ข้อมูลที่นำเสนอใจ

การหาค่าเฉลี่ยเลขคณิตแบบจัดกลุ่ม .....	264
พื้นฐานการหาค่าเฉลี่ยถ่วงน้ำหนัก กับข้อมูลด้วยฟังก์ชัน weighted.mean().....	266
การหาค่าเฉลี่ยถ่วงน้ำหนักของข้อมูล ที่ถูกจัดกลุ่ม .....	267
การตรวจสอบค่าในช่วงที่ระบุ ด้วยฟังก์ชัน between() ของ dplyr.....	268
การคำนวณค่าทางสถิติตามแถวหรือคอลัมน์ ด้วยฟังก์ชัน apply().....	270
การคำนวณค่าทางสถิติเบื้องต้น แบบแบ่งกลุ่มด้วยฟังก์ชัน tapply().....	272
การใส่หมายเลขบรรทัด ด้วยฟังก์ชัน row_number() ของ dplyr.....	273
การจัดลำดับแบบแบ่งกลุ่มข้อมูล ด้วยฟังก์ชัน ntile() ของ dplyr.....	274
การเรียงลำดับข้อมูลแบบ Percent Rank ด้วยฟังก์ชัน percent_rank() ของ dplyr.....	275
การคำนวณค่าสถิติเบื้องต้น ด้วยฟังก์ชัน summary().....	277
การคำนวณค่าสถิติเบื้องต้น ด้วยฟังก์ชัน summarise() ของ dplyr .....	279



CHAPTER

# 13 ▶ เมทริกซ์ (Matrix)

โครงสร้างข้อมูลแบบเมทริกซ์ (Matrix).....	248
การแปลงเวกเตอร์เป็นเมทริกซ์ ด้วยฟังก์ชัน dim().....	249
พื้นฐานการจัดการข้อมูลในเมทริกซ์ ด้วยฟังก์ชัน sweep().....	251
การทำงานกับข้อมูลในเมทริกซ์ แนวแถวหรือคอลัมน์.....	252
การหาผลรวมคอลัมน์ ด้วยฟังก์ชัน colSums().....	255
การหาผลรวมแถว ด้วยฟังก์ชัน rowSums().....	256
การหาค่าเฉลี่ยเลขคณิตตาม แนวคอลัมน์และแถวด้วยฟังก์ชัน colMeans() กับ rowMeans().....	257
การเพิ่มคอลัมน์ใหม่ในเมทริกซ์ ด้วยฟังก์ชัน cbind().....	258
การเพิ่มแถวใหม่ในเมทริกซ์ ด้วยฟังก์ชัน rbind().....	260
การหาค่ามัธยฐาน (Median) ในเมทริกซ์.....	261

CHAPTER

# 14 ▶ การทำงานกับ ข้อมูลที่นำเสนอใจ

การหาค่าเฉลี่ยเลขคณิตแบบจัดกลุ่ม .....	264
พื้นฐานการหาค่าเฉลี่ยถ่วงน้ำหนัก กับข้อมูลด้วยฟังก์ชัน weighted.mean().....	266
การหาค่าเฉลี่ยถ่วงน้ำหนักของข้อมูล ที่ถูกจัดกลุ่ม .....	267
การตรวจสอบค่าในช่วงที่ระบุ ด้วยฟังก์ชัน between() ของ dplyr.....	268
การคำนวณค่าทางสถิติตามแถวหรือคอลัมน์ ด้วยฟังก์ชัน apply().....	270
การคำนวณค่าทางสถิติเบื้องต้น แบบแบ่งกลุ่มด้วยฟังก์ชัน tapply().....	272
การใส่หมายเลขบรรทัด ด้วยฟังก์ชัน row_number() ของ dplyr.....	273
การจัดลำดับแบบแบ่งกลุ่มข้อมูล ด้วยฟังก์ชัน ntile() ของ dplyr.....	274
การเรียงลำดับข้อมูลแบบ Percent Rank ด้วยฟังก์ชัน percent_rank() ของ dplyr.....	275
การคำนวณค่าสถิติเบื้องต้น ด้วยฟังก์ชัน summary().....	277
การคำนวณค่าสถิติเบื้องต้น ด้วยฟังก์ชัน summarise() ของ dplyr .....	279




สามารถยืมและติดตามหนังสือใหม่ได้ที่ ระบบห้องสมุดอัตโนมัติ Walai Autolib

<https://lib.rmutp.ac.th/catalog/Bibltem.aspx?BibID=b00106741>



### เรียนรู้ Data science ด้วยภาษา R / ศุภชัย สมพานิช.

Author	ศุภชัย สมพานิช
Published	กรุงเทพฯ : โปริชั่น, 2564
Edition	พิมพ์ครั้งที่ 1
Detail	320 หน้า : ภาพประกอบ ; 23 ซม
Subject	การประมวลผลข้อมูลอิเล็กทรอนิกส์(+) การเรียนรู้ของเครื่อง(+) โปรแกรมประยุกต์บนเว็บ(+) อาร์ (ภาษาคอมพิวเตอร์)(+)
ISBN	9786162047831
ประเภทแหล่งที่มา	 Book



สำหรับการศึกษาระดับปริญญาตรีและการอ้างอิง