

อรวิน ประทีปสุกรี

ผู้เขียนหนังสือ Bestseller ด้าน Programming

PROVISION

Python สำหรับงาน Data Science Data Visualization และ Machine Learning

วิธีใช้งาน Library ยอดนิยม Pandas, NumPy, Matplotlib และ Scikit-learn
ตัวอย่างการใช้ Machine Learning : ทำนายยอดขาย • วิเคราะห์แนวโน้มระยะเร่ง • อนุมัติเงินกู้
Sentimental Analysis (วิเคราะห์ข้อความ) • Image Recognition • Customer Segmentation



มหาวิทยาลัยเทคโนโลยีราชมงคลพระนคร

ห้องสมุดพระนครเหนือ



501032041



DOWNLOAD

[provision.co.th/
Python_Datascience](http://provision.co.th/Python_Datascience)

Contents

CHAPTER

01 ทำความรู้จักกับ Data Science

Big Data และ Data Science คืออะไร.....	12
ขั้นตอนการทำ Data Science (Data Science Pipeline).....	13
Obtain.....	14
Scrub.....	14
Explore.....	15
Model.....	15
Intepret.....	16
Data Science กับ Python.....	17
ไลบรารีพื้นฐานสำหรับการทำ Data Science ของ Python.....	18
Pandas.....	18
NumPy.....	18
SciPy.....	19
Matplotlib.....	19
Seaborn.....	19

CHAPTER

02 การใช้งาน Jupyter Notebook และ Google Colab

ทำความรู้จักกับ Jupyter Notebook.....	21
Jupyter Notebook กับ Google Colab ต่างกันอย่างไร.....	21
การติดตั้งโปรแกรม Anaconda.....	22
การใช้งาน Jupyter Notebook.....	25
การสร้าง folder เพื่อจัดเก็บโค้ดโปรแกรม.....	27
การเขียนโปรแกรมลงบน Jupyter Notebook.....	28
การลบและเพิ่มเซลล์.....	29

การคัดลอก/ตัดเซลล์เดิมไปไว้ยังเซลล์ใหม่.....	31
การสลับที่เซลล์.....	33
การรวมเซลล์และแตกเซลล์.....	34
การบันทึกโค้ดโปรแกรมใน Jupyter Notebook.....	36
การเปิดไฟล์โค้ดโปรแกรมใน Jupyter Notebook.....	38
การเรียกใช้งาน Google Colab.....	39

CHAPTER

03 พื้นฐานไพธอน

การเขียนคอมเมนต์ในไพธอน.....	43
กฎการตั้งชื่อของภาษาไพธอน.....	43
Keywords ในภาษาไพธอน.....	44
ทำความรู้จักกับตัวแปร.....	44
การรับและแสดงผลข้อมูล.....	45
ชนิดข้อมูลของไพธอน.....	45
Numeric.....	45
Integers.....	45
Floating point numbers.....	45
Complex numbers.....	46
Boolean.....	47
String.....	47
การหาความยาวสตริง.....	47
การค้นหาตัวอักษรและกลุ่มตัวอักษรในสตริง.....	49
List.....	49
การหาความยาวลิสต์.....	49
การค้นหาตัวอักษรและกลุ่มตัวอักษรในลิสต์.....	49
การแก้ไขข้อมูลในลิสต์.....	51



การเพิ่มข้อมูลในลิสต์.....	51	คำสั่งทำซ้ำ.....	81
การลบข้อมูลในลิสต์.....	52	คำสั่งทำซ้ำ while	81
Tuple.....	52	การทำงานของคำสั่ง while	ร่วมกับคำสั่ง break และ continue.....
Set	53	การทำงานของคำสั่ง while	ร่วมกับคำสั่ง else
การหาความยาวเซต.....	53	คำสั่งทำซ้ำ for.....	83
การค้นหาข้อมูลในเซต.....	53	การทำงานของคำสั่ง for	ร่วมกับฟังก์ชัน range().....
การเพิ่มข้อมูลในเซต.....	54	การทำงานของคำสั่ง for	ร่วมกับคำสั่ง break และ continue.....
การลบข้อมูลในเซต.....	54	การทำงานของคำสั่งทำซ้ำ for	ร่วมกับคำสั่ง else
Dictionary.....	55	การสร้างและเรียกใช้ฟังก์ชัน.....	86
การหาความยาวดิกชันนารี.....	56	วิธีเรียกใช้งาน Built-in modules.....	88
การค้นหาข้อมูลในดิกชันนารี.....	56	เรียกใช้โมดูลด้วยคำสั่ง import.....	88
การแก้ไขข้อมูลในดิกชันนารี.....	57	เรียกใช้โมดูลด้วยคำสั่ง from ... import....	88
การเพิ่มข้อมูลในดิกชันนารี.....	57		
การลบข้อมูลในดิกชันนารี.....	57		
ตัวดำเนินการ (Operator) ในไพธอน.....	58		
ตัวดำเนินการทางคณิตศาสตร์			
(Arithmetic operators).....	58		
ตัวดำเนินการทางตรรกศาสตร์ (Logical operators)...	59		
ตัวดำเนินการระดับบิต (Bitwise operators).....	60		
ตัวดำเนินการกำหนดค่า (Assignment operators)....	63		
ตัวดำเนินการเปรียบเทียบ (Comparison operators)...	64		
ตัวดำเนินการเอกลักษณ์ (Identity operators).....	65		
ตัวดำเนินการสมาชิก (Membership operators)...	66		
ลำดับความสำคัญของตัวดำเนินการ.....	67		
การแปลงชนิดข้อมูล.....	69		
Implicit type conversion	69		
Explicit type conversion	69		
คำสั่งเงื่อนไข.....	74		
คำสั่งเงื่อนไข if	74		
คำสั่งเงื่อนไข if-else	75		
คำสั่งเงื่อนไข if-elif-else	77		
คำสั่งเงื่อนไข nested-if	79		
		การสร้าง ndarray	
		จากข้อมูลชนิดอื่นของ Python.....	91
		การใช้ฟังก์ชันต่างๆ ของ NumPy	
		สร้าง ndarray.....	92
		zeros() และ ones().....	92
		full().....	93
		arange().....	94
		linspace().....	96

CHAPTER
04 รู้จักกับไลบรารี
NumPy

มาทำความรู้จักกับไลบรารี NumPy กัน.....	90
การติดตั้ง NumPy.....	90
การสร้าง ndarray.....	90
การสร้าง ndarray	
จากข้อมูลชนิดอื่นของ Python.....	91
การใช้ฟังก์ชันต่างๆ ของ NumPy	
สร้าง ndarray.....	92
zeros() และ ones().....	92
full().....	93
arange().....	94
linspace().....	96

eye()	98
empty()	99
การสร้าง ndarray จากเลขสุ่ม ด้วยโมดูล Random	100
การเรียกใช้โมดูล Random	100
rand()	101
randint()	103
choice()	105
การกำหนดการกระจายของเลขสุ่มด้วย choice() (Random Distribution)	106
การสลับตำแหน่งสมาชิกแบบสุ่มใน ndarray (Shuffling & Premutation)	107
โอเปอเรชันกับ ndarray (Vectorization & Broadcasting)	108
Vectorization	108
Broadcasting	110
กฎการทำโอเปอเรชันของ ndarray	111
ndarray 2 ตัว มีจำนวนมิติเท่ากัน แต่จำนวนสมาชิกไม่เท่ากันในบางมิติ	112
ndarray 2 ตัว มีจำนวนมิติไม่เท่ากัน แต่จำนวนสมาชิกเท่ากันในมิติหนึ่ง หรือหลายมิติ	114
ndarray 2 ตัว มีจำนวนมิติเท่ากัน และจำนวนสมาชิกในต่างมิติเท่ากัน	117
ndarray 2 ตัว ซึ่งตัวใดตัวหนึ่ง มีจำนวนสมาชิกเป็น 1	118
ndarray 2 ตัวที่ไม่สามารถนำมา ทำโอเปอเรชันกันได้	118
การเข้าถึงสมาชิกใน ndarray (Array Indexing)	119
การเข้าถึงสมาชิกใน ndarray มากกว่า 1 มิติ	119
Index Slicing	120

การทำ Index Slicing ด้วย operator	121
การทำ Index Slicing ด้วย Integer Index Arrays	124
การทำ Index Slicing ด้วย Boolean Index Arrays	127
ฟังก์ชันการทำงานอื่นๆ ของ NumPy	128
round()	128
abs()	129
power(), log() และ exp()	129
sum()	130
sqrt()	131
logical_and(), logical_or(), logical_not() และ logical_xor()	131
all() และ any()	132
min(), max(), mean() และ median()	132
sort()	133

CHAPTER

05

โครงสร้างข้อมูล
ของ Pandas

Pandas คืออะไร	136
รู้จักกับโครงสร้างข้อมูลของ Pandas	136
Series	136
DataFrame	137
Panel	138
เรียนรู้การสร้างข้อมูลแบบ Series	138
การสร้างข้อมูลแบบ Series จาก List	139
การสร้างข้อมูลแบบ Series โดยกำหนด index	140
การสร้างข้อมูลแบบ Series โดยกำหนด dtype	141
การสร้างข้อมูลแบบ Series จาก Tuple	141
การสร้างข้อมูลแบบ Series จาก Dictionary	142

การสร้างข้อมูลแบบ Series
จาก ndarray ของ NumPy..... 142

เรียนรู้การสร้างข้อมูลแบบ DataFrame..... 143

การสร้างข้อมูลแบบ DataFrame
จาก Series 144

*การสร้าง DataFrame
 จากข้อมูล Series ชุดเดียว..... 144*

*การสร้าง DataFrame
 จากข้อมูล Series หลายชุด..... 144*

การสร้างข้อมูลแบบ DataFrame จาก List ... 146

การสร้างข้อมูลแบบ DataFrame
จาก List ของ Dictionary 146

การสร้างข้อมูลแบบ DataFrame
จาก ndarray ของ NumPy..... 147

การสร้างข้อมูลแบบ DataFrame
จากไฟล์ Excel..... 148

การอ่านข้อมูลจากไฟล์ Excel..... 149

การสร้างข้อมูลแบบ DataFrame
จากไฟล์ CSV 156

การอ่านและเขียนข้อมูลไฟล์ CSV..... 156

 การกำหนดค่า na_position..... 166

 การกำหนดค่า inplace 166

การเพิ่มข้อมูลใน Series 167

การลบข้อมูลใน Series..... 169

การอัปเดตข้อมูลใน Series 170

การดึงข้อมูลใน DataFrame..... 170

 การดึงข้อมูลด้วย head() และ tail()..... 170

 การดึงข้อมูลด้วย at() และ iat()..... 172

 การดึงข้อมูลด้วย loc() และ iloc() 172

 การดึงข้อมูล 1 คอลัมน์ของทุกแถว..... 173

 การดึงข้อมูลหลายคอลัมน์ของทุกแถว 173

 การดึงข้อมูล 1 แถวของทุกคอลัมน์..... 174

 การดึงข้อมูลหลายแถวของทุกคอลัมน์ 175

 การดึงข้อมูล 1 ค่าจากแถวและคอลัมน์ 176

 การดึงข้อมูลหลายแถวหลายคอลัมน์..... 177

 การดึงแถวข้อมูลโดยกำหนดเงื่อนไข 178

 การดึงแถวข้อมูลด้วยฟังก์ชัน iteritems()... 179

 การดึงแถวข้อมูลด้วยฟังก์ชัน iterrows() 180

 การดึงแถวข้อมูลด้วยฟังก์ชัน itertuples()... 181

การเรียงลำดับข้อมูลใน DataFrame..... 182

 การเรียงลำดับค่า index 182

 การเรียงลำดับค่าข้อมูล
 โดยพิจารณาจากคอลัมน์เดียว..... 183

 การเรียงลำดับค่าข้อมูล
 โดยพิจารณาจากหลายๆ คอลัมน์..... 184

การเปลี่ยนชื่อคอลัมน์ใน DataFrame..... 185

การเพิ่มคอลัมน์ใน DataFrame 186

 เพิ่มคอลัมน์ด้วยวิธีสร้างลิสต์ของคอลัมน์ 186

 เพิ่มคอลัมน์ด้วยฟังก์ชัน insert()..... 187

 เพิ่มคอลัมน์ด้วยฟังก์ชัน assign() 188

CHAPTER
06 การจัดการกับข้อมูล
Series และ DataFrame
ด้วย Pandas

การดึงข้อมูลใน Series 159

การดึงกลุ่มข้อมูลใน Series 162

 การดึงกลุ่มข้อมูลโดยระบุตำแหน่ง
 ของข้อมูล 164

 การดึงกลุ่มข้อมูลด้วย loc()..... 164

การเรียงลำดับข้อมูลใน Series..... 165

 การกำหนดค่า ascending..... 166

Contents

การเพิ่มแถวใน DataFrame.....	188
การลบคอลัมน์และแถวใน DataFrame.....	190
การอัปเดตข้อมูลใน DataFrame.....	192
การจัดกลุ่มข้อมูลด้วย groupby().....	193
การจัดกลุ่มข้อมูลใน Series.....	195
การจัดกลุ่มข้อมูลใน DataFrame ด้วยคอลัมน์เดียว.....	196
การจัดกลุ่มข้อมูลใน DataFrame ด้วยคอลัมน์หลายคอลัมน์.....	199

CHAPTER

07 ฟังก์ชันของ Series และ DataFrame ใน Pandas

ฟังก์ชันการทำงานที่น่าสนใจของ Series.....	202
ฟังก์ชันทางคณิตศาสตร์และสถิติ.....	202
ฟังก์ชันตรวจสอบข้อมูลและทำงานกับข้อมูลใน Series.....	204
ฟังก์ชันเกี่ยวกับการทำงานกับ missing values.....	207
missing values คืออะไร.....	207
ฟังก์ชันจัดการกับ index.....	209
ฟังก์ชันการทำงานที่น่าสนใจของ DataFrame.....	212
ฟังก์ชันทางคณิตศาสตร์และสถิติ.....	212
ฟังก์ชันตรวจสอบข้อมูลและทำงานกับข้อมูลใน DataFrame.....	214
ฟังก์ชันเกี่ยวกับการทำงานกับ missing value.....	220
ฟังก์ชันจัดการกับ index.....	221

CHAPTER

08 การสร้าง Pivot Table ด้วย Pandas

รู้จักกับ Pivot Table ใน Excel.....	223
การสร้างรายงานด้วย PivotTable.....	224
การเลือกข้อมูลเพื่อแสดงใน PivotTable.....	225
Pivot Table ใน Pandas.....	227
สร้าง Pivot Table ด้วย index ตัวเดียว.....	228
สร้าง Pivot Table ด้วย index หลายตัว.....	229
การเพิ่มคอลัมน์ลงใน PivotTable.....	230
การจัดการกับ missing value ใน PivotTable.....	231
การแสดงค่าผลรวมของกลุ่มใน PivotTable.....	232
การกรองข้อมูล (Filter) ใน PivotTable.....	233
สร้าง Pivot Table โดยไม่กำหนดค่า values.....	234
สร้าง Pivot Table โดยกำหนด aggfunc หลายตัว.....	235

CHAPTER

09 การทำ Data Visualization ด้วย Pandas

Data Visualization คืออะไร.....	238
Library สำคัญในการทำ Data Visualization.....	238
Matplotlib.....	238
Seaborn.....	239
การเลือกรูปแบบ Visualization ให้เหมาะสมกับข้อมูล.....	239
การพล็อตกราฟด้วยฟังก์ชัน plot().....	239
kind : กำหนดประเภทกราฟ.....	240
x : กำหนดค่าที่จะแสดงในแกน x.....	
y : กำหนดคอลัมน์ที่ต้องการพล็อตกราฟร่วมกับแกน x.....	240

grid : ตีเส้นตารางบนกราฟ.....	241
legend : กำหนดคำอธิบายสัญลักษณ์ ในกราฟ.....	242
subplots : พล็อตกราฟแยกตามคอลัมน์ ...	242
use_index : กำหนดลำดับข้อมูล บนแกน x ด้วย index.....	243
fontsize : กำหนดขนาดฟอนต์.....	244
rot : กำหนดมุมในการหมุนเลเบล ของแกน x.....	244
stacked : พล็อตกราฟแบบวางซ้อนกัน เป็นชั้น.....	245
title : กำหนดหัวเรื่องกราฟ.....	245
layout : กำหนดเลย์เอาต์ในการแสดงกราฟ	246
figsize : กำหนดขนาดกราฟ.....	246
xlabel และ ylabel : กำหนดชื่อเลเบล ของแกน x,y.....	247
position : กำหนดตำแหน่งการจัดวาง ของ bar charts.....	247
xticks และ yticks : กำหนดค่า ลำดับข้อมูลของแกน x,y.....	249
color : กำหนดสีของกราฟ.....	249
ประเภทของกราฟหรือแผนภูมิใน Pandas.....	250
Bar charts.....	250
Horizontal bar charts.....	252
Line graphs.....	253
Area plots.....	256
Pie charts.....	257
Box plots.....	260
Histogram.....	264
Scatter plots.....	269
Hexbin plots.....	273
Kernel Density Estimate charts.....	278

รู้จักกับไลบรารี Matplotlib.....	282
การบันทึกกราฟออกมาเป็นไฟล์.....	284
เรียนรู้การทำงานของฟังก์ชัน plot() ใน Matplotlib.....	286
linestyle : กำหนดรูปแบบเส้นกราฟ.....	286
linewidth : กำหนดความหนา ของเส้นกราฟ.....	287
marker : กำหนดตัวเน้นจุดพิกัด.....	287
markersize : กำหนดขนาดของ ตัวเน้นจุดพิกัด.....	293
markeredgecolor : กำหนดสีขอบ ของตัวเน้นจุดพิกัด.....	294
markerfacecolor : กำหนดสีด้านใน ของตัวเน้นจุดพิกัด.....	294
color : การกำหนดสีของกราฟ.....	295
format strings : กำหนดรายละเอียด ให้กับ marker แบบย่อ.....	295
title : กำหนดหัวเรื่องกราฟ.....	296
xlabel และ ylabel : กำหนดชื่อเลเบล ของแกน x,y.....	296
xticks และ yticks : กำหนดค่าลำดับ ข้อมูลของแกน x,y.....	297
figure : ปรับแต่ง container สำหรับการพล็อตกราฟ.....	298
การปรับขนาดรูปภาพกราฟ.....	298
การพล็อตกราฟหลายตัวด้วย subplots.....	299
legend : กำหนดคำอธิบายสัญลักษณ์ ในกราฟ.....	302

Contents

กำหนดสไตล์ของกราฟด้วยฟังก์ชัน style().....	304
กำหนดใช้ฟอนต์ภาษาไทยในกราฟ.....	306
การกำหนดฟอนต์สำหรับออบเจ็กต์ต่างๆ.....	306
การกำหนดฟอนต์สำหรับใช้งานกับกราฟทุกตัว.....	307
การสร้าง Bar charts.....	308
การกำหนดรูปแบบลายของแท่งกราฟ.....	310
การจัดกลุ่มแท่งกราฟ.....	314
การสร้าง Horizontal bar charts.....	318
การสร้าง line graphs.....	319
การสร้าง Area plots.....	320
การสร้าง Pie charts.....	321
การสร้าง Box plots.....	323
การสร้าง Histogram.....	327
การสร้าง Scatter plots.....	328
การสร้าง Hexbin plots.....	330
การสร้าง Kernel Density Estimate charts.....	332

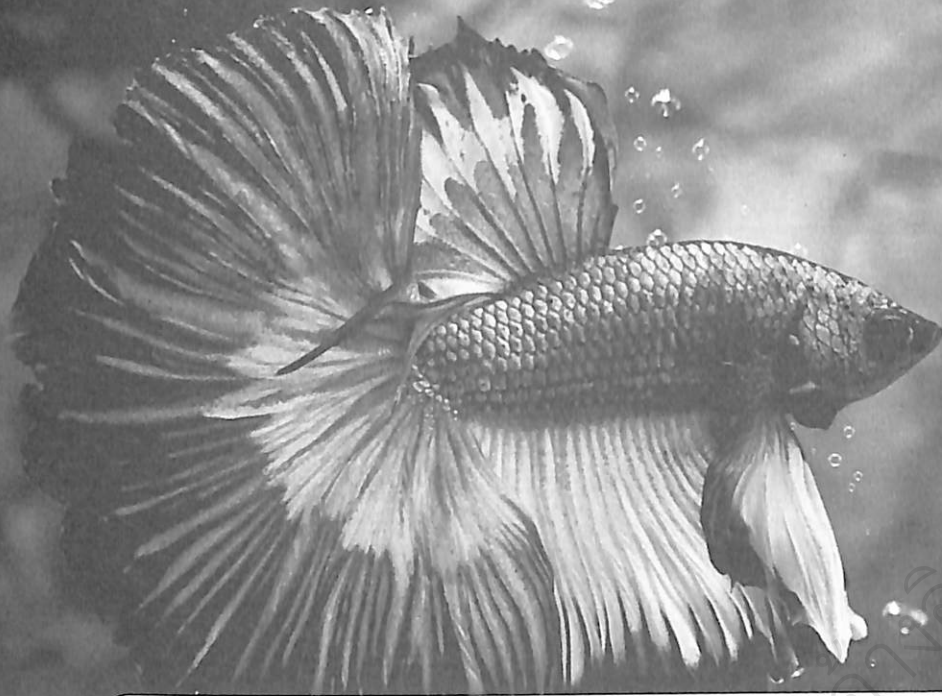
CHAPTER

11

Machine Learning (ML)

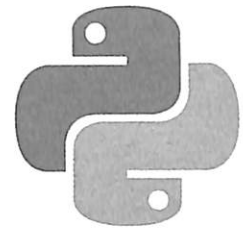
Machine Learning คืออะไร.....	335
ประเภทของ Machine Learning.....	337
Supervised Learning.....	337
Unsupervised Learning.....	340
Reinforcement learning.....	340
ไลบรารี Scikit-learn กับ Machine Learning Algorithm.....	341

Linear Regression.....	341
การทำงานของ Simple Linear Regression.....	341
ตัวอย่าง : ทำนายยอดขายด้วย Simple Linear Regression.....	347
การทำงานของ Multiple Linear Regression.....	348
ตัวอย่าง : ทำนายเงินเดือนด้วย Multiple Linear Regression.....	348
Logistic Regression.....	349
ทำความรู้จักกับ Training data และ Testing data.....	352
ตัวอย่าง : วิเคราะห์แนวโน้มการเป็นมะเร็งเต้านมด้วย Logistic Regression.....	353
การแสดง Confusion Matrix แบบกราฟิก.....	359
การแสดง Classification Report.....	360
Support Vector Machine (SVM).....	364
ตัวอย่าง : วิเคราะห์การอนุมัติเงินกู้ด้วย SVM.....	367
Naïve Bayes Classification.....	372
ตัวอย่าง : Sentimental Analysis วิเคราะห์ความรู้สึกจากข้อความด้วย Naïve Bayes.....	377
K-Nearest Neighbors (K-NN).....	381
ตัวอย่าง : Image Recognition วิเคราะห์และทำนายรูปภาพด้วย K-NN.....	384
การลดมิติข้อมูลด้วย PCA.....	391
K-Means Clustering.....	394
ตัวอย่าง : การทำ Customer Segmentation ด้วย K-Means Clustering.....	397



CHAPTER

01



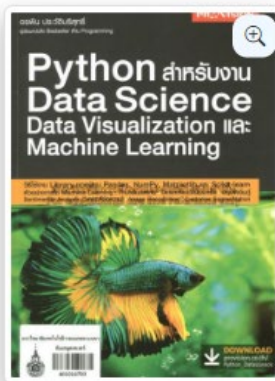
ทำความรู้จักกับ
Data Science

สามารถยืมและติดตามหนังสือใหม่ได้ที่ ระบบห้องสมุดอัตโนมัติ Walai Autolib

<https://lib.rmutp.ac.th/bibitem?bibid=b00107046>

B Python สำหรับงาน Data Science Data Visualization และ Machine Learning / อรพิน ปรวัฑบริสุทธ์.
อรพิน ปรวัฑบริสุทธ์.

My list 



Subject [ไพธอน \(ภาษาคอมพิวเตอร์\)](#)

Details

Published กรุงเทพฯ : ไพรวิชั่น, 2564.
Edition พิมพ์ครั้งที่ 1.
Detail 400 หน้า : ภาพประกอบ ; 23 ซม.
ISBN 9786162047886

 24  29  0

 MARC

 Export

 Save

 Share

สำหรับการศึกษาระดับปริญญาโทและปริญญาเอก