



สำนักพิมพ์มหาวิทยาลัยธรรมศาสตร์



การเขียนโปรแกรม
ภาษาไพธอนเบื้องต้น
สำหรับงานวิทยาศาสตร์ข้อมูล

ฉบับพิมพ์ครั้งที่ 2

ดร.จิรพล สังข์โพธิ์
ดร.ศราวุธ แรมจันทร์

การเขียนโปรแกรมภาษาไพธอนเบื้องต้น

สำหรับงานวิทยาศาสตร์ข้อมูล



ถ่ายเอกสารแทนการใช้หนังสือ
คือการทำลายภูมิปัญญาสร้างสรรค์

การเขียนโปรแกรมภาษาไพธอนเบื้องต้น สำหรับงานวิทยาศาสตร์ข้อมูล

ฉบับพิมพ์ครั้งที่ 2

รองศาสตราจารย์ ดร.จิรพล สังข์โพธิ์
รองศาสตราจารย์ ดร.ศราวุธ แรมจันทร์
วิทยาลัยนวัตกรรม มหาวิทยาลัยธรรมศาสตร์

สำนักพิมพ์มหาวิทยาลัยธรรมศาสตร์

2565

จิรพล สังข์โพธิ์, ศราวุธ แรมจันทร์.

การเขียนโปรแกรมภาษาไพธอนเบื้องต้นสำหรับงานวิทยาศาสตร์ข้อมูล.

1. ไพธอน (ภาษาคอมพิวเตอร์).

OA76.73

ISBN 978-616-314-901-5

ISBN (E-BOOK) 978-616-314-928-2

ลิขสิทธิ์ของรองศาสตราจารย์ ดร.จิรพล สังข์โพธิ์, รองศาสตราจารย์ ดร.ศราวุธ แรมจันทร์
สงวนลิขสิทธิ์

ฉบับพิมพ์ครั้งที่ 2 เดือนสิงหาคม 2565

จำนวน 50 เล่ม

ฉบับอิเล็กทรอนิกส์ (e-book) เดือนตุลาคม 2565

จัดพิมพ์และจัดจำหน่ายโดยสำนักพิมพ์มหาวิทยาลัยธรรมศาสตร์

ท่าพระจันทร์: อาคารธรรมศาสตร์ 60 ปี ชั้น U1 มหาวิทยาลัยธรรมศาสตร์

ถนนพระจันทร์ กรุงเทพฯ 10200 โทร. 0-2223-9232

ศูนย์รังสิต: อาคารโดมบริหาร ชั้น 3 ห้อง 317 มหาวิทยาลัยธรรมศาสตร์

ตำบลคลองหนึ่ง อำเภอคลองหลวง จังหวัดปทุมธานี 12120

โทร. 0-2564-2859-60 โทรสาร 0-2564-2860

<http://thammasatpress.tu.ac.th>, e-mail: unipress@tu.ac.th

พิมพ์ที่โรงพิมพ์มหาวิทยาลัยธรรมศาสตร์

ภาพปกโดย BiZkettE1/Freepik

พิมพ์ครั้งที่ 1 เดือนพฤศจิกายน 2562 จำนวน 50 เล่ม

พิมพ์ครั้งที่ 1 เดือนเมษายน 2563 จำนวน 100 เล่ม (ฉบับพิมพ์เพิ่มครั้งที่ 1)

พิมพ์ครั้งที่ 1 เดือนสิงหาคม 2563 จำนวน 100 เล่ม (ฉบับพิมพ์เพิ่มครั้งที่ 2)

พิมพ์ครั้งที่ 2 เดือนสิงหาคม 2565 จำนวน 50 เล่ม

ราคาเล่มละ 200.- บาท

สารบัญ

คำนำ	(15)
บทที่ 1 บทนำ (Introduction)	1
แนะนำภาษาไพธอน (Python)	2
การติดตั้งสื่อการเรียนภาษาไพธอน: PyCharm Edu	4
การเริ่มใช้งานสื่อศึกษาภาษาไพธอนด้วย PyCharm Edu	8
สรุป	11
แบบฝึกหัด	12
บทที่ 2 ตัวแปรภายในภาษาไพธอน (Variables)	13
การประกาศตัวแปร (Undefined Variable)	15
ชนิดตัวแปร (Variable Type)	15
การแปลงค่าภายในตัวแปร (Type Conversion)	16
ตัวดำเนินการทางคณิตศาสตร์ (Arithmetic Operators)	18
การกำหนดค่าภายในตัวแปร (Assignment)	19
การดำเนินการเชิงตรรกศาสตร์ (Boolean Operator)	21
การดำเนินการเชิงเปรียบเทียบ (Comparison Operator)	23
สรุป	25
แบบฝึกหัด	26
บทที่ 3 ข้อมูลอักขระ: (String)	27
การเชื่อมต่อข้อมูลตัวอักขระ (Concatenation)	28
การสร้างข้อมูลอักขระตามจำนวนที่กำหนด (String Multiplication)	29
การค้นหาลำดับอักขระ (String Indexing)	30
การค้นหาลำดับอักขระด้วยจำนวนเชิงลบ (String Negative Indexing)	31
การค้นหาลำดับอักขระด้วยช่วงตัวเลข (String Slicing)	32
การค้นหาข้อความด้วยอักขระ (In Operator)	34

การนับจำนวนอักขระ (String Length)	36
การจัดรูปแบบการแสดงผลอักขระด้วยเครื่องหมาย Backslash (Character Escaping)	37
เมธอดพื้นฐานเกี่ยวกับอักขระ (Basic String Method)	39
รูปแบบการแสดงผลข้อมูลอักขระ (String Formatting)	40
สรุป	41
แบบฝึกหัด	42
บทที่ 4 โครงสร้างข้อมูล (Data Structure)	43
ลำดับ (List)	43
ตัวดำเนินการเชิงลำดับ (Lists Operation)	45
การระบุลำดับ (List Items)	47
ทูเพิล (Tuple)	49
ดิกชันนารี (Dictionary)	50
ดิกชันนารี คีย์และแวลู (Dictionary Key and Value)	52
คีย์เวิร์ด (Keyword)	53
สรุป	55
แบบฝึกหัด	56
บทที่ 5 เงื่อนไข (Condition Expressions)	57
ตัวดำเนินการเชิงตรรกะ (Boolean Expression)	58
ลำดับของการดำเนินการ (Boolean Operator Order)	61
เงื่อนไข (If Statement)	63
เงื่อนไขแบบซับซ้อน (Else, elseif part in if statement)	65
สรุป	68
แบบฝึกหัด	69
บทที่ 6 การวนซ้ำ (Loops)	71
การวนซ้ำด้วยคำสั่ง For (For Loop)	72
การวนซ้ำด้วยข้อมูลที่เป็นข้อความภายใต้คำสั่ง For (For Loops Using String)	75
การวนซ้ำด้วยคำสั่ง While (While Loop)	77

การออกจากการทำงานด้วยคำสั่ง Break (Break Keyword)	80
การยกเว้นการประมวลผลภายในการทำงานด้วยคำสั่ง Continue (Continue Keyword)	82
สรุป	84
แบบฝึกหัด	85
บทที่ 7 ฟังก์ชัน (Functions)	87
การพัฒนาฟังก์ชัน	87
พารามิเตอร์และการเรียกอาร์กิวเมนต์ (Parameter and Call Arguments)	89
การส่งค่ากลับ (Return Value)	91
การกำหนดค่าให้กับพารามิเตอร์ (Default Parameter)	93
สรุป	96
แบบฝึกหัด	97
บทที่ 8 ไลบรารี Numpy	99
อาร์เรย์ (Array)	100
คำสั่ง .arange	102
คำสั่ง .reshape	103
คำสั่ง .linspace	104
ชนิดของข้อมูลในอาร์เรย์	105
ตัวดำเนินการภายในอาร์เรย์	106
Indexing and Slicing	110
Vectorization	112
การทำ Indexing และ Slicing ในอาร์เรย์ 2 มิติ	114
สรุป	116
แบบฝึกหัด	117
บทที่ 9 เหมือนข้อมูล (Data Mining)	119
การสร้าง Series และ Data Frame	119
การนำข้อมูลจากภายนอกเข้าสู่การประมวลผล	122
การเข้าถึงข้อมูลและการประมวลผลข้อมูล	124

Indexing, Selection และ Filtering	128
การดำเนินการกับข้อมูลภายใน Data frame	133
การวิเคราะห์ข้อมูลโดยใช้ชุดคำสั่งจากไลบรารี Scikit-Learn	145
สรุป	148
แบบฝึกหัด	149
บทที่ 10 Data Visualization	151
Bar	152
Histogram	153
Density Plot	155
การแสดงผล Histogram ร่วมกับ Density Plot	156
Box Plot	157
Scatter Plot	160
Word Cloud	161
สรุป	163
แบบฝึกหัด	164
บรรณานุกรม	165
ดัชนี	167

สารบัญภาพ

ภาพที่ 1.1	การแปลภาษาจากชุดคำสั่งไปเป็นภาษาที่อุปกรณ์ดิจิทัลสามารถเข้าใจได้โดย Interpreter	3
ภาพที่ 1.2	การเลือกดาวน์โหลดซอฟต์แวร์ PyCharm	4
ภาพที่ 1.3	ตัวติดตั้งซอฟต์แวร์ PyCharm	4
ภาพที่ 1.4	เริ่มเข้าสู่กระบวนการติดตั้งซอฟต์แวร์ PyCharm	5
ภาพที่ 1.5	การติดตั้งซอฟต์แวร์ PyCharm ลงสู่เครื่องคอมพิวเตอร์	5
ภาพที่ 1.6	ตัวเลือกในการติดตั้งซอฟต์แวร์ PyCharm	6
ภาพที่ 1.7	เลือกแสดงเมนูในการเข้าสู่ซอฟต์แวร์ PyCharm	6
ภาพที่ 1.8	เริ่มการติดตั้งซอฟต์แวร์	7
ภาพที่ 1.9	เสร็จสิ้นกระบวนการติดตั้ง	7
ภาพที่ 1.10	การเลือกบทบาทสำหรับการใช้เครื่องมือ PyCharm ในการศึกษาภาษาไพธอน	8
ภาพที่ 1.11	การเลือกสื่อการศึกษาผ่านซอฟต์แวร์ PyCharm	8
ภาพที่ 1.12	การเลือกศึกษาการเขียนโปรแกรมภาษาไพธอนผ่านซอฟต์แวร์ PyCharm	9
ภาพที่ 1.13	สื่อการทดลองการเขียนโปรแกรมไพธอน	9
ภาพที่ 1.14	การเขียนชุดคำสั่งครั้งแรก	10
ภาพที่ 1.15	การทดสอบชุดคำสั่ง	10
ภาพที่ 1.16	การแปลภาษาด้วย Interpreter เพื่อทำการทดสอบชุดคำสั่ง	11
ภาพที่ 2.1	การบรรจุข้อมูลลงบนตัวแปร	13
ภาพที่ 2.2	การดำเนินการด้วยตัวแปร	14
ภาพที่ 2.3	ตัวอย่างการแสดงผลผ่านตัวแปรด้วยคำสั่ง print	15
ภาพที่ 2.4	แสดงผลข้อมูลภายในตัวแปรชนิดตัวเลข	16
ภาพที่ 2.5	แสดงผลการแปรรูปตัวแปรชนิดเลขจำนวนจริงให้กลายเป็นตัวแปรชนิดเลขจำนวนเต็ม	17
ภาพที่ 2.6	ผลลัพธ์จากการประมวลผลผ่านตัวดำเนินการทางคณิตศาสตร์	19
ภาพที่ 2.7	แสดงผลค่าภายในตัวแปรหลังจากใช้ตัวดำเนินการกำหนดค่า	21
ภาพที่ 2.8	ผลลัพธ์จากการดำเนินการด้วยตัวดำเนินการเชิงตรรกศาสตร์	22
ภาพที่ 2.9	ผลลัพธ์ในการประมวลผลด้วยการดำเนินการเชิงเปรียบเทียบ	24

ภาพที่ 3.1	แสดงข้อมูลจากตัวแปรที่ถูกบรรจุด้วยข้อมูลอักขระที่ผ่านการเชื่อมต่อ	29
ภาพที่ 3.2	แสดงผลการสร้างข้อมูลอักขระซ้ำตามจำนวนที่กำหนด	30
ภาพที่ 3.3	แสดงผลข้อมูลจากการค้นหาอักขระ	31
ภาพที่ 3.4	แสดงผลการค้นหาข้อมูลอักขระด้วยจำนวนเชิงลบ	32
ภาพที่ 3.5	แสดงผลข้อมูลจากการค้นหาโดยการกำหนดช่วงตัวเลข	34
ภาพที่ 3.6	แสดงผลการค้นหาข้อความอักขระด้วยข้อความที่กำหนดไว้	35
ภาพที่ 3.7	การนับจำนวนข้อมูลอักขระด้วยคำสั่ง len	36
ภาพที่ 3.8	การจัดรูปแบบการแสดงผลอักขระโดยใช้เครื่องหมาย Backslash	38
ภาพที่ 3.9	การแปรรูปอักขระด้วยเมธอด .lower() และ .upper()	40
ภาพที่ 3.10	แสดงผลข้อมูลร่วมกับข้อความผ่านตัวดำเนินการ	41
ภาพที่ 4.1	แสดงผลจากโครงสร้างข้อมูลแบบลำดับ	44
ภาพที่ 4.2	แสดงผลการดำเนินการกับข้อมูลเชิงโครงสร้างแบบลำดับ	46
ภาพที่ 4.3	การลบข้อมูลออกจากชุดข้อมูลเชิงโครงสร้างแบบลำดับ	48
ภาพที่ 4.4	แสดงจำนวนของข้อมูลที่อยู่ภายในทูเพิล	49
ภาพที่ 4.5	การเขียนชุดคำสั่งโดยใช้ดิกชันนารี	51
ภาพที่ 4.6	แสดงผล Keys และ Values ในดิกชันนารี	53
ภาพที่ 4.7	แสดงผลการตรวจสอบคีย์เวิร์ดภายในชุดข้อมูลเชิงโครงสร้างแบบลำดับ และดิกชันนารี	55
ภาพที่ 5.1	ตัวดำเนินการ And	58
ภาพที่ 5.2	ตัวดำเนินการ Or	59
ภาพที่ 5.3	ผลลัพธ์จากการดำเนินการผ่านตัวดำเนินการ And และ Or	60
ภาพที่ 5.4	ลำดับความสำคัญของตัวดำเนินการ	62
ภาพที่ 5.5	การทำงานของคำสั่ง If	63
ภาพที่ 5.6	แสดงผลตามเงื่อนไขโดยใช้คำสั่ง If	65
ภาพที่ 5.7	กระบวนการทำงานของเงื่อนไขแบบซับซ้อน	66
ภาพที่ 5.8	การดำเนินการตามคำสั่ง If แบบซับซ้อน	68
ภาพที่ 6.1	การวนซ้ำโดยใช้คำสั่ง For	72
ภาพที่ 6.2	แสดงผลข้อมูลโดยใช้คำสั่ง For	74

ภาพที่ 6.3	แสดงผลการวนซ้ำด้วยข้อมูลที่เป็นข้อความภายใต้คำสั่ง For	77
ภาพที่ 6.4	แสดงผลการวนซ้ำโดยใช้คำสั่ง While	79
ภาพที่ 6.5	การออกจากการวนซ้ำโดยใช้คำสั่ง Break	82
ภาพที่ 6.6	แสดงผลลัพท์จากการการยกเว้นการประมวลผลโดยใช้คำสั่ง Continue	84
ภาพที่ 7.1	แสดงผลการพัฒนาและเรียกใช้ฟังก์ชัน	89
ภาพที่ 7.2	การกำหนดพารามิเตอร์และอาร์กิวเมนต์	91
ภาพที่ 7.3	แสดงผลการส่งค่ากลับไปยังอาร์กิวเมนต์ที่เรียกใช้งานฟังก์ชันผ่านคำสั่ง Return	93
ภาพที่ 7.4	แสดงผลจากฟังก์ชันที่มีการกำหนดค่าให้กับพารามิเตอร์	95
ภาพที่ 8.1	ตัวอย่างข้อมูลแบบอาร์เรย์	100
ภาพที่ 8.2	แสดงผลข้อมูลอาร์เรย์ 1 มิติ	101
ภาพที่ 8.3	แสดงผลข้อมูลอาร์เรย์ 2 มิติ	102
ภาพที่ 8.4	แสดงผลข้อมูลโดยใช้คำสั่ง .arange ในการกำหนดข้อมูล	103
ภาพที่ 8.5	แสดงผลการแปรรูปข้อมูลให้อยู่ในรูปแบบอาร์เรย์	104
ภาพที่ 8.6	แสดงผลข้อมูลโดยใช้คำสั่ง .linspace ในการสร้างข้อมูล	105
ภาพที่ 8.7	การกำหนดชนิดข้อมูลให้กับอาร์เรย์	106
ภาพที่ 8.8	ผลลัพธ์จากการดำเนินการข้อมูลอาร์เรย์	107
ภาพที่ 8.9	แสดงผลข้อมูลจากการสุ่มข้อมูลภายในตัวแปรอาร์เรย์	108
ภาพที่ 8.10	แสดงผลรวมจากข้อมูลตัวเลขภายในตัวแปรอาร์เรย์	109
ภาพที่ 8.11	แสดงผลข้อมูลที่น้อยที่สุดในตัวแปรอาร์เรย์	109
ภาพที่ 8.12	แสดงผลข้อมูลที่มากที่สุดในตัวแปรอาร์เรย์	110
ภาพที่ 8.13	แสดงข้อมูลโดยระบุตำแหน่งที่อยู่ของข้อมูลในตัวแปรอาร์เรย์	111
ภาพที่ 8.14	Slicing ข้อมูลอาร์เรย์	112
ภาพที่ 8.15	การใช้งานคำสั่ง .vectorize เพื่อทำการประมวลผลข้อมูลอาร์เรย์	113
ภาพที่ 8.16	การทำ Indexing และ Slicing จากตัวแปรอาร์เรย์ขนาดใหญ่	116
ภาพที่ 9.1	การสร้าง Series	120
ภาพที่ 9.2	การสร้าง Data frame	121
ภาพที่ 9.3	การนำเข้าและแสดงผลข้อมูลไฟล์ csv	123
ภาพที่ 9.4	การนำเข้าไฟล์ csv จากแหล่งข้อมูลทางอินเทอร์เน็ต	124

ภาพที่ 9.5	การเข้าถึงข้อมูลภายใน Data frame	126
ภาพที่ 9.6	การประมวลผลข้อมูลภายใน Data Frame	128
ภาพที่ 9.7	การทำ Indexing ร่วมกับ Data frame	130
ภาพที่ 9.8	การทำ Selection ร่วมกับ Data frame	131
ภาพที่ 9.9	การทำ Filtering ร่วมกับ Data frame	133
ภาพที่ 9.10	การประมวลผลแบบ Broadcasting ร่วมกับ Data frame	134
ภาพที่ 9.11	Sorting ข้อมูลภายใน Data frame	136
ภาพที่ 9.12	การจัดการกับ Missing Value	137
ภาพที่ 9.13	การแทนที่ข้อมูลใหม่ลงไปใน Missing Value	139
ภาพที่ 9.14	การประมวลผลข้อมูลภายในซีรีย์ด้วยสถิติพื้นฐาน	141
ภาพที่ 9.15	การทำ Label Encoding	142
ภาพที่ 9.16	การทำ One-Hot Encoding	144
ภาพที่ 9.17	Machine Learning แบบ Supervised Learning	145
ภาพที่ 9.18	Decision Tree	148
ภาพที่ 10.1	แสดงผล Bar	153
ภาพที่ 10.2	Histogram	154
ภาพที่ 10.3	Density Plot	155
ภาพที่ 10.4	Histogram ร่วมกับ Density Plot	157
ภาพที่ 10.5	ตัวอย่าง Box Plot	158
ภาพที่ 10.6	แสดงผล Box Plot	159
ภาพที่ 10.7	แสดงผล Scatter Plot	161
ภาพที่ 10.8	แสดงผลแบบ WordCloud	162

สารบัญตาราง

ตารางที่ 2.1	ตัวอย่างการใช้ตัวดำเนินการทางคณิตศาสตร์ตามไวยากรณ์ของภาษาไพธอน	18
ตารางที่ 2.2	ไวยากรณ์และตัวอย่างการกำหนดค่าภายในตัวแปรของภาษาไพธอน	19
ตารางที่ 2.3	ตัวดำเนินการเชิงตรรกศาสตร์ภายในภาษาไพธอน	22
ตารางที่ 2.4	ตัวดำเนินการเชิงเปรียบเทียบในภาษาไพธอน	23

คำนำ

หนังสือเล่มนี้เขียนขึ้นมาเพื่อเป็นแนวทางสำหรับผู้สนใจฝึกการเขียนโปรแกรมโดยใช้ภาษาไพธอนเป็นสื่อและใช้สำหรับผู้สนใจการประยุกต์ใช้ภาษาไพธอนสำหรับงานวิทยาศาสตร์ข้อมูล โดยหนังสือเล่มนี้กล่าวถึง การฝึกหัดเขียนชุดคำสั่งโดยใช้ซอฟต์แวร์ PyCharm ตัวแปรภายในภาษาไพธอน ข้อมูลอักขระ โครงสร้างข้อมูล เงื่อนไข การวนซ้ำ ฟังก์ชัน ไลบรารี Numpy เหมืองข้อมูล (Data Mining) และ Data Visualization

จากเนื้อหาดังกล่าว หนังสือเล่มนี้ได้ปูพื้นฐานเริ่มต้นตั้งแต่ผู้เริ่มฝึกการเขียนโปรแกรมไปจนถึงการวิเคราะห์ข้อมูลด้วยเทคนิคทางวิทยาศาสตร์ข้อมูลซึ่งนำภาษาไพธอนไปประยุกต์ใช้ในการเรียบเรียงขั้นตอนวิธี ผู้เขียนหวังเป็นอย่างยิ่งว่าหนังสือเล่มนี้จะช่วยให้ผู้ที่สนใจในการเริ่มต้นฝึกเขียนโปรแกรมด้วยภาษาไพธอนและผู้สนใจการประยุกต์ใช้ภาษาไพธอนในงานวิทยาศาสตร์ข้อมูลเกิดทักษะเบื้องต้นในการเขียนโปรแกรมเพื่องานวิทยาศาสตร์ข้อมูลต่อไป

ผู้ช่วยศาสตราจารย์ ดร.จิรพล สังข์โพธิ์

ดร.ศราวุธ แรมจันทร์



1

บทนำ (Introduction)

วิทยาศาสตร์ข้อมูล (Data Science) ได้ทวีความสำคัญขึ้นอย่างมากในหลากหลายอุตสาหกรรมที่ต้องการการวิเคราะห์ข้อมูลขั้นสูงในการตัดสินใจทางธุรกิจ (Durrant-Whyte, 2015) ข้อมูลที่ไม่มีโครงสร้าง (Unstructured Data) (Korfatis, 2019) ได้กลายเป็นเป้าหมายสำคัญในการนำมาหลอมรวมเข้ากับข้อมูลขนาดใหญ่ (Big Data) (Grover V, 2018) ในการแปรรูปให้ข้อมูลที่ไม่มีโครงสร้างให้กลับมาอยู่ในรูปแบบที่องค์กรสามารถนำมาหลอมรวมเข้ากับข้อมูลที่มีโครงสร้าง (Structured Data) (Bacciu D, 2018) และข้อมูลกึ่งโครงสร้าง (Semi-Structured Data) (Scriney M, 2019) ที่องค์กรมีอยู่เพื่อเตรียมข้อมูลให้พร้อมต่อการนำไปวิเคราะห์ (Data Preparation) (Benhar H, 2019) สามารถทำได้ผ่านเครื่องมือทางดิจิทัลที่มีความหลากหลาย และหนึ่งในเครื่องมืออันทรงพลังที่สามารถแปรรูปข้อมูลที่ไม่มีโครงสร้างซึ่งกระจัดกระจายอยู่ในเครือข่ายสังคม (Social Network) และแฝงตัวอยู่ในธุรกรรมทางดิจิทัลระหว่างคู่ค้าของแต่ละธุรกิจให้กลายมาเป็นข้อมูลที่พร้อมใช้งานคือ ไพธอน (Python) (Holman JO, 2018)

หนังสือเล่มนี้จึงได้เรียบเรียงการเขียนโปรแกรมด้วยภาษาไพธอนโดยใช้เครื่องมือในการทดลองคือซอฟต์แวร์ PyCharm Edu (JetBrains, 2018) เพื่อให้ผู้อ่านสามารถนำเอาเนื้อหาที่ได้รับจากการอ่านหนังสือเล่มนี้ไปทำงานทดลองผ่านซอฟต์แวร์ดังกล่าว จนกระทั่งได้รับองค์ความรู้พื้นฐานที่เหมาะสมต่อการศึกษาเกี่ยวกับการสกัดและจัดเตรียมการวิเคราะห์ข้อมูลที่ไม่มีโครงสร้างให้อยู่ในรูปแบบที่เหมาะสมต่อการวิเคราะห์ด้วยกระบวนการวิทยาศาสตร์ข้อมูลต่อไป

แนะนำภาษาไพธอน (Python)

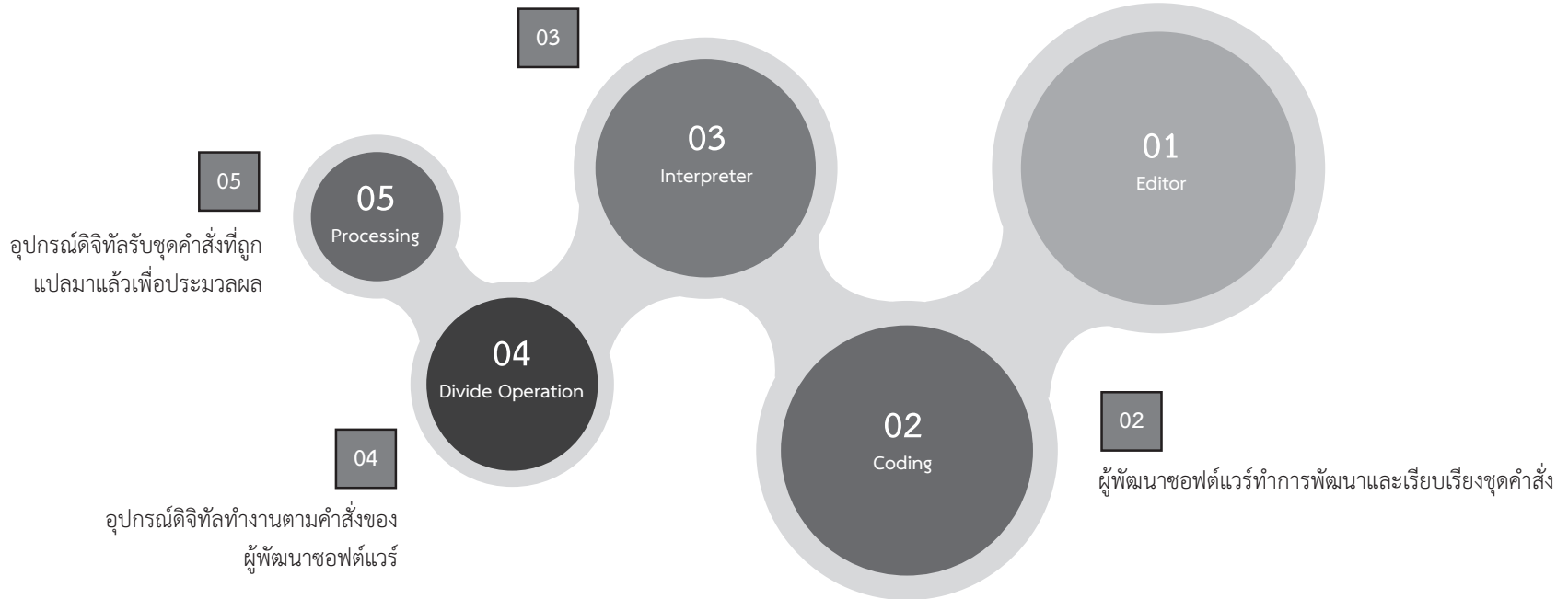
ภาษาไพธอน (Python) ได้มีการพัฒนาขึ้นโดย Buido van Rossum ในปี 1980 เพื่อใช้ในการพัฒนาซอฟต์แวร์ที่ตอบสนองต่อหลากหลายวัตถุประสงค์ (The Python Software Foundation, 2018) โดยเป็นภาษาระดับสูงซึ่งหมายถึงไวยากรณ์และภาษาที่ใช้ในการเขียนโปรแกรมมีความใกล้เคียงกับภาษาปกติที่มนุษย์ใช้สื่อสารกัน อีกทั้งยังถูกพัฒนาให้มีไวยากรณ์ที่ผ่อนคลายและเป็นมิตรกับผู้พัฒนาซอฟต์แวร์มากกว่าภาษาอื่นๆ ในอดีต อาทิ ละเว้นการเขียนเครื่องหมายก้ามปู [] และเครื่องหมายอัฒภาค (;) หรือ Semicolon ซึ่งอยู่ในภาษาสำหรับการเขียนโปรแกรมอื่นๆ และมีความเหมาะสมต่อการพัฒนาซอฟต์แวร์ทั้งในลักษณะการเขียนโปรแกรมเชิงโครงสร้าง (Structured Programming) และการเขียนโปรแกรมเชิงวัตถุ (Object-Oriented Programming) (Kanaki, 2018) อีกทั้งไพธอนมีลักษณะเป็น Open-Source (Upasani, 2016) ซึ่งนักพัฒนาซอฟต์แวร์สามารถนำไปใช้ได้โดยปราศจากลิขสิทธิ์อันเป็นสาเหตุให้นักพัฒนาซอฟต์แวร์จากทั่วโลกล้วนมีส่วนช่วยในการพัฒนาให้ไพธอนเป็นภาษาที่ทรงพลังในการประมวลผลข้อมูลภายใต้การโปรแกรมที่รองรับต่อการพัฒนาซอฟต์แวร์ต่างๆ โดยเวอร์ชันปัจจุบันคือ Python 3

ไพธอนเองมีความสามารถในการจัดการหน่วยความจำอย่างยืดหยุ่นและลดความซับซ้อนของการพัฒนาซอฟต์แวร์ในอดีตที่จำเป็นต้องพัฒนาโครงสร้างข้อมูลที่ซับซ้อนเพื่อเพิ่มประสิทธิภาพในการประมวลผลข้อมูลผ่านหน่วยความจำและหน่วยประมวลผลกลาง โดยภาษาไพธอนสามารถถูกแปรให้กลายเป็นคำสั่งที่ใช้ควบคุมให้อุปกรณ์ดิจิทัลทำงานผ่านตัวแปลภาษาในลักษณะ Interpreter ดังภาพที่ 1.1

จากภาพที่ 1.1 ผู้พัฒนาซอฟต์แวร์จะใช้ Software Editor ในการเขียนชุดคำสั่งด้วยภาษาไพธอน โดย Software Editor ที่ได้รับความนิยมจะรองรับต่อการเขียนโปรแกรม การทดสอบโปรแกรม การบันทึกโปรแกรม ตลอดจนการนำเข้า Library สำเร็จรูปต่างๆ ที่ผู้พัฒนาซอฟต์แวร์สามารถเรียกใช้คำสั่งสำหรับการเขียนโปรแกรมที่มีวัตถุประสงค์เฉพาะเจาะจงได้ จากนั้นผู้พัฒนาซอฟต์แวร์จะทำการพัฒนาและเรียบเรียงชุดคำสั่งให้เป็นไปตามวัตถุประสงค์ในการใช้งานซอฟต์แวร์ และเมื่อทำการทดสอบซอฟต์แวร์ ไพธอนต้องอาศัย Interpreter ซึ่งเป็นตัวแปลภาษาไพธอนที่มีความใกล้เคียงภาษามนุษย์ให้กลายเป็นภาษาที่เครื่องคอมพิวเตอร์สามารถเข้าใจได้ (Machine Language) (Hassitt A, 1973) โดยอุปกรณ์ดิจิทัลจะประมวลผลคำสั่งและข้อมูลเพื่อให้อุปกรณ์ดิจิทัลสามารถปฏิบัติตามคำสั่งที่ผู้พัฒนาซอฟต์แวร์ได้เรียบเรียงไว้ได้ในที่สุด

ตัวแปลภาษาแปลชุดคำสั่งให้อยู่ในรูปแบบที่อุปกรณ์ดิจิทัล
สามารถเข้าใจได้

01 ผู้พัฒนาซอฟต์แวร์ใช้ Software Editor ในการเขียนชุดคำสั่ง

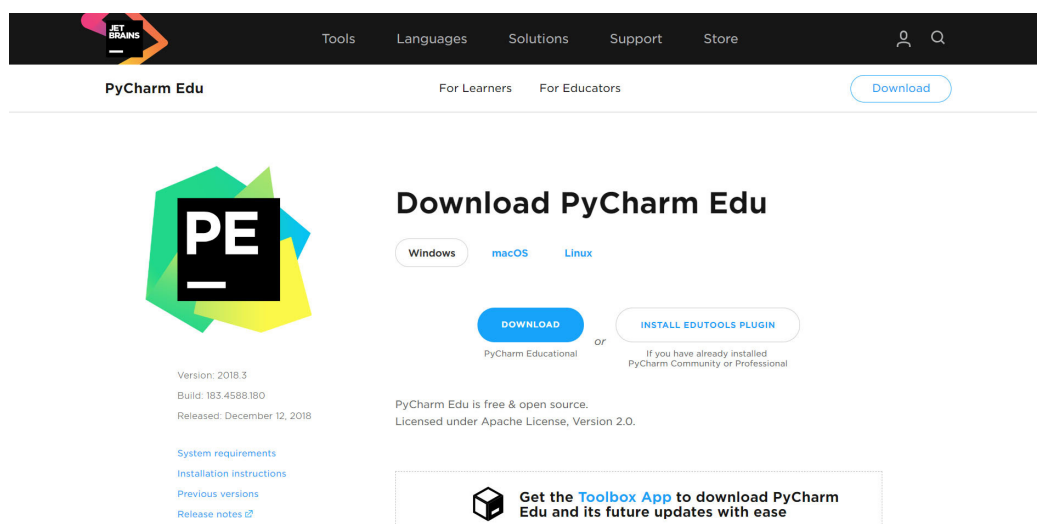


ภาพที่ 1.1 การแปลภาษาจากชุดคำสั่งไปเป็นภาษาที่อุปกรณ์ดิจิทัลสามารถเข้าใจได้โดย Interpreter

การติดตั้งสื่อการเรียนรู้ภาษาไพธอน: PyCharm Edu

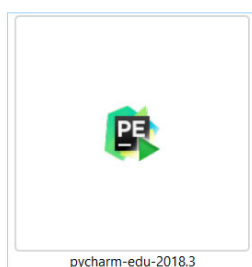
หนังสือเล่มนี้ได้ใช้วัตถุดิบประกอบการศึกษาตำราโดยใช้ซอฟต์แวร์ที่เป็นสื่อการสอนคือ PyCharm Edu โดยซอฟต์แวร์ดังกล่าวจะถูกนำมาใช้สำหรับผู้อ่านที่ต้องการทำการศึกษาและทดลองการเขียนคำสั่งโดยใช้ภาษาไพธอนจากหนังสือเล่มนี้

ผู้อ่านสามารถดาวน์โหลดซอฟต์แวร์ผ่านเว็บเบราว์เซอร์ใดๆ อาทิ Google Ghorm โดยเข้าสู่ URL: ดังนี้ <https://www.jetbrains.com/pycharm-edu/download/#section=windows>



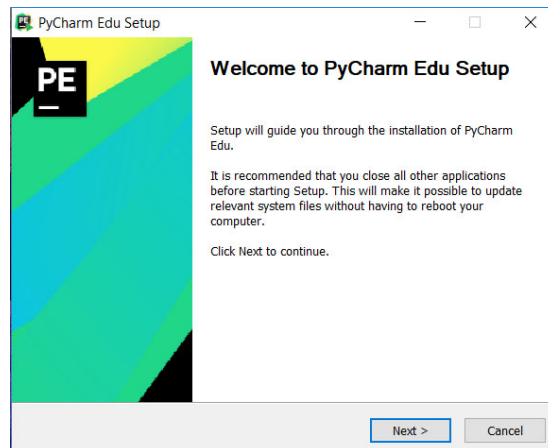
ภาพที่ 1.2 การเลือกดาวน์โหลดซอฟต์แวร์ PyCharm

จากภาพที่ 1.2 ผู้อ่านสามารถเลือกดาวน์โหลดซอฟต์แวร์ PyCharm ที่เหมาะสมสำหรับการติดตั้งบนระบบปฏิบัติการ Windows ระบบปฏิบัติการ macOS และระบบปฏิบัติการ Linux จากนั้นให้เลือกการดาวน์โหลดที่ PyCharm Edu โดยหนังสือเล่มนี้จะสาธิตการติดตั้งซอฟต์แวร์ PyCharm สำหรับระบบปฏิบัติการ Windows



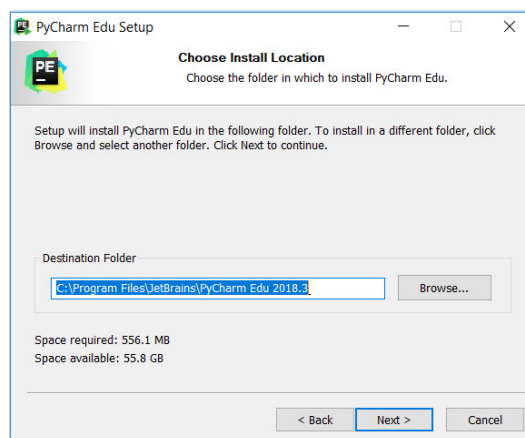
ภาพที่ 1.3 ตัวติดตั้งซอฟต์แวร์ PyCharm

หลังจากการดาวน์โหลดตัวติดตั้ง ให้ผู้อ่าน Double Click ที่ไฟล์ตามภาพที่ 1.3 เพื่อติดตั้งซอฟต์แวร์ PyCharm ลงบนเครื่องคอมพิวเตอร์บนระบบปฏิบัติการใดๆ



ภาพที่ 1.4 เริ่มเข้าสู่กระบวนการติดตั้งซอฟต์แวร์ PyCharm

จากภาพที่ 1.4 ตัวติดตั้งซอฟต์แวร์ จะเริ่มติดตั้งซอฟต์แวร์ลงบนเครื่องคอมพิวเตอร์ของผู้อ่าน ให้ผู้อ่านกดปุ่ม Next > เพื่อเข้าสู่กระบวนการติดตั้ง หรือกดปุ่ม Cancel เพื่อยกเลิกกระบวนการติดตั้ง



ภาพที่ 1.5 การติดตั้งซอฟต์แวร์ PyCharm ลงสู่เครื่องคอมพิวเตอร์

ภาพที่ 1.5 ตัวติดตั้งจะแนะนำไดรฟ์และโฟลเดอร์ที่เหมาะสมต่อการติดตั้งซอฟต์แวร์ PyCharm โดยผู้อ่านสามารถเลือกติดตั้งลงในแหล่งอื่นๆ โดยคลิกเลือกที่ปุ่ม Browse... และหากผู้อ่านต้องการติดตั้งตามคำแนะนำของตัวติดตั้งให้กดปุ่ม Next >