



Python

เรียนรู้ Concept และพัฒนา Coding

Data Science

- เรียนรู้หลักการและวิธีการ Coding เกี่ยวกับ Data Science เบื้องต้นอย่างครบถ้วน
- แนะนำการ Coding กับ NumPy, Pandas, MongoDB, Matplotlib และ PandasGUI
- หากมีพื้นฐาน Python เบื้องต้นมาแล้ว สามารถศึกษาด้วยตนเอง Step-by-Step ได้อย่างรวดเร็ว
- เหมาะสำหรับนักศึกษาและผู้สนใจจะเพิ่ม Skill การ Coding เกี่ยวกับ Data Science



www.boonamee.com
<https://www.boonamee.com/>
9786164872104

ศุภชัย สมพานิช

บทที่ 1	เตรียมความพร้อมก่อนเข้าสู่โลกของข้อมูล	1
	ภาษาไพธอนกับงานด้านข้อมูล	2
	การดาวน์โหลดและติดตั้ง Python SDK.....	2
	การเขียนโค้ดภาษาไพธอนด้วย PyCharm	5
	การดาวน์โหลดและติดตั้งโปรแกรม PyCharm	5
	เริ่มต้นเขียนโค้ดภาษาไพธอนในโปรแกรม Pycharm.....	7
	การรันภาษาไพธอนในโปรแกรม Pycharm.....	9
	การเขียนโค้ดภาษาไพธอนด้วย Visual Studio Code	11
	การดาวน์โหลดและติดตั้งโปรแกรม Visual Studio Code และส่วนขยายสำหรับ ภาษาไพธอน	11
	เริ่มต้นเขียนโค้ดภาษาไพธอนใน Visual Studio Code.....	13
	การรันโค้ดภาษาไพธอนในโปรแกรม Visual Studio Code.....	14
	การ Config โปรแกรม Visual Studio Code ให้สามารถใช้ Package ของไพธอน เพิ่มเติม.....	16
	การเขียนโค้ดภาษาไพธอนในรูปแบบ Jupyter Notebook.....	18
บทที่ 2	พื้นฐานการทำงานกับข้อมูลด้วย NumPy	21
	การดาวน์โหลดและติดตั้ง NumPy.....	21
	การใช้งาน NumPy กับโครงสร้างข้อมูลแบบอาร์เรย์ (Array)	22
	การตรวจสอบชนิดข้อมูลพื้นฐานในอาร์เรย์ของ NumPy.....	26
	การหาค่าเฉลี่ยเลขคณิต, ค่าสูงสุด และค่าต่ำสุดของ NumPy.....	29
	พื้นฐานการสุ่มตัวเลขทศนิยม.....	30
	พื้นฐานการสุ่มตัวเลขจำนวนเต็มด้วยฟังก์ชัน randint()	31
	พื้นฐานการสุ่มตัวเลขจำนวนเต็มด้วยฟังก์ชัน randrange().....	31
	การสุ่มตัวเลขจำนวนเต็มหลายจำนวน.....	33
	พื้นฐานการใช้งานอาร์เรย์ 2 มิติ ของ NumPy.....	34
	การหาผลรวมในอาร์เรย์ด้วยฟังก์ชัน sum()	35

การสร้างอาร์เรย์ด้วยฟังก์ชัน zeros() แบบมีข้อมูลเลข 0 เท่านั้น	38
การสร้างอาร์เรย์ด้วยฟังก์ชัน ones() แบบมีข้อมูลเลข 1 เท่านั้น	41
การเลือกข้อมูลแบบเป็นช่วงจากอาร์เรย์แบบ 1 มิติ ของ NumPy	44
การเลือกข้อมูลแบบเป็นช่วงจากอาร์เรย์แบบ 2 มิติ ของ NumPy	46
การเลือกข้อมูลจากอาร์เรย์แบบมีเงื่อนไขด้วยฟังก์ชัน where()	49
การรวมอาร์เรย์ด้วยฟังก์ชัน append()	51

บทที่ 3 พื้นฐานการทำงานกับข้อมูลด้วย Pandas 55

การดาวน์โหลดและติดตั้ง Pandas	55
การแสดงผลข้อมูลจากอาร์เรย์ 2 มิติ แบบตารางด้วยฟังก์ชัน DataFrame()	56
การตัดคอลัมน์หรือแถวออกจาก DataFrame ด้วยฟังก์ชัน drop()	57
การตัดค่าซ้ำออกจาก DataFrame ด้วยฟังก์ชัน drop_duplicates()	59
การเปลี่ยนชื่อคอลัมน์ใน DataFrame ด้วยฟังก์ชัน rename()	61
การสร้างฟังก์ชันทำงานกับข้อมูลโดยอาศัยฟังก์ชัน apply()	62
การสร้างแลมด้า (Lambda) ทำงานกับข้อมูลโดยอาศัยฟังก์ชัน apply()	64
การทำงานกับข้อมูลตามแกน X หรือแกน Y	65
การทำงานกับข้อมูลตามแนวคอลัมน์ด้วยฟังก์ชัน applymap()	66
โครงสร้างข้อมูลแบบ Series ของ Pandas	67
Series แบบกำหนดค่า Index เอง	70
การสร้าง Series จากโครงสร้างข้อมูล Dictionary (dict)	71
การอ่านข้อมูลจาก Series	72
การอ่านข้อมูลจาก Series แบบเป็นช่วงข้อมูล	75

บทที่ 4 DataFrame ของ Pandas 79

พื้นฐานการสร้าง DataFrame ของ Pandas	79
การสร้าง DataFrame แบบหลายคอลัมน์	81
การสร้าง DataFrame จากรายการ list แบบหลายคอลัมน์	83
การตรวจสอบข้อมูลเบื้องต้นของ DataFrame	86
การตรวจสอบชนิดข้อมูลของ DataFrame ด้วยคุณสมบัติ dtypes	89
การสร้าง DataFrame จาก Dictionary (dict)	92
การรวม DataFrame โดยอาศัยฟังก์ชัน append()	95
การรวม DataFrame ด้วยฟังก์ชัน concat()	98

	พื้นฐานการอ่านข้อมูลจาก DataFrame ด้วยคุณสมบัติ <code>loc</code>	100
	พื้นฐานการจัดกลุ่มข้อมูลใน DataFrame ด้วยฟังก์ชัน <code>groupby()</code>	105
บทที่ 5	การทำงานกับไฟล์ Excel	109
	การอ่านข้อมูลจากไฟล์ Excel ด้วย <code>xrd</code>	109
	การแสดงรายชื่อคอลัมน์ที่ได้จากไฟล์ Excel	113
	การเลือกข้อมูลเป็นช่วงจากไฟล์ Excel	114
	การค้นหาข้อมูลด้วยฟังก์ชัน <code>match()</code> หรือฟังก์ชัน <code>contains()</code> ของ Pandas	118
	การสร้างไฟล์ Excel จาก DataFrame ด้วย <code>xlsxwriter</code>	121
	การ Export ข้อมูลจาก DataFrame ไปสู่โครงสร้างข้อมูลแบบอื่นๆ	123
	การสร้างไฟล์ Excel และเพิ่มข้อมูลด้วย <code>openpyxl</code>	129
	การใส่ข้อมูลแบบมีเงื่อนไข	132
บทที่ 6	ทำงานกับข้อมูลในไฟล์ CSV	135
	พื้นฐานการอ่านข้อมูลจากไฟล์ CSV ด้วยฟังก์ชัน <code>read_csv()</code> ของ Pandas	135
	การระบุเครื่องหมายอื่นๆ ในไฟล์ CSV	137
	การเลือกบางคอลัมน์หรือบางแถวในไฟล์ CSV	138
	การเปลี่ยนชนิดข้อมูลคอลัมน์	143
	ชนิดข้อมูลพื้นฐานของไพธอนกับ Pandas	146
	การเปลี่ยนเป็นชนิดข้อมูลจัดกลุ่ม <code>category</code> ของ Pandas	146
	การเลือกข้อมูลจากไฟล์ CSV ด้วยคุณสมบัติ <code>iloc</code>	148
บทที่ 7	ทำงานกับข้อมูลในฐานข้อมูล SQL Server	153
	การดาวน์โหลดและติดตั้งฐานข้อมูล SQL Server	154
	การสร้างตารางเก็บรายการสินค้า	160
	การติดตั้ง ODBC เชื่อมต่อไพธอนเข้ากับฐานข้อมูล SQL Server	165
	การกำหนดให้ฐานข้อมูล SQL Server เข้ารหัสภาษาไทย	168
	การจัดการข้อมูล (CRUD) ในฐานข้อมูล SQL Server ด้วยไพธอน	169
	การอ่านข้อมูลจากฐานข้อมูล SQL Server	169
	การเพิ่มข้อมูลใหม่ลงในฐานข้อมูล SQL Server	172
	การแก้ไขข้อมูลในฐานข้อมูล SQL Server	174
	การลบข้อมูลในฐานข้อมูล SQL Server (แบบตารางเดียว)	176



บทที่ 8	พื้นฐานการใช้งานระบบฐานข้อมูล MongoDB	179
	ระบบฐานข้อมูล NoSQL แบบ MongoDB คืออะไร	179
	การดาวน์โหลดและติดตั้งฐานข้อมูล MongoDB	180
	พื้นฐานการเชื่อมต่อกับฐานข้อมูล MongoDB ด้วย MongoDB Compass	182
	การสร้างฐานข้อมูล MongoDB ใหม่	183
	การดาวน์โหลดและติดตั้ง pymongo	189
	การจัดการข้อมูล (CRUD) ในฐานข้อมูล MongoDB	189
	การเพิ่มข้อมูลใหม่เข้าไปในฐานข้อมูล MongoDB	190
	การอัปเดตข้อมูลในฐานข้อมูล MongoDB	193
	การลบข้อมูลในฐานข้อมูล MongoDB	195
บทที่ 9	การทำงานกับข้อมูลสูญหายและค่าผิดปกติ	197
	ปัญหาของข้อมูลสูญหาย	197
	การถอดข้อมูลสูญหายออกด้วยฟังก์ชัน dropna()	198
	การถอดค่าข้อมูลสูญหายแบบใช้ตัวแปรเติม	199
	การยืนยันยอมให้มีค่าสูญหาย	200
	การถอดค่าแบบกำหนดแทน X หรือแทน Y	201
	การคำนวณปริมาณข้อมูลสูญหาย	202
	การคำนวณปริมาณข้อมูลสูญหาย (หน่วยเปอร์เซ็นต์)	203
	การตรวจสอบข้อมูลสูญหายด้วยฟังก์ชัน isnull()	204
	พื้นฐานการเติมค่าที่สูญหายด้วยฟังก์ชัน interpolate()	205
	การเติมค่าสูญหายด้วยฟังก์ชัน fillna()	206
	ทำความเข้าใจกับข้อมูลที่มีค่าผิดปกติ (Outliers)	210
	การคำนวณค่าเฉลี่ยเลขคณิตแบบตัดค่าสูญหายด้วยฟังก์ชัน nanmean()	212
บทที่ 10	การใช้งาน Pandas แบบมีส่วนแสดงผล (PandasGUI)	215
	การดาวน์โหลดและติดตั้ง PandasGUI	216
	พื้นฐานการแสดงผลข้อมูลจาก DataFrame ด้วย PandasGUI	216
	การกรองข้อมูล (เช่น Filters)	218
	การ Import ข้อมูลจากไฟล์ CSV และ Excel	220
	การ Export ข้อมูลออกเป็นไฟล์ CSV	220

บทที่ 11	ค่ากลางของข้อมูล (Measures of Central Tendency)	223
	การหาค่าเฉลี่ยเลขคณิต (Arithmetic Mean) ด้วยฟังก์ชัน mean() ของ NumPy	223
	การหาค่าเฉลี่ยเลขคณิตใน DataFrame ด้วยฟังก์ชัน mean() ของ Pandas.....	225
	การหาค่าเฉลี่ยเลขคณิตแบบจัดกลุ่มข้อมูลใน DataFrame ของ Pandas.....	227
	การหาค่าเฉลี่ยเลขคณิตกับตัวเลขทศนิยมด้วยฟังก์ชัน fmean() ของ statistics.....	230
	ค่าเฉลี่ยเรขาคณิต (Geometric Mean).....	231
	การหาค่าเฉลี่ยถ่วงน้ำหนัก (Weight Mean) ด้วยฟังก์ชัน average() ของ NumPy.....	234
	การหาค่าเฉลี่ยถ่วงน้ำหนักใน DataFrame ของ Pandas.....	236
	การหาค่าเฉลี่ยฮาร์โมนิก (Harmonic Mean).....	238
	การหาค่ามัธยฐาน (Median).....	239
	ฐานนิยม (Mode).....	242
	การคำนวณค่ามัธยฐานแบบตัดค่าสูญหายด้วยฟังก์ชัน nanmedian().....	243
	ฐานนิยมแบบหลายค่าด้วยฟังก์ชัน multimode() ของ statistics.....	244
บทที่ 12	การหาค่าแห่งของข้อมูล (Measures of Position)	247
	การหาค่าแห่งแบบควอร์ไทล์ (Quartile).....	247
	การหาค่าแห่งแบบเดซิส์ (Decile).....	251
	การหาค่าแห่งแบบเปอร์เซ็นต์ไทล์ (Percentile).....	253
	วิธีการคำนวณค่าผิดปกติ Outliers.....	256
บทที่ 13	การวัดการกระจายของข้อมูล (Measures of Dispersion)	261
	เป้าหมายของการวัดการกระจายของข้อมูล.....	261
	การวัดการกระจายแบบสัมบูรณ์ (Absolute Variation).....	262
	พิสัย (Range).....	263
	ส่วนเบี่ยงเบนมาตรฐาน (Standard Deviation).....	264
	ส่วนเบี่ยงเบนเฉลี่ย (Mean Deviation).....	266
	ส่วนเบี่ยงเบนควอร์ไทล์ (Quartile Deviation).....	270
	การวัดการกระจายสัมพัทธ์ (Relative Variation)	273
	สัมประสิทธิ์ของพิสัย (Coefficient of Range).....	273
	สัมประสิทธิ์ของการแปรผัน (Coefficient of Variation).....	275
	สัมประสิทธิ์ของส่วนเบี่ยงเบนเฉลี่ย (Coefficient of Average Deviation)	280

	สัมประสิทธิ์ของส่วนเบี่ยงเบนควอร์ไทล์ (Coefficient of Quartile Deviation).....	283
	การคำนวณค่าทางสถิติเบื้องต้นด้วยฟังก์ชัน describe() ของ Pandas.....	286
	การแสดงผลทางสถิติเบื้องต้นของ PandasGUI.....	288
บทที่ 14	พื้นฐานการแสดงผลข้อมูลแบบกราฟด้วย Matplotlib.....	291
	การดาวน์โหลดและติดตั้ง Matplotlib.....	291
	พื้นฐานการสร้างกราฟจุด (Scatter Plot).....	292
	การสร้างกราฟจุดจากข้อมูลหลายชุด.....	294
	การกำหนดขนาดจุด.....	297
	การกำหนดความเข้มจางของจุดข้อมูล.....	298
	การสร้างกราฟแท่ง (Bar Graph).....	299
	การสร้างกราฟวงกลม (Pie Graph).....	302
	การแบ่งสัดส่วนข้อมูลออกจากส่วนของกราฟวงกลม.....	303
	พื้นฐานการสร้างกราฟเส้น (Line Graph).....	304
	การกำหนดข้อมูล 1 แกน.....	308
	การสร้างกราฟเส้นและ Mark จุดข้อมูล.....	309
	การปรับแต่งกราฟเป็นเส้นประ.....	310
	การปรับแต่งรูปแบบกราฟเส้นด้วย linestyle.....	313
	กราฟเส้นกับข้อมูลหลายชุด.....	315
	การแยกกราฟเส้นจากข้อมูลแต่ละชุด.....	316
	การสร้างกราฟ BoxPlot.....	319
บทที่ 15	การสร้างกราฟของ PandasGUI.....	325
	การสร้างกราฟแท่ง (Bar Graph).....	326
	การสร้างกราฟเส้น (Line Graph).....	329
	การสร้างกราฟกระจายของข้อมูล หรือกราฟจุด (Scatter Graph).....	330
	การสร้างกราฟวงกลม (Pie Graph).....	331