



# Daily Merge App Report by Python



# 積み上げファイルと日別のファイルを毎朝マージしたい

---

積み上げファイル



+

前日のファイル



=

最新ファイル



マージ → データ型、欠損値、異常値を確認するアプリを開発

---



# アプリに**積み上げファイル**と**前日のファイル**をUP

## Cumulative Data Merge & Validation App

### ① ファイルアップロード

#### 積み上げデータ



Drag and drop file here

Limit 200MB per file • CSV, XLSX

Browse files



チャット用サンプルデータ積み上げ.xlsx 1.5MB



#### 昨日のデータ



Drag and drop file here

Limit 200MB per file • CSV, XLSX

Browse files



チャット用サンプルデータ前日.xlsx 11.8KB



# 積み上げデータの内訳とデータ型が表示される

## 【一昨日までのデータ】

全体プレビュー：

	店舗	注文タイプ	商品カテゴリー	商品タイプ	日付	数量	単価	仕入コスト	売上	利益
0	さいたま	オンライン	コンピューター	タブレット	2024-01-01 00:00:00	1	48100	41366	48100	6734
1	さいたま	オンライン	ソフトウェア	ソフトウェア	2024-01-01 00:00:00	8	15200	85120	121600	36480
2	さいたま	オンライン	ソフトウェア	ソフトウェア	2024-01-01 00:00:00	1	14300	10010	14300	4290
3	さいたま	オンライン	周辺機器	ヘッドセット	2024-01-01 00:00:00	1	10000	6500	10000	3500
4	さいたま	オンライン	ソフトウェア	ソフトウェア	2024-01-01 00:00:00	1	63000	44100	63000	18900

データ型：

	0
店舗	object
注文タイプ	object
商品カテゴリー	object
商品タイプ	object
日付	datetime64[ns]
数量	float64
単価	int64

# 積み上げデータの欠損値が表示される

欠損値：

↓ 🔍 🔄

	0
店舗	0
注文タイプ	0
商品カテゴリー	0
商品タイプ	0
日付	0
数量	0
単価	0
仕入コスト	0
売上	0
利益	0

# 前日のファイルの内訳とデータ型も表示される

## 【昨日のデータ】 ⇄

プレビュー：

	店舗	注文タイプ	商品カテゴリー	商品タイプ	日付	数量	単価	仕入コスト	売上	利益
0	さいたま	オンライン	ソフトウェア	ソフトウェア	2025-06-30 00:00:00	111	465000	325500	465000	139500
1	さいたま	オンライン	ソフトウェア	ソフトウェア	2025-06-30 00:00:00	1	257000	179900	257000	77100
2	さいたま	オンライン	ソフトウェア	ソフトウェア	2025-06-30 00:00:00	2.35	255000	419475	599250	179775
3	さいたま	オンライン	周辺機器	モニター	2025-06-30 00:00:00	1	223000	144950	223000	78050
4	さいたま	オンライン	ソフトウェア	ソフトウェア	2025-06-30 00:00:00	1	54000	37800	54000	16200

データ型：

	0
店舗	object
注文タイプ	object
商品カテゴリー	object
商品タイプ	object
日付	datetime64[ns]
数量	float64
単価	int64

# 前日のファイルの欠損値とデータ型の変更の有無が表示される

欠損値：

📄 🔍 🔄

	0
店舗	0
注文タイプ	0
商品カテゴリー	0
商品タイプ	0
日付	0
数量	0
単価	0
仕入コスト	0
売上	0
利益	0

データ型に変更があります

Excelファイルを閲覧中、間違えてデータ型に変更が加えられてしまった場合、検知できる



### ③ グループ別 前日比(%) の算出

[illegible]

# 前日のファイルの入力ミスなど、異常値を検知できる

## ④ 異常値検知

異常値検知の閾値(%)



異常な変動が見られるグループ：

	店舗	注文タイプ	商品カテゴリー	商品タイプ	数量_pct_change	単価_pct_change	仕入コスト_pct_change	売上_pct_change	利益_pct_change
0	さいたま	オンライン	ソフトウェア	ソフトウェア	507%	153%	-30%	-30%	-30%
3	東京	オンライン	ソフトウェア	ソフトウェア	100%	208%	686%	686%	1476%
6	東京	店舗	周辺機器	ヘッドセット	350%	-24%	199%	199%	199%

スライダーで閾値を200%に設定

3グループに異常値があるとアラートが出た

# 前日のファイルの入力ミスなど、異常値を検知できる

	A	B	C	D	E	F	G	H	I	J
1	店舗	注文タイプ	商品カテゴリ	商品タイプ	日付	数量	単価	仕入コスト	売上	利益
2	さいたま	オンライン	ソフトウェア	ソフトウェア	2025/6/30	111	465000	325500	465000	139500
3	さいたま	オンライン	ソフトウェア	ソフトウェア	2025/6/30	1	257000	179900	257000	77100
4	さいたま	オンライン	ソフトウェア	ソフトウェア	2025/6/30	2	255000	419475	599250	179775
5	さいたま	オンライン	周辺機器	モニター	2025/6/30	1	223000	144950	223000	78050
6	さいたま	オンライン	ソフトウェア	ソフトウェア	2025/6/30	1	54000	37800	54000	16200
7	さいたま	オンライン	コンピューター	デスクトップP	2025/6/30	2	90000	116100	135000	18900
8	さいたま	オンライン	コンピューター	デスクトップP	2025/6/30	1	80000	68800	80000	11200
9	東京	オンライン	ソフトウェア	ソフトウェア	2025/6/30	1	22900	16030	22900	687000

実は検証のため、異常値になるように2か所の値を2桁多く設定していた  
結果、2グループが正しく異常値として検出  
なお、残りの1グループはたまたま大量購入があった模様

6	東京	店舗	周辺機器	ヘッドセット	350%
---	----	----	------	--------	------



That' s a Wrap