



SAS® FORUM Japan 2020

Experience, Innovation and Resilience
#SFJ2020



Python・Rも自由自在 SAS Enterprise Guideの新しい活用法

SAS Institute Japan株式会社
Enterprise Analytics Platform Group

マネージャ
張 重陽



SAS® FORUM
Japan 2020

#SFJ2020



自己紹介

- **名前**

張 重陽(ちょう ちょうよう)

- **所属**

SAS Institute Japan

Enterprise Analytics Platform Group

- **経歴**

企業のミッションクリティカル基盤など大規模プラットフォームの研究開発に長年従事
シリコンバレーでインメモリ・データベースのコア設計・開発に数年間従事
顧客にデジタルトランスフォーメーション(DX)の課題解決、価値創出を中心に活動中



#SFJ2020

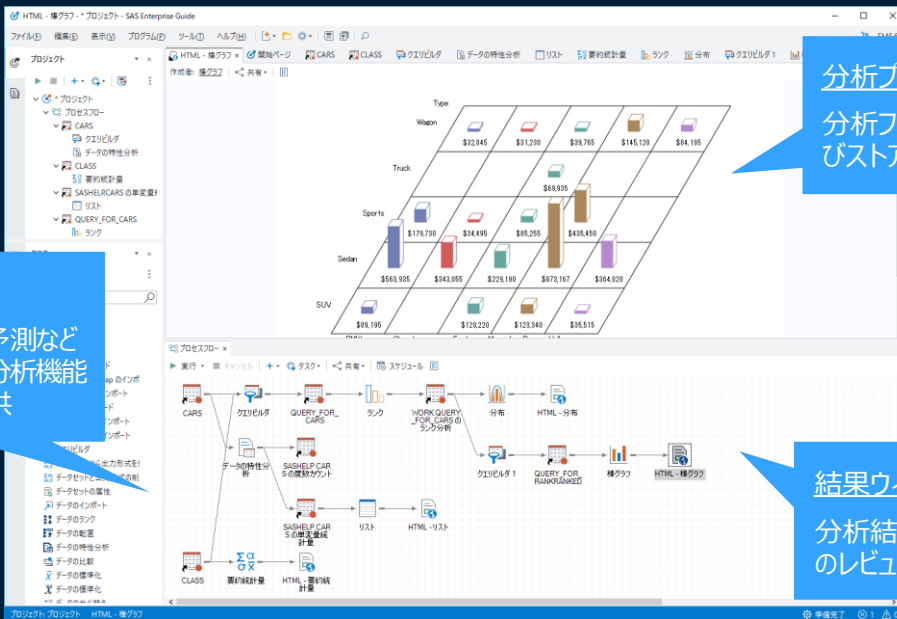
Copyright © SAS Institute Inc. All rights reserved.



SAS Enterprise Guide(EG)

データ加工/集計/可視化をドラッグ操作で実現

- データの準備から加工・変換、集計・分析、レポート作成に至る一連の作業を標準分析用タスクを活用し、ドラッグ操作で実現
- 透過的にデータにアクセスし、分析プロセスはグラフィカルなフロー図として可視化され、共有が容易に実現



分析タスクリスト
クエリや集計、予測などのSASの高度な分析機能をメニュー化し提供

分析プロセスフロー図
分析フローの設計/管理及びストアドプロセスの登録

結果ウィンドウ
分析結果の表示やデータのレビューの実施

主な機能

- データ加工処理
 - 統計手法を活用した高度な分析
 - クイック統計量(データプロファイル)
 - アドホックなデータ探索
 - 様々なグラフ
 - レポート作成
 - ストアドプロセス作成
 - 分析プロセスフロー共有
 - SASコードの作成/編集/実行
- 等



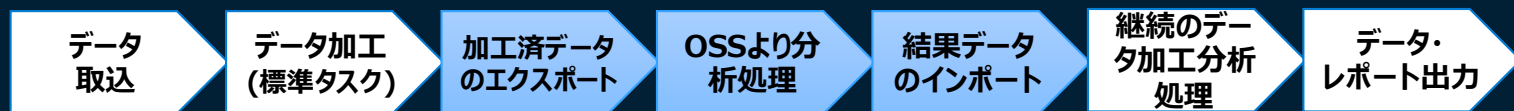
#SFJ2020



EGとオープンソース（OS）の連携について

Python・Rとのコラボレーションの現状と期待

EGでOSSの実行、バッチ処理まで簡単に実行



EGユーザーの声：



オープンソースのバッチ処理を効率化したい



Python・Rの様々なライブラリを使いたい



Python・Rの実行をEGのフローでやりたい



既存のオープンソースコードをEGに取り込みたい



ネット上のPython・Rコードを簡単に使いたい

#SFJ2020

EGから利用可能なタスク

標準タスク + カスタムタスク

EGで利用できるタスク

標準タスク

ビルドイン・タスク

SAS Studio
タスク

カスタムタスク

EG カスタムタスク

OpenSourceTask

SAS Studio
カスタムタスク



#SFJ2020

Copyright © SAS Institute Inc. All rights reserved.



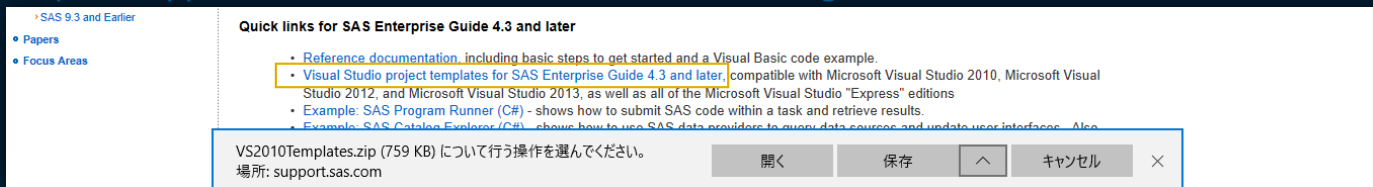
EGのカスタムタスク

ユーザ自身で機能を定義できるタスク

◆ カスタムタスクの作成:

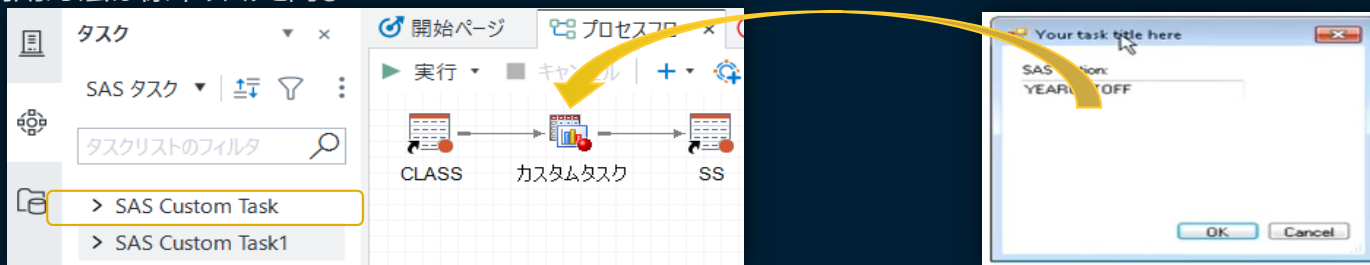
- 業務ニーズに応じ分析手法、機能から、タスクのUIまでカスタマイズ
- SASのサポートサイトから、テンプレートをダウンロード、誰でも入手可

<https://support.sas.com/documentation/onlinedoc/guide/customtasks/>



◆ カスタムタスクの使用:

- 事前準備は該当フォルダへコピーのみ
- 利用方法は標準タスクと同じ



#SPJ2020

OSTの紹介

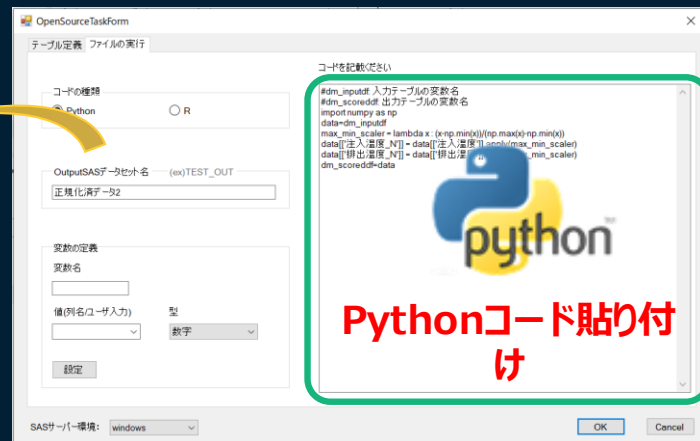
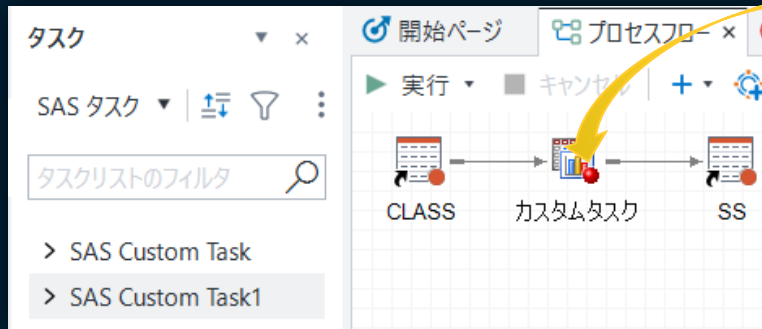
EGで簡単にオープンソースコードを実行できるカスタムタスク

◆ OSTとは？

- EGカスタムタスクのテンプレートに基づいて開発したタスク
- EGフロー上のデータセットに対し、オープンソースのコードを実行、その結果をフロー上に出力

◆ OSTの特徴

- 標準タスクと同じ使い方で使用可能
- Pythonのコードを貼り付け、実行をクリックするだけ
- 任意のOSバージョン、ライブラリを利用可能
- デバッグ用のエラーメッセージ、ログも確認可能



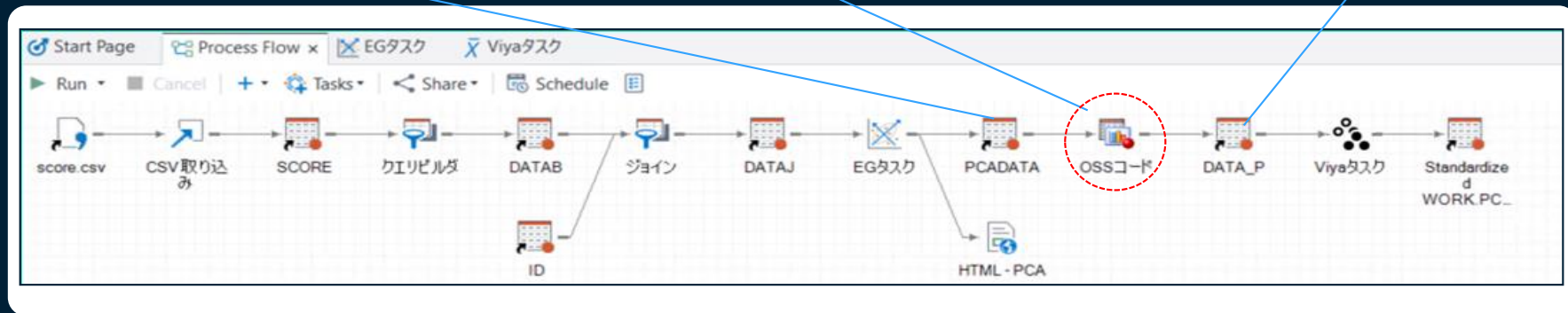
OSTの利用イメージ

GUI操作でオープンソースコードを活用

OSTの入力に任意のデータセットを指定

Pythonであらかじめ作成されたモデリングコードなどをEGのプロセスフローから実行

Pythonの処理結果をそのままデータセットでフローに出力



OSS処理を含む分析フローを保存し、チームに共有し、再利用も可能
EGのスケジューリング機能を使い、Pythonコードを簡単に自動実行

#SFJ2020

錠剤の製造データ

出荷検査と製造条件の関連性分析の前処理

APIロット番号	API0001	API0004	API00017	API00036	API00055	API00058
API粒子サイズ	Small	Small	Medium	Small	Large	Large
粉碎時間	27	11	20	13	13	19
スクリーンサイズ	4	5	4	3	5	4
ステアリン酸マグネシウム 供給業者	Smith Ind	Jones Inc	Jones Inc	Smith Ind	Smith Ind	Smith Ind
乳糖供給業者	James Ind	James Ind	Bond Inc	Bond Inc	James Ind	Bond Inc
砂糖供給業者	Sour	Sour	Sour	Sweet	Sweet	Sweet
タルク供給業者	Rough	Smooth	Rough	Smooth	Smooth	Rough
攪拌開始時刻	2017/01/10 12:00:00	2017/01/10 13:00:00	2017/01/10 14:00:00	2017/01/10 15:00:00	2017/01/10 16:00:00	2017/01/10 17:00:00
攪拌終了時刻	2017/01/10 12:16:03	2017/01/10 13:14:22	2017/01/10 14:14:28	2017/01/10 15:14:25	2017/01/10 16:16:07	2017/01/10 17:12:52
攪拌速度	59.88738347	59.79302953	60.79748693	59.39862446	59.92163909	59.39309584
圧縮機	Compress2	Compress2	Compress2	Compress1	Compress2	Compress2
圧力	25.4666912	24.93406224	25.50390751	24.78763191	25.31677487	24.62765021
滑沢剤供給業者	Mac	Mac	Down	Mac	Down	Mac
滑沢剤の粘度	105.3611582	100.2324839	100.2716029	98.0081339	97.58263028	94.14173417
注入温度	107.9452278	107.5401013	106.6058611	106.0937342	108.2573814	106.3463243
排出温度	70.45899545	70.8236838	69.19734969	68.80908707	69.4473854	69.14887509
噴霧量	404.6032939	407.3671634	399.3018683	403.7339267	396.7093465	404.702108
噴霧器圧力	61.00706763	60.57121956	59.08822803	58.84462958	59.58913672	60.39150307
溶出	75.85	73.77	73.43	72.86	74.22	69.46
ロット受け入れ	Accept	Accept	Accept	Accept	Accept	Reject

#SFJ2020

デモの紹介

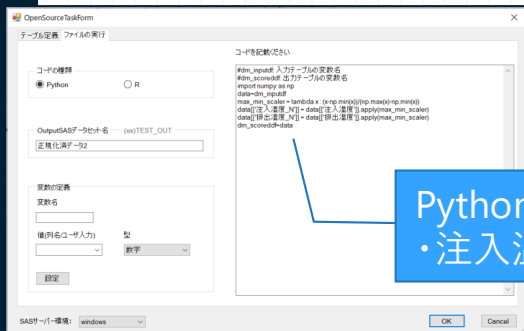
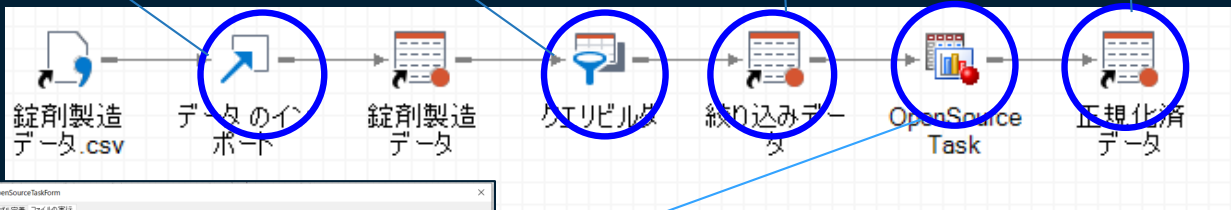
PythonコードでEG上のデータセットを標準化

csvデータインポート

攪拌時間の計算と
データの絞り込み

OSTの処理対象
データ

OSTの処理結果
データ



Pythonコード処理：
・注入温度と排出温度の正規化処理



#SFJ2020

Copyright © SAS Institute Inc. All rights reserved.





SAS® FORUM
Japan 2020

ご視聴ありがとうございました！

#SFJ2020





SAS® FORUM Japan 2020

Experience, Innovation and Resilience
#SFJ2020

 https://www.sas.com/ja_jp/home.html

 <https://www.facebook.com/SASJapan>

 <https://twitter.com/SASJapan>

