

データ基盤プロジェクト - POS #197

NEC 生データ整形環境検討・構築

2023/11/15 11:35 - 武田 遼河

ステータス:	完了	開始日:	2023/11/15
優先度:	通常	期日:	2024/04/30
担当者:	武田 遼河	進捗率:	80%
カテゴリ:		予定工数:	15.00時間
対象バージョン:		作業時間:	0.00時間

説明
現行は日別・店別で各データを取り込んでいるが、取込ファイル数が大量になることでランニングコストが増大している。コストカットのため、日別単位で各データを整形してs3にアップロードすることを検討する。

【TODO】

- ・ 整形環境の検討
 オンプレ？ EC 2 ？
- ・ 実行環境の検討
 bat？ 期間DB経由？ アプリ？
- ・ 整形仕様検討
 営業日が複数存在してはならない等の制約はある？

履歴

#1 - 2023/11/22 13:34 - 武田 遼河

- 期日を 2023/11/24 から 2023/12/15 に変更

【方針】

- ・ 整形環境についてはひとまずオンプレ（ファイル共有サーバー）で構築

【進め方】

1. CU: 取込が必要なファイルを選定して村上さん側に通知 済
2. 村上さん: 整形後フォーマットを指定してCU通知
3. CU: 実装
4. 村上さん: 自動連携開始

【実装方法】

- ・ 以下 3 つの方法で検討しています。最もメンテがしやすいものを採用しようと思っています。
- ・ C# .exe形式
- ・ batch(Powershell)
- ・ python .exe形式

#2 - 2023/11/27 09:45 - 武田 遼河

- ステータスを 新規 から 進行中 に変更

<<<期日の設定消える現象テスト>>>

#3 - 2023/11/27 09:51 - 武田 遼河

- 期日を 2023/12/15 にセット

- 進捗率を 100 から 20 に変更

期日再設定

#4 - 2023/11/27 19:08 - 武田 遼河

- ファイル 446_yyyymmdd (UTF-8 CSV) .xlsx を追加
- ファイル 447_yyyymmdd (UTF-8 CSV) .xlsx を追加
- ファイル 448_yyyymmdd (UTF-8 CSV) .xlsx を追加
- ファイル 492_yyyymmdd (UTF-8 CSV) .xlsx を追加
- ファイル 493_yyyymmdd (UTF-8 CSV) .xlsx を追加
- ファイル Z09_yyyymmdd (UTF-8 CSV) .xlsx を追加

- ファイル 001_yyyymmdd (UTF-8 CSV) .xlsx を追加
- ファイル 102_yyyymmdd (UTF-8 CSV) .xlsx を追加
- ファイル 401_yyyymmdd (UTF-8 CSV) .xlsx を追加
- ファイル 434_yyyymmdd (UTF-8 CSV) .xlsx を追加
- 進捗率を 20 から 40 に変更
- 予定工数を 15.00時間 にセット

添付のフォーマットと以下仕様で合意しております。

ファイル名称
[ファイル名]_[yyyymmddhhMM].csv
日付はs3アップロード日時
コンマ区切りcsv

ファイル内容
「店舗コード」「営業日付」を各レコードにカラムとして追加
1 ファイルに複数営業日の混在あり

【TODO】

- ・管理方法決定（バージョン管理ツール導入？）
- ・11/9以降データを新整形で送付
- ・過去欠損調査（444）
- ・マスタファイル連携検討

#5 - 2023/12/15 17:46 - 武田 遼河

- 期日を 2023/12/15 から 2024/01/31 に変更
- 進捗率を 40 から 70 に変更

TODO単位で進捗を記載します。
12/15期限としていた整形スクリプトの作成まで完了

【TODO】

整形スクリプト作成	済	
整形ファイルAWSアップロード処理作成	50%	
45サーバー機能移管	0%	
整形仕様ドキュメント作成	0%	

#6 - 2023/12/26 12:34 - 武田 遼河

- 進捗率を 70 から 80 に変更

【TODO】

整形スクリプト作成	済	
整形ファイルAWSアップロード処理作成	済	
45サーバー機能移管	80%	
整形仕様ドキュメント作成	0%	

AWSアップロード機能と自動化を実装いたしました。
一部backup処理が12サーバーに残留しているため、完全移行まで作業予定です。

#7 - 2024/02/01 19:11 - 武田 遼河

- 期日を 2024/01/31 から 2024/02/29 に変更

CC売上データ収集の仕組み構築を優先いたしますので期限を延長いたします。

#8 - 2024/03/25 12:13 - 武田 遼河

- 期日を 2024/02/29 から 2024/04/30 に変更

タスク整理の上期限を延長します。

#9 - 2024/05/03 14:01 - 武田 遼河

- ステータスを 進行中 から 完了 に変更

データ基盤整形環境についてはまかせてネットへ機能移管の方向で調整中とのことですので、本チケットの残タスクはいったん取り止めとクローズします。

ファイル

446_yyyymmdd (UTF-8 CSV) .xlsx	10.4 KB	2023/11/27	武田 遼河
447_yyyymmdd (UTF-8 CSV) .xlsx	10.6 KB	2023/11/27	武田 遼河
448_yyyymmdd (UTF-8 CSV) .xlsx	10.6 KB	2023/11/27	武田 遼河
492_yyyymmdd (UTF-8 CSV) .xlsx	10.4 KB	2023/11/27	武田 遼河
493_yyyymmdd (UTF-8 CSV) .xlsx	11.4 KB	2023/11/27	武田 遼河
Z09_yyyymmdd (UTF-8 CSV) .xlsx	10.9 KB	2023/11/27	武田 遼河
102_yyyymmdd (UTF-8 CSV) .xlsx	11.7 KB	2023/11/27	武田 遼河
001_yyyymmdd (UTF-8 CSV) .xlsx	10.5 KB	2023/11/27	武田 遼河
401_yyyymmdd (UTF-8 CSV) .xlsx	12 KB	2023/11/27	武田 遼河
434_yyyymmdd (UTF-8 CSV) .xlsx	10.5 KB	2023/11/27	武田 遼河