

【データ基盤】レコードが重複している

2024/08/22 18:15 - 島貫 圭太

ステータス:	完了	開始日:	2024/08/22
優先度:	高め	期日:	2024/09/30
担当者:		進捗率:	0%
カテゴリ:		予定工数:	0.00時間
対象バージョン:		作業時間:	0.00時間

説明

86038KK堺筋本町店の監査データについて
8/19が重複している おそらくFFPにて修正された事でトランザクションデータが2回連携された。

```
1 SELECT
2   col_3,col_5,col_14,col_15,col_17,col_72,col_73,col_75
3 FROM
4   "redshift_db"."dwh"."dwh_nec_tbl_493"
5   where col_5 BETWEEN '20240801' and '20240831'
6   and col_3 = '086038'
7   order by col_5 asc;
```

Result 1 (22)

col_3	col_5	col_14
086038	20240805	000001
086038	20240810	000004
086038	20240811	000002
086038	20240812	000000
086038	20240813	000003
086038	20240814	000001
086038	20240815	000000
086038	20240816	000001
086038	20240817	000003
086038	20240818	000001
086038	20240819	000002
086038	20240819	000002
086038	20240820	000003
086038	20240821	000002

2回目以降のデータが正となる様な制御が必要 DMへのデータ生成時でもOK

履歴

- #1 - 2024/08/26 16:25 - 武田 遼河
- ステータスを新規から完了に変更

2024/8/26 Mtg

- ・ 該当の重複についてはFFP（POS管理システム）側で修正が発生した場合に自動でデータ基盤S3部分まで連携される仕様
他通常データと区別なく1ファイルに整形されて連携

【対応】

- ・ 改修について

まかせてネット側への要求仕様として、修正データはS3別領域へ別ファイルとして出力する仕様として依頼中
現行システム構成での対応はせずまかせて移管時の用件として調整する

- ・ 本チケット該当データのリカバリについて

dwh領域は取り込まれた全ファイルデータが共存 PUBLIC,DM段階で店舗コード&営業日付でユニークとなるようメンテナンスしていく方針となります。

該当データ種別についてはDM化がまだなので、今回については削除等対応は実施せず、DM化の際にまとめてデータクレンジングを実施します

。

本件としてはクローズします。

ファイル

clipboard-202408221811-dlyvi.png	55.3 KB	2024/08/22	島貫 圭太
----------------------------------	---------	------------	-------