

# STREAMLINED BATCH CORRECTION APPLICATION

COMMODITY FLOW SURVEY

U.S. Census Bureau

Patrick Nguyen – Supervisor

Cecile Murray – Technical Mentor

# BACKGROUND

## Departments

- Bureau of Transportation Statistics
- US Census Bureau

## Uses

- Transportation Facilities
- Energy Use
- Safety Risk
- Environmental Concerns

# PROJECT OVERVIEW

## Survey Volume

- 2017 – Previous Survey Year: ~6 million rows of data collected
- 2022 – Most Recent Survey: ~80 million rows of data collected

## Collection Method

- Respondus Portal: digitized submission of shipment data
- Inaccuracies in data

# PREVIOUS SYSTEMS

## 2017 System 1

- required searching of individual shipment/destination IDs and manual updating
- Slow, Non-Scalable, Unintuitive

CFS 2022 FROGS (Production)    Precanvass   Main Menu ▾   Mail Group   Prior Surveys ▾   Macro Tool   Close

Enter ID and press enter: <input type="text"/>										Quarter	01 ▾	Panel	<input type="text"/>	Enter Line Number: <input type="text"/>	⌂	Select Change Source ▾	Save Changes
Status	Establishment	Shipments	Destination	Weighting	Shipment Flags	Imputation Flags	Remarks	Address/Contacts	Audit Report	Search	Admin						
EHM																	

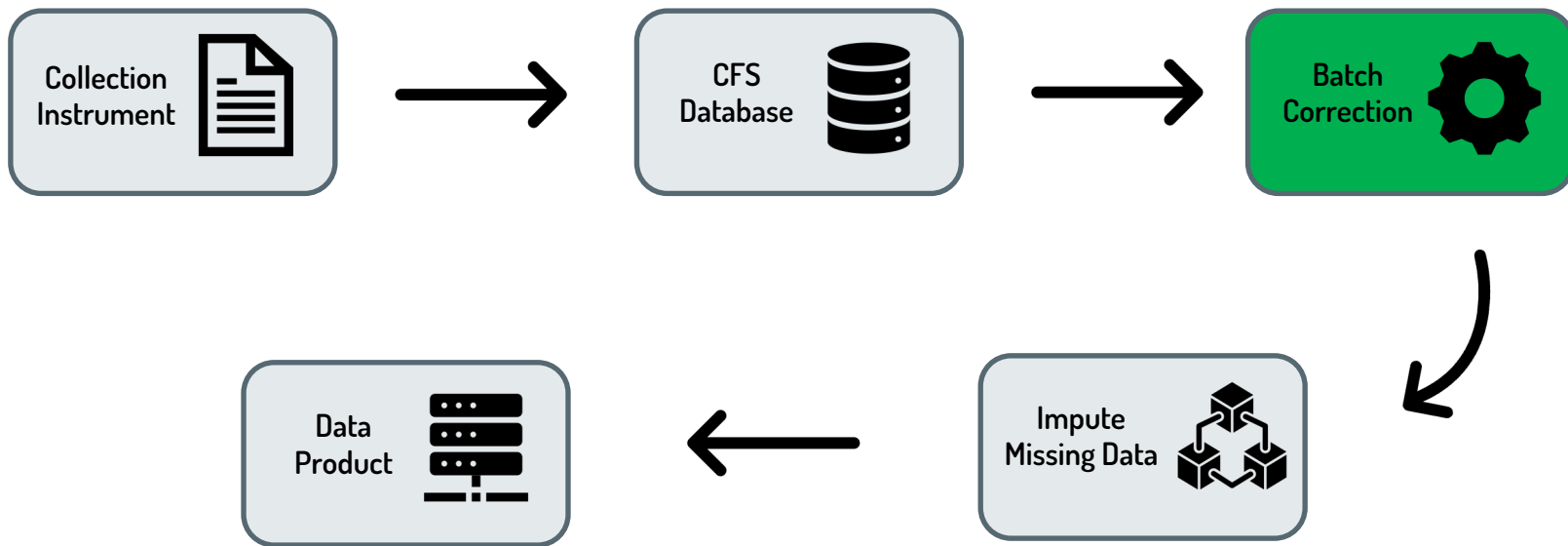
## 2017 System 2

- Batch correction which required a lengthy and slow process, and over a week for changes to apply

# PROJECT GOAL

- Replace the outdated correction system for the Commodity Flow Survey Team by developing an application that permits analysts to overwrite incorrect data in batches based on custom filters
- Reduce turnaround time for database corrections from over a week to ~24 hours
- Improve upon previous CDF's application

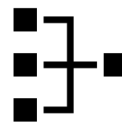
# CFS Architecture



# BATCH CORRECTION CONCEPT



Incorrect Input



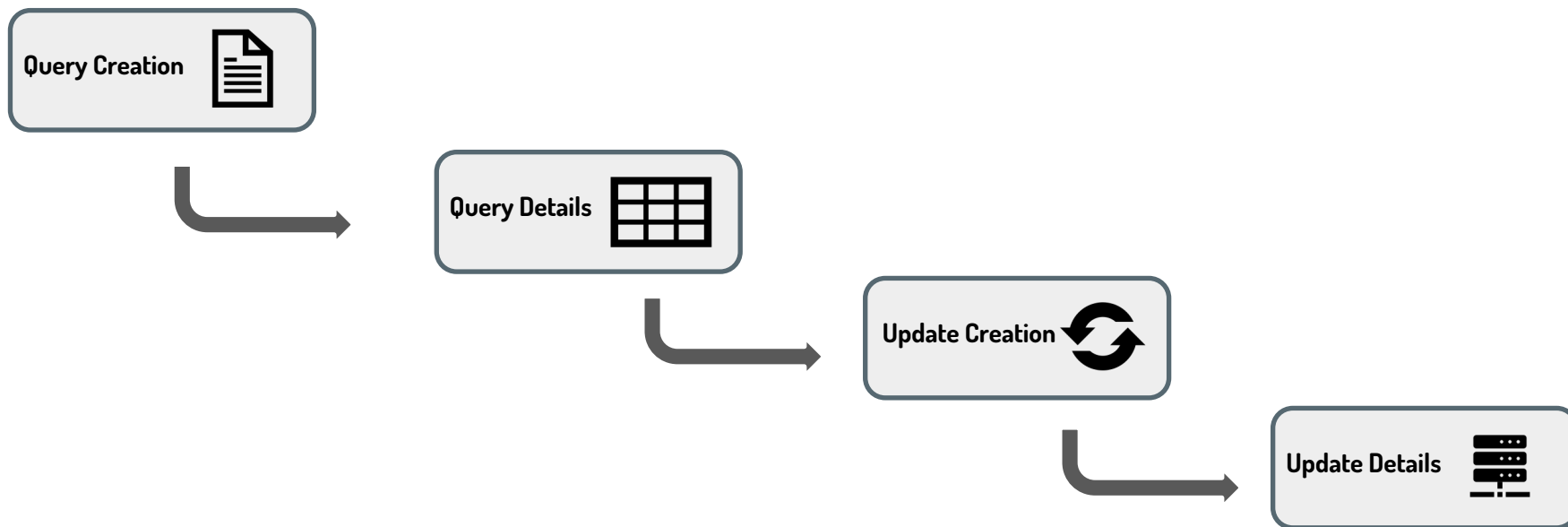
Batch Correction



Corrected Data

Shipment Date	["27/03/2022", "2022/03/27", "March 27, 2022", "3/27/22"]	→	03/27/2022
Value, Weight	[-1, None, ∞, "0"]	→	ex: 400
Export Code	[None, null, N/A, 0, False]	→	N
Shipment Day	[twenty seventh, Sunday, 0]	→	27
Hazmat Code	[empty, 0, 1, True, None, N/A]	→	ex: 1202

# BATCH CORRECTION FLOW





# QUERY CREATION

Disclosure Prohibited, Confidential, Title 26 U.S.C. [HOME](#) [CREATE QUERY](#) [QUERIES](#) [UPDATES](#) [LOGOUT](#)

## CREATE QUERY

### REQUIRED

**TITLE**

SELECT TABLE ▼

**DESCRIPTION**

enter query description

**TABLE NAME**

SELECT TABLE ▼

### QUERY PARAMETERS

ID_NUMBER	<input type="text"/>	<input type="radio"/>	HAZMAT CODE	<input type="text"/>	<input type="radio"/>
EXPORT CODE	<input type="text"/>	<input type="radio"/>	WEIGHT	<input type="text"/>	<input type="radio"/>
SHIPMENT DATE	<input type="text"/>	<input type="radio"/>	SHIPMENT DAY	<input type="text"/>	<input type="radio"/>
SHIPMENT YEAR	<input type="text"/>	<input type="radio"/>	SHIPMENT VALUE	<input type="text"/>	<input type="radio"/>

**QUARTER**

SELECT QUARTER ▼

[PREVIEW QUERY](#)

QUERY PREVIEW

SELECT FROM WHERE

# NEXT STEPS

## Deployment

- Internal testing and approval
- Integrate into the current process workflow
- Push to deployment, receive feedback and iterate upon the application fixing bugs and adding requested functionality

## Functionality

- Add an approval/rejection system for leadership to view incoming updates
- Develop functionality that permits analysts to perform complex lookups across datasets

# THANK YOU!