

Franchise Disclosure Document Roster Table Extraction

Business Frame Team

U.S. Census Bureau

Jessica Wellwood — Assistant Division Chief

INTRODUCTION & BACKGROUND

- Franchise Disclosure Document (aka FDD): Released every year or so by franchised enterprises.
- Usually around 300-500 pages long.
- One item is a roster table of establishments (stores) under the franchise.

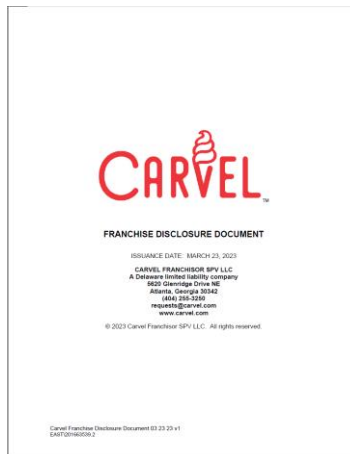


TABLE OF CONTENTS	
	PAGE
ITEM 1 THE FRANCHISOR AND ANY PARENTS, PREDECESSORS AND AFFILIATES	1
ITEM 2 BUSINESS EXPERIENCE	15
ITEM 3 LITIGATION	18
ITEM 4 BANKRUPTCY	19
ITEM 5 INITIAL FEES	19
ITEM 6 OTHER FEES	21
ITEM 7 ESTIMATED INITIAL INVESTMENT	33
ITEM 8 RESTRICTIONS ON SOURCES OF PRODUCTS AND SERVICES	42
ITEM 9 FRANCHISEE'S OBLIGATIONS	46
ITEM 10 FINANCING	48
ITEM 11 FRANCHISOR'S ASSISTANCE, ADVERTISING, COMPUTER SYSTEMS AND TRAINING	49
ITEM 12 TERRITORY	64
ITEM 13 TRADEMARKS	68
ITEM 14 PATENTS, COPYRIGHTS AND PROPRIETARY INFORMATION	70
ITEM 15 OBLIGATION TO PARTICIPATE IN THE ACTUAL OPERATION OF THE FRANCHISE BUSINESS	72
ITEM 16 RESTRICTIONS ON WHAT THE FRANCHISEE MAY SELL	73
ITEM 17 GENERAL, TERMINATION, TRANSFER AND DISPUTE RESOLUTION	74
ITEM 18 PUBLIC POLICES	80
ITEM 19 FINANCIAL PERFORMANCE REPRESENTATIONS	80
ITEM 20 OUTLETS AND FRANCHISEE INFORMATION	81
ITEM 21 FINANCIAL STATEMENTS	85
ITEM 22 CONTRACTS	88
ITEM 23 RECEIPT	88

EXHIBITS	
EXHIBIT A FINANCIAL STATEMENTS	
EXHIBIT B CARVEL FRANCHISE AGREEMENT AND RELATED AGREEMENTS	
Schedule A - Franchise Specific Terms	
Schedule B - Personal Guarantees	
Schedule C - Guarantees of Payment and Performance	
Schedule D - State Law Acknowledgment (If Required)	
Schedule E - Multi-Unit Acknowledgment (If Applicable)	
EXHIBIT C OTHER AGREEMENTS	
Carvel Express Schedule (to replace Schedule A of the Franchise Agreement)	
Carvel Hotspot Express Schedule (to replace Schedule A of the Franchise Agreement)	
Carvel Ice Cream Truck Schedule (to replace Schedule A of the Franchise Agreement)	
Carvel Branded Truck Schedule (to replace Schedule A of the Franchise Agreement)	
Carvel Branded Shoppe Schedule (to replace Schedule A of the Franchise Agreement)	
Carvel Branded Support Services Agreement	
EXHIBIT D INFORMATION ON FRANCHISEES	
EXHIBIT E INFORMATION ON FORMER FRANCHISEES	
EXHIBIT F STATE ADMINISTRATORS	
EXHIBIT G AGENTS FOR SERVICE OF PROCESS	
EXHIBIT H STATE ADDENDA TO DISCLOSURE DOCUMENT	
EXHIBIT I FRANCHISE DISCLOSURE ACKNOWLEDGEMENT	

OBJECTIVE & SCOPE

Goal: Given an FDD, extract its roster table(s) and format them for the Business Register.

Obstacles:

- PDF file format is unstructured.
- Different FDDs are formatted differently. Rosters are in different locations, have different column names, etc.



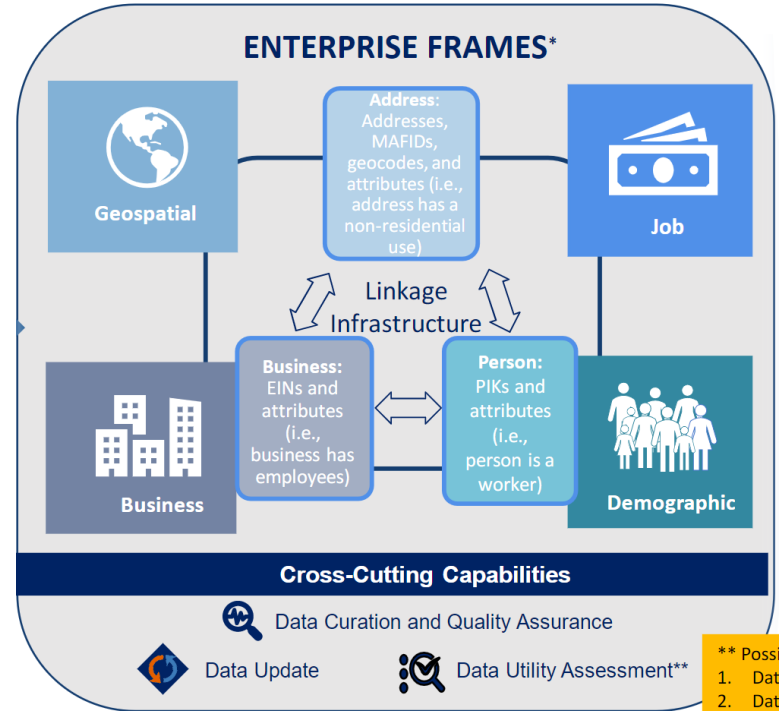
NAME1	STREET	CITY	STATE_ABBR	ZIP	CONTACT_PHONE
0 PRETZEL KING, LLC (CA)	5701 Outlets at Tejon Pkwy	Arvin	California	93203	(818) 634-2449
1 ZHOOR, INC.	7 Fairfax Road	Bakersfield	California	93307	(661) 847-9949
2 JEDO INC.	825 Buena Vista Road	Bakersfield	California	93311	(661) 847-9949
3 OLUF Incorporated	11037 Santa Monica Boulevard	Los Angeles	California	90025	(310) 444-0011
4 Khyber LLC	401 West Main Street	Avon	Connecticut	06001	(860) 409-0582
5 Fresh Worx, Inc.	650 Farmington Ave.	Bristol	Connecticut	06010	(860) 845-8736
6 Zip Enterprises, LLC	20-A Killingworth Turnpike	Clinton	Connecticut	06413	(860) 235-4537
7 Joseph Kozakwicz	1 Padanaram Road	Danbury	Connecticut	06811-4820	(203) 792-3428
8 Shailesh Saoji, Hufen Zhong	2864 - A Main Street	Glastonbury	Connecticut	06033	(860) 633-2881
9 MJB Concepts, LLC	1 Civic Center Plaza	Hartford	Connecticut	06103	(203) 318-1566
10 Michael Amone	1081 Bridgeport Avenue	Millford	Connecticut	06460-3147	(203) 874-1427
11 Connecticut Creamery, LLC	778 New Haven Road	Naugatuck	Connecticut	06770	(203) 723-7770
12 Albert Yaldek	14 Danbury Rd	New Milford	Connecticut	06776-3402	(860) 354-3554
13 RSJSL, LLC	1018 Main Street	Newington	Connecticut	06111-2908	(860) 667-1467
14 A & K International Inc.	456 Main Avenue	Norwalk	Connecticut	06851-1057	(203) 846-3002
15 A & K International Inc.	324 Connecticut Avenue	Norwalk	Connecticut	06854	(203) 831-0300
16 Wei Wen Chen	113 Danbury Road	Ridgefield	Connecticut	06877-4113	(203) 438-7424
17 A. L. LLC	1060 Long Ridge Road	Stamford	Connecticut	06903	(203) 569-3600
18 Justino Inc.	1575 Barnum Ave.	Stratford	Connecticut	06614	(203) 380-0033
19 Torrington Ice Cream, Inc.	650 S Main St.	Torrington	Connecticut	06790-6923	(860) 489-8405
20 FK&ST LLC	243 Hartford Turnpike	Vernon	Connecticut	06066	(860) 875-2278
21 GG Creamery, LLC	700 Watertown Avenue	Waterbury	Connecticut	06708-2031	(203) 754-5518
22 Watertown Creamery, LLC.	1300 Main St	Watertown	Connecticut	06795-3110	(860) 274-1462

MOTIVATION

Part of the *Frames Program*: an overarching project to improve Census databases with record linkage.

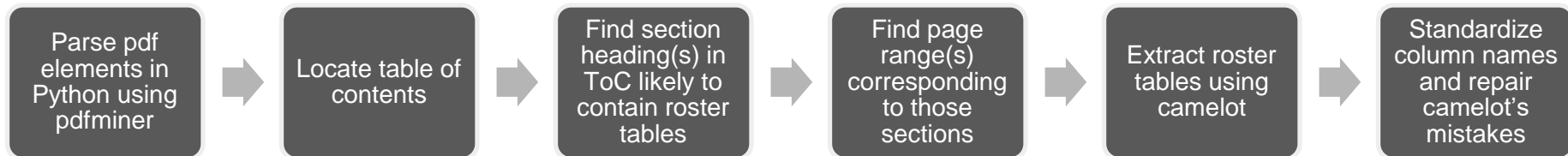
- *record linkage*: connecting entries (possibly in different datasets) that refer to the same entity

Extracting data from FDD roster tables supports the Frames Program by adding to the Business Register, which is being used as a model for the Business Frame.



PROJECT DETAILS & PROCESS

Original plan: Automate (1) identifying relevant pages *and* (2) extracting tables.



PROJECT DETAILS & PROCESS

Original plan: Automate (1) identifying relevant pages *and* (2) extracting tables.

Problem: Step 1 is slower and less reliable than a human. **47% success rate, ~1 minute per FDD.**

- PDF parser package takes a long time to run
- PDF parser package sometimes misses entire elements
- Some sections in the ToC are missing, reordered, or renamed in the actual FDD

New plan (as of August 4th): Have a human identify relevant pages and what layout the tables use.

- Knowing the table layout in advance speeds up extraction and improves accuracy.
- Faster than automatically inferring the table layout
- Will likely improve accuracy since we know the layout
- Still WIP; using a fellowship extension

NEXT STEPS & CONCLUSION

Bias testing

- Identify protected attributes
- Get group/individual/causal fairness metrics
- Improve fairness using preprocessing, inprocessing, postprocessing

Identify permanently closed stores

- Typically in a separate table. How to identify? What if the page is shared with another table?