



United States Department of Agriculture

IRI InfoScan: retailer-based scanner data

Robert Wood Johnson Foundation/USDA, Economic Research Service
Grants Competition Informational Session

March 1, 2022

Patrick McLaughlin

The analysis, findings, and conclusions expressed in this presentation should not be attributed to IRI.



Proprietary Scanner Data

- Consumer purchase and sales data
 - **Retail point-of-sale data**
 - Household panel data
 - See presentation on IRI Consumer Network
 - Weekly COVID-19 response data
 - See separate presentation
- Used by industry for marketing research
- Used by USDA for
 - Research projects, including external *collaborations*
 - Program evaluations, regulatory impact analyses
 - Data products, both public facing and internal value-added

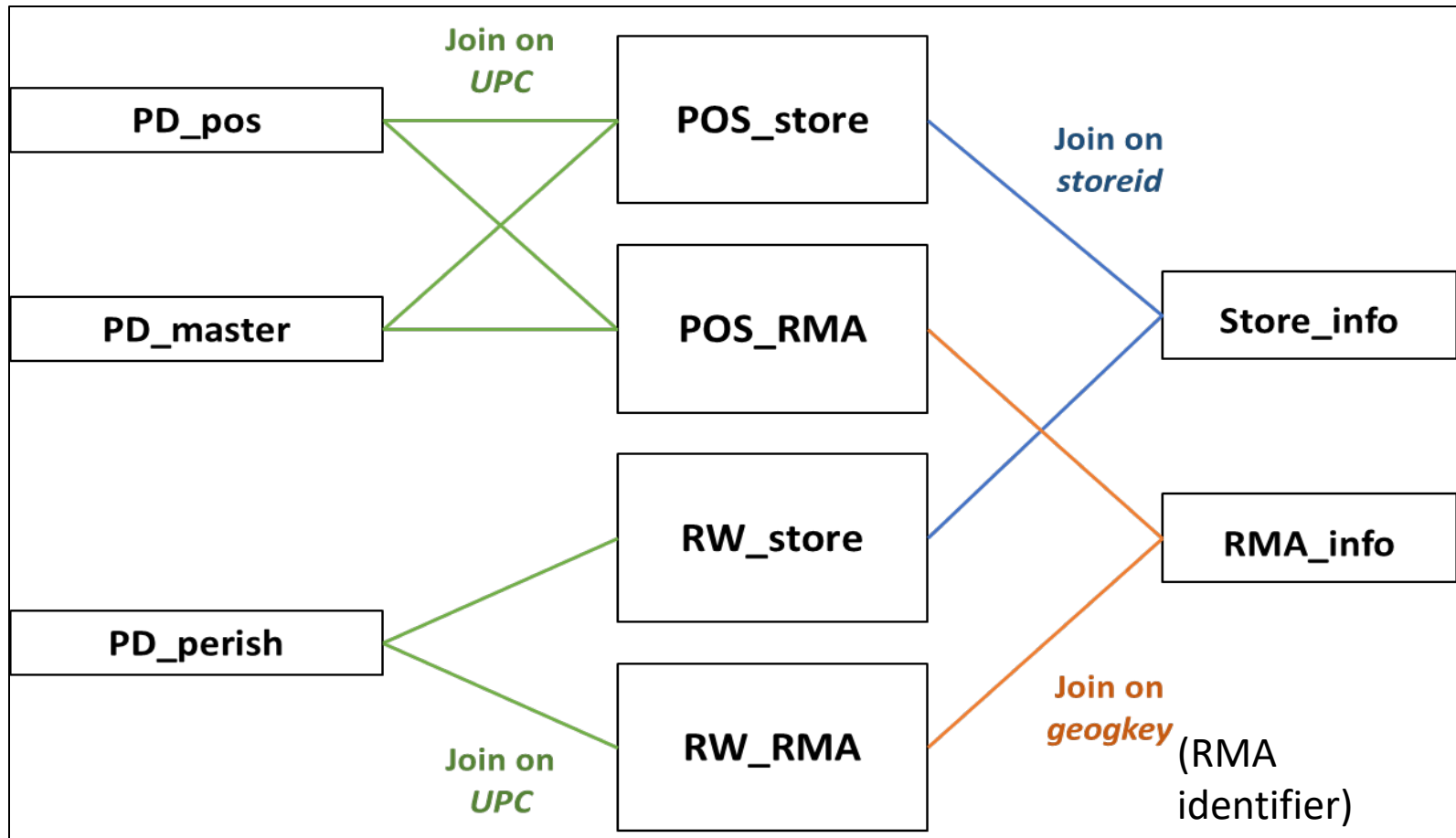


Retail Scanner Data

- IRI InfoScan retail data (2008-2020)
- What are sales value and quantities of food sold in retail stores?
 - Dollars = total dollars of a food item sold in a week, units = quantity
 - Compute prices
- Weekly, item-level sales data for about 60,000 stores for each year
 - 6 to 8 billion sales records per year
 - Proficiency with basic SQL queries strongly recommended
- Covers about 20% of store locations, 50% of total food sales
- Store types include grocery, club, supercenter, convenience, dollar, drug, military
- Detailed product information
- Store information and locations
 - Weights for stores through 2018
 - See separate presentation
 - Linkable to Nielsen TDLinx
 - See separate presentation



InfoScan Table Schema, 2008-2018



Retailer Geography

- Roughly half of stores provide store-level sales data
 - I.E., sales data from individual outlets of a chain
- The other half provides sales data at the retail marketing area (RMA)
 - E.G., aggregated across all outlets of a chain in the Washington, DC/Baltimore, MD region
 - RMA makeup varies by chain, and may vary by year
- “Store_info” and “rma_info” tables provide store-level characteristics
 - In both: store name, geographic location (address, county, State, Census tract and block), channel (e.g., grocery)
 - In rma_info: which RMA the store belongs to



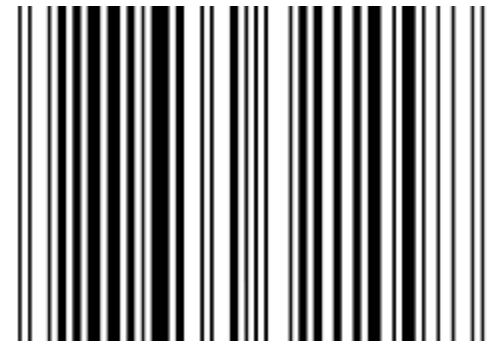
Two Broad Types of Food Items

- Point-of-sale (POS) or UPC food items (roughly half of all sales)
 - E.G. packaged and durable/shelf-stable goods like cans of soup or boxes of cereal
 - All of a “fixed weight”
 - Collection and reporting of sales essentially static across 2008-2020
- Perishables (sometimes known as random weight or RW)
 - E.G., fruits, vegetables, meat, seafood, deli + bakery items
 - Two types
 - Random weight (RW) – typically sold by weight (e.g., pounds of chicken breast) or by unit (e.g., number of lemons)
 - Fixed weight (FW) – packaged perishable items (e.g., 5 pound bag of apples, container of berries)
 - Collection and reporting varies between 2008 – 2018 vs 2019 – 2020
- All food item information in product dictionaries



POS Product Dictionary, 2008 - 2018

- Product characteristics for over 1 million UPCs per year
- Key characteristics
 - Description
 - Brand
 - Size/weight
 - Product categorization
- Attributes, health, and nutrition-related claims
 - Type, flavor
 - Nutrition Facts Panel (limited)
 - Organic, gluten-free, type of sugar/artificial sweetener



Perishables Product Dictionary, 2008 - 2018

- Product characteristics for roughly 30-50K “above-UPC” food items per year
- Key characteristics
 - Form (e.g., cut of meat)
 - Preparation (e.g., cut fruit)
 - Package size (if FW)
 - Product categorization
- Attributes, health, and nutrition-related claims
 - Claims like organic or kosher
 - Color

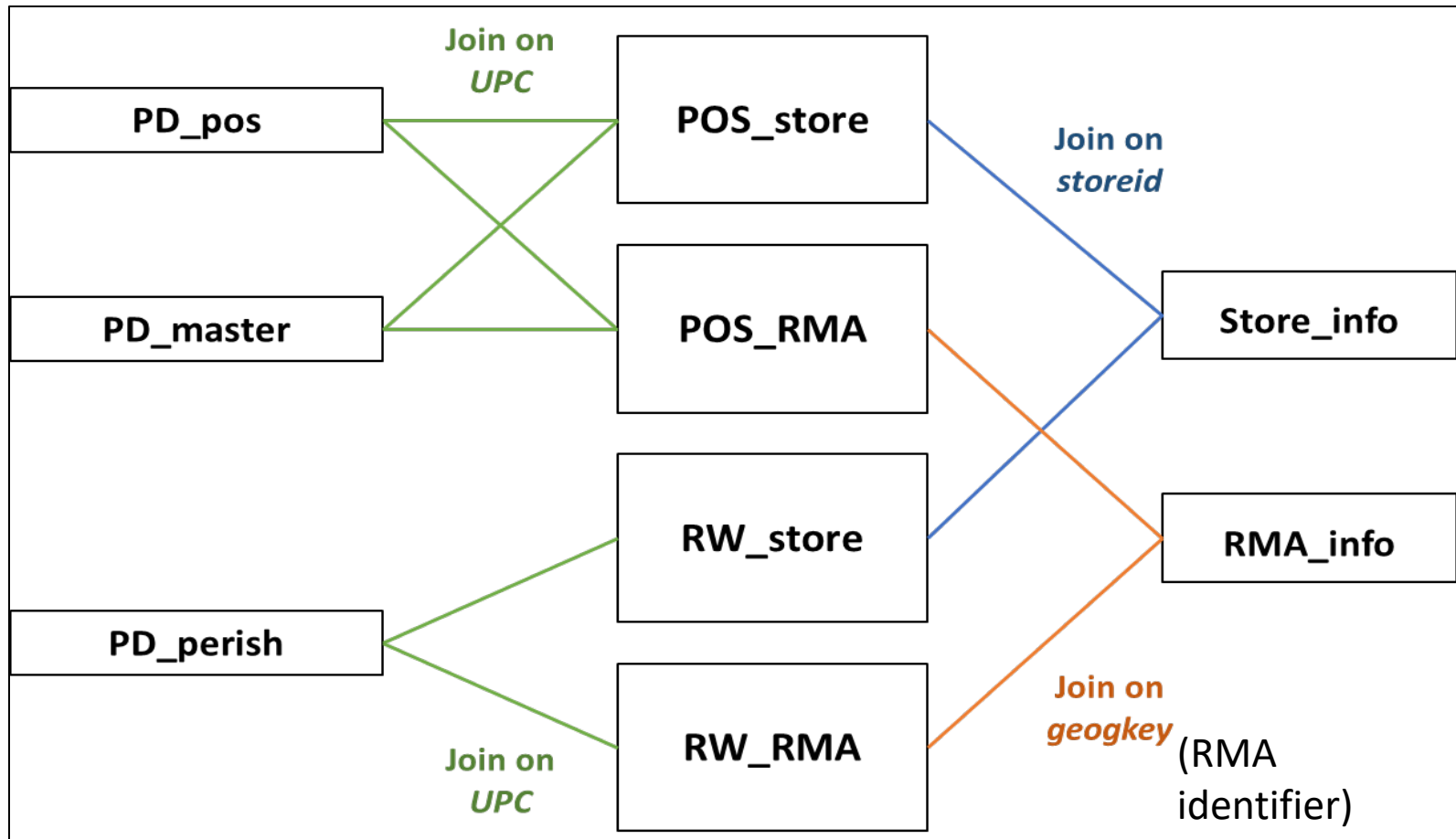


POS + Perishables Dictionary, 2019-2020

- Product characteristics for over 1 million POS UPCs per year and 150,000 perishable products
- Same columns and information as POS product dictionary, 2008-2018
 - New: brand information and description for perishable products
- Some fields not populated for perishables products
 - Closer to perishables dictionary, 2008-2018



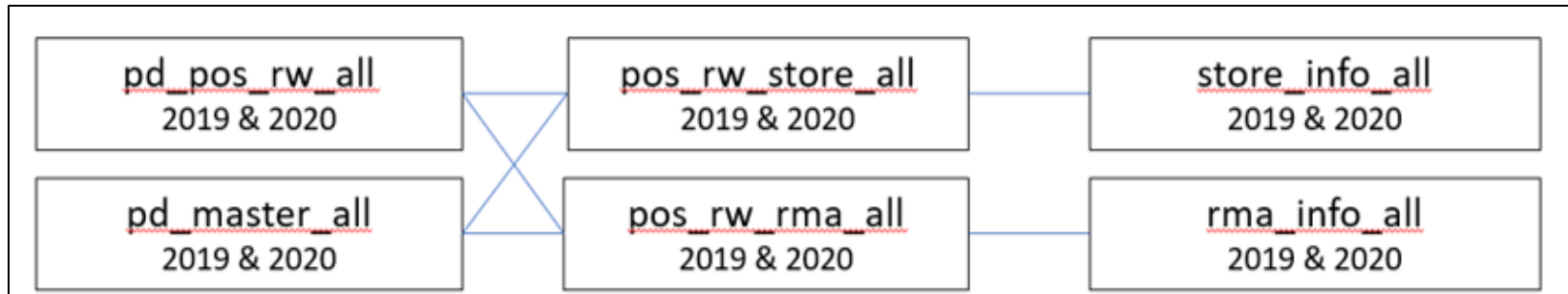
InfoScan Table Schema, 2008-2018



InfoScan Table Schema, 2019-2020

Join on UPC

Join on storeid



Join on geogkey
(RMA identifier)



For more information

- [Using Scanner Data](#) on the ERS website
 - Legal language for third-party access agreement (TPA)
 - Business offices **must** sign as is, no exceptions
- Technical bulletins
 - [Understanding IRI Household-Based and Store-Based Scanner Data](#) (TB-1942) by Mary Muth (RTI) and others
 - [Food-at-Home Expenditures: Comparing Commercial Household Scanner Data From IRI and Government Survey Data](#) (TB-1946)
 - [Examining Food Store Scanner Data: A Comparison of the IRI InfoScan Data with Other Data Sets, 2008–2012](#) (TB-1949)





United States Department of Agriculture

ERS Scanner Team

Patrick McLaughlin

Patrick.McLaughlin@usda.gov

Alex Stevens

Alexander.Stevens@usda.gov

Xiao Dong

Xiao.Dong@usda.gov

[Using Scanner Data](#)

