

Instacart data dictionary

orders (3.4m rows, 206k users):

- order_id: order identifier
- user_id: customer identifier
- eval_set: not relevant to this project
- order_number: the order sequence number for this user (1 = first, n = nth)
- order_dow: the day of the week the order was placed on 0 = saturday
- order_hour_of_day: the hour of the day the order was placed on
- days_since_prior: days since the last order, capped at 30

products (50k rows):

- product_id: product identifier
- product_name: name of the product
- aisle_id: foreign key
- department_id: foreign key
- prices: sale price to customer

departments (21 rows):

- department_id: department identifier
- department: the name of the department

order_products__prior (30m+ rows):

- order_id: foreign key
- product_id: foreign key
- add_to_cart_order: order in which each product was added to cart
- reordered: 1 if this product has been ordered by this user in the past, 0 otherwise

customers (206k+ rows):

- user_id: foreign key
- First Name: customer first name
- Surnam: customer last name
- Gender: customer gender
- STATE: customer US state of residence
- Age: customer age in years
- date_joined: date customer joined
- n_dependants: number of dependents reported by customer
- fam_status: values are married, single, divorced/widowed, living with parent
- income: self reported customer household income