

Overview

Introduction

Data preparation

Exploratory analysis

Model deployment

Findings & Suggestions

Potential Improvements

Predict Customer Success

- Focus:
 - B2B, local restaurants in Utah
- Target Variables
 - Longevity ★
 - Total Gross Profit Dead Net ★
- Prediction Models
 - Patterns, trends, relationships

Sales Data

- Change data type
 - MAX & MIN_POSTING_DATE converted to Date
- Group customer sales records by CUSTOMER_NUMBER_BLIENDED
 - Modify MAX & MIN_POSTING_DATE
 - Extract total values (GROSS_PROFIT_DEAD NET, DISCOUNT, INVOICE_PRICE, etc.)

Customer Data

- Clean data
 - Simplify ADDRESS_ZIP_CODE to 5 digits
- Filter data
 - Utah only
 - Include only “Eating & Drinking” customers from CUSTOMER_ACTIVITY_CLUSTER
 - Include only “DSD” customers from BUSINESS_TYPE_EXTENSION_DESCRIPTION

American Community Survey (2017-2021)

- Population
- Household income
- Household size
- Year structure built
- Housing costs
- Rent
- Selected owner costs
- Number of cars used to commute
- Number of households without internet

Customer Longevity Calculation

**Maximum
Posting Date**

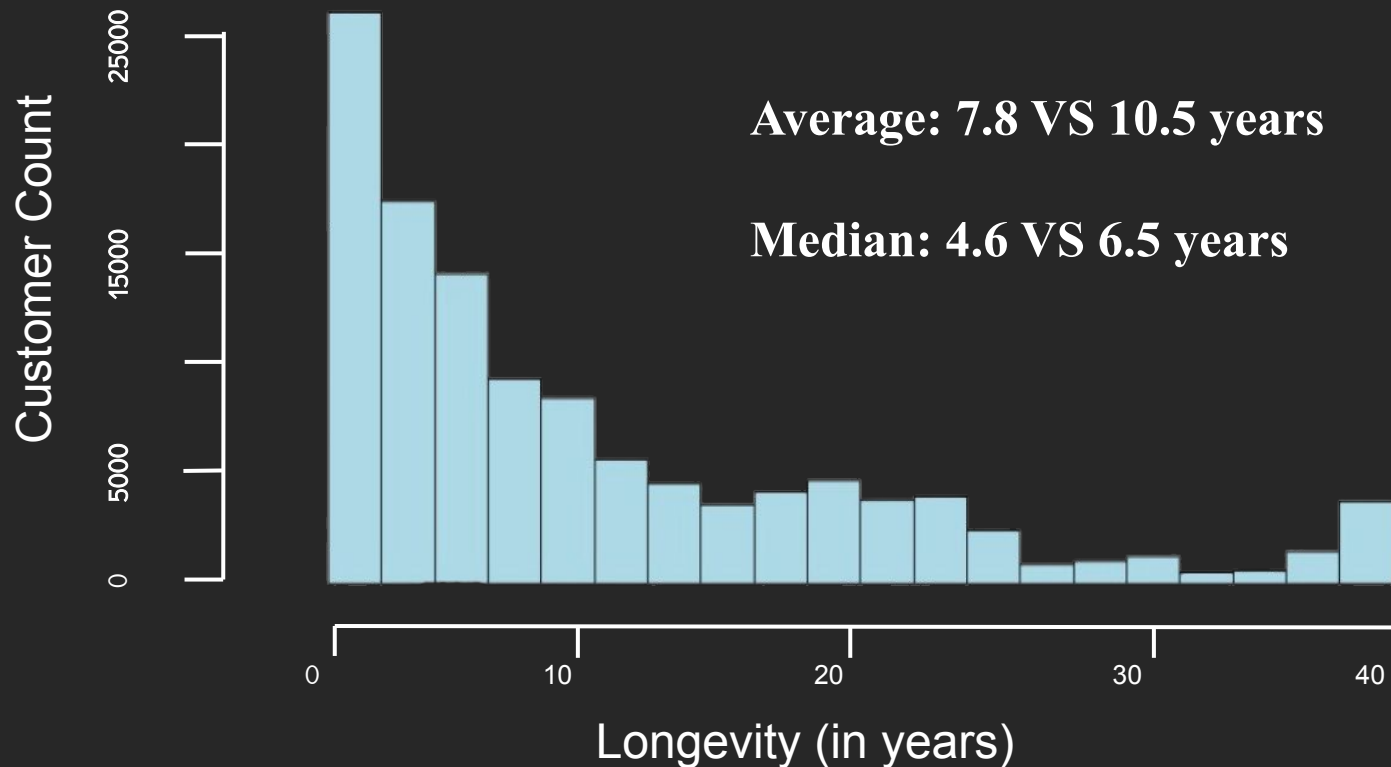
—

**Onboarding
Date**

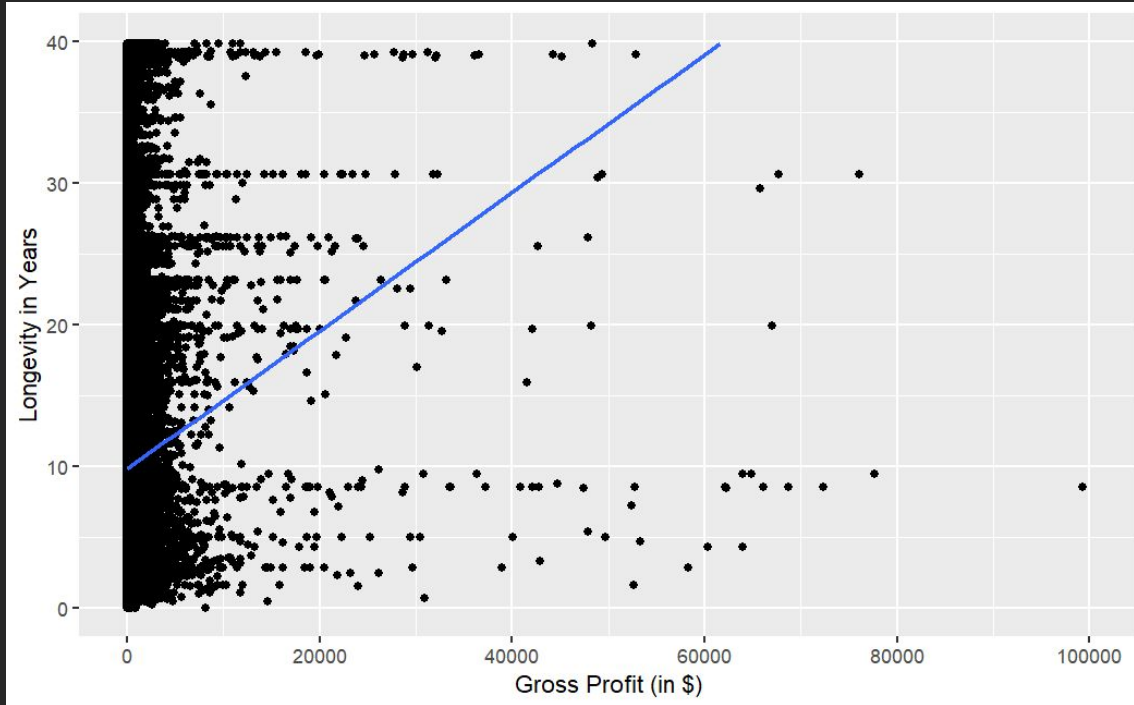
=

**Customer
Longevity**

Customer Longevity Distribution



Customer Longevity vs. Profit



The relationship between customer longevity and gross profit is not obvious by plotting

Customer Average Lifetime Value

\$19.6K

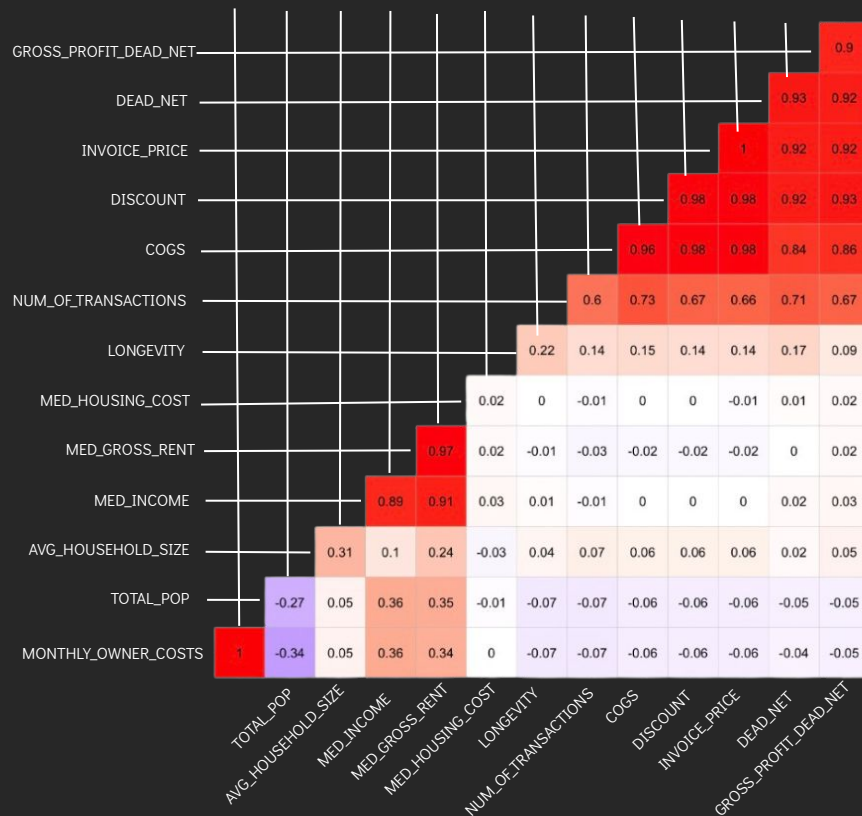
Revenue per customer
(2021 and 2022)

\$4.99K

Average customer longevity
7.85 years

$4.99 / 2 * 7.85 \approx 19.6k$

Numeric correlations



Customer longevity is not highly correlated with any other variables

Gross profit is highly correlated with dead net, invoice price, discount, etc.

Two models

Predict customer longevity

- Using support vector regression

Predict customer gross profit

- Using gradient boosting regression

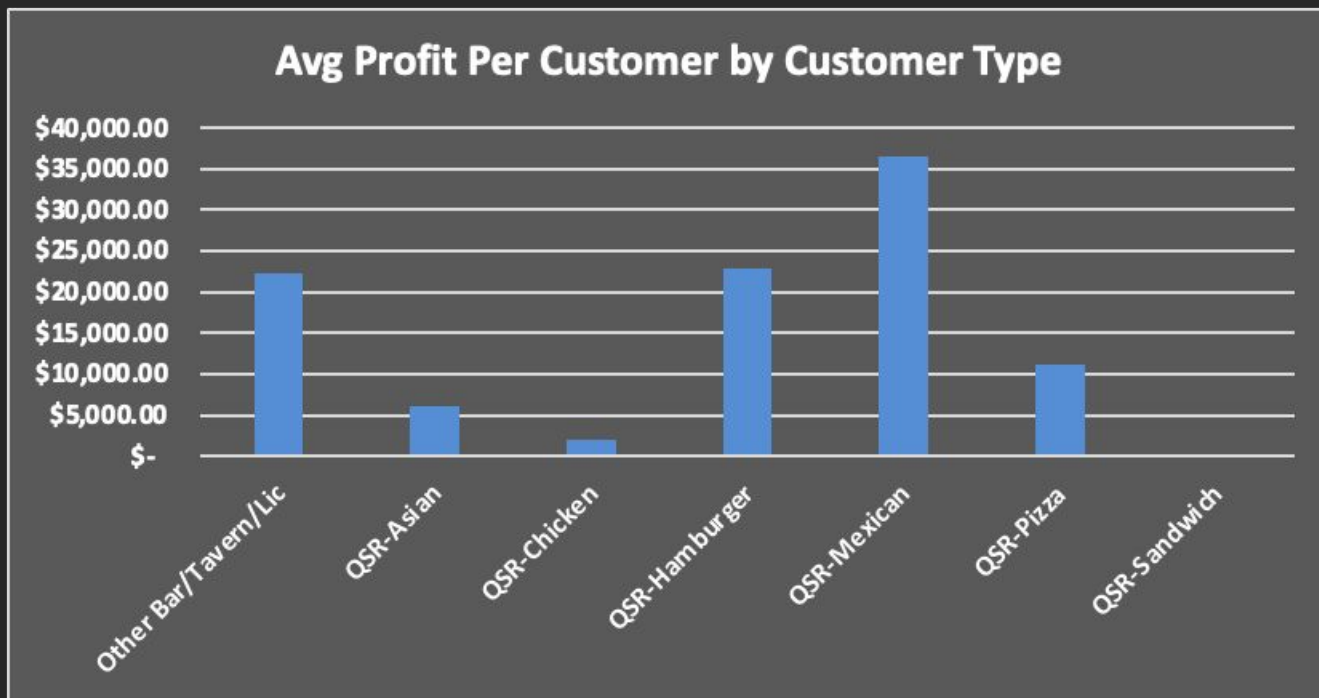
Predicting customer longevity

- Numeric prediction (years)
- Method generalizes well with highly dimensional data
- Simple and stable implementation
- Best fit $R^2 = 0.28$
- Needs improvement

Predicting customer gross profit

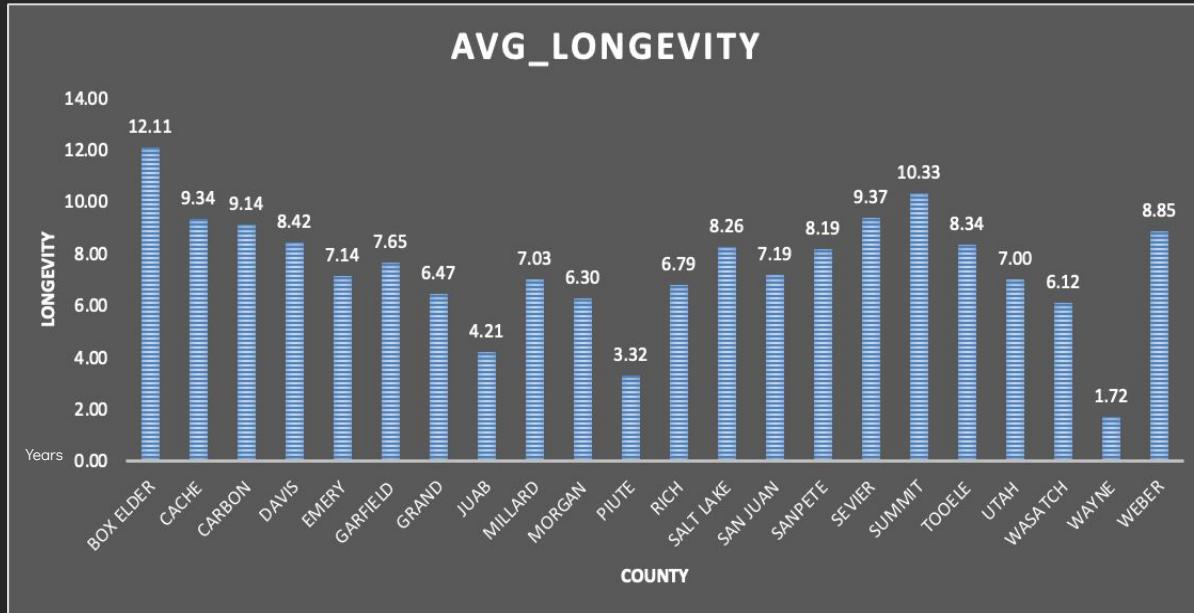
- Method captures complex patterns
- Feature selection
- Best fit $R^2 = 0.84$
- Includes discount as predictor

Findings & Suggestions



Provide promotions or discounts for customers with high average profit

Findings & Suggestions



1. BOX ELDER 12.11 years ★

2. SUMMIT 10.33 years ★

3. SEVIER 9.37 years

4. CACHE 9.34 years

5. CARBON 9.14 years

Ogden Sales Office

Improvements

- Find data google review (address/name...)
- Customer Clustering
- Different ACS variables

Questions?