

Pulse of Engagement

Visual Analytics for Economic Health in Engagement, OH

VAST Challenge 2022 – Challenge 3

Thomas Gantz Michal Sterzel Jan Marxen

December 2025



Introduction

VAST Challenge 2022 – Challenge 3

Introduction

The Challenge

- Analyze economic health of a fictional city
- Dataset: ~120 million data points
- 15 months of 5-minute granularity data

Three Questions

1. Business Prosperity
2. Resident Financial Health
3. Employer Health & Turnover

Economic Analysis

Q1: Business

Q2: Residents

Q3: Employers

Question 1: Business Prosperity



Dashboard Overview

Question 1: Business Prosperity

VAST Challenge 3: Economics Dashboard

Business Prosperity

Resident Financial Health

Employer Health & Turnover

Venue Analytics Dashboard

Restaurant & Pub Performance Intelligence

Analysis Period

Mar 1 — May 31, 2023

TOTAL REVENUE



\$7.27M

TOTAL VISITS



802.980

AVG. PER VISIT



\$9.06

ACTIVE VENUES



32

PROSPERING



9

STRUGGLING



23

Filters

Venue Type

All Types

Venue

All Venues

Customer

All Customers

Metric

Total Spending

Start Date

01.03.2022

End Date

31.05.2023

Sort Top N By

Total Spending

Top N Venues

32

Venue Type: ● Restaurant ● Pub

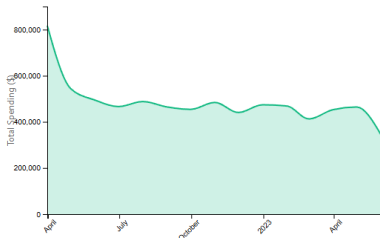
Growth Analysis

Question 1: Business Prosperity

Revenue & Traffic Trends

Check-ins and spending over time

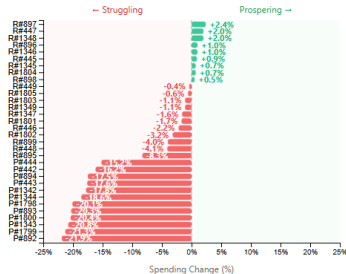
Resolution: Monthly



Business Health Analysis

Prospering vs struggling venues

Comparing spending: Mar 01 - Oct 12 vs Oct 12 - May 25 9 prospering | 23 struggling



Key Insight: Revenue drops in April · Business health is heterogeneous:
 ~1/3 growth ↑ ~1/3 slight decline ↓ ~1/3 significant decline ↓

Market Concentration

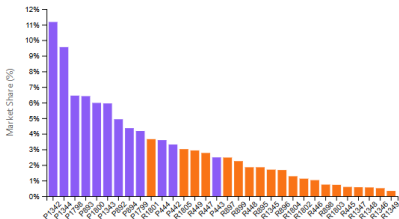
Question 1: Business Prosperity

Market Share Distribution

Revenue breakdown by venue

Chart: Bar Chart

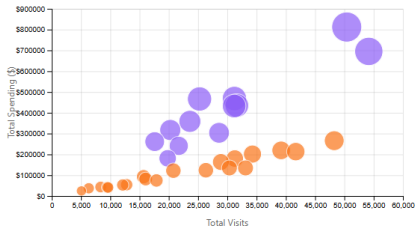
Total spending: \$7273740.82 | Showing top 32 venues



Performance Matrix

Venue comparison overview

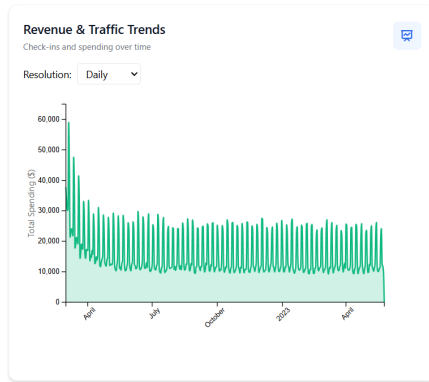
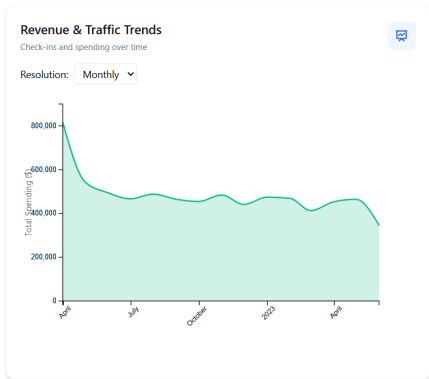
Showing top 32 venues | Bubble size = market share



Key Insight: Two pubs capture 20% of total spending · Pubs dominate restaurants

Temporal Trends

Question 1: Business Prosperity



💡 **Key Insight:** Weekend oscillation distinguishes cyclical from structural decline

Individual Customer Patterns

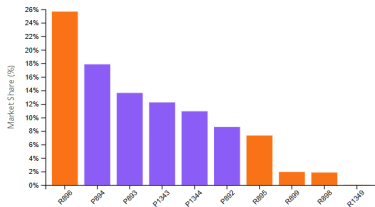
Question 1: Business Prosperity

Market Share Distribution

Revenue breakdown by venue

Chart: Bar Chart

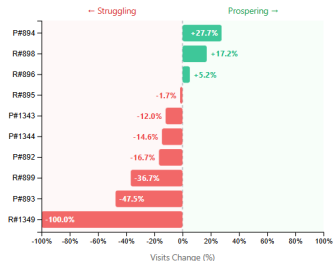
Total spending: \$14297.66 | Showing top 32 venues



Business Health Analysis

Prospering vs struggling venues

Comparing visits: Mar 01 - Oct 12 vs Oct 12 - May 24 3 prospering | 7 struggling



💡 Micro-level signals: ❤️ R#896: 26% share · 🦋 P#894: +27.7% growth
· 🚫 R#1349: abandoned



Key Findings

Question 1: Business Prosperity

Prosperous

- ✓ Pubs outperform restaurants
- ✓ P#1342, P#1344 dominate market




Struggling

- ✗ Top performers decline in H2
- ✗ ~1/3 show substantial drops







Overall: Aggregate spending declining over 15 months

Visualization Progression

-  **Overview** → establish baseline context
-  **Temporal filtering** → identify patterns over time
-  **Individual detail** → surface micro-level signals

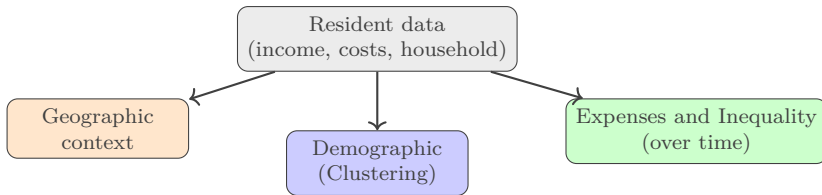
Key Design Decisions

-  **Coordinated views:** hover-linking for cross-chart exploration
-  **Split-period comparison:** quantifies growth directly
-  **Global filters:** all-to-all, one-to-all, one-to-one analysis
-  **Dual metrics:** visits and spending reveal correlation

Question 2: Resident Financial Health

Q2: Analysis Approach

Question 2: Resident Financial Health



- Building heatmap
- Savings by location
- Identify red zones

- Demographic, expense and salary features
- K-Means clustering
- Personas & drivers

- Expense dynamics over time
- Inequality trends (Gini)
- Income vs. expenses

Geographic Financial Health

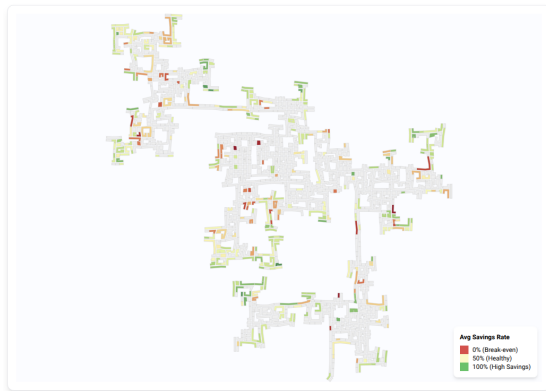
Question 2: Resident Financial Health

Building-Level Heatmap

- Colors by average savings rate
- Red: break-even or negative
- Yellow: moderate savings
- Green: high savings

Insights

- “Red pockets” persist over time
- Chronic, not worsening, conditions
- Mini-clusters suggest local stressors



Clustering Summary ($k = 3$)

Question 2: Resident Financial Health

Model Choice

- Elbow plot used to balance fit vs. complexity
- Smallest stable solution selected: $k = 3$ **clusters**
- Used all participant data (+ financial journal)

Affluent Achievers
High income, strong buffer

Stretched Households
Kids & tight budgets

Lean Savers
Moderate income, lower costs

How It's Used in the Dashboard

- **Cluster** filter applied to the PCP
- Same filter applied to the **Living Gap scatterplot**
- Global time slider

Resident Profile: Affluent Achievers

Question 2: Resident Financial Health

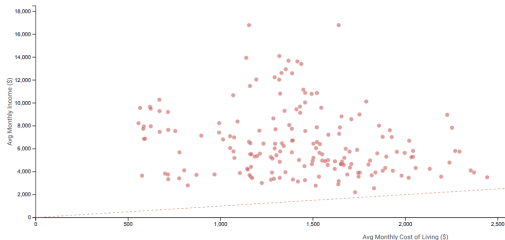
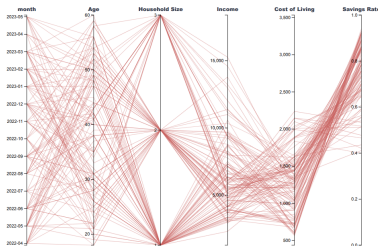
Affluent Achievers

Main Characteristics

- High income levels
- Predominantly graduate education
- Significant financial buffer

Median Statistics (Apr 2022)

- **Income:** \$5,756
- **Cost:** \$1,419
- **Savings:** 76.6%



Resident Profile: Stretched Households

Question 2: Resident Financial Health

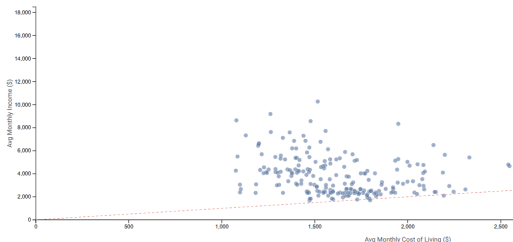
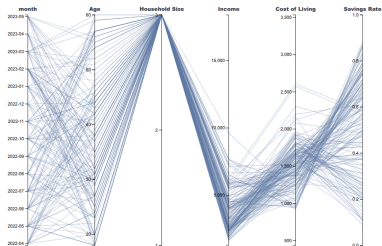
Stretched Households

Main Characteristics

- Larger households, often with children
- Tightest budget constraints
- "Living Gap" pressure is highest here

Median Statistics (Apr 2022)

- **Income:** \$2,869
- **Cost:** \$1,405
- **Savings:** 51.0%



Resident Profile: Lean Savers

Question 2: Resident Financial Health

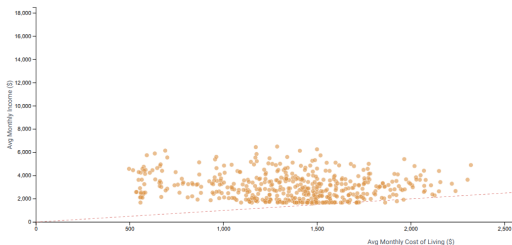
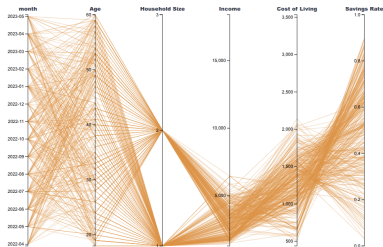
Lean Savers

Main Characteristics

- Smaller households
- Typically without children
- Moderate income, but lower costs than families

Median Statistics (Apr 2022)

- **Income:** \$3,352
- **Cost:** \$1,586
- **Savings:** 54.5%



What Drives Savings?

Question 2: Resident Financial Health

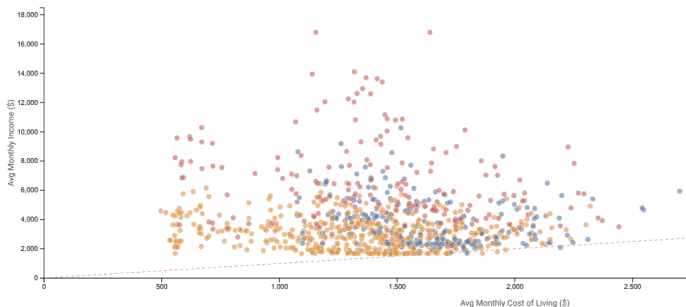
Demographic Drivers

Savings rate predictors (ΔR^2)

- Cost of living (0.828)
- Income (0.408)
- Household size (0.376)
- Has kids (0.127)

Cluster separators (η^2)

- Has kids (83.1%)
- Graduate education (72.0%)
- Household size (61.9%)
- Income (38.0%)



Inequalities Over Time

Question 2: Resident Financial Health

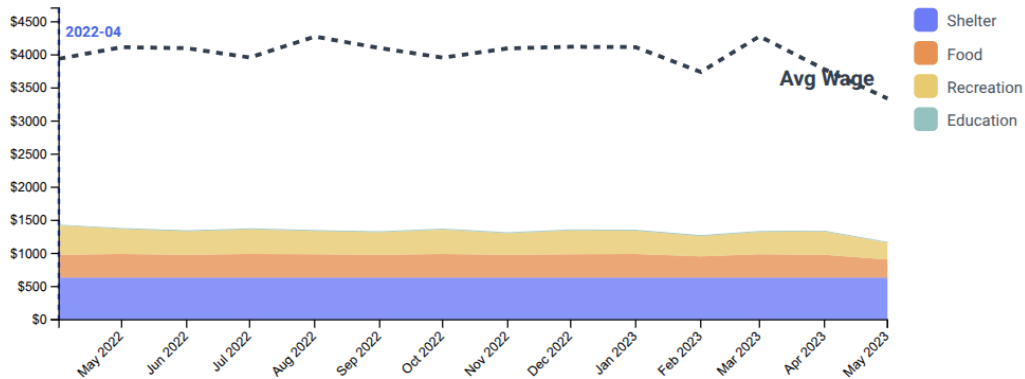


Inequality Trends

- Gini coefficient tracks disparity
- Income inequality stable over time
- Savings inequality slightly higher

Expense Dynamics Over Time

Question 2: Resident Financial Health



Question 3: Employer Health

Employer Health: Methodology & Approach

Question 3: Employer Health

Workforce Dynamics

- Turnover, hires, quits, net change
- Identifies extreme churn and instability
- Focus on employer-level change intensity

Stability & Retention

- Turnover vs. average tenure
- Headcount as contextual factor
- Distinguishes stable vs. high-risk employers

Mobility & Context

- Job-to-job flows between employers
- Geographic concentration of churn
- City-level employment and economic scale

Employer Turnover Ranking

Question 3: Employer Health

Ranking Methodology

- Ranks employers by turnover, hires, quits, or net change
- Focuses on upper tail of workforce churn
- Highlights extreme instability cases

Key Observation

- Some employers exceed 100% turnover
- Driven by small average headcounts
- Indicates intense, concentrated churn



Turnover vs. Workforce Tenure

Question 3: Employer Health

Inverse Relationship

- Clear separation into two regimes
- **Stability Zone:** low turnover, high tenure
- **Instability Zone:** high turnover, low tenure

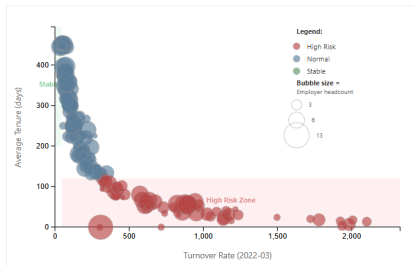
Bubble Size

- Encodes employer headcount
- Size alone does not guarantee stability
- High-risk large employers amplify impact

Employer Stability Overview

Employer Stability Overview

Multi-dimensional view: bubble size = headcount, position = turnover/tenure, color = stability category.



● **High Risk**
Turnover > 45%
Tenure < 120d

● **Normal**
Between thresholds

● **Stable**
Turnover < 45%
Tenure > 200d

Worker Mobility Between Employers

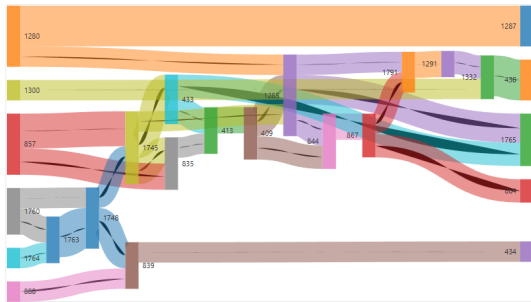
Question 3: Employer Health

Job-to-Job Flows

- Directional worker movement
- Some employers act as labor sources
- Others function as recipients

Network Effects

- Worker movement links employers
- Instability propagates through network
- Asymmetry reveals labor market structure



Short-Term Workforce Growth & Decline

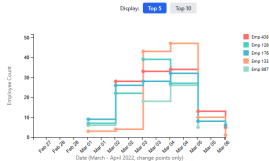
Question 3: Employer Health

Non-highlighted View

Employee Count Trends

Employee Count Trends

Tracks workforce size changes over time for selected employers (only days with changes shown). Click lines or legend to highlight specific employers independently. Compare growth and decline patterns.

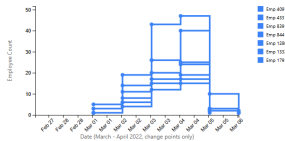


Highlighted View

Employee Count Trends

Employee Count Trends

Tracks workforce size changes over time for selected employers (only days with changes shown). Click lines or legend to highlight specific employers independently. Compare growth and decline patterns.



- Step-line chart: only days with changes
- Abrupt spikes and drops
- Reflects short-term, episodic workforce adjustments within the observed period

Geographic Concentration of Employer Instability

Question 3: Employer Health

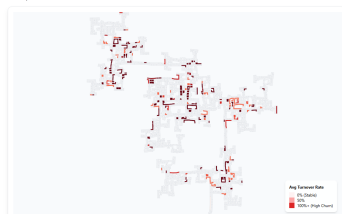
Building-Level Aggregation

- Turnover aggregated to building level
- Clear spatial hotspots visible
- Identifies localized instability clusters

Implications

- Workforce instability concentrates spatially
- Not uniformly distributed across city
- Suggests local economic stressors

Geographic Turnover Heatmap



CURRENT SELECTION

295.6%

City-wide Average Turnover

This map highlights "churn hotspots" across the city. Buildings are colored by the average turnover rate of employees located within them.

Dark Red areas indicate locations with high workforce instability. These may be commercial hubs with high turnover industries (e.g., retail, dining) or specific large employers facing retention issues.

Hover over any building to see specific turnover metrics and activity levels.

Design Decisions

Frontend

- **React 18:** Component architecture
- **D3.js v7:** Visualization rendering
- **TailwindCSS:** Styling
- **Axios:** API communication

Infrastructure

- **Docker Compose:** Orchestration
- **Nginx:** Reverse proxy

Backend

- **Python 3.11:** Core language
- **Flask:** REST API
- **Pandas/NumPy:** Data processing
- **Scikit-learn:** K-Means clustering
- **Pytest:** Testing



Team Organization

Work Organization

Team Organization

Division of Work

- One question per team member
- Shared infrastructure setup
- Code reviews via Git

Thomas	Q1: Business Prosperity
Jan	Q2: Resident Financial Health
Michal	Q3: Employer Health

Shared Components

- Docker infrastructure
- API structure
- Test framework

Communication

- Regular syncs and feedback
- Clear API contracts

Lessons Learned




Lessons Learned

Lessons Learned

What Worked Well

- ✓ Docker for reproducibility
- ✓ Clear question separation
- ✓ Caching for large datasets
- ✓ Test-driven development

Challenges

-  Large dataset with uneven distributions
-  Balancing detail vs. overview
-  Pre-processing for real-time usage

Would Do Differently

-  More upfront data profiling

Thank You!


Questions?

Thomas Gantz Jan Marxen Michal Sterzel

Q1: Business

Q2: Residents

Q3: Employers

 github.com/janmarxen/VAST-challenge

Data Visualization – EUMaster4HPC – December 2025