Optimizing Food Bank Resources in the DMV
Capital Area Food Bank

Client

- Serves 380,000 individuals across the DMV
- Largest Food Bank in the mid-Atlantic region
- Sources and distributes over 30 million meals annually

Opportunity

- Identify areas of most opportunity in the DMV area.
- Recommend census tracts in greatest need of CAFB resources.

Tools

- Python
- Statsmodels
- Pandas
- NumPy
- Folium

Data

DMV Hunger Heat Map

Processing

OLS Regression Model

<table>
<thead>
<tr>
<th>Target: Partner Count</th>
<th>Significant Predictors:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Food Insecure Rate</td>
<td>2019 F1 Rate, 2019 F1 Population, PublicTransportPercent, PovertyPercent, HealthInsurancePercent, HighSchoolEducationPercent</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Coeff</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>-2.364</td>
<td>0.002</td>
</tr>
<tr>
<td>-0.829</td>
<td>0.886</td>
</tr>
<tr>
<td>0.800</td>
<td>0.000</td>
</tr>
<tr>
<td>0.817</td>
<td>0.000</td>
</tr>
<tr>
<td>0.086</td>
<td>0.561</td>
</tr>
<tr>
<td>-0.436</td>
<td>0.685</td>
</tr>
</tbody>
</table>

Results

- 5 census tracts with significant negative residual values indicate areas where CAFB resources are needed
- 50 census tracts with significant positive residual values indicate areas where CAFB resources can be reallocated

Next Steps

- Identify partnerships with healthcare facilities
- Find opportunities for CAFB expansion in Federal Opportunity Zones
- Analyze the relationship between housing prices and density shifts in food insecure populations
- Retrieve recent and time-sensitive data

Special Thanks

- Jake Erlich, Project Champion
- Tim Le, Data Specialist at CAFB
- Joshua Kohn, Faculty Advisor
- Dr. Joseph Bailey, Course Instructor

DataDigestion

Sapna Bagalkotkar
Aaron Kurtz
Ben Lin
Natalie Wolfe
Compology:
- San Francisco-based startup providing dumpster sensors collecting data on waste production and removal statuses

Hypotheses:
- Fill level at pick up will **average about 50%**
- **Reduced recycling** in State of Emergency
- Denton dumpster will fill fastest

Key Insights:
- Gosset has highest refuse fill rates
- All buildings produce **more recycling** than refuse except Gosset and Denton
- Recycling **fill rates are inflated** by unfolded boxes in dumpsters

Weekly Collections by Location (Actual vs. Proposed)

<table>
<thead>
<tr>
<th>Location</th>
<th>Current</th>
<th>Proposed</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Denton Refuse</td>
<td>2</td>
</tr>
<tr>
<td>1</td>
<td>Denton Recycle</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>VMH Refuse</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>VMH Recycle</td>
<td>2</td>
</tr>
</tbody>
</table>

Data Visualization:
- Average Refuse vs. Recycle Fill Levels
- Average Fill Rates per Dumpster

Facilities Management:
- Manages and maintains University of Maryland’s campus, including refuse & recycling removal

Recommendations:
- Campaign to **break down boxes** in order to decrease rapid recycling fill
- **Optimize** collection routes based on average and maximum fill rates observed
- Host a **sustainability competition** between Freshmen dorms
- Install **additional Compology sensors** into high-traffic dumpsters

Impact:
- Reduce risk of over/underfilling dumpster
- **400 kg of CO₂ saved every semester** by reducing weekly collections of each dumpster by 1

Special Thanks:
- Joseph Bailey, Faculty Advisor | Jason Gates, Compology CEO
- William Guididas, Assistant Director of FM

Victoria Miske — Civil and Environmental Engineering
Ryan Geros — Mechanical Engineering
Amod Mathur — Computer Science
Daniel Eshel-On — Computer Science and Finance
Allie Strouse — Finance and OMBA
Flywheel Digital
“Driving Growth in Digital Retail”

Client and Project Introduction

Flywheel Digital uses customer search queries on Amazon to develop insights for manufacturers, in order to best approach brand, supply chain, and sales disciplines.

**PROJECT DATA & SCOPE:**

- **GIVEN:** search query data from Amazon.com in the form of 7 CSV (40GB of data)
- **GOAL:** Analyze trends within the flywheel dataframes and external data sources to predict trend patterns for specified products

Results

**Search Frequency x Smart Watches**

- Google Trends for Smart Watches JUN-JUL 2017
- The Google Trends results give a nuanced look into the data, preceding trend spikes in the search frequency pattern histograms

**Search Frequency x Kitchen Products**

Goals and Methodology

1. **CREATE A MANAGEABLE DATAFRAME**
   - SAMPLED 20% of the data provided
   - STANDARDIZED formatting across multiple sample CSV files
   - CLEANSED data

2. **ANALYZE TRENDS AMONG SIMILAR PRODUCTS**
   - CREATED AND COMPARED histograms for related product categories

3. **BENCHMARK WITH EXTERNAL DATA**
   - COMPARED identified trend patterns
   - CORRELATED predictions with Google Trends data

Recommendations & Next Steps

**SHORT TERM:**

- Implement the histograms and Google Trends into trend predicting models for products relating to Smart Watches and Kitchen Products
- Use current models to calculate inventory, estimate release dates, determine product “shelf life” for manufacturers

**LONG TERM:**

- Implement better natural language processing (NLP) to refine search terms and create more accurate trend models
- Use the team’s code and create trend prediction models for a number of different product categorizations

Special Thanks to:
- Project Champions: Christine Gambino and Sam Sussan
- Instructors: Dr. Joe Bailey, Josh Kohn, and Ekansh Vinaik
- Faculty Mentor: Josh Kohn

Lydia Runnals

Mansi Reddy

John Dutan

Ayo Awobajo
Openbarre is a barre studio that was opened in 2017 in the College Park Shopping Center by Q17 alum Lauren Filocco. Her studio has expanded tremendously since first opening and with this growth Lauren presented us with questions pertaining to her current strategy:

- Are classes currently optimally scheduled or should shifts be made?
- Are the current client recruitment methods effective and profitable?
- Is 2 hours notice a fair late cancellation? Should it be more?

The data received was split into attendance and cancellation history by client. Data clean-up included converting class times to dateTime objects, standardizing first name last name format, and removing hidden unicode characters.

First timers free class is more effective in retaining long term clients
Prioritize and increase marketing of first timers free class
Increase frequency of class as interest increases

Avoid Barre Cardio, Barre Stretch, Glutes and Thighs, and classes similar to Barre 101 between 5:00-6:00 pm (~3-4 attendance)
Schedule popular Barre and Arms and Abs between 5:30-6:00 pm

“I think that the 2 hours is fair. A stricter policy would make me less likely to sign up in advance. Everything usually depends on my work schedule so I might not sign up for the class until the last minute.” - Client

95% agreed the policy was fair
45% would sign up last minute if changed

After evaluating other studio policies and this data, Lauren should maintain her current protocol.
Improving Customer Retention Rates
Oscar Health

What is Oscar Health?
Oscar Health is a technology-focused health insurance company founded in 2012. Their main focus is on affordable and understandable healthcare through quality member experiences.

Opportunity
- Analyze if resolution time has an impact on retention
- Suggest prioritization for improvement on specific issue type
- Identify correlation between escalations and retention

Hypothesis
Faster response to claims and eligibility escalations leads to higher customer retention

Resolution Time
5 day difference → 4.618% decrease in customer retention
$1,560,000 loss in member premiums

Issue Prioritization
Eligibility Issues - 44.6% retention rate

Predictive Model
- Neural Network
- 58.60% Accuracy

1. Focus on resolving tickets faster to save money in the future
2. Allocate more resources towards resolving eligibility issues
3. Incorporate more customer-centric metrics into predictive model

Want Some Data?
Tania Arya, Kene Azinge, Simran Chertara, Srijay Kasturi, Jonathan Zhang

Special Thanks To:
Nate Elencweig, Dr. Kylie King, Dr. Joseph Bailey, Joshua Kohn
Improving Long-Term Forecast

Client & Objective

- PepsiCo Beverages North America (PBNA)
- Demand Planning Team
- Direct Beverage Delivery to Stores
- Weekly Product Forecasts

Process

1. Approach
   - Understand objective: Compare the PreVail and long term forecasts and when Pepsi should utilize each

2. Data Cleansing
   - Stripped data down to the necessary information: Removed null and zero values

3. Transformation
   - Made our numbers into information by calculating R Squared and MAPE (Mean Average Percent Error)

4. Modeling
   - Tried using open source tools like Facebook’s Prophet but it was too computationally intense
   - Created a per-item Linear Regression model as well as a more encompassing Ensemble model and tested both against existing forecasts

Data Analysis & Findings

- MSE analysis suggests that a simple linear regression outperforms a complicated ensemble model as well as Pepsi’s forecasts
- Plotting forecasts shows even more room for improvement

Recommendations

- Linear Combination of 2 Forecasts
  - As described in our data analysis, we found that the most optimal forecasting method is a linear combination of 2 of the given forecasts: OP and PreVail. While not all items produce more accurate forecasts with our method, many do, and so overall the linear combination optimizes the forecasting accuracy. We recommend an array of functions, one unique to each item, that will provide a more accurate forecast for that particular item. Each of these functions would be a variant on the linear combination we provide, tailored to optimize the forecast for that particular item.

Next Steps

- Brand Performance
  - Attach the inventory IDs to their package or brand to determine its performance based on the mean absolute percent error

- Update Forecast Comparison
  - Current data: Only includes one forecast for each item
  - New data: Includes all three forecasts for each item
  - The team will use the new data to better analyze the performance of each forecast on the different items.

Acknowledgements

- Clients: Elizabeth Gillum & Andrew Yerkes
- Project Champion: Dr. Pamela Armstrong
- Faculty Adviser: Dr. Joseph Bailey