EvoArm
- Small robot arm
- 3 degrees of freedom
- 3D printable
- Controlled with a Python App
- … where to go from here?

Customization!
- Every 3D-printed arm can be different
- Change mechanics for different purposes

Exploring Possible Configs
- Tedious and impractical to try many designs
- Different people need different capabilities
- Can the exploration process be made accessible?

Vis Idea
- Interactive exploration of design space
- Data: Calculated online
  - Reachable points
  - Max load (across reachable space)
  - Max velocity (across reachable space)
- Design:
  - Spatial data -> spatial display?
  - Derive attributes?
  - Combine certain parameters?

Vis can help!
- Want to rapidly iterate – new/different configs

Config IK Solution
Dimensions
Design constraints
Servo constraints

Want to rapidly iterate – new/different configs

PITCH:
VISUALIZING THE ENERGY PERFORMANCE OF A BUILDING
ARASH SHADKAM

WHAT
- ENERGY PERFORMANCE DATA OF A BUILDING (FOR NOW THE BUILDING IS THE CENTER FOR INTERACTIVE RESEARCH ON SUSTAINABILITY “CIRS”)
- TIME-SERIES DATA FROM SENSORS INCLUDING TEMPERATURE AND OCCUPANCY DATA (IF POSSIBLE)
- DERIVED: NORMALIZED ENERGY PERFORMANCE DATA

WHY
- BETTER UNDERSTANDING OF THE BUILDING’S ENERGY PERFORMANCE
- DISCOVERING TRENDS AND CORRELATIONS IN THE ENERGY PERFORMANCE DATA AND IDENTIFY POTENTIAL OPTIMIZATION OPPORTUNITIES IN THE BUILDING’S PERFORMANCE

HOW
- FACETS: MULTI-FORM OVERVIEW-DETAIL VIEWS/LINKED HIGHLIGHTING
- MANIPULATE: SELECT
- REDUCE: FILTER/RANGE SLIDERS FOR DIFFERENT TIME SPANS

THANKS!
A VISUALIZATION TOOL FOR COMPUTER PROGRAM PERFORMANCE DEBUGGING

Augustine Wong

WHAT IS COMPUTER PROGRAM PERFORMANCE DEBUGGING?

Diagnosing why a computer program is running slowly

MY PC IS SLOW

HOW DO VISUALIZATION TOOLS HELP?

Let's look at an existing visualization tool...

PROJECT OBJECTIVES

Create a visualization tool which:

- Uses the “search, show context, expand on demand” approach

Visualizes “patterns” of computer program behavior

Evaluates which patterns are good starting points for initially exploring the computer program

WHAT IS COMPUTER PROGRAM PERFORMANCE DEBUGGING?

Diagnosing why a computer program is running slowly

Let's look at an existing visualization tool...

HOW DO VISUALIZATION TOOLS HELP?

Visualizations For Justifying Machine Learning Predictions

David Johnson

Visualizations For Justifying Machine Learning Predictions

David Johnson

Motivation

Strengths of ML allowed expansion to diverse fields

Fields and contexts far removed from traditional ML

Users not trained in ML

PROJECT OBJECTIVES

Create a visualization tool which:

- Uses the “search, show context, expand on demand” approach

Visualizes “patterns” of computer program behavior

Evaluates which patterns are good starting points for initially exploring the computer program

WHAT IS COMPUTER PROGRAM PERFORMANCE DEBUGGING?

Diagnosing why a computer program is running slowly

Let's look at an existing visualization tool...

HOW DO VISUALIZATION TOOLS HELP?

Visualizations For Justifying Machine Learning Predictions

David Johnson

Visualizations For Justifying Machine Learning Predictions

David Johnson

Motivation

Strengths of ML allowed expansion to diverse fields

Fields and contexts far removed from traditional ML

Users not trained in ML

PROJECT OBJECTIVES

Create a visualization tool which:

- Uses the “search, show context, expand on demand” approach

Visualizes “patterns” of computer program behavior

Evaluates which patterns are good starting points for initially exploring the computer program

WHAT IS COMPUTER PROGRAM PERFORMANCE DEBUGGING?

Diagnosing why a computer program is running slowly

Let's look at an existing visualization tool...

HOW DO VISUALIZATION TOOLS HELP?

Visualizations For Justifying Machine Learning Predictions

David Johnson

Visualizations For Justifying Machine Learning Predictions

David Johnson

Motivation

Strengths of ML allowed expansion to diverse fields

Fields and contexts far removed from traditional ML

Users not trained in ML

PROJECT OBJECTIVES

Create a visualization tool which:

- Uses the “search, show context, expand on demand” approach

Visualizes “patterns” of computer program behavior

Evaluates which patterns are good starting points for initially exploring the computer program

WHAT IS COMPUTER PROGRAM PERFORMANCE DEBUGGING?

Diagnosing why a computer program is running slowly

Let's look at an existing visualization tool...

HOW DO VISUALIZATION TOOLS HELP?

Visualizations For Justifying Machine Learning Predictions

David Johnson

Visualizations For Justifying Machine Learning Predictions

David Johnson

Motivation

Strengths of ML allowed expansion to diverse fields

Fields and contexts far removed from traditional ML

Users not trained in ML

PROJECT OBJECTIVES

Create a visualization tool which:

- Uses the “search, show context, expand on demand” approach

Visualizes “patterns” of computer program behavior

Evaluates which patterns are good starting points for initially exploring the computer program

WHAT IS COMPUTER PROGRAM PERFORMANCE DEBUGGING?

Diagnosing why a computer program is running slowly

Let's look at an existing visualization tool...

HOW DO VISUALIZATION TOOLS HELP?

Visualizations For Justifying Machine Learning Predictions

David Johnson

Visualizations For Justifying Machine Learning Predictions

David Johnson

Motivation

Strengths of ML allowed expansion to diverse fields

Fields and contexts far removed from traditional ML

Users not trained in ML

PROJECT OBJECTIVES

Create a visualization tool which:

- Uses the “search, show context, expand on demand” approach

Visualizes “patterns” of computer program behavior

Evaluates which patterns are good starting points for initially exploring the computer program

WHAT IS COMPUTER PROGRAM PERFORMANCE DEBUGGING?

Diagnosing why a computer program is running slowly

Let's look at an existing visualization tool...

HOW DO VISUALIZATION TOOLS HELP?

Visualizations For Justifying Machine Learning Predictions

David Johnson

Visualizations For Justifying Machine Learning Predictions

David Johnson

Motivation

Strengths of ML allowed expansion to diverse fields

Fields and contexts far removed from traditional ML

Users not trained in ML

PROJECT OBJECTIVES

Create a visualization tool which:

- Uses the “search, show context, expand on demand” approach

Visualizes “patterns” of computer program behavior

Evaluates which patterns are good starting points for initially exploring the computer program
Who is Scandio?

- 2016:
  - 40 employees
  - 82 clients
  - 176 projects

Projects:
- Fixed price ("client pays what's estimated")
- Time and material ("client pays the hours")

What is a fixed price project at Scandio?

- Efforts range from 5 days - 100 days
- Duration ranges from 3 weeks – 1 year
- Before project starts: effort estimation
- Generally higher risk of "failure"
  - If over estimation in the end, company mostly has to pay (sometimes compromises with client)

What are the project results?

- Total amount of efforts in the end
  - Exactly as estimated (rare)
  - Less than estimated (sometimes) 😊
  - More than estimated (sometimes) 😞

What are the key attributes?

1. Hours worked
   - Employees track time on project in web app
2. Degree of completion (DOC)
   - Estimated monthly by project lead
3. Hourly rate for project
   - Determined in the beginning dependent on budget and total effort
   - Changes retrospectively depending on 1 and 2

What are the project results?

- Total amount of efforts in the end
  - Exactly as estimated (rare)
  - Less than estimated (sometimes)
  - More than estimated (sometimes)

Time Tracking

Project results (good)

Project results (bad)

Visualizing Internal Components of a Convolutional Neural Network

Mahdi Ghodsi - Hooman Shariati

Convolutional Neural Network

- The idea has been around since 1980s
- But Introduction of GPU computing with 30x speed up gave DNNs a boost
- Very Popular Research Area

Background:

- What is Machine Learning
- Machine Learning is taking over.
- Applied to many fields: Bioinformatics, Gaming, Medical diagnosis, Marketing, Machine Vision, ….
- Convolutional Neural Network
  - ImageNet Competition
  - Google Deep Dream

However ...
**Convolutional Neural Network**

How researchers see CNNs

Convolutional Neural Network

"Neural networks have long been known as "black boxes" because it is difficult to understand exactly how any particular neural network functions due to the large number of interacting, non-linear parts."

Yajin Zhou

Department of Computer Science North Carolina State University

**Motivation**

- Understanding the behaviour of distributed systems is hard
- Developers need tools for comprehending their systems
- Most distributed systems are designed around FSM
- FSMs are often how developers think of their systems
- Can an FSM be generated from an execution so developers can check their mental models?

**Concept**

- Collect distributed snapshots (state from across the whole system)
- Calculate a distance between each snapshot (xor distance)
- Plot each snapshot at its relative distance using clustering
- Connect each snapshot with time curve

**Visualizing and Understanding Convolutional Networks**

Visualizing and making sense of CNNs in literature:

- Collect distributed snapshots (state from across the whole system)
- Calculate a distance between each snapshot (xor distance)
- Plot each snapshot at its relative distance using clustering
- Connect each snapshot with time curve

**Alternatives to Error Bars**

- Violin Plots
- Box Plots
- Dynamic Icon Arrays
- Gradient Plots

**Case Scenario**

- Imagine you are Betty
- Just finished chemo for breast cancer
- Typical post-chemo therapy is Tamoxifen for 5 years

**Tamoxifen 10-year risk estimates compared to 5-year risk estimates (out of 1000)**

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Breast cancer recurrence</td>
<td>-28</td>
</tr>
<tr>
<td>Death from breast cancer</td>
<td>-28</td>
</tr>
<tr>
<td>Development of gallstones</td>
<td>+2</td>
</tr>
<tr>
<td>Development of endometrial cancer</td>
<td>+16</td>
</tr>
<tr>
<td>Stroke</td>
<td>+2</td>
</tr>
</tbody>
</table>

**Project**

- Design new visualization to present ambiguity to patients
- Interactivity
  - Adjust bounds of error
  - Show best & worst case scenarios
  - Show how risk estimates might change given different samples

**Visualizing and Understanding Convolutional Networks**

Visualizing and making sense of CNNs in literature:

- Collect distributed snapshots (state from across the whole system)
- Calculate a distance between each snapshot (xor distance)
- Plot each snapshot at its relative distance using clustering
- Connect each snapshot with time curve

**Visualizing and Making Sense of CNNs**

Visualizing and understanding Convolutional Networks (By M. Zeiler NYU)

**With confidence intervals**

5-year vs. 10-year Tamoxifen Therapy

**Tamoxifen 10-year risk estimates compared to 5-year risk estimates (out of 1000)**

- Breast cancer recurrence: -28
- Death from breast cancer: -28
- Development of gallstones: +2
- Development of endometrial cancer: +16
- Stroke: +2

**Project**

- Design new visualization to present ambiguity to patients
- Interactivity
  - Adjust bounds of error
  - Show best & worst case scenarios
  - Show how risk estimates might change given different samples

**Dviz**

Visualizing Distributed Systems with Stewart Grant and Jodi Spacek

**Motivation**

- Understanding the behaviour of distributed systems is hard
- Developers need tools for comprehending their systems
- Most distributed systems are designed around FSM
- FSMs are often how developers think of their systems
- Can an FSM be generated from an execution so developers can check their mental models?

**Concept**

- Collect distributed snapshots (state from across the whole system)
- Calculate a distance between each snapshot (xor distance)
- Plot each snapshot at its relative distance using clustering
- Connect each snapshot with a time curve

**etcd (distributed key value store) puts -> gets**

**Limitations**

- States are not labeled meaningfully
- Semantics of state transitions are not clear
- FSMs require both

**Visualizing Ambiguity**

James Hicklin

**With confidence intervals**

5-year vs. 10-year Tamoxifen Therapy

**Alternatives to Error Bars**

- Violin Plots
- Box Plots
- Dynamic Icon Arrays
- Gradient Plots

**Case Scenario**

- Imagine you are Betty
- Just finished chemo for breast cancer
- Typical post-chemo therapy is Tamoxifen for 5 years

**Tamoxifen 10-year risk estimates compared to 5-year risk estimates (out of 1000)**

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Breast cancer recurrence</td>
<td>-28</td>
</tr>
<tr>
<td>Death from breast cancer</td>
<td>-28</td>
</tr>
<tr>
<td>Development of gallstones</td>
<td>+2</td>
</tr>
<tr>
<td>Development of endometrial cancer</td>
<td>+16</td>
</tr>
<tr>
<td>Stroke</td>
<td>+2</td>
</tr>
</tbody>
</table>

**Project**

- Design new visualization to present ambiguity to patients
- Interactivity
  - Adjust bounds of error
  - Show best & worst case scenarios
  - Show how risk estimates might change given different samples

**Dviz**

Visualizing Distributed Systems with Stewart Grant and Jodi Spacek

**Motivation**

- Understanding the behaviour of distributed systems is hard
- Developers need tools for comprehending their systems
- Most distributed systems are designed around FSM
- FSMs are often how developers think of their systems
- Can an FSM be generated from an execution so developers can check their mental models?

**Concept**

- Collect distributed snapshots (state from across the whole system)
- Calculate a distance between each snapshot (xor distance)
- Plot each snapshot at its relative distance using clustering
- Connect each snapshot with a time curve

**etcd (distributed key value store) puts -> gets**

**Limitations**

- States are not labeled meaningfully
- Semantics of state transitions are not clear
- FSMs require both
**Background**
- Project originated as an MA thesis in the CS department
- New technique that applied low-stretch trees to network visualization
- Implemented novel edge-bundling technique
- Does not rely on fixed vertices/fixed layout or explicit hierarchical data structure

**Problem**
Physician researchers are often interested in data exploration before committing to a project.

*Generally use descriptive statistics to see if there are any obvious signals.*

Is there any specific group of patients that have the worse outcome compared to the rest?
Are there natural groupings in the dataset?
Is there an underlying structure to the data?

**Proposed solution**
Cluster visualization
Use dimensionality reduction methods such as t-sne.
Plot resulting clusters.
Draw survival plots by cluster membership.
Allow investigation of cluster membership.

**Research Questions**
- Scatterplots? Occlusion? Continuous scatterplots?
- Interaction?
- Spatial aggregation? Does it make sense?
- Dimensionality reduction? Too much information?
- Effective color coding?
- Dimensional Ordering, Spacing, and Filtering Approach (DOSFA)? Similarities show patterns?
Background
- Two iterations submitted for publication:
  1. Graph Drawing (technique focused)
  2. Pacific Vis (more emphasis on motivation and visualization application)
- Both rejected 😞
- Reviewer comments largely yearning for a deeper/more defined motivation

Resurrection Pitch
- Find the motivation!
- Develop and execute a user study
- Revise and resubmit paper

Why?
- Lots of potential!
- De-hairball a cluttered network:
- Novel, layout free network idioms:

Next Steps
- Complete literature review of network idioms, tasks and taxonomies
- Brainstorm new cases where “set” or intuitive network layout is not optimal or necessary for a given task

Questions?

Automatic Grading Service Dataset
Nick Bradley
nbrad11@cs.ubc.ca

Background
Continuous grading service
5.5 GB from 13K test result records (more coming everyday)
Some data fields (don’t worry if these don’t mean anything to you)
◦ Grade for every commit each student made
◦ Test metrics: # tests pass/fail, coverage, duration
◦ Code metrics: LOC, build failures
◦ Grade requests: timestamp
◦ More data can be pulled from GitHub (diffs, history, branches,…)

Current Instructor Dashboard
Current Operational Dashboard

Idea + Impact
Student-facing dashboard:
- Expanded to CS110, CS210, and CS310 + their corresponding MOOC offerings
- User interface will be used by 1000s of students in the production system
- Instructor-facing dashboard:
- Design study with domain expert (current CPSC310 instructor)
- Challenge: need to scale to 1000s of students

Analysis tool:
- Probably only if you are interested in software engineering
- Likely end up as a SE paper

nbrad11@cs.ubc.ca
Visual Methods for Analyzing Motifs in Time-Oriented Data
Soheil Kianzad
PhD student CS

Stock technical analysis
Yuan Li , GrammarViz, 2012

ViSoccer
Visualizing European soccer players
Yann Dubois

ViSccer
Visualizing European soccer players
Yann Dubois

Why?
Other sports
By region
World cup
By game
Why?
Other sports
By region
World cup
By game

What?
+25,000 matches
+10,000 players
11 European leagues
Players and Teams' attributes
Detailed match events
Betting odds

+ sports page scrapping

How?
D3
P5.js
Tableau

WHO ARE WE?
- Responsible for academic oversight and support for approx. 300 graduate degree programs
- Strategic leaders in graduate education at UBC
- Support for faculty, programs & students
- Contact hub for everything related to graduate students
  - Communications & Recruitment
  - Admission
  - Awards
  - Theses & Dissertations
  - Doctoral Exams
  - Professional Development
- Approx. 10,000 graduate students in Vancouver

DATA PROJECTS
- Option 1: Canadian Graduate & Professional Student Survey (CGPSS)
  - Satisfaction levels in 13 sections, e.g. general, PD, research experience, financial support, social life
  - Breakdown by discipline, year of study, degree level, gender, etc.
- Option 2: Graduate School data
  - Application data
  - Enrolment statistics
  - Graduation statistics
  - Time in program and completion rates

Desired Outcomes:
1. Visualize key findings from 2016 study
2. Time comparison: 2010 to 2013 to 2016
3. Benchmarking: program vs. UBC vs. Canada

Audiences:
- Students
- Units (access controlled), e.g. program or department dashboard
- Department Head
- Program Director
- Faculty

CGPSS

<table>
<thead>
<tr>
<th></th>
<th>2010</th>
<th>2013</th>
<th>2016</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Doctoral</td>
<td>13 812</td>
<td>18 377</td>
<td>18 822</td>
<td>51 011</td>
</tr>
<tr>
<td>Research Masters</td>
<td>13 593</td>
<td>17 546</td>
<td>18 086</td>
<td>49 225</td>
</tr>
<tr>
<td>Other Masters</td>
<td>11 213</td>
<td>15 741</td>
<td>16 834</td>
<td>43 788</td>
</tr>
<tr>
<td>Total</td>
<td>38 618</td>
<td>51 664</td>
<td>53 742</td>
<td>144 024</td>
</tr>
</tbody>
</table>

univ | 38 | 48 | 50

TEAM
Louise Mol
Systems and Data Analysis Manager

Jens Locher
Assistant Dean

GRADUATE STUDIES – DATA VISUALIZATIONS
JENS LOCHER, ASSISTANT DEAN – STRATEGIC TECHNOLOGIES AND BUSINESS INITIATIVES

2

WHO ARE WE?
- Responsible for academic oversight and support for approx. 300 graduate degree programs
- Strategic leaders in graduate education at UBC
- Support for faculty, programs & students
- Contact hub for everything related to graduate students
  - Communications & Recruitment
  - Admission
  - Awards
  - Theses & Dissertations
  - Doctoral Exams
  - Professional Development
- Approx. 10,000 graduate students in Vancouver

DATA PROJECTS
- Option 1: Canadian Graduate & Professional Student Survey (CGPSS)
  - Satisfaction levels in 13 sections, e.g. general, PD, research experience, financial support, social life
  - Breakdown by discipline, year of study, degree level, gender, etc.
- Option 2: Graduate School data
  - Application data
  - Enrolment statistics
  - Graduation statistics
  - Time in program and completion rates

Desired Outcomes:
1. Visualize key findings from 2016 study
2. Time comparison: 2010 to 2013 to 2016
3. Benchmarking: program vs. UBC vs. Canada

Audiences:
- Students
- Units (access controlled), e.g. program or department dashboard
  - Department Head
  - Program Director
  - Faculty

CGPSS

<table>
<thead>
<tr>
<th></th>
<th>2010</th>
<th>2013</th>
<th>2016</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Doctoral</td>
<td>13 812</td>
<td>18 377</td>
<td>18 822</td>
<td>51 011</td>
</tr>
<tr>
<td>Research Masters</td>
<td>13 593</td>
<td>17 546</td>
<td>18 086</td>
<td>49 225</td>
</tr>
<tr>
<td>Other Masters</td>
<td>11 213</td>
<td>15 741</td>
<td>16 834</td>
<td>43 788</td>
</tr>
<tr>
<td>Total</td>
<td>38 618</td>
<td>51 664</td>
<td>53 742</td>
<td>144 024</td>
</tr>
</tbody>
</table>

univ | 38 | 48 | 50

TEAM

2

WHO ARE WE?
- Responsible for academic oversight and support for approx. 300 graduate degree programs
- Strategic leaders in graduate education at UBC
- Support for faculty, programs & students
- Contact hub for everything related to graduate students
  - Communications & Recruitment
  - Admission
  - Awards
  - Theses & Dissertations
  - Doctoral Exams
  - Professional Development
- Approx. 10,000 graduate students in Vancouver

DATA PROJECTS
- Option 1: Canadian Graduate & Professional Student Survey (CGPSS)
  - Satisfaction levels in 13 sections, e.g. general, PD, research experience, financial support, social life
  - Breakdown by discipline, year of study, degree level, gender, etc.
- Option 2: Graduate School data
  - Application data
  - Enrolment statistics
  - Graduation statistics
  - Time in program and completion rates

Desired Outcomes:
1. Visualize key findings from 2016 study
2. Time comparison: 2010 to 2013 to 2016
3. Benchmarking: program vs. UBC vs. Canada

Audiences:
- Students
- Units (access controlled), e.g. program or department dashboard
  - Department Head
  - Program Director
  - Faculty

CGPSS

<table>
<thead>
<tr>
<th></th>
<th>2010</th>
<th>2013</th>
<th>2016</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Doctoral</td>
<td>13 812</td>
<td>18 377</td>
<td>18 822</td>
<td>51 011</td>
</tr>
<tr>
<td>Research Masters</td>
<td>13 593</td>
<td>17 546</td>
<td>18 086</td>
<td>49 225</td>
</tr>
<tr>
<td>Other Masters</td>
<td>11 213</td>
<td>15 741</td>
<td>16 834</td>
<td>43 788</td>
</tr>
<tr>
<td>Total</td>
<td>38 618</td>
<td>51 664</td>
<td>53 742</td>
<td>144 024</td>
</tr>
</tbody>
</table>

univ | 38 | 48 | 50

TEAM

Louise Mol
Systems and Data Analysis Manager

Jens Locher
Assistant Dean

GRADUATE STUDIES – DATA VISUALIZATIONS
JENS LOCHER, ASSISTANT DEAN – STRATEGIC TECHNOLOGIES AND BUSINESS INITIATIVES

2

WHO ARE WE?
- Responsible for academic oversight and support for approx. 300 graduate degree programs
- Strategic leaders in graduate education at UBC
- Support for faculty, programs & students
- Contact hub for everything related to graduate students
  - Communications & Recruitment
  - Admission
  - Awards
  - Theses & Dissertations
  - Doctoral Exams
  - Professional Development
- Approx. 10,000 graduate students in Vancouver

DATA PROJECTS
- Option 1: Canadian Graduate & Professional Student Survey (CGPSS)
  - Satisfaction levels in 13 sections, e.g. general, PD, research experience, financial support, social life
  - Breakdown by discipline, year of study, degree level, gender, etc.
- Option 2: Graduate School data
  - Application data
  - Enrolment statistics
  - Graduation statistics
  - Time in program and completion rates

Desired Outcomes:
1. Visualize key findings from 2016 study
2. Time comparison: 2010 to 2013 to 2016
3. Benchmarking: program vs. UBC vs. Canada

Audiences:
- Students
- Units (access controlled), e.g. program or department dashboard
  - Department Head
  - Program Director
  - Faculty
**Who are we?**

Q.I. Leap Analytics
- Team of data scientists
- Solutions for retail stores
- 2 products
  - Recommender System
  - Interactive Dashboard

---

**What is a recommender system?**

**What’s the visualization task?**

End user: Business that is using the Recommender System

End user desires:
- Which items recommended
- Trends in item recommendations
- Cluster users with similar purchase history
- Cluster items with similar buying history

---

**What kind of data would you have to work with?**

Transaction data for online store
- 50,000 transactions
- 2,000 unique items
- 13,000 unique customers
- With time, date, city of purchase

Generated recommendation data
- Customer, item viewing history, top 10 recommended items (with scores)

---

**Benefits beyond the classroom**

- Implemented in our dashboard product so customers would get to see how their recommender system is being used
- Possibility of internship on completion of project
- Talk to me afterwards if interested in the project!

lauren.fratamico@qileap.com