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
Dimensions
Design constraints
Servo constraints
Dimensions
Design constraints
Servo constraints

Config
IK
Solution
Exploring Possible Configs

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

Changing one of dozens of parameters
Want to rapidly iterate - new/different configs

- Config
- IK
- Solution

- Dimensions
- Design constraints
- Servo constraints
Vis can help!
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?
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
HOW

- FACET: 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
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
Quantum Annealing Visualization

Austin Wallace
5th year undergraduate student
Integrated Science-Machine Learning

Chimera Graph

1QBit
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
Motivation

Biggest factor for users is understanding how predictions occur

Particularly important in\textsuperscript{1}:

- High risk applications like medicine
- Consumer-facing applications such as Recommender Systems
- Context-Aware applications

\textsuperscript{1}Biran, McKeown. 2014. Justification Narratives For Individual Classification
Justification Visualization

Visualizations present important evidence for a prediction

Intensions are to tie in to thesis work
Yelp Visualization Tool

Dilan Ustek
Matthew Chun
Motivation

- Target User: Yelp end-users
- Comparing businesses
- Filtered visualization
The Dataset

https://www.yelp.com/dataset_challenge/dataset

```json
yelp_academic_dataset_business.json
{
  "business_id": "encrypted business id",
  "name": "business name",
  "neighborhood": "hood name",
  "address": "full address",
  "city": "city",
  "state": "state -- if applicable --",
  "postal code": "postal code",
  "latitude": "latitude",
  "longitude": "longitude",
  "stars": "star rating, rounded to half-stars",
  "review_count": "number of reviews",
  "is_open": "0/1 [closed/open]",
  "attributes": "[an array of strings: each array element is an attribute]",
  "categories": "[an array of strings of business categories]",
  "hours": "[an array of strings of business hours]",
  "type": "business"
}
```

```json
yelp_academic_dataset_user.json
{
  "user_id": "encrypted user id",
  "name": "first name",
  "review_count": "number of reviews",
  "yelping_since": "date formatted like "2009-12-19"
  "friends": "[an array of encrypted ids of friends]",
  "useful": "number of useful votes sent by the user",
  "funny": "number of funny votes sent by the user",
  "cool": "number of cool votes sent by the user",
  "fans": "number of fans the user has",
  "elite": "[an array of years the user was elite]",
  "average_stars": "floating point average like 4.31",
  "compliment_hot": "number of hot compliments received by the user",
  "compliment_more": "number of more compliments received by the user",
  "compliment_profile": "number of profile compliments received by the user",
  "compliment_cute": "number of cute compliments received by the user",
  "compliment_list": "number of list compliments received by the user",
  "compliment_note": "number of note compliments received by the user",
  "compliment_plain": "number of plain compliments received by the user",
  "compliment_cool": "number of cool compliments received by the user",
  "compliment_funny": "number of funny compliments received by the user",
  "compliment_writer": "number of writer compliments received by the user",
  "compliment_photos": "number of photo compliments received by the user",
  "type": "user"
}
```
Scope

- One city but yet to be decided
- Focus on the end users, aka the people who use the Yelp site/app
- Data features to consider ... it depends but theme of holistic/detailed comparison
  - Discover the “nuances” behind the existing Yelp data eg. distribution of 5 star restaurants in different price categories
  - More informed decisions for end users
Project Pitch

Information Visualization 2017

Felix Grund
Munich
Do you need advice regarding your current IT solution or targeted assistance for a specific project? Our experienced consultants uncover the optimisation potential of your existing strategy and cater to your individual needs. We offer Requirements Engineering, UI/UX Consulting, Project Management and Cloud Strategy from a single source.

EXACTLY WHAT I NEED
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
• When do estimation and degree of completion conflict?
• When are our hourly rates too low?
• How do hourly rates change retrospectively?
• What tendencies can we observe over multiple projects?
• When interfere to maintain project success?
• How can we identify wrong estimations on DOC?
• How do project leads differ in their monthly estimations?
• ...
Is there still time?
## Time Tracking

### Tages-Informationen

<table>
<thead>
<tr>
<th>Datum</th>
<th>von</th>
<th>bis</th>
<th>Pause</th>
<th>Stunden</th>
</tr>
</thead>
<tbody>
<tr>
<td>13.05.2014</td>
<td>09:45</td>
<td>17:30</td>
<td>0,75</td>
<td>7,00</td>
</tr>
</tbody>
</table>

### Buchungs-Informationen

<table>
<thead>
<tr>
<th>Projektkennung</th>
<th>Tätigkeit</th>
<th>Issue ID</th>
<th>Bemerkung</th>
<th>Stunden</th>
</tr>
</thead>
<tbody>
<tr>
<td>DKMS_OPT1</td>
<td>Dashboard</td>
<td>DKMSSUPP</td>
<td>Dashboard Anpassungen</td>
<td>1,00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>95</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DKMS_OPT1</td>
<td>Dashboard</td>
<td>DKMSSUPP</td>
<td>Dashboard styling</td>
<td>0,75</td>
</tr>
<tr>
<td></td>
<td></td>
<td>96</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DKMS_OPT1</td>
<td>News-Umstrukturierung</td>
<td>DKMSSUPP</td>
<td>News Styling</td>
<td>1,25</td>
</tr>
<tr>
<td></td>
<td></td>
<td>102</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BSH_SUPP2014</td>
<td>PIWIK Support</td>
<td>BSHP-22</td>
<td>Test System Sync</td>
<td>0,50</td>
</tr>
<tr>
<td>BSH_SUPP2014</td>
<td>Corporate Wiki Support</td>
<td>BSHSUPP-226</td>
<td>Test Upgrade</td>
<td>0,25</td>
</tr>
<tr>
<td>INF_SEC2014</td>
<td>POC - User Switch</td>
<td>INFPRJ-18</td>
<td>Diskussionen Tests/Implementierung</td>
<td>1,00</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Tomcat Proxy</td>
<td></td>
</tr>
<tr>
<td>DKMS_OPT1</td>
<td>TPL</td>
<td>DKMSSUPP</td>
<td>Telko mit Guido zu Projektstand</td>
<td>1,00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>119</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Project results (good)

### Mitarbeiterstunden

<table>
<thead>
<tr>
<th>Mitarbeiter</th>
<th>Stunden</th>
</tr>
</thead>
<tbody>
<tr>
<td>fgrund</td>
<td>60,00</td>
</tr>
<tr>
<td>gschmidl</td>
<td>12,50</td>
</tr>
<tr>
<td><strong>Summe</strong></td>
<td><strong>72,50</strong></td>
</tr>
</tbody>
</table>

### Projektstunden

<table>
<thead>
<tr>
<th>Tätigkeit</th>
<th>Plan</th>
<th>Ist</th>
<th>Rest</th>
</tr>
</thead>
<tbody>
<tr>
<td>DKMS_OPT1: Blogposts-Plugin</td>
<td>16,00</td>
<td>14,75</td>
<td>1,25</td>
</tr>
<tr>
<td>DKMS_OPT1: Dashboard</td>
<td>16,00</td>
<td>15,25</td>
<td>0,75</td>
</tr>
<tr>
<td>DKMS_OPT1: IE10 Anpassungen</td>
<td>8,00</td>
<td>2,75</td>
<td>5,25</td>
</tr>
<tr>
<td>DKMS_OPT1: News-Migration</td>
<td>8,00</td>
<td>0,75</td>
<td>7,25</td>
</tr>
<tr>
<td>DKMS_OPT1: News-Umstrukturierung...</td>
<td>8,00</td>
<td>0,25</td>
<td>-2,25</td>
</tr>
<tr>
<td>DKMS_OPT1: Patient Faces</td>
<td>24,00</td>
<td>11,50</td>
<td>12,50</td>
</tr>
<tr>
<td>DKMS_OPT1: TPL</td>
<td>4,00</td>
<td>17,50</td>
<td>-13,50</td>
</tr>
<tr>
<td><strong>Summe</strong></td>
<td><strong>84,00</strong></td>
<td><strong>72,50</strong></td>
<td><strong>11,50</strong></td>
</tr>
</tbody>
</table>
### Mitarbeiterstunden

<table>
<thead>
<tr>
<th>Mitarbeiter</th>
<th>Stunden</th>
</tr>
</thead>
<tbody>
<tr>
<td>fgrund</td>
<td>52,00</td>
</tr>
<tr>
<td>jgrabski</td>
<td>10,75</td>
</tr>
<tr>
<td>jstadler</td>
<td>9,00</td>
</tr>
<tr>
<td><strong>Summe</strong></td>
<td><strong>71,75</strong></td>
</tr>
</tbody>
</table>

### Projektstunden

<table>
<thead>
<tr>
<th>Tätigkeit</th>
<th>Plan</th>
<th>Ist</th>
<th>Rest</th>
</tr>
</thead>
<tbody>
<tr>
<td>DKMS_SSO: SSO Link</td>
<td>0,00</td>
<td>14,25</td>
<td>-14,25</td>
</tr>
<tr>
<td>DKMS_SSO: SSO Plugin</td>
<td>32,00</td>
<td>57,50</td>
<td>-25,50</td>
</tr>
<tr>
<td><strong>Summe</strong></td>
<td><strong>32,00</strong></td>
<td><strong>71,75</strong></td>
<td><strong>-39,75</strong></td>
</tr>
</tbody>
</table>
Thanks.
Visualizing Internal Components of a Convolutional Neural Network

Mahdi Ghodsi - Hooman Shariati
Background:

What is Machine Learning

Machine Learning is taking over.

Applied to many fields:
  Bioinformatics, Gaming, Medical diagnosis,
  Marketing, Machine Vision, ....
Convolutional Neural Network

Machine Learning Approaches

- Deep Neural Network (DNN)
- Decision Trees
- Support Vector Machine (SVM)
- Bayesian Networks
The idea has been around since 1980s
But Introduction of GPU computing with 30x speed up gave DNNs a boost

Convolutional Neural Network

ImageNet Competition

Google Deep Dream
Convolutional Neural Network

Very Popular Research Area
Convolutional Neural Network

However ...
Convolutional Neural Network

How researchers see CNNs

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

Yajin Zhou

Department of Computer Science North Carolina State University
Convolutional Neural Network

How researchers see CNNs

How CNNs looks like
Visualizing and making sense of CNNs in literature:

Visualizing and Understanding Convolutional Networks By M. Zeiler (NYU)
Visualizing Ambiguity

James Hicklin
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>
Point estimates...

- Imagine Betty only cared about her chance of dying from breast cancer and her chance of developing endometrial cancer.
With confidence intervals...

5-year vs. 10-year Tamoxifen Therapy

Number of People (out of 1000)

Decrease in deaths from BC

Developing Endometrial Cancer
Alternatives to Error Bars

Violin Plots

http://www.datavizcatalogue.com/methods/violin_plot.html

Box Plots

http://www.datavizcatalogue.com/methods/box_plot.html

Dynamic Icon Arrays

Gradient Plots
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
- FSM 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 it’s 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
- FSM’s require both
Extensions to Project

Improving Visualization
Interaction Extension

FSM would provide a higher level on which users could zoom in on...
Filtering the Clusters

- Partitioning: intrinsic meaning
- Collect data invariants: filter to show aggregate data using existing tool set
- Label: Represent clusters by their invariants
- Visualize transitions: use the diff of cluster invariants
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?
Why this project is neat

- Stems from an existing body of work
- Has practical applications for debugging distributed systems
- No end of data to represent, can easily be extended after the course
Visualizing patient clusters

Lovedeep Gondara
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.
Thanks
Spanner, Resurrected.

CPSC 547 Project Pitch

Madison Elliott
February 16, 2017
Background

• Project originated as an MA thesis in the CS department
Background

• Project originated as an MA thesis in the CS department

• New technique that applied low-stretch trees to network visualization
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
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
Background

• Two iterations submitted for publication:
  1. Graph Drawing (technique focused)
  2. Pacific Vis (more emphasis on motivation and visualization application)
Background

• Two iterations submitted for publication:
  1. Graph Drawing (technique focused)
  2. Pacific Vis (more emphasis on motivation and visualization application)

• Both rejected 😞
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!
Resurrection Pitch

• Find the motivation!

• Develop and execute a user study
Resurrection Pitch

• Find the motivation!

• Develop and execute a user study

• Revise and resubmit paper
Why?

• Lots of potential!

Figure 3: Comparison between an arbitrary spanning tree and a low-stretch spanning tree for an 8-by-8 grid graph.
Why?

- De-hairball a cluttered network:
Why?

- Novel, layout free network idioms:
Next Steps

• Complete literature review of network idioms, tasks and taxonomies
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

<table>
<thead>
<tr>
<th>Date</th>
<th>Repo</th>
<th>#Sec</th>
<th>% overall</th>
<th>% pass</th>
<th>% cover</th>
<th>#P</th>
<th>#F</th>
<th>#S</th>
<th>#LOC</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>02/18 @ 04:40:53</td>
<td>cpsc310project_team40</td>
<td>12.0</td>
<td>72.8</td>
<td>66</td>
<td>95.89</td>
<td>33</td>
<td>17</td>
<td>0</td>
<td>316</td>
<td>green</td>
</tr>
<tr>
<td>02/18 @ 02:25:33</td>
<td>cpsc310project_team21</td>
<td>16.3</td>
<td>32.09</td>
<td>22</td>
<td>67.43</td>
<td>11</td>
<td>39</td>
<td>0</td>
<td>608</td>
<td>green</td>
</tr>
<tr>
<td>02/16 @ 01:33:41</td>
<td>cpsc310project_team80</td>
<td>9.2</td>
<td>44</td>
<td>30</td>
<td>95.84</td>
<td>15</td>
<td>35</td>
<td>0</td>
<td>505</td>
<td>green</td>
</tr>
<tr>
<td>02/16 @ 01:17:14</td>
<td>cpsc310project_team104</td>
<td>10.6</td>
<td>32.28</td>
<td>22</td>
<td>68.4</td>
<td>11</td>
<td>39</td>
<td>0</td>
<td>250</td>
<td>green</td>
</tr>
<tr>
<td>02/18 @ 01:10:42</td>
<td>cpsc310project_team65</td>
<td>108.6</td>
<td>25.04</td>
<td>12</td>
<td>72.21</td>
<td>6</td>
<td>44</td>
<td>0</td>
<td>439</td>
<td>green</td>
</tr>
<tr>
<td>02/18 @ 01:09:20</td>
<td>cpsc310project_team78</td>
<td>16.0</td>
<td>77.6</td>
<td>72</td>
<td>95.84</td>
<td>38</td>
<td>14</td>
<td>0</td>
<td>505</td>
<td>green</td>
</tr>
<tr>
<td>02/16 @ 01:08:58</td>
<td>cpsc310project_team17</td>
<td>18.2</td>
<td>87.2</td>
<td>84</td>
<td>95.73</td>
<td>42</td>
<td>8</td>
<td>0</td>
<td>539</td>
<td>green</td>
</tr>
</tbody>
</table>
Current Operational Dashboard

- AutoTest Queue: 0
- Reference UI: UP
- Geocoder: UP
- Class Portal: UP
Idea + Impact

Student facing dashboard
- Expanded to CS110, CS210, and CS310 + their corresponding MOOC offerings
- Vis will be used by 1000s of students in production system
- Challenge: make it engaging + promote ‘good’ behaviour
- Feedback: prototype can be made available to current students

Instructor facing dashboard
- Design study with domain expert (current CPSC310 instructor)
- Challenge: needs 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
EMAIL
Visual Methods for Analyzing Motifs in Time-Oriented Data

Soheil Kianzad
PhD student CS
Stock technical analysis

Yuan Li, GrammarViz, 2012

120 days

https://nonb-abcc.ncifcrf.gov/apps/site/help_visualize

-10%  +12%

Stock price $
ViSoccer

Visualizing European soccer players

Yann Dubois
Why?

Other sports

World cup

By region

https://buckets.peterbeshai.com/


http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=7042477

By game

What?

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

+ Sports page scraping
How?

- D3
- P5.js
- Tableau

http://www.optasports.com/about/what-we-do/visualisation.aspx
http://www.soccerstatistically.com/blog/2015/6/18/valuwing-the-mls-superdraft
GRADUATE STUDIES – DATA VISUALIZATIONS

JENS LOCHER, ASSISTANT DEAN – STRATEGIC TECHNOLOGIES AND BUSINESS INITIATIVES
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
• Central hub for everything related to graduate students
  • Communications & Recruitment
  • Admission
  • Awards
  • Thesis & 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
<table>
<thead>
<tr>
<th>CGPSS</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>
<tr>
<td>univ</td>
<td>38</td>
<td>48</td>
<td>50</td>
<td></td>
</tr>
</tbody>
</table>
CGPSS

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
## GRADUATE SCHOOL DATA (CURRENT)

<table>
<thead>
<tr>
<th>REGISTRATION PERIOD (EXTRACT DATE)</th>
<th>APPLICATIONS</th>
<th>OFFERS</th>
<th>ACCEPTS (CAME)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1995 (March-01-1996)</td>
<td>7,411</td>
<td>2,793</td>
<td>1,732</td>
</tr>
<tr>
<td>1996 (March-01-1997)</td>
<td>7,301</td>
<td>2,892</td>
<td>1,621</td>
</tr>
<tr>
<td>1997 (March-01-1998)</td>
<td>7,562</td>
<td>2,885</td>
<td>1,786</td>
</tr>
<tr>
<td>1998 (March-01-1999)</td>
<td>6,864</td>
<td>2,879</td>
<td>1,826</td>
</tr>
<tr>
<td>1999 (March-01-2000)</td>
<td>7,550</td>
<td>3,278</td>
<td>2,074</td>
</tr>
<tr>
<td>2000 (March-01-2001)</td>
<td>7,420</td>
<td>3,207</td>
<td>2,066</td>
</tr>
<tr>
<td>2001 (March-01-2002)</td>
<td>7,899</td>
<td>3,305</td>
<td>2,156</td>
</tr>
<tr>
<td>2002 (March-01-2003)</td>
<td>10,170</td>
<td>3,841</td>
<td>2,548</td>
</tr>
<tr>
<td>2003 (March-01-2004)</td>
<td>11,778</td>
<td>3,089</td>
<td>2,728</td>
</tr>
<tr>
<td>2004 (March-01-2005)</td>
<td>10,339</td>
<td>4,090</td>
<td>2,802</td>
</tr>
<tr>
<td>2005 (March-01-2006)</td>
<td>9,729</td>
<td>3,933</td>
<td>2,671</td>
</tr>
<tr>
<td>2006 (March-01-2007)</td>
<td>9,935</td>
<td>4,069</td>
<td>2,690</td>
</tr>
<tr>
<td>2007 (March-01-2008)</td>
<td>9,720</td>
<td>4,042</td>
<td>2,672</td>
</tr>
<tr>
<td>2008 (March-01-2009)</td>
<td>9,809</td>
<td>4,378</td>
<td>2,907</td>
</tr>
<tr>
<td>2009 (March-01-2010)</td>
<td>11,787</td>
<td>4,810</td>
<td>3,363</td>
</tr>
</tbody>
</table>

![UBC Graduate Studies - Application Process Numbers from 1995 to 2015](image)
GRADUATE SCHOOL DATA (CURRENT)
GRADUATE SCHOOL DATA (ALTERNATIVE EXAMPLE)

The University of Michigan offers a remarkably broad and rigorous array of graduate degree programs that are among the very best in the country in each field of study. The U-M attracts outstanding students to graduate study, and prepares them to make leading contributions to society through successful careers in professions and academic disciplines. Interdisciplinary study and joint degrees are a special strength of U-M’s programs. The vibrant community of graduate and professional students on campus is highly diverse in citizenship, demographic background, and intellectual perspective. The Rackham Graduate School works together with faculty in the schools and colleges of the University to sustain this diversity, understanding it as critical to our dynamic intellectual climate.

In order to make the activity and culture of graduate programs more visible, we provide basic statistics about the Ph.D. programs at the University of Michigan. The data and variables were selected to offer a more accurate and helpful picture than those provided by external sources. We also encourage you to visit specific graduate program websites to learn more about the intellectual life, successes, and opportunities in each program.

**ADMISSIONS**

**Enrollment**

**DOCTORAL FUNDING**

*In many cases, students received funding from external sources and did not need additional financial support.*
TEAM

Louise Mol
Systems and Data Analysis Manager

Jens Locher
Assistant Dean
Visualizing Trends in Product Recommendations

Q.I. Leap Analytics
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!

Q.I. Leap
lauren.fratamico@qileap.com